

The background is a dark blue gradient with a starry texture. On the left side, there are several overlapping circular elements. One prominent feature is a large circular scale with tick marks and numbers ranging from 140 to 260. Other circles contain curved lines and arrows, suggesting a sense of motion or data flow.

# ISKB49 - DIGITÁLNÍ KURÁTORSTVÍ VÝZNAMNÉ VLASTNOSTI DIGITÁLNÍCH OBJEKTŮ

MICHAL KONEČNÝ

KABINET INFORMAČNÍCH STUDIÍ A KNIHOVNICTVÍ

6. BŘEZNA 2019

# CO NÁS ČEKÁ?

1. Úvod (pro ty, co tu nebyli).
2. Opakování (pro ty, co tu byli).
3. Vyhodnocení úkolů.
4. Povíme si, jak model OAIS souvisí s významnými vlastnostmi
5. Představíme si framework InSPECT
6. Zkusíme zjistit, jak poznáme, co je významné
7. A jak to zapíšeme pomocí metadat
8. Workshop!



# DIGITÁLNÍ KURÁTORSTVÍ

V širším pojetí se jedná o údržbu, zhodnocování a důvěryhodnost digitálních informací pro současné a budoucí použití.

*David Giaretta*

# DLOUHODOBÉ UCHOVÁVÁNÍ

- *Digital preservation, long term preservation...*
- Pořizování, uchovávání a zpřístupnění dat takovým způsobem, aby byla zajištěna jejich použitelnost, dostupnost a autenticita



# OPEN ARCHIVAL INFORMATION SYSTEM

- Otevřený Archivační Informační Systém
- Terminologický a konceptuální rámec pro:
  - popis a porovnávání architektury,
  - volbu strategií LTP,
  - porovnání datových modelů digitálních archivů,
  - odhadování změny v průběhu času.



# VÝZNAMNÉ VLASTNOSTI DIGITÁLNÍCH OBJEKTŮ

Vlastnosti digitálních objektů, které musí být zachovány v průběhu času, aby byla trvale zajištěna jejich přístupnost, použitelnost a účel.

*Andrew Wilson*

*Je potřeba, aby u obrázku byla převedena i barva? Je nutné u textu převést do nového formátů i rozdělení na odstavce nebo verše? Neztratí se smysl díla, když tyto vlastnosti nebudou převedeny?*



