

Představení projektu

TA ČR: Vytvoření standardů pro komunikaci informačního systému digitálního archivu s jeho okolím

Petr Pytelka, LightComp v.o.s.

NTM 2024 – Odborné skupiny specializovaných archivů České
archivní společnosti

LightComp

Cíle projektu

TITSMV301: Vytvoření standardů pro komunikaci informačního systému digitálního archivu s jeho okolím

Certifikované metodiky:

Tvorba a forma výměnných balíčků

Komunikační protokol umožňující přenos a zpracování dat z DA

Ověření výsledku – aplikace:

Validátor výměnných balíčků

Zpracování digitálních archiválií v Elza

Dokončení:



Způsob řešení



Program **Beta2**



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY

Úzká spolupráce se zástupci archivů:



Národní archiv



ARCHIV

HLAVNÍHO
MĚSTA PRAHY



ÚSTAV DĚJIN UNIVERZITY
KARLOVY A ARCHIV
UNIVERZITY KARLOVY
Univerzita Karlova

Řešitelé:

LightComp
Enhance Your Technology



MARBES







T A
Č R

Dříve dokončené výsledky
pro archivy

ELZA

CAM

Metodiky a standardizované
postupy:

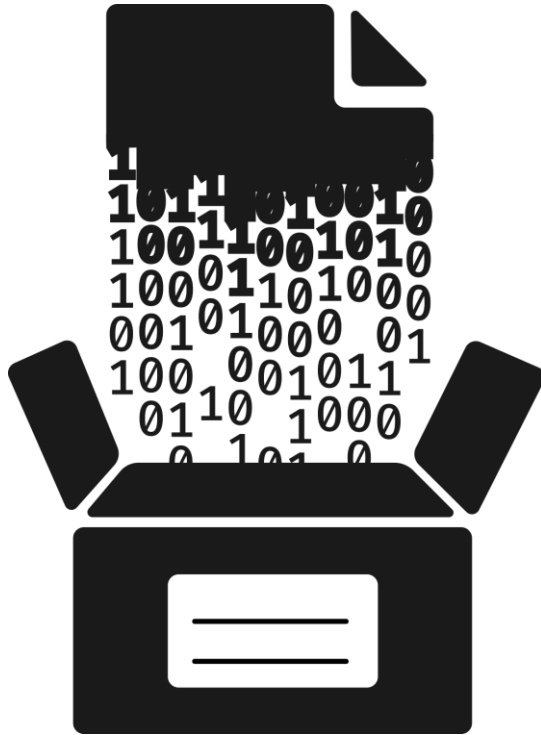
Metodika implementace ELZA

Technická rozhraní

EAD pro archivní popis

CAM – dopad na ZP

Otázky a oblasti řešení



Příjem digitálních archiválií a jejich uložení je dokončeno.

Jaká data jsou uchovávána v archivních balíčcích?

Jaká je granularita a podoba archivního balíčku v jednotlivých DA?

Rozsah metadat v balíčku?

Jak zajistit úplné předání a správné porozumění příjemcem?

Jak zpracovávat digitálních archiválie?

Program **Beta2**

T A
Č R

TITSMV301: Vytvoření standardů pro komunikaci informačního systému digitálního archivu s jeho okolím

Východiska:

Referenční model OAIS

Specifikace E-ARK (=European Archival Records and Knowledge Preservation, viz dilcis.eu)

