

**Manuál pro databázové zpracování knižní ilustrace
v historických fondech paměťových institucí
(parametry digitalizace)**

e-ilustrace.cz

Praha 2021

Vysvětlivky zkratk:

KPS - Knihopis českých a slovenských tisků do roku 1800

BCBT - Bibliografie cizojazyčných bohemikálních tisků do roku 1800

NK ČR - Národní knihovna ČR

KNAV - Knihovna Akademie věd ČR

NDK - Národní digitální knihovna

MARC21 - Bibliografický formát MARC21 (Machine Readable Cataloging)

RDA - Katalogizační pravidla RDA (Resource Description and Access)

IIIF - Standard pro sdílení obrazových dat (International Image Interoperability Framework)

Obsah

Úvod	3
Zpracování a prezentace obrazových dat v bibliografických databázích	4
Prezentace obrazových dat v prostředí e-ilustrace	5
Parametry digitalizace	6
Standardy mobilní digitalizace	7

Úvod

Manuál byl vytvořen v rámci projektu TAČR TL05000157 *Česká knižní ilustrace v raném novověku* (Technologická agentura ČR – Program na podporu aplikovaného společenskovedního a humanitního výzkumu, experimentálního vývoje a inovací ÉTA). Informace o projektu naleznete na stránkách www.e-ilustrace.cz.

Základem *Manuálu (parametry digitalizace)* je vymezení souboru obrazových dokumentů digitalizovaných a zpracovávaných pro prezentaci knižní ilustrace v databázích národní retrospektivní bibliografie tisků do roku 1800 ([KPS](#) a [BCBT](#)) a nově vznikajícím prostředí e-ilustrace. Dále zahrnuje stručný přehled standardů digitalizace v Národní knihovně ČR, stanovuje parametry pro mobilní digitalizaci v institucích, které nedisponují vlastním digitalizačním centrem, a poskytuje tak základní, obecné vodítko pro pořízení a zpracování obrazového materiálu ilustrací v historických knižních fondech a sbírkách.

Řešitelský tým projektu TAČR TL05000157
Kontaktní formulář <https://e-ilustrace.cz/>

Zpracování a prezentace obrazových dat v bibliografických databázích

Národní retrospektivní bibliografickou databází tisků vydaných do roku 1800 tvoří dvě samostatně budované databáze (KPS a BCBT), jejichž tvorbu, údržbu a rozvoj zajišťují v České republice dvě pracoviště. Databáze KPS je vytvářena v Oddělení rukopisů a starých tisků NK ČR v automatizovaném katalogizačním systému Aleph. Databázi BCBT vytváří a spravuje Knihovnědné oddělení KNAV v systému Verbis.

Struktura bibliografických záznamů podléhá [pravidlům pro katalogizaci starých tisků ve formátu MARC21/RDA](#). Obě databáze jsou on-line přístupné jak samostatně, tak v rámci společného rozhraní v portálu www.knihoveda.cz pomocí vyhledávacího nástroje VuFind.

V dosavadním zpracování národní retrospektivní bibliografie bylo samostatné zpracování obrazových dat spíše výjimečné, normou se však stala praxe odkazování na digitalizáty exemplářů evidovaných vydání přístupné v domácích i zahraničních digitálních knihovnách. Tento způsob on-line zpřístupnění prvotisků a starých tisků však není systematizovaný a neumožňuje kompletaci obrazových dat ilustrací pro vybraná časová období ani samostatné vyhledávání a další vědeckou práci s nimi.

Základem projektu *Česká knižní ilustrace v raném novověku* je proto kompletace ilustračního materiálu v bohemikálních tiscích v letech 1475–1550, v podobě digitalizace jednotlivých ilustrací a jejich následného bibliografického zpracování a zpřístupnění v bibliografických databázích KPS a BCBT.

