

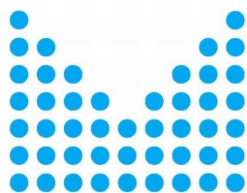


Transformace

digitálních prostorových dat

pro účely trvalého uložení
v digitálním archivu

kolektiv autorů
verze 1.0



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY

Transformace digitálních prostorových dat pro účely trvalého uložení v digitálním archivu

Verze 1.0

Ministerstvo vnitra ČR
Archivní správa
Nad Štolou 936/3
Poštovní schránka 21
170 34 Praha 7

T A

Tento projekt je financován se státní podporou
Technologické agentury ČR v rámci
Programu Éta

Č R

www.tacr.cz
Výzkum užitečný pro společnost

Název projektu:

Geografické informační systémy a plnění povinností legislativy v oblasti archivnictví

Č. projektu: 2020TL03000364

Řešitelé projektu:

Výzkumný ústav meliorací a ochrany půdy v.v.i.

Žabovřeská 250, 156 27 Praha 5 – Zbraslav

ČR - Národní Archiv

Archivní 2257/4, 149 00 Praha 4

Doba řešení: 1. 7. 2020 – 31. 12. 2022

Důvěrnost a dostupnost: veřejně dostupný (URL: <https://www.mvcr.cz/clanek/metodiky.aspx>)

Datum revize:

Program Éta

Program na podporu aplikovaného společenskovedního a humanitního výzkumu, experimentálního vývoje a inovací ÉTA (dále jen „program“) byl schválen usnesením vlády České republiky č. 37 ze dne 16. 1. 2017.

Národní archiv

Oddělení kontroly výkonu spisové služby, fondů státní správy po roce 1992 a elektronických dokumentů

Řešitelský tým:

Rechtarik Martin, Mgr.

Bernas Jiří, Bc.

Kunt Miroslav, Ing.

Nimrichtrová Pavlína, Mgr.

Vojáček Milan, PhDr., Ing., Ph.D.



Obsah

1.	Úvod	6
1.1.	Účel metodiky.....	6
1.2.	Cíl metodiky.....	6
2.	Pravidla pro trvalé ukládání prostorových dat.....	7
2.1.	Prostorová data	8
2.2.	Geografický informační systém/GIS	9
2.3.	Formáty pro předávání a trvalé uchování dat Geografických informačních systémů a prostorových dat	10
3.	Datový balíček pro trvalé uložení obsahující prostorová data (tzv. GeoSIP)	12
3.1.	Datová struktura EARK_Geospatial včetně popisu povinných i volitelných částí	14
3.1.1.	Metadata pro archivní katalogový systém (archivní metadata)	14
3.1.2.	Datasety – tzv. reprezentace.....	15
3.1.3.	Rastrový model dat	15
3.1.4.	Vektorový model dat.....	17
3.1.5.	Prostorová metadata/Geometadata.....	18
3.1.6.	Souřadnicové referenční systémy	18
3.1.7.	GIS (data Geografického informačního systému).....	19
3.1.8.	Databáze a tabulková data	20
3.1.9.	Documentation.....	22
3.1.10.	Schemas.....	22
4.	Postup výběru a ukládání prostorových dat.....	23
4.1.	Předání ve formě GeoSIP.....	25
4.2.	Předání ve formě datové struktury EARK_Geospatial.....	26
4.3.	Předání ve formě volných souborů v připravené struktuře	27
5.	Validace datová a metadatová.....	27
6.	Ověření postupu na příkladu GIS - data aplikace eKatalog BPEJ.....	28
6.1.	Popis zdrojových dat aplikace	28
6.2.	Příprava datových zdrojů.....	28
6.3.	Tvorba GeoSIP balíčku a datová validace	28
7.	Závěr	30
8.	Informační zdroje	31
8.1.	Odborná literatura.....	31
8.2.	Normy, směrnice, metodiky	32
8.3.	Webové zdroje	33
8.4.	Použité veřejné datové zdroje.....	35

Příloha č. 1 – Glosář.....	36
Příloha č. 2 – Přehled právních předpisů k prostorovým datům.....	42
Analýza klíčových právních předpisů.....	43
Analýza podpůrných právních předpisů.....	45
Příloha č. 3 – Dublin Core	48
Příloha č. 4 – Soubory METS.xml v EARK_Geospatial.....	49
METS.xml uvnitř adresáře /komponenty	49
Příloha č. 5 – mets.xml v GeoSIP	50
Popis vnitřních elementů datasetu	50
Ukázka souboru mets.xml Pro datový balíček GeoSIP	52
Příloha č. 6 – Návrh převodu elementů prostorových metadat do archivního prostředí.....	55
Příloha č. 7 – Ukázka archseznam.xml	57
Příloha č. 8 – Popis testovacích datasetů.....	58
1. GIS Data eKatalogu BPEJ.....	58
2. Topografická databáze České republiky (Data200) - vrstva Vodstvo.....	59
3. Austrálie – územně správní hranice	60
4. Charakteristiky toků a povodí ČR - Jizera DIBAVOD	61
Příloha č. 9 – Vzory licenčních dohod o převzetí/předání archiválií	62
Autorskoprávní otázky při předávání dat do archivu k trvalému uložení.	62
Dohoda o podmínkách převzetí archiválií.....	63
Zápis k úřednímu záznamu o předání archiválií	66
Příloha č. 10 – GIS formáty a jejich trvalé ukládání (katalog).....	68
GeoPackage	68
GML (Geographic Information - Geography Markup Language).....	69
KML (Keyhole Markup Language)	70
EO-GeoJSON, CityJSON (JavaScript Object Notation JSON)	71
Formáty Interlis 1 a Interlis 2	72
Formát TIFF (Tagged Image File Format nebo Tag Image File Format).....	73
GEOTIFF	75
JPEG (Joint Photographic Experts Group Interchange Format)	76
JPEG 2000 (Joint Photographic Experts Group 2000)	77
PNG (Portable Network Graphics).....	78
GIF (Graphics Interchange Format)	79
Esri Shapefile	80
Esri Grid	81
Esri BIL, BIP, BSQ.....	82

Formát Lidar/LiDAR/LADAR.....	83
GPJ, GRD, GRF, GSB, GSI, GSR, SRF (Formáty Golden Software)	84
SIARD (Software-Independent Archiving of Relational Databases)	85
CSV (Comma Separated Values).....	86
Formáty PDF/A	87
Formát PDF Geospatial.....	87
Formáty CAD (.dwg, .dwf, .dxf, .dgn a další)	89
Příloha č. 11 – Výběrový přehled formátů dle Formátových pravidel NAČR	90

1. Úvod

Archivnictví je humanitní obor, který je zaměřený na dlouhodobou péči o dokumenty vzniklé lidskou činností. Digitální archivnictví¹ se zabývá trvalým uchováním a zpřístupňováním digitálních záznamů a dokumentů. Základním principem je zajistit čitelnost, přístupnost a autenticitu dat, což v sobě zahrnuje zajištění bitstream ochrany (tzn. datovou identitu a integritu) ale i ochrany logické (čitelnost, srozumitelnost pro určitou komunitu budoucích uživatelů, možnost dalšího využívání, zajištění ochrany citlivých dat aj.).

1.1. Účel metodiky

Zákon č. 499/2004 o archivnictví a spisové službě (dále též „Zákon“) zcela jasně ukládá ve svém § 3 všem veřejnoprávním původcům povinnost uchovávat dokumenty a nabízet je k výběru archiválií. Termín dokument je definován v navazujících § 4 a § 5 a prostorová data zcela prokazatelně naplňují definici dle § 4 bod b) a § 5 odst. (2), a to v podobě celého datasetu. Všeobecný zájem o využívání prostorových dat jejich informační hodnotu a význam jen potvrzuje. Zmíněný zákon však pochopitelně už podrobněji nedefinuje, jak se to má v praxi realizovat. Metodika proto vyplňuje prostor, který vytvořil Zákon pro oblast prostorových dat a je určena jako pomůcka

- › archivům a archivářům
- › veřejnoprávním původcům
- › soukromoprávním původcům
- › a dalším tvůrcům a vlastníkům prostorových dat.

1.2. Cíl metodiky

Metodika si klade za cíl postihnout následující klíčové body pro úspěšné zvládnutí výběru, předání, uložení a trvalého uchovávání prostorových dat a GIS:

- › doporučení vhodných formátů pro předávání a trvalé uchování ve smyslu „nejlepší praxe“
- › definování logické struktury dat v podobě datasetů, tzv. informační balíčků, pro předávání k trvalému uložení v digitálním archivu (datových balíčků GeoSIP)
- › navržení doporučených postupů předávání dat s ohledem na vstupní možnosti předávajícího
- › transformace atributů metadat prostorových dat do archivních metadat sloužících k vyhledávání archiválií v archivních katalogových systémech a pomůckách

Autoři si dobře uvědomují omezení, které je spojeno s dalším rozvojem prostorových dat, používaných formátů, jejich zpracováním a sdílením. Proto lze očekávat, že metodiku bude nutné podrobovat revizím a úpravám, a to zejména v oblasti doporučených formátů a způsobu datového transferu. Stejně tak bude nutné metodiku adaptovat a rozvíjet v prostředí vzniklém v rámci vznikající Strategie rozvoje infrastruktury pro prostorové informace v České republice po roce 2020 (GeoInfostrategie 2020+).

¹ Dnes je již překonaná představa, že ekvivalentem digitálního archivu je zálohování. Digitální archiv v sobě obsahuje souhrn postupů, činností a organizačních opatření, díky nimž je možné pečovat o svěřené digitální entity. Pojem, který je na místě v této souvislosti použít, je „digital preservation“, srv. např. Trevor Owens: The Theory and Craft of Digital Preservation, 2018; Geoffrey Yeo, Records, information and data, 2018, a další

2. Pravidla pro trvalé ukládání prostorových dat

Hlavním úkolem archivů v České republice je výběr archiválií z nabízených dokumentů, péče o ně, a snaha o jejich zpřístupnění. Elektronické dokumenty se staly archiváliemi nového tisíciletí a mezi ně jednoznačně náleží prostorová data.²

Po počáteční fázi digitální archivace, kdy hlavním cílem byl výběr, formátové konverze a ukládání spíše menších datových objektů, stojí nyní archivnictví před novou fází, a to akvizice a uchování rozsáhlých datasetů z různých oblastí lidské činnosti, mnohdy ve formátech nevhodných pro dlouhodobé ukládání. Takovým případem jsou prostorová data, která mají významný potenciál pro budoucí využívání, reinterpretaci a nové vytěžování ve veřejném zájmu. S ohledem na členství České republiky v Evropské unii a její plány stát se lídrem v datově řízené společnosti a ekonomii,³ je trvalé ukládání digitálních datasetů a jejich zpřístupnění archivy klíčovou otázkou blízké budoucnosti.

Celosvětově byla již od počátku patrná snaha postihnout možnosti uchování prostorových dat pro příští generace⁴ a zároveň byla identifikována základní úskalí, která plynou z jejich charakteristik:

- › jedná se o komplexní a specifický typ digitálních dat
- › použití složitých a rozmanitých datových formátů, z nichž většina je proprietárních, a tedy i nevhodných pro trvalé ukládání
- › dynamičnost a pomíjivost dat
- › standardy pro geoprostorová metadata byly definovány jak na národní, tak na mezinárodní úrovni, přesto se metadata často od dat oddělují, nebo jsou nesprávná nebo nestandardní povahy
- › při uchování prostorových dat je třeba zahrnout souřadnicové referenční systémy, kartografické reprezentace, topologii, soubory projektu a strukturu dat
- › pokrývají širokou škálu datových struktur (vektorové a rastrové datové modely; nestrukturované a topologické; domény diskrétní i spojitě; formáty podporují různé podmnožiny a aspekty datové struktury). Nelze tedy hovořit o „prostorových datech“ jako o jednom druhu informací, které lze všechny zvládnout stejnými aplikacemi a formáty
- › prostorová data je možné zpracovávat na různých prostorových úrovních, data lze agregovat a členit, přitom každá z úrovní má své použití. Výběr úrovně prostorových dat může být jedním z kroků v procesu uchování dat
- › jejich objem neustále roste
- › masivně se využívají se od 90. let 20. stol., cenné mohou být informace i z letitých GIS systémů

Aby mohla být taková data předávána k trvalému uložení s perspektivou jejich budoucí dostupnosti a využitelnosti, je vhodné, aby byl v maximální možné míře zachován jejich technický a sociální kontext, aby byla logicky strukturována, dostatečně popsána příslušnými

² KALINA, Tomáš: Elektronické dokumenty – archiválie nového tisíciletí, in Archivy a zítřek, Sborník příspěvků z konference uspořádané ve dnech 22. – 23. září 1999, Jindřichův Hradec 2000, s. 111–117.

³ https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

⁴ <https://www.dpconline.org/docs/technology-watch-reports/2265-twr-09-01-gis-mcgarva/file>

metadata a doplněna nezbytnou dokumentací. Uložení objektů ve standardizované adresářové struktuře umožňuje kromě jiného i datovou a strukturální validaci dat.

Data by dále měla být předávána v datových formátech vhodných pro trvalé uchování. Takovými formáty jsou formáty, které jsou dobře zdokumentovány, jejich zdrojové kódy jsou volně dostupné, nebo formáty, které jsou lidsky čitelné, příp. jsou u nich známé a snadno dostupné migrační strategie s co nejmenším rizikem datové ztráty.

Z podstaty prostorových dat a GIS dále plyne, že často nelze jejich výběr provádět plošně automatizovaným způsobem a je třeba vyhodnocovat optimální strategii trvalého uchování případ od případu. Tam, kde je to možné, je žádoucí začlenit postupy výběru a předávání archiválií do životního cyklu dat.

2.1. Prostorová data

Pojmem prostorová data (nebo též *geografická data/informace, geoinformace, geoprostorová data/informace, geodata; angl. spatial data/information, geospatial data/information, georeferenced data/information; něm. Räumliche Daten, Geodaten, Geographischen Daten*) rozumíme „*Informace týkající se jevů implicitně nebo explicitně přidružených k místu vztahenému k Zemi. Zahrnuje údaje týkající se vzhledu těchto jevů, zejména jejich časoprostorové charakteristiky (popis polohy, geometrického tvaru a topologických vztahů), doplněné popisnými údaji s kvalitativními a kvantitativními charakteristikami. Nejčastějšími typy prostorově vztahených a reprezentovaných jevů jsou geometrické, příp. objemové objekty, geografické objekty a oblasti a astronomické objekty (nebeská tělesa). Jev může být modelován (zobrazen) bodem, křivkou, plochou nebo prostorovým tělesem. Prostorovou identifikaci objektů umožňují souřadnice, tj. soubor čísel určujících přímou lokalizaci bodu a topologie, tj. vztah k poloze jiných objektů.*“⁵ Jan Kolář jako prostorová data označuje: „*...údaje, jejichž výskyt či platnost je možné přiřadit určitému místu v prostoru. Pro prostorové informace následně získané z těchto dat obdobně platí, že jsou polohově vázána neboli lokalizována k určitému místu. Je-li pak toto místo speciálně na zemském povrchu nebo v jeho blízkosti, nazývají se příslušná prostorová data geografická data a informace z nich získané geografické informace.*“⁶

Prostorová data jsou využívána k poznávání minulosti, současnosti a modelování budoucnosti. Zpravidla jsou určena geometrickým objektem a prostorovými vztahy s okolím (topologií), dále jsou založena na ortofotografiích, satelitních snímcích, snímcích z dronů, na databázích obsahujících informace o zeměpisné délce a šířce nebo také adresních bodech. Geografické informační systémy nám pak umožňují takováto data zpracovávat a získávat z nich geografické informace.

Nejčastěji se prostorová data dělí na vektorové a rastrové datové modely. Detailnějšímu popisu těchto dat se věnuje kapitola [3.1](#) metodiky a její podkapitoly, jednotlivé pojmy a

⁵ Citováno z https://aleph.nkp.cz/F/?func=direct&doc_number=000000397&local_base=KTD, zdroj ČSN EN ISO 19101-1. Geografická informace - Referenční model. Část 1: základy. Praha: Úřad pro technickou normalizaci, metrologii a státní zkušebnictví, leden 2018, str. 48.

⁶ KOLÁŘ, Jan: Geografické informační systémy 10, Praha, ČVUT, 2003, str. 10-11; dále např. ŠTYCH, Přemysl a kol.: Vybrané funkce geoinformačních systémů, CITT Praha, Akademie kosmických technologií, Oblast Galileo, GMES, z podpory ESF, MPSV, Praha 2008, str. 10–12.

termíny jsou vysvětleny v [Glosáři](#). Pro práci s prostorovými daty je využívána široká škála geografických informačních systémů, které umožňují jejich tvorbu, lokalizaci, úpravy, vyhodnocování a prezentaci.

2.2. Geografický informační systém/GIS

Geografický informační systém⁷ (dále jen „GIS“; angl. *geographic information system*, něm. *Geographisch Informationssystem, Räumlich Informationssystem*) je počítačový informační systém, který umožňuje ukládat, spravovat a analyzovat geografická data a informace.

První geografický informační systém vyvinul v roce 1963 geograf Roger Tomlinson pro kanadskou vládu za účelem dohledu a správy nad přírodními zdroji. K již dříve digitalizovaným mapám se mu podařilo připojit počítačový systém, díky kterému šlo získávat a analyzovat lokalizovaná data. Tak byl de facto vytvořen koncept GIS, který je založen na propojení geografických entit a jejich vlastností, tzv. atributů (popisných dat).⁸

Původně tyto systémy sloužily kartografům k vytváření map, ale s postupující digitalizací se rozšířily do mnoha sektorů a jsou využívány k analýzám a práci s daty. Dnes je nalézáme v centrální státní správě,⁹ v samosprávě,¹⁰ na univerzitách,¹¹ ve vědecko-výzkumných ústavech¹² a samozřejmě v soukromém sektoru.¹³

GIS jsou využívány v nepřeberném množství podob k nejrůznějším účelům, např. k modelování změn krajiny v souvislosti s klimatickou změnou, k vytváření územních plánů samosprávou, pro optimální instalaci inženýrských sítí atd.¹⁴ Pro tyto činnosti mohou být využívány specifické mapové podklady, na kterých jsou modelována geometrická data, a to v podobě bodů, linií, ploch, anebo i trojrozměrné objekty, ale také rastrové objekty jako např. eroze půdy,¹⁵ teplota, úhrny srážek¹⁶ atd. Tyto informační systémy de facto nahradily původně používané průhledné fólie, které byly zakreslovány nad analogovými mapami.

Data z geografického informačního systému chápeme z hlediska zákona o archivnictví buď jako jeden dokument, tedy jako jeden celek, ale lze je též hodnotit jako jednotlivé dokumenty,

⁷ https://aleph.nkp.cz/F/BIKN68EGRYMF15RBDVKIBQ5N6QGCA6JNKA1DMDHTVTMLJ13H9G-37562?func=find-b&find_code=WTD&x=0&y=0&request=gis&adjacent=N

⁸ <https://www.britannica.com/technology/GIS>

⁹ Např. Národní geoportál Inspire, dostupné z: <https://geoportal.gov.cz/web/guest/home/>; geoportál ČÚZK, dostupné z <https://geoportal.cuzk.cz/>; mapy pokrytí datovými signály ČTÚ, dostupné z: <https://digi.ctu.cz/>; geoportál Ministerstva zdravotnictví <https://geoportal.mzcr.cz/portal/home/> atd.

¹⁰ Nejen velká města jako např. hl. m. Praha, dostupné z: [Mapové aplikace | Geoportál hl. m. Prahy \(geoportalpraha.cz\)](#); Plzeň, dostupné z: <https://mapy.plzen.eu>, ale i menší, např. Říčany, dostupné z: <https://ricany.gepro.cz> a též obce, např. Lejšovka, dostupné z: <https://www.lejsovka.cz/obec/geoportal>

¹¹ Např. Masarykova Univerzita, dostupné z: <https://geoportal.muni.cz/>; Univerzita Palackého v Olomouci, dostupné z: <https://gis.upol.cz/>

¹² Např. VÚMOP, dostupné z: <https://geoportal.vumop.cz/>; Národní památkový ústav, dostupné z: <https://geoportal.npu.cz/web/>; Krkonošský národní park, dostupné z: <https://ags.krnapp.cz/mapy/prohlizecka/>

¹³ Např. ČEZ a.s., dostupné z <https://geoportal.cezdistribuce.cz/geoportal/> Gepro spol. s.r.o., dostupné z: <http://www.gepro.cz>

¹⁴ Např. Digitální technická mapa Hlavního města Prahy, dostupné z: <https://app.iprpraha.cz/apl/app/dtmp/index.html>

¹⁵ Aplikace Monitoring Eroze půdy, VÚMOP. Dostupné z: <https://me.vumop.cz/app/>

¹⁶ Např. radarové modely Českého hydrometeorologického ústavu, dostupné z https://www.chmi.cz/files/portal/docs/meteo/rad/data_israview.html

podle zvoleného výběru a podle zvolené vrstvy. Je nutné poznamenat, že pro trvalé ukládání prostorových dat je zcela nezbytné předávat data v komplexní podobě. Rozpad datasetu na jednotlivé komponenty by zásadně snížil budoucí výpovědní hodnotu a omezil by interoperabilitu a možnosti obnovy a zpřístupnění dat.

2.3. Formáty pro předávání a trvalé uchování dat Geografických informačních systémů a prostorových dat

Volba vhodných, zdokumentovaných a vyzkoušených formátů je pro oblast trvalého ukládání digitálních dat naprosto klíčovou otázkou. Na těchto vlastnostech totiž závisí dlouhodobá čitelnost uložených dat¹⁷. Právě z tohoto důvodu digitální archiv Národního archivu na vstupu všechny formáty identifikuje¹⁸ a kontroluje dle zveřejněných formátových pravidel.¹⁹

Formáty používané v oblasti prostorových dat nejsou výjimkou a jejich dlouhodobá čitelnost ve spojení s předpokládanou interoperabilitou je zatím velkou neznámou. Obecně známý formát Shapefile, který je nyní nejčastěji používaný k předávání nebo zveřejňování prostorových dat, je vzorovým příkladem nevhodného formátu pro trvalé uchování.²⁰ Na druhou stranu při konverzích do formátů vhodných pro dlouhodobé ukládání může docházet ke ztrátě informací, a tím k jejich znehodnocení. Je proto nezbytné hledat, popsat a testovat vhodné formáty pro dlouhodobé uložení prostorových dat a lze v této oblasti očekávat další vývoj. Na základě testovacích datových setů, vyzkoušených postupů, získaných znalostí jsou v datasetech vyžadovány **tyto výstupní formáty pro prostorová data**:

1. **Formát GML** (Geography Markup Language `.gml`) pro vektorová data nebo objekty s ohledem na dlouhodobou datovou čitelnost a široké využití v oblasti EU.
2. Pro rastrová data veškeré formáty dle § 23 odst. 3 vyhlášky č. 259/2012 Sb., o podrobnostech výkonu spisové služby v aktuálním znění.²¹ Doporučeným formátem je **TIFF** (vč. verze pro prostorová data **GEOTIFF**)
3. Pro relačně databázové části systémů GIS doporučujeme využít ověřený formát **SIARD**, tabulková data lze také ukládat ve formátu **CSV**, kódovaná v UTF-8 a opatřená metadaty, které umožňují zpětný import do databáze (např. soubor `createdb.sql`)
4. Dokumentaci, která je přílohou ukládaného datového setu, doporučujeme konvertovat dle § 23 odst. 2–4 Vyhlášky.

Akceptovanými formáty jsou:

5. Formát GeoPackage (`.gpkg`), dokumentace tohoto formátu je otevřená a zveřejněná, je možná jeho datová validace²² a navíc není vyloučeno ani budoucí označení tohoto formátu jako vhodného pro trvalé ukládání prostorových dat. Tento formát

¹⁷ Další možnou strategií trvalého uchování, emulaci, v současné době Národní archiv standardně nevyužívá a nemá zařazenu do běžných postupů, zejména z důvodu značných nároků na technické a licenční zajištění. I když ani tyto postupy nejsou vyloučeny, je tato metodika plně orientovaná výhradně na metodu migrace formátů.

¹⁸ <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>.

¹⁹ Bližší informace o formátech pro prostorová data obsahuje příloha 9, Formátová pravidla Národního archivu jsou blíže popsána v příloze 10.

²⁰ <http://switchfromshapefile.org>, <https://www.gis-blog.com/geopackage-vs-shapefile/>

²¹ Dále jen Vyhláška.

²² <https://github.com/PDOK/geopackage-validator>

doporučujeme využít jako druhou reprezentaci dat v případě, kdy velikost vstupních dat znemožňuje transformaci do formátu .gml. Objemově se jedná o jednotky gigabytů a s tím spojená technická omezení pro zpracování velkého objemu dat v textové podobě. Pokud z nějakého důvodu proběhlo nebo proběhne rozdělení datasetu na menší části do formátu .gml, reprezentaci ve formátu .gpkg je vhodné uložit jako celek, tj. nerozdělenou. Jeho využití jako jediné reprezentace je nutno konzultovat s archivem, protože dlouhodobá čitelnost tohoto formátu zatím není ověřena.

6. Formát JPEG 2000 pro rastrová data v případě bezztrátové komprese.²³
7. Soubory ve formátech obsahující popisné informace k ukládaným datům, např. soubory ve formátech SLD, LYR, LYRX případně i jiné. Objekty tohoto typu, které jsou lidsky čitelné a editovatelné pomocí dostupných textových editorů, doporučujeme uložit též ve formátu PDF/a a zajistit tak jejich integritu.
8. Formáty rodiny LIDAR.
9. Formáty rodiny CAD.

Přijímané formáty, u nichž lze předpokládat možnou ztrátu čitelnosti v dlouhodobém časovém horizontu a formátové konverze do nové datové reprezentace buď během ukládacího procesu²⁴, nebo během doby uložení:

10. ESRI Shapefile (jako sestava povinných komponent .shp, .shx a .dbf a navázaných komponent ve formátech .prj, .xml, .cpg atp.)
11. Formáty rodiny Golden Software (.gpj, .grd, .grf, .gsf, .gsi, .gsr, .srf)
12. Formáty rodiny CAD (.dwg, .dxf, .dgn a další)

²³ <https://knihovnavue.nkp.cz/archiv/2020-2/recenzovane-prispevky/posuzovani-souborovych-formatu-z-hlediska-dlouhodobeho-uchovavani-a-navrh-metodiky-pro-narodni-knihovnu-ceske-republiky>

²⁴ Tímto se rozumí, že objekty datasetu budou uloženy v originálních formátech i s vědomím, že šance na jejich dlouhodobou čitelnost je velmi malá, a proto bude předmětem uložení i datová reprezentace v některém z výstupních nebo alespoň akceptovaných formátů i za cenu datové ztráty (pokud možno co nejmenší), neboť data uložená v proprietární podobě není možné ve všech případech migrovat do formátů doporučených v [Příloze č. 10](#).

3. Datový balíček pro trvalé uložení obsahující prostorová data (tzv. GeoSIP)

Na základě zjištěných informací, mezinárodních zkušeností a doporučení **Metodika definuje nový typ informačního balíčku SIP**, ve kterém budou předávána prostorová data do digitálního archivu, tzv. **GeoSIP**. Prostorová data není vhodné uchovávat odtrženě od jejich kontextu, a proto je zcela nezbytné, aby byla opatřena popisnými metadaty a dokumentací, a to vše bylo předáváno k trvalému uložení v přehledné a srozumitelné struktuře. Za tímto účelem byla definována struktura datasetů, tzv. informačních balíčků pro přenos a ukládání v digitálním archivu.

Princip datových informačních balíčků je založen na referenčním modelu OAIS²⁵, který předpokládá použití informačních balíčků pro přenos do archivu (tzv. SIP), balíčků pro trvalé uchování v digitálním archivu (AIP) a balíčků pro zpřístupnění (DIP).²⁶ Tato metodika se zabývá datovým balíčkem GeoSIP pro transformaci a přenos prostorových dat a komplexu dat z Geografických informačních systémů do digitálního archivu.

Struktura datového informačního balíčku pro trvalé uložení prostorových dat GeoSIP respektuje tradiční strukturu českých datových balíčků SIP, která je závazná v České republice od r. 2017 a je zakotvena Národním standardem pro elektronické systémy spisové služby (dále NSESSS)²⁷. Z důvodu mezinárodní kompatibility bylo rozhodnuto, že vnitřní struktura GeoSIP bude plně odpovídat evropské specifikaci CITS Geospatial pro informační balíček s prostorovými daty (dále jen EARK_Geospatial)²⁸.

GeoSIP má podobu adresáře, v něm vloženého souboru mets.xml²⁹ a adresáře /Komponenty. Ten obsahuje validní strukturu EARK_Geospatial dle následující kapitoly 3.1. Pro tvorbu GeoSIP slouží nástroj pro automatickou transformaci prostorových dat pro účely trvalého uložení – ArchiGIS.³⁰

Nástroj ArchiGIS vytváří informační balíčky GeoSIP dle této kapitoly, strojově je opatřuje metadaty, která zajistí datovou integritu a autenticitu dle příloh č. 4 a č. 5., extrahuje/vytváří popisná metadata do souboru dc.xml; pokud nejsou obsažena ve struktuře předávaných dat, vyžaduje jejich vyplnění. Umožňuje tvorbu archivního seznamu archseznam.xml dle přílohy č. 7 pro potřeby automatizace procesů předání datových balíčků v prostředí Národního archivního portálu (srv. kapitolu 4 této metodiky). Zmíněný seznam dále umožňuje předání více balíčků v rámci jedné předávky.