Oborové standardy METS, PREMIS, EAD

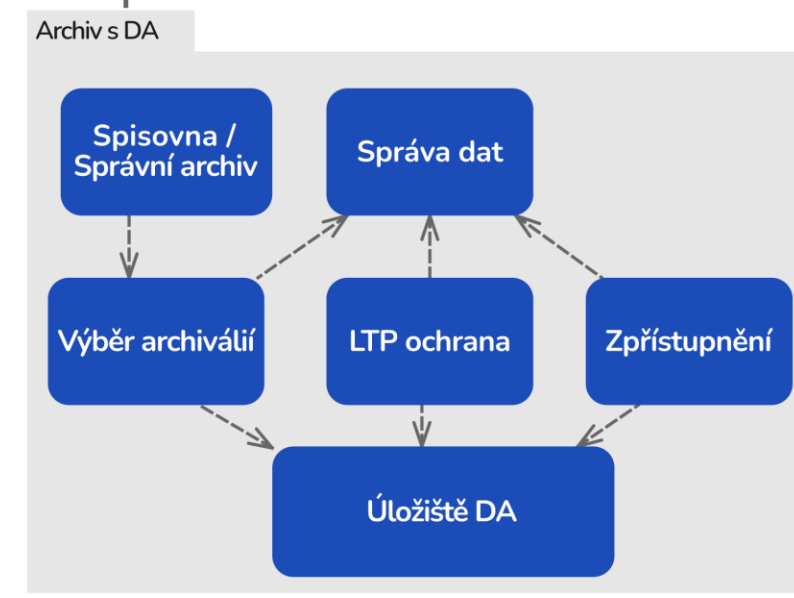
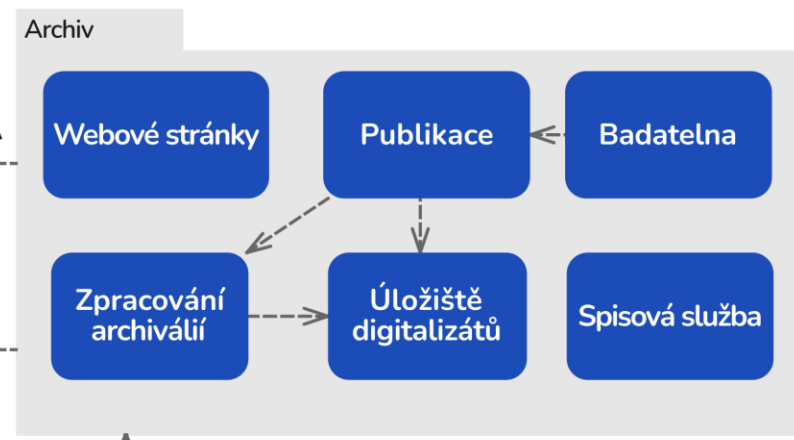
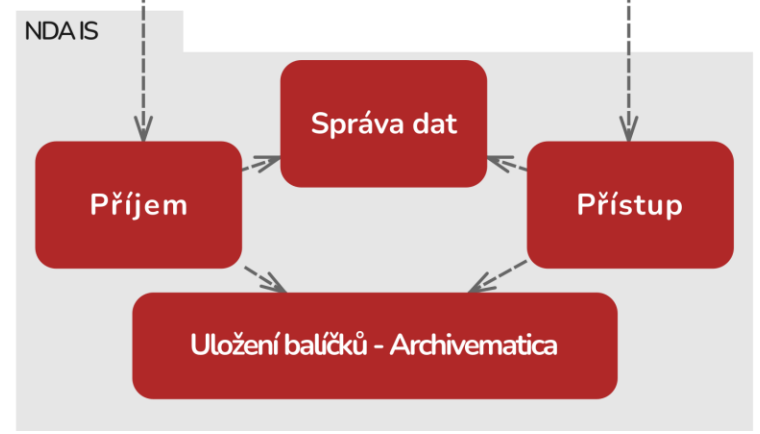
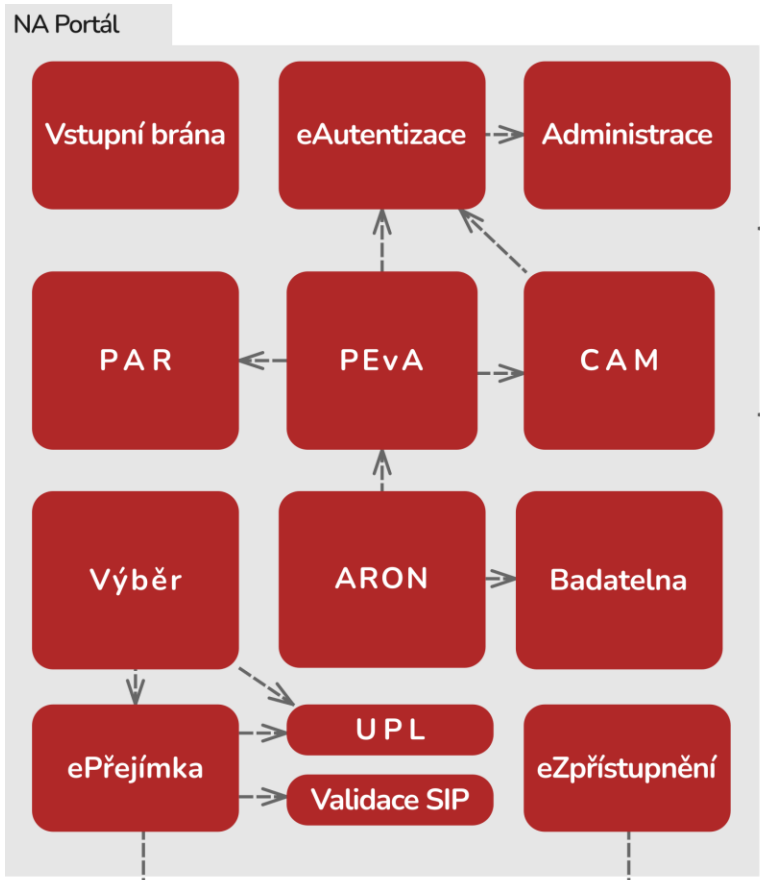