Pro tento účel byl navržen *Manuál pro popis knižní ilustrace a Pravidla pro bibliografické zpracování obrazových dokumentů ve formátu MARC21/RDA*. Výsledkem budou bibliografická metadata k jednotlivým obrazovým dokumentům, která umožňují sdílení dat ve formátu XML pomocí protokolu OAI-PMH. Obrazová data jsou pak uložena ve vnitřních digitálních uložistiích katalogizačních systémů ve formátu JPEG, což zajišťuje jejich zveřejnění formou náhledů (celé strana s ilustrací a výřez ilustrace) v rámci samotných bibliografických databází a zároveň sdílení ve vyhledávacím rozhraní VuFind a převzetí obrazových dat, metadat a bibliografických metadat do nově vyvíjené aplikace e-ilustrace.

Prezentace obrazových dat v prostředí e-ilustrace

Katalogizační systémy obou bibliografických databází slouží především k tvorbě bibliografických metadat a jejich provázání s obrazovými daty. Jejich prezentační možnosti jsou však omezené a neumožňují další práci s obrazovým materiálem. Vyvíjené prostředí e-ilustrace díky využitým nástrojům naopak poskytne zejména širší variabilitu náhledů a úprav zveřejněných obrazových dokumentů, včetně tvorby vlastních kolekcí v prezentačním a vyhledávacím prostředí Mirador a samostatné rozhraní pro vědeckou práci s digitalizovanými ilustracemi v software VISE.

Vzhledem k tomu, že jsou digitalizáty zpřístupňovány na více úrovních, vyvstává zde potřeba dvojího formátování obrazových dat. Zatímco pro zveřejnění obrazových dokumentů v rámci bibliografických databází je podporován formát JPEG, pro využití v prostředí Mirador je nutný formát JPEG2000, který umožňuje sdílení podle stanovených standardů IIIF.

Z uvedeného vyplývá, že užití formátů JPEG a JPEG2000 zajišťuje jednak potřebnou variabilitu pro různá prezentační prostředí, ale zároveň je možné druhý zmiňovaný využít také pro archivaci obrazových dokumentů, byť tento požadavek není povinnou součástí řešeného projektu.

Parametry digitalizace

Parametry digitalizace dodržované v digitalizačním centru NK ČR kombinují doporučení pro data vzniklá v podprogramu VISK 7, [standardy NDK](#) a reflektují doporučení pro kódování souborů JP2 v rámci zahraničních projektů zabývajících se digitalizací a dlouhodobým uchováváním obrazových dat. Pro skenování je vyžadováno rozlišení minimálně 300 PPI. Barevná hloubka pro tyto dokumenty je fixně nastavena na 24 bitů, barva (RGB).

Standardy NDK předpokládají využívání grafického formátu JPEG2000 jak pro archivní kopie, tak pro kopie určené pro zpřístupňování. Vytváření obrazových dokumentů ve formátu JPEG2000 zajišťuje v NK ČR software Kakadu 7. Pro generování uživatelských kopií v grafickém formátu JPEG je zvolena ztrátová až vizuálně mírně ztrátová komprese. Kromě uživatelské kvality (Photoshop hodnota 3, RGB, JPG, 300 DPI) poskytuje digitalizační centrum digitalizáty také v tiskové kvalitě (Photoshop hodnota 11-12, sRGB, JPG (možno TIFF), 300 DPI nebo 400 DPI).

U výřezu obrazových dokumentů je doporučeno dodržovat cca 1 mm vně okraje dokumentu. Veškeré úpravy obrazů je třeba provádět na archivních souborech, uživatelský soubor se bude generovat až po všech úpravách. Uživatelská i archivní kopie musí mít stejný rozměr (v pixelech) a stejné rozlišení (v DPI).

Pro potřeby prezentace obrazových dat v bibliografických databázích NK ČR a KNAV, potažmo v rozhraní VuFind, vytváříme vždy dva obrazové dokumenty, na které odkazujeme v bibliografickém záznamu ilustrace. Odlišujeme je pomocí identifikátoru ilustrace a slovního vyjádření v příslušném podpoli bibliografického záznamu. Identifikátor, který vychází z čísla KPS či BCBT, obsahuje informaci o pořadí ilustrace v evidovaném vydání (př. K04275_IL015) a pomocí číselného dodatku odlišuje výřez ilustrace (př. K04275_IL015_1.jpg) od celé stránky dokumentu, na níž je ilustrace umístěna (př. K04275_IL015_2.jpg).

