



National Archives
Czech Republic

Das nationale Archivportal: der standardisierte Weg zur digitalen Archivierung strukturierter und unstrukturierter Daten

XArchiv?
Automatisierung und Standardisierung
bei der Digitalen Archivierung

Zbyšek Stodůlka

München, 27. 1. 2021

Nationaler Standard für elektronische DMS/VBS

- seit 2009/2010 (erste Version)
 - Dokumentenaustauschs zwischen DMS/VBS (Beilage Nr. 1)
 - Übergabe an das Archiv (Beilage Nr. 2)
- seit 2012 (zweite Version)
 - neue Konstruktion des SIPs – 1 Akte/ Vorgang/ Dokument = 1 SIP (Beilage Nr. 2+3)
 - neue Beilagen:
 - METS (Beilage Nr. 3)
 - XML-Schema zur Übermittlung der Bewertungsentscheidung und zur Übermittlung der ID AIP im digitalen Archiv (Beilage Nr. 4)
 - XML-Schema zur Export und Import des Akten- und Aussonderungsplans (beilage Nr. 5)
- seit 2017 (dritte Version)
 - neue Beilage Nr. 1
 - zwei Sorten von SIP (LABEL) - 1) für die Bewertung nur Metadaten - Dateien zu konkreten Objekten erst nach Anforderung (mit Beilage Nr. 4); 2) Übergabe an das Archiv - Dateien immer
 - XML-Schema zum Import und Export des Transaktionsprotokoll - im SIP um die Autentizität zu stärken (Beilage Nr. 6)