²⁵ ČSN ISO 14721. [Systémy pro přenos dat a informací z kosmického prostoru – Otevřený archivační informační systém – Referenční model.](#)

²⁶ Vysvětlení pojmů SIP, AIP a DIP srv. [přílohu 1 - Glosář](#)

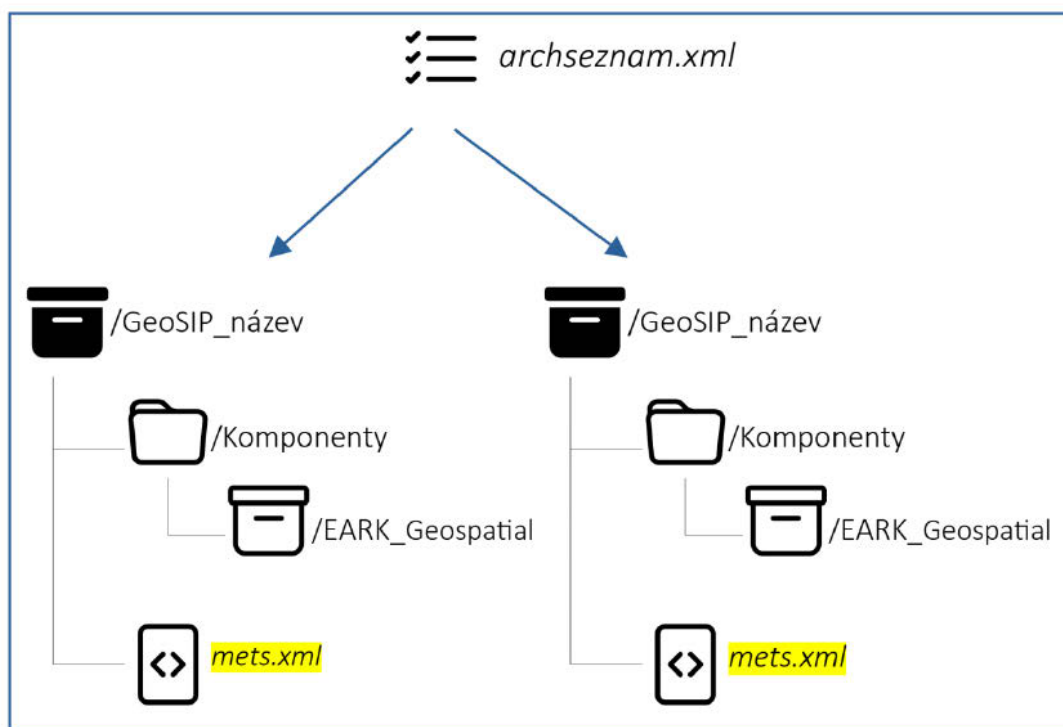
²⁷ <https://www.mvcr.cz/soubor/vestnik-mv-57-2017-oznameni-ministerstva-vnitra-kterym-se-zverejnuje-narodni-standard-pro-elektronicke-systemy-spisove-sluzby.aspx>

²⁸ [Specification for digital geospatial data records archiving, verze 3.0.0 z 31. 8. 2021, dostupné z: https://github.com/DILCISBoard/CITS-Geospatial.git](https://github.com/DILCISBoard/CITS-Geospatial.git)

²⁹ Srv. [Přílohu č. 5](#) této metodiky

³⁰ Vývoj nástroje ArchiGIS je úkolem Výzkumného ústavu meliorací a ochrany půdy jako součást společného projektu z podpory Technické agentury ČR, srv. str. 2

Obrázek 1 – Schéma podoby GeoSIP pro modul ePředání Národního archivního portálu



Obrázek 2 - Schéma předání dvou balíčků GeoSIP v rámci jedné předávky

3.1. Datová struktura EARK_Geospatial včetně popisu povinných i volitelných částí

Datová struktura EARK_Geospatial je povinným obsahem datového balíčku GeoSIP a je plně tvořena podle specifikace EARK_Geospatial³¹.

EARK_Geospatial má vždy podobu adresářové struktury. Kořenový adresář nese název zvolený původcem, doporučujeme použít název shodný s názvem ukládaného datasetu bez diakritiky a bez mezer, které je vhodné nahradit podtržítkem nebo pomlčkou. Využití velkých písmen např. na začátku názvu adresáře je povoleno, ale není vyžadováno. Kořenový adresář obsahuje vždy povinnou adresářovou strukturu a soubor *METS.xml*. Podadresáře /Metadata a /Representation obsahují volitelně-povinnou podadresářovou strukturu, která bude validována a podadresář /Representation obsahuje navíc další soubor *METS.xml*.³² Přípustné varianty zobrazují schémata struktur na obrázcích 4 až 6, **závazné jsou pouze adresáře v povinné části struktury.**

Tuto strukturu lze rovněž nahrát do Národního archivního portálu, s ohledem na plánovanou automatizaci však doporučujeme použití datového balíčku GeoSIP.

3.1.1. Metadata pro archivní katalogový systém (archivní metadata)

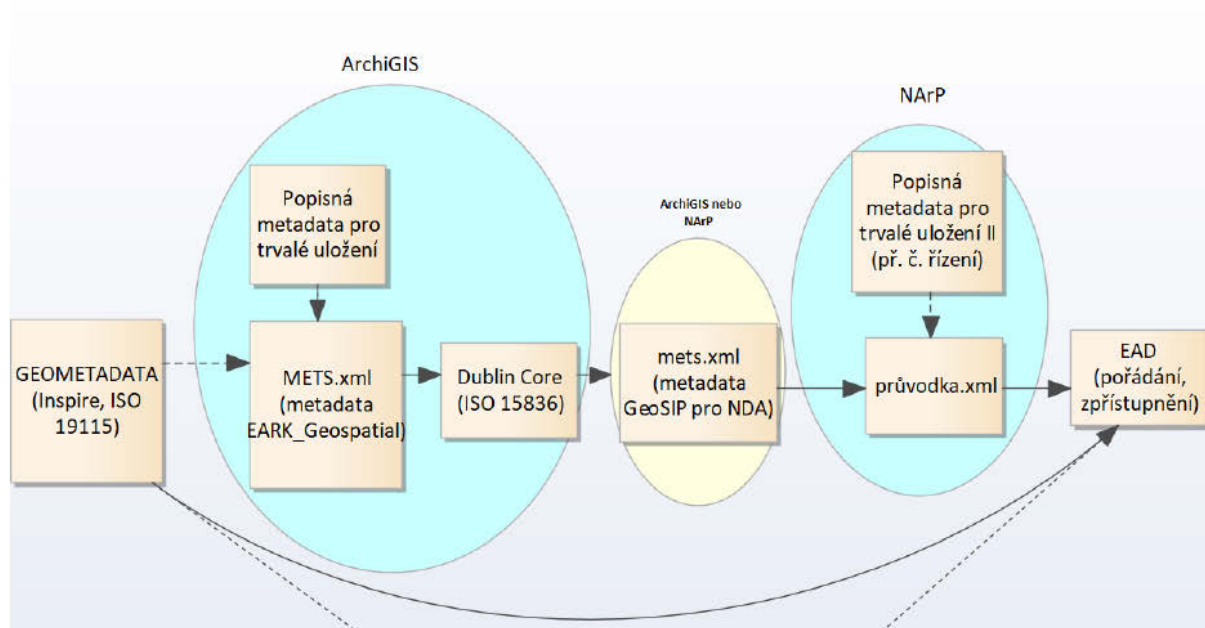
Metadata pro archivní katalogový systém jsou strojově vytěžována ze souboru mets.xml na úrovni celého balíčku GeoSIP:

- a) Soubor [mets.xml](#) je vytvořen původcem (při použití nástroje ArchiGIS) a uložen na stejné úrovni jako adresář /Komponenty v datovém balíčku GeoSIP
- b) nebo se vytváří automatizovaně v prostředí NArP vytěžením popisných metadat ze souboru dc.xml dle [přílohy č. 3](#) při nahrávání struktury EARK_Geospatial
- c) nebo se vytváří automatizovaně v prostředí NArP z metadat vyplněných původcem při popisování ukládaného datasetu.

Základními vytěžovanými metadaty jsou: Název datového balíčku, komentář, který obsahuje bližší popis datasetu, datum uzavření balíčku, pořadové číslo v případě že datový balíček obsahuje více datasetů, časový rozsah konkrétního datasetu, typ omezení v případě že dataset obsahuje chráněné informace (např. osobní údaje, autorská práva aj.), číslo jednacích žádostí o předání, k němuž se bude vázat úřední záznam o převzetí s identifikátorem archivního balíčku, tzv. AIPu. Dále mezi archivní metadata patří druh evidenční jednotky a počet jednotek v balíčku. Detailní rozbor a ukázkou souboru mets.xml obsahuje [příloha č. 5](#).

³¹ [Specification for digital geospatial data records archiving, verze 3.0.0 z 31. 8. 2021, dostupné z https://github.com/DILCISBoard/CITS-Geospatial.git](https://github.com/DILCISBoard/CITS-Geospatial.git); Zároveň byla vydána i metodika obsahující podrobnosti stavby a využití struktury EARK_Geospatial Guideline for CITS Geospatial, dostupné z <https://github.com/DILCISBoard/CITS-Geospatial>.

³² Blíže příloha č. 4.



Obrázek 3 - Využití popisných a geoprostorových metadat pro trvalé uložení

3.1.2. Datasets – tzv. reprezentace

Datasets se ukládají v rámci jednoho balíčku SIP vždy do adresáře /nazen_datasetu do podadresáře /representations ve struktuře popsané detailně v rámci kapitoly 4 a vzorových příkladů. Adresářová struktura je strojově kontrolována vůči souboru mets.xml a zpravidla obsahuje následující datové objekty v různých formátech.

3.1.3. Rastrový model dat³³

Rastrová data zachycují hodnoty sledovaných jevů v konkrétních lokalitách. Na rozdíl od vektorových modelů objekty jako takové neexistují, je pouze zaznamenán atribut v konkrétní poloze. Tyto informace jsou ukládány do polí mřížky (tzv. pixely), které jako celek tvoří obraz. Rastrové modely se využívají pro zobrazení veličin jako např. srážky, teplota, vegetace, typy půd atd. ve sledovaném prostoru.

Pro rastrové modely je důležitá volba vhodného formátu, který v dostatečné míře dochová kvalitu barevné škály.³⁴ Z výstupních formátů dle Zákona lze využít kterýkoliv z formátů zveřejněných v § 23 čl. odst. 3³⁵ vyhlášky č. 259/2012 v aktuálním znění a doplnit ho potřebnými informacemi o souřadnicích, pokud nejsou přímo jeho součástí. Většina rastrových formátů je dobře dokumentovaná, softwarové nástroje jsou dostupné, proto většina problémů vzniká při ukládání proprietárních formátů a vlastností obrázků.³⁶

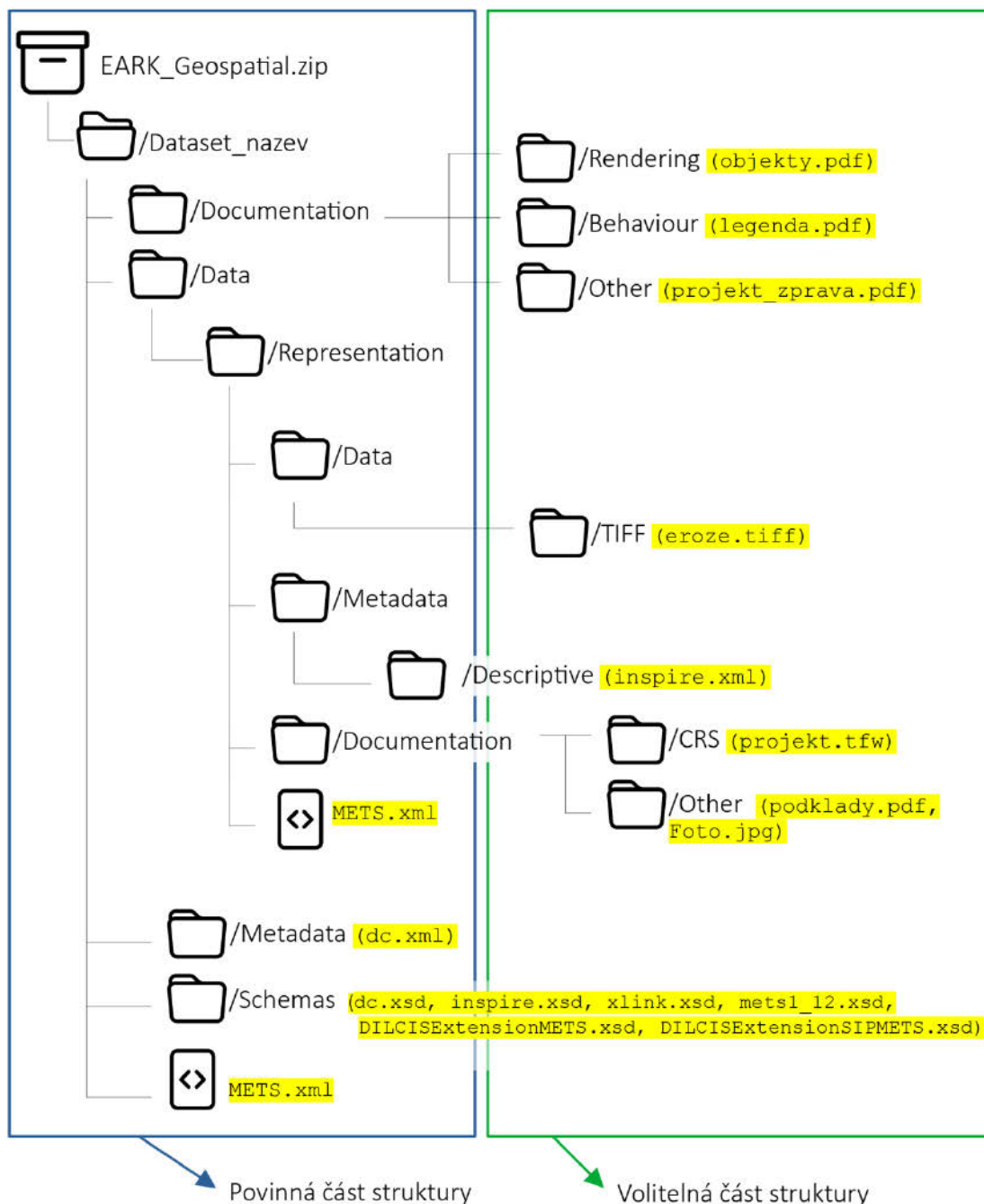
³³ Více o rastrovém datovém modelu in ŠTYCH, Přemysl a kol.: Vybrané funkce ..., str. 14–17; dále např. KOLÁŘ, Jan: Geografické inf... kapitola 3.4 str. 36–42, ČVUT, Praha 2003

³⁴ Srv. poznámka 23

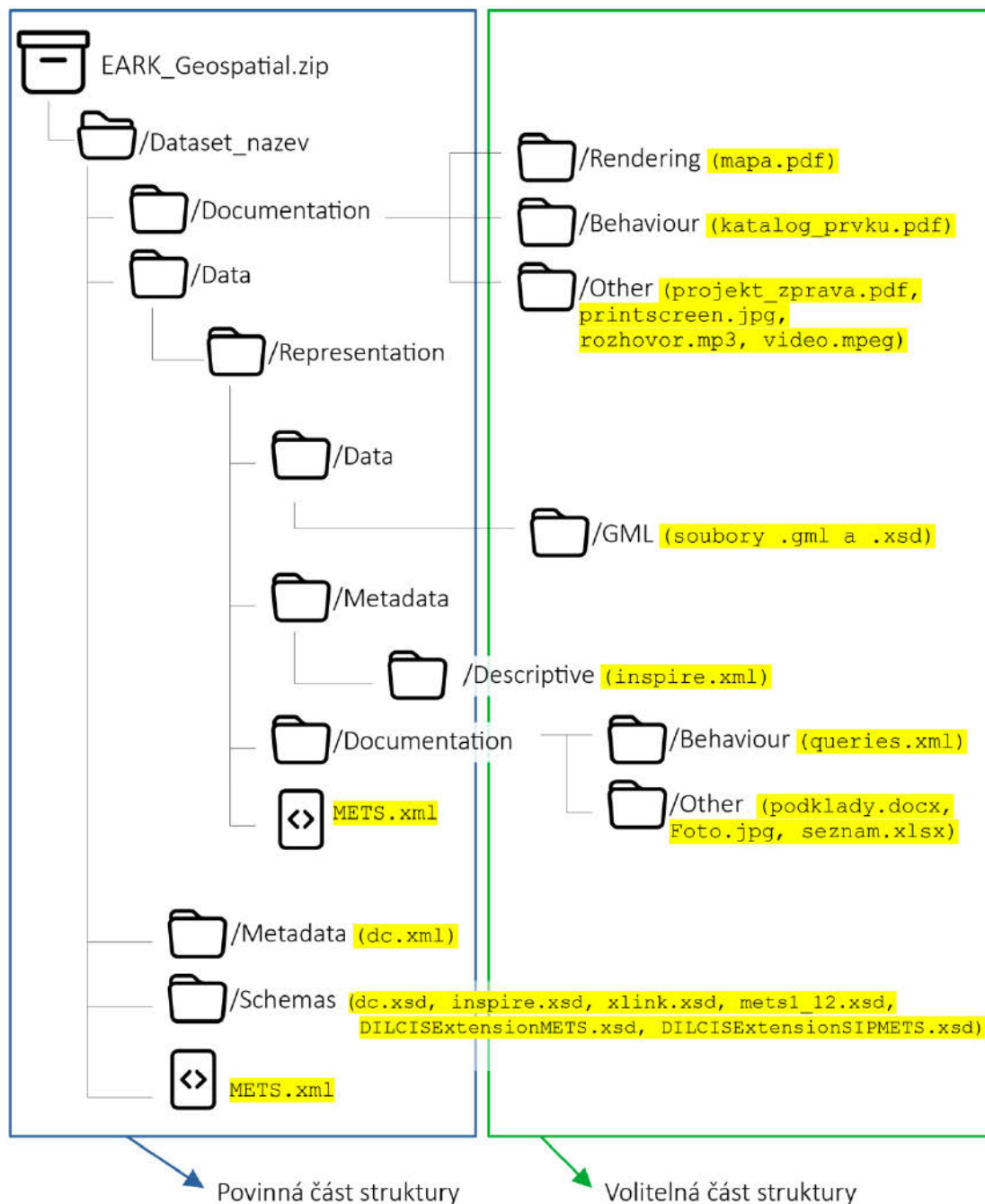
³⁵ Upozorňujeme však na možnou ztrátovost formátu JPEG, srv. Příloha č. 10 [JPEG \(Joint Photographic Experts Group Interchange Format\)](#) a [JPEG 2000 \(Joint Photographic Experts Group 2000\)](#)

³⁶ Preserving Raster Images, Data Types Series, Artefactual Systems and the Digital Preservation Coalition, 2021, Dostupné on-line: <https://www.dpconline.org/docs/technology-watch-reports/2478-preserving-raster-images/file> Cit. 26. 8. 2021

Pokud byl pro tvorbu rastrového modelu použit jakýkoliv mapový podklad, je naprosto zásadní, aby i tento podklad byl dalším objektem trvalého uložení a součástí předávaného datasetu. Schéma struktury uložení rastrových dat v balíčku GeoSIP zobrazuje obrázek 4.



Obrázek 4 - Prostorová data uložená ve struktuře EARK_Geospatial obsahující rastrová data v jedné reprezentaci



Obrázek 5 - Prostorová data uložená ve struktuře EARK_Geospatial obsahující vektorová data v jedné reprezentaci

3.1.4. Vektorový model dat³⁷

Vektorové objekty jsou sestavou geometrických elementů reprezentující objekty reálného světa jako prvky, které jsou reprezentovány body, liniemi a polygony.³⁸ Tyto objekty vyplňují prostor, zatímco okolní prostor není definován. Vlastnosti těchto geometrických prvků, tzv. atributy, jsou s prvky provázány v podobě popisných dat. Vektorový datový model se

³⁷ Více o vektorovém datovém modelu in ŠTYCH, Přemysl a kol.: Vybrané funkce geoinformačních systémů, CITT Praha, Akademie kosmických technologií, Oblast Galileo, GMES, z podpory ESF, MPSV, Praha 2008, str. 12–13; dále např. KOLÁŘ, Jan: Geografické inf... kapitola 3.3, str. 31–35.

³⁸ Používány jsou také ekvivalentní termíny: tečka, úsečka a plocha.

v systémech GIS dále dělí na 1. Špagetový model, 2. Hierarchický model a 3. Topologický model. – link na literaturu

Ačkoliv se jedná nejčastěji o tabulku nebo o sestavy tabulek s daty a datovými typy, které popisují prostorové objekty, není ve většině případů možný snadný export takové tabulky např. do souboru CSV. To je dáno přítomností informací o poloze objektů (polohových dat). Pro export ze systémů GIS je proto nutné využít vhodné formáty, zejména GML, GeoPackage aj., aby nedošlo k datové ztrátě a tak k nevratnému poškození prostorových dat.³⁹

Pokud byl pro tvorbu vektorového modelu použit jakýkoliv mapový podklad, je naprosto zásadní, aby i tento podklad byl dalším objektem trvalého uložení a součástí předávaného datasetu. Schéma struktury s vektorovými daty zobrazuje obrázek 5.

3.1.5. Prostorová metadata/Geometadata⁴⁰

Prostorová data jsou v balíčku dokumentována ve formě prostorových metadat, která obsahují běžné popisy dat a také popisy specifické pro prostorová data (přesnost, počet řádků, měřítko, měrné jednotky, informace o souřadnicovém systému atd.). Jedná se o metadata, která vytváří původce k bližšímu popisu vytvořeného datasetu. V původních systémech lze geoprostorová metadata ukládat různými způsoby (databáze, standardizované soubory .xml, běžná dokumentace atd.).

Metadata jsou obsažena buď na úrovni celého balíčku, nebo pro konkrétní reprezentaci, podle potřeby původce dat. Do datového balíčku GeoSIP by měla být přidána ve strojově čitelném formátu (např. .xml) do adresáře /Metadata/Descriptive pro celý balíček nebo do adresáře Representation/Rep.../Data/Metadata/Descriptive, pokud jsou specifická pro jednotlivou reprezentaci.

Dataset předaný k trvalému uložení bude dále detailněji popsán vytěžením příložených metadat, jak je naznačeno na obrázku 3, a to dle připraveného mapování prostorových metadat do archivního standardu Encoded Archival Description⁴¹ v [příloze č. 6](#), aby byl archivní popis v souladu s Metodickým pokynem č. 2/2020 pro evidenci a zpřístupňování map v archivech České republiky.⁴²

Nepovinný podadresář /Other je vyhrazen pro jakoukoli jinou, nestandardizovanou, strojově nebo lidsky čitelnou dokumentaci obsahující kontextové informace, jako jsou soubory projekt_zprava.pdf, projekt_rozhovor.mp3. Tento adresář je doporučený.

3.1.6. Souřadnicové referenční systémy

Jedná se o matematická pravidla, která umožňují ke každé hodnotě souřadnic přiřadit polohu objektu a naopak – souřadnicový systém (coordinate reference systém – CRS). Znalost souřadnicového referenčního systému je důležitá pro přesné použití geoprostorových dat.

³⁹ [Specification for digital geospatial data records archiving, verze 3.0.0 z 31. 8. 2021, dostupné z https://github.com/DILCISBoard/CITS-Geospatial.git](https://github.com/DILCISBoard/CITS-Geospatial.git)

⁴⁰ Legislativa EU - Národní geoportál INSPIRE (gov.cz) <https://geoportal.gov.cz/web/guest/legislativa-eu>

⁴¹ <https://www2.archivists.org/groups/technical-subcommittee-on-encoded-archival-standards-ts-eas/encoded-archival-description-ead>

⁴² <https://www.mvcr.cz/soubor/metodicky-pokyn-c-2-2020-pro-evidenci-a-zpristupnovani-map-v-archivech-ceske-republiky.aspx>

Řídí se normou ČSN 19111 (979830) Geografická informace – Vyjádření prostorových referencí souřadnicemi.

Souřadnicových systémů existuje v geoinformatice velké množství, přičemž každý z nich se různou měrou hodí pro práci v různých měřítkách a na různých územích v závislosti na tom, jaké referenční těleso a případně kartografické zobrazení používá. Pro Českou republiku jsou závazné následující systémy: S-JTSK,⁴³ dále systém ETRS⁴⁴ pro území členských států Evropské unie, dále se lze setkat se souřadnicovým systémem WGS⁴⁵, který zahrnuje celý svět, a také se systémem S-1942.⁴⁶ Kromě těchto závazných systémů existuje i velké množství jiných,⁴⁷ které nástroje pro práci s prostorovými daty umí vzájemně migrovat.

Některé formáty geoprostorových dat (rastrové i vektorové) přímo obsahují informace o souřadnicovém referenčním systému, ze kterého datová sada vychází, buď vložené jako součást samotného souboru (např. GeoTIFF), nebo jako další samostatný soubor (např. „.prj“ pro ESRI Shapefile). Rizikové jsou varianty, kdy formát souřadnicové informace neobsahuje (např. TIFF image file) a jsou obsaženy v samostatném souboru (např. TIFF World File – TFW). Proto je nutné, aby tyto informace byly zahrnuty v metadatech datové sady. Dle specifikace jsou uložena v podadresáři `/documentation/CRS`.⁴⁸

3.1.7. GIS (data Geografického informačního systému)

Podadresář GIS náleží do adresáře s datovými reprezentacemi a původce/archivář ho vytváří v případě archivace kompletního nebo i několika částí systému GIS. V takovém případě je zcela klíčová podrobná dokumentace, která se ukládá do vlastního podadresáře `/documentation`. Podadresář dokumentace obsahuje především seznam datasetů spravovaných v GIS, zjednodušenou vizualizaci, případně více vizualizací dat, které datový balíček obsahuje, a vzorky výstupů, které bylo možné ze systému GIS získat.

Podadresář `/metadata` by měl obsahovat datový model, informace o relacích objektů, popisy vrstev, vlastnosti vrstev, souřadnicové transformace, popisky prostorových objektů, dotazy (queries), kterými jsou objekty propojeny s tabulkovými daty, a popis tvorby výstupů, které bylo možné ze systému GIS získat. Dále do tohoto podadresáře náleží potřebná metadata pro obnovení funkčnosti systému GIS.

Podadresář `/data` je doporučeno hlouběji strukturovat, aby byla zachována logická struktura dat archivovaného systému. Vhodným příkladem je např. struktura v rámci adresáře

⁴³ S-JTSK je závazným geodetickým referenčním systémem na území ČR dle nařízení vlády č. 430/2006 Sb. v platném znění.

⁴⁴ ETRS je závazným geodetickým referenčním systémem na území ČR dle nařízení vlády č. 430/2006 Sb. v platném znění, a dále závazným souřadnicovým referenčním systémem na území EU dle směrnice INSPIRE.

⁴⁵ WGS84 je závazným geodetickým referenčním systémem na území ČR dle nařízení vlády č. 430/2006 Sb. v platném znění.

⁴⁶ S-42/83 je závazným geodetickým referenčním systémem na území ČR dle nařízení vlády č. 430/2006 Sb. v platném znění.

⁴⁷ [https://geoportal.cuzk.cz/\(S\(gcktcmwdvexj4bomil3luntm\)\)/Default.aspx?mode=TextMeta&side=sit.trans&text=souradsystemy](https://geoportal.cuzk.cz/(S(gcktcmwdvexj4bomil3luntm))/Default.aspx?mode=TextMeta&side=sit.trans&text=souradsystemy)

⁴⁸ Srv. poznámka. 27 a [Specification for digital geospatial data records archiving, verze 3.0.0 z 31. 8. 2021, dostupné z CITS-Geospatial/CITS Geospatial v3.pdf at master · DILCISBoard/CITS-Geospatial · GitHub](#), str. 26.

s reprezentací prostorových dat, která je strukturována dle typů dat a podadresáře jsou nazvány dle mimetype objektů v nich uložených. Schéma pro data obsahuje obrázek 6.

