

Specifikace obsahu díla

Zhotovitel se zavazuje provést pro objednatele za podmínek sjednaných touto smlouvou digitalizaci naskenovaných knihovních dokumentů podle aktuálních standardů a formátů používaných Národní digitální knihovnou, která spočívá ve zpracování dokumentů objednatele provedením těchto hlavních činností:

- Úprava obrazových souborů (ořez a narovnání).
- Konverze obrazových souborů pomocí OCR do textového formátu ALTO XML.
- Vytvoření popisných údajů – bibliografických, technických a administrativních metadat (standard MODS a DublinCore).
- Záznam komplexních dokumentů na výměnný harddisk (struktura PSP balíčků ve formátu METS).
- Dodání jedné kopie kompletního reformátovaného dokumentu pro archivaci.
- Dodání jedné kopie pro zpřístupnění v systému Kramerius.
- Přeprava dokumentů a zpracovaných dat.

Standards digitalizace:

<http://www.ndk.cz/standardy-digitalizace>

https://www.ndk.cz/standardy-digitalizace/dmf_monografie_1-3-2

Zhotovitel bude při plnění díla vycházet z dokumentu „Definice metadatových formátů pro digitalizaci monografických dokumentů, verze 1.3.2, ze dne 8. 10. 2018, který si obstará sám, např. stažením z výše uvedeného internetového odkazu. Pokud NDK provede během roku 2019 (do předání díla zhotovitelem) úpravy a doplnění předepsaných standardů, je zhotovitel povinen tyto změny akceptovat včetně analytické struktury.

Parametry obrazových souborů:

<http://www.ndk.cz/standardy-digitalizace/standardy-pro-obrazova-data>

Specifikace definované na webové adrese <http://www.ndk.cz/standardy-digitalizace>. Obrazové soubory pro archivní účely musí splňovat veškeré parametry uvedené ve sloupci Master Copy. Obrazové soubory pro uživatelské účely pak musí splňovat veškeré parametry uvedené ve sloupci Production Master Copy. Všechny tyto podmínky jsou pro předkladatele projektu závazné. Obrazové soubory v obou formátech (ve všech stanovených parametrech) lze generovat prostřednictvím nástroje Kakadu (<http://www.kakadusoftware.com/>), který je NK ČR doporučen jako optimální.

Úprava obrazových souborů:

- ořez cca 1 mm vně okraje dokumentu, narovnání podle řádků textu.

Veškeré úpravy obrazů se budou provádět na archivních souborech, uživatelský soubor se bude generovat po všech úpravách. Uživatelská i archivní kopie musí mít stejný rozměr (v pixelech) a stejné rozlišení (v DPI).

Konverze obrazových souborů pomocí OCR

do textového formátu s úspěšností rozpoznávání min. 95%. V rámci konverze do textového formátu je požadováno vytvoření souborů ve formátu ALTO XML podle specifikace uvedené na webové adrese <http://www.ndk.cz/standardy-digitalizace/standardy-pro-obrazova-data>.

Soubor ALTOxml poskytující informace o souřadnicích textových informací vhodných pro vyhledávání bude generován z uživatelské kopie. Po té již nesmí být obrazové soubory upravovány. V případě nutnosti, musí být znovu po úpravě znovu vytvořen soubor ALTO xml.

Vytvoření popisných údajů

– bibliografických metadat a technických a administrativních metadat (bibliografického záznamu a struktury dokumentu) ve standardu MODS a DublinCore musí plně respektovat veškerá ujednání, která jsou definovaná na webové adrese:

https://www.ndk.cz/standardy-digitalizace/dmf_monografie_1-3-2

Digitální dokument musí obsahovat metadata pro všechny části (strany) dokumentu. Indexní údaje musí odkazovat na správné části dokumentu (např. strana 4 v metadatech musí odpovídat straně 4 v dokumentu).