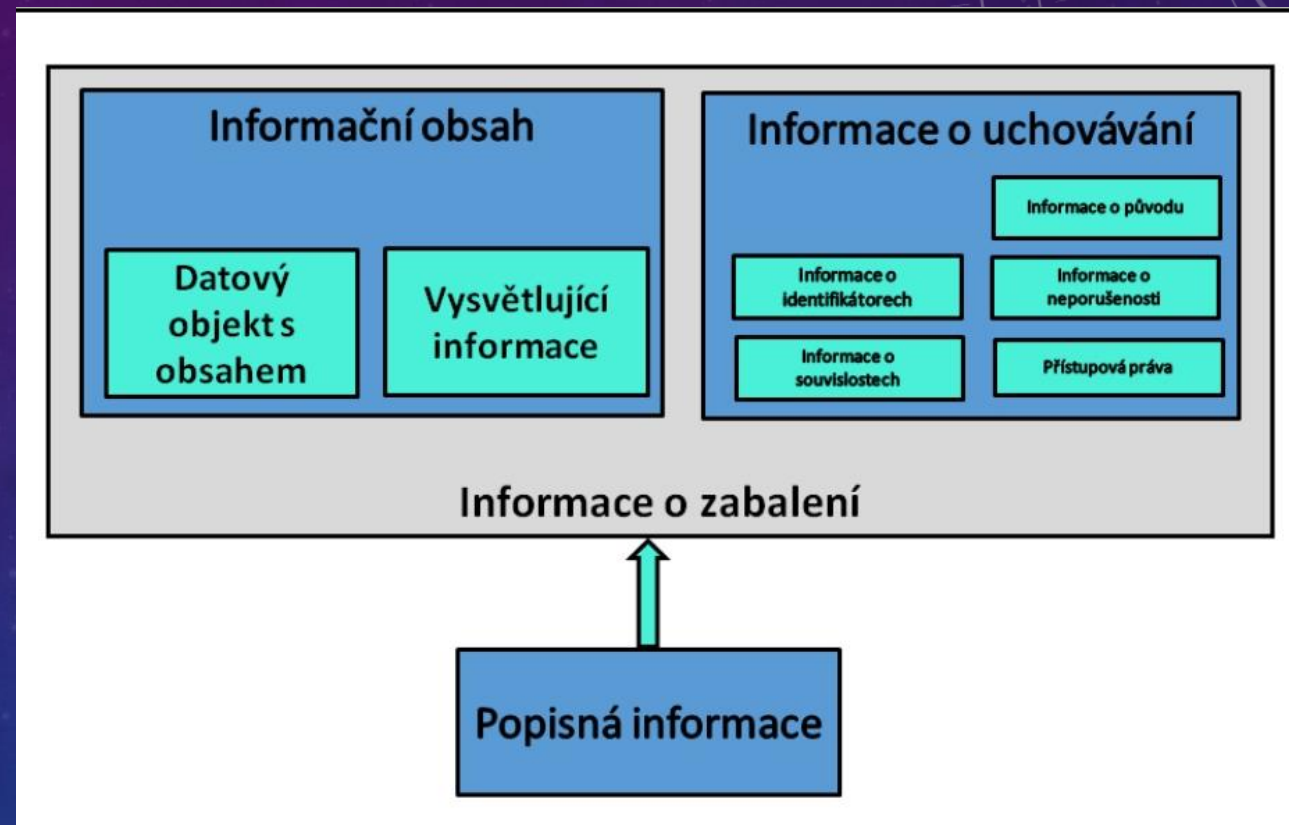
# URČENÁ SKUPINA

Stanovená skupina možných koncových uživatelů, kteří by měli být schopni porozumět konkrétní množině informací



# ARCHIVNÍ INFORMAČNÍ BALÍČEK

- Objekt, který je předmětem ochrany
- Informace nezbytné pro jeho ochranu, zpřístupnění a zajištění srozumitelnosti
- “Obal” který vše drží pohromadě
- Popisné informace abychom ho poznali a našli





# INSPECT PROJECT

- Projekt realizovaný v letech 2007 - 2009
- Investigating the Significant Properties of Electronic Content Over Time
- Vyvrcholení snah o formalizaci postupů pro určování významných vlastností
- Vychází z frameworku FBS (Function-Behaviour-Structure)
- Používá analýzu objektů a stakeholderů

# FUNCTION-BEHAVIOUR-STRUCTURE

- **FUNKCE:**  
Zamýšlený záměr, účel nebo smysl, který lze provést (funkcí textu je přenos informace).
- **CHOVÁNÍ:**  
Výstup založený na funkci a struktuře (například interpretace obsahové informace).
- **STRUKTURA:**  
Stavební prvky Objektu, které umožňují stakeholderům dosáhnout požadovaného chování.

# ANALÝZA OBJEKTŮ

Hodnotitel analyzuje reprezentativní vzorek objektů, identifikuje funkce a chování, které lze aplikovat, a vlastnosti nezbytné pro jejich provedení.

1. Výběr typu objektu pro analýzu
2. Analýza struktury
3. Identifikace účelu technických vlastností
4. Určení očekávaného chování
5. Klasifikace chování do funkcí
6. Propojení chování a struktury
7. Revize a dokončení

# 1. VÝBĚR TYPU OBJEKTU PRO ANALÝZU

Hodnotitel může zvolit obecný typ objektu (bitmapu, audio nahrávku, webovou stránku) nebo konkrétnější podtyp, který má specifické charakteristiky. Podtypy mohou být vybrány na základě různých kritérií.

## Příklady:

- E-mailová zpráva
  - Podle formátu: Zpráva v prostém textu / Formátovaná zpráva
  - Podle účelu: Osobní zpráva / Pracovní e-mail
- Audio nahrávka
  - Podle obsahu: Hudební záznam / Záznam řeči
  - Podle datového obsahu: Pouze audio / Audio s texty / Audio s metadaty

## 2. ANALÝZA STRUKTURY OBJEKTU

Hodnotitel analyzuje objekt a získá úplný seznam technických vlastností. Cílem je pochopení druhu technických vlastností a jejich možných hodnot.

### Příklady:

- Bitmapa
  - Formát: JPEG, TIFF, PNG
  - Barevný prostor: RGB, CMYK, BW
- Audio nahrávka
  - Bitový tok: pevný BR, variabilní BR v rozsahu od-do
  - Frekvenční rozsah:

# 3. IDENTIFIKACE ÚČELU TECHNICKÝCH VLASTNOSTÍ

Smyslem je určení role, kterou má vlastnost v rámci Datového Objektu. Pokud vlastnost přispívá k znovuvytvoření Informačního Objektu, je vhodné zaznamenat její hodnotu pro pozdější vyhodnocení výsledku uchovávacích aktivit.