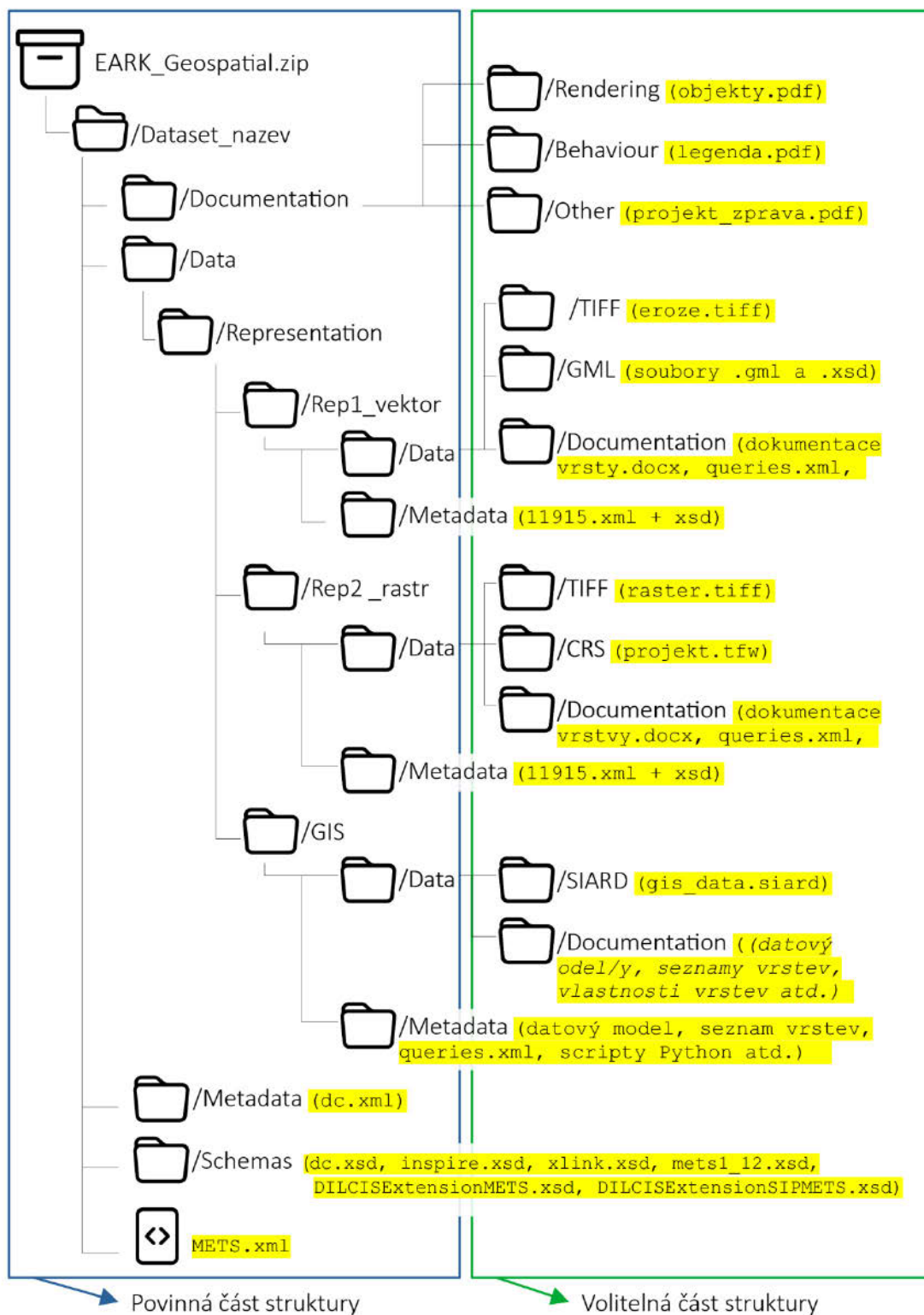
3.1.8. Databáze a tabulková data

Jako databázi označujeme strukturovanou, ale i nestrukturovanou sestavu dat v počítačové podobě, která může být uložena v jediném nebo více počítačích, ale dostupná různými způsoby.⁴⁹ V oblasti prostorových dat se prakticky výhradně využívají databáze pro strukturovaná data a jsou de facto analytickým jádrem systémů GIS.

Pro práci s prostorovými daty lze využít širokou škálu databázových prostředí, z nichž mnohá jsou dostupná jako open source software a pro specifickou práci s prostorovými daty je stačí doplnit vhodným rozšířením (plug-in). V České republice je hojně využíváno prostředí PostgreSQL+PostGIS, Oracle-Spatial, ale v podstatě všechny běžně využívané relační databáze lze vybavit rozšířením a využít je pro práci s prostorovými daty.

V případě malých jednoduchých datasetů lze předpokládat, že nebude třeba uložení rozsáhlého systému dat a bude zcela dostatečné exportovat datové sestavy v podobě souborů CSV, tedy postup, který se běžně využívá pro datové migrace. Tabulkové sestavy je však nutné řádně validovat a ověřit, zda bylo využito kódování v široce využívaném standardu UTF-8. Na datovém příjmu bude datová validita tabulkových dat vždy kontrolována. Vzhledem k očekávané datové obnově a interoperabilitě v dlouhodobém horizontu je datová validita zcela kardinální požadavek.

⁴⁹ SIARD, možnost archivace relačních databází, Rehtorik, M. in Archivní časopis 2/2020, str. 128–154.



Obrázek 6. Prostorová data ve struktuře EARK_Geospatial obsahující vektorová data, rastrový podklad a související informace v databázové části GIS

3.1.9. Documentation

Podadresář /Documentation je povinný jak na kořenové úrovni, tak na úrovni datasetových reprezentací. Obsahuje doporučenou adresářovou strukturu, která obsahuje podadresáře:

/Structure s popisem originální struktury datasetu/ů; např. vztahů mezi dvěma nebo více složkami datasetu tak, aby bylo možné rekonstruovat původní funkcionality. Například rastrový objekt a jeho asociace na souřadnicový soubor (world file), vektorová datová sada kombinovaná s tabulkou, popis struktury vrstev geodat pro tvorbu mapy, konfiguraci webových služeb, definování mash-up WMS atd. Tyto informace by měly nejlépe poskytnuty pomocí standardizovaných strojově čitelných souborů.

/Rendering pro data popisující podobu zobrazování dat; jakékoli informace, které přispívají k obnovení informačního objektu, např. soubor ve formátu .sld, (barevná mapa hodnot pixelů rastru; styled layer descriptor pro webové služby), dokumentace popisující projekt kartografické mapy, návrhy zpráv atd. Adresář není povinný, měl by být obsažen podle potřeby buď na úrovni celého balíčku, nebo pouze na úrovni reprezentace.

/Behaviour pro popis chování, resp. získávání dat, např. v podobě používaných SQL Queries, algoritmů vykreslování, dokumentace funkčnosti systému, analýzy funkcionalit, uživatelské příručky, školicí materiály, videa o používání systému aj. Adresář není povinný.

/Other v případě potřeby dataset doplnit o další neupřesněnou, ale důležitou dokumentaci datasetu/ů. Je doporučeno, aby dokumentace obsažená v tomto adresáři byla ve výstupních formátech dle § 23 bodu (1) odst. c) vyhlášky č. 259/2012 Sb. Tento Adresář není povinný.

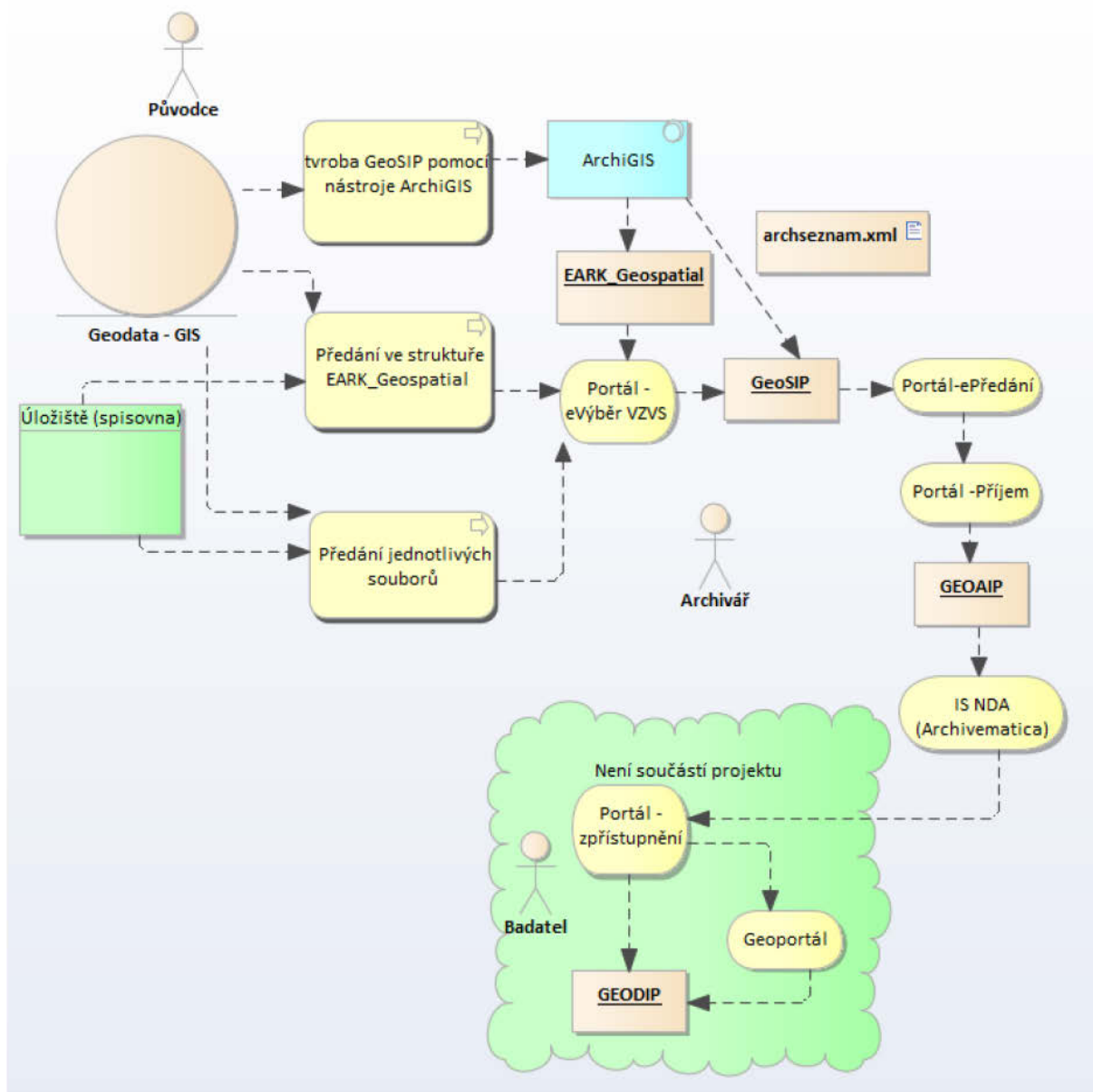
3.1.10. Schemas

Podadresář Schemas je povinný v případě použití XML schémat, musí v něm být umístěna všechna XML schémata geoprostorových metadat použitých v balíčku. Povinným obsahem adresáře /Schemas pro EARK_Geospatial jsou soubory dle specifikace CITS-Geospatial⁵⁰ (aktuálně: dc.xsd, inspire.xsd, xlink.xsd, mets1_12.xsd, DILCISExtensionMETS.xsd, DILCISExtensionSIPMETS.xsd).

⁵⁰ Srv. poznámka 39.

4. Postup výběru a ukládání prostorových dat⁵¹

Výběr a předání GIS a prostorových dat je nutno provádět v souladu s § 3, § 4 a § 7–15 Zákona a dle doporučení této metodiky.



Obrázek 7 - Schéma výběru a předávání prostorových a GIS dat do národního digitálního archivu a využití informačních balíčků

⁵¹ Doporučený postup vychází ze situace platné v listopadu 2021. Avšak právě na základě této metodiky se připravuje automatizace tohoto procesu jak celku v podobě nástroje ArchiGIS, který vyvíjí VÚMOP pro export datového balíčku, ale také jeho jednotlivých částí na úrovních Národního archivního portálu v podobě validátorů a nástrojů pro konverzi metadat.

Výběr a předání prostorových dat archivu (celkový přehled):

- › Původce (tvůrce, vlastník) dat nejprve osloví příslušný archiv, kterému **nabídne data** k výběru archiválií s jejich popisem, např. datovou zprávou pomocí Informačního systému datových schránek. Popis nabízeného datasetu musí vždy obsahovat: Název, stručný popis obsahu, časový rozsah, druh systému, z kterého byl/bude export proveden, formát/y datasetu. Archiv na základě popisu určí data k uložení do archivu a dohodne s původcem podrobnosti exportu, způsob předání dat k výběru či přímo k uložení do digitálního archivu, tj. která data budou vybrána, v jaké struktuře je lze připravit, v jakých formátech mají být předána.

Nejvhodnějším způsobem je příprava dat v podobě GeoSIP balíčku

Pokud se jedná o data, která jsou průběžně aktualizovaná, je vhodné určit vhodnou periodicitu předávání datasetů. Např. nový legislativní předpis, nové technologické zpracování, zásadní změny ve sledované oblasti atd.

Po ukončení výběru archiválií archiv vyhotoví protokol o výběru archiválií.

- › Původce **provede export** dat podle dohodnuté strategie. Buď vytvoří datové balíčky GeoSIP, strukturu EARK_Geospatial, nebo připraví data pro předání ve formě volných souborů (srv. bod 4.3 této metodiky). Zároveň podle domluvy provede převod formátů dle platných formátových pravidel NA.

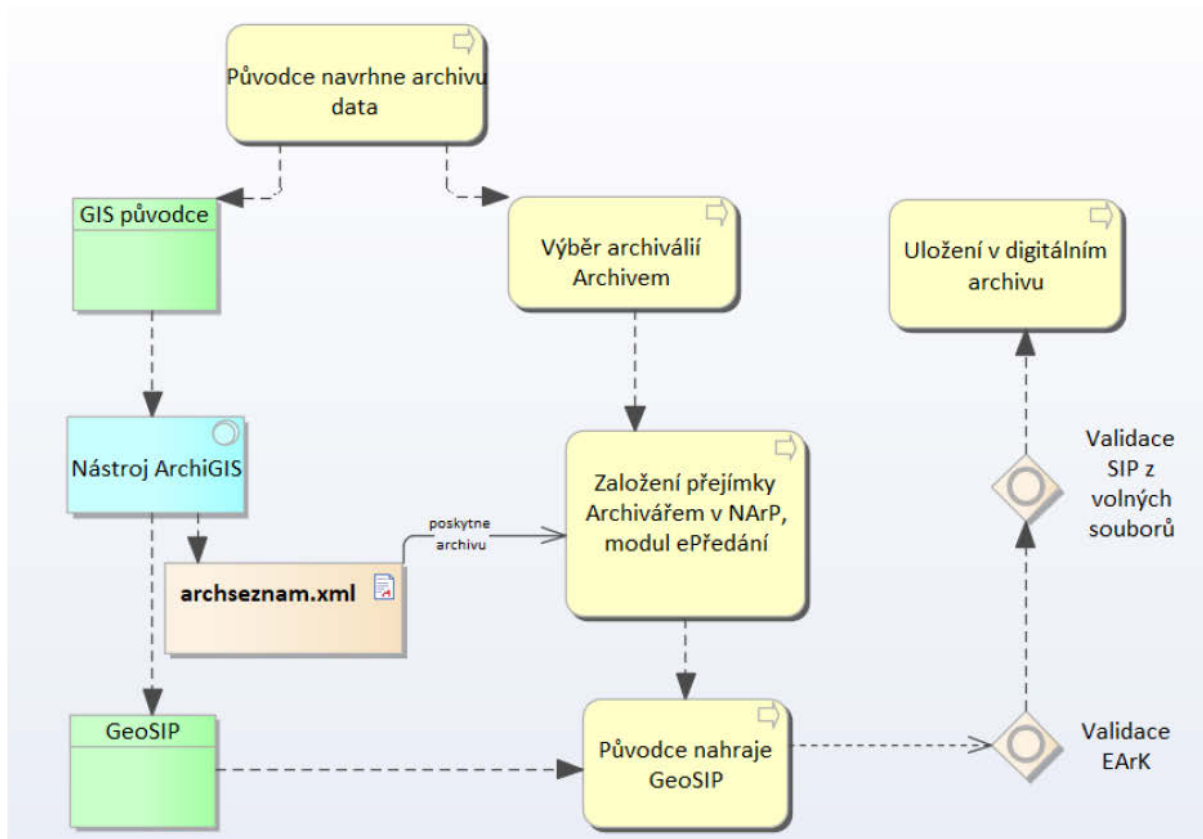
Předává-li původce data ve formě GeoSIP balíčků, je doporučeno provést datovou validaci např. zpětným importem nebo jiným nástrojem.

- › Digitální dokumenty a data se předávají k uložení v národním digitálním archivu prostřednictvím **Národního archivního portálu**⁵². Archivář připraví prostor pro nahrání dat, příp. pro jejich opatření metadaty, a původce data nahraje.
- › Následně archivář provede zařazení do archivního souboru a data jsou uložena v LTP systému digitálního archivu.
- › Po předání k **trvalému uložení** vyhotoví archiv úřední záznam o předání archiválií, kterým se stvrzuje, že dataset byl vybrán za archiválii a uložen do digitálního archivu. Součástí záznamu je jednoznačný identifikátor AIP. Datová autenticita, integrita a důvěryhodnost je zajištěna systémem digitálního archivu.

V následujících bodech jsou popsány možné scénáře předání prostorových dat.

⁵² <https://portal.nacr.cz>; testovací instance je dostupná na <https://portaltest.nacr.cz>.

4.1. Předání ve formě GeoSIP



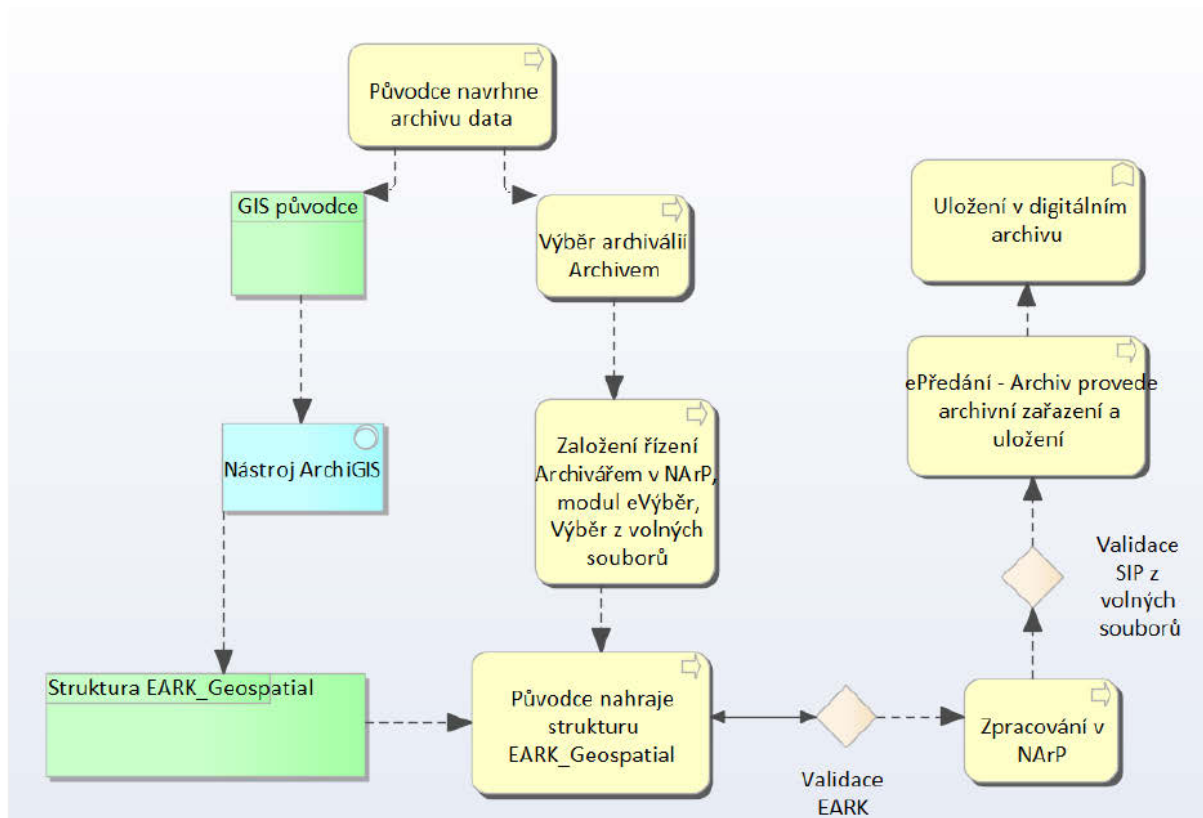
Obrázek 8 - Předávání dat ve formě GeoSIP

Původce pomocí nástroje ArchiGIS a prostorových dat do podoby datového balíčku SIP vytvoří jeden či více GeoSIP a soubor `archseznam.xml`. GeoSIP je variantou SIPu pro výběr z volných souborů a pro jeho vytvoření bude původce k řádnému vyplnění metadat řízení potřebovat od příslušného archivu: číslo jednacích výběrů, datum uzavření výběru a kód původce z databáze Interpi. Soubor `archseznam.xml` poskytne archivu, který v NArP v modulu ePředání připraví tzv. přejímku, do které původce nahraje připravený GeoSIP/y.

Archivář do struktury balíčků nezasahuje. Balíček GeoSIP je podroben validaci a uložen v digitálním archivu.

- › *Postup je určen pro živé, plně podporované systémy. Je vhodné jej zařadit do standardního workflow dat. Takto lze předávat více GeoSIP najednou.*

4.2. Předání ve formě datové struktury EARK_Geospatial



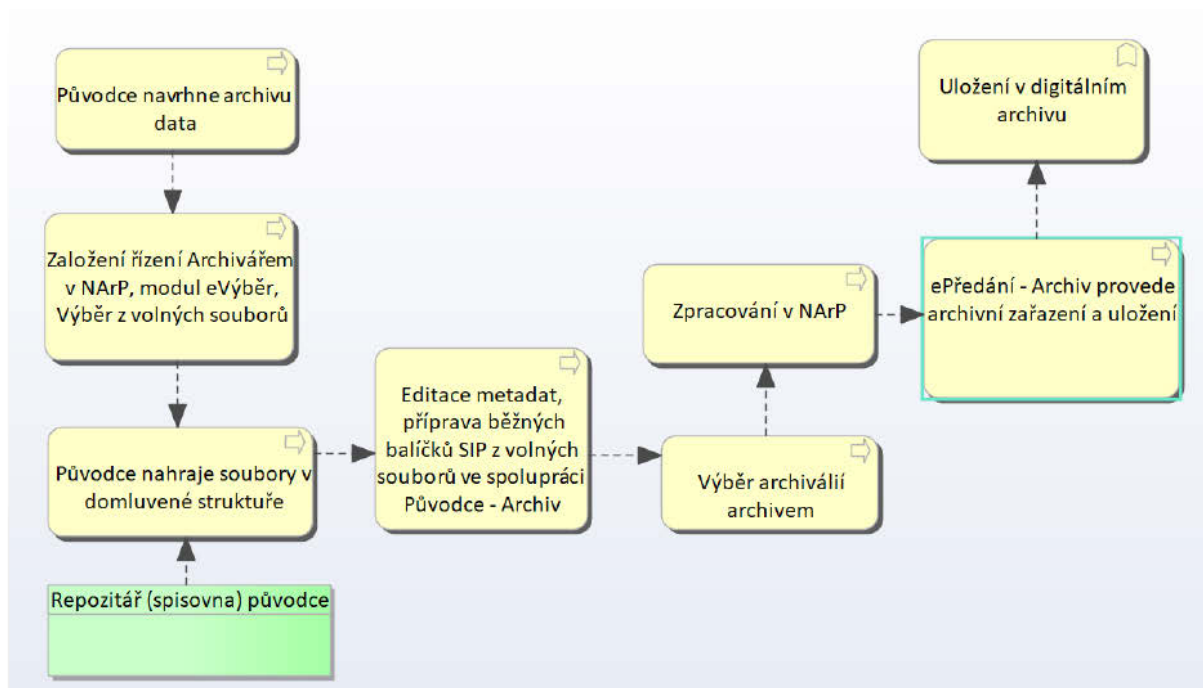
Obrázek 9 - Předání dat ve formě EARK_Geospatial

Původce pomocí nástrojů (Nástroj ArchiGIS, [Roda-In](#), [eatb - eArchiving Tool Box](#)) vytvoří strukturu [EARK Geospatial](#), který nahraje prostřednictvím NArP v modulu eVýběr formou nahrávání z volných souborů. V modulu eVýběr následně původce nebo archivář (podle vzájemné dohody) vytvoří balíček, do kterého nahraný EARK_Geospatial vloží.

Archivář ani původce do struktury balíčku opět nezasahují. Archivář dokončí provedení výběru archiválií, portál EARK_Geospatial transformuje na GeoSIP, ten je po validaci předán k uložení.

- › *Postup je určen pro živé, podporované systémy. Tímto způsobem lze předávat k výběru více GeoSIP najednou, je možné připojit základní popisná metadata na úrovni celého GeoSIP balíčku. Lze využít, pokud původce balíčky vytváří mimo nástroj ArchiGIS. Je vhodné předání dat k trvalému uložení zařadit do standardního workflow dat.*

4.3. Předání ve formě volných souborů v připravené struktuře



Obrázek 10 - Předání dat mimo GeoSIP balíčky

Původce si připraví data do struktury blízké EARK_Geospatial, ideálně uloží prostorová data do adresáře `nazev_GeoSIP/reprezentace/data/`, metadata datasetu `nazev_GeoSIP/reprezentace/metadata/` a související dokumentaci do adresáře `nazev_GeoSIP/dokumentace`. Archivář založí v NArP v modulu eVýběr řízení pro výběr z volných souborů. Původce nahraje připravený dataset a takto připravený dataset musí být popsán v prostředí NArP, kde bude vytvořen běžný balíček SIP z volných souborů. Archivář následně dokončí provedení výběru archiválií a vytvořený balíček SIP je předán k uložení.

- › *Postup je určen pro zastaralé, nepodporované systémy, kde nelze vytvořit jeden z doporučených GeoSIP, příp. nabízí-li původce data uložená mimo GIS. Tímto způsobem je možné připojit základní popisná metadata na úrovni celého balíčku, v rámci výběru archiválií je možné vybrat pouze určité soubory k trvalému uchování.*

5. Validace datová a metadatová

Validace datových balíčků je aktuálně plánována ve dvou úrovních. Na první úrovni bude nejprve validován datový balíček typu GeoSIP validátorem NAČR pro SIPy z volných souborů, ten prochází vývojem a kontroly budou definovány na základě této metodiky a získaných zkušeností. Druhá úroveň validace bude prováděna na úrovni datového balíčku typu EARK_Geospatial, který je vyvíjen pro konsorcium E-ARK pod dohledem DILCIS Board. Po jeho zveřejnění bude implementován do pracovních postupů v rámci trvalého ukládání prostorových dat.

6. Ověření postupu na příkladu GIS - data aplikace eKatalog BPEJ

6.1. Popis zdrojových dat aplikace

Aplikace eKatalog BPEJ (Bonitovaná půdně ekologická jednotka) interpretuje jednotlivé části kódu BPEJ a jednoduchou formou vysvětluje informace uvedené ve vyhlášce Ministerstva zemědělství ČR č. 227/2018 Sb. ze dne 4. října 2018, kterou se stanoví charakteristika BPEJ a postup pro jejich vedení a aktualizaci. Dále pak řeší fiskální účely využití dat BPEJ, jejich využití pro realizaci Společné zemědělské politiky a využití při realizaci ochrany zemědělského půdního fondu.

eKatalog BPEJ slouží k hodnocení absolutní i relativní produkční schopnosti zemědělských půd a podmínek jejich nejúčelnějšího využití. BPEJ je charakterizována pětímístným kódem. První číslice vyjadřuje příslušnost ke klimatickému regionu. Druhá a třetí číslice určuje zařazení půdy do hlavní půdní jednotky klasifikační soustavy (HPJ). Čtvrtá stanovuje stupeň sklonitosti a příslušnou expozici ke světovým stranám ve vzájemné kombinaci, pátá číslice vyjadřuje hloubku půdy a skeletovitost půdního profilu ve vzájemné kombinaci. Základní soustava vymezuje 2140 BPEJ, pro které jsou k dispozici i ekonomické charakteristiky a nově vymezených 138 kódů, pro které je nutné nejprve ekonomické charakteristiky vyhodnotit. Dohromady tedy 2278 kódů BPEJ.⁵³

6.2. Příprava datových zdrojů

VÚMOP se pomocí aplikace QGIS Desktop LTR 3.14-Pí připojil k datům uloženým v systému GIS, provedl datový export do formátu .gml, přiložil metadata ve formátu .xml a .pdf a zaslal Národnímu archivu. Pilotní datovou sadu tvořilo celkem 33 objektů, z toho 15 souborů ve formátu .gml, 15 schémat .xsd, 1 soubor .jpg zobrazující schéma uložení dat v databázi a popisaná metadata ve formátu .xml a .pdf.

6.3. Tvorba GeoSIP balíčku a datová validace

Datový balíček byl vytvořen nejprve ve formě doporučené struktury EARK_Geospatial pomocí nástroje Roda-In tool v.2.4.0 a byla manuálně vyplněna metadata dle schématu Dublin Core.⁵⁴

Datová struktura byla následně nahrána do testovací instance Národního archivního portálu, kde byl opět manuálně vytvořen datový balíček GeoSIP v řízení „z volných souborů“, kdy kompletní data datasetu měla velikost 2,353 GB, struktura EARK_Geospatial v komprimované podobě .zip dosahovala velikosti 660 MB. Nahrání dat na portál bylo bezproblémové, ale závislé na rychlosti připojení. Soubor `CZNDA_potvrzeni_nahranych_dat.csv` potvrzující úspěšné nahrání dat byl řádně vygenerován.

V rámci pracovních činností v prostředí Národního archivního portálu byla ověřena možnost stáhnout si kterýkoliv objekt k nahlédnutí a ověření datové validity a poté byla vytvořena přejímka dat do digitálního archivu.

Datová validace byla testována formou náhledu exportovaných prostorových dat v nástroji FME Data Inspector v. FME(R) 2021.1.2.0 (20210903 - Build 21627 - WIN64) a v aplikaci QGIS

⁵³ <https://bpej.vumop.cz>

⁵⁴ <https://github.com/keeps/roda-in/releases>

Desktop LTR 3.16.4-Hannover, parametry testovací PC sestavy: Intel Pentium Core i5 – 3470 @ 3,20GHz, RAM 16 GB, 64 Bit Windows 7, SP 1.

Největší objekt datasetu dosahoval velikosti 1,74 GB a prohlížení takto uložených dat bylo možné jen v omezené podobě s velkým prodlením, které bylo způsobeno načítáním. Validita jednotlivých objektů byla dále ověřena pomocí nástrojů Altova XMLSpy Enterprise Edition SP1 2018 (x64) a Notepad++ v. 8.1.5 (64bit) s rozšířením XML Tools, které strojově validují data obsažená v souborech formátu GML vůči příslušnému schématu XSD. Metadatový soubor ve formátu PDF/A byl validován pomocí validátoru zveřejněného na webových stránkách Národního archivu.⁵⁵

Po ukončení výběru ze strany Národního archivu byl řádně vygenerován protokol o výběru archiválií a byla vytvořena testovací přejímka. Datový balíček GeoSIP byl následně zařazen a předán k uložení do systému Národního digitálního archivu.