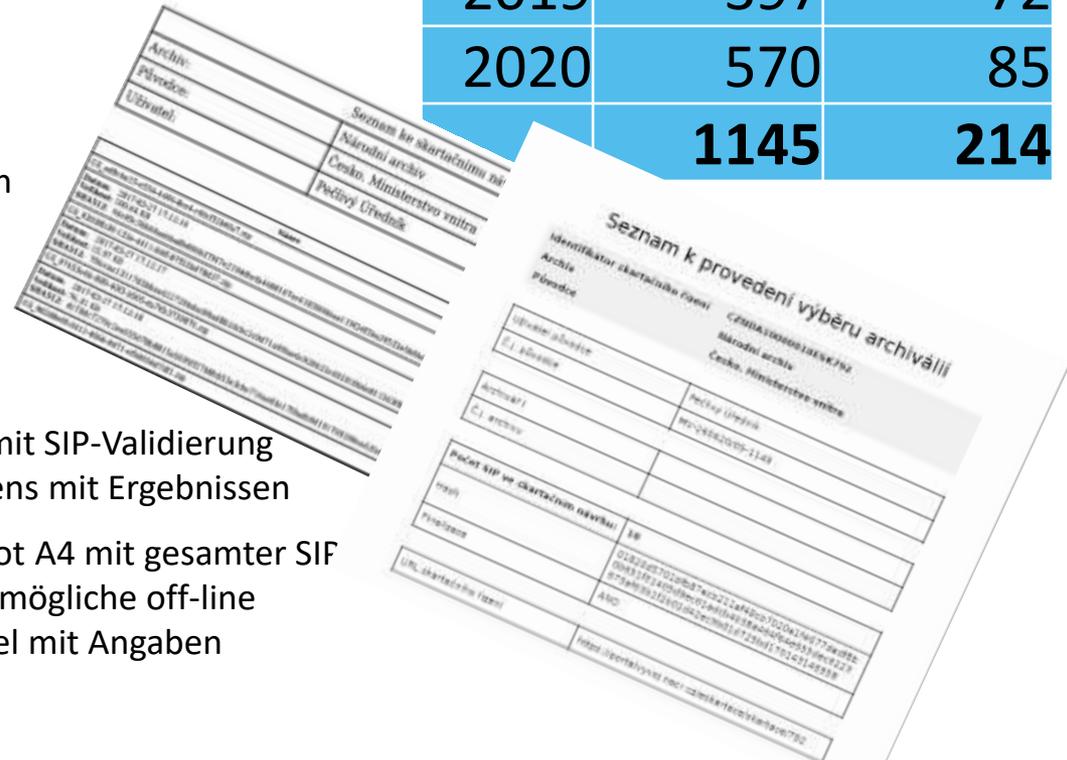
```
<?xml version="3.0" encoding="UTF-8"?>
<meta:mets xmlns:mets="http://www.loc.gov/METS/" xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
http://www.mvcr.cz/nasas/v3 http://www.mvcr.cz/nasas/v3/nasas.xsd http://www.public.cz/erna_trans/v_01_01
TransakcniProtokolNavrh_verze1.2.xsd" TYPE="asis" LABEL="Datový balíček pro předávání dokumentů a jejich metadat do archivu" OBJID="GS_7c24fce-
7443-4c3b-ac97-0db1dbdcb050" ID="GS_7c24fce-7443-4c3b-ac97-0db1dbdcb050" xmlns:xsi="http://www.w3.org/2001/XMLSchema" xmlns:link="http://www.w3.org/1999/xlink"
xmlns:erms="http://www.nerms.cz/nasas/v3" xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:link="http://www.w3.org/1999/xlink"
GSAESR01#484.11.40ESR0148411X01" LASTMODDATE="2020-12-23T15:27:43.929641Z" CREATEDATE="2020-12-23T15:27:43.929641Z">
  <meta:agent TYPE="ORGANIZATION" ID="CRDUSE000A13" ROLE="CREATOR">
    <meta:name>Pečlivý úřad, Úřední 1, 10100</meta:name>
    <meta:note>eSSL</meta:note>
  </meta:agent>
  <meta:agent TYPE="INDIVIDUAL" ID="CRDUSR00A1ZV" ROLE="CREATOR">
    <meta:name>Novák Václav</meta:name>
    <meta:note>eSSL</meta:note>
  </meta:agent>
  </meta:metsHdr>
  <meta:dmdSec ID="dmd001">
    <meta:dmdRpt LABEL="Schéma XML pro předávání dokumentů a jejich metadat do archivu" MIMEType="text/xml" MDTypeVersion="3.0"
    OTHERMDType="NSES5" MDType="OTHER">
      <meta:xmldata>
        <nsas:Spis ID="CRDUX005YF47">
          <nsas:EvidenčníÚdaje>
            <nsas:Identifikace>
              <nsas:Identifikátor zdroj="eSSL">CRDUX005YF47</nsas:Identifikátor>
            </nsas:Identifikace>
            <nsas:Popis>
              <nsas:Nazev>Modernizace trati Píseň-Ostrava, 1.stavba úsek Okružní Křivoklát</nsas:Nazev>
              <nsas:Komentář>Modernizace trati Píseň-Ostrava, 1.stavba úsek Okružní Křivoklát</nsas:Komentář>
            </nsas:Popis>
            <nsas:Přístupnost>