## **Kategorizace technických vlastností:**

- **Obsahové:** vztahují se k informačnímu obsahu (text, statické obrazy, pohyblivé obrazy, zvuk). Příklad: délka, počet znaků, počet kapitol.
- **Kontextové:** Popisují vlastnosti prostředí v němž byl obsah vytvořen nebo které mají vliv na jeho smysl. Příklad: datum vzniku, jméno autora.
- **Poskytovací:** Informace přispívající k znovuvytvoření. Příklad: použité písmo, velikost, bitová hloubka.
- **Strukturální:** Vztahy mezi dvěma nebo více částmi Obsahu. Příklad: titulky k filmu, OCR přepis.
- **Behaviorální:** Popisují interakci obsahu. Příklad: hypertextové odkazy.

## 4. URČENÍ OČEKÁVANÉHO CHOVÁNÍ

Hodnotitel posuzuje různé typy aktivit které uživatel může chtít provádět. V této fázi analýzy jsou brány v potaz všechny způsoby používání Objektu bez vazby na konkrétní stakeholdery.

Chování je popisováno z pohledu možného uživatele, který chce jeho prostřednictvím dosáhnout určitého výsledku.

### **Příklady:**

- Chci zobrazit historický text v původní podobě.
- Chci text statisticky analyzovat.
- Chci použít ilustrace ve své vědecké práci.

# 5. KLASIFIKACE CHOVÁNÍ DO FUNKCÍ

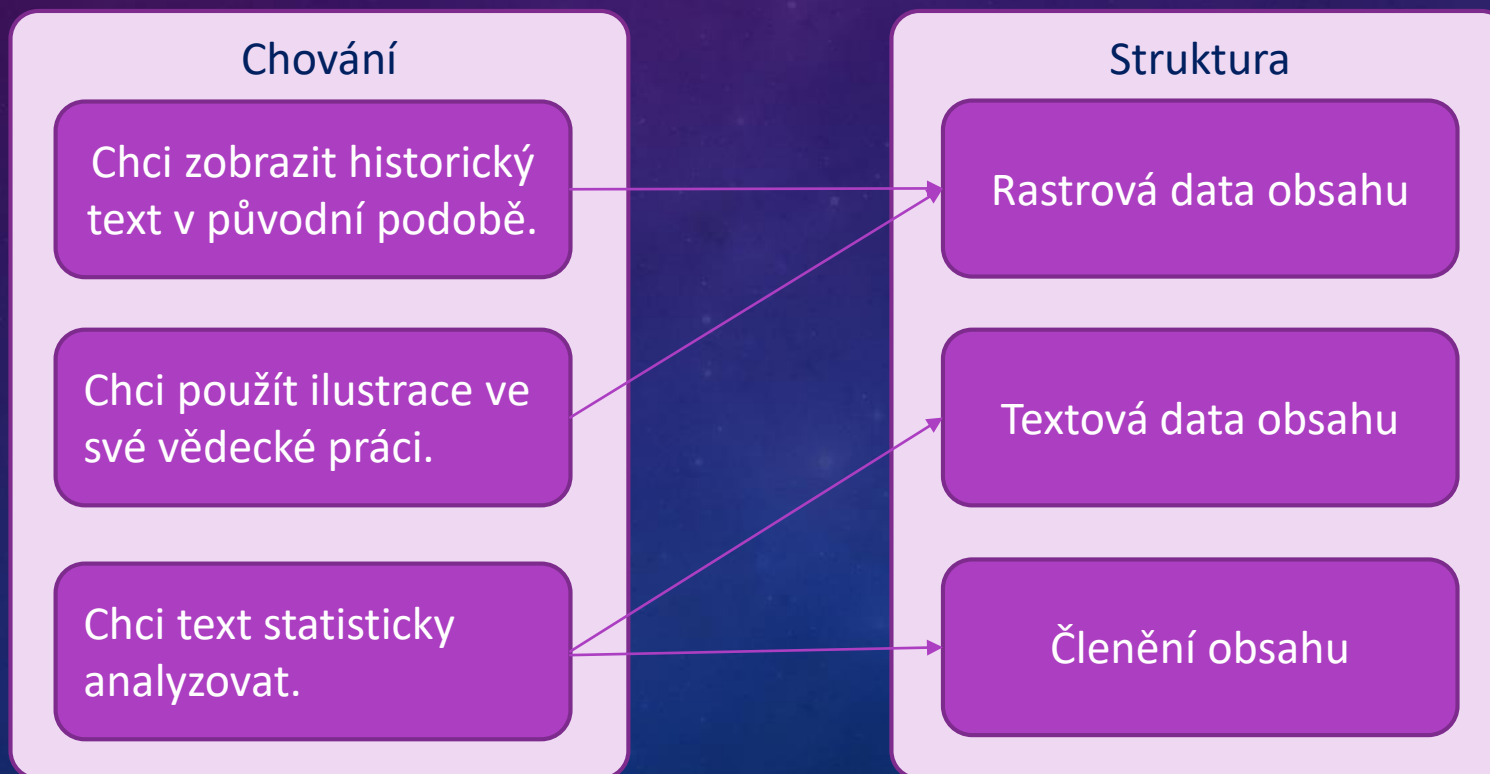
Chování identifikované v předchozím kroku je převedeno do soustavy funkcí. Funkce mohou posloužit jako základ pro přizpůsobení budoucí formy Informačního Objektu potřebám stakeholderů





