

# projekt\_1840\_Pristup\_k\_projektu\_detailny

## PRÍSTUP K PROJEKTU

Identifikovanie požiadaviek **na technickú časť riešenia.**

### Identifikácia projektu

<b>Povinná osoba</b>	Ministerstvo vnútra Slovenskej republiky
<b>Názov projektu</b>	Digitalizácia a modernizácia elektronických služieb štátnych archívov v európskom kontexte
<b>MetaIS kód</b>	project_1840
<b>Zodpovedná osoba za projekt</b>	Ing. Štefan Mačuga – projektový manažér
<b>Realizátor projektu</b>	Ministerstvo vnútra Slovenskej republiky
<b>Vlastník projektu</b>	Ing. Adrián Jenčo – generálny riaditeľ sekcie verejnej správy MV SR Ing. Igor Sibert – generálny riaditeľ sekcie informatiky, telekomunikácií a bezpečnosti MV SR

### Schvaľovanie dokumentu

Položka	Meno a priezvisko	Organizácia	Pracovná pozícia	Dátum	Podpis (alebo elektronický súhlas)
Vypracoval	Ing. Štefan Mačuga	MV SR	Projektový manažér	1.6.2023	

Obsah

- 1. POPIS ZMIEN DOKUMENTU
  - 1.1 História zmien
- 2. ÚČEL DOKUMENTU
  - 2.1 Použité skratky
  - 2.2 Ďalšie definície pojmov, skratky a terminológia
- 3. POPIS NAVRHOVANÉHO RIEŠENIA
- 4. ARCHITEKTÚRA RIEŠENIA PROJEKTU
  - 4.1 AS IS Architektúra
  - 4.2 TO BE Architektúra
  - 4.3 Biznis vrstva
    - 4.4. Aplikačná vrstva
      - 4.4.1 Správa evidencií
      - 4.4.2 ArchivLog
      - 4.4.3 Subsystém Digitalizácia
      - 4.4.4 Migrácia
      - 4.4.5 Rozsah informačných systémov
      - 4.4.6 Využívanie nadrezortných centrálnych blokov (SaaS)
    - 4.5 Dátová vrstva
      - 4.5.1 Údaje v správe organizácie
      - 4.5.2 Dátový rozsah projektu
      - 4.5.3 Kvalita a čistenie údajov
      - 4.5.4 Role a predbežné personálne zabezpečenie pri riadení dátovej kvality
      - 4.5.5 Referenčné údaje
      - 4.5.6 Otvorené údaje
      - 4.5.7 Analytické údaje
      - 4.5.8 Moje údaje
      - 4.5.9 Prehľad jednotlivých kategórií údajov
      - 4.5.10 Využívanie údajov
      - 4.5.11 Správa údajov
      - 4.5.12 Riadenie údajov
  - 4.6 Technologická vrstva
    - 4.6.1 Prehľad technologického stavu.
    - 4.6.2 Architektúra infraštruktúry
    - 4.6.3 Bezpečnosť a Riadenie
    - 4.6.4 Špecifické technologické potreby projektu
    - 4.6.5 Požiadavky na výkonnostné parametre, kapacitné požiadavky
    - 4.6.6 Návrh riešenia technologickej architektúry
    - 4.6.7 Využívanie služieb z katalógu služieb vládneho cloudu
    - 4.6.8 Jazyková lokalizácia
  - 4.7 Bezpečnostná architektúra
    - 4.7.1 Riziká pri súčasnom stave online zverejňovania
- 5. ZÁVISLOSTI NA OSTATNÉ ISVS / PROJEKTY
- 6. ZDROJOVÉ KÓDY
- 7. PREVÁDZKA A ÚDRŽBA
  - 7.1 Požiadavky na podporu, údržbu a rozvoj
    - 7.1.1 Úroveň podpory používateľov
    - 7.1.2 SLA parametre
  - 7.2 Požadovaná dostupnosť IS
- 8. POŽIADAVKY NA PERSONÁL
- 8.1 Pracovné náplne
- 9. IMPLEMENTÁCIA A PREBERANIE VÝSTUPOV PROJEKTU
- 10. PRÍLOHY

# 1. POPIS ZMIEN DOKUMENTU

## História zmien

Verzia	Dátum	Zmeny	Meno
--------	-------	-------	------

1.0	24.8.2022	Vypracovanie dokumentu	Štefan Mačuga

## 2. ÚČEL DOKUMENTU

V súlade s Vyhláškou 85/2020 Z.z. o riadení projektov je dokument *Pristup k projektu* pre iniciačnú fázu určený na rozpracovanie detailných informácií prípravy projektu z pohľadu budúceho stavu a navrhovaného riešenia.

Dokument popisuje v zmysle vyššie uvedenej vyhlášky riešenie projektu v oblasti požiadaviek na:

1. architektúru riešenia – biznis vrstva, aplikačná vrstva, technologická vrstva;
2. dátový model, dátové konverzie a migrácie;
3. použité dátové centrá a prípadne zdôvodnenie ich použitia;
4. požiadavky na HW, SW a licencie;
5. bezpečnosť riešenia;
6. testovanie a akceptačné kritériá;
7. prevádzku riešenia, výkonnostné parametre, dostupnosť a zálohovanie;
8. integrácie, rozhrania a spoločné komponenty;
9. dokumentáciu a školenia.

### 2.1 Použité skratky

ID	SKRATKA	POPIS
1.	MDV SR	Ministerstvo dopravy a výstavby Slovenskej republiky
2.	IS	Informačný systém
3.	Strategický plán	Strategický plán rozvoja
4.	EU	Európska únia
5.	AI	Umelá inteligencia
6.	API	Aplikačné rozhranie
7.	RTVS	Rozhlas a televízia Slovenska
8.	MV SR	Ministerstvo vnútra Slovenskej republiky
9.	IKT	Informačno-komunikačné technológie
10.	IS	Informačný systém
11.	ISEA	Elektronický archív ministerstva vnútra SR
12.	RV	Riadiaci výbor projektu
13.	VO	Verejné obstarávanie
14.	PM	Projektový manažér
15.	IT	Informačné technológie
16.	DC MV SR	Dátové centrum MV SR
17.	SITB MV SR	Sekcia informatiky, telekomunikácií a bezpečnosti MV SR
18.	SVS MV SR	Sekcia verejnej správy MV SR
19.	ISVS	Informačný systém verejnej správy

### 2.2 Ďalšie definície pojmov, skratky a terminológia

**Digitálny priestor** – informačné a komunikačné technológie vrátane internetu, mobilných a pridružených technológií a zariadení, ako aj digitálne siete, databázy, obsahy a služby.

**Digitalizácia** – proces naskenovania, vyextrahovania textu zo skenu a online publikovania kópií archívnych dokumentov, ktoré sú uložené v archíve v listinnej forme.

**DMS** – informačný systém (alebo súbor programov) pre správu dokumentov používaný na ukladanie a prezentovanie elektronických dokumentov online.

**OCR** – metóda umožňujúca preklad obrazu tlačенých alebo písaných znakov zo skenu do textovej, editovateľnej formy (Optical Character Recognition).

**ArchivLog** – modul aplikácie ISEA pre evidenciu a manažment procesov týkajúcich sa životného cyklu archívnych dokumentov v závislosti od sledovanej úrovne popisu v rámci archívneho fondu.

**Bežný meter (bm)** – základná merná jednotka vyjadrujúca množstvo vedľa seba zvisle postavených archívnych škatúl s archívnymi dokumentmi na polici archívneho regálu. Spravidla 1 bm tvorí 8 archívnych škatúl, pričom množstvo listov v 1 archívnej škatuli sa pohybuje od 500 do 1100 v závislosti od veľkosti a typu nosiča.

**1 km dokumentov** – reprezentuje 1000 bm (bežných metrov), resp. 8000 archívnych škatúl s celkovým obsahom 4,000,000-8,800,000 listov.

**Elektronický archív ministerstva vnútra SR (ISEA)** – informačný systém, vybudovaný projektom realizovaným v rokoch 2013-2017, v dvoch fázach: prvej (Kód ITMS: 21110120036) a druhej (Kód ITMS: 311071A268).

### 3. POPIS NAVRHOVANÉHO RIEŠENIA

Predkladaný projekt je integrálnym pokračovaním predchádzajúcich etáp budovania Elektronického archívu MV SR: „Elektronický archív Ministerstva vnútra SR (isvs\_202)“ – ďalej len „ISEA“, projekt Elektronický archív MV SR (projekt\_46) – 1. fáza OPIS a 2. fázu projektu.

Prvé dve etapy pokrývali pilotné ukladanie a dlhodobú archiváciu zdigitalizovaných archívnych dokumentov, vytvorenie predpokladov na ďalšie digitálne napĺňanie archívu a systém elektronických služieb digitálneho archívu pre občanov.

Predkladaný projekt predpokladá v priamej nadväznosti masovú digitalizáciu existujúcich archívnych dokumentov a vytvorenie systému automatického popisovania metadát novo-prichádzajúcich dokumentov. Novovytvorený IS bude využívať existujúce úložné a procesorové kapacity projektu ISEA vrátane existujúcich dátových štruktúr a bude prirodzeným spôsobom dopĺňať už pokryté archivačné procesy. Systém bude napojený na modul dlhodobého uchovávanía archívnych dokumentov, modul vyhľadávania, modul agendového systému a evidencií v ISEA a bude tiež umožňovať sprístupňovanie archívnych dokumentov prostredníctvom elektronických služieb pre občanov. Novou funkcionalitou pre zabezpečenie riadenia modernizácie archívnych procesov je monitorovanie digitalizačného procesu.

**Ochrana archívnych dokumentov** sa čoraz viac spája s ich digitalizáciou, pričom jedným z pozitív procesov digitalizácie je skutočnosť, že sa s fyzickými archívnymi dokumentami po ich sprístupnení ďalej nemanipuluje. Do procesu digitalizácie sa preferuje prednostne zaradiť archívne dokumenty, ktoré budú mať v najbližšom období preukázateľne problémy so životnosťou a prečítaním ich obsahu (na základe analýz morfológie papiera a ostatných súčastí dokumentu), ale stále vykazujú kompaktnosť s minimálnou mierou viditeľného poškodenia.

V súlade so zákonom o archívoch a registratúrach má štátny archív predkladať na štúdium originál archívneho dokumentu iba výnimočne. Kópiu archívneho dokumentu predkladá vždy, ak ide o archívny dokument, ktorý vznikol pred rokom 1526, obsahuje informácie, ku ktorým je obmedzený prístup alebo by pri štúdiu mohlo dôjsť k jeho poškodeniu, resp. možnosti jeho poškodenia. Štátne archívy spravujú 39 013 stredovekých listín.

Priority výberu z hľadiska historickej hodnoty sú vzácne tlače a rukopisy s vysokou kultúrnou a historickou hodnotou, ktoré sú vystavené riziku fyzického poškodenia súvisiaceho s ich vekom, materiálovým zložením či spôsobom využívania a vyžadujú špeciálne zaobchádzanie. V prípade rukopisov je potrebné aplikovať tiež technologické a vedecké inovácie v oblasti rozpoznávania textu v rukopisoch, založené na využívaní prvkov umelej inteligencie. Celkovo projekt pokrýva:

1. výrazné navýšenie úložných a spracovateľských kapacít na vyriešenie archívneho dlhu;
2. upgrade ISEA o systém automatizovaného popisovania metadát archívnych dokumentov, manažmentu dát a systém riadenia digitalizácie,
3. program digitalizácie vybraných archívnych dokumentov (50 km) na nasledujúce roky a jeho nákladovú a personálnu udržateľnosť.

Realizácia projektu je plánovaná v dvoch inkrementoch:

- Inkrement 1 (pilot) – vytvorenie podmienok pre aktívnu digitalizáciu, rekuperácia a rozvoj ISEA, overenie prínosnosti projektu zriadením prototypného digitalizačného pracoviska
- Inkrement 2 (plné riešenie) – dobudovanie kapacít pre masovú digitalizáciu, ukladanie a sprístupňovanie archívnych dokumentov, optimalizácia procesov a prechod na celoštátnu prevádzku riešenia.

Projekt je významný z hľadiska postupného prechodu na archivovanie elektronicky vzniknutých dokumentov (e-born), pričom reflektuje programové vyhlásenie vlády na roky 2021 – 2024, časť

Efektívna verejná správa, s. 27, [https://www.nrsr.sk/web/Dynamic/DocumentPreview.aspx?DocID=477513#30\\_0](https://www.nrsr.sk/web/Dynamic/DocumentPreview.aspx?DocID=477513#30_0)

v kontexte s UZNESENÍM VLÁDY SLOVENSKEJ REPUBLIKY č. 456 z 13. júla 2022 k Národnej koncepcii rozvoja štátnych archívov s výhľadom do roku 2028, kde sú okrem samotnej digitalizácie a celkovej modernizácie štátnych archívov rozpracované tiež nasledovné strategické priority:

1. elektronizácia štátnych archívov a nové moderné technológie;
2. efektívnosť prístupu k archívnym dokumentom;
3. systém sledovania produktivity;
4. záchrana archívnych dokumentov pred degradáciou;
5. ochrana informácií obsiahnutých v archívných dokumentoch pred ich zničením (napr. vojnový konflikt, prírodné katastrofy a pod.).

## 4. ARCHITEKTÚRA RIEŠENIA PROJEKTU

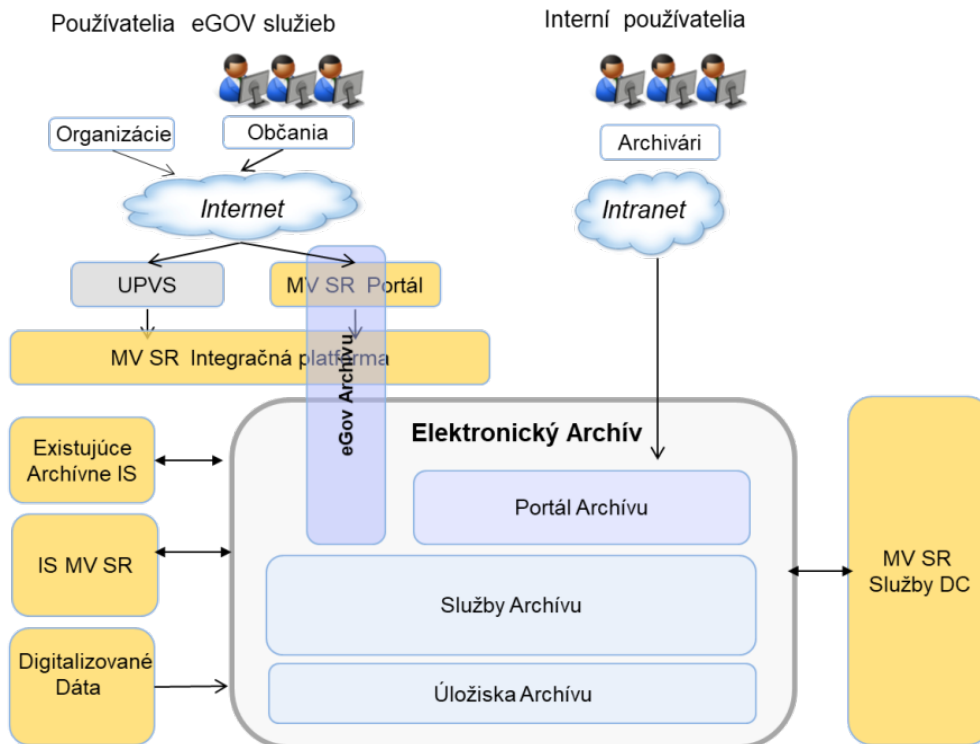
Hlavným cieľom projektu je vytvoriť podmienky pre komplexnú digitalizáciu novo prichádzajúceho archívneho obsahu a pilotnú digitalizáciu historického obsahu v podobe digitalizačných pracovísk a inteligentného informačného systému. Riešenie bude zakomponované do existujúceho informačného systému Elektronického archívu ministerstva vnútra SR (ISEA) a bude vybudované na otvorenej platforme pre orchestráciu virtuálnych kontajnerov a serverov.