              <nsas:BezpečnostníKategorie>
                <nsas:Identifikátor zdroj="eSSL">0</nsas:Identifikátor>
                <nsas:Název>Dokument ke zveřejnění</nsas:Název>
                <nsas:BezpečnostníStupeň>Dokument ke zveřejnění</nsas:BezpečnostníStupeň>
              </nsas:BezpečnostníKategorie>
            </nsas:Přístupnost>
            <nsas:Příslušnost>
              <nsas:IniciaceDokument>CRDUP003H6CO</nsas:IniciaceDokument>
            </nsas:Příslušnost>
            <nsas:EvidenčníÚdaje>
              <nsas:EvidenčníCíle>
                <nsas:EvidenčníCíle>MP-SOP1057/13</nsas:EvidenčníCíle>
                <nsas:PoraďovéCíle>1057</nsas:PoraďovéCíle>
                <nsas:UrcenéCasoveObdabi>
                  <nsas:Rok>2013</nsas:Rok>
                  <nsas:UrcenéCasoveObdabi>
                    <nsas:NázevEvidenceDokumentu>SOP</nsas:NázevEvidenceDokumentu>
                  </nsas:UrcenéCasoveObdabi>
                </nsas:UrcenéCasoveObdabi>
                <nsas:Převod>
                  <nsas:DatumVytvoreni datum="2013-07-08T13:58:54.000Z">2013-07-08</nsas:DatumVytvoreni>
                </nsas:Převod>
              </nsas:EvidenčníÚdaje>
            </nsas:EvidenčníÚdaje>
            <nsas:Identifikace>
              <nsas:Identifikátor zdroj="eSSL">CRDUX000AFMD</nsas:Identifikátor>
            </nsas:Identifikace>
            <nsas:Popis>
              <nsas:Nazev>ostatní (soukromé stavby v ochranném pásmu dráhy, vedení apod.)</nsas:Nazev>
              <nsas:Komentář>ostatní (soukromé stavby v ochranném pásmu dráhy, vedení apod.)</nsas:Komentář>
            </nsas:Popis>
            <nsas:Převod>
              <nsas:DatumVytvoreni datum="2009-01-13T10:49:47.000Z">2009-01-13</nsas:DatumVytvoreni>
            </nsas:Převod>
          </nsas:Spis>
        </meta:xmldata>
      </meta:dmdRpt>
    </meta:dmdSec>
  </meta:mets>
</meta:mets>
```



