

# Formáty uložení dat

- Používat pokud možno otevřené formáty s dobrou podporou
- Podstatná je možnost migrace na nový formát
- Ukládat “Master soubory” s pokud možno nejúplnější informací o procesu digitalizace
- Pro běžnou manipulaci je vhodné používat menší soubory kde není problém s velikostí uložení a dobou zpracování

# Nejběžnější formáty

- Grafické formáty
  - TIFF, JPEG, JPEG2000, PNG
  - + snadno zpracovatelné a mezi sebou převoditelné
  - bez informace o procesu digitalizace
- RAW Soubory
  - RAW soubory výrobců, DNG, RAW TIFF
  - + Obsahují nejdetailnější informace o digitalizaci a lze z nich kdykoliv vytvářet grafické formáty
  - Jejich zpracování vyžaduje znalosti o daném zařízení a jejich morální životnost je tak nižší. Navíc často není možné dávkové zpracování.

# TIFF

- Nejlépe podporovaný grafický formát v oblasti DTP vůbec
- Často nativní formát programů pro digitalizaci
- Kontrolovaný firmou Adobe Systems
- Stabilní, poslední modifikace jsou z roku 1992. Je to ale starý formát, složitý na implementaci a má mnoho omezení. (například podporu UNICODE)
- Nepodporuje soubory nad 2GB, varianta BigTIFF
- Patent algoritmu LZW již vypršel
- Nejčastěji používaný formát pro master soubory, často bez další komprese.

# PNG (Portable Network Graphics)

- Otevřený formát, reakce na patentovou politiku algoritmu LZW v roce 1995
- Verze 1.0 je z roku 1996, verze 1.2 standardizována jako ISO/IEC 15948:2004
- Nabízí kompresi algoritmem Deflate (ZIP), která je ale dostupná i v některých implementacích TIFF
- Je výrazně méně podporován než TIFF.

# JPEG

- Klasický formát podporující ztrátovou kompresi cca 10:1
- ISO 10918-1:1994
- Základním problémem je chybící podpora 16ti bitové hloubky
- Kompresní algoritmus z roku 1986 byl již překonán. Hlavním problémem jsou čtvercové artefakty na hranicích makrobloků
- Pro svou širokou podporu je JPEG vhodný zejména pro náhledové soubory. Použití pro master soubory komplikuje ztrátovost komprese a bitová hloubka

# JPEG2000

- Zamýšlen jako následovník populárního formátu JPEG
- ISO/IEC 15444-1:2000
- Podpora větších bitových hloubek
- Kvalitní komprese pomocí waveletů bez blokových artefaktů
- Soubor je uložen po hladinách postupně se zvětšujícího rozlišení. Nahrávání je tedy velmi rychlé
- Volitelně i kvalitní bezztrátová komprese (úspora 30% oproti PNG/TIFF). Tim Vitale a další ji často kritizují pro ztrátovost, není to ale pravda.
- Metadata ve formátu EXIF
- Odolnost vůči poškození dat



Uncompressed  
378 KiB  
1:1



JPEG JFIF  
11.2 KiB  
1:33.65  
IJG q 30



JPEG 2000  
11.2 KiB  
1:33.65

# JPEG2000

- Problémy jsou zejména s kvalitou informace
  - Není nativně podporován většinou programů pro digitalizaci: není to ale problém, kompresi náročnou na výpočetní čas je vhodné provádět dávkově přes noc.
  - Není nativně podporován v Adobe Photoshopu, existuje ale několik pluginů včetně j2k, který je zdarma
  - Volně šiřitelná implementace Japster
  - Několik komerčních velmi kvalitních a výkonných kodeků
- Využití JPEG2000 vyžaduje opatrnost a studium
  - Používáno v mikroskopii, výzkumu vesmíru (mars rover), zpracování filmu, lékařských datech a dalších oblastech kde možnost dat a kvalita komprese je kritická.
- Riziko patentovaných algoritmů



# RAW Soubory

- Podstatné zejména u one-shot systémů, kde každý pixel obsahuje informaci jen o jedné barevné složce. RGB data se interpolují.
- Uložení dat se snímačů v původní podobě včetně informace o nastavení zařízení (expozici, citlivosti atd).
- Tyto informace lze použít při převodu do grafického formátu (například citlivost pro redukci šumu)
- Využití RAW souborů pro archivaci je ale omezené, protože jejich morální životnost končí s prodejem daného zařízení.
- Soubory bývají závislé na výrobci

# DNG (Digital NeGative)

- Standarizovaný RAW formát, snaží se řešit závislost na výrobci.
- ISO standard se připravuje, není ale plně otevřeným formátem, některé informace jsou závislé na výrobci
- Kvalitní software pro zpracování DNG (Adobe Lightroom, Aperture, ACDSee, Vuescan atd).
- Obsahuje původní data se snímačů, profil kamery a další charakteristiky hardware
- Formát je relativně nový (z roku 2004), jeho hodnotu pro archivaci bude třeba ještě zhodnotit.
- Implementace jsou ne vždy stejně dobré jako implementace RAW

# Příklad uložení v našem projektu

- Master soubory:
  - Uložené ve formátu RAW TIFF VuSCANu převedén mezztrátově do JPEG2000 a metadat. Převod šetří 40% prostoru a je možné jej vrátit zpět.
- Náhledové soubory:
  - Ve formátu JPEG o šířce 1200 bodů automaticky generované z RAW TIFFů. Pro internet zmenšeny na rozlišení 640 bodů po delší straně.
- Odvozené soubory (příprava pro tisk apod.)
  - Uložené ztrátově ve formátu JPEG2000