⁵⁵ <https://digi.nacr.cz/validatorPDF/>

7. Závěr

Metodika Transformace digitálních prostorových dat pro účely trvalého uložení v digitálním archivu implementuje schválenou strukturu informačního balíčku pro trvalé ukládání prostorových dat, kterou připravilo evropské konsorcium archivů EARK za podpory Evropské komise.⁵⁶ Tím ji uvádí do souladu s platnou českou legislativou, aby prostorová data mohla být trvale ukládána v podobě kompatibilní, jak s dosavadními postupy a již vybudovanou infrastrukturou, tak i s ohledem na možnosti budoucího využití nejenom v České republice.

Fakticky se jedná o první definici pro komplexní datové sety v České republice, která umožní hlubší strukturální, formátovou a datovou validaci. Představuje a stanovuje nové postupy, zejména s ohledem na maximálně automatizovaný proces tvorby a příjmu datových balíčků do digitálního archivu. Metodika dále navrhuje transformaci metadatových prvků používaných v oblasti správy a sdílení prostorových dat na prvky používané k popisu v archivních informačních systémech tak, aby byly v souladu se Základními pravidly pro zpracování archiválií⁵⁷ a došlo k uchování těchto popisných informací v maximální možné míře. Metodika se též zabývá formáty, které jsou využívány pro výměnu nebo pro krátkodobé uchování dat z pohledu jejich možného využití pro trvalé uchování a navrhuje možné alternativy formátových migrací pro formáty, které jsou proprietární a nelze u nich zaručit jejich dlouhodobou čitelnost, což je jednou z klíčových otázek digitálního archivnictví.

Metodika je založena na nejnovějších poznatcích, čerpá z odborných zdrojů a pokouší se propojit doposud sobě vzdálené disciplíny. Do oblasti prostorových dat, která je exaktní a zaměřená na přírodní vědy a poznání a popis těchto fenoménů, přináší archivnictví humanitní prvek, kterým je uchování cenných informací vytvořených člověkem. To se sice již dělo v minulosti, ale postupující a vysoce dynamická digitalizace naší společnosti umožňuje zpracování informací v detailech, které v době analogové prakticky nebyly možné, a tento rozvoj dále postupuje. Tím se sice zvyšuje hodnota a objem těchto informací, ale na druhé straně vytváří problém, jak tyto informace dlouhodobě nejenom uchovat, ale především rychle a efektivně poskytovat. Přitom právě předpokládaná interoperabilita bude klást zvýšené nároky na znalosti a schopnosti archivářů, bude muset dojít k hlubšímu propojení zmíněných disciplín i na praktické úrovni. Naopak tvůrcům prostorových dat musí být srozumitelně vysvětlen význam a důležitost této transformace pro potřeby trvalého uložení, neboť úplná datová obnova mnohdy nebude možná. Již nyní si to mnozí původci a tvůrci prostorových dat uvědomují nezávisle na tom, zda jsou z veřejné nebo soukromé sféry. Metodika však nemůže být neměnným dogmatem, naopak je to výchozí bod, který bude nutno průběžně revidovat a upravovat dle nových poznatků, technologického rozvoje, vylepšených postupů a úprav v legislativě.

⁵⁶ Srv. poznámka 39.

⁵⁷ <https://www.mvcr.cz/soubor/zakladni-pravidla-pro-zpracovani-archivalii-2015-cervene-vyznaceny-mi-zmenami.aspx>

8. Informační zdroje

8.1. Odborná literatura

DPC Zpráva o sledování technologií – Uchování CAD (2013), Dostupné z: <https://www.dpconline.org/docs/technology-watch-reports/896-dpctw13-02-pdf/file>

DPC Zpráva o sledování technologií – Uchování GIS (2009), Dostupné z: <https://www.dpconline.org/docs/technology-watch-reports/2265-twr-09-01-gis-mcgarva/file>

DPC Zpráva o sledování technologií – JPEG 2000 jako archivní formát (2008),

JPEG 2000 - a Practical Digital Preservation Standard? DPC Technology Watch Report 08-01 2008, Dostupné z: <http://www.dpconline.org/docs/technology-watch-reports/87-jpeg-2000-a-practical-digital-preservation-standard/file>

LEŠKA, Rudolf: Archivy a autorské právo – symbióza nebo kolize? In: Archivní časopis. Praha: Ministerstvo vnitra, 2016, roč. 66, č. 2, str. 5–28. ISSN 0004-0398

KALINA, Tomáš: Elektronické dokumenty – archiválie nového tisíciletí, in: Archivy a zítřek, Sborník příspěvků z konference uspořádané ve dnech 22.–23. září 1999, Jindřichův Hradec 2000, str.111–117.

KOLÁŘ, Jan: Geografické informační systémy 10, ČVUT, Praha 2003

Kol. autorů: Digitální archivnictví, Brno 2019, dostupné on-line <https://munispace.muni.cz/library/catalog/book/1407>

KUNT, Miroslav (red.). Možnosti a formy zpřístupnění archivních fondů nebo jejich součástí veřejnosti v elektronické podobě. Výzkumná zpráva projektu výzkumu a vývoje VE20072009004. Praha 2010. Dostupné z: http://www.nacr.cz/wp-content/uploads/2015/11/moznosti_02.pdf Cit. 22. 11. 2018

OWENS, T.: The Theory and Craft of Digital Preservation, 2018

Preserving Raster Images; Data Types Series, Artefactual Systems and the Digital Preservation Coalition (2021) Dostupné z: <https://www.dpconline.org/docs/technology-watch-reports/2478-preserving-raster-images/file>

Preserving GIS, Data Types Series Artefactual Systems and the Digital Preservation Coalition (2021) Dostupné z: <https://www.dpconline.org/docs/technology-watch-reports/2481-preserving-gis/file>

RECHTORIK, M: SIARD, možnost archivace relačních databází, in Archivní časopis 2/2020, str. 128–154

ŠTYCH, Přemysl a kol.: Vybrané funkce geoinformačních systémů, CITT Praha, Akademie kosmických technologií, Oblast Galileo, GMES, z podpory ESF, MPSV, Praha 2008

YEO, G.: Records, information and data, London 2018

8.2. Normy, směrnice, metodiky

ČSN EN ISO 19101-1. Geografická informace – Referenční model. Část 1: Základy. Praha: Úřad pro technickou normalizaci, metrologii a státní zkušebnictví, leden 2018

ČSN 19111 (979830) Geografická informace – Vyjádření prostorových referencí souřadnicemi

ČSN ISO 14721. Systémy pro přenos dat a informací z kosmického prostoru – Otevřený archivační informační systém – Referenční model. Praha 2014

Dokumentace NDA, verze 1.13, interní dokument

Dublin Core, dostupné z: <https://www.dublincore.org>

EAD (Encoded Archival Description), dostupné z: <https://www2.archivists.org/groups/technical-subcommittee-on-encoded-archival-standards-ts-eas/encoded-archival-description-ead>

E-ARK CSIP Common Specification for Information Packages, verze 2.0.4, publikováno 12. 6. 2020, Dostupné z <https://earkcsip.dilcis.eu>

Guideline for CITS Geospatial, Appendix 3, Mapping of INSPIRE metadata descriptions to archival descriptive metadata standards, dostupné z: https://citsgeospatial.dilcis.eu/guideline/Guideline_CITS_Geospatial_Appendix_3_v1.pdf

ISO/IEC 15444-1:2004. Information technology -- JPEG 2000 image coding system -- Part 1: Core coding system, Annex I: JP2 file format syntax (formal name) JPEG 2000 jp2 file format (common name)

Metodický pokyn č. 2/2020 pro evidenci a zpřístupňování map v archivech České republiky. Dostupné z: <https://www.mvcr.cz/soubor/metodicky-pokyn-c-2-2020-pro-evidenci-a-zpristupnovani-map-v-archivech-ceske-republiky.aspx>

Národní standard pro elektronické systémy spisové služby, In: Věstník MVČR č. 57/2017, Dostupné z: <https://www.mvcr.cz/soubor/vestnik-mv-57-2017-oznameni-ministerstva-vnitraterym-se-zverejnuje-narodni-standard-pro-elektronicke-systemy-spisove-sluzby.aspx>

Provozní řád NDA, verze 1.0, interní dokument

Směrnice Evropského parlamentu a Rady 96/9/ES ze dne 11. března 1996 o právní ochraně databází. Dostupné z <https://eur-lex.europa.eu/legal-content/cs/TXT/?uri=CELEX:31996L0009>

Směrnice Evropského parlamentu a Rady 2007/2/ES ze dne 14. března 2007 o zřízení Infrastruktury pro prostorové informace v Evropském společenství (INSPIRE). Dostupné z <http://data.europa.eu/eli/dir/2007/2/oj>, Cit. 28. 8. 2021

Specification for digital geospatial data records archiving, verze 3.0.0, publikováno 31. 8. 2021, Dostupné z <https://github.com/DILCISBoard/CITS-Geospatial.git>

Spezifikation Geo-SIP und Geo-Dossier, dostupné z: <https://www.bar.admin.ch/bar/de/home/archivierung/digitale-unterlagen/archivierung-von-geodaten.html> ; <https://www.bar.admin.ch/dam/bar/de/dokumente/kundeninformation/Spezifikation%20Geo->

[SIP%20und%20Geo-Dossier.pdf.download.pdf/Spezifikation_Geo-SIP_und_Geo-Dossier_V1.0_2016-12-05.pdf](#)

Vyhláška č. 259/2012 Sb., o podrobnostech výkonu spisové služby

Vyhláška č. 645/2004 Sb., kterou se provádí některá ustanovení zákona o archivnictví a spisové službě a o změně některých zákonů

Základní pravidla pro zpracování archiválií, metodický pokyn č. 1/2021
<https://www.mvcr.cz/soubor/zakladni-pravidla-pro-zpracovani-archivalii-ver-3-0-priloha-metodickeho-pokynu-c-1-2021.aspx>

Zákon č. 499/2004 Sb., o archivnictví a spisové službě

8.3. Webové zdroje

Datasets, otevřená data a jejich trvalé ukládání. Dostupné z: http://ipres2013.sysresearch.org/t3_1-iPresDataTutorial.pdf

Datový balíček SIP z volných souborů včetně GeoSIP (aktuální popis a schéma). Dostupné z: <http://www.nacr.cz/sipy>

DROID, dostupné z: <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>

eKatalog BPEJ (Bonitovaná půdně ekologická jednotka) poskytovaný veřejnosti VÚMOP, dostupné z: <https://bpej.vumop.cz>

Evropská unie a její strategie v oblasti dat, dostupné z: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

FADGI – metodika pro digitalizáty (rastrové soubory) (září 2016), dostupné z: http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image_Tech_Guidelines_2016.pdf

Formátová doporučení Kongresové knihovny v oblasti GIS, geoprostorových a dalších kartografických formátů, dostupné z: <https://www.loc.gov/preservation/resources/rfs/geocarto.html> Cit. 26. 8. 2021

Formát Shapefile (Proč musí Shapefile umřít. Jáchym Čepický, 2015), dostupné z: https://www.cagi.cz/upload/documents/givs2015/cepicky_shapefile_must_die.pdf

Formáty – porovnání GeoPackage vs. Shapefile, dostupné z: <http://switchfromshapefile.org>,
<https://www.gis-blog.com/geopackage-vs-shapefile/>,
<https://www.cagi.cz/upload/documents/givs2018/30-givs-2018-cepicky.pdf>

Geografická Informace, dostupné z: https://aleph.nkp.cz/F/?func=direct&doc_number=000000397&local_base=KTD

Geopackage Validátor, dostupné z: <https://github.com/PDOK/geopackage-validator>

GIS, dostupné z: <https://www.britannica.com/technology/GIS>,
https://aleph.nkp.cz/F/BIKN68EGRYMF15RBDVKIBQ5N6QGCA6JNKA1DMDHTVTMLJ13H9G-37562?func=find-b&find_code=WTD&x=0&y=0&request=gis&adjacent=N

INSPIRE – metadata a katalogové služby, dostupné
z: <https://inspire.ec.europa.eu/training/metadata-and-catalogue-services>

INSPIRE – terminologie, dostupné
z: [https://geoportal.cuzk.cz/\(S\(mkpsvthtyg3yj2fxvbncrctb\)\)/Default.aspx?lng=CZ&mode=TextMeta&text=INSPIRE terminologie&head_tab=sekce-04-gp&side=INSPIRE pristup](https://geoportal.cuzk.cz/(S(mkpsvthtyg3yj2fxvbncrctb))/Default.aspx?lng=CZ&mode=TextMeta&text=INSPIRE%20terminologie&head_tab=sekce-04-gp&side=INSPIRE_pristup)

JPG a JPEG2000, dostupné z: <https://knihovnavrue.nkp.cz/archiv/2020-2/recenzovane-prispevky/posuzovani-souborovych-formatu-z-hlediska-dlouhodobeho-uchovavani-a-navrh-metodiky-pro-narodni-knihovnu-ceske-republiky>; JPEG (Joint Photographic Experts Group Interchange Format) a JPEG 2000 (Joint Photographic Experts Group 2000); <https://www.dpconline.org/docs/technology-watch-reports/2478-preserving-raster-images/file>

MICKA, dostupné z: <https://micka.bnhelp.cz/>

Národní archiv, dostupné z: www.nacr.cz

Česká implementace EAD, prozatím interní dokument

Národní archivní portál v produkční a testovací instanci, dostupné z: <https://portal.nacr.cz>;
<https://portaltest.nacr.cz>

OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 2: SQL, dostupné z: https://portal.ogc.org/files/?artifact_id=25354

Otevřená data ČR, dostupné z: <https://ofn.gov.cz/prostorova-data/2019-08-22/#uvod>

PostgreSQL: Documentation: 14: 38.6. Function Overloading dostupné
z: <https://www.postgresql.org/docs/14/xfunc-overload.html>

Prostorová data (Lukáš Holman, Jan Kříž), dostupné
z: <https://www.natur.cuni.cz/geografie/geoinformatika-kartografie/ke-stazeni/projekty/moderni-geoinformacni-metody-ve-vyuce-gis-a-kartografie/prostorova-data/>

Roda-In, nástroj pro tvorbu datových balíčků, dostupné z: <https://github.com/keeps/roda-in/releases>

Souřadnicové systémy (ČUZAK), dostupné
z: [https://geoportal.cuzk.cz/\(S\(gcktcmwvdxj4bomil3luntm\)\)/Default.aspx?mode=TextMeta&side=sit.trans&text=souradsystemy](https://geoportal.cuzk.cz/(S(gcktcmwvdxj4bomil3luntm))/Default.aspx?mode=TextMeta&side=sit.trans&text=souradsystemy)

Souřadnicové systémy (Jan Šimbera), dostupné
z: <https://www.natur.cuni.cz/geografie/geoinformatika-kartografie/ke-stazeni/projekty/moderni-geoinformacni-metody-ve-vyuce-gis-kartografie-a-dpz/souradnicove-systemy/?fbclid=IwAR2pj9HOam25QZY9fVU7YzjOW3c5HYa2ioOGrjQ7DXwlczGly2oXUIfxR0o>

Validátor datových balíčků SIP podle NSESSS, dostupné z: <https://validatorsip.nacr.cz>

Validátor PDF/A NAČR, dostupné z: <https://digi.nacr.cz/validatorPDF/>

Výzkumný ústav meliorací a ochrany půdy, v.v.i., dostupné z: <https://www.vumop.cz/>

Zpráva k možnostem digitální archivace prostorových dat z roku 2009, dostupné z: <https://www.dpconline.org/docs/technology-watch-reports/2265-twr-09-01-gis-mcgarva/file>

Webové zdroje a specifikace formátů jsou uvedeny v profilu konkrétního formátu.

Všechny použité zdroje byly dostupné ke dni 15. 11. 2021 a byly uchovány v dlouhodobě čitelné formě jako součást dokumentace této metodiky.

Grafické prvky získány z: <https://icons.getbootstrap.com/>

8.4. Použité veřejné datové zdroje

Topografická databáze České republiky (Data200) – vrstva vodstvo, dostupné z: [https://geoportal.cuzk.cz/\(S\(t1u0bztrcdopsvdob5d2ktcv\)\)/Default.aspx?mode=TextMeta&metadataID=CZ-CUZK-DATA200-VODSTVO-V&metadataXSL=Full&side=mapy_data200](https://geoportal.cuzk.cz/(S(t1u0bztrcdopsvdob5d2ktcv))/Default.aspx?mode=TextMeta&metadataID=CZ-CUZK-DATA200-VODSTVO-V&metadataXSL=Full&side=mapy_data200)

Austrálie, správní hranice – ASGS (2016 Edition) – Boundaries, dostupné z: <https://www.data.gov.au/data/dataset/asgs-2016-edition-boundaries/resource/48a33591-0b0a-4712-a60b-39ef4b564f2b>

Charakteristiky toků a povodí ČR – Jizera DIBAVOD , dostupné z: <https://www.dibavod.cz/24/charakteristiky-toku-a-povodi-cr.html>

Příloha č. 1 – Glosář

Archiválie	Dokument, který byl vzhledem k době vzniku, obsahu, původu, vnějším znakům a trvalé hodnotě dané politickým, hospodářským, právním, historickým, kulturním, vědeckým nebo informačním významem vybrán ve veřejném zájmu k trvalému uchování a byl vzat do evidence archiválií [§ 2, písm. f) Zákona].
Archivní metadata	Jsou popisné informace, které jsou vytěžovány Informačním systémem digitálního archivu.
archseznam.xml	Soubor podle přílohy 4 NSESS obsahující výsledek výběru archiválií. Seznam obsahuje identifikátory kořenové entity v daném GeoSIP. Využití viz kapitola 4.1 předání GeoSIP .
Databáze	Je strukturovaná, ale i nestrukturovaná sestava dat v počítačové podobě, uložená v jediném nebo více počítačích, ale dostupná různými způsoby.
Datový balíček AIP	Informační balíček („Archival Information Package“), který je trvale uchováván v digitálním archivu.
Datový balíček DIP	Informační balíček („Dissemination Information Package“), který je žadateli předložen v případě vyžádání dat z digitálního archivu.
Datový balíček SIP	Informační balíček („Submission Information Package“) určený k exportu nebo přenosu entit, dat, dokumentů nebo datasetů z různých informačních systémů do digitálního archivu.
Datový formát	Datový formát je způsob kódování komponenty, který zajišťuje uložení dokumentu nebo jeho části (částí) pro účely zpracování výpočetní technikou a jeho znázornění. Pojem „datový formát“ se pro účely národního standardu užívá v obdobném významu jako „formát“.
Datový formát akceptovaný	Je takový datový formát, který se archivář rozhodne akceptovat, např. tehdy, kdy originální datový formát nemá alternativu ve výstupních datových formátech podle § 23 Vyhlášky nebo by jeho konverze znamenala významnou datovou ztrátu.
Datový formát výstupní (Výstupní datový formát)	Je datový formát, který je ukotven v legislativě ČR a je vhodný pro trvalé uchování v digitálním archivu. Výstupní formáty jsou zveřejněny v § 23 Vyhlášky č. 259/2012 Sb., o podrobnostech výkonu spisové služby.
Datový set/Dataset	Digitální dataset (datový balíček) je evidenční jednotka označující sadu dat, která byla původně vytvořena nebo spravována prostřednictvím výpočetní techniky a která byla vybrána za archiválii a jejíž replika je uložena v Národním archivu nebo archivu oprávněném k ukládání archiválií v digitální podobě nebo bezpečnostním archivu. Dataset (datový balíček) má obvykle podobu polí a tabulek, ve kterých jsou obsažena a strukturována data (např. spisy,

	výsledky průzkumu, sčítání lidu, inventáře, databáze, prezentace, webové stránky, matematické modely apod.) včetně počtu záznamů a jejich celkové velikosti vyjádřené v bytech. Digitální dataset (datový balíček) je evidenční jednotkou pro nezpracované archiválie.
Digitální	Pojem „digitální“ vyjadřuje způsob zpracování entity představovaný numerickým řetězcem tvořeným čísly „1“ a „0“ (proud bitů) interpretovatelný pomocí výpočetní techniky. Pojem „elektronický“ se pro účely této metodiky užívá obdobně.
Digitální archiv	Je sestava nástrojů, databází a služeb, které zajišťují integritu, autenticitu a dlouhodobou čitelnost digitálních objektů a informací.
Dokument	Dokumentem je každá písemná, obrazová, zvuková nebo jiná zaznamenaná informace, ať již v podobě analogové nebo digitální, která byla vytvořena původcem nebo byla původci doručena [§ 2 písm. e) Zákona].
DROID	Nástroj pro identifikaci formátů, je využíván v rámci NDA. https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/ .
Dublin Core Metadata Set	Základní metadatové schéma skládající se z 15 základních elementů, které je využíváno v různých oborech a je společným jádrem různých metadatových schémat.
EAD (Encoded Archival Description)	Standard XML pro tvorbu strukturovaných archivních pomůcek užívaných v digitálním prostředí založený na archivním popisném standardu ISAD(G). V současné verzi EAD3 jej představuje XML schéma ve třech možných syntaxích: XSD, Relax NG a XSD. Umožňuje zachytit strukturovaný popis obsahu archiválií uvnitř archivních fondů nebo sbírek bez ohledu na jejich rozsah a složitost vnitřní struktury.
Elektronický	Pojem „elektronický“ se pro účely této metodiky používá obdobně jako pojem „digitální“.
Elektronický dokument	Elektronický dokument je jakýkoli obsah uchovávaný v elektronické podobě, zejména jako text nebo zvuková, vizuální nebo audiovizuální nahrávka [čl. 3 bod 35 nařízení Evropského parlamentu a Rady (EU) č. 910/2014 ze dne 23. července 2014 o elektronické identifikaci a službách vytvářejících důvěru pro elektronické transakce na vnitřním trhu a o zrušení směrnice 1999/93/ES].
Export	Export je proces vytvoření repliky elektronických seskupení a dokumentů spojený s vytvořením metadat těchto seskupení a dokumentů, a to za účelem převedení vzniklé repliky do jiného systému. Exportovaná seskupení, dokumenty zůstávají zachovány v původním systému, nejsou tedy na rozdíl od přenosu bezprostředně po jeho

	realizaci smazány. Úspěšný export do digitálního archivu je doplněn záznamem příslušného jednoznačného identifikátoru digitálního archivu.
Geografický informační systém/GIS	Počítačový informační systém, který umožňuje ukládat, spravovat a analyzovat prostorová data.
Geoobjekt	Abstraktní zobrazení jevu reálného světa související s určitou polohou nebo zeměpisnou oblastí.
INSPIRE	Infrastruktura pro prostorové informace v Evropském společenství, pro kterou stanoví směrnice 2007/2/ES obecná pravidla pro zřízení infrastruktury pro prostorové informace v Evropském společenství pro účely politik Společenství v oblasti životního prostředí a politik nebo činností, které mohou mít vliv na životní prostředí. Rozsah prostorových dat je součástí přílohy směrnice 2007/2/ES. Směrnice je v ČR implementována zákonem č. 380/2009 Sb., který novelizuje zákon č. 123/1998 Sb., o právu na informace o životním prostředí, a zákonem č. 200/1994 Sb., o zeměměřičství.
Interoperabilita	Možnost kombinace souborů prostorových dat a vzájemné komunikace mezi službami bez opakovaných ručních zásahů tak, aby bylo dosaženo soudržného výsledku a aby byla zvýšena přidaná hodnota souborů prostorových dat a služeb založených na prostorových datech.
Jednoznačný identifikátor	Jednoznačný identifikátor je znak pevně spojený s entitou zajišťující jeho nezaměnitelnost a jedinečnost. Jednoznačný identifikátor obsahuje vždy označení původce, popřípadě zkratku označení původce, a to ve formě alfanumerického kódu.
Komponenta	Komponentou v digitální podobě se rozumí jednoznačně vymezený proud bitů tvořící počítačový soubor.
Konverze	Konverze je proces transformace jedné nebo více komponent do jiného formátu. Výsledkem konverze je ztvárnění.
LTP systém	Srv. Digitální archiv.
Metadata	Metadata se rozumí data popisující souvislosti, obsah a strukturu dokumentů a jejich správu v průběhu času [§ 2 písm. o) Zákona].
Metadatový katalog	Strukturovaná dokumentace o spravovaných datasetech. Původci pomáhá efektivně spravovat a vyhodnocovat prostorová data a umožňuje uživatelům pohodlnou práci s prostorovými daty.

METS (Metadata Encoding and Transmission Standard)	Standard vytvořený Kongresovou knihovnou, založený na jazyce XML, který slouží pro výměnu popisných, administrativních a strukturálních metadat s ohledem na objekty uvnitř digitálního repozitáře. Umožňuje zahrnout i další metadatová schémata např. Dublin Core, MODS, PREMIS atd. Jako výměnný formát pro tvorbu datových balíčků SIP jej od roku 2012 využívá i Národní standard pro elektronické systémy spisové služby (jeho příloha č. 3).
Model/Náhledová data	V případě datového setu s prostorovými daty se tímto rozumí dokument ve výstupním datovém formátu, který v jednoduché podobě znázorní obsah datasetu, např. v podobě statických obrazových dat ve výstupním datovém formátu PDF/A nebo JPG.
Národní archivní portál/NArP	Archiválie v digitální podobě jsou vybírány, předávány a zpřístupňovány způsobem umožňujícím dálkový přístup prostřednictvím národního portálu pro zpřístupnění archiválií v digitální podobě, jehož správcem je Národní archiv. Portál je dostupný na https://portal.nacr.cz .
OAIS (Open Archival Information System)	Mezinárodně standardizovaný referenční model, užívaný v oblasti uchovávání digitálních dokumentů, je definován normou ISO 14721:2012.
Objekt	Entita spravovaná v informačním systému.
OGC (Open Geospatial Consortium)	Mezinárodní standardizační organizace založená na dobrovolné shodě a za účelem spolupráce na procesu otevřené shody podporující vývoj a implementaci standardů pro geoprostorová data a služby.
Otevřená data/Open data	Jsou informace zveřejňované způsobem umožňujícím dálkový přístup v otevřeném a strojově čitelném formátu, jejichž způsob ani účel následného využití není omezen a které jsou evidovány v národním katalogu otevřených dat. Otevřená data jsou vysoce efektivní způsob zveřejňování informací veřejného sektoru.
PRONOM	PRONOM je on-line informační systém o formátech datových souborů a jejich podpůrných softwarových produktech. Původně vyvinut pro podporu Národního archivu Spojeného království, nyní je veřejně přístupný. Využívá jej nástroj pro identifikaci formátů DROID. http://www.nationalarchives.gov.uk/PRONOM/Default.aspx
Prostorová data	Prostorová data zahrnují všechny typy dat, která jsou prostorově určena, tedy přímo nebo nepřímo odkazují na určitou polohu nebo zeměpisnou oblast. Zpravidla jsou určena geometrickým objektem a prostorovými vztahy s okolím (topologií).

Průvodka NArP	Soubor vznikající v rámci transformace dat portálem, je povinně součástí informačního balíčku. Poskytuje vazbu mezi jednotlivými komponentami v informačním balíčku a jejich popisem a současně obsahuje popisná metadata archiválií ve standardu EAD 3. Tato metadata mohou pocházet ze zdrojového souboru mets.xml a nebo jsou vložena v průběhu předání archiválií. Průvodka je v AIP uložena v adresáři /objects pod názvem pruvodka.xml.
Přenos	Přenos je proces přemístění dokumentů nebo datasetů spolu s jejich metadaty do jiného systému. Účelem přenosu je zejména převést vybrané dokumenty nebo datasety do digitálního archivu.
Původce	Je každý, z jehož činnosti dokument vznikl; za dokument vzniklý z činnosti původce se považuje rovněž dokument, který byl původci doručen nebo jinak předán.
PUID (Pronom Unique Identifier)	Jednoznačný identifikátor souborových formátů v registru PRONOM.
Rastrová data	Jedná se o síť polí v řádcích a sloupcích (tzv. pixelech), kdy pole obsahují hodnotu. Data uložená v rastrové podobě zobrazují vlastnosti nebo hodnoty z reálného světa jako např. eroze krajiny, teplota, úhrny srážek atd.
Soubor METS.xml a jeho elementy	Je soubor, kterým původce předává metadatové informace ve strukturované podobě digitálnímu archivu.
Souřadnicové systémy	Jsou matematická pravidla, která umožňují ke každé hodnotě souřadnic přiřadit polohu objektu a naopak – souřadnicový systém ¹ (coordinate reference system – CRS). Souřadnicových systémů existuje v geoinformatice velké množství, přičemž každý z nich se různou měrou hodí pro práci v různých měřítkách a na různých územích, v závislosti na tom, jaké referenční těleso a případně kartografické zobrazení používá. Řídí se standardem ISO 19111:2019
Vektorová data	Soubor geometrických elementů reprezentující fenomény reálného světa jako diskrétní prvky. Prvky jsou reprezentovány body, liniemi a polygony.
Vektorový model	Vektorový model je nazývaný podle způsobu vyjádření jeho jednotlivých částí – úseků křivek s jasně definovanou velikostí a směrem – vektorem. Model používá k vyjádření (kódování) geometrických dat tři základní grafické elementy
Vrstva	Sestava geografických dat. Jako vrstva mohou být zobrazeny vektorové prvky, rastrové mapy anebo také trojrozměrné scény/modely.
Výběr archiválií	Posouzení hodnoty dokumentů a rozhodnutí o jejich vybrání za archiválie a zařazení do evidence archiválií [§ 2 písm. g) Zákona].
Webová služba	Webová služba je nástroj umožňující komunikaci a výměnu

	informací prostřednictvím sítě Internet.
WMS (Web Map Service)	Webová mapová služba; jedná se o standard vyvinutý a dále rozšiřovaný OGC, umožňuje sdílení geografické informace ve formě rastrových map v prostředí Internetu.
Znázornění	Uživatelsky srozumitelná interpretace dokumentu nebo datasetu v digitální podobě. Znázornění má zpravidla podobu zobrazení na obrazovce, tiskového výstupu, zvukové nebo multimediální prezentace.
Zničení	Zničením se rozumí proces likvidace entit, který znemožňuje jejich rekonstrukci a identifikaci jejich obsahu.
Ztvárnění	Ztvárněním se rozumí výsledek konverze nebo převedení dokumentu, kterým je vyjádřena transformace dokumentu nebo komponenty při použití jednoho nebo více formátů odlišných od původních formátů. Ztvárnění se zpravidla vytvářejí pro uchování dokumentů v digitální podobě za účelem minimalizace rizika ztráty přístupu k jejich obsahu v čase. Například dokumenty vyhotovené v proprietárním datovém formátu musí být uloženy jako ztvárnění ve výstupním datovém formátu stanoveném prováděcím právním předpisem upravujícím podrobnosti výkonu spisové služby (například PDF/A). Výsledkem konverze dokumentu je ztvárnění některých nebo všech jeho komponent.