Nationales Archivportal

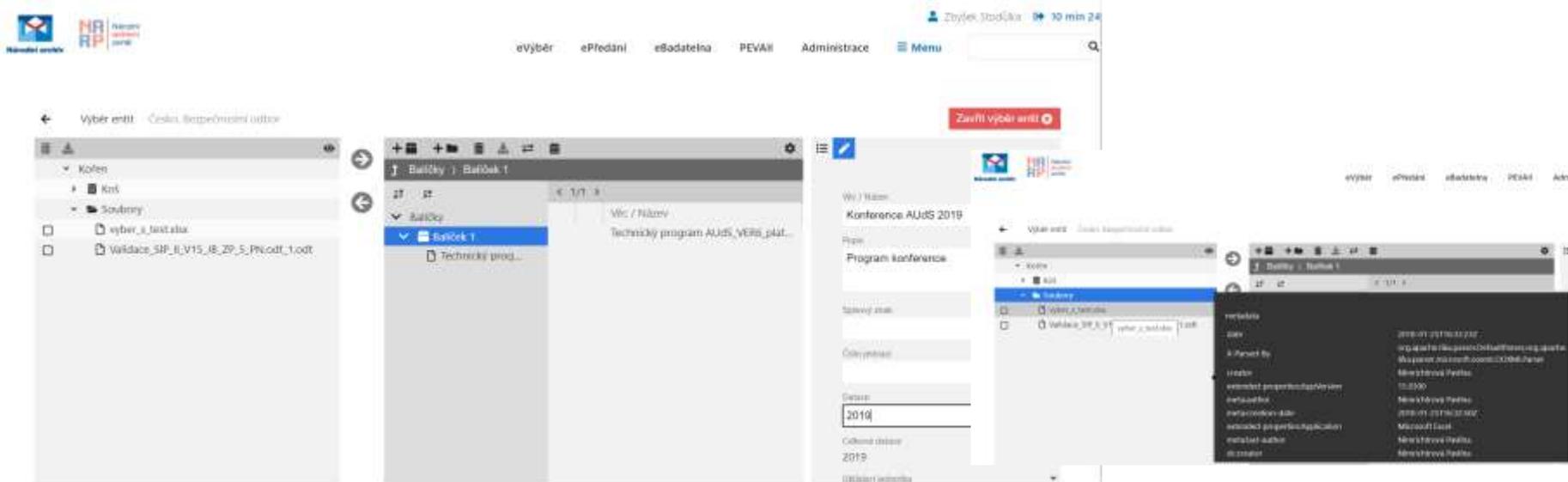
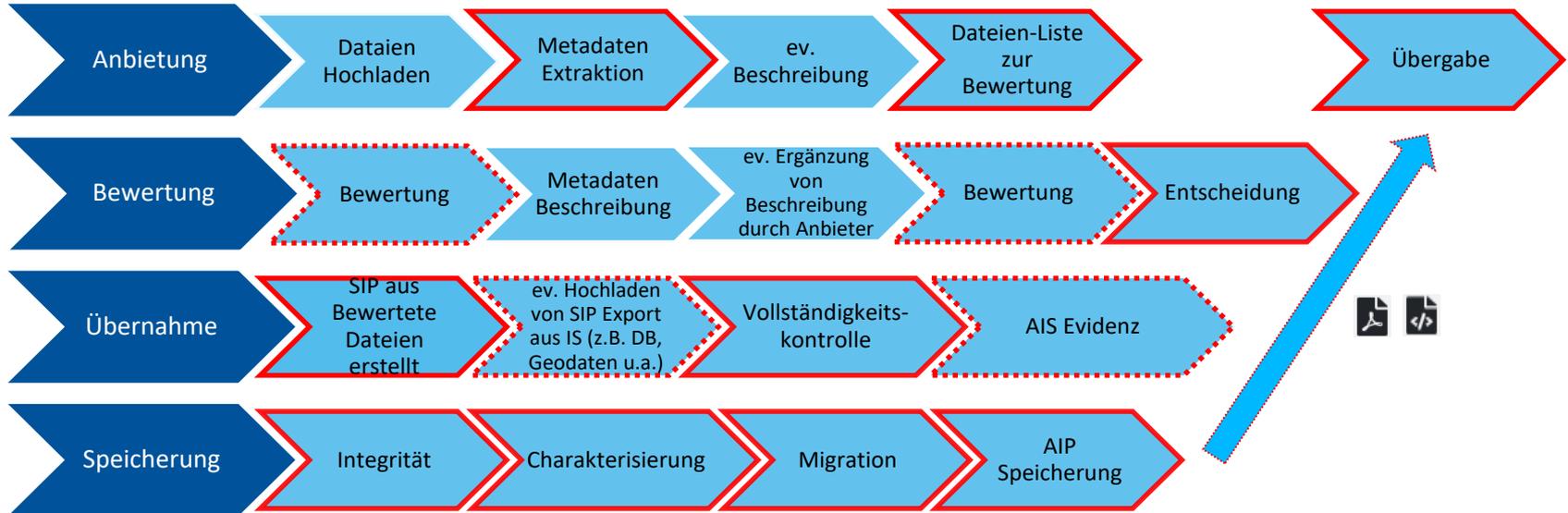
- 2014-2016
- intern entwickelte webbasierte Programme zur Bewertung und Übernahme
- Provenienzstelle lieferte SIPs an DVS / USB / HDD
- Hochladen durch den Archivar
- Provenienzstelle erhielt Ergebnis XML (Beilage Nr. 4) und PDF mit Entscheidung, welche Schriftgutsobjekte zur Kassation und Archivierung bestimmt sind wurden
- 2017-2021
- 2 Versionen – Produktions- und Testversion
- 65 Archive (300+ Archivar*innen)
- 470 Provenienzstellen
- Verlauf des Verfahrens mit Ergebnissen
- Hochladen als ZIP durch Provenienzstelle mit SIP-Validierung
- Provenienzstelle sieht Verlauf des Verfahrens mit Ergebnissen
- Modernisierung 2021 – Bewertungsangebot A4 mit gesamter SIF Zahl; SIP-Grösse 4 GB zum 100 GB erhöht; mögliche off-line Bewertung mit Export und Import von Excel mit Angaben

