

# LTP Mining



# LTP – Long Term Preservation

- Digital preservation je metoda uchovávání digitálních záznamů tak, aby obsah byl v budoucnu čitelný i když zastarává původní hardware a software.
- Hardware – diskety 5,25", 3,5", DAT, CD ...
- Software – SAM, WP, FLC, PIC, XM, SWF, ....



# Metadata ... ?



# Metadata

- Metadata poskytují informace o datech.
- Deskriptivní metadata: Titul, Autor, Popis, ...
- Technická metadata: Formát souboru, velikost souboru, rozlišení, bitrate ...
- Administrativní metadata: Jaká aplikace soubor vytvořila, kdo upravil metadata.
- Strukturální metadata: Zapouzdření všeho dohromady, struktura stránek, ...

# Standardy pro metadata

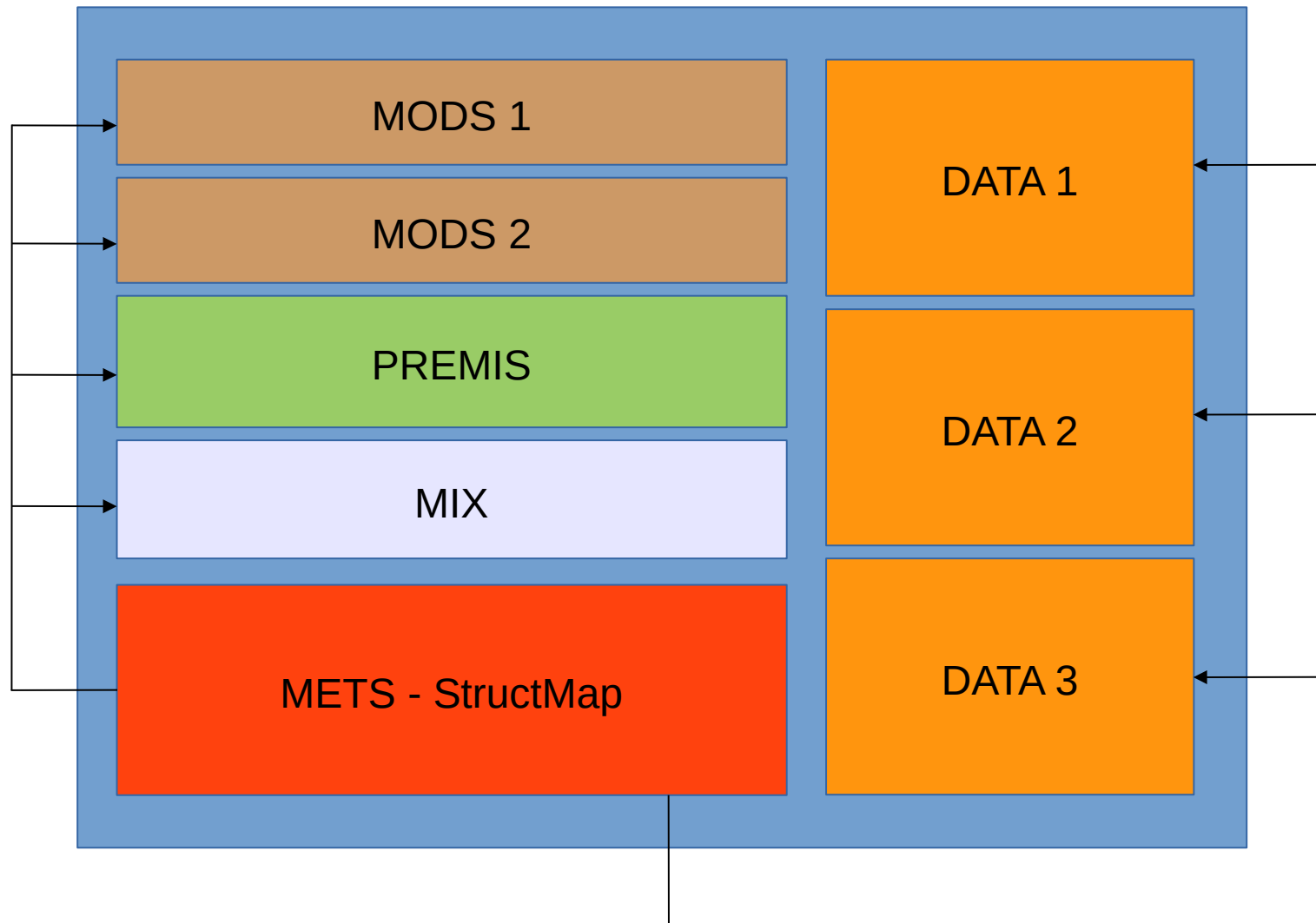
- MODS, Dublin Core, NISO MIX, PREMIS, METS, DDI, EBUCore, EBU CCDM, FOAF, EAD, CDWA, VRA Core, Darwin Core, ONIX, CWM, EML, IEE LOM, CSDGM, ISO 19115, e-GMS, GILS, TEI (P5), <indecs>, MARC, XOBIS, PBCore, MEI, DOI, ISO/EIC 19506, ISO 23081, MoReg2010, DIF, RDF, MDDL, NIEM, SAML, ...

