

Návrh digitálního archivu

Archiv obecních novin

1. Úvod

Obec Fiktivní vydává vlastní obecní noviny, které vychází každý měsíc. Populace obce čítá 1 000 obyvatel, jenž jsou potencionálními čtenáři novin. Cílem tohoto návrhu je naplánovat digitální archiv pro správu, zhodnocení a důvěryhodné poskytnutí novin pro současné i budoucí čtenáře.

1.1 Popis instituce

Vydavatelem novin je obec Fiktivní. Fiktivní jsou obcí ležící v kraji Vysočina. Obec má své vlastní noviny, které komentují lokální události. Organizační struktura vedení obce sestává ze Zastupitelstva obce, Kontrolního výboru a Finančního výboru. Noviny mají svoji redakční radu, pozici korektura a grafika.

Vize

Poskytovat pravdivé, včasné a použitelné informace o lokálním i zasahujícím globálním dění občanům obce.

Mise

Účelem obecních novin je zprostředkovávat lokální i globální zprávy svým občanům v takové podobě, aby byl informace občany použitelné pro jejich chápání a zapojení v roli občana ve veřejném dění obce. Noviny musí podporovat rovný přístup, který povede k demokratickým hodnotám.

Hodnoty

Poskytovat pravdivé informace.
Poskytovat informace v čas.
Poskytovat informace v použitelné podobě.
Zapojovat do tvorby obsahu novin občany.
Poskytovat prostor pro vyjádření všem bez rozdílu.
Zachovávat obsah novin pro současné a budoucí použití.

1.2 Strategie

Dlouhodobá strategie pro rozvoj novin zahrnuje změnu Současného stavu na stav Očekávaný.

Současný stav

Pro popis stávajícího stavu využijeme [matici NDSA](#). NDSA matice stanovuje pět oblastí zájmu (uložiště, integrita, řízení, metadata a obsah) a čtyři úrovně jejich digitálního zachování. Současný stav

odpovídá první úrovni ochrany u všech oblastí, tedy **organizace zná svá data**.

Cíle

Primární

Poskytovat informace občanům obce tak, aby byli v použitelné podobě a zachovaly si autenticitu i pro budoucí využití.

Sekundární

Poskytovat informace občanům i vně obce.

Očekávaný stav

Očekávaným stavem je dosažení čtvrtě úrovně ochrany dat dle matice NDSA, tedy **organizace opravuje svá data**.

2. Implementace archivu

Cílem této části textu je poskytnout šablonu pro naplánování změn, tak aby mohl být Současný stav transformován do Očekávaného stavu. Dokument si klade cíl, být referenčním materiálem, který pomůže při nastavování změn.

2.1 OAIS model

Hlavním referenčním materiálem návrhu archivu je referenční model OAIS. OAIS obsahuje funkční a informační model. **Funkční model** zahrnuje následující komponenty:

- Funkční celek příjem
- Funkční celek správa dat
- Funkční celek archivního uložení
- Funkční celek zpřístupnění
- Funkční celek správa archivu
- Funkční celek plánování uchování dat

Vně archivu náleží Tvůrci, Koncový uživatel a Vedení

Informační model stanovuje tři informační balíčky:

- Submission Information Package (SIP)
- Archival Information Package (AIP)
- Dissemination Information Package (DIP)

Zavazujeme se tímto modelem při návrhu řídit a využít jeho rozdělení komponent a struktur při plánování naší změny.

2.2 Stakeholderi

Pro hladký chod archivu je nezbytné si stanovit jeho lidské okolí. Určit osoby odpovědné za chod archivu a jeho využívání. Popsat zainteresované okolí je cílem následující části návrhu.

2.2.1 Určení odpovědnosti

Za chod archivu je zodpovědná redakční rada novin společně s určenou osobou digitálního kurátora.

Financování archivu je zajištěno finančním výborem obce.

Kontrolní orgán obce řídí transparentnost běhu digitálního archivu.

Rozhodující slovo v záležitostech fungování archivu má zastupitelstvo obce s přihlédnutím na zprávy kontrolního a finančního výboru.

2.2.2 Určená skupina

Určenou skupinou archivu jsou v první řadě obyvatelé obce Fiktivní. Archiv se však zavazuje poskytovat své služby komukoli i mimo obec Fiktivní. Z povahy obsahu je však přednostním uživatelem lokální čtenář novin.

2.3 Určení SIP, AIP a DIP

V této části návrhu se zaměříme na tři klíčové informační balíčky dle OAIS modelu. Balíčky budeme popisovat a zařazovat v kontextu funkčních celků OAIS modelu, tak aby došlo k naplnění našich požadavků na určení struktury archivu odpovídající struktuře referenčního modelu OAIS.

2.3.1 Příjem

Příjem dat (ingest) bude probíhat pravidelně každý měsíc s novým číslem novin. Starší tištěná vydání budou digitalizována. V terminologii OAIS se jedná o balíček SIP. Tvůrcem dat je redakční rada novin a dopisovatelé do novin (tedy jednotliví autoři článků).

2.3.1.1 Typ dat

Analogové (digitalizace)

Tištěné starší vydání budou digitalizována. Výstupním formátem bude PDF/A soubor a ODT soubor.

Digitální

Digital-born vydání budou přijímána ve formátu PDF/A a ODT.

2.3.1.2 Formát

Doporučené formáty na příjmu jsou ODT a PDF/A.

2.3.1.3 Metadata

Doporučená metadata na příjmu jsou popisná metadata ve formátu Dublin Core, vysvětlující informace a informace o původu. Metadata budou zabalena ve formátu METS.