## 6. PROPOJENÍ STRUKTURY A CHOVÁNÍ

Hodnotitel identifikuje část technických vlastností Datového Objektu které přispívají k vytvoření Informačního Objektu. Tyto vlastnosti mohou být následně měřeny a vyhodnocovány po provedení uchovávacích akcí.



# 7. REVIZE A DOKONČENÍ

Hodnotitel posoudí všechny předchozí kroky a zváží, zda je nutné provést revizi. V této fázi by měly být pokládány otázky:

1. Existuje nějaké další chování, které by mělo být zváženo?
2. Lze některou z identifikovaných funkcí rozložit na dvě nebo více funkcí, které budou přesnější?
3. Existují nějaké další vlastnosti, které by měly být propojeny s funkcemi?

# ANALÝZA POŽADAVKŮ STAKEHOLDERŮ

Hodnotitel identifikuje jednoho nebo více stakeholderů kteří mají nějaký vztah k Inofrmačnímu Objektu a analyzuje funkce, které chtějí provádět. Tyto funkce jsou porovnány s funkcemi zjištěnými při analýze objektů.

1. Identifikace stakeholdera
2. Výběr typů objektů pro analýzu
3. Určení skutečného chování
4. Klasifikace chování do funkcí
5. Porovnání s funkcemi objektů
6. Přiřazení přijatelných rozmezí hodnot
7. Revize a dokončení

# 1. IDENTIFIKACE STAKEHOLDERA

Prvním krokem je určení stakeholderů, kteří budou cílem analýzy, a získání jejich spolupráce. Pro získání informací o akcích, které budou stakeholdeři provádět, lze použít běžné výzkumné metody – dotazníková šetření, strukturovaná i nestrukturovaná interview nebo pozorování.

## **Příklady:**

- Konkrétní instituce, která má definované cíle
- Role pracující v určitém prostředí nebo se specifickými softwarovými nástroji
- Role, která má přímý vztah k Informačnímu Objektu (například tvůrce)
- Výzkumník, který bude analyzovat obsah pomocí kvantitativních výzkumných metod
- Student, který si přeje číst informace, ale nepotřebuje nutně editovat nebo analyzovat data

## 2. VÝBĚR TYPU OBJEKTU PRO ANALÝZU

Hodnotitel vybírá objekty, které jsou používány stakeholdery. Doporučený postup je vybírat objekty na základě obecných funkcí definovaných v 5. kroku analýzy objektů.

### **Příklady:**

- Digitalizovaná publikace
- Audio nahrávka

# 3. URČENÍ SKUTEČNÉHO CHOVÁNÍ

Cílem je určení aktivit, které specifická kategorie stakeholderů bude provádět s objektem. V FBS modelu jsou aktivity stakeholderů označovány jako “skutečné” chování a jsou odlišeny od “předpokládaného” chování definovaného v kroku 4 analýzy objektů.

- Skutečné chování může představovat podmnožinu předpokládaného chování
- Může však zahrnovat i chování, které nebylo při analýze objektu definováno
- Předpokládané chování může posloužit jako podklad pro přípravu interview
- Stakeholdeři mohou být požádáni, aby předvedli způsob, jakým budou používat Informační Objekt

## 4. KLASIFIKACE CHOVÁNÍ DO FUNKCÍ

Obdoba kroku č. 5 z analýzy objektů, jehož výstup by hodnotitel měl použít jako podklad při klasifikaci.