### 4.1 AS IS Architektúra

V súčasnosti je v prevádzke archív MV SR pod názvom Elektronický archív MV SR, ktorého integrálnou súčasťou boli aj predchádzajúce už zmienené pilotné projekty Štátneho archívu:

*Elektronický archív ministerstva vnútra SR (ISEA) – informačný systém, vybudovaný projektom realizovaným v rokoch 2013-2017, v dvoch fázach: prvej (Kód ITMS: 21110120036) a druhej (Kód ITMS: 311071A268).*

Systémová architektúra súčasného Elektronického archívu MV SR je na nasledujúcom diagrame:

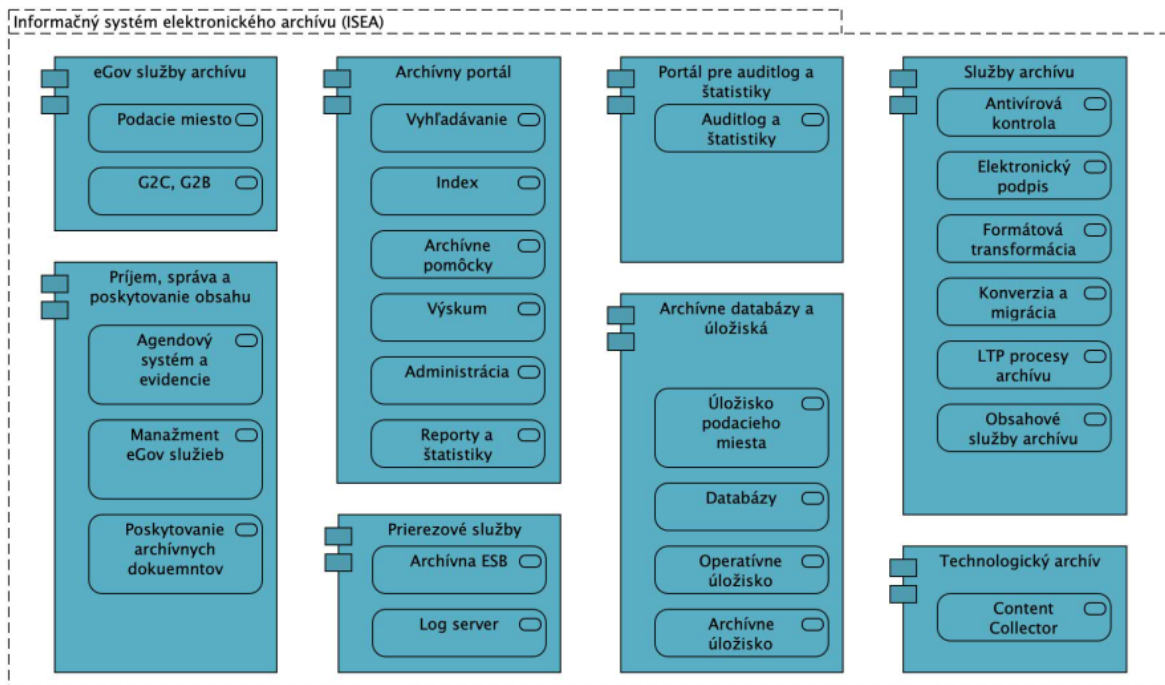


*Systémová architektúra AS-IS EA MV SR*

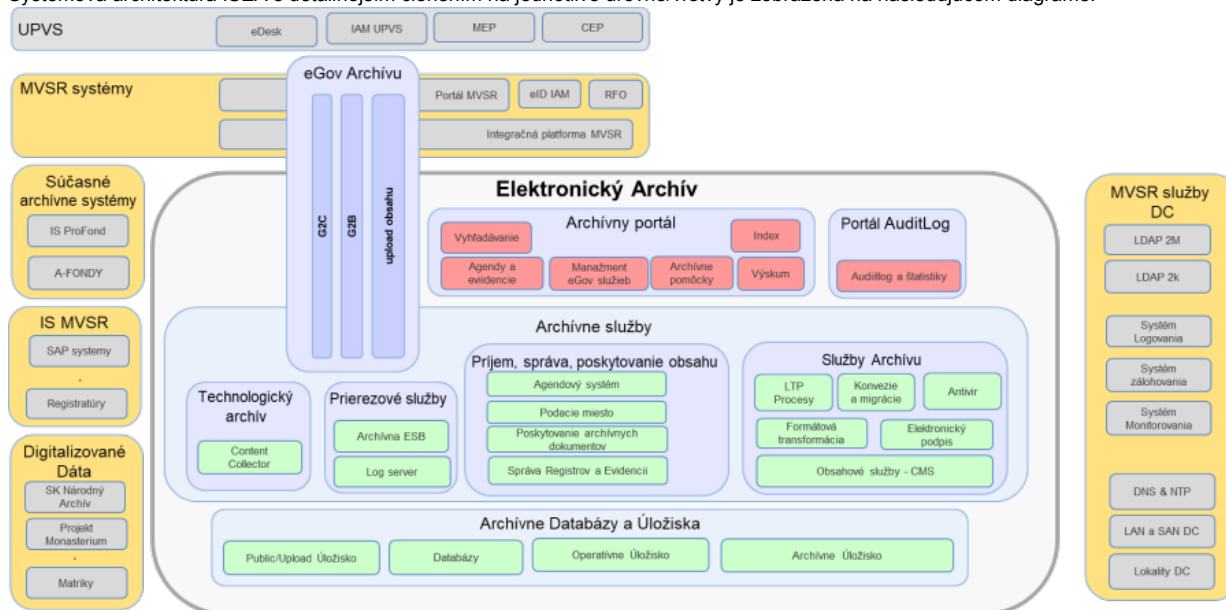
Súčasný systém je implementovaný v DC MV SR a využíva (je integrovaný) aplikačné, integračné a prevádzkové služby DC MV SR. Migrácia „Digitalizované Dáta“ bola realizovaná ako jednorazová špecifická migračná činnosť, ktorá overila základný prístup a postupy pri tvorbe a naplňaní ISEA obsahom.

Integrácia nových nástrojov a prvkov strojového učenia do informačného systému ISEA v rámci popisovaného projektu zmení doterajšie pracovné postupy, ako aj efektívnosť práce archivárov pri tvorbe a správe archívneho obsahu. Architektúra nových modulov digitalizácie rešpektuje podmienky zadania a AS IS stav. Jedná sa teda o doplnenie resp. rozšírenie systému integráciou nových technologických prvkov a rozšírenie kapacít existujúceho systému.

Detailnejší stav súčasnej architektúry a funkčnosti Elektronického archívu ministerstva vnútra SR je znázornený na nasledujúcich diagramoch systémovej aj aplikačnej architektúry ISEA na úrovni komponentov a funkčných modulov:



Systémová architektúra ISEA s detailnejším členením na jednotlivé úrovne/vrstvy je zobrazená na nasledujúcom diagrame:



ISEA sa v súčasnosti skladá z viacerých modulov, v rámci ktorých sú vytvorené aplikačné komponenty. Moduly sú členené do prezentačnej, aplikačnej a persistentnej vrstvy. Dôležitou súčasťou systému je modul pre integračné a logovacie služby.

## 4.2 TO BE Architektúra

Projekt „Digitalizácia a modernizácia elektronických služieb štátnych archívov v európskom kontexte“ nadväzuje na predchádzajúci projekt v zmysle popisu v kapitole 4.1 z pohľadu efektívneho využitia existujúcich zdrojov, avšak nie je jeho integračnou súčasťou. Jedná sa o samostatný projekt a svojim rozsahom pokrýva nasledovné:

Aktivita	Nákup	Vývoj	Integrácia
Vytvoriť podmienky pre komplexnú digitalizáciu novo-prichádzajúceho archívneho obsahu a pilotnú digitalizáciu historického obsahu.	X	X	X
Dobudovať funkcionality a kapacity ISEA a zabezpečiť jeho podporu.	X	X	-
Vybudovať špecializované digitalizačné pracoviská vybavené knižnicami, prietahovými a veľkoformátovými skenermi a skenermi na mikrofilmy vrátane digitalizačného workflow softvéru pre spracovanie digitálnych kópií.	X	-	X

Zaviesť ArchivLog pre evidenciu životného cyklu archívnej jednotky.	-	X	X
Sprístupniť archívne dokumenty z pilotnej prevádzky digitalizačných pracovísk v ISEA.	-	-	-
Implementovať systém automatizácie popisovania metadát s využitím technológie strojového učenia.	X	X	X
Implementovať systém monitorovania procesu digitalizácie a sledovania produktivity práce digitalizačných pracovísk.	X	X	X
Zabezpečiť migráciu všetkých zdigitalizovaných archívnych dokumentov z lokálnych úložísk na centrálné dátové úložisko, zabezpečiť pravidelné zálohovanie archívu s dôrazom na kritické dáta.	-	-	X
Zabezpečiť dostatočnú kapacitu na uloženie digitálnych kópií archívnych dokumentov vyhotovených systematickou digitalizáciou.	X	-	X
Dovybaviť bádateľne o príručné skenery a počítače za účelom zníženia zaťaženia pracovníkov bádateľne.	X	-	-
Oboznámiť verejnosť s prínosmi a fungovaním elektronického archívu, so zameraním na perspektívnych užívateľov.	-	-	-

Ďalej sa zaoberáme výhradne TO BE architektúrou.

## 4.3 Biznis vrstva

Biznis vrstva architektúry bola vytvorená vo väzbe na potreby používateľov pre vytváranie, správu a využívanie obsahu archívu pre riešenie životných situácií.

Projekt sleduje nasledujúce oblasti:

1. komplexnosť informácií pre občanov o archívnom dedičstve uloženom v štátnych archívoch;
2. zavedenie systematickej digitalizácie archívneho dedičstva prostredníctvom vybudovania digitalizačných pracovísk;
3. zavedenie systematického manažmentu údajov pre ISEA vrátane nastavenia príslušných procesov a metodík pre správu celého životného cyklu údajov;
4. úpravu organizačnej štruktúry smerom k vytvoreniu role správcu digitalizácie.

### Funkčné požiadavky podľa oblastí a zaradenia do modulov sa nachádzajú v katalógu požiadaviek, ktorý je súčasťou CBA.

Riešiteľský tím zvažoval použitie cloudových zdrojov pre projekt digitalizácie archívnych dokumentov. K návrhu predkladanej infraštruktúrnej architektúry ho viedli nasledovné technické, bezpečnostné a organizačné obmedzenia (zdroj: stanovisko SITB, MV SR):

1. V aktuálnej implementácii government cloudu nie je k dispozícii prepojenie na interné systémy MV SR. Vytvorenie tohto prepojenia bolo zamietnuté v iniciálnej fáze budovania government cloudu z organizačných a bezpečnostných dôvodov. Prepojenie týchto v súčasnosti navzájom izolovaných prostredí by znamenalo realizáciu samostatného projektu s komplexnou analýzou technických, bezpečnostných a organizačných dopadov s požiadavkami na dodatočné zdroje.
2. Prípadná komplexná relokácia infraštruktúrneho prostredia ISEA do government cloudu vytvára požiadavky na špecifické zdroje a integrácie cieľového prostredia. Ide o:
  - a. Dostupnosť dlhodobého zdieľaného úložiska v kapacite 1PB
  - b. Pásková knižnica s kapacitou 5PB
  - c. Integrácie na existujúce interné systémy MV SR – minimálne REGOB/RFO, adresárové služby používateľov, zálohovací systém, a pod.

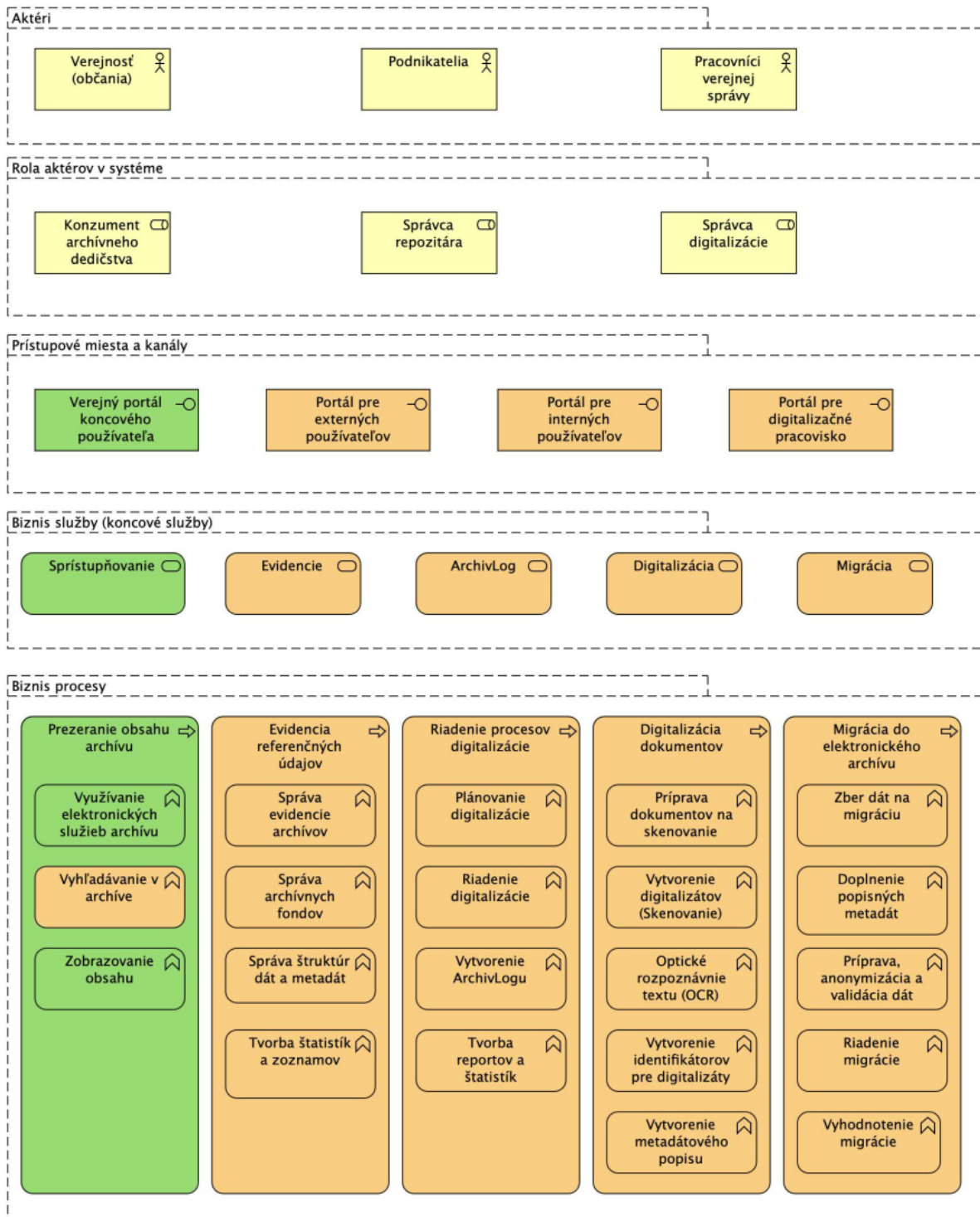
Žiadna z týchto požadovaných funkcionalít v prostredí government cloudu nie je k dispozícii.

1. Pre charakter ukladaných archívnych dokumentov nie je odporúčané ich paušálne spracovanie v cloudovom prostredí vzhľadom na ich citlivosť. Informačný systém Elektronický archív spracováva množstvo osobných údajov a je považovaný za kritický prvok infraštruktúry štátu, čomu zodpovedalo aj jeho architektonické umiestnenie v rámci interných systémov MV SR a z čoho vyplýva aj súčasný návrh, ktorý toto umiestnenie nemení.
2. Okrem uvedených je ďalším zásadným dôvodom potreba portácie aplikácie ISEA do systémového prostredia vládneho cloudu. To pri komplexnosti integrácie na ďalšie interné systémy MV SR nepripadá do úvahy vzhľadom na riziko prípadnej nefunkčnosti kritických systémov štátu (napr. REGOB so svojimi subsystémami akými sú dokladové agendy a ďalšie.) Samotná portácia ISEA do vládneho cloudu by mala väčší rozsah ako tento projekt, ktorý by tak výrazne presiahol dostupnú alokáciu. Navyiac tieto vynaložené prostriedky by neprinesli žiadnu novú funkčnosť ani kvalitu služieb.

Riešiteľský tím taktiež zvažoval využitie existujúcich kapacít pre digitalizáciu z pracovísk Ministerstva kultúry SR. Toto riešenie však nie je dostupné, z nasledujúcich dôvodov (zdroj: stanovisko MK SR):

1. Všetky dlhodobé úložiská a digitalizačné technológie MK SR, získané prostredníctvom eurofondov, v rámci národných projektov Operačného programu Informatizácia spoločnosti, prioritná os 2, sa aktuálne plnohodnotne využívajú.
2. Nebola vykonaná obnova technologického vybavenia viacerých digitalizačných pracovísk, čoho následkom je že mnohé zariadenia sú na konci svojej životnosti. Presun týchto technológií taktiež nie je možný, z dôvodu ich previazania na ostatné procesy pracovísk, ako je mechanická očista a ďalšie, čím by sa narušil kontinuálny a logický proces digitalizácie.
3. Štátne archívy pre povahu a formáty svojich dokumentov potrebujú odlišné typy skenovacích zariadení, menovite najmä prietahové skenery na voľné listy, pričom MK SR disponuje primárne veľkoplošnými skenermi na skenovanie výtvarných diel.

Na nasledujúcom obrázku je náhľad biznis architektúry :



Zelená časť označuje rozšírenie, resp. doplnenie existujúcich biznis procesov a služieb ISEA, ostatné sú nové biznis procesy a služby.