Nezbytnou součástí metadat tvoří údaje o chybějících částech dokumentu (strany, čísla, přílohy apod.), nesprávné údaje o stránkování, o dataci a číslování dílčích částí (denních, týdenních, měsíčních výtisků, ročníků apod.).

Záznam komplexních dokumentů:

Zhotovitel zajistí tvorbu dokumentů ve struktuře PSP balíčků ve formátu METS podle specifikace dostupné na webové adrese

https://www.ndk.cz/standardy-digitalizace/dmf_monografie_1-3-2

Struktura digitálního dokumentu musí respektovat základní strukturu PSP balíčku definovanou tamtéž, přičemž jeho základními prvky musí být:

1. archivní kopie (1 MC pro každou stránku),
2. uživatelské kopie (1 UC pro každou vzniklou MC, tedy stránku),
3. OCR - ALTO XML soubor pro každou stránku,
4. OCR TXT soubor - pro možnost stáhnout si jen text dokumentu (tam, kde kvalita OCR je odpovídající), vyhledávání/indexace,
5. metadata pro MC,
 - a. bibliografická metadata – MODS a DC,
 - b. strukturální metadata – METS,
 - c. technická metadata – MIX, PREMIS,
 - d. administrativní metadata – PREMIS, METS,



6. kontrolní metadatové soubory (s kontrolními součty a údaji o vzniku dat apod.).

Požadavky na kvalitativní parametry:

Čitelnost digitálního dokumentu musí odpovídat čitelnosti originální předlohy (knihy nebo mikrofilmu).

Členění dokumentu na fyzických médiích (v případě potřeby) musí respektovat přirozené členění původního dokumentu podle časových úseků (rok, pololetí, čtvrtletí, měsíc) nebo podle čísel (např. 1 – 10, 11 – 20, 21-30).

Ukládání souborů:

Archivní obrazové soubory budou ukládány ve formátu JPEG 2000 s bitově bezztrátovou kompresí, uživatelské obrazové soubory ve formátu JPEG 2000 s vizuálně bezztrátovou kompresí. Pro textové soubory bude použit formát ALTO XML.

Konzultace:

Konzultace poskytne Odbor správy a digitalizace fondů NK ČR: Mgr. Tomáš Foltýn (tomas.foltyn@nkp.cz).

Konzultace v otázce strukturování digitálních dokumentů (tvorby metadat) lze konzultovat s pracovištěm NK ČR, oddělení LTP úložiště (Zuzana.Kvasova@nkp.cz).



Číslo úkolu ÚZEI 4414

PŘEDÁVACÍ – AKCEPTAČNÍ PROTOKOL

ZHOTOVITEL:	OBJEDNATEL:
sídlo: IČO: zástupce zhotovitele pověřený k předání:	Ústav zemědělské ekonomiky a informací sídlo: Mánesova 1453/75, 120 00 Praha 2 IČO: 00027251 zástupce objednatele pověřený k převzetí:

Číslo smlouvy objednatele	Datum uzavření smlouvy
/2019	

Předmětem akceptačního řízení je ověření plnění díla digitalizace skenů

Datum předání:

Závěr akceptačního řízení: (variantu závěru označte křížkem):

- | | |
|---|--|
| <input type="checkbox"/> akceptováno bez výhrad | <input type="checkbox"/> neakceptováno, k přepracování |
| <input type="checkbox"/> akceptováno částečně, s výhradami | <input type="checkbox"/> neakceptováno s následným odstoupením od smlouvy |

Seznam nedostatků/vad:

Pořadové číslo	Popis nedostatku/vady	Poznámka
1	---	
2	---	
3	---	

Termín pro odstranění nedostatků/vad, v případě akceptace se závěrem „akceptováno částečně s výhradami“ či „neakceptováno k přepracování“: do ---.

Tento protokol byl vyhotoven ve třech stejnopisech, z nichž objednatel obdrží dva a zhotovitel jeden.

Akceptační řízení skončeno dne .

podpis zástupce zhotovitele

podpis zástupce objednatele