# Nový standard ?

HOW STANDARDS PROLIFERATE:  
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)



# AIP – Archival Information Package



# AIP

- Může existovat více verzí AIP v jednom archívu
- Každý typ entity může mít jinou strukturu AIP (monografie, periodika, e-knihy, hudebniny, mapy, softwarové hry, ...). Různé standardy metadat a i metadata bez standardu.
- Každý producent dat může mít různé požadavky, které mohou vynutit více verzí AIP.



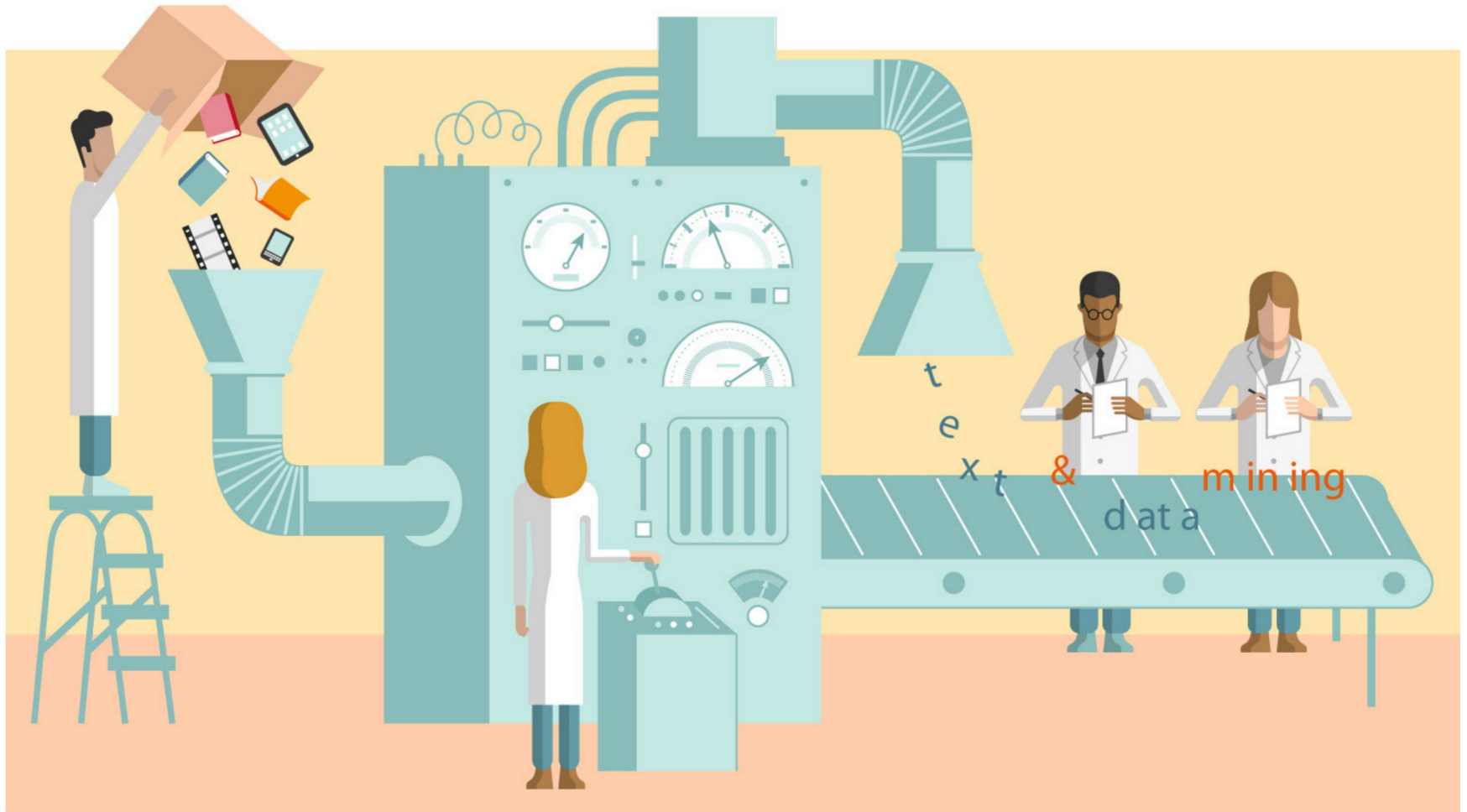
Mnoho různých AIP....