Zohledněná specifika:

Účel užití

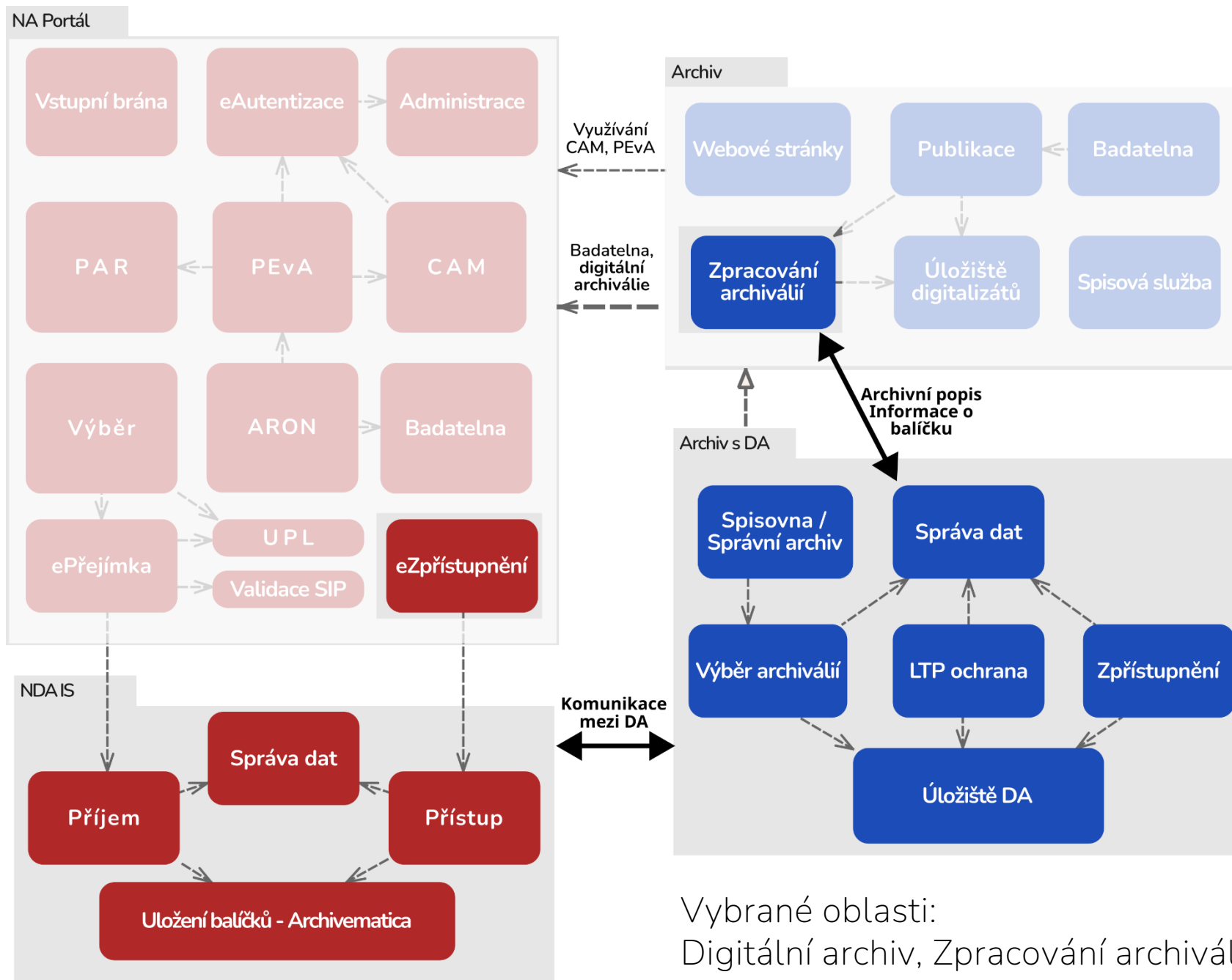
Spisové služby (NSESSS)