Zásadní pomůckou pro vědeckou práci s ilustrací je měřítko, které musí být viditelnou součástí obrazového dokumentu. A to přesto, že je rozměr ilustrace povinným údajem v bibliografických metadatech. V případě obrazového dokumentu pořízeného profesionálním skenerem je tento údaj součástí metadat a je potřeba ho dodatečně vizualizovat pomocí software.

Standardy mobilní digitalizace

S potřebou využití možnosti mobilní digitalizace se často setkáme u projektů, které předpokládají [zapojení](#) širší veřejnosti do procesu kompletace obrazových dat nebo (což je případ projektu *Česká knižní ilustrace v raném novověku*) jsou často unikátně dochované originály obrazových dokumentů [uchovávané](#) v institucích či soukromých sbírkách, které nedisponují možností vlastní digitalizace.

Pro tyto případy je nutné stanovit alespoň základní parametry digitalizačních postupů tak, aby byla zajištěna požadovaná kvalita vytvářených obrazových dat, a to zejména s ohledem na jejich další zpracování a zamýšlené využití. Důležitou roli zde hraje také povaha či typ snímaného objektu, který nezdíka vyžaduje silně individuální přístup. V případě mobilní digitalizace obrazových dokumentů v tiscích vydaných do roku 1800 je samozřejmě nutné využívat neinvazivní a nedestruktivní postupy.

Z hlediska volby vhodného mobilního zařízení, kdy v úvahu připadají fotoaparát a mobilní telefon či tablet vybavený fotoaparátem, hrají klíčovou roli vedle samotné přenosnosti především dva faktory. Prvním je rozlišení fotoaparátu uváděné obvykle v Megapixelech (Mpx), druhým pak formát a velikost získaných obrazových dat.

Zatímco u fotoaparátů a mobilních telefonů je dnes na trhu řada modelů, jejichž rozlišení se pohybuje na škále zhruba od 3-5 Mpx až po téměř 200 Mpx, u tabletů se nejvyšší rozlišení pohybuje kolem 13 Mpx. Pro účely fotografování obrazových dokumentů prezentovaných ve webovém prostředí je [dostačující](#) minimální rozlišení 10 Mpx, které v případě potřeby umožní i kvalitní tisk formátu A3. Výrazně vyšší rozlišení klade nároky na kapacitu datového úložiště a výkon počítače při zpracování.

Obvyklými formáty nasnímaných obrazových dat v případě fotoaparátu jsou RAW, jenž bezpodmínečně vyžaduje následnou datovou konverzi, či JPEG, u mobilních telefonů a tabletů je ukládání dokumentů nastaveno většinou výrobcem ve formátu JPEG a zároveň v rozlišení 72-96 DPI, jenž je dostatečné pro prezentaci obrazových dat v počítačovém prostředí, včetně dostačujícího přiblížení detailu.

Na rozdíl od stacionárních skenerů hraje v případě mobilní digitalizace důležitou roli také nakládání se snímaným dokumentem i mobilním zařízením a pokud je to možné, pak je třeba zajistit co nejstálější světelné podmínky. Pro digitalizaci mobilními telefony jsou k dispozici i poměrně nenákladná technická řešení typu digitalizačního [stanu](#),

kteře zaručí i neměnné světelné podmínky umělým osvětlením. Při běžném využití fotografických přístrojů je vhodné využít denní světlo a kvalitní stativ s příslušnými redukcemi, který zajistí dostatečnou stabilitu fotoaparátu i konstantní, ideálně paralelní polohu vůči snímanému dokumentu tak, aby při digitalizaci nedošlo k nežádoucímu zkreslení. Ze stejného důvodu je vhodné omezit využití pokročilých funkcí fotoaparátů v mobilních zařizenech, například vysoký dynamický rozsah (HDR) a dokumenty umístit na unifikovaný podklad.

Jak již bylo zmíněno výše, pro vědeckou práci s ilustrací musí být viditelnou součástí obrazového dokumentu měřítko, které je nutné přiložit ke snímanému dokumentu. Pro usnadnění orientace při následném zpracování digitalizátů je vhodné přiložit také štítek s identifikátorem ilustrace. Barevná kalibrační tabulka naopak není pro účely digitalizace tohoto typu nezbytná.