Jahr	Bewertung	Übernahme
2015	6	1
2016	8	4
2017	53	15
2018	111	37
2019	397	72
2020	570	85
	1145	214





Nationales Archivportal – Workflow (un)strukturierte Daten



Standardisierung und Automatisierung

- Der SIP-Validator dient am NArP zur technischen und inhaltlichen Kontrolle – erreichbar auch im Web als Dienst (<https://validatorsip.nacr.cz>) mit einer Regelnbeschreibung
 - SIP-Validierung beträgt mehr als 140 Kontrollen
 - Schrittweise Einführung des SIP-Validators wirkte positiv auf die Implementierung der Funktionalität durch die Lieferanten aus
 - Produktions- und Entwicklungsversionen für Lieferantenfeedback verfügbar
 - Die Gültigkeit wird anhand des XML-Schemas, der METS-Struktur (CHECKSUM, SIZE usw.) und des Inhalts (Registrierungsplan, Aufbewahrungsfrist usw.) überprüft.
 - Identifiziert das Format mithilfe des DROID und des PRONOM-Registers (z. B. MIMETYPE)
- Maschinenlesbarkeit für DMS/VBS

```
<?xml version="1.0" encoding="UTF-8" standalone="true"?>
<Sitemap xmlns:loc="http://www.nacr.cz/nsess/v3 http://www.nacr.cz/nsess/v3/nsess-DA.xsd"
  datumvytvoreni="2021-01-26T09:25:47.876+01:00" identifikatorSkartacnihoListu="306 2601/2021-0913"
  identifikatorPravniho="Česká tisková kancelář: 1000196051" identifikatorArchivu="10000010"
  ID="ESK1014_20210126092547_F" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://www.nacr.cz/nsess/v3">
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005YF47</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUP003M6CO</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005YF01</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005VP39</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUP003K7LY</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005VPK4</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005SL6X</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX0059K3</Identifikator>
    <Operace>vybrat za archiválii</Operace>
  </EntitaSeznamu>
```



```
<?xml version="1.0" encoding="UTF-8"?>
<Sitemap datumVytvoreni="2021-01-26T09:32:07.383+01:00" identifikatorPravniho="3002"
  identifikatorSkartacnihoListu="1014" identifikatorPravniho="n000196051" identifikatorArchivu="10000010" ID="3002"
  xmlns="http://www.nacr.cz/nsess/v3">
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005YF47</Identifikator>
    <IdentifikatorDA zdroj="CZ_NDA">ee60cb00-1bb0-430a-9223-12f402a83674</IdentifikatorDA>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUP003M6CO</Identifikator>
    <IdentifikatorDA zdroj="CZ_NDA">ee60cb00-1bb0-430a-9223-12f402a83674</IdentifikatorDA>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005YF01</Identifikator>
    <IdentifikatorDA zdroj="CZ_NDA">ee60cb00-1bb0-430a-9223-12f402a83674</IdentifikatorDA>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005VP39</Identifikator>
    <IdentifikatorDA zdroj="CZ_NDA">845c82b3-3ae4-47a9-84b2-bb7c8ecceb16</IdentifikatorDA>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUP003K7LY</Identifikator>
    <IdentifikatorDA zdroj="CZ_NDA">845c82b3-3ae4-47a9-84b2-bb7c8ecceb16</IdentifikatorDA>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005VPK4</Identifikator>
    <IdentifikatorDA zdroj="CZ_NDA">845c82b3-3ae4-47a9-84b2-bb7c8ecceb16</IdentifikatorDA>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX005SL6X</Identifikator>
    <IdentifikatorDA zdroj="CZ_NDA">845c82b3-3ae4-47a9-84b2-bb7c8ecceb16</IdentifikatorDA>
  </EntitaSeznamu>
  <EntitaSeznamu>
    <Identifikator zdroj="GENIS">CRDUX0059K3</Identifikator>
    <IdentifikatorDA zdroj="CZ_NDA">845c82b3-3ae4-47a9-84b2-bb7c8ecceb16</IdentifikatorDA>
  </EntitaSeznamu>
```

Implementierung und Erfahrungen

- Langsame Änderungen von DMS / VBS durch Lieferanten - etwa 50% haben den Übergang zur vollständigen elektronischen Aktenführung nicht geschafft, besonders bei Aussonderung und SIP-Export (Novellierung des Archivgesetzes 2021 etabliert nach 10 Jahren die Zertifizierung der DMS/VBS)
- Ausbildung für die Archivar*innen seit 2016 mit Schulungen (in Prag und in Regionen von Methodiker der Gebietsarchiven), Materialien im LMS Moodle
- Für Provenienzstellen und Lieferanten Seminare, Informationsblatt (2x jährlich) mit Mustern und Anweisungen (<https://www.nacr.cz/verejnost/2-predarchivni-pece/verejnopravni-puvodci/informacni-list>)
- Modernisierung 2021 – nur on-line Seminare und Videos
- Entwicklungspläne und Hindernisse:
 - Web-services zur automatisierten Kommunikation mit DMS/VBS
 - Hoheheit zur Genehmigung von Akten- und Aussoderungsplänen durch Archiven – ohne dies anspruchsvolle Bewertung
 - ständige Entwicklung z.B. neue Gebiete mit europäischen SIP Standards in Vorbereitung (elektronische Patientenakte, SIARD usw.)





National Archives
Czech Republic

zbysek.stodulka@nacr.cz