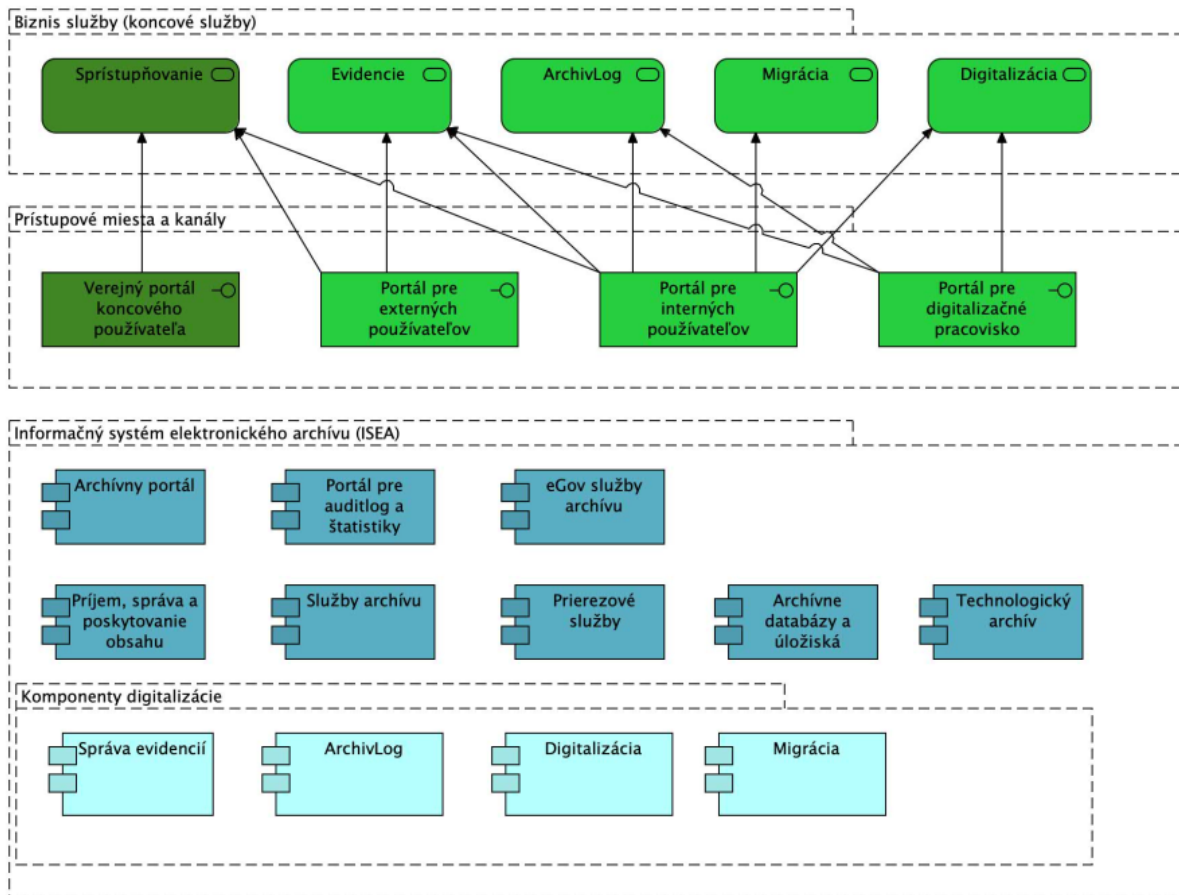
#### 4.4. Aplikačná vrstva

Východiskom je existujúci informačný systém ISEA. Na aplikačnej úrovni je potrebné doplniť funkcionality ISEA s cieľom:

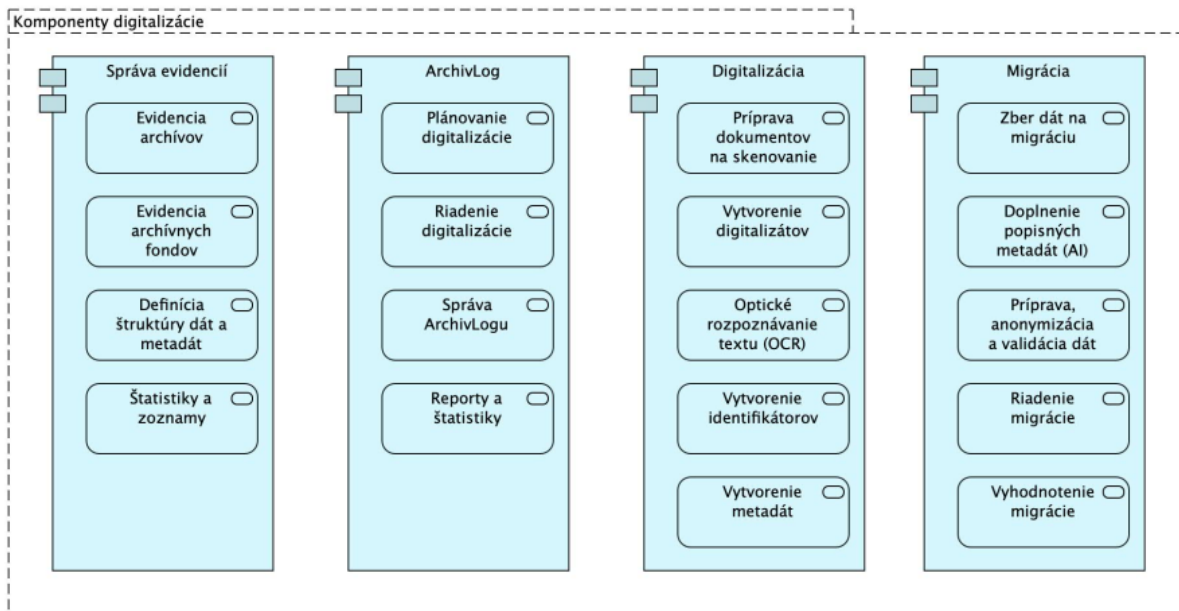


1. implementácie subsystému automatizácie tvorby metadát pre zvýšenie efektivity digitalizačného procesu;
2. implementácie subsystému na sledovanie produktivity práce pre zabezpečenie riadenia digitalizačného procesu;
3. nasadenie sofistikovaného vyhľadávania;
4. podpory digitalizačných procesov systematickej digitalizácie archívneho dedičstva s integráciou digitalizačných pracovísk,;
5. zavedenia systematického manažmentu údajov pre ISEA vrátane nastavenia príslušných procesov a metodík pre správu celého životného cyklu údajov;
6. zvýšenia obsahu informácií pre občanov o archívnom dedičstve uloženom v štátnych archívoch.

Na nasledujúcom obrázku je diagram cieľovej architektúry ISEA. Súčasná aplikačná architektúra ISEA je rozšírená o komponenty digitalizácie:



Komponenty digitalizácie obsahujú nové aplikačné moduly (aplikačné služby) zobrazené v diagrame:



Popis nových aplikačných modulov jednotlivých komponentov digitalizácie je v nasledujúcich kapitolách.

#### 4.4.1 Správa evidencií

Komponent zabezpečuje funkcionality na správu centrálnych evidencií Elektronického archívu MV SR. Skladá sa z nasledovných súčastí:

#### **Evidencia archívov**

Nová aplikačná časť Agendového systému a evidencií ISEA bude slúžiť na doplnenie archívov, aktualizáciu údajov a vyhľadávanie informácií o existujúcich archívoch v evidencii archívov.

Do modulu ASE bude doplnené aj nové používateľské rozhranie na správu existujúcich číselníkov, na zavedenie novej hodnoty do číselníka a jeho aktualizáciu v ISEA (klasifikačné schéma, stav spracovania AF, prístup k AF, zloženie AF, archívne pomôcky (druhy), druhy poškodenia, druh a formát elektronických archívnych dokumentov).

#### **Evidencia archívnych fondov**

Nový aplikačný komponent poskytne funkcionality na vytváranie a definovanie existujúcich a nových archívnych fondov a archívnych zbierok, vrátane správy lokačného plánu uloženia archívnych fondov v archívoch.

Definícia archívneho fondu pozostáva z definície jeho všeobecných a špecializovaných atribútov. Všeobecné atribúty sú spoločné pre každý archívny fond a sú to vlastne základné metadáta, ktorými je popísaný každý archívny fond. Špecializované atribúty sú definované konkrétne pre jednotlivé kategórie archívnych dokumentov.

Súčasťou Evidencie archívnych fondov bude aj sada funkcionalít na delimitáciu archívneho fondu do iného archívu a na vytváranie a správu plánu digitalizácie.

#### **Definícia štruktúry dát a metadát**

Funkcionality na vytvorenie a aktualizáciu štruktúry dát a metadát archívnych dokumentov budú primárne slúžiť na definíciu nasledujúcich oblastí:

- na definíciu štruktúry základných metadát pre archívny fond;
- na definíciu štruktúry rozšírených, špecializovaných metadát archívneho fondu;
- na definíciu štruktúry metadát archívnych dokumentov podľa jednotlivých diplomatických kategórií.

#### **Štatistiky a zoznamy**

Aplikačný komponent bude slúžiť na tvorbu reportov, štatistik a zoznamov zo správy evidencií referenčných údajov.

Vzhľadom na odporúčanie realizovať projekt vo dvoch fázach, predpokladáme pre realizáciu pilotnej fázy nevyhnutné úpravy aplikačného prostredia.

Všetky ďalšie funkcionality budú realizované v druhej fáze implementácie.

## **4.4.2 ArchivLog**

Evidencia, plánovanie procesov konzervovania, reštaurovania, digitalizovania a dislokácie archívnych dokumentov bude novou súčasťou Elektronického archívu ministerstva vnútra SR (ISEA).

ArchivLog bude slúžiť na sledovanie životného cyklu archívnej jednotky pomocou workflow s pokročilými funkciami na spracovanie skenov, ich konverzie, vyťažovania textov, automatizáciu procesov spracovania archívnych dokumentov, analýzu rozloženia a rozpoznávania štruktúry, indexovanie, evidovanie informácií o procese a stave spracovania jednotlivých archívnych fondov a archívnych zbierok, evidovanie informácií o procese a histórii reštaurovania archívnych dokumentov, aby sa procesy dali riadiť a kontrolovať.

#### **Systém bude obsahovať:**

1. integrovanú správu a automatické riadenie pracovných postupov digitalizácie;
2. kontrolu kvality, plánovanie postupov a ich schvaľovanie;
3. správu a kontrolu meta údajov;
4. monitorovanie technologických a procesných chýb, ich včasnú opravu;
5. modul správy a vyhľadávania archívnych dokumentov v databáze;
6. modulárne pripájanie nových typov zariadení;
7. funkcionality vo forme správy procesov, vstupu meta údajov, grafických operácií, ktorými sú orezanie, otočenie, maskovanie, identifikácia súborov, import externých obrazov, výstup meta údajov v XML;
8. prácu s čiarovými a QR kódmi pri identifikácii archívnych dokumentov;
9. prepojenie evidenčných údajov a číselníkov s inými systémami na evidenciu archívneho dedičstva;
10. plánovanie procesu skenovania, prepravy;
11. prehľad voľných a obsadených termínov v plánoch, preplánovanie nadväzujúcich krokov.

#### **ArchivLog umožní:**

1. evidenciu neinvazívnych a invazívnych postupov skúmania archívnych dokumentov;
2. ukladanie výstupov v rôznych formátoch z fyzikálnych, chemických analýz archívnych dokumentov, evidovania reštaurátorských zásahov, postupov ošetrovania, analýz v externých laboratóriách;
3. vytvorenie a tlače záverečných správ a výstupov;
4. priebežné dopĺňanie informácií a hodnotení s fotodokumentáciou k zmenám stavu archívnych dokumentov;
5. plánovanie sterilizácie a očisty archívnych dokumentov pred ich digitalizáciou;
6. vyhotovovanie plánu digitalizácie, operatívne zmeny a prehľad o stave zdigitalizovania jednotlivých archívnych fondov;
7. plánovanie transportov archívnych dokumentov na reštaurovanie a digitalizáciu;
8. prácu s čiarovými a QR kódmi pre identifikáciu archívnych dokumentov.

Žiadny IS momentálne neumožňuje sledovanie a riadenie týchto procesov a ich plánovanie. Štátne archívy disponujú rôznymi aplikáciami, ktoré plnia čiastkové funkcie potrebné pre riadenie procesov. Vzhľadom ku skutočnosti, že iba niektoré funkcie príslušných aplikácií sú predmetom integrácie nie je možné zabezpečiť zapísanie alebo prenos rovnakej informácie do každej aplikácie. Východiskom je zabezpečenie centrálného IS určeného na evidenciu archívneho dedičstva, ktorý by umožnil riadenie nadväzujúcich procesov pri ich správe. V praxi to znamená, že riadenie procesov prebieha buď čiastkovo alebo mimo týchto aplikácií. Za súčasných podmienok je takmer nemožné plánovať dislokáciu archívnych fondov medzi objektmi a plánovať na seba nadväzujúce procesy ako sterilizovanie, ošetrovanie a následné digitalizovanie.

#### **Kontrola a uloženie**

DOtransport Vrutky - MVVytvorenie dávky/Nahratie dokumentuČistenie dokumentuKontrola dokumentuVytvorenie, príprava a odoslanie transportného boxu s dokumentmiTransport MV - VrutkyPrijatie, kontrola a vyprázdnenie transportného boxuPresun transportných boxov z Konzervačného centra do Digitalizačného centraSurová digitalizácia (snímanie tlačenej predlohy)Image ProcessingImage ProcessingPríprava a vrátenie transportného boxuPrijatie transportného boxuKontrola a uloženie dokumentovSledovanie toku digitálnej kópieSledovanie toku fyzického dokumentu dokumentu dokumentuSterilizácia a dokumentu

#### **ArchívLog, zabezpečiť:**

Uchovanie existujúcich informácií migráciou z aplikácií, čím sa aj naďalej umožní presná lokácia v rámci archívov (poschodie, číslo miestnosti, číslo regálov, číslo police, číslo archívnej škatule atď.), vrátane rozšírenia o identifikáciu pomocou QR a čiarových kódov. Taktiež poskytnúť:

1. aktuálne informácie o stave archívnych fondov (biologické poškodenia, fyzikálne poškodenia, popis fondu, stav zdigitalizovania a pod.);
2. plánovanie a kontrolu presunov archívnych fondov;
3. štatistické ukazovatele a zostavy;
4. možnosť plánovať a evidovať sterilizačné, konzervátorské a reštaurátorské zásahy, ukladať výsledky chemických a mikrobiologických výsledkov, tlač záverečných správ;
5. možnosť plánovať preberanie a prevoz archívnych dokumentov na základe QR a čiarového kódu;
6. možnosť plánovania procesu digitalizácie;
7. možnosť prepojenia na úložisko (DMS).

Vzhľadom na odporúčanie realizovať projekt vo dvoch fázach, predpokladáme pre realizáciu pilotnej fázy nevyhnutné úpravy aplikačného prostredia. Všetky ďalšie funkcionality budú realizované v druhej fáze implementácie.

### **4.4.3 Subsystem Digitalizácia**

Odbor archívov a registratúr MV SR potrebuje zaviesť subsystém monitorovania a manažmentu procesu digitalizácie pre dozor nad napĺňaním zákonných a odborných požiadaviek štátnych archívov, zaručenie efektivity využitia nových technológií a účinné napĺňanie cieľov projektu. Subsystem zabezpečuje nadobúdanie objektívnych informácií o napĺňaní projektu, zvýšenie efektivity práce zamestnancov, skvalitnenie operatívnych riešení a strategické plánovanie budúcich investícií.

Popri nasadení tohto softvérového nástroja je nevyhnutná tiež realizácia **organizačných opatrení na úrovni manažmentu a zamestnancov ŠA v nasledujúcich oblastiach:**

1. odborný dozor a konzultačná podpora pre manažment;
2. systém merania skutočne vykonanej práce k naplneniu cieľov digitalizácie
  - úroveň individuálneho zamestnanca, archívu a národná úroveň;
3. analýza údajov zo systému sledovania monitorovania digitalizácie a optimalizácia zdrojov;
4. predpoveď trendov a rizík s pomocou technických nástrojov a odborných metodík;
5. reštrukturalizácia ľudských zdrojov
  - efektívne využitie komerčných firiem pre účely štátnych archívov
  - zvýšenie podielu odbornej práce vykonávanej odborníkmi;
6. vzdelávanie
  - zriadenie programu pre zvyšovanie technologickej kompetencie zamestnancov.

### Nástrojmi monitorovania digitalizácie sú:

1. integrovaný, elektronický systém s osobným, zaheslovaným účtom zamestnancov
  - meranie skutočne vykonanej práce formou naplňovania zadaných kvót
  - história vykonanej práce s meta-údajmi
  - prístup k údajom v systéme na úrovni zamestnancov, archívu a štátnej úrovni;
2. vytvorenie flexibilných metód pre optimalizáciu zdrojov v reálnom čase;
3. nadviazanie kontaktu s komerčnými firmami;
4. reorganizácia práce zamestnancov a spolupráca so sociálnymi programami;
5. pôsobenie odbornej rady archívov a vzdelávací program.

### Predpokladmi pre nasadenie systému sledovania produktivity práce sú:

1. vstupné zariadenia, pripojenie na internet na všetkých pracoviskách
  - osobný identifikátor a školenie pre zamestnancov.

Digitalizácia archívnych dokumentov je základným stavebným kameňom budúceho informačného systému. Obsahuje funkcionality pozostávajúcu z nasledujúcich komponentov:

#### Príprava dokumentov na skenovanie

Vytvorenie sady funkcionalít v ISEA na riadenie prípravy archívnych dokumentov na skenovanie.