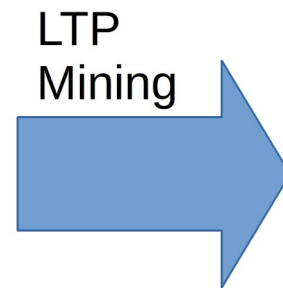
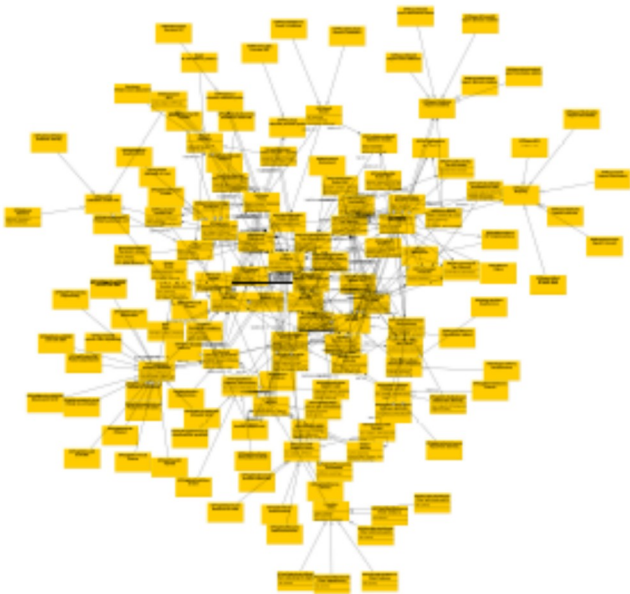
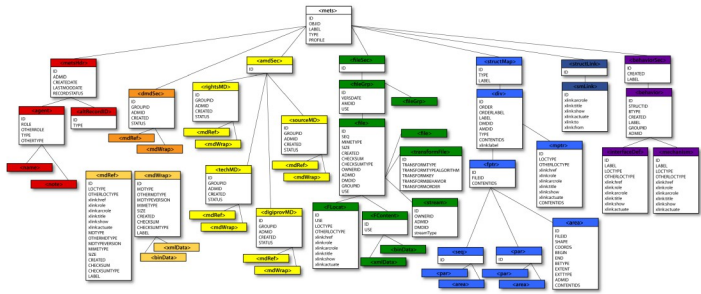
# Úkoly pro LTP archiv

- Jaké AIP byly naskenovány vadným skenerem SC01 v období od října 2015 ?
- Kolik je celkem AIP, které je možné v roce 2016 zveřejnit díky změně autorských práv ?
- Jaká je celková velikost datových souborů, které jsou uloženy jako TIFF bez komprese ?
- U kolika souborů byla provedeno OCR se softwarem Tesseract ve verzi 3.04.01 ?