Zkušenosti ze stávající implementací DA



Využívání CAM, PEvA

Badatelna, digitální archiválie

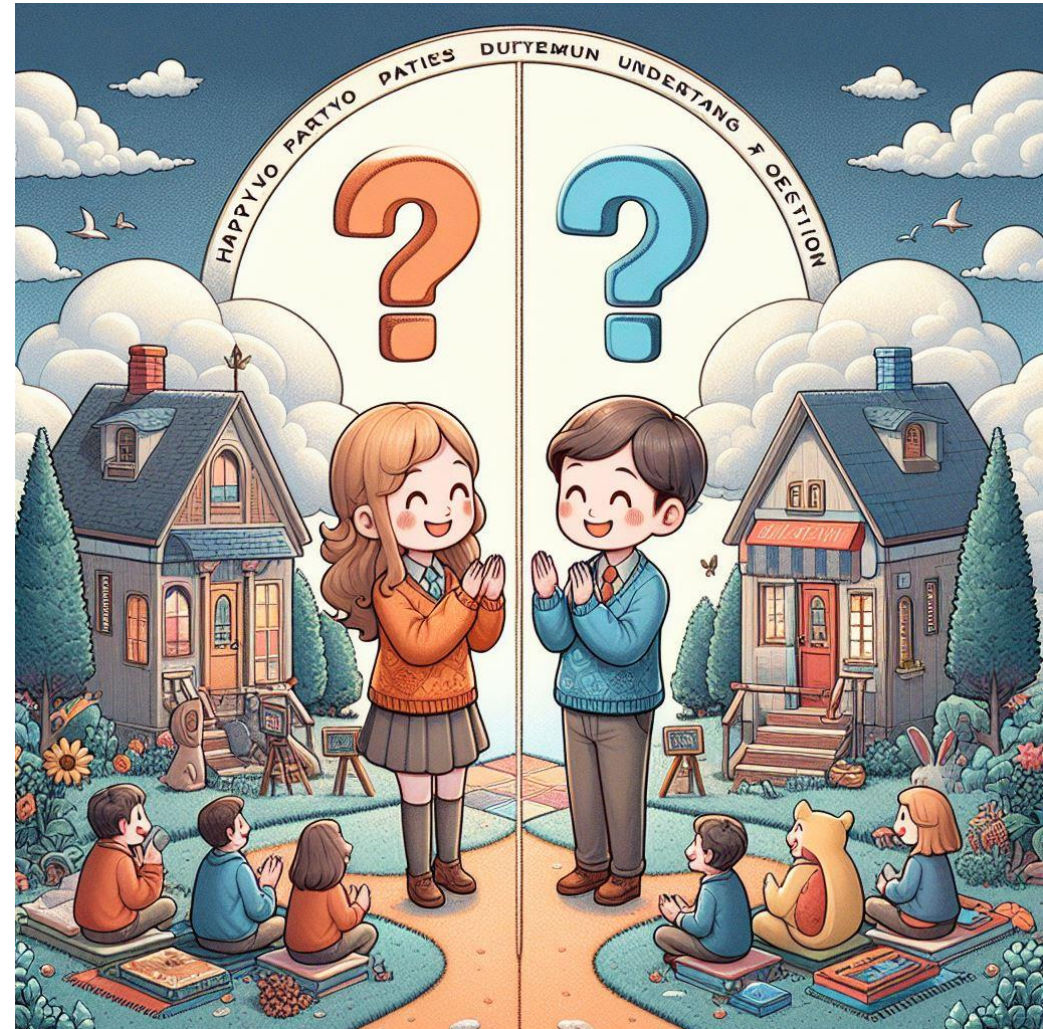


Vybrané oblasti:
Digitální archiv, Zpracování archiválií

Základní princip

Obě strany musí mít shodné porozumění obsahu balíčku.

- Předávají se jen přesně specifikované informace
- Definované typy předávaného obsahu (NSESSS, volné soubory)
- Doplnující informace vyžadují dohodu obou stran