#### Vytvorenie digitalizátov

Vytvorenie aplikačnej podpory na zabezpečenie vytvárania elektronických kópií pôvodne neelektronických archívnych dokumentov, ktoré budú prostredníctvom komponentu pre digitalizáciu archívnych dokumentov načítavané do systému ISEA. Digitalizované skeny archívnych dokumentov budú týmto komponentom uložené a pripravené na ďalšie spracovanie.

#### Optické rozpoznanie textu (OCR)

Aplikačná podpora pre vytvorenie a spracovanie OCR vrstvy pre digitalizáty.

#### Vytvorenie identifikátorov

Vytvorenie aplikačnej podpory pre vytvorenie a spracovanie identifikátora pre digitalizát a identifikátora pre archívny dokument.

#### Vytvorenie metadát

Zavedenie centrálného systému na zabezpečenie automatizovaného vytvorenia a zapísania základných popisných metadát pre všetky digitalizované objekty.

Vzhľadom na odporúčanie realizovať projekt vo dvoch fázach, predpokladáme pre realizáciu pilotnej fázy nevyhnutné úpravy aplikačného prostredia. Všetky ďalšie funkcionality budú realizované v druhej fáze implementácie.

## 4.4.4 Migrácia

Komponent realizuje samotnú migráciu. Orchestruje prípravu migračnej sady, samotnú migráciu a následnú kontrolu migrovaných dokumentov. Poskytuje najmä nasledovné funkcionality:

#### Zber dát na migráciu

Zber dát na migráciu zdigitalizovaných archívnych dokumentov z lokálnych úložísk na centrálné dátové úložisko a zabezpečenie ich pravidelného zálohovania s dôrazom na kritické dáta.

#### Doplnenie popisných metadát (Machine learning)

Komponent extrahuje metadáta z ich zdroja prostredníctvom definícií vytvorených v Evidencii archívnych fondov. Takto extrahované metadáta uloží do interného úložiska a poskytne pre ďalšiu validáciu a samotnú migráciu. Implementovaný systém automatizácie popisovania metadát využije technológie strojového učenia (Machine learning) na zabezpečenie automatizovaného vytvorenia a zapísania pokročilých popisných metadát pre relevantné digitalizované objekty. Zároveň bude vybudované používateľské rozhranie pre archivárov na zabezpečenie validácie a korekcie popisných metadát digitalizátov pred samotnou migráciou do Elektronického archívu Slovenska.

Významom zavedenia automatizácie a strojového učenia v štátnych archívoch je zvýšenie efektivity práce archívov, redukcia nákladov, dodatočná bezpečnosť a kontrola kvality pri spracovaní dokumentov a zvýšenie dostupnosti archívnych dokumentov pre občana. Automatizačné systémy a strojové učenie sú dnes štandardom v moderných pracoviskách v Európe aj vo svete.

Aplikácia technológií prebieha v nasledujúcich oblastiach:

1. automatizácia priradovania metadát (Metadata tagging / 'Self-describing' systems);
2. sofistikované vyhľadávanie (Personalized search);
3. vyhodnocovanie a vizualizácia dát (Metrics);
4. počítačová bezpečnosť – detekcia anomálií a výstražný systém (Cyber security).

Nástrojmi subsystému automatizácie a strojového učenia sú:

1. strojové učenie (Machine learning);
2. automatizácia (Automation) ;
3. softvérové, 'slabé' AI, typu 1 (reaktívne) a typu 2 (limitovaná pamäť);
4. spracovanie prirodzeného jazyka (Natural language processing).

### Príprava, anonymizácia a validácia dát

Komponent, ktorý na základe Evidencie archívnych fondov skontroluje migračné údaje. Kontroluje digitalizáty a kontroluje metadáta. Zabezpečí reporting položiek, ktoré neprešli kontrolou a funkčnosť na ich úpravu a anonymizáciu pred migráciou.

### Riadenie migrácie

Komponent na riadenie a kontrolu priebehu migrácie archívnych dokumentov z centrálného dátového úložiska do ISEA. Prostredníctvom integrácie na ISEA umožní vyhodnotenie úspešnosti migrácie. V prípade neúspechu niektorých položiek ich zdokumentuje a umožní ich úpravu a opätovný pokus o migráciu alebo vyradenie z migračnej dávky.

### Vyhodnotenie migrácie

Tvorba reportov a štatistík z procesov riadenia migrácie dát do ISEA a vyhodnotenie migračných behov.

Vzhľadom na odporúčanie realizovať projekt vo dvoch fázach, predpokladáme pre realizáciu pilotnej fázy nevyhnutné úpravy aplikačného prostredia.

Všetky ďalšie funkcionality budú realizované v druhej fáze implementácie.

## 4.4.5 Rozsah informačných systémov

Kód ISVS (z MetalS)	Názov ISVS	Modul ISVS (zaškrtnite ak ISVS je modulom)	Stav ISVS	Typ ISVS	Kód nadradeného ISVS (v prípade zaškrtnutého checkboxu pre modul ISVS)
N/A	N/A	N/A	N/A	N/A	N/A

Prehľad dotknutých informačných systémov v projekte – súčasný stav.

Kód ISVS (z MetalS)	Názov ISVS	Modul ISVS (zaškrtnite ak ISVS je modulom)	Stav ISVS	Typ ISVS	Kód nadradeného ISVS (v prípade zaškrtnutého checkboxu pre modul ISVS)
isvs_202	Elektronický archív ministerstva vnútra SR		Prevádzkovaný a plánujeme rozvíjať	Agendový	N/A

Prehľad rozvíjaných ISVS v projekte – budúci stav.

Kód AS (z Meta IS)	Názov AS	Poskytovaná na externú integráciu (zaškrtnite ak áno)	Typ cloudovej služby	ISVS/ modul ISVS (kód z MetalS)	Aplikačná služba realizuje KS (kód KS z MetalS)
as_62908	Evidencia nového dokumentu do elektronického archívu		žiadny	isvs_202	Evidovanie nového archívneho dokumentu v elektronickom archíve
as_62909	Evidencia zmeny v dokumente do elektronického archívu		žiadny	isvs_202	Evidovanie zmien archívneho dokumentu v elektronickom archíve
as_62910	Poskytovanie údajov z elektronického archívu 3. stranám		žiadny	isvs_202	Poskytnutie údajov z elektronického archívu 3. stranám
as_62911	Poskytovanie údajov z elektronického archívu verejnosti		žiadny	isvs_202	Poskytnutie údajov z elektronického archívu verejnosti
as_62912	Poskytovanie údajov z elektronického archívu do iných informačných systémov		žiadny	isvs_202	Poskytovanie údajov z elektronického archívu do iných informačných systémov

Prehľad budovaných aplikačných služieb – budúci stav.

Koncové služby, ktoré budú výstupom projektu a ich popis:

Kód KS (z MetalS)	Názov KS	Popis	Používateľ	Úroveň elektronizácie KS	
ks_351408	Evidovanie nového archívneho dokumentu v elektronickom archíve	Služba umožňuje zaevidovať nový objekt/ archívny dokument v elektronickom archíve	G2G, G2E,	Úroveň 4	as_62908
ks_351409	Evidovanie zmien archívneho dokumentu v elektronickom archíve	Služba umožňuje zaevidovať zmeny archívneho dokumentu v elektronickom archíve	G2G, G2E,	Úroveň 4	as_62909
ks_351410	Poskytnutie údajov z elektronického archívu 3. stranám	Služba umožňuje poskytnutie údajov z elektronického archívu 3. stranám v závislosti od používateľskej role	G2G, G2E,	Úroveň 4	as_62910
ks_351411	Poskytnutie údajov z elektronického archívu verejnosti	Služba umožňuje poskytnúť vybrané údaje z elektronického archívu verejnosti	G2G, G2E, G2C, G2B	Úroveň 4	as_62911

Aplikačné služby, ktoré budú výstupom projektu a ich popis:

Kód a názov AS	Popis

Evidencia nového dokumentu do elektronického archívu	Služba zabezpečuje evidenciu údajov o novom archívnom dokumente do elektronického archívu v požadovanom tvare
Evidencia zmeny v dokumente do elektronického archívu	Služba zabezpečuje evidenciu zmeny údajov o archívnom dokumente do elektronického archívu v požadovanom tvare
Poskytovanie údajov z elektronického archívu 3. stranám	Služba zabezpečuje poskytovanie údajov o archívnom dokumente z elektronického archívu 3. stranám na základe používateľských rolí
Poskytovanie údajov z elektronického archívu verejnosti	Služba zabezpečuje poskytovanie vybraných údajov z elektronického archívu verejnosti v rámci zvýšenia kultúrneho povedomia
Poskytovanie údajov z elektronického archívu do iných informačných systémov	Služba zabezpečuje poskytovanie údajov z elektronického archívu do iných informačných systémov cez definované rozhranie

#### 4.4.6 Využívanie nadrezortných centrálnych blokov (SaaS)

Neplánuje sa využívanie nadrezortných centrálnych blokov ani podporných spoločných blokov (SaaS). V rámci projektu nebola identifikovaná funkcionality nadrezortných centrálnych blokov využiteľná pre potreby a výstupy projektu.

##### 4.4.6.1 Prehľad plánovaného využívania podporných spoločných blokov (SaaS)

Neplánuje sa využívanie podporných spoločných blokov (SaaS). V rámci projektu nebola identifikovaná funkcionality spoločných blokov (SaaS) využiteľná pre potreby a výstupy projektu.

##### 4.4.6.2 Prehľad plánovaných integrácií ISVS na nadrezortné centrálné bloky

###### – spoločné moduly

Neplánuje sa využívanie nadrezortných centrálnych blokov – spoločných modulov. V rámci projektu nebola identifikovaná funkcionality spoločných modulov využiteľná pre potreby a výstupy projektu.

##### 4.4.6.3 Prehľad plánovaných integrácií ISVS na nadrezortné centrálné bloky

###### - modul procesnej integrácie a integrácie údajov (IS CSRÚ)

Neplánuje sa integrácia na IS CSRÚ. V rámci projektu nebude realizované poskytovanie alebo konzumovanie referenčných údajov z IS CSRÚ, nie je preto potrebné vytvoriť integračnú väzbu na modul proces integrácie a integrácie údajov podľa § 10 ods. 11 zákona č 305/2013 Z. z. zákon o e-Governmente.

##### 4.4.6.4 Poskytovanie údajov z ISVS do IS CSRÚ

Neplánuje sa integrácia na IS CSRÚ. V rámci projektu nebude realizované poskytovanie referenčných údajov z ISEA do IS CSRÚ, nie je preto potrebné vytvoriť integračnú väzbu na modul proces integrácie a integrácie údajov podľa § 10 ods. 11 zákona č 305/2013 Z. z. zákon o e-Governmente.

##### 4.4.6.5 Konzumovanie údajov z IS CSRÚ

Neplánuje sa integrácia na IS CSRÚ. V rámci projektu nebude realizované konzumovanie referenčných údajov z IS CSRÚ, nie je preto potrebné vytvoriť integračnú väzbu na modul proces integrácie a integrácie údajov podľa § 10 ods. 11 zákona č 305/2013 Z. z. zákon o e-Governmente.

## 4.5 Dátová vrstva

Kapitola definuje objekty evidencie, ktoré spravuje Štátny archív MVSR a poskytuje tiež detailný pohľad na údaje, ktoré budú potrebné v dobudovanom IS.

### 4.5.1 Údaje v správe organizácie

Prehľad o údajoch v správe organizácie predstavuje nasledovná tabuľka:

ID	Objekt evidencie
#1	Listiny

#2	Mapy a plány
#3	Úradné knihy a pomocná evidencia
#4	Spisový materiál
#5	Účtovný materiál
#6	Technická dokumentácia
#7	Matriky
#8	Obrazový materiál a fotografie
#9	Osobné doklady
#10	Osobná korešpondencia
#11	Kroniky
#12	Kartotéky
#13	Pečate
#14	Iné

## 4.5.2 Dátový rozsah projektu

V predchádzajúcej kapitole boli definované dáta v správe organizácie. Z hľadiska rozsahu projektu sú podstatné dáta všetkých objektov evidencie, ktoré budú novým IS využívané, nakoľko sú všetky súčasťou archívnych fondov - archívnej jednotky v ktorej je nutné zachovať vnútornú integritu. Jedná sa o nasledovné:

ID	Objekt evidencie	URI identifikátor	Popis
#1	Listiny	N/A	<i>Písomné dokumenty o právnom úkone, s dlhodobou právnou záväznosťou.</i>
#2	Mapy a plány	N/A	<i>Grafické dokumenty zjednodušene zobrazujúce zemský povrch či oblohu a pôdorysný náčrt diela.</i>
#3	Úradné knihy a pomocná evidencia	N/A	<i>Súbor úradných záznamov postupne vedených pre právne, administratívne, prevádzkové a evidenčné účely úradu.</i>
#4	Spisový materiál	N/A	<i>Súbor dokumentov, ktoré vznikli v súvislosti s vybavovaním daných záležitostí.</i>
#5	Účtovný materiál	N/A	<i>Súbor dokumentov, ktorý vznikol z finančného hospodárenia pôvodcu fondu.</i>
#6	Technická dokumentácia	N/A	<i>Členený súbor podkladov, ktoré umožňujú realizovať fázy výroby diela.</i>
#7	Matriky	N/A	<i>Úradné knihy o narodení (krste), sobáši a úmrtí občanov matričného obvodu.</i>
#8	Obrazový materiál a fotografie	N/A	<i>Výtvarné, fotografické, schematické a iné grafické informatívne dokumenty.</i>
#9	Osobné doklady	N/A	<i>Dokumenty slúžiace na identifikáciu fyzickej osoby.</i>
#10	Osobná korešpondencia	N/A	<i>Súbor listov z písomnej výmeny informácií medzi fyzickými a právnickými osobami.</i>
#11	Kroniky	N/A	<i>Knihy opisujúce historické udalosti v chronologickom poradí, zvyčajne počas dlhého obdobia.</i>
#12	Kartotéky	N/A	<i>Jednotne organizované usporiadané zbierky záznamov.</i>
#13	Pečate	N/A	<i>Odtlačky matrice typára s identifikačným znamením, ktoré jednoznačne identifikuje a osvedčuje majiteľa k zaisteniu právnej platnosti dokumentu.</i>
#14	Iné	N/A	<i>Ostatné významné záznamy, ktoré nemožno zaradiť do typických diplomatických kategórií.</i>

Potrebná je tiež migrácia už existujúcich dát, ich harmonizovanie a zaktualizovanie v kontexte s novými požiadavkami a funkcionalitami nad databázou, ktorej budú základom.

Finálny dátový model bude vytvorený v etape detailného návrhu riešenia.

## 4.5.3 Kvalita a čistenie údajov

Zhodnotenie objektov evidencie (digitalizátov) z pohľadu dátovej kvality.

S ohľadom na charakter projektu nie je možné niektoré dáta označiť za viac alebo menej významné, nakoľko ich kvalita spoločne definuje skenovaný a zaznamenávaný objekt.

Typy dát sú znázornené v tabuľke nižšie:



ID OE	Objekt evidencie	Významnosť kvality 1 (malá) až 5 (veľmi významná)	Citlivosť kvality 1 (malá) až 5 (veľmi významná)	Priorita – poradie dôležitosti
#1	Listiny	5	4	1.
#2	Mapy a plány	5	4	1.
#3	Úradné knihy a pomocná evidencia	5	3	1.
#4	Spisový materiál	3	3	1.
#5	Účtovný materiál	4	3	1.
#6	Technická dokumentácia	3	3	1.
#7	Matriky	4	3	1.
#8	Obrazový materiál a fotografie	5	3	1.
#9	Osobné doklady	3	3	1.
#10	Osobná korešpondencia	3	3	1.
#11	Kroniky	3	3	1.
#12	Kartotéky	3	3	1.
#13	Pečate	5	3	1.
#14	Iné	3	2	1.

Kategorizácia objektov evidencie z pohľadu dátovej kvality – budúci stav.

#### 4.5.4 Role a predbežné personálne zabezpečenie pri riadení dátovej kvality

Rola	Činnosti	Pozícia zodpovedná za danú činnosť (správca ISVS / dodávateľ)
Dátový kurátor	Evidencia požiadaviek na dátovú kvalitu, monitoring a riadenie procesu.	MV SR
Odborník pre IT dohľad, QA	Dohľad a zabezpečenie správnosti implementácie IT riešení, kontrola kvality.	Dodávateľ
Špecialista pre databázy	Analyzuje požiadavky na dáta, modeluje obsah procedúr.	Dodávateľ
Manažér kvality	Spracovanie výstupov merania, interpretácie, zápis biznis pravidiel, hodnotiace správy z merania.	MV SR
Špecialista pre infraštruktúry	Spravovanie IS, zabezpečovanie dostupnosti systémov, profylaktická činnosť.	Dodávateľ
Špecialista pre bezpečnosť IT	Zodpovednosť za dodržanie princípov a štandardov na IT bezpečnosť, za kontrolu a audit správnosti riešenia v oblasti bezpečnosti.	Dodávateľ

Prehľad rolí a personálneho zabezpečenia pre riadenie dátovej kvality.