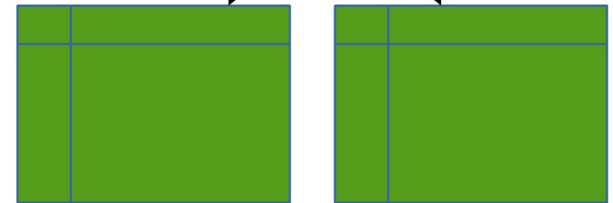
# LTP Mining



# LTP Mining

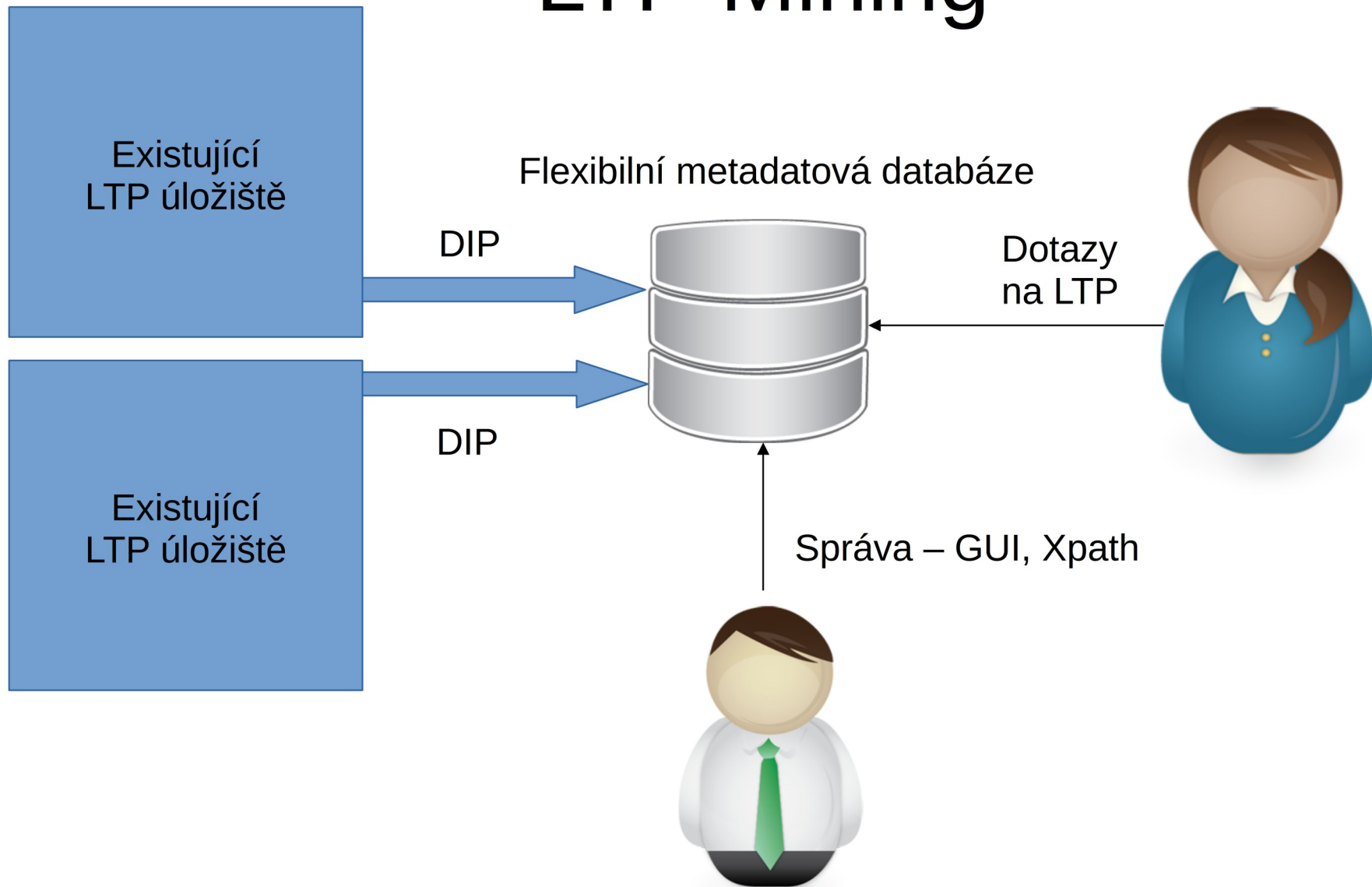


Dotazy



Dynamické tabulky

# LTP Mining



# SAFE LTP Mining

- V souladu s OAIS.
- Založeno na standardech – METS, MODS, PREMIS, MIX, DC ... Přijímá jakékoliv XML.
- Poskytuje podporu pro certifikaci LTP archívu.
- Otevřené API pro software třetích stran. Beží na Linux, Mac OS X, Windows. Možnost běhu v cloudu.
- Standardy: CMIS Query, XPath Query, Jasper reports.
- Optimalizováno na rychlost a velké objemy metadat.

# LTP Mining



Použité obrázky v prezentaci (děkuji):

<http://xkcd.com>

<http://miriadna.com>

<http://copyrightuser.org>

<https://en.wikipedia.org>

<http://ltpmining.com>

[roman.bouchner@aiptsafe.cz](mailto:roman.bouchner@aiptsafe.cz)