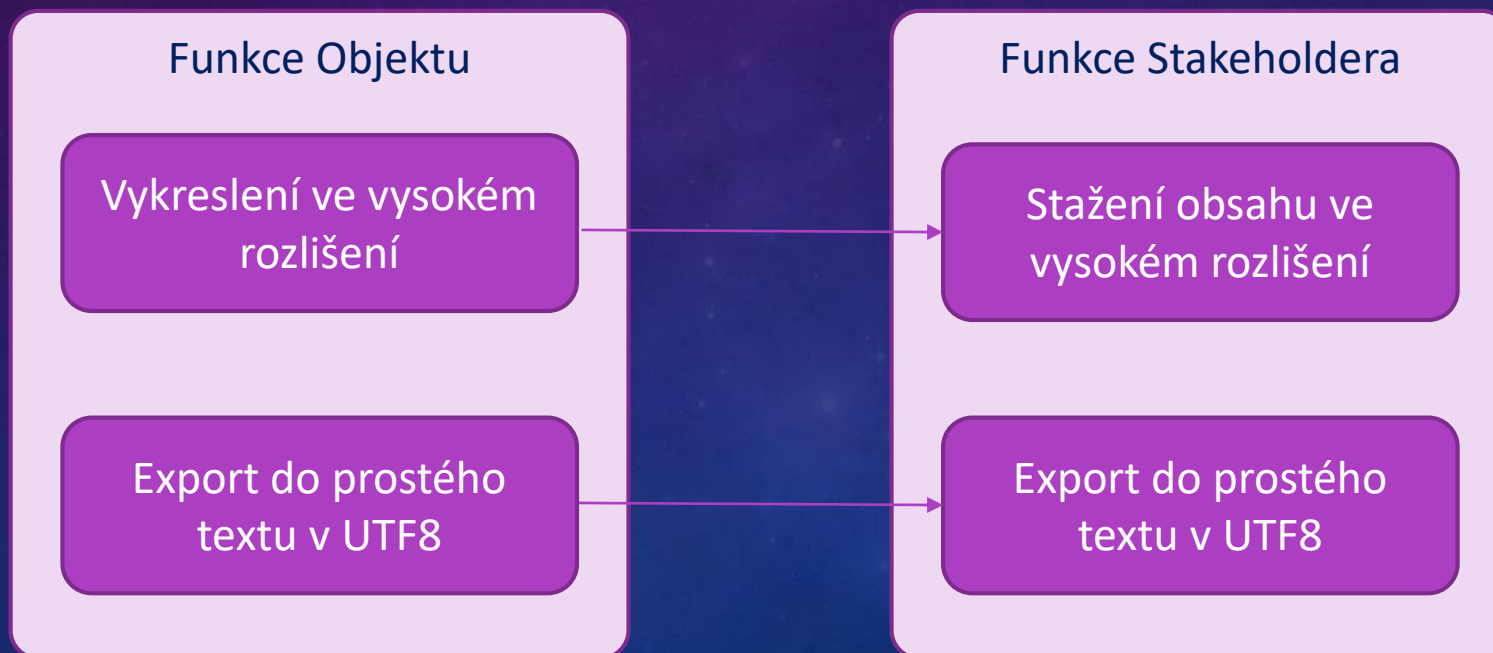
- Mohou být definovány nové funkce, které nebyly zaznamenány v souboru předpokládaného chování
- Funkce jsou definovány z pohledu stakeholdera



# 5. POROVNÁNÍ FUNKCÍ

Cílem je vytvoření seznamu technických vlastností, které jsou významné při provádění funkcí požadovaných stakeholderem.