Příloha č. 2 – Přehled právních předpisů k prostorovým datům

V České republice dosud chybí národní právní předpis, který by uceleně a systematicky zakotvil národní infrastrukturu pro prostorové informace. Vzhledem k mimořádnému významu prostorových informací navrhla Evropská komise zřídit Infrastrukturu pro prostorové informace v Evropském společenství, která má napomáhat tvorbě politik a činností s přímým nebo nepřímým vlivem na životní prostředí.

Obecná pravidla byla stanovena Směrnicí Evropského parlamentu a Rady 2007/2/ES, o zřízení Infrastruktury pro prostorové informace v Evropském společenství (INSPIRE). Směrnice INSPIRE byla transponována v ČR zákonem č. 380/2009 Sb., kterým se mění zák. č. 123/1998 Sb., o právu na informace o životním prostředí, ve znění pozdějších předpisů, a zák. č. 200/1994 Sb., o zeměměřictví. Ministerstvo životního prostředí touto novelou zřizuje Národní geoportál INSPIRE, který široké veřejnosti zpřístupňuje prostorová data týkající se alespoň jednoho z témat přílohy směrnice. Služby na geoportálu umožňují uživateli vyhledávat, prohlížet a stahovat data. V průběhu implementace principů směrnice INSPIRE do národního prostředí se potvrdilo, že je účelné rozšířit principy INSPIRE nad rámec 34 témat prostorových dat směrnice INSPIRE řešených a na jejich základě pak národní infrastrukturu pro prostorové informace (NIPI) koordinovaně rozvíjet.

Tvorba, poskytování, správa, sdílení i využívání prostorových informací probíhá v současné době roztržštěně a bez dostatečné koordinace z centrální úrovně – chybí gestor. V ČR není k dispozici ucelený přehled existujících prostorových informací ani služeb nad nimi poskytovaných. V praxi dochází k opakovanému pořizování obdobných či dokonce stejných prostorových informací (duplicita). Na druhé straně některé významné prostorové informace chybí (např. informace o bytech). Neexistuje jednotná koordinace řízení kvality dat a služeb nad prostorovými daty. Většina právních předpisů neuvádí výslovnou povinnost ke zpracování, sdílení či poskytování prostorových dat, ale pouze zpracování informací. Je třeba respektovat závěry dokumentů vzešlých z dosavadních prací na Strategii rozvoje infrastruktury pro prostorové informace v ČR do roku 2020 (GeoInfoStrategie – usnesení vlády ČR č. 815 ze dne 8. října 2014) a rady expertů z oblasti geodzie a geoinformatiky, jakož i principy eGovernmentu a Smart Administration. Standardizaci a posun v této oblasti lze předpokládat od připravované Geoinfostrategie 2020+.

V oblasti prostorových dat musí být nejprve zřejmé politické zadání, které jasně označí subjekty veřejné správy, které budou odpovědné za koordinaci rozvoje, i míru zapojení soukromého sektoru. Technické řešení se v současnosti jeví jako poměrně jasné a v rámci digitalizace veřejné správy i proveditelné. Povinnost data sdílet a poskytovat není legislativně jednotně definována, respektive jednotlivé právní předpisy neuvádějí výslovně pojem sdílení nebo poskytování, i když obsah jejich ustanovení lze pod pojmy sdílení či poskytování do značné míry podřadit. Tato míra se u jednotlivých právních předpisů liší a je třeba nadefinovat, co se rozumí sdílením a co poskytováním.

Sdílení je zpravidla zajištění přístupu k datům a informacím pro více uživatelů nezávisle na sobě, při sdílení souborů např. modelem peer–to–peer, (Pozn.: rovný s rovným nebo klient–klient, jednotliví klienti/uživatelé spolu komunikují přímo; opakem je klient–server, kde klienti

komunikují vždy s centrálním serverem), uživatelé mají možnost složky stahovat i nahrávat, bez nutnosti požádat o data jejich pořizovatele či správce.

Poskytovat znamená zejména dávat, skýtat – čili jednosměrná záležitost, kdy držitel dat předá jejich soubor dalšímu subjektu, ať už na žádost nebo ze zákona.

Je třeba konstatovat, že řada právních předpisů je těžko zařaditelných – např. RÚIAN (registr územní identifikace, adres a nemovitostí), zřízený zákonem č. 111/2009 Sb., o základních registrech. Jedná se o typický registr, který je na pomezí sdílení a poskytování dat. Převažuje názor, že v tomto případě dochází ke sdílení prostorových dat. V případě povinnosti poskytovat prostorová data veřejnosti se může jednat o právo na informace.

Právní předpisy z oblasti prostorových dat lze rozdělit na: klíčové, podpůrné a méně významné.

Klíčové právní předpisy – buď obsahují právní úpravu týkající se prostorových dat ve velké míře, obsahují duplicity či jiné zásadní problémy, které je třeba odstranit, anebo naopak skýtají velký potenciál k dalšímu rozvoji NIPI, byť se (zatím) prostorovými daty nezabývají. To je např. zákon č. 2/1969 Sb., o zřízení ministerstev a jiných ústředních orgánů státní správy (kompetenční zákon), jehož změna je nutná pro určení gestora za oblast prostorových dat a informací.

Podpůrné právní předpisy – obsahují alespoň nějakou úpravu týkající se prostorových dat či informací, resp. agendy s prostorovým potenciálem.

Méně významné právní předpisy se prostorových dat a informací nedotýkají vůbec nebo jen velmi okrajově.

Většina právních předpisů neuvádí výslovnou povinnost ke zpracování, sdílení či poskytování prostorových dat, nýbrž pouze zpracovávání informací. V současné době jsou respektovány závěry dokumentů vzešlých z dosavadních prací na Strategii rozvoje infrastruktury pro prostorové informace v ČR do roku 2020 (GeoInfoStrategie), rady expertů z oblasti geodézie a geoinformatiky, vč. principů eGovernmentu a Smart Administration.

[Analýza klíčových právních předpisů](#)

Zákon č. 2/1969 Sb., o zřízení ministerstev a jiných ústředních orgánů státní správy (tzv. kompetenční zákon) právní předpis, rámcově stanovuje působnost jednotlivých ministerstev, obsahuje zásady, které se vztahují na všechny ústřední správní úřady. Neupravuje způsob tvorby prostorových dat a informací, správu prostorových dat, podmínky poskytování či sdílení.

Jde o jeden z klíčových právních předpisů ČR, kde by bylo vhodné jasně vymezit odpovědnost některého z ministerstev za oblast národní infrastruktury pro prostorové informace s ohledem na potřebu alokovat pro tuto problematiku potřebné finanční prostředky ze státního rozpočtu.

Zákon č. 200/1994 Sb., o zeměměřictví vymezuje jednotlivé zeměměřické činnosti, které jsou buď samy o sobě pořizováním či zpracováváním prostorových dat, nebo s prostorovými daty souvisí. Zákon upravuje práva a povinnosti při výkonu zeměměřických činností, konkrétně v §

6 ukládá povinnost ke sdílení a poskytování prostorových dat, dále upravuje ověřování výsledků zeměměřických činností a geodetické referenční systémy a státní mapová díla.

Vyhláška č. 31/1995 Sb., kterou se provádí zákon č. 200/1994 Sb., o zeměměřictví. Jde spíše o technickou normu stanovující podmínky správy a poskytování prostorových dat ve veřejném zájmu orgány státní správy na úseku zeměměřictví. Z hlediska úvah o budoucí podobě národní infrastruktury pro prostorová data je možné využít tohoto prováděcího předpisu a doplnit další podrobnosti ohledně formy poskytování dat, rozšiřování dat apod.

Zákon č. 139/2002 Sb., o pozemkových úpravách a pozemkových úřadech, gestorem je ministerstvo zemědělství. Je třeba se zaměřit na lepší využití prostorových dat pořízených v rámci pozemkových úprav a dále nevyužívaných a nesdílených.

Zákon č. 219/2000 Sb., o majetku ČR a jejím vystupování v právních vztazích. Jako nutné se jeví přijetí nařízení vlády specifikující prostorová data a informace a nakládání s nimi dle § 14a tohoto zákona a v rámci NIPÍ počítat se začleněním tohoto registru, který je naplněn a jehož data by mohla sloužit jako podklad pro další činnost veřejné správy. Ke zvážení je, zda má být registr ze zákona veřejný či neveřejný.

Zákon č. 256/2013 Sb., o katastru nemovitostí (katastrální zákon). Zásadní právní předpis v oblasti evidence nemovitostí s výraznými soukromoprávními dopady a definicemi pojmů pozemek, parcela, budova apod.

- Vyhláška č. 357/2013 Sb., o katastru nemovitostí (katastrální vyhláška).
- Vyhláška č. 358/2013 Sb., o poskytování údajů z katastru nemovitostí.

Zákon č. 183/2006 Sb., o územním plánování a stavebním řádu (stavební zákon), má vazbu na řadu obecně závazných právních předpisů. Stavební zákon je klíčovým právním předpisem pro oblast prostorových dat, avšak jeho právní úprava z hlediska sdílení či poskytování dat a informací je zcela nedostatečná. S ohledem na digitalizaci veřejné správy a princip Smart Administration, stejně jako na princip obíhání dat, nikoli občanů, je třeba přijmout odpovídající změny, které umožní maximální využití prostorových dat pořizovaných v jednotlivých procesech dle stavebního zákona. U geometrických plánů, které se v některých případech řízení předkládají, existuje jejich další využití, kdy na jejich podkladě dochází ke změně v katastrální mapě. Dokumentace skutečného provedení stavby naproti tomu žádné další využití v současné době nenachází a prostorová data, která byla pro tyto účely pořízena, tak zůstávají ležet ladem.

Zákon č. 123/1998 Sb., o právu na informace o životním prostředí. Transponuje příslušné předpisy EU a upravuje zabezpečení práva na přístup k informacím o životním prostředí. Ve vztahu k prostorovým datům stanoví pravidla pro zřízení infrastruktury pro prostorová data pro účely politik životního prostředí a politik nebo činností, které mohou mít vliv na životní prostředí a zpřístupňování prostorových dat prostřednictvím síťových služeb na Národním geoportálu

INSPIRE.

Zákon velice široce stanoví povinnost ke zpřístupňování dat, která pořizují povinné subjekty a která mají vztah k životnímu prostředí. Zavádí geoportál pod správou MŽP. Výsledkem je

velice obsáhlý fond dat, která jsou ovšem značně nesourodá a uživatelsky poměrně nepřívětivá – v tomto objemu a rozmanitosti dat není zřejmě možné zpracovat je v jednotné struktuře, přesnosti, podrobnosti, což ani není cílem tohoto zákona. Pro účely zpřístupnění informací o životním prostředí a již pořízených prostorových dat je zřejmě právní úprava dostačující. Chybí požadavky na přesnost – v tomto objemu a při rozlišnosti povahy dat, která jsou zpřístupňována, ani není možné požadovanou přesnost specifikovat – pro jednotlivé povinné subjekty platí, že zpracovávají data tak, aby to primárně vyhovovalo jejich potřebám, nikoli tak, aby byla dále využitelná a zpracovatelná dalšími subjekty.

Analýza podpůrných právních předpisů

Zákon č. 17/1992 Sb., o životním prostředí. Zákon neupravuje způsob tvorby prostorových dat a informací, ani správu prostorových dat, podmínky poskytování, sdílení či definici prostorového objektu. Tento zákon představuje lex generalis k řadě speciálních právních předpisů s ohledem na prostorová data a informace. V § 18 je obecná povinnost sledovat působení vlastní činnosti na životní prostředí a poskytovat informace. Vazby na související právní předpisy:

- Zákon č. 123/1998 Sb., o právu na informace o životním prostředí
- Zákon č. 185/2001 Sb., o odpadech
- Zákon č. 254/2001 Sb., o vodách
- Zákon č. 274/2001 Sb., o vodovodech a kanalizacích pro veřejnou potřebu
- Zákon č. 201/2012 Sb., o ochraně ovzduší
- Zákon č. 25/2008 Sb., o integrovaném registru znečišťování životního prostředí a integrovaném systému plnění ohlašovacích povinností v oblasti životního prostředí

Zákon č. 89/2012 Sb., občanský zákoník, obsahuje základní ustanovení/definice, např. nemovitě věci, ohraničení prostorových objektů, jednotka, společné části, je však třeba sjednotit řadu definic. Neupravuje způsob tvorby prostorových dat a informací, správu, podmínky poskytování, sdílení apod.

Zákon č. 114/1992 Sb., o ochraně přírody a krajiny. Vazby na související právní předpisy:

- Zákon č. 256/2013 Sb., o katastru nemovitostí
- Vyhláška č. 357/2013 Sb., o katastru nemovitostí (katastrální vyhláška)
- Vyhláška č. 395/1992 Sb., kterou se provádějí některá ustanovení zákona č. 114/1992 Sb., o ochraně přírody a krajiny.

V rámci dalších prací na GeoInfoStrategii bude vhodné oblasti prostorových dat a informací pro účely ochrany životního prostředí věnovat zvláštní pozornost, neboť jde o velice rozsáhlý soubor souvisejících prvků, které jsou již v současné době předmětem zpracování v různých resortních informačních systémech a mohly by společně tvořit jeden z pilířů rozvoje architektury NIPI.

Zákon č. 128/2000 Sb., o obcích (obecní zřízení). Vazby na související předpisy:

- Zákon č. 358/2013 Sb., o poskytování údajů z katastru nemovitostí
- Zákon č. 111/2009 Sb., o základních registrech
- Zákon č. 151/1997 Sb., o oceňování majetku
- Zákon č. 183/2006 Sb., o územním plánování a stavebním řádu (stavební zákon)

Z hlediska prostorových dat a informací se u výše uvedených právních předpisů jedná o možnost, **nikoliv povinnost** zpracovávat, poskytovat a sdílet data. Zpracování a předávání prostorových informací, právo na bezplatné využívání údajů katastru nemovitostí ... Spolupráce s orgány státní správy dle výše uvedených právních předpisů by mohla být vedena nad prostorovými daty, která by byla předmětem systému NIPI.

Zákon č. 100/2001 Sb., o posuzování vlivů na životní prostředí. Vazby na související předpisy:

- Zákon č. 123/1998 Sb., o právu na informace o životním prostředí
- Zákon č. 183/2006 Sb., o územním plánování a stavebním řádu (stavební zákon)
- Zákon č. 20/1966 Sb., o péči o zdraví lidu
- Zákon č. 289/1995 Sb., o lesích
- Zákon č. 258/2000 Sb., o ochraně veřejného zdraví
- Zákon č. 201/2012 Sb., o ochraně ovzduší
- Zákon č. 129/2000 Sb., o krajích

Zákon č. 131/2000 Sb., o hlavním městě Praze – z hlediska prostorových dat a informací možnost, **nikoliv povinnost** zpracovávat, poskytovat a sdílet data.

Zákon č. 44/1988 Sb., o ochraně a využití nerostného bohatství (horní zákon). Zákon stanoví povinnost k vedení řady evidencí, které jsou převážně vedeny formou rejstříku bez návaznosti na prostorová data.

Zákon č. 121/2000 Sb., o právu autorském (autorský zákon).

Zákon č. 111/2009 Sb., o základních registrech. Zákon stanoví obecně legislativní podmínky fungování základních registrů.

Zákon č. 106/1999 Sb., o svobodném přístupu k informacím. Upravuje pravidla pro poskytování informací povinnými subjekty.

Zákon č. 240/2000 Sb., o krizovém řízení (krizový zákon), který stanoví podmínky k zajištění připravenosti na krizové stavy, vč. zabezpečení závaznými geografickými podklady. S ohledem na specifičnost oblasti krizového plánování a řízení je otázka, do jaké míry je žádoucí/bezpečné zapojení do budoucí NIPI.

Legislativní analýze slouží výhradně pro potřeby této metodiky, není a ani nemůže být zcela vyčerpávající. Jejím cílem je shrnout současnou legislativu v oblasti prostorových dat, jejich

sdílení, přenosu atp., aby tyto legislativně ukotvené možnosti mohly posloužit při přípravě této metodiky a zároveň v ní být rámcově zohledněny.

Příloha č. 3 – Dublin Core

Takzvaný Dublin Core je metadatové schéma skládající se z 13 základních elementů, využívané v různých oborech. Je společným jádrem různých metadatových schémat. Schéma bylo definováno s cílem popsat uložené informace, a tak usnadnit hledání v databázovém prostředí. Univerzálnost je v jistém smyslu silnou a zároveň slabou stránkou. Na jednu stranu propojuje popisné informace z různých vědních oborů, ale zároveň je kompromisem, který neumožňuje naplno využívat jednotlivé metadatové slovníky, které vědní obory využívají k popisování svých dat a datasetů.

Č.	Název	Popis
1	<Title>	Název
2	<Subject>	Téma
3	<Name>	Jméno/ název tvůrce/vlastníka dat
4	<Description>	Popis obsahu datasetu
5	<Date>	Datum vytvoření datasetu
6	<ResourceType>	Typ popisovaného objektu
7	<Format>	Formát, ve kterém je objekt uložen
8	<ResourceIdentifier>	Identifikátor
9	<Source>	Původ, informace o mateřském objektu
10	<Language>	Jazyk
11	<Relation>	Identifikátor nebo vazba na příbuzný dataset
12	<Coverage>	Prostorové/časové vymezení obsahu datasetu
13	<RightsManagement>	Práva a omezení spojená s využíváním datasetu

Do souboru METS.xml se ze souboru dc.xml budou vytěžovat tyto elementy

	Elementy Dublin Core	Elementy METS.xml v GeoSIP
1	<Title>	<sip:nazev>
4	<Description>	<sip:komentar>
5	<Date>	<sip:datumUzavreni>
6	<ResourceType>	<sip:druh>
8	<ResourceIdentifier>	<sip:identifikator>
12	<Coverage>	<sip:odRoku> <sip:doRoku>
13	<RightsManagement>	<sip:typOmezeni>

Do souboru pruvodka.xml se ze souboru dc.xml budou vytěžovat tyto elementy:

	Elementy Dublin Core	Elementy pruvodka.xml v GeoSIP
2	<Subject>	<index>
9	<Source>	<index>
11	<Relation>	<origination>

Příloha č. 4 – Soubory METS.xml v EARK_Geospatial

METS.xml uvnitř adresáře /komponenty

Soubor je sestaven plně podle EARK specifikace GEOSIP a CSIP EARK⁵⁸). V EARK_Geospatial musí vždy existovat alespoň jedna reprezentace dat, a proto musí balíček obsahovat minimálně dva soubory METS.xml. Jednak METS.xml pro celý balíček, kde je uvedeno, že balíček obsahuje prostorová data. Dále METS.xml pro každou reprezentaci, který popisuje hlavní datové typy v souboru reprezentace (podrobně viz specifikace GEOSIP a CSIP EARK).

METS.xml pro celý GEOSIP balíček

- › prvek mets/@TYPE. **musí** obsahovat text „Geospatiala Data“
- › hodnota prvku mets/@csip:CONTENTINFORMATIONTYPE **musí** být „citsgeospatial_v3_0“ a **nesmí** být „OTHERCONTENTINFORMATIONTYPE“
- › v prvku mets/@PROFILE **musí** být hodnota [https://citsgeospatial.dilcis.eu/profile/E-ARK-GEOSPATIAL- ROOT.xml](https://citsgeospatial.dilcis.eu/profile/E-ARK-GEOSPATIAL-ROOT.xml)
- › v mets/fileSec/fileGrp[@USE='Representations'] /@csip:CONTENTINFORMATIONTYPE musí být hodnota „citsgeospatial_v3_0“
- › pro každý fileGrp/@csip:CONTENTINFORMATIONTYPE s hodnotou "citsgeospatial_v3_0" musí být odpovídající @div- reprezentace ve StructMap

METS.xml pro balíček reprezentace

- › prvek mets/@TYPE. **musí** obsahovat text „Geospatiala Data“
- › hodnota prvku mets/@csip:CONTENTINFORMATIONTYPE **musí** být „citsgeospatial_v3_0“ a **nesmí** být „OTHERCONTENTINFORMATIONTYPE“
- › v prvku @PROFILE **musí** být hodnota "[https://citsgeospatial.dilcis.eu/profile/E-ARK-GEOSPATIAL- REPRESENTATION.xml](https://citsgeospatial.dilcis.eu/profile/E-ARK-GEOSPATIAL-REPRESENTATION.xml)"

⁵⁸ <https://dilcis.eu/specifications/common-specification>

Příloha č. 5 – mets.xml v GeoSIP

Popis vnitřních elementů datasetu⁵⁹

<sis:dataset ID="dataset_1">

Pole <sis:dataset > je rodičovská entita metadatového prostoru, která obsahuje povinné podřízené entity. Zároveň obsahuje informaci identifikátoru v rámci řízení.

<sis:identifikator>CZNDA100000010MSK573_1</sis:identifikator>

Pole <sis:identifikator> je prostor, kam je na portálu uložen unikátní identifikátor datového balíčku, příp. obsahuje unikátní identifikátor, který do pole vloží původce.

<sis:nazev>Natura 2000 celá ČR</sis:nazev>

Pole <sis:nazev> je textový prostor pro název datového setu, který bude sloužit pro pořadací software a hledání na Národním archivním portálu.

<sis:komentar>Dataset obsahuje informace všech chráněných území typu Natura 2000 v podobě vektorového zakreslení</sis:komentar>

Pole <sis:komentar> je prostor pro základní popis datasetu.

<sis:datumUzavreni>2021-03-25</sis:datumUzavreni>

Pole <sis:datumUzavreni> je prostor pro informaci o datu vytvoření datového balíčku SIP.

<sis:poradoveCislo>1</sis:poradoveCislo>

Pole <sis:poradoveCislo> je prostor pro numerickou informaci o pořadovém čísle balíčku v případě, že původce data rozdělil do více balíčků (např. dle jednotlivých krajů).

⁵⁹ Aktuální popis a schéma pro vytváření SIPů z volných souborů vč. GeoSIP je dostupný na adrese <http://www.nacr.cz/sipy>

<sip:casovyRozsah>

<sip:odRoku>2000</sip:odRoku>

<sip:doRoku>2005</sip:doRoku>

</sip:casovyRozsah>

<sip:omezeniPristupu>

<sip:typOmezeni>neurčeno</sip:typOmezeni>

</sip:omezeniPristupu>

<sip:cisloJednaci>AOPK21SGJS152SSF</sip:cisloJednaci>

<sip:spisovyZnak/>

<sip:celkovyPocetSipXml>1</sip:celkovyPocetSipXml>

<sip:evidencyJednotka>

<sip:druh>dataset</sip:druh>

<sip:pocet>1</sip:pocet>

</sip:evidencyJednotka>

<sip:poznamka/>

</sip:dataset>

Pole <sip:casovyRozsah> je prostor pro numerickou informaci o časovém období, kterým jsou vymezeny informace v datasetu.

Pole <sip:typOmezeni> je vyhrazeno pro metadatovou informaci o tom, zda je k datasetu omezen přístup.

Pole <sip:cisloJednaci> je prostor pro metadatovou informaci o čísle jednacím přejímky.

Pole <sip:celkovyPocetSipXML> obsahuje informaci o počtu balíčků SIP v rámci jednoho výběru archiválií.

Pole <sip:druh> je vyhrazeno pro metadatovou informaci o druhu evidenční jednotky. Zde se použije vždy „dataset“.

Pole <sip:pocet> je prostor pro metadatovou informaci o počtu balíčků SIP v jednom datasetu.

Ukázka souboru mets.xml Pro datový balíček GeoSIP

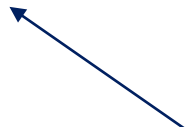
```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:sip="http://www.nacr.cz/sipy"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" ID="CZNDA100000010ESK1669" LABEL="Datový balíček Národního archivu pro
výběr archiválií z datových souborů mimo ERMS – Submission Information Package (SIP)"
xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd http://www.nacr.cz/sipy
http://digi.nacr.cz/sipy/msk01/mimoskart_v05.xsd info:lc/xmlns/premis-v2 http://www.loc.gov/standards/premis/v2/premis-v2-
2.xsd">
  <mets:metsHdr CREATEDATE="2021-09-06T14:19:35.000+02:00" LASTMODDATE="2021-09-06T14:32:59+02:00">
    <mets:agent ROLE="CREATOR" TYPE="ORGANIZATION">
      <mets:name>Národní archiv</mets:name>
      <mets:note>100000010</mets:note>
    </mets:agent>
    <mets:agent ROLE="CREATOR" TYPE="INDIVIDUAL">
      <mets:name>Josef Kanárek</mets:name>
      <mets:note></mets:note>
    </mets:agent>
  </mets:metsHdr>
  <mets:dmdSec ID="msk1">
    <mets:mdWrap MDTYPE="OTHER">
      <mets:xmlData>
        <sip:rizeni ID="rizeni_1">
          <sip:identifikator>CZNDA100000010ESK1669</sip:identifikator>
          <sip:nazev>Ukázka výběru GeoSIP</sip:nazev>
          <sip:komentar></sip:komentar>
          <sip:datumUzavreni>2021-09-06</sip:datumUzavreni>
          <sip:cisloJednaci>JBa 0609/2021-1432</sip:cisloJednaci>
          <sip:celkovyPocetSipXml>2</sip:celkovyPocetSipXml>
          <sip:puvodce>
            <sip:nazev>Bernas Jiří</sip:nazev>
            <sip:interpi>n000531213</sip:interpi>
          </sip:puvodce>
          <sip:archiv>
            <sip:nazev>Národní archiv</sip:nazev>
            <sip:cArchivu>100000010</sip:cArchivu>
          </sip:archiv>
          <sip:poznamka></sip:poznamka>
        </sip:rizeni>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmdSec>
</mets:mets>
```

```

<sis:dataset ID="dataset_1">
  <sis:identifikator zdroj="MSK">MSK_NA_20210906_123008</sis:identifikator>
  <sis:nazev>GeoSIP obsahující zjednodušený formát geodat</sis:nazev>
  <sis:komentar></sis:komentar>
  <sis:datumUzavreni>2021-09-06</sis:datumUzavreni>
  <sis:poradoveCislo>2</sis:poradoveCislo>
  <sis:casovyRozsah>
    <sis:odRoku>2004</sis:odRoku>
    <sis:doRoku>2021</sis:doRoku>
  </sis:casovyRozsah>
  <sis:omezeniPristupu>
    <sis:typOmezeni>neurčeno</sis:typOmezeni>
  </sis:omezeniPristupu>
  <sis:cisloJednaci>Spis02/2021</sis:cisloJednaci>
  <sis:celkovyPocetSipXml>0</sis:celkovyPocetSipXml>
  <sis:evdencniJednotka>
    <sis:druh>dts</sis:druh>
    <sis:pocet>1</sis:pocet>
  </sis:evdencniJednotka>
  <sis:poznamka></sis:poznamka>
</sis:dataset>
<sis:adresar ID="adr1">
  <sis:puvodniNazevSouboruAdresare>Zaměření areálových sítí -
Chodovec</sis:puvodniNazevSouboruAdresare>
  <sis:puvodniCesta></sis:puvodniCesta>
  <sis:novaCesta>komponenty/Arealove_site</sis:novaCesta>
  <sis:nazevDokumentu>Zamereni_arealovych_siti.dwg</sis:nazevDokumentu>
  <sis:obsahDokumentu>Zaměření areálových sítí archivního areálu v Praze na Chodovci.</sis:obsahDokumentu>
  <sis:casovyRozsah>
    <sis:odRoku>2004</sis:odRoku>
    <sis:doRoku>2021</sis:doRoku>
  </sis:casovyRozsah>
  <sis:evdencniJednotka>
    <sis:druh>dts</sis:druh>
    <sis:pocet>1</sis:pocet>
  </sis:evdencniJednotka>
</sis:adresar>
</mets:xmlData>

```

```
</mets:mdWrap>
</mets:dmdSec>
<mets:structMap TYPE="logická" LABEL="Logická struktura v rámci řízení">
  <mets:div DMDID="řízení_1" TYPE="řízení">
    <mets:div DMDID="dataset_1" TYPE="dataset">
      <mets:div DMDID="adr1" TYPE="adresář" LABEL="GeoSIP">
        </mets:div>
      </mets:div>
    </mets:div>
  </mets:div>
</mets:structMap>
</mets:mets>
```



Atribut **LABEL=""** slouží k označení adresáře jako balíčku SIP s prostorovými daty a jsou pro něj povoleny pouze hodnoty: "GeoSIP".