2.3.2 Zpracování

Funkční celek správy dat zajišťuje správu dat v archivu. Tato činnost zpracování je též spojena s funkčním celkem archivního uložení. Zpracovávaná dat jsou balíček AIP. Uložiště bude vytvořeno pod open-source systémem Islandora založeného na Fedora Commons.

2.3.2.1 Popisná metadata

Odpovědná osoba v archivu (digitální kurátor) zkontroluje metadata a vysvětlující informace a v případě nedostatečných metadat či vysvětlujících informace doplní chybějící sama či si vyžádá od tvůrce dat doplnění.

2.3.2.2 Vysvětlující informace

Vysvětlující informace jsou součástí balíčku SIP. Tyto informace mohou být dle potřeby obohaceny kurátorem archivu.

2.3.2.3 Informace o uchování

Dle OAIS modelu zařazujeme mezi informace o uchování informace o identifikátorech, informace o souvislostech, informace o původu, informace o neporušenosti a přístupová práva. Tyto informace budou přiřazeny k datovému objektu digitálním kurátorem ve spolupráci s redakční radou.

2.3.2.4 Data

Daty jsou samotná vydání novin ve zmiňovaném formátu ODT, PDF/A a EPUB.

2.3.3 Zprostředkování

Funkční celek zpřístupnění dat bude uskutečněn pomocí webové infrastruktury obce, kde bude vyčleněno místo pro výstupy archivu tzv. DIP. Balíček DIP bude zahrnovat soubory ve formátu PDF a EPUB, které budou přes webové rozhraní přístupné uživateli.

Funkční celek správa archivu bude realizován v systému Islandora. Odpovídající funkční celek bude přístupný vedení obce (tedy zastupitelstvu, kontrolnímu a finančnímu výboru).

Funkční celek plánování uchování dat bude též součástí systému Islandora a bude přístupný digitálním kurátorovi a redakci novin.

2.4 Strategie dlouhodobé ochrany

Strategie dlouhodobé ochrany dat archivu bude zvolena následovně:

2.4.1 Formátová migrace

Bude prováděna formátová migrace dle aktuálního technologického okolí. Momentálně je doporučena migrace do formátu ODT, PDF/A a EPUB.

2.4.1 Normalizace

Data archivu budou normalizována do třech formátů ODT, PDF/A a EPUB.

2.4.3 Fyzická migrace

Kromě logické ochrany dat výše, bude archiv zálohovat data bitovým kopírováním na dvě různá místa, která budou od sebe fyzicky vzdálena tak, aby v případě živelné pohromy nedošlo ke zničení obou záloh.

2.5 Nástroje

Tato část implementace archivu si klade cíl zvolit odpovídající vhodné nástroje, jimiž bude předchozí spíše abstraktní plánování naplněno v realitu.

2.5.1 Identifikace formátu

Pro identifikaci formátů bude zvolen nástroj SIEGFRIED.

2.5.2 Validace formátu

Pro validaci formátů bude zvolen nástroj FITS.

2.5.3 Migrace formátů

Pro migraci formátů bude zvolen nástroj UNOCONV. Pomocí tohoto nástroje též budeme normalizovat formáty.

2.5.4 Repozitář

Pro zřízení repozitáře bude použit systém Islandora. Tento systém splňuje požadavky na finanční nenáročnost. Při jeho nasazení bude v případě potřeby využita metodická pomoc.

3. Hodnocení a revize archivu

Posledním oddílem plánu návrhu archivu je oddíl, který definuje, kdo bude vyhodnocovat archiv. Do vyhodnocení se zahrnuje využití archivu uživateli i vnitřní fungování archivu.

3.1 Monitoring změn

Digitální kurátor se zavazuje ke sledování technologických změn v oblasti dlouhodobého uchovávání dat a kurátorských kompetencí.

3.1.1 Technologické změny

Sledováním technologických změn je myšlena především kontrola stávajících metadatových standardů a nástrojů pro manipulaci s daty archivu.

3.1.2 Sledované informační zdroje

Kurátor bude sledovat fóra pro digitální kurátory (např. Digital Curation Google Group), odborná periodika (např. International Journal of Digital Curation) a bude-li možné navštěvovat oborové konference (např. OPF Events).

3.2 Vyhodnocování archivu

Vnitřní hodnocení archivu vykonává kontrolní orgán obce a výsledek předává zřizovateli (zastupitelstvu obce). Vnější (uživatelské) hodnocení archivu předkládá digitální kurátor zřizovateli.

3.2.1 Vnitřní kontrola

Vnitřní kontrola se zabývá především kontroly čerpání financí a etickými a právními aspekty fungování archivu.

3.2.2 Vnější kontrola

Vnější kontrola zahrnuje vyhodnocení využívání archivu uživateli. Výsledky kontroly slouží k obhajobě existence archivu.

3.3 Certifikace důvěryhodnosti

Cílem archivu je stát se důvěryhodným archivem.

3.4 Metodika PLATTER

Při návrhu archivu se zavazujeme dodržovat metodiku, která vznikla v rámci projektu DigitalPreservationEurope (DPE). Tato metodika stanovuje deset bodů, kterými se máme při návrhu důvěryhodného archivu řídit. Český překlad metodika [je dostupný zde](#).

3.5 Typ certifikátu

Cílem naplnění metodiky PLATTER je dosažení certifikace nestor Seal for Trustworthy Digital Archives. Více informací o certifikaci [naleznete zde](#).