#### 4.5.5 Referenčné údaje

Neplánuje sa poskytovanie údajov vo forme referenčných údajov po ich naplnení do ISEA z rôznych dostupných dátových zdrojov a ani konzumácia referenčných údajov z iných IS.

#### 4.5.6 Otvorené údaje

Plánujeme poskytnutie otvorených údajov verejnosti. Pôjde o vybrané údaje v súlade s požiadavkami a v kontexte s pravidlami o poskytovaní údajov v zmysle zákona. Plánujeme poskytovať nasledujúce údaje:

- Názov archívu
- Referenčný kód
- Názov archívneho fondu
- Časový rozsah
- Rozsah v bm
- Názov pôvodcu
- Podmienky prístupu
- Dátum zápisu (zmeny)
- Archívne pomôcky

- Druh archívnej pomôcky
- Názov archívnej pomôcky
- Časový rozsah
- Evidenčné číslo archívnej pomôcky
- Rok vzniku
- Autor

Nakoľko Štátny archív nedisponuje zharmonizovanými a kompletnými údajmi o digitalizovaných objektoch, bude presný typ informácií, vhodných pre poskytnutie verejnosti finálne definovaný až vo fáze Analýzy a špecifikácie. Využitie OpenAPI technológie bude v rámci projektu taktiež predmetom analýzy.

Za účelom poskytovania otvorených údajov budeme využívať službu:

- Poskytnutie údajov pre [data.gov.sk](http://data.gov.sk) MVS SR ISEA (as\_63239), naviazanú na "sluzba\_is\_48063" Zápis a aktualizácia údajov z dátového zdroja.

## 4.5.7 Analytické údaje

Neplánuje sa sprístupnenie dátových zdrojov na analytické účely, keďže v rámci ISEA nebudú generované údaje vhodné na analytické spracovanie.

## 4.5.8 Moje údaje

Plánuje sa poskytnutie údajov pre službu Moje dáta, keďže v rámci projektu sa uvažuje s konaniami, ktoré sa týkajú fyzickej osoby alebo právnickej osoby, evidovaním údajov, vrátane osobných údajov, viažucich sa k fyzickej osobe alebo právnickej osobe, resp. evidenciou údajov obsiahnutých v návrhu na začatie konania, žalobe, rozhodnutí, žiadosti, sťažnosti, vyjadrení, stanovisku a ohlásení alebo inom dokumente, ktorý vydáva v konaní povinný subjekt, viažuci sa ku konkrétnej fyzickej osobe alebo právnickej osobe.

## 4.5.9 Prehľad jednotlivých kategórií údajov

ID	Register / Objekt evidencie	Referenčné údaje	Moje údaje	Otvorené údaje	Analytické údaje
#1	Listiny	-	X	X	-
#2	Mapy a plány	-	X	X	-
#3	Úradné knihy a pomocná evidencia	-	X	X	-
#4	Spisový materiál	-	X	X	-
#5	Účtovný materiál	-	X	X	-
#6	Technická dokumentácia	-	X	X	-
#7	Matriky	-	X	X	-
#8	Obrazový materiál a fotografie	-	X	X	-
#9	Osobné doklady	-	X	X	-
#10	Osobná korešpondencia	-	X	X	-
#11	Kroniky	-	X	X	-
#12	Kartotéky	-	X	X	-
#13	Pečate	-	X	X	-
#14	Iné	-	x	X	-

## 4.5.10 Využívanie údajov

Z pohľadu využívania údajov navrhované riešenie zabezpečí:

1. prípravu a naplnenie dát požadovaných širokou verejnosťou na prehľadávanie a umožnenie prístupu na základe žiadostí súvisiacich so životnými situáciami občanov a podnikateľov;
2. prehľadávanie a umožňovanie prístupu referenčných údajov po ich naplnení v ISEA z rôznych zdrojov v štátnych archívoch;
3. samotné prehľadávanie a umožňovanie prístupu k archívnemu dedičstvu bude zabezpečené existujúcimi službami ISEA pre bádateľskú a správnu agendu;
4. vytvorenie nástrojov pre anonymizáciu, ktoré zabezpečia, že údaje osobného charakteru bude možné používať bez narušenia súkromia.

## 4.5.11 Správa údajov

Správa údajov pri digitalizácii archívnych dokumentov bude riešená centrálné, v kompetencii odboru archívov a registratúr MV SR.

Riešenie bude poskytovať nasledovné funkcie:

1. na riadenie údajov vrátane funkcií a aplikácií pre naplnenie referenčných údajov a evidencií, riadenie prístupov a oprávnení, riadenie kvality, riadenie prístupu k osobným údajom, riešenie incidentov pri plnení archívu a riadenie životného cyklu metadát;
2. pre jednotnú formu unikátnych identifikátorov ľubovoľných dátových údajov pre ich adresovateľnosť - každý údaj bude jednoznačne adresovateľný cez jednoznačný identifikátor, pričom tento identifikátor prideluje systém, ktorý stráži celkovú integritu referenčných identifikátorov;
3. na vytvorenie centrálného modelu dát pre plnenie ISEA;
4. na správu evidencie digitalizácie;
5. na vytváranie a aktualizáciu plánu digitalizácie a plnenia ISEA;
6. na riadenie plnenia ISEA dátami z archívnych fondov;
7. na definovanie štruktúry metadát k archívnym dokumentom;
8. na definovanie pravidiel na doplnenie popisných metadát vytvorených z podkladov.

## 4.5.12 Riadenie údajov

Za riadenie plnenia údajov a ich kvalitu budú zodpovedať príslušné pracoviská organizačne zaradené pod odbor archívov a registratúr MV SR.

Pre úspešné riadenie údajov budú vytvorené funkcie na:

1. dávkové naplnenie ISEA dátami z uložených digitalizovaných archívnych dokumentov;
2. prípravu elektronických obrazov archívnych dokumentov z archívnych fondov v neelektronickej forme (digitalizácia neelektronických archívnych dokumentov uložených v štátnych archívoch);
3. prehľadávanie a spracovávanie tlačených a písaných znakov do textovej editovateľnej formy (OCR) pre digitalizované archívne dokumenty pred naplnením do ISEA;
4. optické rozpoznávanie obsahu archívnych dokumentov (funkcia OCR) pred vložením do archívu;
5. zrevidovanie a doplnenie popisných metadát vytvorených z podkladov – napr. z názvu skenu, z metadát z iných zdrojov, z metadát z xls podľa diplomatických kategórií;
6. vytváranie a editovanie nových metadát pred naplnením do archívu;
7. kontrolu kvality dát pred a počas procesu migrácie dát;
8. vyhodnotenie úspešnosti behu plnenia dát a na zopakovanie migrácie chybných dát po ich oprave;
9. uloženie archívnych dokumentov a ich metadát v ISEA;
10. anonymizáciu údajov, na ktoré sa vzťahuje ochrana osobných údajov, teda pri sprístupňovaní dát archívnych dokumentov bude dostupná funkcia pre anonymizáciu dát;
11. sprístupnenie archívnych dokumentov na ďalšie spracovanie.

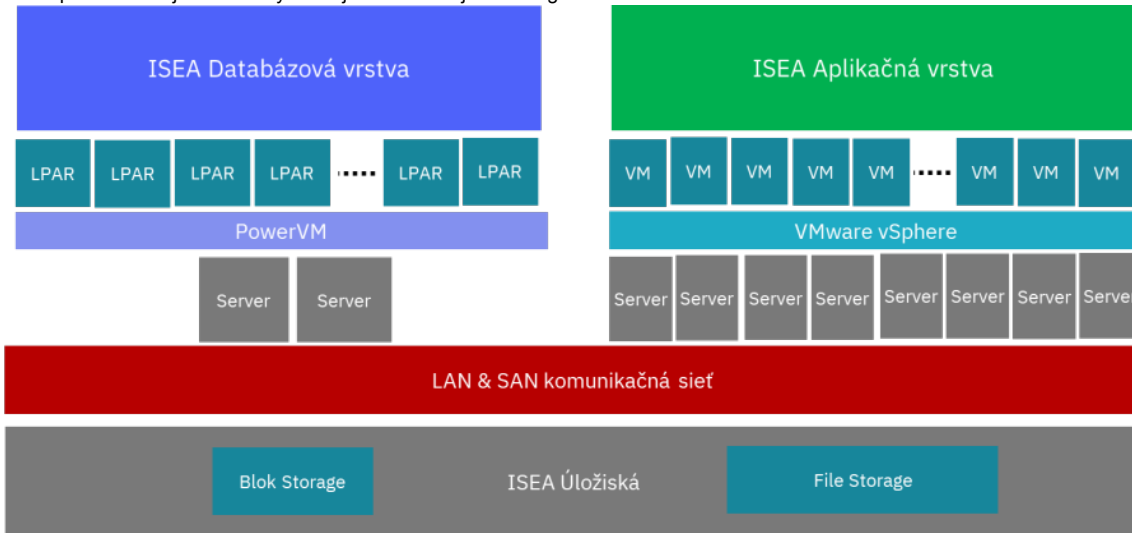
## 4.6 Technologická vrstva

### 4.6.1 Prehľad technologického stavu.

Z pohľadu **AS IS** stavu je systém ISEA postavený nasledovne:

Aplikačné komponenty systému sú vytvorené integráciou vývoja (PHP a Java), SW produktov (SW middleware) a technológií . Ako aplikačné serve sa využívajú Apache HTTP, Apache Tomcat, WebSphere Liberty. Vyhľadávanie je založené na platforme ElasticSearch.

Koncept Technickej architektúry ISEA je na nasledujúcom diagrame:



Technické riešenie je postavené na virtuálnej platforme VMware (prezentačná a aplikačná vrstva) a technológii PowerVM (perzistentná vrstva), a je plne integrované do prostredia DC MV SR. Technické riešenie pre Úložiská archívu využíva diskové polia - operatívne a archívne úložiská, páskové knižnice - archívne úložiská.

Z dôvodu konzistencie cieľového riešenia a kompatibility jednotlivých komponentov bude vhodné uvažovať o použití identických technológií aj v prípade To-Be riešenia.

Súčasnú technické riešenie má nasledovné charakteristiky:

1. prevádzkový režim systému 24x7;
2. vysoká dostupnosť riešenia zabezpečená na aplikačnej a infraštruktúrnej úrovni redundanciou komponentov a automatizovaného riadenia činností;
3. infraštruktúra je umiestnená v dvoch centrálnych geograficky oddelených DC MV SR;
4. riešenia sú implementované v štandardnej 3-vrstvovej architektúre:
  - prezentačná vrstva,
  - servisná a aplikačná vrstva,
  - dátová vrstva;
5. umiestnenie aplikačných a technologických komponentov je v súlade s princípmi bezpečnostnej zonácie DC MV SR (Demilitarized zone, Untrusted zone, Trusted zone);
6. každý server (aj virtualizovaný) je umiestnený v jedinej bezpečnostnej zóne (DMZ vs. Untrusted vs. Trusted);

1. riešenie má 2 prostredia – produkčné a testovacie, pričom:

*Produkčné prostredie (PROD) – 2-lokalitné:*

- primárna lokalita DC 1 – kompletne HA (High availability) riešenie,
- sekundárna lokalita DC 2 – DR lokalita v plnej funkčnosti, ale s redukovanou serverovou infraštruktúrou a páskovým typom dlhodobého archívneho úložiska;

*Testovacie prostredie (TEST) – 1-lokalitné, umiestnené v DC 1:*

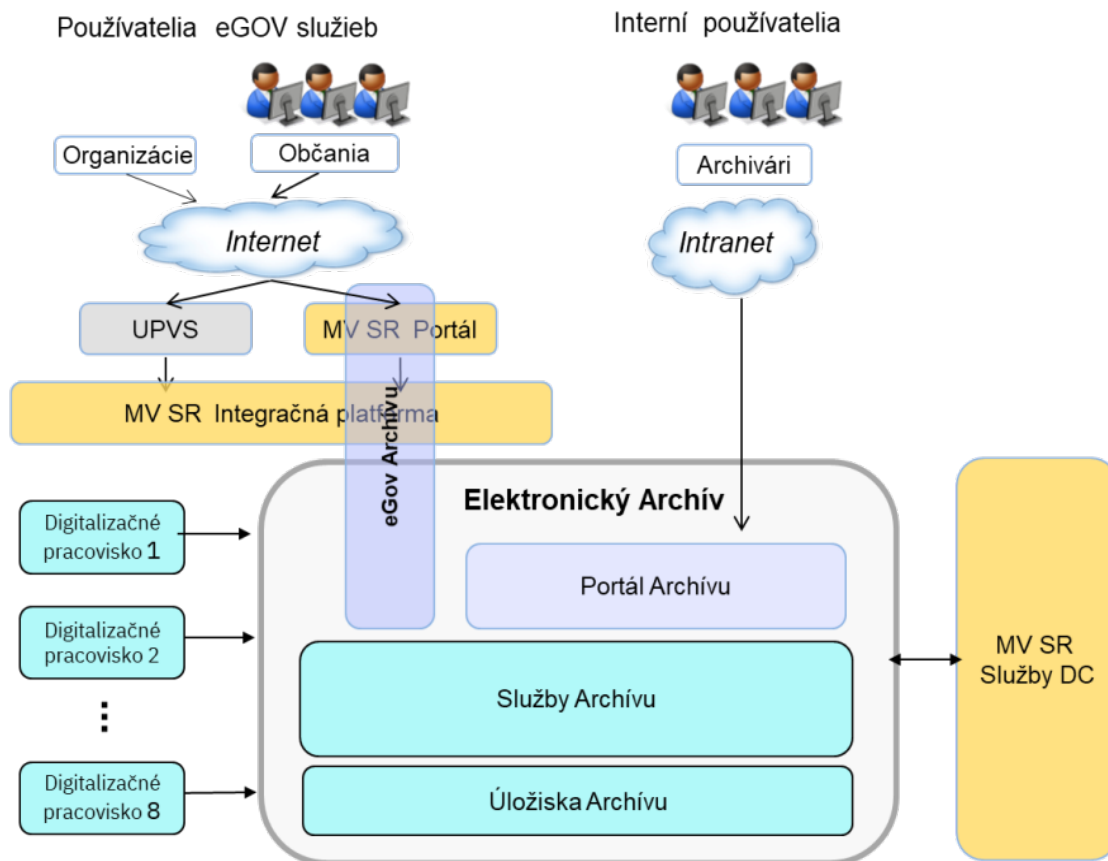
- plne nezávislé voči PROD,
- bez DR lokality,
- kapacitne redukované v porovnaní s PROD, schopné zabezpečiť plnohodnotnú prevádzku IS v obmedzenom režime z pohľadu zaťažiteľnosti.

1. archívne úložisko je z pohľadu oboch lokalít navzájom technologicky rozdielne, aby dáta neboli uložené len na jednom type média:
  - primárna lokalita – dôveryhodné diskové úložisko
  - sekundárna lokalita – dôveryhodné páskové úložisko;
2. je implementovaná úplná virtualizácia HW infraštruktúry a dynamická re-distribúcia výkonu (LA - load balancing) na úrovni hypervizora;
3. vysoká dostupnosť (HA) je zabezpečená v rámci primárnej lokality predovšetkým nasledovnými technológiami:
  - serverová infraštruktúra
  - VMware vSphere, funkcionálna HA,

- IBM Power, funkcionálna HA;
- pre dátové úložiská:
  - redundancia kontrolerov/radičov a ďalších interných komponentov,
  - nezávislé napájanie,
  - RAID technológie na úrovni fyzických diskov;
- v prípade katastrofickej udalosti fyzických serverov (DR) v DC 1 je „failover“ činnosti medzi DC realizovaný predpripraveným plánom DR;
- pre zabezpečenie doby obnovy sú potrebné dáta zrkadlené medzi DC na úrovni diskových polí cez SAN v synchrónnom režime;
- prechod prevádzky do záložnej lokality (DR) – riadeným prepnutím na základe rozhodnutia kompetentnej technickej autority a v súlade s krokmi DR scenára, obdobne aj nasledovný návrat späť ;
- relokácia VMware virtuálnych serverov HP Intel do druhej lokality je riadená cez Site Recovery Manager (VMware SRM);
- synchrónne zrkadlenie vybraných dát medzi oboma DC je realizované na aplikačnej a storage úrovni (SAN) s využitím existujúcej infraštruktúry prepojenia oboch DC;
- riešenie využíva IBM TSM servery v archívnom móde pre ukladanie a sprístupňovanie dokumentových balíčkov z páskových technológií;
- archívne TSM servery sú umiestnené na IBM Power serveroch s OS AIX v dedikovaných LPAR-och;
- riešenie je začlenené do existujúcej sieťovej infraštruktúry na úrovni SAN, LAN, WAN, load balancerov a firewall infraštruktúry v zmysle definovaných komunikačných štandardov DC MV SR;
- riešenie je začlenené do základnej environmentálnej infraštruktúry s využitím jej kapacít (oblasti: napájanie, chladenie, záložné zdroje – UPS a motor generátor);
- riešenie využíva infraštruktúrne služby v DC ako sú DNS, NTP, LDAP , Monitoring CA, zálohovací systém TSM;
- zálohovanie systému využíva existujúcu centrálnu TSM infraštruktúru v oboch DC;
- eGov služby ISEA využívajú existujúcu infraštruktúru Portálu a Integračnú platformu MV SR a súvisiacu infraštruktúru (napr. EKR, MDA, DMS, LDAP2M).