Příloha č. 6 – Návrh převodu elementů prostorových metadat do archivního prostředí

Název elementu **Vysvětlení obsahu elementu** **Základní pravidla pro zpracování EAD průvodka NArP archiválií, metodický pokyn č. 1/2021 - zpracování map v archivech**

Název	Název	Formální název jednotky popisu (ZP 4.2.4)	unittitle
Abstrakt	Popis obsahu prostorových dat	Obsah, regist (ZP 4.2.3)	abstract
Typ	Informace, zda se jedná o dataset, aplikaci, službu aj.	vždy bude dataset	physdescstructured (vždy dts)
Odkaz	URL odkaz na webové stránky tvůrce	Edice a literatura (ZP 5.2.11)	origination
Jedinečný identifikátor	Informace o identifikátoru přiděleným tvůrcem	Jiné označení (ZP 4.2.2)	localtype
Vázaný zdroj	Vazba na zdrojová data	Odkazy na příbuzné dokumenty, archiválie a pomůcky pro vyhledávání (ZP 4.5.2)	
Jazyk	Informace o jazyku datasetu	Jazyk, písmo (ZP 5.2.10)	
Tematická kategorie	Informace o základních tematických kategoriích datasetu	Druh mapy	index
Klíčové slovo	Upřesňující informace o tematických kategoriích dle standardizovaných metadat	Archivní autoritní záznamy jako přístupové body (ZP 6.3)	index
Geografické ohraničení	Informace o prostorovém ohraničení datasetu v podobě souřadnic	Souřadnice (ZP 5.2.6)	
Referenční datum	Informace o datu vzniku datasetu	Datace vzniku jednotky popisu (ZP 4.2.5)	unitdatestructured /daterange
Časový rozsah	Informace o časovém rozsahu dat v datasetu	Jiné datace jednotky popisu než datace vzniku jednotky popisu (ZP 5.2.1)	
Původ	Informace o důvodu a účelu vzniku datasetu	Dějiny jednotky popisu (ZP 4.3.2)	didnote
Prostorové rozlišení	Měřítko, ve kterém jsou prostorová data zpracována	Měřítko (ZP 5.2.5)	index

Specifikace	Informace o kompatibilitě s uznávanými standardy pro prostorová data	Pravidla zpracování jednotky popisu (ZP 4.7.2)	
Podmínky pro přístup a užití	Informace o autorských právech	Podmínky přístupu, práva k jednotce popisu a její reprodukci (ZP 4.4.1)	
Omezení veřejného přístupu	Informace o omezení přístupu k datům	Možnost zveřejnění reprodukce jednotky popisu (ZP 4.4.3)	accessrestrict
Zodpovědná organizace	Kontaktní informace na tvůrce datasetu	Veřejná poznámka (ZP 4.6.2)	origination
Typ prostorové prezentace	Informace o druhu dat vektor/rastr	Technika, adjustace, nosič a látka záznamu (ZP 5.2.8)	index
Souřadnicový referenční systém (název)	Název souřadnicového systému použitého při tvorbě datasetu	Rozměry, hmotnost, velikost, množství (ZP 5.2.4)	index
Údržba	Informace jak je dataset aktualizován	-	
Nadřízený	Související osoby/archiválie	Archivní autoritní záznamy jako přístupové body (ZP 6.3)	index
Podřízený	Související osoby/archiválie	Archivní autoritní záznamy jako přístupové body (ZP 6.3)	index
Formát	Informace o datovém formátu, ve kterém jsou prostorová data dostupná	Technika, adjustace, nosič a látka záznamu (ZP 5.2.8)	
Identifikátor metadatového záznamu	Identifikátor metadatové informace na geoportal.gov.cz	Jiné označení (ZP 4.2.2)	
Metadata kontakt	Kontakt na tvůrce datasetu	Služební poznámka (ZP 4.6.1)	origination
Norma metadat	Informace o standardu metadat použitého k popisu datasetu	Pravidla zpracování jednotky popisu (ZP 4.7.2)	
Datum metadat	Datum aktualizace metadat na geoportal.gov.cz	-	
Jazyk metadat	Informace o jazyku, ve kterém jsou zpracována metadata datasetu	Jazyk, písmo (ZP 5.2.10)	