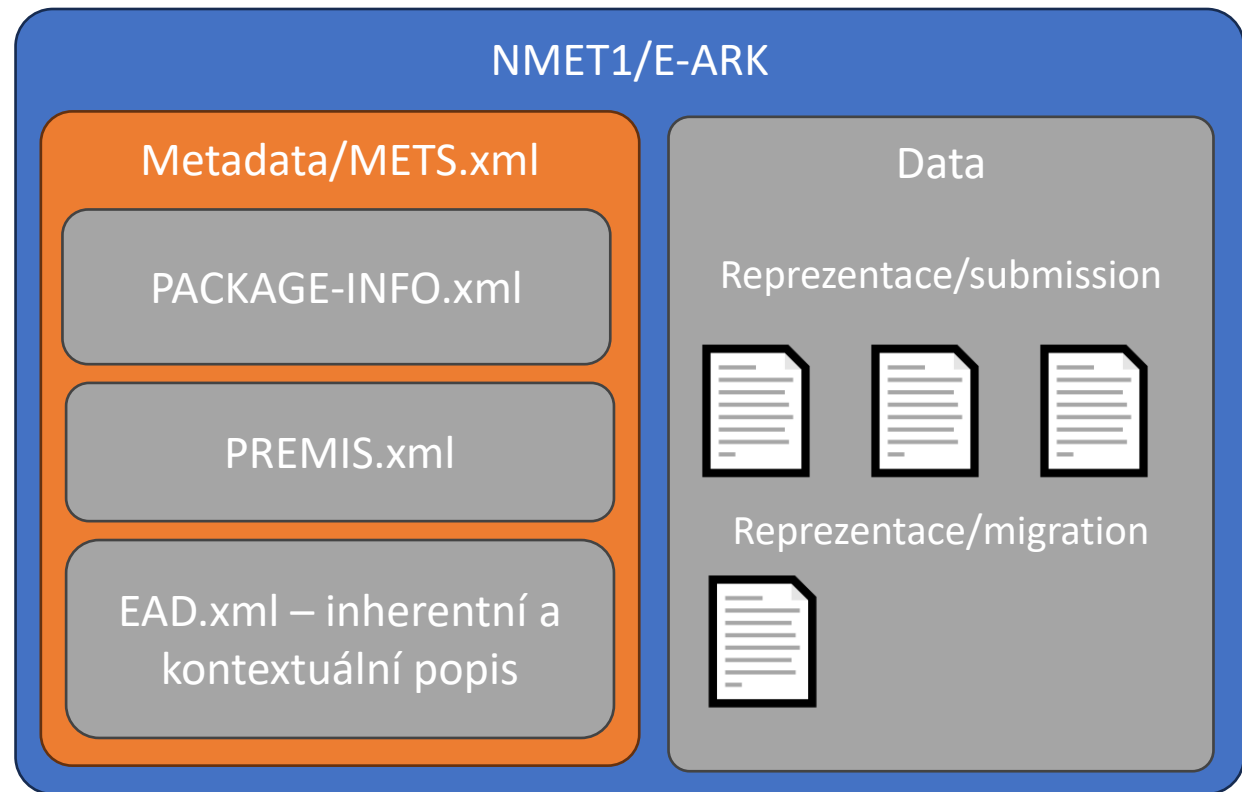
První výsledky a závěry I

Připravena specializace E-ARK, pravidla pro METS.xml

Metadata o uchování
(formát PREMIS),
pro balíček a jednotlivé
komponenty.

Popisná metadata:
(EADv3 a ČR Profil)