#### 4.6.2 Architektúra infraštruktúry

Technické a technologické riešenie v **TO BE** stave – znamená rozšírenie o komponenty a moduly Digitalizácie - musí byť plne integrované do súčasného prostredia DC MV SR. To znamená virtualizovateľné, nezávislé na hardvérovom prostredí, prevádzkovateľné v DC MV SR, pričom špecifikácia systémových požiadaviek bude v súlade s aktuálne poskytovanou úrovňou služieb v rámci dotknutých IS a DC.

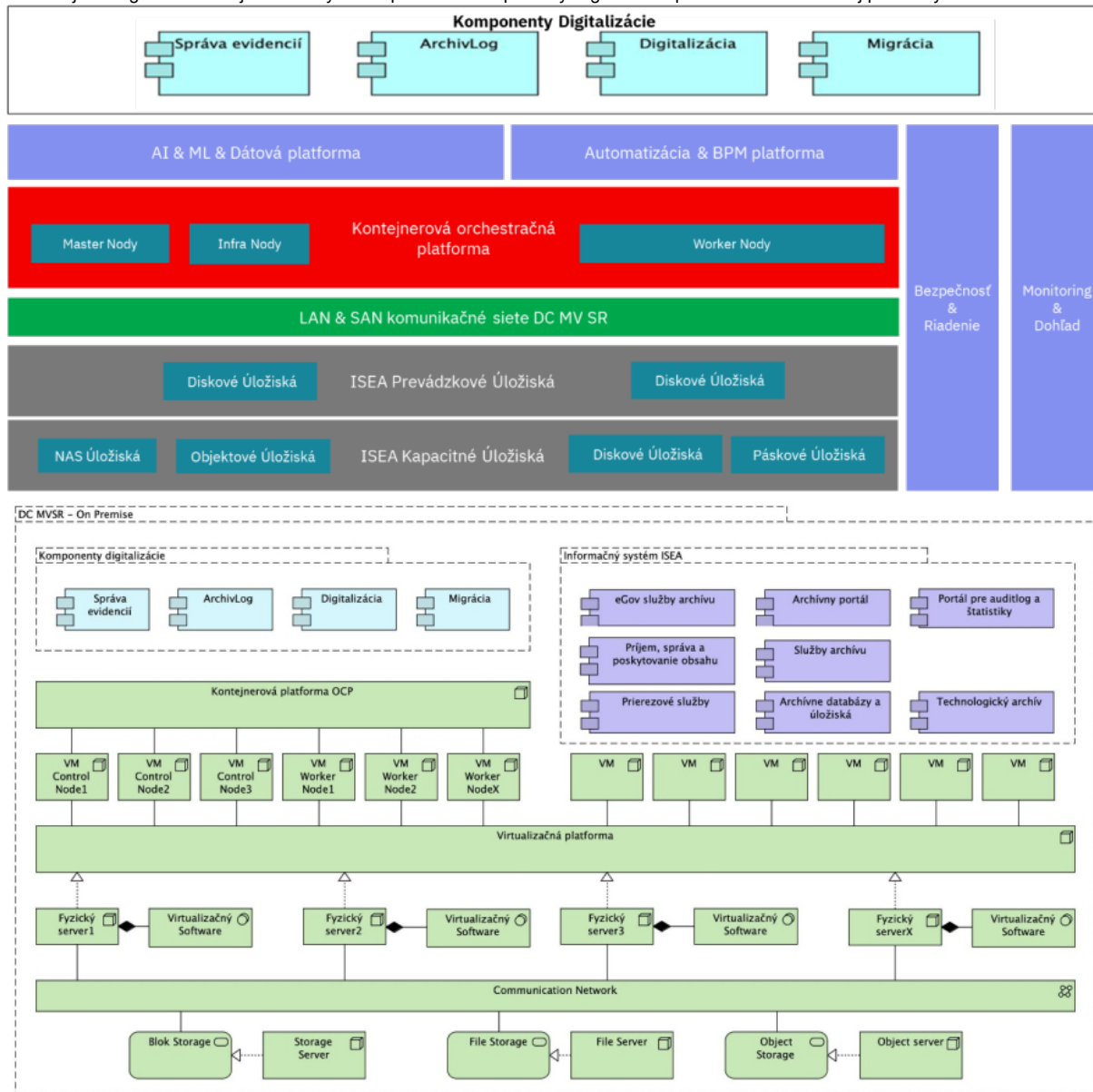


#### Systémová architektúra s rozšírením Digitalizácia

Systém ISEA bude doplnený o Digitalizačné pracoviská, ktoré primárne zabezpečia prípravu a prevod zo zdrojovej formy (papier a mikrofílm) dokumentov do digitálnej formy nasadením špecializovaných skenerov. Na Digitalizačných pracoviskách sa zrealizujú aj základné postprocessingové činnosti špecifikované podľa typu vstupných dokumentov a pripraví sa dávka na odoslanie do centrálného systému Elektronický Archív. Centrálny systém ISEA bude doplnený o funkcionálnu a potrebné infraštruktúrne kapacity na úrovni Služieb a Úložisk Archívu.

Pre technologickú platformu rozšírenia Centrálného systému ISEA je doporučovaná otvorená platforma orchestrácie kontajnerov a virtuálnych serverov,

ktorá umožňuje zabezpečiť dostupnosť aj odolnosť nasadzovaných aplikácií. Zároveň s využitím doporučených prístupov a vzorov, pri tvorbe a nasadzovaní aplikácií (Cloud Native applications) vytvorí moderné, otvorené kontajnerové prostredie pre požadovanú úroveň prevádzky systému. Otvorená platforma zároveň umožňuje integráciu so súčasnou prevádzkou systému v datacentre MV SR (on-premise prostredie). Nasledujúci diagram zobrazuje technický návrh pre nové Komponenty Digitalizácie pre rozšírenie súčasnej platformy ISEA:



### Návrh Technickej architektúry Digitalizácie ISEA

Technický návrh je rozdelený do viacerých úrovní a technologických častí:

- Storage úroveň je tvorená vrstvami a technológiami s ohľadom na požiadavky na výkonnosť, kapacitu a optimalizáciu nákladov. Je tvorená vysokovýkonnou vrstvou, ktorá poskytuje maximálny výkon úložiska s nízkou latenciou pre IO operácie (IOPS) a veľkou šírkou pásma pre celkovú priepustnosť (Gbps). Druhá Kapacitná vrstva je optimalizovaná s ohľadom na náklady/kapacitu. Vyžaduje vysokú garanciu odolnosti, dostupnosti a spoľahlivosti s možnosťami geo-replikačných schopností. Kapacitná vrstva môže byť zostavená z viacerých rôznych úložných architektúr a technológií, napr. pevné disky (HDD), sieťové úložisko (NAS), objektové úložisko a páskové technológie.
- Pre sieťovú (komunikačnú) úroveň sa predpokladá využitie súčasného sieťového prostredia (SDN) DC MV SR. Primárne rozdelenie infraštruktúry a prístupu do bezpečnostných zón bude realizované na sieťovej úrovni a doplnené na každej úrovni riešenia.
- Pre sieťovú vrstvu medzi oboma prevádzkovými výpočtovými centrami MVSR a vývojovým prostredím projekt počíta s posilnením prepojenia na báze DWDM.
- Úroveň Compute tvorí otvorená kontajnerová platforma Red Hat OpenShift a riešenie na automatizované nasadzovanie aplikácií, resp. správu kontajnerov - Kubernetes (K8S). Red Hat OpenShift s dostatočným počtom Master, Infra a Worker Nodov, kde Worker Nody predstavujú prevádzkové prostredie pre aplikačné kontajnery riešenia. Pre aplikačné prostredie sú navrhnuté dve platformy nad ktorými bude vytvorené/implementované požadované funkčné riešenie. AI & ML - Dátová platforma je určená pre masívnu automatizáciu spracovania vstupných dokumentov a na elimináciu manuálneho prístupu. Platforma umožní určenie typu dokumentov a charakteristických entít použitých v dokumentoch bez toho, aby dokument musel čítať a analyzovať pracovník archívu. Platforma AI & ML - poskytnúť:

1. klasifikácia dokumentov a identifikácia dokumentov podľa podobných vlastností, tréningovanie modelov klasifikácie, sledovanie a vyhodnotenie kvality klasifikácie, spätná väzba klasifikácie na vzorkách dokumentov, detekcia kvality modelu a identifikácia potreby pretrénovania modelu
2. identifikácia entít v dokumentoch, zostavenie metadát popisujúcich entity v dokumente, správa entít, manuálne doplnenie entít do indexu platných entít
3. spracovanie jedného dokumentu - prípadne dávky (digest) dokumentov
4. hromadné / dávkové spracovanie predložených dokumentov - prípadne častí (digest) dokumentov

Platforma pre Automatizáciu a BPM - poskytne (okrem iného):

1. prostredie pre návrh a správu procesov a ostatných artefaktov prostredníctvom webovej aplikácie, ktorá podporuje prehliadače Google Chrome 69+, Mozilla Firefox 60+, Microsoft Internet Explorer 11+, Microsoft Edge 25+, Apple Safari 11+
2. prostredie umožniť definovať procesy za použitia prvkov štandardu BPMN 2.0 formou drag and drop
3. prostredie podporuje tvorbu úloh pre používateľov a pre systém, ktorý ich vykonáva na pozadí
4. prostredie umožňuje staticky aj dynamicky definovať SLA, prioritu a dátum spracovania úloh a procesov
5. prostredie podporuje dynamické filtrovanie používateľov, ktorí majú pracovať na úlohe
6. prostredie disponuje návrhárom grafických rozhraní pre používateľské úlohy a neprocesné funkcionality
7. prostredie umožňuje definíciu orchestračných služieb formou drag and drop
8. prostredie podporuje definíciu tímov, ktoré je možné priradiť k používateľským úlohám procesu
9. prostredie obsahuje nástroje pre správu a integráciu.

### 4.6.3 Bezpečnosť a Riadenie

Systém musí spĺňať všetky požiadavky na bezpečnosť v DC podľa odsúhlasených smerníc MV SR.

### 4.6.4 Špecifické technologické potreby projektu

#### 4.6.4.1 Digitalizácia a vyťaženie textu z naskenovaných archívnych dokumentov

Vybavenosť štátnych archívov IKT charakterizuje morálna a technická zastaranosť pracovných staníc aj skenerov, ktorých je nedostatok, alebo sú po životnosti.

Technológia na extrahovanie, následné prehľadávanie a spracovávanie tlačенých a písaných znakov do textovej editovateľnej formy (OCR) je aplikovateľná na novodobých archívnych dokumentoch, ktoré vznikali prostredníctvom písacích strojov alebo boli vytlačené v tlačiarňach. Zároveň významne urýchli vyhľadávanie informácií vo vyextrahovanom texte a ich následné spracovanie. Na vysporiadanie sa s chybovosťou dokážeme použiť technológiu vyhľadávania s prvkami umelej inteligencie, ktorá umožňuje prípadné chyby nesprávne rozpoznávaných alfanumerických znakov nahradiť znakmi významovo najbližšími a ponúknuť súvisiaci obsah.

Pri rukopisoch je potrebné aplikovať technologické a vedecké inovácie v oblasti rozpoznávania textu v rukopisoch založené na využívaní umelej inteligencie, ktoré súvisia s rozvojom metodológie digitálnych humanitných vied v Európe. Cieľom je implementácia najnovších poznatkov z výskumu automatického textového rozpoznávania historických dokumentov na Slovensku. Môžeme nadviazať na APVV projekt realizovaný na UMB "Inovatívne sprístupnenie písomného dedičstva Slovenska prostredníctvom systému automatickej transkripcie historických rukopisov" (tzv. "SKRIPTOR"). Platforma Transkribus je založená na mechanizme tzv. Neural engine, t. j. strojovom učení, kedy sa obrazovej predlohe znaku na rukou písanom origináli priraduje konkrétny alfanumerický znak. Opakovaním tohto procesu, teda učením, sa množina možných riešení, ktoré Transkribus priraduje obrazovej predlohe znaku postupne znižuje až na hodnotu rovnú skutočnej hodnote znaku. Transkribus umožňuje ukladať hotové modely pre konkrétny typ písma a zdieľať ho navzájom medzi vedeckými inštitúciami aj zo zahraničia a využiť už hotové modely.

#### 4.6.4.2 Ciele výberu a digitalizácie najstarších rukopisných a tlačенých dokumentov

V súlade so zákonom o archívoch a registratúrach štátny archív má predkladať na štúdium originál archívneho dokumentu iba výnimočne. Kópiu archívneho dokumentu predkladá vždy, ak ide o archívny dokument, ktorý vznikol pred rokom 1526, obsahuje informácie, ku ktorým je obmedzený prístup alebo by pri štúdiu mohlo dôjsť k jeho poškodeniu, resp. možnosti jeho poškodenia. Štátne archívy spravujú 39 013 stredovekých listín.

Priority výberu archívnych dokumentov na digitalizáciu:

1. vzácne tlače a rukopisy s vysokou kultúrnou a historickou hodnotou, ktoré sú vystavené vysokému riziku fyzického poškodenia súvisiaceho s ich vekom, materiálovým zložením či spôsobom využívania v priereze ich existencie, a vyžadujú špeciálne zaobchádzanie pri používaní;
2. archívne dokumenty bádateľsky často využívané formou štúdia alebo formou kópie, identifikované na základe analýzy údajov v žiadostiach o prístup k archívny dokumentom;
3. popularizácia vedy a histórie online prístupom tematicky zameraných publikačných celkov (významné osobnosti, udalosti, budovy) v ponímaní prístupu ako Open data sety;
4. umožnenie prístupu k obsahovo zaujímavým archívny dokumentom, ktoré môžu byť voľne prístupné online a zároveň tvoria významnú súčasť kultúrneho dedičstva Slovenska, Európy a sveta;
5. archívne dokumenty sprístupnené archívnu pomôckou (katalóg, inventár);
6. archívne dokumenty formou i obsahom vhodné pre extrahovanie textu a metadát pre rýchle vyhľadávanie podľa kľúčových slov;
7. Archívne dokumenty využiteľné pre vedecko-výskumnú činnosť s potenciálom využitia v osvetovej činnosti, vo vzdelávacom procese na všetkých stupňoch alebo v širšom spoločenskom kontexte (médiá, školy a pod.).

#### 4.6.4.3 Súčasný stav online sprístupnenia zdigitalizovaných archívnych dokumentov

Štátne archívy nemajú špecializovaných digitalizačných zamestnancov vyčlenených na skenovanie a následné postprocessing úpravy naskenovaných dokumentov. Z uvedeného dôvodu sa prístup k archívny dokumentom rieši najmä osobným štúdiom archívnych dokumentov v bádateľniach štátnych archívov. Pri osobnom štúdiu dochádza k manipulácii s archívny dokumentami, ktorej výsledkom môže byť ich poškodenie.

Jedným z ďalších preventívnych opatrení pri kontaktnom štúdiu archívnych dokumentov ustanovenom vo vyhláske MV SR č. 628/2002 Z. z., ktorou sa vykonávajú niektoré ustanovenia zákona o archívoch a registratúrach a o doplnení niektorých zákonov v znení neskorších predpisov (ďalej len „vyhláska č. 628/2002 Z. z.“) je ochrana pred ich odcudzením. Bádateľne štátnych archívov nie sú vybavené kamerovým systémom pre zabezpečenie zvýšenej ochrany archívnych dokumentov pred odcudzením a poškodením.

Umožnenie prístupu k archívny dokumentom v rámci bádateľskej a správnej agendy si vyžaduje takmer 30% všetkého času potrebného na zabezpečenie odborných archívnych činností.

Štátne archívy nemajú žiadny multimediálny formát so základnými informáciami o svojej činnosti a archívnych súboroch na prezentačné účely a systematickú prípravu pútavého digitálneho obsahu pre širokú verejnosť.

Najdôležitejšou súčasťou digitalizácie je okrem skenovania stoviek miliónov dokumentov vyextrahovanie textu zo skenov aj samotný spôsob ich sprístupnenia verejnosti.

Bez nasadenia výkonných technologických nástrojov je teda sprístupnenie všetkých dostupných archívnych materiálov verejnosti drahé, neefektívne a s vysokým rizikom znehodnotenia, či odcudzenia cenných historických predmetov, či písomností.

#### 4.6.4.4 Hardvérová infraštruktúra

Pre ukladanie zdigitalizovaných archívnych dokumentov je potrebné spočítať celkový objem ukladaných dát, požadovanú priepustnosť systému a tomu zodpovedajúci požadovaný výkon pre spracovanie zdigitalizovaných a e-born archívnych dokumentov. Dôležitým parametrom je čas potrebný na digitalizáciu existujúceho obsahu.