- Funkce identifikované v předchozím kroku jsou porovnány s funkcemi objektu





## 6. PŘIŘAZENÍ PŘIJATELNÝCH ROZMEZÍ HODNOT

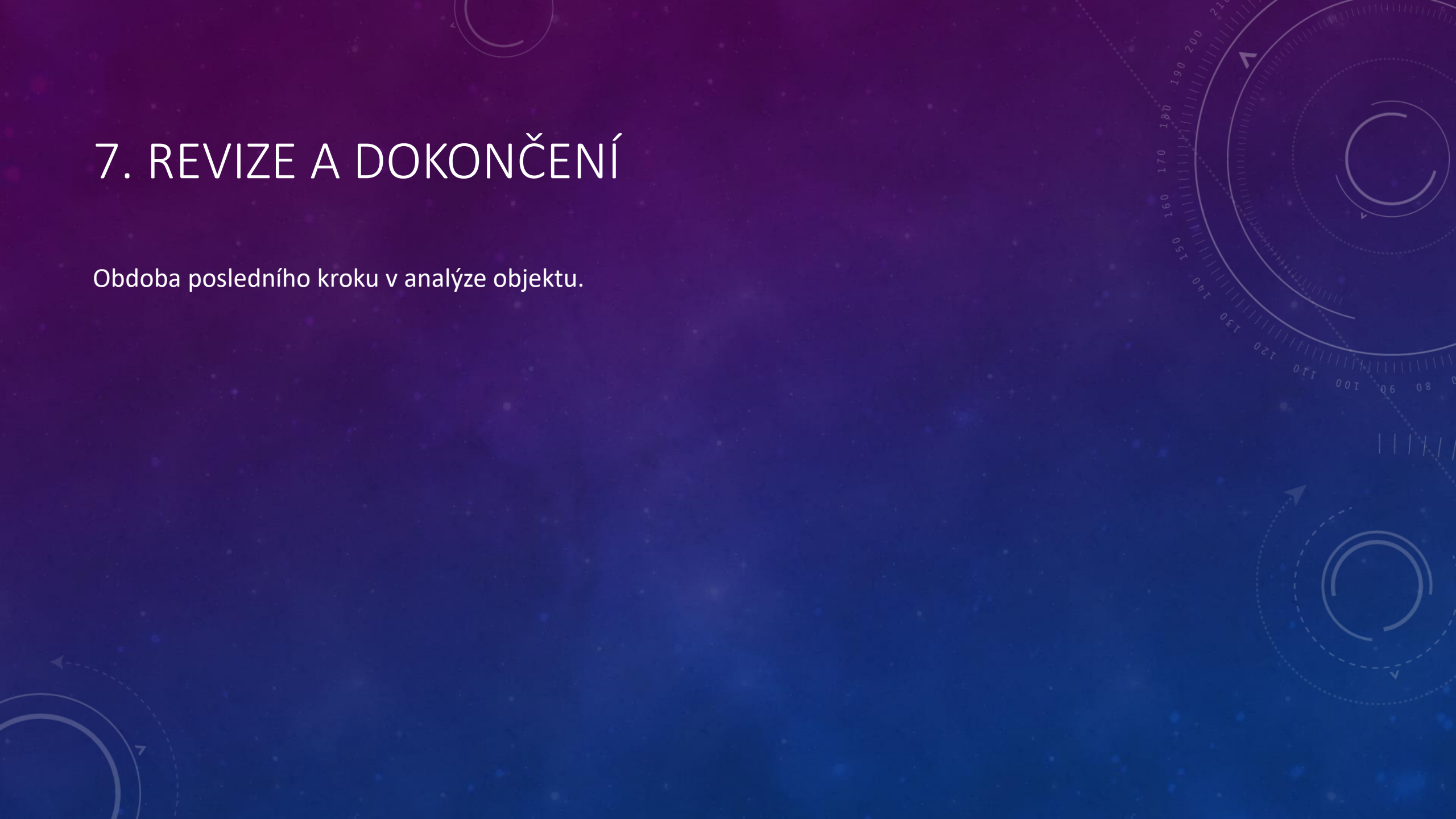
Cílem je určení mezních hodnot vlastností, které jsou pro stakeholdera přijatelné. Tyto mezní hodnoty mohou být použity při vyhodnocení uchovávacích akcí při vytváření Informačního Objektu pro konkrétního stakeholdera. V některých případech není možné (nebo praktické) zachovávat všechny hodnoty původního objektu při jeho přetváření, proto je nutné určit míru přesnosti.

Používají se čtyři typy mezních hodnot:

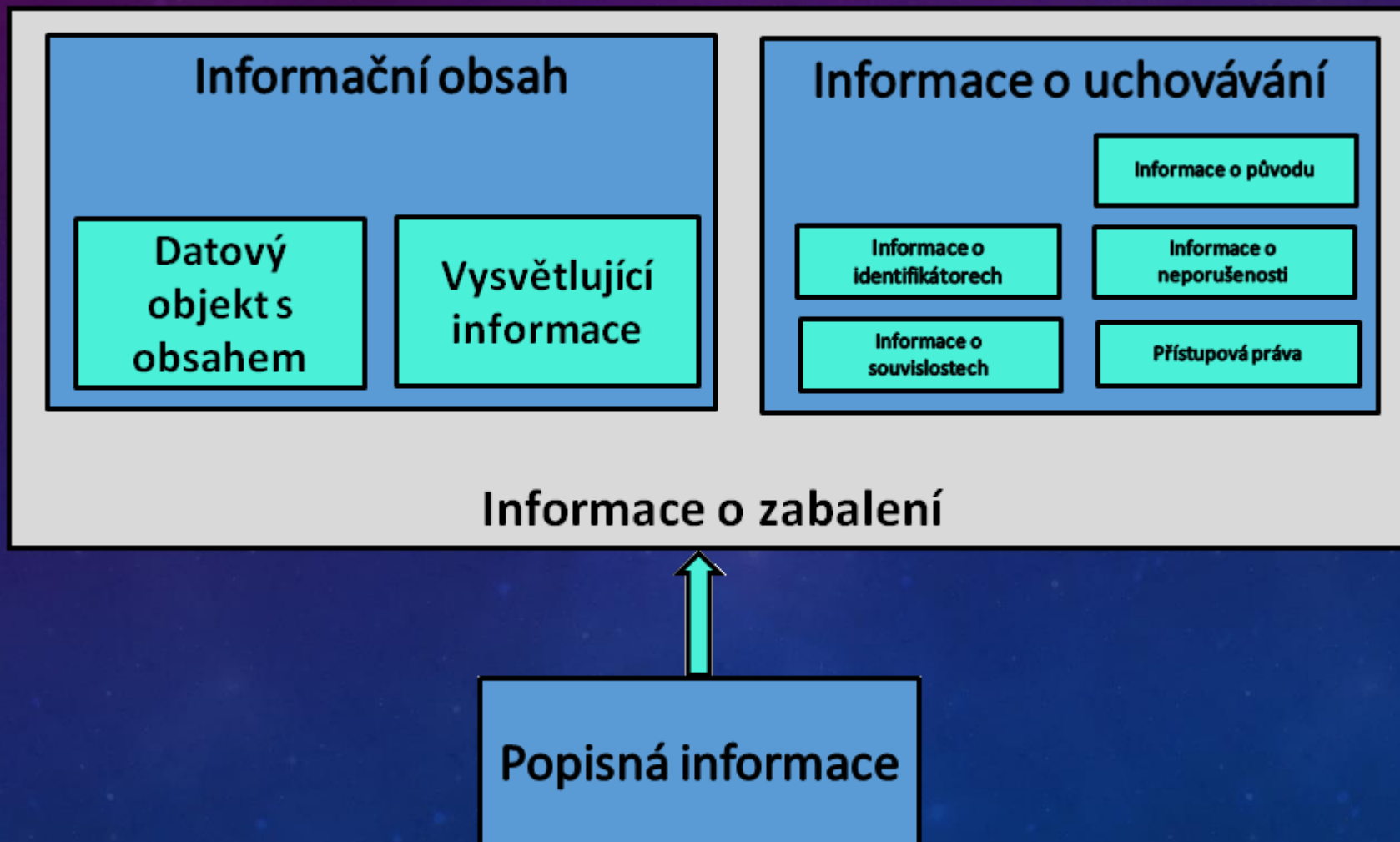
1. **Rovnost:** vlastnost uložená v záznamu musí být přesně shodná s hodnotami uloženými v metadatech – příklad: DPI rasterového obrazu musí být přesně 300
2. **Minimum:** nejnižší číselná hodnota, která je přípustná – příklad: DPI musí být nejméně 200. Minimum může být použito pouze v kombinaci s maximem.
3. **Maximum:** nejvyšší číselná hodnota, která je přípustná
4. **Rozsah:** hodnota odpovídá jedné z více možných hodnot

# 7. REVIZE A DOKONČENÍ

Obdoba posledního kroku v analýze objektu.



# ARCHIVNÍ INFORMAČNÍ BALÍČEK



# ADMINISTRATIVNÍ METADATA

OAIS archivy často využívají metadatový slovník PREMIS

- Podpora integrace s METS
- Obsahuje *semantic unit 1.4 significantProperties*:
  - significantPropertiesType
  - significantPropertiesValue
  - significantPropertiesExtension

```
▼<premis:premis xmlns:premis="http://www.loc.gov/premis/v3" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="ht
xsi:schemaLocation="http://www.loc.gov/premis/v3 http://www.loc.gov/standards/premis/premis-3-0-draft.xsd" version=
▼<premis:object xsi:type="premis:file">
  ▼<premis:objectIdentifier>
    <premis:objectIdentifierType>local</premis:objectIdentifierType>
    <premis:objectIdentifierValue>001</premis:objectIdentifierValue>
  </premis:objectIdentifier>
  ▼<premis:preservationLevel>
    <premis:preservationLevelType>logical preservation</premis:preservationLevelType>
    <premis:preservationLevelValue>full preservation</premis:preservationLevelValue>
    <premis:preservationLevelRole authority="preservationLevelRole" authorityURI="http://id.loc.gov/vocabulary/pre
valueURI="http://id.loc.gov/vocabulary/preservation/preservationLevelRole/int">intention</premis:preservationL
    <premis:preservationLevelRationale>institutional policy</premis:preservationLevelRationale>
    <premis:preservationLevelDateAssigned>2015-02-23</premis:preservationLevelDateAssigned>
  </premis:preservationLevel>
  ▼<premis:significantProperties>
    <premis:significantPropertiesType>ImageWidth</premis:significantPropertiesType>
    <premis:significantPropertiesValue>2464</premis:significantPropertiesValue>
  </premis:significantProperties>
  ▼<premis:significantProperties>
    <premis:significantPropertiesType>ImageHeight</premis:significantPropertiesType>
    <premis:significantPropertiesValue>3248</premis:significantPropertiesValue>
  </premis:significantProperties>
  ▼<premis:significantProperties>
    <premis:significantPropertiesType>XResolution</premis:significantPropertiesType>
    <premis:significantPropertiesValue>300</premis:significantPropertiesValue>
  </premis:significantProperties>
  ▼<premis:significantProperties>
    <premis:significantPropertiesType>YResolution</premis:significantPropertiesType>
    <premis:significantPropertiesValue>300</premis:significantPropertiesValue>
  </premis:significantProperties>
  ▼<premis:significantProperties>
    <premis:significantPropertiesType>SamplesPerPixel</premis:significantPropertiesType>
    <premis:significantPropertiesValue>3</premis:significantPropertiesValue>
  </premis:significantProperties>
  ▼<premis:significantProperties>
    <premis:significantPropertiesType>BitsPerSample</premis:significantPropertiesType>
    <premis:significantPropertiesValue>8</premis:significantPropertiesValue>
  </premis:significantProperties>
```