dva typy popisu

Vzorové balíčky



První výsledky a závěry II

Návrh komunikačních rozhraní.

Povinnosti DA pro zpracování archiválií.

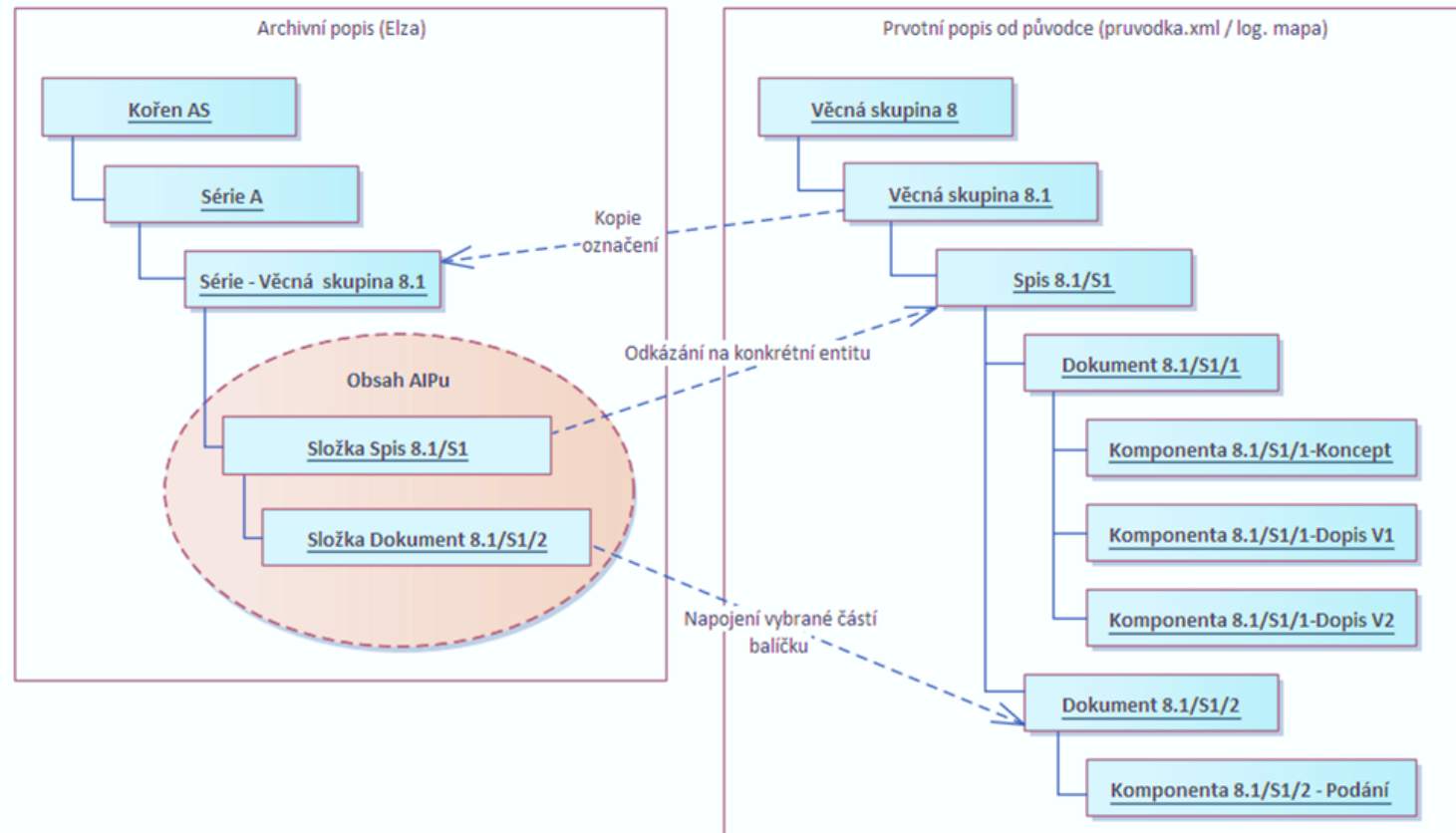
Metody předávání balíčků a rozsah metadat.