Digitalizácia predpokladá spracovanie 50 kilometrov z celového aktuálneho objemu 200 kilometrov nespracovaných archívnych dokumentov. Pri predpoklade 40 MB na jednu stranu archívneho dokumentu pri vysokom rozlíšení je výsledkom 20 PB dát, čo predstavuje neakceptovateľne vysoký objem dát a tomu zodpovedajúce náklady. Preto bol zvolený koeficient, ktorý umožní optimalizovať zdroje, potrebné na digitalizáciu požadovaného rozsahu. Skúsenosti s masovou digitalizáciou na Slovensku opodstatňujú použitie tohto koeficientu. Pre účely digitalizácie štátnych archívov je použitý koeficient 0,3. To znamená, že 30 percent obsahu bude digitalizované vo vysokom rozlíšení a zvyšných 70 percent v nižšom rozlíšení (10 MB na stranu). Optimalizácia nemá vplyv na povinnosť predkladať originálny dokument na právne úkony. Pri prípadnej zmene legislatívy bude potrebné zohľadniť použiteľnosť digitalizovaných dokumentov s nižším rozlíšením.

Pri použití horeuvedených parametrov je celkový ročný objem ukladaných dát 1,14PB. V diapazóne 10 rokov to predstavuje 11,5 PB digitalizovaného archívneho obsahu. Vzhľadom na skúsenosti z projektov, realizovaných v slovenských podmienkach, je odporúčané použiť etapovitý prístup a predovšetkým v oblasti médií začať s kapacitou na úrovni 25 percent cieľovej kapacity. Pre účely dlhodobej archivácie sa odporúča použiť technológie s čo najdlhšou dobou životnosti. Jednotlivé komponenty musia byť škálovateľné do 150 percent 10-ročného horizontu projektu.

Architektúra predpokladá digitalizáciu v 10 centrách a finálnym centralizovaným spracovaním. Ukladanie dát bude hierarchické na diskovom poli s kapacitou 10 percent cieľového objemu uložených dát. Celkový výkon počítačov, ktoré spracovávajú digitalizované dokumenty musí zodpovedať požiadavkám na priepustnosť systému. Celkový systém musí byť odolný a schopný režimu disaster recovery (Návrat do pôvodného stavu v prípade poruchy).

Navrhovaná digitalizačná technika pre potreby digitalizácie archívnych dokumentov v štátnych archívoch:

##### **Digitalizácia viazaných dokumentov:**

*Knižný skener pre viazané dokumenty do formátu A1+*

1. Knižný skener s vysokým rozlíšením, motorickou knižnou kolískou, prítlačným sklom, nožným spínačom, ovládacím SW;
2. Skener slúži na digitalizáciu viazaných dokumentov, kníh, script, zväzkov, máp, výkresov a ďalších dokumentov do formátu A1+. Operátor chráni originál dokumentu a získava optimálnu reprodukčnú kvalitu;
3. Rozmery digitalizovaných dokumentov: 850x635mm;
4. Optické rozlíšenie: 600x600dpi;
5. Rýchlosť skenovania: 2.8 sekúnd pre DIN A1 pri 400dpi;
6. Farebná hĺbka: 48 bit color, 16 bit gray scale (škála šedi);
7. Výstup skenov: 24 bit color, 8 bit gray scale, bitonal.

##### **Príklad: Bookeye5 V1A.**

*Knižný skener formátu A1 pre veľmi hrubé knihy, skenovacia plocha 914 x 1270 mm*

1. Profesionálny produkčný skener A1 s knižnou kolískou na digitalizáciu viazaných dokumentov, máp a plánov do formátu A1+;
2. Veľkoplošný skener s prítlačným sklom a knižnou kolískou;
3. Formát predlohy: minimálne 914mm x 635mm;
4. Hrúbka predlohy: do 35 cm (možnosť upgradu do hrúbky 50cm);
5. Váha predlohy do 25kg (možnosť upgradu do váhy 35kg);
6. Optické rozlíšenie min. 600 dpi - stabilné rozlíšenie nezávislé od formátu predlohy;
7. Farebná hĺbka minimálne 48bit interne, 24bit externe;
8. Rýchlosť skenovania: 6 sec. pri 300dpi farbe, 12 sec. pri 600 dpi farbe na celú plochu;
9. Snímanie pomocou CCD kamery;
10. Systém denného svetla, teda osvetlenie okolia nemá žiadny vplyv na kvalitu skenovania;
11. Možnosť upgradu na plno robotický skener s automatickým otáčaním strán pre predlohy do A1;
12. Osvetlenie – biele LED svetlo;



13. Minimálna manipulácia so skenovaným dokumentom a jeho zvýšená ochrana keďže ide o cenné historické dokumenty, ovládanie vertikálneho pohybu knižnej koľsky nožným alebo ručným spínačom, horizontálny výjazd knižnej koľsky pred skener, prítlačné sklo.

Príklad: SMA Book scanner Scan Master 1.

#### **Pracovisko pre digitalizovanie voľných listov:**

*Vysokovýkonný prietahový dokumentový skener na voľné listy do formátu A3+:*

1. Dokumentový obojstranný skener pre neviazané dokumenty a voľné listy s hromadným podávačom s nastaviteľnými ultrazvukovými senzormi pre kontrolu dvojítych papierov;
2. Kamerový systém: 2x trilineárna CCD kamera v prachotesnom puzdre;
3. Prechod papiera skenerom: na výber 180 ° otočenie, alebo rovný prechod bez ohnutia pomocou Bypass funkcie, vhodný najmä pre hrubé listy, kartičky, poskladané dokumenty;
4. Kontrola papiera na celej dráhe transportu skenerom;
5. Výstupné formáty: komprimovaný a nekomprimovaný TIFF a JPG, PDF;
6. Systém skenera chránený heslom;
7. Ergonomická pracovná plocha s rýchlym prístupom ku všetkým transportným systémom;
8. Multistream skenovanie – v jednom kroku skener dodáva výstup vo farbe, v odtieňoch šedej a čierno-bielej;
9. Možnosť ovládať základné funkcie skenera ako sú Štart, Stop, Pauza pomocou dotykového displeja na skeneri bez nutnosti použitia PC;
10. Softvérový ovládač: TWAIN, ISIS;
11. Spoľahlivé oddeľovanie listov spracovávaných dokumentov;
12. Možnosť skenovať poskladané papiere a papiere so spinkami;
13. Číselné parametre:
  - rýchlosť skenovania: min. 220 A4 listov/min. pri rozlíšení 200 a 300 dpi;
  - možnosť zníženia rýchlosti na 100, 130, 170 A4 listov/min.;
  - minimálny formát papiera: 60 x 100 mm;
  - maximálna skenovacia plocha: 315mm x nekonečná dĺžka;
  - maximálny formát papiera: 320 mm x nekonečná dĺžka;
  - váha papiera: 30-260 g/m<sup>2</sup>;
  - možnosť skenovať kartón do hrúbky min. 2 mm.;
  - podávač pre min. 500 listov s funkciou pamäte vložených listov;
  - počet senzorov pre kontrolu dvojítych papierov: min. 3;
  - optické rozlíšenie: min. 600 dpi;
  - výstupné rozlíšenie: 75, 100, 150, 200, 300, 400, 600 dpi;
  - výstupná farebná hĺbka: min. 24 bit pre farbu, 8 bit pre odtiene šedej, čiernobiela;
  - výstupná USB konektivita: min. 1x USB 3.0 port.

Príklad: XINO S722.

*Plochý príručný A3 skener určený ako druhý skener k prietahovému skeneru:*

1. Formát A3;
2. skenovanie od kraja ;
3. určený napr. knihy a zviazané dokumenty, fotografie, čiže pre dokumenty, ktoré nie sú nevhodné do prietahového skenera.

Príklad: Skener Avision FB 6280E.

#### **Digitalizovanie veľkých listov a máp:**

*Veľkoformátový skener na historické mapy do veľkosti A0+ so skenovacou plochou 1270 x 915mm*

1. Skener s plochým lôžkom pre skenovanie máp, dokumentácií, výkresov, plánov, krehkých originálov atď.;
2. Skenovacia plocha: A0+ (1270 x 915mm);
3. Skenovací mód: Farba / Škála šedi / Čierno-bielo;
4. Minimálne optické rozlíšenie: 600 dpi;
5. Skener musí byť vrátane operačného softvéru;
6. funkcie Roll away / ľahkého prenosu na iné miesto;
7. nožného spínača;
8. skladateľný do vertikálnej polohy kvôli ľahkému prenosu;
9. s implementovanou 3D technológiou (zachytáva štruktúru 3D objektov-pečatí, mincí, a pod.);

Príklad: SMA Versascan3650 A0 Flatbed Scanner.+

*Veľkoformátový skener 2 x A0, skenovacia plocha 2540 x 914 mm*

1. denná dávka spracovaných dokumentov: 300 listov;
2. Rozsah skenovania 914 x 2540 mm (36 x 100 palcov);
3. Optické rozlíšenie 600 ppi v celom rozsahu skenovania;
4. 48 bitová farebná hĺbka;
5. IP zariadenie s individuálnou IP adresou;
6. Rolovací držiak s funkciou zloženia;
7. Dostupné aj s motorickým nastavením výšky;

8. Bez IR a UV žiarenia;
9. Virtuálna 3D technológia (V3D);
10. Obojsmerné skenovanie.

Príklad: Double A0 Flatbed Scanner VERSASCAN 36100.

*Prietaňový skener jednostranný*

1. veľkoformátový prietaňový skener, jednostranný;
2. rozlíšenie 1200x600dpi;
3. maximálna skenovacia šírka: 1220mm;
4. maximálna šírka vloženého dokumentu: 1270mm;
5. maximálna hrúbka dokumentu: 3mm;;
6. automatické rozpoznanie veľkosti
7. obslužný softvér.

Príklad: WideTEK 48 CCD.

#### **Digitalizovanie starých mikrofilmov:**

*Mikrofilmový skener*

Filmový skener na digitalizáciu 16 a 35mm mikrofilmov, mikrofišov, mikrofilmových kariet, farebných negatívov a pozitívov.

1. farebný CCD senzor 14,22 megapixelov;
2. biele LED svetlo;
3. interne 36bit;
4. maximálna skenovacia veľkosť: 47mm x 34mm;
5. výstup: 24-bit farba, 8-bit škála šedi, 1-bit čiernobielo;
6. skenovacie rozlíšenie: 75/100/150/200/240/300/400/600 dpi;
7. zväčšenie objektívu: optický zoom: 6.8 – 54x, elektrický zoom: 54 – 150x;
8. zaostrovanie: automatické a manuálne;
9. rýchlosť skenovania: 1 s;
10. podporované formáty: JPEG, TIFF, PNG, BMP, PDF, TIFF Multi-Page, PDF Multi-Page s/bez OCR;
11. výstupný formát: A2, A3, A4, škálovateľný (Portrait/Landscape);
12. rozhranie: USB 2.0 High Speed;
13. vrátane ovládacieho a skenovacieho softvéru;

Príklad: Skener SL1000.

*Samoobslužný stojanový skener pre verejnosť:*

Formát A3, skenovanie od kraja, určený napr. knihy a zviazané dokumenty, aj voľné listy.

Špecifikácie skenera:

1. Technológia skenovania 12-megapixelový snímač CMOS;
2. Maximálne rozlíšenie 4032 x 3024 pixelov;
3. Veľkosť dokumentu A3 (420 x 290 mm/na šírku);
4. Rýchlosť skenovania 1,5 sekundy na stranu vo farebnom režime - 30 str./min;
5. Automatické zaostrenie;
6. Vstavané LED svetlá 4ks;
7. Výstupné rozlíšenie 300 DPI;
8. Výstupná úroveň Farba / Čiernobiela / Odtiene šedej;
9. Rozhranie/ovládače USB 2.0 typu B x 1 (pre pripojenie k počítaču);
10. USB 2.0 Type-A x 1 (pre USB rozšírenie);
11. Napájanie 5V cez USB;
12. Výstupné formáty dokumentu: JPG, PDF (obrázok), PDF (vyhľadateľný text s OCR), PDF (Txt), Word, Text, Excel. Elektronická kniha: EPUB;
13. Rozpoznanie väzby kníh, automatické odstránenie prstov na okraji skenov;
14. Video: AVI, MPEG, FLV, WMV.

Príklad: Canon skener IRISCAN DESK 5 PRO.

## **4.6.5 Požiadavky na výkonnostné parametre, kapacitné požiadavky**

Nižšie sú uvedené požiadavky na výkonnostné a kapacitné parametre systému v TO BE stave:

Parameter	Jednotky	Predpokladaná hodnota
Počet interných používateľov	Počet	4
Počet súčasne pracujúcich interných používateľov v špičkovom zaťažení	Počet	4
Počet externých používateľov (internet)	Počet	20

Počet externých používateľov používajúcich systém v špičkovom zaťažení	Počet	5
Počet transakcií (podaní, požiadaviek) za obdobie	Počet/obdobie	100/rok
Objem údajov na transakciu	Objem/transakcia	N/A
Objem existujúcich kmeňových dát	Objem	N/A
Počet externých používateľov prístupujúcich cez mobilnú aplikáciu	Počet	5000

#### Prehľad vybraných kapacitných a výkonových požiadaviek – budúci stav

Vzhľadom na odporúčanie realizovať projekt vo dvoch fázach, predpokladáme pre realizáciu pilotnej fázy vybudovanie pilotného digitalizačného pracoviska. Finančná alokácia je 15 percent z nákladov na digitalizačnú techniku. Presnú špecifikáciu určí analýza potrieb pilotného riešenia.

## 4.6.6 Návrh riešenia technologickej architektúry

### HW, SW a komunikačná infraštruktúra

Nakoľko je počítačové vybavenie pracovníkov archívov zastaralé, uvažuje sa so zabezpečením nového vybavenia (vstupné zariadenia) pre digitalizačné pracoviská. Operátor bude pristupovať cez webové rozhranie, využívajúc existujúce štandardné internetové pripojenie.

#### HW infraštruktúra prostredia – centrálny segment (backend)

Riešenie počíta s prevádzkou troch serverových prostredí na centrálnej kontajnerizačnej platforme s nasledovnými základnými požiadavkami:

Server	Počet	Processor (vCPU)	Operačná pamäť (RAM)	Pevný disk	Popis
Aplikačný server	1	32	2TB	3,2TB	Server primárne určený pre aplikáciu (backend, frontend) a proxy, ktorý bude zabezpečovať bezpečné spojenie používateľov a aplikácií s celým riešením
Databázový server	1	32	2TB	3,2TB	Server určený na trvalé ukladanie dát
Monitoring	1	32	2TB	3,2TB	Server určený pre podporne aplikácie slúžiace na monitoring a reporting systému a virtuálnych serverov

Toto prostredie bude zrkadlené do záložného výpočtového strediska.

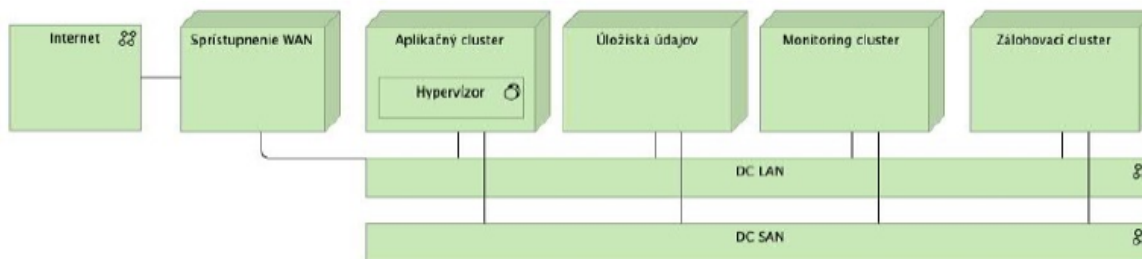
Jednotlivé navrhované servery budú bežať pod operačným systémom Linux. Pripojenie výpočtových stredísk bude realizované prostredníctvom DWDM multiplexorov vzhľadom na výmenu veľkých objemov dát. Bude doplnená existujúca DWDM infraštruktúra prepojenia oboch DC.

Vzhľadom na odporúčanie realizovať projekt vo dvoch fázach, predpokladáme pre realizáciu pilotnej fázy vybudovanie pilotného HW prostredia.

Pilotné prostredie bude vytvorené bez záložného pracoviska iba na jednej lokalite. Pozostáva z 2 serverov, špecifikovaných v horeuvedenej tabuľke, pričom realizácia aplikačného, databázového a monitorovacieho prostredia bude na virtuálnej úrovni. Pre overenie funkcionality sú potrebné k dispozícii prevádzkové úložné kapacity diskového poľa a cieľové riešenie páskovej knižnice s polovičnou kapacitou médií. Všetky ďalšie systémy koncového riešenia budú implementované v druhej fáze.

#### SW infraštruktúra – centrálny segment (backend)

Odporúča sa prevádzka virtuálnych serverov na virtualizačnej platforme VMware vSphere. Na správu serverov sa používa VMware vCenter. Pre vCenter použijeme operačný systém Windows Server.



Model technologickej architektúry

## 4.6.7 Využívanie služieb z katalógu služieb vládneho cloudu

Riešenie vzhľadom na charakter projektu nebude využívať služby vládneho cloudu.

## 4.6.8 Jazyková lokalizácia

Požaduje sa slovenská a anglická jazyková lokalizácia.