Příloha č. 7 – Ukázka archseznam.xml

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<Seznam xmlns="http://www.mvcr.cz/nsesss/v3"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" ID="ESK1644"
  identifikatorArchivu=" sestaveno nástrojem ArchiGIS "
  identifikatorPuvodce="ČÚZAK"
  identifikatorSkartacnihoRizeni="NA-2021/7229"
  datumVytvoreni="2021-08-12T13:25:59+02:00"
  xsi:schemaLocation="http://www.mvcr.cz/nsesss/v3
http://www.mvcr.cz/nsesss/v3/nsesss-DA.xsd">
  <EntitaSeznamu>
    <Identifikator zdroj="MSK">CZND10000010ESK1669_1382492</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="individual">MSK_NA_20210906_123008</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
</Seznam>
```

Vysvětlivky

identifikatorArchivu: defaultní hodnota="sestaveno nástrojem ArchiGIS" nebo „neuvedeno“

identifikatorPuvodce: identifikátor původce poskytnutý archivem

identifikatorSkartacnihoRizeni: číslo jednací výběru archiválií původce, příp. jiné evidenční číslo řízení

datumVytvoreni: datum sestavení seznamu v příslušném formátu

xsi:schemaLocation: přípustná hodnota je "http://www.mvcr.cz/nsesss/v3
http://www.mvcr.cz/nsesss/v3/nsesss-DA.xsd">

<EntitaSeznamu> obsahuje pouze identifikátory celých GeoSIP

<Identifikator zdroj="VUMOP-123">5cb74630-5840-4214-825a-d103611fda54</Identifikator>:

zdroj – zkratka zdrojového systému pro tvorbu seznamu;

identifikátor: identifikátor uvedený v balíčku GeoSIP

<Operace>: pro možnost předání GeoSIP digitálnímu archivu musí být hodnota „vybrat za archiválii“

Příloha č. 8 – Popis testovacích datasetů

Pro účely přípravy metodiky byly využity datové sety, vycházejí z reálné podoby prostorových dat.

1. GIS Data eKatalogu BPEJ⁶⁰

	Název atributu	Popis
1	<Title>	Bonitované půdně ekologické jednotky
2	<Subject>	Zemědělství
3	<Name>	Výzkumný ústav meliorací a ochrany půdy, v.v.i.
4	<Description>	Mapa bonitovaných půdně ekologických jednotek, mapa velkého měřítká
5	<Date>	03. 04. 2017
6	<ResourceType>	Dataset
7	<Format>	GML
8	<ResourceIdentifier>	VUMOP:CZ-VUMOP-BPEJ
9	<Source>	Státní pozemkový úřad
10	<Language>	Český jazyk
11	<Relation>	SPU # CZ-01312774-BPEJ_aplikace
12	<Coverage>	03. 04. 2017
13	<RightsManagement>	žádné podmínky neplatí, žádné omezení veřejného přístupu
	Počet objektů celkem	16
	Datový objem celkem	2,1 GB
	Typ dat	Vektor
	Počet vrstev	1

S daty BPEJ byl dále proveden pilotní pokus migrace dat do archivního formátu SIARD⁶¹. Datový soubor byl sice vytvořen, ale nebyl validní vůči svému schématu resp. standardu. Standard specifikace pro archivní formát SIARD není kompatibilní v oblasti rutin s funkcemi, které umožňují přetěžování databáze, zatímco standard SIARD vyžaduje, aby každá rutina použitá v databázi byla unikátní.⁶² Do popisného souboru metadata.xml dále nedojde k popsání informací o extenzích, které původce využíval pro specifickou práci s prostorovými daty v originálním systému a ve výsledku není možná obnova dat do produkčního stavu v systému GIS.

⁶⁰ blíže popsáno v [kapitole č. 6](#)

⁶¹ Srv. Příloha č. 10, formát SIARD, str. 85.

⁶² Srv. Bod 5.15. specifikace [SIARD](#) str. 35.

2. Topografická databáze České republiky (Data200) - vrstva Vodstvo⁶³

	Název atributu	Popis
1	<Title>	Topografická databáze České republiky (Data200) - vrstva Vodstvo
2	<Subject>	Vodstvo
3	<Name>	Český úřad zeměměřický a katastrální
4	<Description>	Topografická databáze České republiky (Data200) - vrstva Vodstvo. Minimální polohová přesnost je 100 m a podrobnost odpovídá měřítku 1:200 000. Vrstva Vodstvo obsahuje vodní toky, vodní plochy, ostrovy, hráze, jezy, bažiny.
5	<Date>	01. 01. 2020
6	<ResourceType>	Dataset
7	<Format>	Shape File
8	<ResourceIdentifier>	https://cuzk.cz/CZ-00025712-CUZK_DATA200_VODSTVO
9	<Source>	https://geoportal.cuzk.cz
10	<Language>	Český jazyk
11	<Relation>	
12	<Coverage>	01. 01. 2020
13	<RightsManagement>	Licence Creative Commons CC BY 4.0
	Počet objektů celkem	86
	Datový objem celkem	15,6 MB
	Typ dat	Vektor
	Počet vrstev	10

⁶³ [https://geoportal.cuzk.cz/\(S\(t1u0bztrcdopsvdob5d2kctv\)\)/Default.aspx?mode=TextMeta&metadataID=CZ-CUZK-DATA200-VODSTVO-V&metadataXSL=Full&side=mapy_data200](https://geoportal.cuzk.cz/(S(t1u0bztrcdopsvdob5d2kctv))/Default.aspx?mode=TextMeta&metadataID=CZ-CUZK-DATA200-VODSTVO-V&metadataXSL=Full&side=mapy_data200)

3. Austrálie – územně správní hranice⁶⁴

	Název atributu	Popis
1	<Title>	ASGS (2016 Edition) - Boundaries
2	<Subject>	Územní správa
3	<Name>	ABS Geospatial Solutions
4	<Description>	Hranice územní správy
5	<Date>	29. 07. 2020
6	<ResourceType>	Dataset
7	<Format>	Geopackage
8	<ResourceIdentifier>	32adc1ef-5bac-4eaa-9521-a116792f32a1
9	<Source>	https://data.gov.au/data/dataset/32adc1ef-5bac-4eaa-9521-a116792f32a1
10	<Language>	Anglický jazyk
11	<Relation>	
12	<Coverage>	29. 07. 2020
13	<RightsManagement>	Creative Commons Attribution 2.5 Australia
	Počet objektů celkem	1
	Datový objem celkem	1,01 GB
	Typ dat	vektor
	Počet vrstev	8

Tento dataset byl dále využit i k pilotním testům formátových migrací pomocí nástroje QGIS 3.16 Hannover s těmito výsledky:

GML	Počet objektů: 16, celkový objem: 2,40 GB
SHP File	Počet objektů: 32, celkový objem: 1,12 GB
GeoJSON	Počet objektů: 8, celkový objem 2,45 GB
Interlis 1	Počet objektů: 8, celkový objem 1,85 GB (migrace nebyla úspěšná)

⁶⁴ <https://www.data.gov.au/data/dataset/asgs-2016-edition-boundaries/resource/48a33591-0b0a-4712-a60b-39ef4b564f2b>, zahraniční dataset byl zvolen z důvodu omezené dostupnosti formátu Geopackage v ČR

4. Charakteristiky toků a povodí ČR - Jizera DIBAVOD⁶⁵

	Název atributu	Popis
1	<Title>	Charakteristiky vodních toků a povodí ČR - Jizera
2	<Subject>	Vodstvo
3	<Name>	Výzkumný ústav vodohospodářský TGM
4	<Description>	V rámci tohoto projektu jsou pro povodí 100 nejdelších vodních toků ČR a povodí dalších 35 toků, pro které byly v roce 2006 vydány Atlasy záplavových území 1:10 000, publikovány technologické možnosti modelování vybraných prostorových charakteristik. Modelování využívá jako základní komponentu geometricko-topologický model říční sítě odvozený ze Základní báze geografických dat 1:10 000 (ZABAGED®), upravený v konvenci Digitální báze vodohospodářských dat 1:10 000 (DIBAVOD) a doplněný dalšími prvky vodopisu, kterými jsou vrstvy vodních nádrží a jezer a dále rozvodnice určující hydrologické členění. Vybraná výškopisná data (ZABAGED®), vrstevnice a prvky bodového pole obsahující hodnotu nadmořské výšky, jsou využity pro konstrukci rastrového modelu reliéfu. Popis použitých zdrojových dat a použitých metod najdete v textové části tohoto projektu.
5	<Date>	07. 01. 2021
6	<ResourceType>	Dataset
7	<Format>	GeoTIFF, PDF
8	<ResourceIdentifier>	
9	<Source>	https://www.dibavod.cz/24/charakteristiky-toku-a-povodi-cr.html
10	<Language>	Český jazyk
11	<Relation>	
12	<Coverage>	07. 01. 2021
13	<RightsManagement>	
	Počet objektů celkem	5
	Datový objem celkem	13,2 MB
	Typ dat	Raster
	Počet vrstev	4

⁶⁵ <https://www.dibavod.cz/24/charakteristiky-toku-a-povodi-cr.html>

Příloha č. 9 – Vzory licenčních dohod o převzetí/předání archiválií

Autorskoprávní otázky při předávání dat do archivu k trvalému uložení.

Data geografických informačních systémů lze podle okolností charakterizovat z hlediska autorskoprávního jako díla kartografická, ale také jako databáze.⁶⁶ Databází je po právní stránce „soubor děl, údajů nebo jiných nezávislých prvků, které jsou systematicky nebo metodicky uspořádány a které jsou jednotlivě přístupné elektronickými nebo jinými prostředky;“⁶⁷ tato definice plně vystihuje i podstatu vektorových dat.

V praxi lze narazit na nedůvěru původců a nejasnosti při předávání vybraných informací v digitální podobě do archivu. I když výběr probíhá na základě zákona č. 499/2004 Sb., jeví se vhodným ošetřit některé otázky týkající se zejména autorského práva a práv souvisejících samostatným dokumentem, který shrne práva a povinnosti původce i archivu – jde zejména o problematiku licence.⁶⁸ Proto je navržena vzorová dokumentace, která má zejména:

- › shrnout práva a povinnosti vyplývající ze zákona č. 499/2004 Sb. zejména pokud jde o nahlížení,
- › zaručit původci, že předaná data v digitální podobě může dále využívat (předává jejich repliku),
- › archivu udělit licenci, přičemž vedle zákonných možností lze stanovit i některé další podrobnosti,
- › neohrožovat případné komerční zájmy původce v případě dalšího využívání dat.

Vzorové dokumenty:⁶⁹

1. Zápis k úřednímu záznamu o předání archiválií: použije se v případě, že se jedná o vlastnictví České republiky, kdy původce i archiv jsou její organizační složky, oba jednájí jménem České republiky. Alternativně lze záznam použít v případě, že archiv sám je organizační složkou původce (např. archivy územních samospráv).
2. Dohoda o podmínkách převzetí archiválií: použije se ve všech ostatních případech, zejména pokud je původcem příspěvková organizace, veřejná výzkumná instituce, státní podnik, státní organizace, organizační složka veřejnoprávního původce nebo soukromoprávní původce (ve smyslu zákona č. 499/2004 Sb.).

⁶⁶ Srv. § 58 odst. 7 zákona č. 121/2000 Sb., autorský zákon.

⁶⁷ Směrnice Evropského parlamentu a Rady 96/9/ES ze dne 11. března 1996 o právní ochraně databází.

⁶⁸ Srv. LEŠKA, Rudolf: Archivy a autorské právo – symbióza nebo kolize? In: *Archivní časopis*. Praha: Ministerstvo vnitra, 2016, roč. 66, č. 2, s. 5–28. ISSN 0004-0398.

⁶⁹ Dokumenty byly připraveny ve spolupráci s advokátní kanceláří „Štáidl Leška advokáti“.

Dohoda o podmínkách převzetí archiválií

Archiv XY

IČO:
se sídlem
zastoupený
(dále jen „Archiv“)

a

Právnícká osoba YZ

IČO:
sídlo:
zastoupena:
(dále jen „původce“)

I. Preambule

1. Konstatuje se, že shora uvedené smluvní strany jsou subjekty s plnou právní způsobilostí k právním jednáním, oprávněnými jednat a podepisovat navenek.
2. Archiv je organizační složkou Působnost Archivu vyplývá zejména ze zákona č. 499/2004 Sb., o archivnictví a spisové službě a o změně některých zákonů, ve znění pozdějších předpisů (dále jen „zákon“). Archiv přejímá v souladu se zákonem od původců dokumenty ve skartačním nebo mimo skartační řízení k trvalému uložení.
3. Původce je, jehož činnost spočívá v
4. Účelem této dohody je mj. zvýšení právní jistoty původce (nad rámec záruk stanovených zákonem) při přebírání, zpřístupňování a využívání archiválií Archivem s tím, že původce z titulu tohoto smluvního vztahu nemá povinnost odstraňovat ze svých informačních systémů předávaná data.
5. Pro odstranění pochybností smluvních stran se sjednává, že v případě licence se jedná o bezúplatnou, časově neomezenou licenci, jejíž podmínky vyplývají z právního řádu České republiky a z této dohody. Licence je poskytována původcem na všechna data, dokumenty, archiválie či vybrané elektronické databáze předané do Archivu do podpisu této dohody, a rovněž další, včetně přírůstkových, pokud se smluvní strany nedohodnou jinak.
6. Cílem sjednání této dohody je zajistit uchování dat k jejich dalšímu využití v celospolečenském zájmu.

II. Předmět dohody

1. Předmětem této dohody je úprava vzájemných práv a povinností smluvních stran při nakládání s archiváliemi a udělení licence původcem, s ohledem na postavení smluvních stran definované shora.

2. Dále je předmětem této dohody závazek smluvních stran zajistit výkon činností dle této dohody tak, aby byl naplněn účel této dohody.

III. Povinnosti Archivu

1. Archiv přebírá od původce archiválie vybrané ve skartačním řízení nebo mimo skartační řízení v souladu se zákonem.
2. Tato dohoda upravuje postup při přebírání a využívání archiválií a vztahuje se na všechny archiválie převzaté Archivem od původce kdykoli před nebo po podpisu této dohody, a to až dokud nebude stranami dohody písemně sjednáno jinak.

IV. Nakládání s archiváliemi

1. Archiv nakládá s archiváliemi a s metadaty archiválií v souladu se zákonem, zejména:
 - a) vede evidenci archiválií, vytváří a spravuje metadata archiválií;
 - b) zpřístupňuje archiválie a jejich metadata v digitální podobě způsobem umožňujícím dálkový přístup prostřednictvím portálu pro zpřístupnění archiválií v digitální podobě;
 - c) veřejnosti umožňuje nahlížet do archiválií po uplynutí zákonné ochranné doby (nyní třicet let) a do všech zveřejněných archiválií;
 - d) vystavuje archiválie;
 - e) na žádost pořizuje výpisy, opisy anebo kopie archiválií v analogové podobě nebo replik archiválií v digitální podobě.
2. Výslovně se uvádí, že převzetím archiválie Archivem nedochází k omezení práva původce nakládat s replikou archiválií nekomerčně i komerčně a jakkoli ji využívat jako své vlastnictví.
3. Strany se dohodly, že archiválie určené původcem v soupisu předávaných archiválií nebudou Archivem s ohledem na jejich citlivou nebo obchodní povahu zveřejněny nebo vystaveny dříve než uplyne X let od jejich převzetí Archivem. Uvedené nevylučuje zákonnou povinnost Archivu zpřístupnit archivně zpracované archiválie veřejnosti, tedy umožnit nahlížení do archiválií starších třiceti let a do všech zveřejněných archiválií v analogové podobě nebo digitální podobě registrovaným uživatelům na dálku prostřednictvím portálu pro zpřístupnění archiválií v digitální podobě, a na žádost badatele pořídit výpis, opis anebo kopii takové archiválie nebo její repliky, to vše po řádné identifikaci badatele v souladu se zákonem a badatelským řádem Archivu.
4. Archiv nakládá s osobními údaji obsaženými v archiváliích v souladu s platnými právními předpisy na ochranu osobních údajů (t. č. GDPR a zákon).
5. Uvedená pravidla nakládání s archiváliemi se plně vztahují i na databáze.

V. Licence

1. Původce tímto uděluje Archivu bezúplatnou nevýhradní licenci k užití archiválií a metadat archiválií, které jsou předmětem autorských práv nebo práv souvisejících s autorskými právy, pro následující způsoby užití:
 - a) rozmnožování v analogové nebo digitální podobě a rozšiřování, zejména v rámci pořizování výpisu, opisu anebo kopie archiválie a jejich předávání badatelské veřejnosti, změně datového formátu a datové struktury za účelem trvalého uchování obsahu archiválie (archivace) nebo sdělování obsahu archiválie, uložení v příslušném digitálním archivu, zpravidla národním digitálním archivu jako

- součástí Národního archivu či pořizování replik archiválie za účelem zpřístupnění nahlížením;
- b) vystavování;
 - c) sdělování veřejnosti způsobem, že kdokoli může mít k archiváliím a metadatům archiválií přístup v místě a v čase podle své vlastní volby, zejména počítačovou nebo obdobnou sítí, a to prostřednictvím portálu pro zpřístupnění archiválií v digitální podobě;
 - d) zařazení do databáze – evidence archiválií.
2. Licencí dle této dohody nejsou dotčeny zákonné licence dle autorského zákona.
 3. V případě, že některý ze souborů převzatých Archivem bude naplňovat znaky databáze ve smyslu autorského zákona, poskytuje původce tímto Archivu nevýhradní oprávnění na vytěžování obsahu databáze nebo její části a na jejich využití pro účely uvedené v odst. 1. tohoto článku. Jakékoli komerční zpřístupnění databáze podléhá souhlasu původce.
 4. Licence dle tohoto článku se poskytují bez omezení.
 5. Archiv je oprávněn licenci dle této dohody postoupit na jinou organizaci pověřenou úkoly v oblasti archivnictví nebo poskytnout třetí osobě pověřené úkoly v oblasti archivnictví podlicenci.
 6. Oprávnění poskytovaná tímto článkem se považují za licenci, jestliže k předmětu ochrany vykonává původce práva (např. jako zaměstnavatel nebo nositel práv výrobce). V případě, kdy je původce nabyvatelem licence s oprávněním poskytnout podlicenci, považují se tato oprávnění za podlicenci a poskytují se v rozsahu, v jakém je původce oprávněn takovou podlicenci poskytnout.
 7. Není-li původcem určeno jinak, vztahuje se tento článek na všechny předané archiválie a datové soubory tvořící databáze v právním smyslu.

VI. Závěrečná ustanovení

1. Tato dohoda nabývá platnosti a účinnosti okamžikem jejího podpisu oběma smluvními stranami. Vzhledem k tomu, že tato dohoda pouze doplňuje a vyjasňuje poměry smluvních stran při výkonu činnosti Archivu, považují smluvní strany její hodnotu za ekonomicky zanedbatelnou.
2. Veškeré změny a doplňky této dohody musí být učiněny formou písemných číslovaných dodatků.
3. Tato dohoda je sepsána ve dvou stejnopisech s platností originálu, z nichž každá ze stran dohody obdrží po jednom výtisku.

V

za Archiv:

V

za původce:

Zápis k úřednímu záznamu o předání archiválií

Původce: Česká republika - *název organizační složky*

IČO:

sídlo:

zastoupení:

1. Archiv v souladu se zákonem o archivnictví a spisové službě (dále jen „Zákon“) přebírá od Původce archiválie vybrané ve skartačním řízení nebo mimo skartační řízení. Součástí tohoto zápisu je soupis předávaných archiválií.
2. Národní archiv nakládá s archiváliemi a s metadaty archiválií v souladu se Zákonem zejména následovně:
 - f) vede evidenci archiválií a vytváří a spravuje metadata archiválií;
 - g) zpřístupňuje archiválie a jejich metadata v digitální podobě způsobem umožňujícím dálkový přístup prostřednictvím portálu pro zpřístupnění archiválií v digitální podobě;
 - h) veřejnosti umožňuje nahlížet do archiválií po uplynutí zákonné ochranné doby (nyní třicet let) a do všech zveřejněných archiválií;
 - i) vystavuje archiválie;
 - j) na žádost pořizuje výpisy, opisy anebo kopie archiválií v analogové podobě nebo replik archiválií v digitální podobě.
3. Výslovně se uvádí, že převzetím archiválie Národním archivem nedochází k omezení práva původce nakládat s archiválií nekomerčně i komerčně a jakkoli ji využívat jako své vlastnictví.
4. Archiválie určené Původcem v soupisu předávaných archiválií nebudou Archivem s ohledem na jejich citlivou nebo obchodní povahu zveřejněny nebo vystaveny dříve než uplyne X let od jejich převzetí Národním archivem. Uvedené nevylučuje zákonnou povinnost Archivu zpřístupnit archiválie veřejnosti, tedy umožnit nahlížení do archiválií starších třiceti let a do všech zveřejněných archiválií v analogové podobě nebo digitální podobě registrovaným uživatelům na dálku prostřednictvím portálu pro zpřístupnění archiválií v digitální podobě, a na žádost badatele pořídít výpis, opis anebo kopii takové archiválie nebo její repliky, to vše po řádné identifikaci badatele a v souladu se zákonem a badatelským řádem Archivu.
5. Archiv nakládá s osobními údaji obsaženými v archiváliích v souladu s platnými právními předpisy na ochranu osobních údajů (nyní GDPR a zákon).
6. Uvedená pravidla nakládání s archiváliemi se plně vztahují i na archiválie, které jsou považovány za databáze.
7. Pokud jsou převzaté archiválie předmětem autorských práv nebo práv souvisejících s autorskými právy, je Národní archiv oprávněn archiválie a metadata archiválií užít v tomto rozsahu:
 - e) rozmnožování v analogové nebo digitální podobě a rozšiřování, zejména v rámci pořizování výpisu, opisu anebo kopie archiválie a jejich předávání badatelské veřejnosti, změně datového formátu a datové struktury za účelem trvalého uchování obsahu archiválie (archivace) nebo sdělování obsahu archiválie, uložení v příslušném digitálním archivu, zpravidla národním digitálním archivu jako součásti Národního archivu či pořizování replik archiválie za účelem zpřístupnění nahlížením;
 - f) vystavování;

g) sdělováním veřejnosti způsobem, že kdokoli může mít k archiváliím a metadatům archiválií přístup v místě a v čase podle své vlastní volby, zejména počítačovou nebo obdobnou sítí, a to zejm. prostřednictvím národního portálu;

h) zařazením do databáze – evidence archiválií.

Tímto nejsou dotčeny zákonné licence vztahující se na činnost Archivu dle autorského zákona.

8. V případě, že některá z archiválií nebo některý soubor převzatý Archivem bude naplňovat znaky databáze ve smyslu autorského zákona, je Archiv oprávněn vytěžovat obsah databáze nebo její části a zužítkovat ho způsoby uvedenými v odst. 7 tohoto článku, nesmí je však jakkoli komerčně nabízet nebo využívat.

9. Odst. 7 a 8 se vztahují pouze na takové archiválie, ke kterým je Česká republika nositelem uvedených práv nebo nabyvatelem licence (a v rozsahu takové licence). Není-li Původcem určeno jinak, vztahuje se tento článek na všechny předané archiválie a datové soubory tvořící databáze v právním smyslu.

V dne

V dne

.....
za Archiv

.....
za Původce

Příloha č. 10 – GIS formáty a jejich trvalé ukládání (katalog)⁷⁰

GeoPackage⁷¹

Mimetype: application/geopackage+sqlite

Extenze: .gpkg

PUID: neuveden (příp.fmt/ 729)

Stručná charakteristika:

Otevřený, na platformě nezávislý a na standardech založený datový formát pro přenos informací, implementovaný jako databázový kontejner SQLite. Byl definován Open Geospatial Consortium (OGC), publikovaný v roce 2014.

Pravidla a požadavky na kontejner SQLite se řídí standardem GeoPackage Encoding Standard, který definuje schéma, včetně definic tabulek, integrity, omezení formátu a omezení obsahu. Je možná jeho validace.

Archivní zhodnocení:

- › obsahuje vektorová i rastrová data
- › textový soubor – ne
- › kontejner - ano
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém - ano
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven (GDAL, QGIS, GeoServer, Esri, ArcGIS, Compusult, Hexagon, Envitia MapLink, GeoTools, National Geospatial-Intelligence Agency, FME aj.)
- › rozšíření – formát má podporu vládních, komerčních i open source organizací
- › velikost – je mírně úspornější než Shapefile, poměrně rychlý při prostorových a databázových filtrech⁷²
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – ponechat nebo převod do GML.
- › ponechat původní formát – ano

Závěr:

Formát je vhodný pro trvalé uchování.

Tento formát doporučujeme využít jako druhou reprezentaci dat v případě, kdy data ve formátu .gml přesahují velikost 1GB. Pokud z nějakého důvodu, proběhlo nebo proběhne rozdělení datasetu na menší části, např. dle krajů, reprezentaci ve formátu .gpkg je vhodné uložit jako celek, tj. nerozdělenou.

⁷⁰ Doporučené formáty a možnosti více datových reprezentací shrnuje [kapitola 2](#) metodiky.

⁷¹ <https://www.ogc.org/standards/geopackage> ; <https://www.geopackage.org/> ; <https://www.loc.gov/preservation/digital/formats/fdd/fdd000520.shtml>; <https://kost-ceco.ch/cms/geopackage.html>

⁷² <https://www.cagi.cz/upload/documents/givs2018/30-givs-2018-cepicky.pdf>
<http://switchfromshapefile.org/>

GML⁷³ (Geographic Information - Geography Markup Language)

Mimetype: application/gml+xml

Extenze: .gml, .xml

PUID: fmt/1047⁷⁴, x-fmt/227

Stručná charakteristika:

Otevřený, na platformě nezávislý formát založený na specifikaci XML, definován Open Geospatial Consortium (OGC) pro vyjádření geografických rysů. Jedná se o modelovací jazyk pro geografické systémy a také o otevřený formát pro výměnu geografických transakcí a prostorových dat. Je založený na standardu ISO 19136: 2007, vytvořen byl v roce 1998. Umožňuje popis schémat a datových sad, výběr profilů pro určité komunity (např. aplikační schéma pro počasí, letectví aj.). Významná je jeho schopnost integrovat všechny formy geografických informací, včetně konvenčních vektorových nebo spojitých objektů, pokrytí a senzorových dat.

Archivní zhodnocení:

- › obsahuje vektorová data
- › textový soubor – ano
- › kontejner – ne
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém – ano, umožňuje mít pro každou geometrii definovaný jiný systém
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven (GDAL, QGIS, Esri, ArcGIS, FME, Geomedia, GeoServer s plug-inem pro GML aj.)
- › rozšíření – tvůrci GIS dat není běžně využíván, ale je rozšířen. Jeho použití pro archivní účely je podporováno aktivitami EU, je na něm vystavěna výměna INSPIRE dat.
- › Velikost – jeho velikost není limitována, ale soubory s velikostí nad 1 GB, lze velmi omezeně zpracovat. Řešením tak může být rozdělení do více souborů (např. dle krajů).
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – ponechat
- › ponechat původní formát – ano

Závěr:

Formát je doporučený pro trvalé uchování, jedná se výstupní formát dle § 23 odst 2 Vyhlášky.

Tento formát doporučujeme využít jako formát, do nějž jsou převáděna vektorová data z méně perspektivních formátů.

⁷³ <https://www.ogc.org/standards/gml>; https://en.wikipedia.org/wiki/Geography_Markup_Language; <https://www.loc.gov/preservation/digital/formats/fdd/fdd000296.shtml> ; https://is.muni.cz/th/p42lc/text_DP.pdf

⁷⁴

<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=1852>

KML⁷⁵ (Keyhole Markup Language)

Mimetype: application/kml+xml

Extenze: .kml, .xml; .kmz

PUID: fmt/244; fmt/724

Stručná charakteristika:

Formát původně vyvinula společnost Keyhole inc. pro prohlížeč prostorových dat Keyhole Earth Viewer. Keyhole inc. byla zakoupena v roce 2004 společností Google, která technologii zahrnuje do portfolia svých produktů. Samotný formát je založený na specifikaci XML, jeho dokumentace byla společností Google zveřejněna a adoptována konsorciem OGC jako standard. Součástí rodiny tohoto formátu je i kontejnerová forma s extenzí .kmz.

Archivní zhodnocení:

- › obsahuje vektorová data
- › textový soubor – ano
- › kontejner – ne, povoleno pro extenzi .kmz
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém – ne, vyžadována příloha A dle specifikace
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven (GDAL, QGIS, Esri, ArcGIS, FME, GeoServer aj.)
- › rozšíření – tvůrci GIS dat není běžně využíván s ohledem na velikostní omezení, ale je dostupný jako vhodný formát k snadnému zobrazování prostorových dat pomocí aplikace Google Earth
- › velikost – limitována na 10MB pro KML, 3MB pro kontejner KMZ
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – ponechat
- › ponechat původní formát – ano

Závěr:

Formát je akceptovaný pro trvalé uchování.

⁷⁵<https://www.ogc.org/standards/kml>;
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000340.shtml>;
<https://www.nationalarchives.gov.uk/PRONOM/fmt/244>;
<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=1523>

EO-GeoJSON, CityJSON (JavaScript Object Notation JSON)⁷⁶

Mimetype: application/geo+json

Extenze: .geojson

PUID: fmt/1367

Stručná charakteristika:

Formát Javascript Object Notation je formát využívaný pro datovou výměnu prostorových prvků včetně neprostorových atributů pomocí všeobecně známého kódování JSON. Implementace GeoJSON podporuje geometrické vektorové typy odvozené ze specifikace OGC Simple Features. Výhodou tohoto formátu je lidská čitelnost, nevýhodou pak absence schématu. Z tohoto důvodu je formát založen na normativním kontextu JSON-LD, který umožňuje explicitně definovat každou vlastnost jako URI, a jeho kódování je určeno schématem JSON [OR18] pro potřeby validace.

Archivní zhodnocení:

- › obsahuje vektorová data
- › textový soubor – ano
- › kontejner – ne
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém – ano, umožňuje mít pro každou geometrii definovaný jiný systém
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven (GDAL, QGIS, Esri, ArcGIS, FME, Mapserver, GeoServer aj.)
- › rozšíření – tvůrci GIS dat není běžně využíván, ale je rozšířen. Jeho použití pro archivní účely je podporováno aktivitami EU, je na něm vystavěna výměna INSPIRE dat.
- › velikost – není limitována, ale soubory s velikostí nad 1 GB, lze velmi omezeně zpracovat. Řešením tak může být rozdělení do více souborů podobně jako u GML, příp. využít rozdělení pomocí sítě.
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – ponechat
- › ponechat původní formát – ano

Závěr:

Formát je vhodný pro trvalé uchování a je možné jej využít jako alternativu preferovaného výstupního formátu GML.

⁷⁶<https://www.loc.gov/preservation/digital/formats/fdd/fdd000382.shtml>;
<http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=2185>;
<https://datatracker.ietf.org/doc/html/rfc7946>; <https://www.ogc.org/standards/eo-geojson>;
<https://docs.ogc.org/cs/20-072r2/20-072r2.html>; <https://github.com/glenrobertson/leaflet-tilelayer-geojson>

Formáty Interlis 1 a Interlis 2⁷⁷

Mimetype: application/Interlis

Verze 1 Extenze: verze 1 .ili (model), .itf (výměna dat)

PUID: verze 1 fmt/1013; fmt/1014;

Verze 2 Extenze: .xtf (model); .xml (výměna dat)

PUID: : verze 2.2 fmt/1012; verze 2.3 fmt/654; fmt/1011; verze 2.3 fmt/653;

Stručná charakteristika:

Formát spravuje instituce KOGIS (koordinace, geoinformace a služby) ve spolupráci s odbornou skupinou eCH INTERLIS; KOGIS poskytuje potřebné informace a dokumenty (více na <https://www.kogis.ch/>). Spolkový topografický úřad provozuje také vlastní informační stránku INTERLIS (<http://interlis.ch/>), kde jsou k dispozici všechny specifikace formátu.

Oproti jiným výměnným formátům pro vektorová data je v INTERLIS datový model uložen odděleně od vlastních prostorových dat, tj. jednotný datový model lze použít pro několik datasetů. U datasetu se tak musí shodovat verze formátu přenosového souboru INTERLIS 1 (.itf Transfer File) a souvisejícího modelu INTERLIS 1 (.ili Model File).

Archivní zhodnocení:

- › obsahuje vektorová data
- › textový soubor – ano
- › kontejner – ne
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém – ano, umožňuje mít pro každou geometrii definovaný jiný systém
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – QGIS, FME, ESRI GDB, PostGIS, nástroje vyvíjené na území Švýcarské federace (ili2c, umleditor, eg/Check*, ilivalidator, ili2db, ili2fme),
- › rozšíření – Jeho využití je prakticky výhradně omezeno na území Švýcarské federace a je využíván i pro archivní účely.
- › velikost – není limitována, ale u souborů s velikostí nad 1 GB, lze předpokládat obdobné obtíže jako u formátů GML a GeoJSON.
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – ponechat
- › ponechat původní formát – ano

Závěr:

Formát je vhodný pro trvalé uchování a je akceptovaný. S ohledem na malou rozšířenost tohoto formátu doporučujeme původcům a tvůrcům prostorových dat využívat formát GML.

⁷⁷ <https://www.interlis.ch>; <https://kost-ceco.ch/cms/interlis-1.html>; <https://kost-ceco.ch/cms/INTERLIS-2.html>; <http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=1819>; <https://www.swisstopo.admin.ch/de/swisstopo/organisation/kogis.html>

Formát TIFF⁷⁸ (Tagged Image File Format nebo Tag Image File Format)

Mimetype: image/tiff

Extenze: .tiff, .tif., .xml

PUID: fmt/353⁷⁹, (fmt/7; fmt/8; fmt/9; fmt/10)

Stručná charakteristika:

TIFF je de facto standardem a zároveň proprietárním formátem společnosti Adobe. Původně byl koncipován jako výměnný formát pro digitalizaci obrázků (vytvořen v roce 1986), poté se díky své univerzálnosti vyvinul v preferovaný formát pro obrazová data.

Definuje jej soubor značek (tagů), které popisují vlastnosti obrázku nebo více obrázků, jež obsahuje. Je velmi flexibilní, může např. využívat různé komprese (např. Packbits, CCITT G3&4, JPEG, LZW⁸⁰).

Specifikace TIFF 6.0⁸¹ definuje tak zvaný „Baseline TIFF“ (základní TIFF), který by měly podporovat všechny aplikace pracující s TIFF a v této verzi je i vyžadován pro trvalé uchovávání.

Archivní zhodnocení:

- › obsahuje rastrová data
- › textový soubor – ne, je binární
- › kontejner – ano
- › obsahuje GIS metadata – ne
- › obsahuje souřadnicový systém – ne, je nutné předávat alespoň se souborem World File TFW⁸²
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – je velmi rozšířený,
- › podpora řady nástrojů a knihoven (GDAL, QGIS, Esri, ArcGIS, FME, GeoServer, GRASS aj.)
- › rozšíření – je běžně využíván pro statickou prezentaci GIS dat, jako rastrový podklad a pro georeferencování
- › velikost – sice umožňuje širokou škálu komprese, pro trvalé uchování jsou akceptována pouze nekomprimovaná data, příp. komprimována bezztrátově.⁸³
- › možnost další interoperability (strojové vytěžení) – ne
- › doporučení pro převod do jiného formátu – ponechat, pokud možno doplnit o související soubory (viz níže) nebo převést do GeoTIFF
- › ponechat původní formát – ano

⁷⁸ <https://www.adobe.io/open/standards/TIFF.html>;

<https://www.loc.gov/preservation/digital/formats/fdd/fdd00022.shtml>;

[https://kost-](https://kost-ceco.ch/cms/kad_tiff_de.html)

[ceco.ch/cms/kad_tiff_de.html](https://training.gismentors.eu/open-source-gis/formaty/rastr.html); <https://training.gismentors.eu/open-source-gis/formaty/rastr.html>

⁷⁹

<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=1099>

⁸⁰ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000135.shtml>

⁸¹ <https://www.adobe.io/content/dam/udp/en/open/standards/tiff/TIFF6.pdf>

⁸² Externí soubor obsahující 6 řádků textu: velikost pixelu ve směru osy X, rotace okolo osy Y, rotace okolo osy X, velikost pixelu ve směru osy Y, souřadnice X středu levého-horního pixelu, souřadnice Y středu levého-horního pixelu. Pokud se soubor worldfile jmenuje stejně jako soubor TIFF a má koncovku .tfw, tak jej většina GIS automaticky použije; <https://www.loc.gov/preservation/digital/formats/fdd/fdd000287.shtml>

⁸³ <https://www.dpconline.org/docs/technology-watch-reports/2478-preserving-raster-images/file>

Závěr:

Formát TIF/TIFF v revizi 6, nekomprimovaný, je vhodný pro trvalé uchování statických obrázků, pro něž je již dnes tzv. výstupním datovým formátem podle § 23 odst. 3 Vyhlášky. Dále je doporučen pro trvalé ukládání digitalizátů⁸⁴ Pro prostorová data je však vhodnější formát, který z TIFF vychází, a to GeoTIFF.

Při ukládání geoprostorových rastrových dat ve formě georeferencovaných rastrových obrázků a souborů se sítí, je doporučeno vytvořit následující sestavu⁸⁵:

1. soubor TIFF pro rastrový objekt
2. soubor TFW pro lokalizaci rastrového objektu
3. soubor PRJ pro uložení a popis použitého koordinačního a referenčního systému (CRS)
4. soubor např. Geopackage⁸⁶ pro uložení dlaždicového schématu nebo prostorového pokrytí

⁸⁴ <https://www.nacr.cz/wp-content/uploads/2019/05/metodika2015.pdf>, str. 21; metodika FADGI, Dostupné z: http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image_Tech_Guidelines_2016.pdf

⁸⁵ Srv. metodika EARK: Guideline for CITS Geospatial, Appendix 2, Long-Term preservation format Profile for Geospatial Raster data using TIFF baseline 6.

⁸⁶ <https://gdal.org/programs/gdaltindex.html>

GEOTIFF⁸⁷

Mimetype: image/tiff

Extenze: .tiff, .tif

PUID: fmt/155⁸⁸

Stručná charakteristika:

GeoTIFF je založen na standardním formátu TIFF v revizi 6, ke kterému jsou přidána metadata určující jeho prostorové umístění a souřadnicový systém (matice čísel je doplněna o georeferenci, tzn. umístění v geografickém prostoru). Formát byl vyvinut na začátku 90. let

Archivní zhodnocení:

- › obsahuje rastrová data
- › textový soubor – ne, je binární
- › kontejner – ano
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém – ano
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven (GDAL, QGIS, Esri, ArcGIS, FME, GeoServer, GRASS aj.)
- › rozšíření – rozšířený, jde o obvyklý výměnný formát pro georeferencované rastrové objekty mezi různými GIS nástroji.
- › velikost – pro trvalé uchování jsou akceptována nekomprimovaná data, příp. komprimována bezztrátově.⁸⁹
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – TIFF nebo ponechat
- › ponechat původní formát – ano

Závěr:

GeoTIFF je vhodný pro trvalé uchování rastrových prostorových dat, ale nepatří mezi současné výstupní datové formáty podle § 23 odst. 3 Vyhlášky.

⁸⁷ <https://www.ogc.org/standards/geotiff>;

<https://www.loc.gov/preservation/digital/formats/fdd/fdd000279.shtml>

⁸⁸ <https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=798>

⁸⁹ <https://www.dpconline.org/docs/technology-watch-reports/2478-preserving-raster-images/file>

JPEG⁹⁰ (Joint Photographic Experts Group Interchange Format)

Mimetype: image/jpeg, application/jpg; application/x-jpg

Extenze: .jpg, .jpeg, .jpe, .jfif, .jfi, .jif

PUID: fmt/42⁹¹; fmt/43⁹²; fmt/44⁹³

Stručná charakteristika:

JPEG je otevřený formát založený na standardu ISO 10918-1:1994, poprvé byl publikován 1992. Jde o nejpoužívanější formát pro výměnu obrázků. Zároveň jde o ztrátovou metodu komprese obrázků do bitového toku a o formát souboru, ve kterém je tento bitový tok uložen.

Archivní zhodnocení:

- › obsahuje rastrovou grafiku
- › textový soubor – ne, je binární
- › kontejner – ne
- › obsahuje GIS metadata – ne
- › obsahuje souřadnicový systém – ne
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven
- › rozšíření – rozšířený jako formát pro export a vizualizace⁹⁴; je využíván např. pro fotografie, není vhodný pro zobrazení textu, linií, ikon, log.