Možnost uložení archivního popisu zpět do digitálního archivu.

```
da-api > tsp main.tsp > {} DA > {} Updates > updates
1 import "@typespec/http";
2 import "@typespec/rest";
3 import "@typespec/openapi3";
4
5 using TypeSpec.Http;
6
7 @service({
8   title: "Služby DA",
9   version: "2024-03-22",
10 })
11 namespace DA {
12   // Identifikátor AIP
13   scalar AipId extends string;
14
15   /**
16    * Identifikátor verze AIP
17    * - je povinný z důvodu kontroly archivního popisu při zaslání
18    */
19   scalar AipVersion extends string;
20
21   @route("/packages")
22   namespace Packages {
23
24     /**
25      * Stažení jednotlivého balíčku v poslední verzi.
26      * Metoda vrací ZIP s daty.
27
28      * @dipType Požadovaná podoba balíčku, přípustné hodnoty jsou
29      * - package_info - PACKAGE-INFO.xml
30      * - archdesc - PACKAGE-INFO.xml a EAD (inherentní a konte
31      * - metadata_base - METS.xml, PACKAGE-INFO.xml, EAD (inhe
32      * - aip_base - metadata_base + binární data
```

Soulad se specifikací nutný pro akreditaci DA

Dopady na archivní popis



Druhy popisu: inherentní, kontextuální
Metody popisu: celý AIP, celá úroveň popisu, vybraná úroveň, komponenta či soubor.

Informace o balíčku (PACKAGE-INFO)

Formou agentů, objektů a událostí dokumentuje historii balíčku (PREMIS v3.0).

vznik (creation), vložení do DA (ingest), delimitace (transfer), předání (export)

```
<!-- Vznik dokumentu/objektu z něž se stala archiválie-->
<event>
  <eventIdentifier>
    <eventIdentifierType>local</eventIdentifierType>
    <eventIdentifierValue>uuid-27c8a259-7cf5-47e2-9bce-3b3da7f7c063</eventIdentifierValue>
  </eventIdentifier>
  <!-- Určení, že se jedná o vznik -->
  <eventType>cre</eventType>
  <!-- Datace archiválie uložené v balíčku -->
  <eventDateTime>2020-10-12/2023-05-16</eventDateTime>

  <linkingAgentIdentifier>
    <linkingAgentIdentifierType>local</linkingAgentIdentifierType>
    <linkingAgentIdentifierValue>uuid-916d7dc2-b19e-43d4-ba28-4c87b5d7c697</linkingAgentIdentifierValue>
  <!-- původce -->
```

Proč projekt sledovat?

Inspirace pro realizaci digitálního archivu (vybraných aspektů).

Přímý dopad na způsob zpracování digitálních archiválií.

Výsledek bude letos (technické specifikace, software).

Technika, jak se dostat k digitálním archiváliím v NDA.

Změny viditelné v Elza, referenční implementace úložiště.

Za řešitele:

Budeme rádi za zpětnou vazbu. Jsme otevřeni připomínkám.

Děkuji za pozornost

Petr Pytelka, LightComp v.o.s.
petr.pytelka@lightcomp.cz

LightComp