# DOPORUČENÉ ZDROJE A ODKAZY

- [Significant Properties, Authenticity, Provenance, Representation](#)  
*Článek Davida Giarety o významných vlastnostech z pohledu modelu OAIS*
- 1. [Framework InSpect](#)  
*Detailní popis frameworku InSPECT*
- 2. [Závěrečná zpráva projektu InSPECT](#)  
*Shrnutí projektu a jeho výstupů*
- 3. [Testovací zprávy projektu InSPECT](#)  
*Určování významných vlastností audiozáznamů, strukturovaného textu, rastrových obrazů a e-mailů.*
- 4. [PREMIS Data Dictionary](#)  
*Standard pro metadata uchování*

DĚKUJI ZA VAŠI POZORNOST

... a rád zodpovím vaše dotazy

Michal Konečný

[jsem@digitalnikurator.cz](mailto:jsem@digitalnikurator.cz)



# POUŽITÉ ZDROJE

- DAPPERT, A., Farquhar, A. 2008. *Significance is in the eye of the stakeholder*. In International Conference on Theory and Practice of Digital Libraries (s. 297-308). Springer, Berlin, Heidelberg.
- DCC Curation Lifecycle Model. 2015. *DCC: because good research needs good data* [online]. Digital Curation Centre [cit. 2018-11-03]. Dostupné z: <http://www.dcc.ac.uk/resources/curation-lifecycle-model>
- GIARETTA, David. 2008. DCC Approach to Digital Curation. *DCC Development* [online]. 2008-07-14 [cit. 2018-11-14]. Dostupné z: <http://twiki.dcc.rl.ac.uk/bin/view/OLD/DCCApproachToCuration>
- GIARETTA, David, Matthews, B., Bicarregui, J., Lambert, S., Guercio, M., Michetti, G., & Sawyer, D. 2009. *Significant Properties, Authenticity, Provenance, Representation Information and OAIS Information*. UC Office of the President: California Digital Library [online]. 2009-05-10 [cit. 2018-11-14]. Dostupné z <https://escholarship.org/uc/item/0wf3j9cw>
- KNIGHT, Gareth, Grace, S., & Montague, L. *Framework for the definition of significant properties*. 2008. JISC. Dostupné z: <https://web.archive.org/web/20160416050908/http://www.significantproperties.org/inspect-finalreport.pdf>
- KONEČNÝ, Michal. 2016. *Návrh kompetenčního modelu a kurikula digitálního kurátorství* [online]. Brno. Diplomová práce. Masarykova univerzita, Filozofická fakulta. Vedoucí práce Miroslav Bartošek Dostupné z: [http://is.muni.cz/th/426710/ff\\_m/](http://is.muni.cz/th/426710/ff_m/).
- LAVOIE, Brian F. a Marek MELICHAR (překl.). 2015. *Úvod do referenčního modelu otevřeného archivačního informačního systému (OAIS)*. Výstup projektu LTP-Pilot. Brno. Dostupné z: <https://drive.google.com/file/d/0BzOLuOh094X8ejZONHVOY3VuOTg/>
- ŠVÁSTOVA, Pavla. 2014. *MODS a další metadatová schémata v oblasti digitalizace dokumentů*. In: *Knihovny současnosti ...: sborník z ... konference, konané ve dnech ... v ..* Brno: Sdružení knihoven ČR, 2014. ISBN 978-80-86249-54-4. ISSN 1805-6962.