- › velikost – pro obrázky s hladkými plynulými přechody je vhodný a díky kompresi má při zachování kvality menší velikost než např. PNG
- › možnost další interoperability (strojové vytěžení) – ne
- › doporučení pro převod do jiného formátu – ponechat
- › ponechat původní formát – ano

Závěr:

JPEG je vhodný pro trvalé uchování statických obrázků, např. statické vizualizace a náhledy prostorových dat, lze jej použít pro dokumentaci. Patří mezi tzv. výstupní datové formáty podle § 23 odst. 3 Vyhlášky. Je třeba počítat s tím, že kompresní mechanismus je ztrátový.

⁹⁰ ISO/IEC 10918 <https://www.iso.org/standard/18902.html> ;

<https://www.loc.gov/preservation/digital/formats/fdd/fdd000017.shtml> , <https://kost-ceco.ch/cms/jpeg.html>

⁹¹ <https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=667>

⁹² <https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=668>

⁹³ <https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=669>

⁹⁴ Např. používán jako exportní formát z Geoportálu ČUZK

JPEG 2000⁹⁵ (Joint Photographic Experts Group 2000)

Mimetype: image/jp2, image/jpeg2000, image/jpeg2000-image, image/x-jpeg2000-image

Extenze: . .jp2; .jp2, .jpg2, .j2c (.jpc, .j2k)

PUID: x-fmt/392⁹⁶

Stručná charakteristika:

JPEG 2000 je standard pro kompresi obrazu založený na vlnkové transformaci. Soubor vyvinutý skupinou Joint Photographic Experts Group (JPEG) pro statické snímky používá základní kódování JPEG 2000. Byl navržen tak, aby kompenzoval několik nevýhod formátu JPEG. Nabízí zejména metodu ztrátové i bezztrátové komprese. V části 2 normy je definován rozšířený formát souboru (.jpx), který zahrnuje mechanismy pro animaci a pro spojování různých proudů kódu do jednoho obrázku. Je založen na standardu ISO/IEC 15444.

Archivní zhodnocení:

- › obsahuje rastrovou grafiku
- › textový soubor – ne, je binární
- › kontejner – ano
- › obsahuje GIS metadata – ne
- › obsahuje souřadnicový systém – ne
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven
- › rozšíření – v GIS rozšířený jako formát pro export a vizualizace; jinak pro digitalizaci a ukládání obrázků⁹⁷
- › velikost – Při stejném kompresním poměru poskytuje lepší kvalitu nežli starší standard JPEG. Naopak při použití bezztrátové metody dosahuje většinou lepšího kompresního poměru než formát PNG.
- › možnost další interoperability (strojové vytěžení) – ne
- › doporučení pro převod do jiného formátu – převést do výstupního datového formátu dle § 23 odst. 3 Vyhlášky
- › ponechat původní formát – ano

Závěr:

JPEG 2000 je velmi nadějným a akceptovaným formátem⁹⁸ pro trvalé uchování statických obrázků, vizualizace, náhledy prostorových dat a pro další součásti dokumentace. Není však výstupním datovým formátem podle § 23 odst. 3 Vyhlášky, a proto by měl být v datovém balíčku doplněn o reprezentaci ve výstupním formátu.

⁹⁵ ISO/IEC 15444-1:2004: Information technology — JPEG 2000 image coding system: Core coding system

<https://www.iso.org/standard/37674.html> ;

<https://www.loc.gov/preservation/digital/formats/fdd/fdd000143.shtml>; <https://kost-ceco.ch/cms/jpeg2000.html>

⁹⁶ <https://www.nationalarchives.gov.uk/PRONOM/x-fmt/392>

⁹⁷ metodika FADGI, Dostupné

z: http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image_Tech_Guidelines_2016.pdf

⁹⁸ Např. specifikace PDF/A 2 umožňuje jeho použití (srv. ISO 19005-2:2011), je vhodný pro trvalé ukládání digitalizátů; JPEG 2000 - a Practical Digital Preservation Standard? DPC Technology Watch Report 08-01 2008 Dostupné z: <http://www.dpconline.org/docs/technology-watch-reports/87-jpeg-2000-a-practical-digital-preservation-standard/file>

PNG⁹⁹ (Portable Network Graphics)

Mimetype: image/png

Extenze: .png

PUID: x-fmt/11; x-fmt/12; x-fmt/13

Stručná charakteristika:

PNG byl vytvořen s cílem poskytovat bezztrátový, přenosný, komprimovaný formát pro uložení rastrových dat. Lze jím nahradit jiné formáty jako např. TIFF. Je vhodný pro použití v online prostředí, podporuje indexování barev, stupně šedi, gama, chromatičnost a detekuje poškození souboru.

Archivní zhodnocení:

- › obsahuje rastrovou grafiku
- › textový soubor – ne, je binární
- › kontejner – ne
- › obsahuje GIS metadata – ne
- › obsahuje souřadnicový systém – ne
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven
- › rozšíření – rozšířený jako formát pro export a vizualizace; tak i pro digitalizaci a ukládání obrázků
- › velikost – Při použití bezztrátové metody dosahuje většinou horšího kompresního poměru než formát JPEG 2000.
- › možnost další interoperability (strojové vytěžení) – ne
- › doporučení pro převod do jiného formátu – ponechat, výstupní formát dle § 23 odst. 3 Vyhlášky
- › ponechat původní formát – ano

Závěr:

PNG je vhodný pro trvalé uchování statických obrázků, např. statické vizualizace a náhledy prostorových dat, lze jej použít pro dokumentaci. Patří mezi tzv. výstupní datové formáty podle § 23 odst. 3 Vyhlášky. Je doporučený od Library of Congress pro uchování digitálních fotografií, digitálních obrázků a dále 2D a 3D počítačově vytvořených rasterových obrázků.

⁹⁹ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000153.shtml>;
<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=666>;
<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=1740>;
<https://gdal.org/drivers/raster/png.html> - raster-png

GIF¹⁰⁰ (Graphics Interchange Format)

Mimetype: image/gif

Extenze: .gif

PUID: x-fmt/3; x-fmt/4

Stručná charakteristika:

Formát Graphics Interchange format je bitmapovým obrázkem široce využívaným na síti internet. Jedná se prakticky o kontejner, který obsahuje záhlaví, popisnou část, kterou následuje jedna nebo více obrazových sekcí a ukončovacích částí. Podporována je 1 až 8 bitová hloubka barev, prokládání a průhlednost, obrázky jsou ukládány pomocí bezztrátové komprese LZW.

Archivní zhodnocení:

- › obsahuje rastrovou grafiku
- › textový soubor – ne, je binární
- › kontejner – ano
- › obsahuje GIS metadata – ne
- › obsahuje souřadnicový systém – ne
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven
- › rozšíření – využívaný jako formát pro ukládání obrázků obsahujících nízký počet barev nebo pro malé dynamické obrazové dokumenty (animace a klipy).
- › možnost další interoperability (strojové vytěžení) – ne
- › doporučení pro převod do jiného formátu – ponechat
- › ponechat původní formát – ano

Závěr:

GIF je výstupním formátem pro dynamické obrazové dokumenty dle § 23 odst. 4 Vyhlášky a jedná se vhodný formát k uložení dokumentace prostorových dat, kdy je součástí dokumentace např. webová stránka, která data zveřejňovala nebo propagovala.

¹⁰⁰[http://giflib.sourceforge.net/;](http://giflib.sourceforge.net/)
<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=619;>
[https://www.loc.gov/preservation/digital/formats/fdd/fdd000133.shtml;](https://www.loc.gov/preservation/digital/formats/fdd/fdd000133.shtml)
<https://gdal.org/drivers/raster/gif.html> - raster-gif

Esri Shapefile¹⁰¹

Mimetype: pro .shp a .shx x-gis/x-shapefile; pro dbf application/octet-stream nebo text/plain
Extenze: .shp + .shx + .dbf a další (.sbn; .sbn; .atx; .fbi; .fbi; .ain; .aih; .ixs; .mxx; .prj; .xml; .cpg.)

PUID: x-fmt/235102, fmt/277103

Stručná charakteristika:

Jedná se o proprietární formát společnosti Esri, ale dokumentace byla v roce 1998 zveřejněna a licence uvolněna s cílem podpořit interoperabilitu mezi aplikacemi GIS. Formát Shapefile uchovává netopologické geometrické a atributové informace o prostorových prvcích v datové sadě. Skládá se minimálně z hlavního souboru *.shp - geometrie (shape), indexového souboru *.shx - propojení geometrie a atributů (index) a tabulky dBASE *.dbf - atributy (databáze). Soubory mají stejný název a liší se koncovkou.

Archivní zhodnocení:

- › obsahuje vektorová data
- › textový soubor – ne, je binární
- › kontejner – ne, jedná se o skupinu formátů
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém – ano
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven (GDAL, QGIS, Esri, ArcGIS, FME, GeoServer, GRASS aj.)
- › rozšíření – Je podporován prakticky všemi nástroji GIS a ve své době se stal v podstatě oborovým standardem a formátem pro výměnu dat díky uvolnění dokumentace a možnosti jeho implementace v software třetích stran. Je však zastaralý (problémy plynoucí z kombinace více souborů, omezení v názvech atributů, omezení celkové velikosti, neznámá znaková sada, neobsahuje topologii)¹⁰⁴.
- › velikost – omezena na 2 GB. Formát je nejužitečnější pro rychlý zápis jednoduchých prvků a atributů, protože ve formátu Shapefile existují omezení týkající se geometrie i atributů. Tato omezení však mohou způsobit ztrátu dat při použití Shapefile pro obsah nebo výměnu složité geometrie nebo atributů.
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – GML, příp. GeoPackage; model objektu v PDF/A
- › ponechat původní formát – ano

Závěr:

Esri Shapefile je sice velice rozšířený, není však vhodný pro trvalé ukládání. Proto je vyžadována jeho migrace do GML nebo GeoPackage, vizualizaci náhledu lze uložit ve výstupním datovém formátu (§ 23 odst. 2 a 3 Vyhlášky).

¹⁰¹<https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/whitepapers/pdfs/shapefile.pdf>;
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000280.shtml>; <https://training.gismentors.eu/open-source-gis/formaty/vektor.html#shapefile>

¹⁰²

<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=328>

¹⁰³

<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=1017>
¹⁰⁴ https://www.cagi.cz/upload/documents/givs2015/cepicky_shapefile_must_die.pdf

Esri Grid¹⁰⁵

Mimetype: image /adf;

Extenze: .adf,

PUID: x-fmt/218

Stručná charakteristika:

Esri Grid je proprietární formátem používaným v software ArcGis pro rastrová data. Jedná se o standardní rastrová data v podobě sítě polí v řádcích a sloupcích, tzv. dlaždice, které reprezentují prostorové fenomény. Rastrová data v tomto formátu v sobě obsahují souřadnicová data. Jednotlivé bloky jsou uloženy v komprimované podobě a ve vysoké kvalitě.

Archivní zhodnocení:

- › obsahuje rastrová data
- › textový soubor – ano, může mít i ASCII podobu
- › kontejner – ne
- › obsahuje GIS metadata – ne
- › obsahuje souřadnicový systém – ano
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven (GDAL, QGIS, Esri, ArcGIS, FME, GeoServer, GRASS aj.)
- › rozšíření – Je podporován prakticky všemi nástroji GIS a ve své době se stal v podstatě oborovým standardem a formátem pro výměnu dat díky uvolnění dokumentace a možnosti jeho implementace v software třetích stran.
- › velikost – omezena povolenou kvalitou 4 000 000 x 4 000 000 polí
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – TIFF, GeoTIFF
- › ponechat původní formát – ano

Závěr:

Esri Grid není vhodný pro trvalé uložení, jedná se sice o nativní formát využívaný v GIS systému, který je široce využíván po celém světě a přestože se jedná o textový formát, doporučujeme datový balíček obohatit o datovou reprezentaci v některém z výstupních formátů dle § 23 odst. 2 nebo odst. 3 Vyhlášky.

¹⁰⁵ <https://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/esri-grid-format.htm>;
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000421.shtml>;
<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=309>;
<https://gdal.org/drivers/raster/aig.html> - raster-aig

Esri BIL, BIP, BSQ¹⁰⁶

Mimetype: image/BIL; image/BIP; image/BSQ

Extenze: .bil; .bip; .bsq

PUID: x-fmt/1238; x-fmt/1239; x-fmt/1240

Stručná charakteristika:

BIL, BIP a BSQ samy o sobě nejsou obrazovými formáty, ale jsou to pouze schémata pro ukládání skutečných hodnot pixelu obrázku do souboru. Tyto soubory podporují zobrazení jednopásmových a vícepásmových obrázků a zpracovávají černobílá data, data ve stupních šedi, pseudobarvy, věrné barvy a multispektrální obrazová data.

Data ve formátech BIL, BIP a BSQ mají binární podobu a aby byla správně interpretována ze strany software, musí mít přidružené záhlaví souboru ASCII. Tento soubor záhlaví obsahuje pomocná data o obrázku, jako je počet řádků a sloupců v obrázku, pokud existuje barevná mapa, a zeměpisná šířka a délka. Na základě implementace formátu BIP je založen známý formát TIFF.¹⁰⁷

Archivní zhodnocení:

- › obsahuje rastrová data
- › textový soubor – ne, jedná se o binární data
- › kontejner – ne
- › obsahuje souřadnicový systém – ano
- › obsahuje GIS metadata – ano
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – ArcGIS Desktop/PRO, FME, QGIS)
- › rozšíření – Formát je známý, ale nahrazen modernějším TIFF
- › velikost – není limitována
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – migrovat do GeoTIFF/TIFF
- › ponechat původní formát – ano

Závěr:

Přestože jsou Esri formáty BIL, BIP a BSQ de facto předchůdci výstupního formátu TIFF, se nejedná o formáty vhodné k trvalému uchování. Proto doporučujeme datový balíček obohatit o datovou reprezentaci v některém z výstupních formátů dle § 23 odst. 2 nebo odst. 3 Vyhlašky zejména z důvodu zastaralosti těchto formátů.

¹⁰⁶ <https://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/bil-bip-and-bsq-raster-files.htm>; <https://www.loc.gov/preservation/digital/formats/fdd/fdd000305.shtml>; <https://www.loc.gov/preservation/digital/formats/fdd/fdd000304.shtml>; <https://www.loc.gov/preservation/digital/formats/fdd/fdd000306.shtml>;

<https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=2056>; <https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=2057>; <https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=2058>;

¹⁰⁷ Conference Proceedings of Mining multimedia and complex data: KDD Workshop MDM/KDD 2002, str. 104 dostupné z <https://books.google.com/books?id=Yiw1EaNUOQ8C&pg=PA104&lpg=PA104&dq=TIFF+BIP+band-interleaved>.

Formát Lidar/LiDAR/LADAR¹⁰⁸

Mimetype: Application/Lidar

Extenze: .laz, .las

PUID: fmt/367; fmt/368; fmt/369

Stručná charakteristika:

Binární datový formát využívaný v mnoha vědeckých oborech pro výměnu dat dálkového optického výzkumu Země a také k tvorbě rastrových map ve vysokém rozlišení.

Archivní zhodnocení:

- › obsahuje rastrová data
- › textový soubor – ne, je binární
- › kontejner – ne
- › obsahuje souřadnicový systém – ano
- › obsahuje GIS metadata – ano
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – ArcGIS Desktop/PRO, DWG True View, FME, Global Mapper, QGIS)
- › rozšíření – Formát je rozšířený
- › velikost – není limitována
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – migrovat do GeoTIFF/TIFF
- › ponechat původní formát – ano
- › limitovány pouze Variabilní datové typy na 65,535 bytů, rozšířené mohou mít větší objem

Závěr:

Formát není vhodný pro trvalé uchování. Pro tento formát je vyžadována migrace druhé datové reprezentace do formátu GeoTIFF nebo TIFF, příp. uložení ve verzi las 1.4.

¹⁰⁸<https://www.loc.gov/preservation/digital/formats/fdd/fdd000418.shtml>;
<https://library.carleton.ca/guides/help/lidar-formats>;
<http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=listReport>;
<https://www.asprs.org/divisions-committees/lidar-division/laser-las-file-format-exchange-activities>;
http://www.asprs.org/wp-content/uploads/2019/03/LAS_1_4_r14.pdf

GPJ, GRD, GRF, GSB, GSI, GSR, SRF ¹⁰⁹ (Formáty Golden Software)

Mimetype: application/Dataset

Extenze: .gpj; .grd; .gsb; .gsi; .gsr; .gsr2; .srf

PUID: není

Stručná charakteristika:

Proprietární datové formáty používané nástroji vyvíjenými společnostmi Golden Software. Nástroje Surfer, Grapher, Strater, Voxler jsou využívány pro tvorbu prostorových dat, jako je modelování zemského povrchu, půdní stratigrafie, důlní díla, znečištění půdy, znečištění podzemních vod, seismologie, meteorologie, archeologie atd. a umožňují analýzu včetně 3D modelování těchto fenoménů.

Archivní zhodnocení:

- › formáty se dělí na soubory s mřížkou, datové sestavy, rastrové mapy a vektorové mapy
- › textový soubor – ne, soubory jsou binární, výjimkou jsou pouze mřížková data .grd, která mohou mít formu ASCII
- › kontejner – ano
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém – ne (výjimkou jsou soubory .gsr)
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – omezený – omezená podpora (nástroje Golden Software, GDAL, FME)
- › rozšíření – Golden Software je v České republice omezeně využíván pro mapování, modelování, projektování a analýzy prostorových fenoménů
- › velikost – limity nezjištěny
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – ponechat
- › ponechat původní formát – ano

Závěr:

Formáty používané nástroji Golden Software nejsou vhodné pro trvalé uchování, ale s ohledem na rizika datové ztráty spojené s datovou konverzí je doporučujeme do datového balíčku SIP uložit jako jednu datovou reprezentaci v originálních formátech. Datový balíček SIP zároveň vyžadujeme obohatit o datovou reprezentaci v některém z metodikou doporučených formátů. Pro statická rastrová data dle § 23 odst. 3 Vyhlášky, pro dynamická rastrová data dle § 23 odst. 4 Vyhlášky, PDF/A-4. Vektorové mapy lze exportovat do formátu SHP File, a ten pomocí open source nástroje QGIS konvertovat např. do formátu Geopackage. Pro datové soubory s hodnotami je možné využít např. formát SIARD nebo CSV ve znakové sadě UTF8 a doplnit metadata, která umožní import dat do databáze.

¹⁰⁹ <https://support.goldensoftware.com/hc/en-us/articles/226629527-The-difference-between-GPJ-and-GRF-files-in-Grapher>; http://surferhelp.goldensoftware.com/subsys/subsys_gsi_grd_file_format.htm; http://grapherhelp.goldensoftware.com/subsys/subsys_gsb_goldensoftware_boundary_file_desc.html; http://surferhelp.goldensoftware.com/subsys/subsys_gsigi_hid_gsigi_filedesc.htm; <https://gdal.org/drivers/raster/gsg.html>; <https://support.goldensoftware.com/hc/en-us/articles/227882008-Surfer-Supported-File-Formats>; <https://support.goldensoftware.com/hc/en-us/articles/115005530868-What-File-Format-Should-I-Export-from-Surfer->

SIARD (Software-Independent Archiving of Relational Databases)¹¹⁰

Mimetype: application/database

Extenze: .siard

PUID: verze 1.0 fmt/161; verze 2.0 fmt/995, verze 2.1. fmt/1196)

Stručná charakteristika:

Na software nezávislý formát pro archivaci relačních databází, který je otevřený, založený na standardech, je využíván k trvalému uložení relačních databází a umožňuje migrace mezi relačními prostředími založenými na SQL Standardu. Formát byl vyvinut ve Švýcarsku pro potřeby Švýcarského federálního archivu. Jedná se o kontejner ve formátu ZIP, který obsahuje tabulková data, potřebná metadata ve formátu XML a volitelně extrahuje objekty BLOB a CLOB.

Archivní zhodnocení:

- › obsahuje tabulková data a popisná metadata ve výstupním formátu XML
- › textový soubor – ano
- › kontejner - ano
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém - ano
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven (Database Preservation Toolkit, SIARDGui, Full Convert)
- › rozšíření – formát má podporu vládních, komerčních i open source organizací
- › velikost – závislá na velikosti ukládané databáze
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu – ponechat
- › ponechat původní formát – ano

Závěr:

Formát je doporučený pro trvalé uchování. Tento formát doporučujeme využít pro části GIS, které jsou běžnou relační databází založenou na SQL Standardu bez extenzí pro práci s prostorovými daty a obsahují běžná tabulková data. Tento formát prozatím není vhodný pro vlastní prostorová data, protože SQL Standard neumožňuje některé z funkcí, které jsou obvyklé v systémech GIS. Testování tohoto formátu a případné úpravy pro potřeby trvalého uložení GIS jsou teprve na svém počátku.

¹¹⁰<https://www.loc.gov/preservation/digital/formats/fdd/fdd000426.shtml>;
<http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=2006>;
https://kost-ceco.ch/cms/siard_de.html; <https://github.com/DILCISBoard/SIARD>;
<https://github.com/keeps/dbptk-developer/releases>; <https://github.com/sfa-siard/SiardGui/releases>;

CSV (Comma Separated Values)¹¹¹

Mimetype: text/csv

Extenze: .csv

PUID: x-fmt/18, CSV schema fmt/800

Stručná charakteristika:

Jednoduchý textový formát využívaný pro ukládání tabulkových dat numerických i textových. Jednotlivé řádky jsou odděleny zalomením a hodnoty ve sloupcích oddělovačem. Každý řádek by měl obsahovat stejné množství polí. Pokud pole obsahuje speciální znak, musí být označeno uvozovkami „pole“. Je povoleno, aby první řádek obsahoval hlavičku s názvem jednotlivých sloupců.

Archivní zhodnocení:

- › obsahuje tabulková data a volitelně popisná metadata
- › textový soubor – ano
- › kontejner – ne
- › obsahuje GIS metadata – ne
- › obsahuje souřadnicový systém – ne
- › neexistuje specifikace
- › rozšíření – formát je široce využíván
- › velikost – závislá na velikosti uložené tabulky
- › možnost další interoperability (strojové vytěžení) – ano
- › doporučení pro převod do jiného formátu
- › ponechat původní formát – ano

Závěr:

Formát je akceptovaný pro trvalé uchování a je možné jej využít pro uložení tabulkových hodnot využívaných ve spojení s prostorovými daty. Pro tabulkové hodnoty, které jsou využívány v systémech GIS, doporučujeme využít formátu SIARD.

¹¹¹<https://www.loc.gov/preservation/digital/formats/fdd/fdd000323.shtml>;
<http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=45>;
<http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=1600>

Formáty PDF/A¹¹²

Mimetype: application/pdf

Extenze: .pdf

PUID: fmt/95, fmt/354, fmt/476, fmt/477, fmt/478, fmt/479, fmt/480, fmt/481

Stručná charakteristika:

PDF/A je standardizovaná verze formátu PDF (Portable Document Format) podle normy ISO 19005, která je určena pro archivaci a trvalé uchování elektronických dokumentů. PDF/A se od PDF liší tím, že zakazuje funkce nevhodné pro trvalé uchování, jako je propojování písem (na rozdíl od vkládání písem), šifrování, komprese, použití JavaScriptu a spustitelných souborů. Požadavky ISO na prohlížeče souborů PDF/A zahrnují pokyny pro správu barev, podporu vložených písem a uživatelské rozhraní pro čtení vložených poznámek. Dosud existují 4 základní verze založené na standardech, všechny jsou výstupním datovým formátem pro textové nebo textové a obrazové dokumenty. Od verze PDF/A-2 může být formát použit jako kontejner, pak je v České republice vyžadováno, aby soubory, které obsahuje, odpovídaly pro trvalé uložení Formátovým pravidlům NA.

Standardy PDF/A:

PDF/A-1 - ISO 19005-1 (založen na PDF 1.4)

PDF/A-2 - ISO 19005-2 (založen na PDF 1.7, ISO 32000-1:2008)

PDF/A-3 - ISO 19005-3 (založen na PDF 1.7, ISO 32000-1:2008 s podporou vložených souborů)

PDF/A-4 - ISO 19005-4 (založen na PDF 2.0, ISO 32000-2:2020)

Archivní zhodnocení:

- › textový soubor – ne, je binární
- › kontejner – od verze PDF/A 2 ano
- › obsahuje GIS metadata – ne
- › obsahuje souřadnicový systém – ne
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný – podpora řady nástrojů a knihoven
- › rozšíření – rozšířený jako formát pro export a vizualizace, celosvětově podporovaný formát pro trvalé uchování
- › možnost další interoperability (strojové vytěžení) – pokud obsahuje strojově čitelnou vrstvu
- › doporučení pro převod do jiného formátu – je výstupním datovým formátem podle formátu dle § 23 odst. 2 Vyhlášky, lze využít pro statické vizualizace, náhledy
- › ponechat původní formát – ano

Závěr:

Jedná se o archivační formát, již dnes patří mezi tzv. výstupní datové formáty pro texty, kombinované texty s obrázkem. Není určen pro uložení strojově zpracovatelných dat. Při použití PDF/A-3 pro trvalé uložení je vyžadováno, aby soubory, které obsahuje, odpovídaly Formátovým pravidlům NA.

Formát PDF Geospatial¹¹³

¹¹² [https://www.pdflib.com/pdf-knowledge-base/pdfa/the-pdf-a-standards/?gclid=EAlaIqobChMlr_SgtqrV8gIVE5_VCh3zfAzhEAMYASAAAEgKiOvD_BwE](https://www.pdflib.com/pdf-knowledge-base/pdfa/the-pdf-a-standards/?gclid=EAlaIqobChMlr_SgtqrV8gIVE5_VCh3zfAzhEAMYASAAAEgKiOvD_BwE;); <https://www.loc.gov/preservation/digital/formats/fdd/fdd000030.shtml> ; <https://kost-ceco.ch/cms/pdf-a-2.html> ; <https://kost-ceco.ch/cms/pdf-a-1.html> ; [kost-ceco | PDF/A-3](#)

Mimetype: application/pdf

Extenze: . pdf

PUID: fmt/1129

Stručná charakteristika:

Formát PDF Geospatial je založen na standardu ISO 32000-1:2020, jedná se o kontejnerové PDF ve verzi 2.0 a skládá se ze záhlaví, těla, které obsahuje objekty tvořící obsah dokumentu, tabulku s odkazy a závěrečné sekce, která umožňuje čitelnost jednotlivých částí kontejnerového PDF.

Archivní zhodnocení:

- › textový soubor – ne, je binární
- › kontejner – ano
- › obsahuje GIS metadata – ano
- › obsahuje souřadnicový systém – ano
- › dostupnost (nástroje na tvorbu, prohlížení, převod) – dostupný
- › rozšíření – Lze využít jako formát pro export a vizualizace
- › možnost další interoperability (strojové vytěžení) – pokud obsahuje strojově čitelnou vrstvu
- › doporučení pro převod dle § 23 odst. 2 Vyhlášky, nelze využít pro statické vizualizace, náhledy
- › ponechat původní formát – ne

Závěr:

Formát není vhodný pro trvalé uložení v digitálním archivu, je vyžadována datová konverze do PDF/A-4.

¹¹³ <https://helpx.adobe.com/acrobat/using/geospatial-pdfs.html> - about geospatial pdfs;
<https://gdal.org/drivers/raster/pdf.html>; <https://www.nationalarchives.gov.uk/PRONOM/fmt/1129>;
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000474.shtml>; <https://www.pdflib.com/pdf-knowledge-base/pdf-20/>

Formáty CAD (.dwg, .dwt, .dxf, .dgn a další)

CAD (Computer-Aided Design¹¹⁴) označuje počítačem podporované návrhy různých výrobků, jako jsou automobily, letadla a budovy. V užším smyslu se jím rozumí kreslení stavebních a architektonických plánů s počítačovou podporou. Produktem je obvykle 2D nebo 3D vektorová grafika s metadaty pro jednotlivé vektorové prvky. Archivace tedy představuje podobný problém jako ukládání dat ze světa GIS, s tím rozdílem, že je využíván jiný software a pro výměnu mezi programy slouží jiné formáty.

Pokud budou nabízena i data ve formátech CAD, pak je doporučeno ponechat původní formát a k němu vytvořit reprezentaci (model) v obecně známém výstupním datovém formátu (PDF/A, příp. JPEG, PNG, srv. Formátová pravidla NA¹¹⁵). Jako velmi perspektivní pro trvalé uložení se jeví proprietární, ale značně rozšířený formát DWG (Drawing).¹¹⁶ Postup ukládání a příp. tvorby datových balíčků pro data z CAD systémů není v současné době však podrobně definován, konkrétní možnosti je třeba konzultovat ad hoc (podle účelu a způsobu budoucího využití, možností původce a dalších okolností).¹¹⁷

¹¹⁴ https://cs.wikipedia.org/wiki/Computer_aided_design

¹¹⁵ Dostupné ze stránky <https://validatorsip.nacr.cz/>; odkaz na aktuální Formátová pravidla NA

¹¹⁶ <https://www.loc.gov/preservation/digital/formats/fdd/fdd000445.shtml>; <https://kost-ceco.ch/cms/dwg.html>;

¹¹⁷ DPC Zpráva o sledování technologií – Uchování CAD (2013), Dostupné z: <https://www.dpconline.org/docs/technology-watch-reports/896-dpctw13-02-pdf/file>;

Příloha č. 11 – Výběrový přehled formátů dle Formátových pravidel NAČR

Veřejnoprávní původci jsou povinni ukládat digitální dokumenty ve výstupních datových formátech podle § 23 Vyhlášky. Původci, kteří nevykonávají spisovou službu v elektronické podobě v elektronických systémech spisové služby, převádějí dokumenty v digitální podobě určené k výběru archiválií mimo skartační řízení do výstupního datového formátu a opatří ho metadaty nejpozději při přípravě výběru archiválií.¹¹⁸

Národní archiv zveřejňuje registr tzv. Formátových pravidel¹¹⁹, kde doporučuje vhodné migrační strategie, upřesňuje, kdy není převod do výstupního datového formátu vyžadován, příp. kdy je doporučeno předávat k uložení komponentu i v originálním formátu.

Legenda

PUID: identifikátor PRONOM

Kategorie: pomocná položka pro specifikaci typu formátu

„Výstupní formát“ a „Výstupní formát alternativně I a II“: pokud jsou vyplněny, uveden odkaz na odstavec § 23 Vyhlášky, podle něž je nutné daný formát převést na výstupní datový formát; nebo je uveden jiný doporučený formát; nebo je vyžadováno rozbalení a poté převedení jednotlivých komponent.

Originál vždy (doporučeno) – pokud je uvedeno „ano“, vyžaduje se předání i nativního formátu

Komentář – specifikuje podrobnosti

Odpovídá Formátovým pravidlům = tzn. ve Formátových pravidlech je alespoň v jednom ze sloupců „Výstupní formát“ nebo „Výstupní formát alternativně“ označeno jako výstupní datový formát nebo možno ponechat. Neodpovídá Formátovým pravidlům = tzn. ve Formátových pravidlech je v obou sloupcích

Verze 42, zveřejněno 7. 6. 2021

¹¹⁸ § 13 odst. 5 Zákona.

¹¹⁹ <http://validatorsip.nacr.cz/>

PUID	Název	Verze	Přípona	Kategorie	Výstupní formát	Výstupní formát alternativně I	Výstupní formát alternativně II	Originál vždy (doporučeno)	Komentář
fmt/277	ESRI Arc/View Shapefile Index		shx	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	Reprezentaci ve VDF (§ 23 odst. 2, 3) či GML společně pro celý Shapefile.
x-fmt/57	Ventura Publisher Vector Graphics		gem	Image (Vector)	§ 23 odst. 2			ano	
fmt/319	ESRI Spatial Index File		sbn, sbx	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	Reprezentaci ve VDF (§ 23 odst. 2, 3) či GML společně pro celý Shapefile.
fmt/320	ESRI Shapefile Projection (Well-Known Text) Format		prj	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	Reprezentaci ve VDF (§ 23 odst. 2, 3) či GML společně pro celý Shapefile.
fmt/321	ESRI Shapefile Header Index		aih	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	Reprezentaci ve VDF (§ 23 odst. 2, 3) či GML společně pro celý Shapefile.
fmt/332	ESRI Arc/View Project	3.x	apr, def	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	
fmt/149	JTIP (JPEG Tiled Image Pyramid)			Image (Raster)	§ 23 odst. 3				
fmt/150	JPEG-LS		jls	Image (Raster)	výstupní formát				
Fmt/151	JPX (JPEG 2000 part 2)		jpg, jpx	Image (Raster)	§ 23 odst. 3				
Fmt/153	Tagged Image File Format for Image Technology		tif, tiff	Image (Raster)	§ 23 odst. 3				

	(TIFF/IT)								
fmt/154	Tagged Image File Format for Electronic Photography (TIFF/EP)		tif, tiff	Image (Raster)	§ 23 odst. 3				
fmt/155	Geographic Tagged Image File Format (GeoTIFF)		tif, tiff	Image (Raster)	§ 23 odst. 3				
fmt/156	Tagged Image File Format for Internet Fax (TIFF-FX)		tfx, tif,tiff	Image (Raster)	§ 23 odst. 3				
fmt/161	SIARD (Software-Independent Archiving of Relational Databases)	1.0	siard	Database	§ 23 odst. 6	individuální posouzení archivem			
fmt/353	Tagged Image File Format		tif, tiff	Image (Raster)	§ 23 odst. 3				
x-fmt/109	Scalable Vector Graphics Compressed		svgz	Image (Vector)	§ 23 odst. 2			ano	
Fmt/362	GSSI SIR-10 RADAN data file		dzt	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	

fmt/ 367	ESRI World File Format		bilw, blw, bpw, btw, jgw, jpgw, pgw, rasterw, tfw, tiffw	Image (Raster)	§ 23 odst. 3				
x- fmt/ 150	Visual FoxPro Database Container File		dcx	Database	§ 23 odst. 6	individuální posouzení archivem			
fmt/ 402	Truevision TGA Bitmap	2.0	icb, tga, vda, vst	Image (Raster)	§ 23 odst. 3				
fmt/ 3	Graphics Interchange Format	87a	gif	Image (Raster)	§ 23 odst. 3				
fmt/ 413	Scalable Vector Graphics Tiny	1.2	svg	Image (Vector)	§ 23 odst. 2			ano	
x- fmt/ 185	Raw Bitmap		raw	Image (Raster)	§ 23 odst. 3				
fmt/ 450	VectorWorks	12.5	vwx	Image (Vector)	§ 23 odst. 2			ano	
fmt/ 451	VectorWorks	2009	vwx	Image (Vector)	§ 23 odst. 2			ano	
x- fmt/ 218	ESRI Arc/Info Binary Grid		adf	GIS	ponechat	individuální posouzení archivem	GML	ano	
fmt/ 463	JPM (JPEG 2000 part 6)		jpm	Image (Raster)	§ 23 odst. 3				
fmt/ 4	Graphics Interchange Format	89a	gif	Image (Raster)	výstupní formát				
x-fmt /225	ESRI MapInfo Data File		mid	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	Reprezentaci ve VDF společně pro .mid a .mif

x- fmt/ 226	ESRI Arc/Info Export File		e00, e01, e02, e03, e04, e05, e06, e07, e08, e09, e10, e11, e12, e13, e14, e15, e16, e17, e18, e19, e20, x00	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	
fmt/ 41	Raw JPEG Stream		jpe, jpeg, jpg	Image (Raster)	výstupní formát				
x- fmt/ 231	ESRI MapInfo Export File		mif	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	Reprezentaci ve VDF společně pro .mid a .mif
fmt/ 42	JPEG File Interchange Format	1.00	jpe, jpeg, jpg	Image (Raster)	výstupní formát				
x- fmt/ 235	ESRI Arc/View ShapeFile		shp	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	Reprezentaci ve VDF (§ 23 odst. 2, 3) či GML společně pro celý Shapefile.
fmt/ 43	JPEG File Interchange Format	1.01	jpe, jpeg, jpg	Image (Raster)	výstupní formát				
fmt/ 44	JPEG File Interchange Format	1.02	jpe, jpeg, jpg	Image (Raster)	§ 23 odst. 3				
fmt/ 529	JPEG Network Graphics		jng	Image (Raster)	§ 23 odst. 3				
x- fmt/ 317	ESRI Arc/View Project	2.x	apr	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	
fmt/ 590	JPEG Extended		jxr, wdp	Image (Raster)	§ 23 odst. 3				

	Range								
x- fmt/ 357	Scanstudio 16-Colour Bitmap		adc	Image (Raster)	§ 23 odst. 3				
x- fmt/ 367	Truevision TGA Bitmap	1.0	afi, bpx, icb, tga, vda, vst	Image (Raster)	§ 23 odst. 3				
fmt/ 684	Vectorworks	2015	vwx	Image (Vector)	§ 23 odst. 2			ano	
fmt/ 686	Vectorworks	2010	vwx	Image (Vector)	§ 23 odst. 2			ano	
fmt/7	Tagged Image File Format	3		Image (Raster)	výstupní formát				
fmt/8	Tagged Image File Format	4		Image (Raster)	výstupní formát				
fmt/9	Tagged Image File Format	5		Image (Raster)	výstupní formát				
Fmt/ 10	Tagged Image File Format	6		Image (Raster)	výstupní formát				
fmt/ 727	Cartesian Perceptual Compression image format		cpc, cpi	Image (Raster)	§ 23 odst. 3				
fmt/ 729	SQLite Database File Format	3	db, db3, sqlite, sqlite3	Database	§ 23 odst. 6	individuální posouzení archivem			Pod tímto PUID je identifikován formát GeoPackage. V současné době nemá v PRONOM vlastní PUID
fmt/ 91	Scalable Vector Graphics	1.0	svg	Image (Vector)	§ 23 odst. 2			ano	
fmt/ 92	Scalable Vector Graphics	1.1	svg	Image (Vector)	§ 23 odst. 2			ano	

x- fmt/ 392	JP2 (JPEG 2000 part 1)		jp2	Image (Raster)	§ 23 odst. 3				
fmt/ 916	ESRI ArcMap Document		mxd, mxt	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	
fmt/ 988	ESRI ArcScene Document		sxd	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	
fmt/ 989	ESRI ArcGlobe Document		3dd	GIS	§ 23 odst. 2, 3, 4	individuální posouzení archivem	GML	ano	
fmt/ 990	ESRI File Geodatabase			Database	§ 23 odst. 6	individuální posouzení archivem			
fmt/ 995	SIARD (Software- Independent Archiving of Relational Databases)	2.0	siard	Database		individuální posouzení archivem			
Fmt/ 887	SafeGuard Encrypted Virtual Disk		hdr, vol	Image (Raster)	§ 23 odst. 3				
Fmt /1135	SQLite Database File Format	2	db, sqlite	Database	§ 23 odst. 6	individuální posouzení archivem			
Fmt /1139	VectorWorks	9	vwx	Image (Vector)	§ 23 odst. 2			ano	
fmt/ 1140	VectorWorks	10	vwx	Image (Vector)	§ 23 odst. 2			ano	
fmt/ 1141	VectorWorks	11	vwx	Image (Vector)	§ 23 odst. 2			ano	

fmt/ 1196	SIARD (Software- Independent Archiving of Relational Databases)	2.1	siard	Database	individuální posouzení archivem				
fmt/ 1252	Raw Flux Image		rfi	Image (Raster)	§ 23 odst. 3				
fmt/ 1366	ESRI Published Map Format		pmf	GIS	§ 23 odst. 2, 3	individuální posouzení archivem		ano	
fmt/ 1367	GeoJSON		geojson	GIS	§ 23 odst. 2, 3	individuální posouzení archivem	GML	ano	
x- fmt/ 227	Geography Markup Language		gml	Text(Mark- up)	ponechat				

Transformace digitálních prostorových dat pro účely trvalého uložení v digitálním archivu

Kolektiv autorů

Odpovědný redaktor: Martin Rechterik

Odborná revize:

Mgr. Martin Buchlovský, Ústřední archiv zeměměřictví a katastru – ČÚZK

Mgr. Petr Dušánek, Zeměměřický úřad

Ing. Mgr. Vladimír Zýka, Výzkumný ústav Silva Taroucy pro krajinu a okrasné zahradnictví, veřejná výzkumná instituce

Odborná revize ze strany hlavního řešitele projektu: Výzkumný ústavu meliorací a ochrany půdy v.v.i.

Mgr. Jiří Brázda

Bc. Miroslav Drozen

Ing. Martin Mistr, Ph.D.

Ing. Vladimír Papaj, Ph.D.

Sazba: Martin Rechterik

Grafický návrh obálky: Eva Tamchynová

Vydal Národní Archiv, Archivní 4/2257, Praha 4 149 00 v roce 2022

Vydání první.

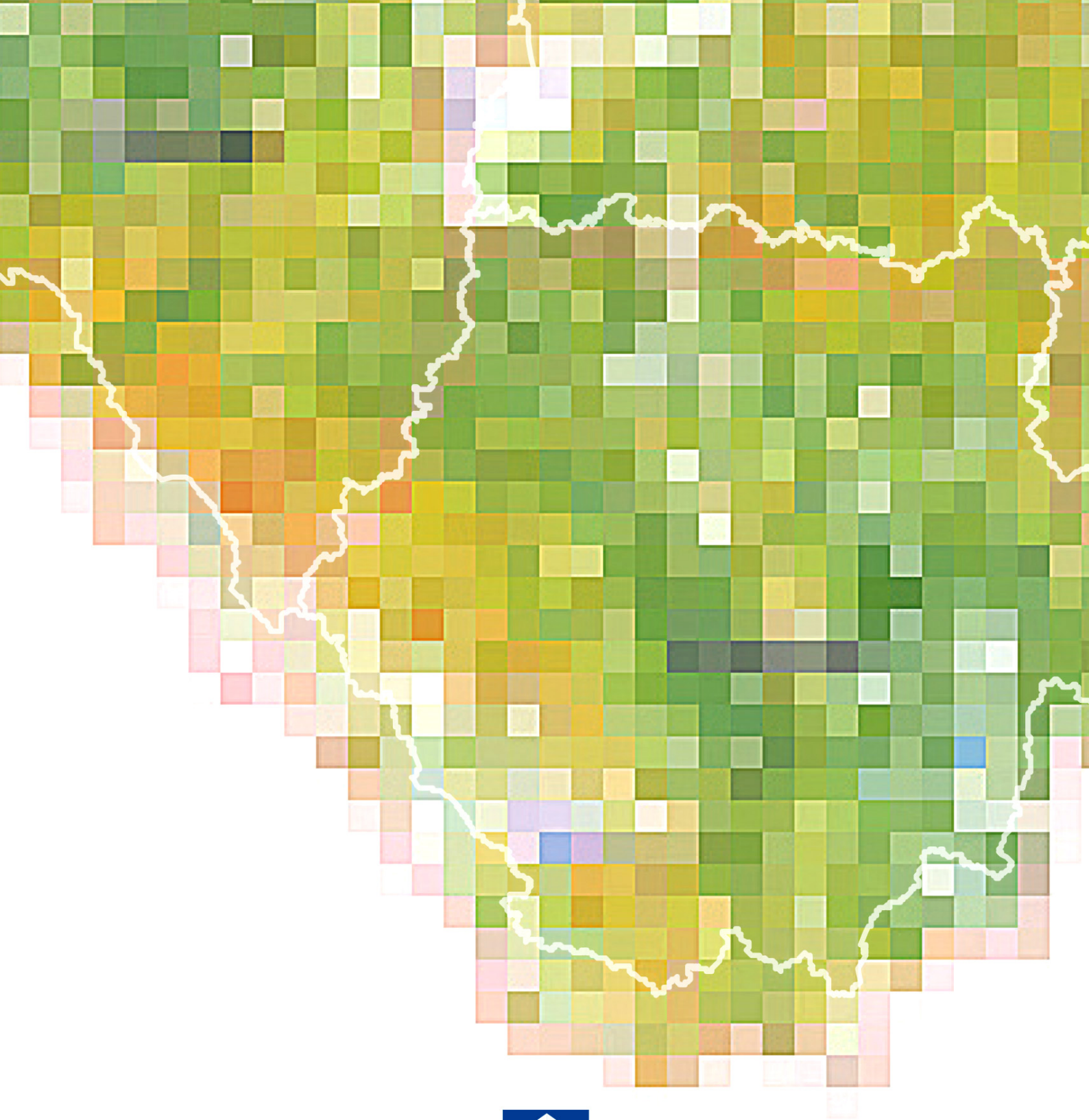
Verze 1.0

Publikace prošla jazykovou úpravou.

neprodejné

www.nacr.cz

ISBN 978-80-7469-122-5 (pdf)



Národní archiv

Název projektu: Geografické informační systémy a plnění povinností legislativy v oblasti archivnictví
č. projektu: 2020TL03000364

T **A** Tento projekt je financován se státní podporou
Č **R** Technologické agentury ČR v rámci Programu Éta



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY



**Výzkumný ústav meliorací
a ochrany půdy, v.v.i.**