## 4.7 Bezpečnostná architektúra

V kontexte bezpečnosti boli identifikované nasledujúce globálne a MV SR štandardy:

*Globálne štandardy:*

1. STN ISO/IEC 27002: Informačné technológie, Bezpečnostné metódy, Pravidlá dobrej praxe riadenia informačnej bezpečnosti
2. Zákon č. 95/2019 Z. z. o informačných technológiách vo verejnej správe a o zmene a doplnení niektorých zákonov

*MV SR štandardy:*

1. Bezpečnostný štandard pre operačný systém AIX a HMC
2. Bezpečnostné princípy pre siete Datacentra a DMZ MV SR
3. Bezpečnostný projekt ISEA

Najdôležitejšie oblasti bezpečnosti, ktoré bude Informačného systému pre digitalizáciu archívnych dokumentov (ISDAD) využívať, sú nasledujúce:

### **Identifikácia a autentizácia používateľov**

Na identifikáciu a autentizáciu používateľov sa budú využívať existujúce služby na podporu práce interných a externých používateľov tak ako je definované v Bezpečnostnom projekte ISEA, pričom všetky ostatné aspekty rozšírenia riešenia vypracuje dodávateľ v rámci realizácie projektu.

### **Dátové centrá a fyzická bezpečnosť**

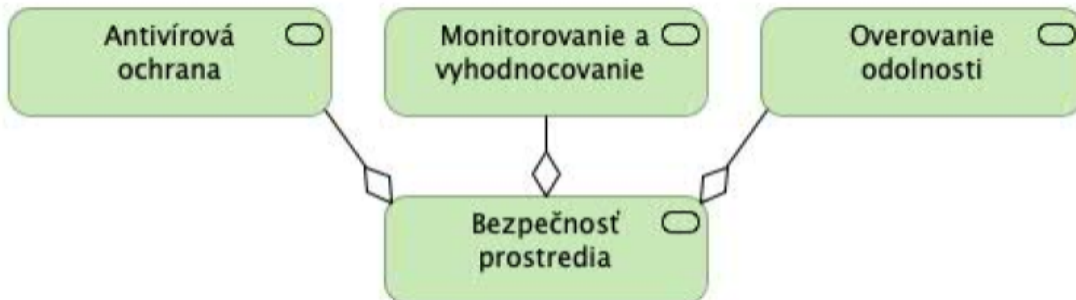
Systém bude prevádzkovaný v existujúcich dátových centrách MV SR. Bezpečnostné opatrenia súvisiace s dátovými centrami (napr. fyzická bezpečnosť, prepojenie dátových centier z hľadiska dostupnosti služieb archívu a pod.) sú týmto predpokladom dané.

### **Sieťová infraštruktúra**

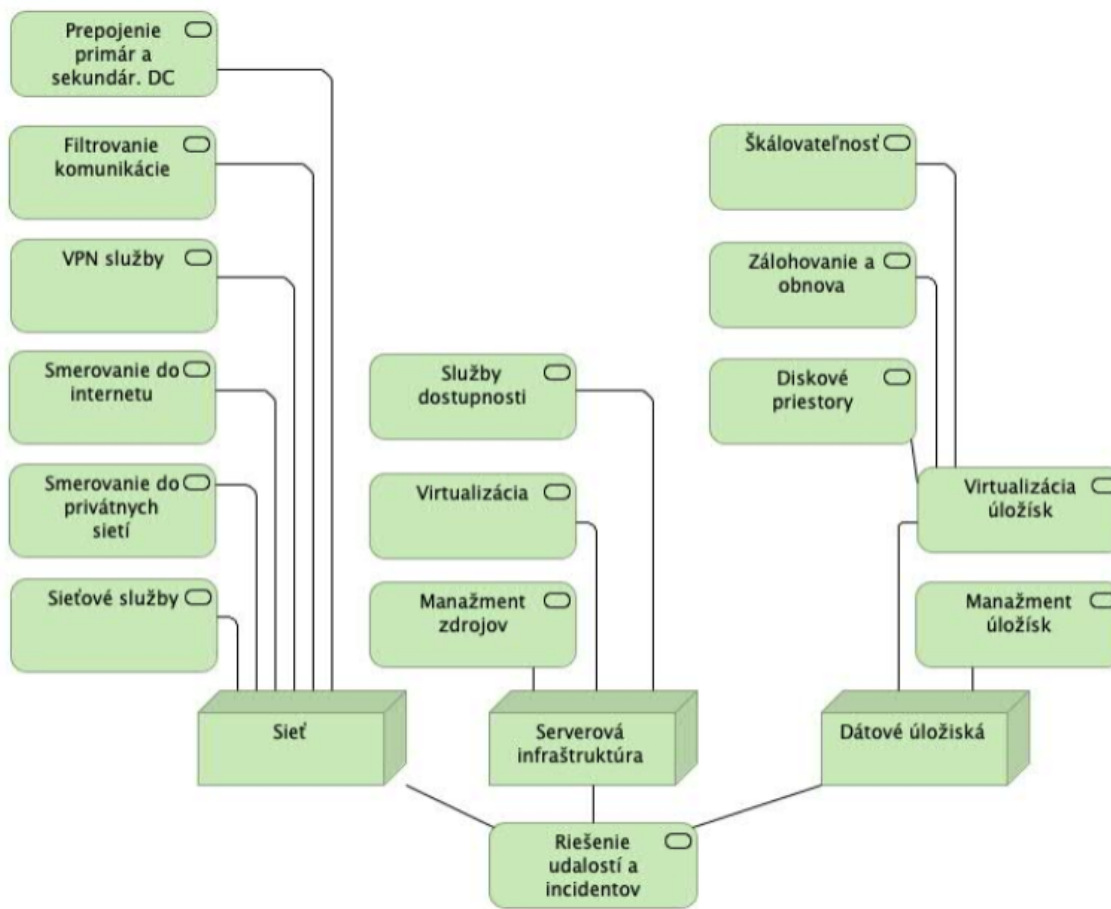
Rovnako, ako v prípade dátových centier, aj sieťová infraštruktúra a s ňou súvisiace mechanizmy bezpečnosti preberú existujúcu architektúru riešení v súlade s Bezpečnostným projektom ISEA.

Všetky rozhrania si budú vyžadovať pripojenie pomocou SSL. Zabezpečený bude monitoring sieťových prístupov, bezpečnosti údajov na diskových poliach, logovanie prístupov a zmien, ako aj služba poskytovania bezpečnej prístupovej siete. V rámci samotného IS budú využívané analytické nástroje pre monitorovanie a vyhodnocovanie bezpečnosti. V rámci IKT vybavenia bude zabezpečené nástroje pre ochranu proti škodlivému softvéru. IKT vybavenie v rámci jednotlivých pracovísk bude využívať VPN prepojenie. Pred spustením IS do prevádzky budú realizované penetračné testy.

Povinnosťou bude preukázať súlad so zákonom č. 95/2019 zákona o informačných technológiách vo verejnej správe a o zmene a doplnení niektorých zákonov. Pre úspešnú realizáciu projektu je potrebné zabezpečiť dodržanie pravidiel stanovených Vyhláškou č. 78/2020 (resp. jej novelizácií) Z. z. o štandardoch pre informačné technológie verejnej správy. Z hľadiska ochrany osobných údajov bude dátový manažment realizovaný v súlade so zákonom č. 18/2018 Z.z. o ochrane osobných údajov a o zmene a doplnení niektorých zákonov. Implementácia a prevádzka systému musí v oblasti bezpečnosti brať do úvahy aj Zákon 69/2018 Z. z. o kybernetickej bezpečnosti, v znení neskorších predpisov. Bude vypracovaný bezpečnostný projekt rešpektujúci tieto pravidlá.



Model bezpečnostnej architektúry



Model bezpečnostnej architektúry – bezpečnosť prostredia

#### 4.7.1 Riziká pri súčasnom stave online zverejňovania

MV SR nedisponuje dostatočnými finančnými zdrojmi pre zabezpečenie pravidelnej výmeny a údržby hardvéru, na ktorom ISEA funguje a ani rozvoj nových funkcionalít a integrácií s inými IS. Servisné zásahy zo strany dodávateľa alebo sekcie informatiky, telekomunikácie a bezpečnosti MV SR sa týkajú iba odstraňovania porúch vo funkčnosti systému. Vzhľadom k tomu, že projekt vybudovania ISEA nepočítal s ďalšími finančnými prostriedkami počas obdobia udržateľnosti projektu, absentoval akýkoľvek rozvoj IS a dopracovávanie nových funkcionalít v nadväznosti na integráciu s ďalšími IS. V súčasnosti po uplynutí doby udržateľnosti ISEA nemá uzatvorenú servisnú zmluvu s jeho dodávateľom pre operatívne riešenie a odstraňovanie vzniknutých technických problémov. Na jeho softvérové súčasti nebol vykonaný update na vyššie verzie, ani upgrade na novšie verzie, čím nedošlo k čiastkovým vylepšeniam alebo odstráneniu chýb.

## 5. ZÁVISLOSTI NA OSTATNÉ ISVS / PROJEKTY

Stakeholder	Kód projektu (z MetaIS)	Názov projektu	Termín ukončenia projektu	Popis závislostí
MV SR	projekt_46	Elektronický archív Ministerstva vnútra Slovenskej republiky (OPIS)	2017	Bez priamej závislosti.

Prehľad projektov, ktoré sú v štádiu vývoja a v korelácii s pripravovaným projektom.

Prostredníctvom ISEA nebudú publikované služby pre spracovanie elektronických podaní, (a súvisiace (pomocné) služby pre úspešné vyplnenie a prípravu podania) do multikanálového prostredia (publikáciu na API GateWay).

## 6. ZDROJOVÉ KÓDY

Vzhľadom na skutočnosť, že na trhu v súčasnosti neexistuje hotové, plne vyhovujúce COTS riešenie pre digitalizáciu a elektronickú archiváciu, predpokladáme, že bude obstarané riešenie, resp. úprava a doplnenie systému ISEA aplikačne aj technologicky, čo je v zásade HW/SW riešenie na mieru, ako aj COTS SW pre potreby centrálného segmentu (operačný systém, databázový systém). Štátny archív obstará licencie na jeho vývoj, používanie a udržiavanie.

Obstaranie a užívanie COTS SW sa bude riadiť štandardnými zmluvnými podmienkami dodávateľa. Pokiaľ ide v celku, alebo čiastočne o na mieru vyvinuté SW riešenie, ktoré bude zohľadňovať špecifické potreby a podmienky Štátneho archívu. Zdrojový kód, vytvorený počas zhotovovania, bude otvorený v súlade s licenčnými podmienkami verejnej softvérovej licencie Európskej únie podľa osobitného predpisu a to v rozsahu, v akom zverejnenie tohto kódu nemôže byť zneužitá na činnosť smerujúcu k narušeniu alebo k zničeniu informačného systému. Zmluva s dodávateľom bude pripravená tak, aby po skončení zmluvného vzťahu, v rámci ktorého bolo unikátne SW dielo vytvorené a po istú dobu prevádzkované pôvodným dodávateľom, disponoval Štátny archív všetkými oprávneniami potrebnými pre ďalšiu prevádzku a rozvoj tohto diela bez závislosti na pôvodnom dodávateľovi, teda i prostredníctvom nového dodávateľa vybraného v neobmedzenej súťaži dodávateľov.

## 7. PREVÁDZKA A ÚDRŽBA

### 7.1 Požiadavky na podporu, údržbu a rozvoj

Požaduje sa poskytovanie služieb podpory prevádzky, údržby a rozvoja IS ISEA a všetkých podporných modulov po dobu 5 rokov od odovzdania diela do prevádzky v rozsahu popísanom v nasledujúcich podkapitolách.

Prevádzka projektu bude zabezpečená za rovnakých podmienok ako prevádzka súčasných informačných systémov prevádzkovaných v DC MV SR a za parametrov uvedených v ďalších podkapitolách.

Úroveň L1 a L2 bude zabezpečená internými zamestnancami MV SR. L3 bude zabezpečená na základe Zmluvy o podpore riešenia - SLA (Service Level Agreement) spĺňajúcej parametre SLA uvedené v ďalších podkapitolách.

#### 7.1.1 Úrovně podpory používateľov

1. **L1 podpora IS** (Level 1, priamy kontakt zákazníka) - jednotný kontaktný bod verejného obstarávateľa – Call Centrum MV SR.
1. **L2 podpory IS** (Level 2, postúpenie požiadaviek od L1) - vybraná skupina garantov a oddelenie aplikačnej podpory OA SITB MV SR, so znalosťou IS (zabezpečuje prevádzkovateľ IS – verejný obstarávateľ).
1. **L3 podpory IS** (Level 3, postúpenie požiadaviek od L2) - na základe zmluvy o podpore IS (zabezpečuje úspešný uchádzač).

#### 7.1.2 SLA parametre

Dostupnosť L3 podpory pre IS je 8x5 (8 hodín x 5 dní od 8.00 do 16:00 počas pracovných dní).

##### Riešenie incidentov – SLA parametre

Za incident je považovaná chyba IS, t.j. správanie sa v rozpore s prevádzkovou a používateľskou dokumentáciou IS. Za incident nie je považovaná chyba, ktorá nastala mimo prostredia IS napr. výpadok poskytovania konkrétnej služby alebo komunikačnej infraštruktúry.

##### 1. Označenie naliehavosti incidentu:

Označenie naliehavosti incidentu	Závažnosť incidentu	Popis naliehavosti incidentu
A	Kritická	Kritické chyby, ktoré spôsobia úplné zlyhanie systému ako celku a nie je možné používať ani jednu jeho časť, nie je možné poskytnúť požadovaný výstup z IS.
B	Vysoká	Chyby a nedostatky, ktoré zapríčinia čiastočné zlyhanie systému a neumožňuje používať časť systému.
C	Stredná	Chyby a nedostatky, ktoré spôsobia čiastočné obmedzenia používania systému.
D	Nízka	Kozmetické a drobné chyby.

možný dopad:

Označenie závažnosti incidentu	Dopad	Popis dopadu
1	katastrofický	katastrofický dopad, priamy finančný dopad alebo strata dát,
2	značný	značný dopad alebo strata dát
3	malý	malý dopad alebo strata dát

## 2. Výpočet priority incidentu

Matica priority incidentov		Dopad		
		Katastrofický - 1	Značný - 2	Malý - 3
Naliehavosť	Kritická - A	1	2	3
	Vysoká - B	2	3	3
	Stredná - C	2	3	4
	Nízka - D	3	4	4

### Vyžadované reakčné doby:

Označenie priority incidentu	Reakčná doba <sup>(1)</sup> od nahlásenia incidentu po začiatok riešenia incidentu	Doba konečného vyriešenia incidentu od nahlásenia incidentu (DKVI) <sup>(2)</sup>	Spôľahlivosť <sup>(3)</sup> (počet incidentov za mesiac)
1	8 hod.	48 hodín	1
2	24 hod.	140 hodín	2
3	24 hod.	280 hodín	10
4	48 hod.	Vyriešené a nasadené v rámci plánovaných release	

- (1) Reakčná doba je čas medzi nahlásením incidentu verejným obstarávateľom (vrátane užívateľov IS, ktorí nie sú v pracovnoprávnom vzťahu s verejným obstarávateľom) na helpdesk úrovne L3 a jeho prevzatím na riešenie.
- (2) DKVI znamená obnovenie štandardnej prevádzky - čas medzi nahlásením incidentu verejným obstarávateľom a vyriešením incidentu úspešným uchádzačom (do doby, kedy je funkčnosť prostredia znovu obnovená v plnom rozsahu). Doba konečného vyriešenia incidentu od nahlásenia incidentu verejným obstarávateľom (DKVI) sa počíta počas celého dňa. Do tejto doby sa nezaráta čas potrebný na nevyhnutnú súčinnosť verejného obstarávateľa, ak je potrebná pre vyriešenie incidentu. V prípade potreby je úspešný uchádzač oprávnený požadovať od verejného obstarávateľa schválenie riešenia incidentu.
- (3) Maximálny počet incidentov za kalendárny mesiac. Každá ďalšia chyba nad stanovený limit spoľahlivosti sa počíta ako začatý deň omeškania bez odstránenia vady alebo incidentu. Duplicitné alebo technicky súvisiace incidenty (zadané v rámci jedného pracovného dňa, počas pracovného času 8 hodín) sú považované ako jeden incident.
- (4) Incidenty nahlásené verejným obstarávateľom úspešnému uchádzačovi v rámci testovacieho prostredia:
  - Majú prioritu 3 a nižšiu
  - Vzťahujú sa výhradne k dostupnosti testovacieho prostredia
  - Za incident na testovacom prostredí sa nepovažuje incident vzťahujúci k práve testovanej funkcionalite.

Vyššie uvedené SLA parametre nebudú použité pre nasledovné služby:

- Služby systémovej podpory na požiadanie (nad paušál)
- Služby realizácie aplikačných zmien vyplývajúcich z legislatívnych a metodických zmien (nad paušál)

Pre tieto služby budú dohodnuté osobitné parametre dodávky.

## 7.2 Požadovaná dostupnosť IS

Popis	Parameter	Poznámka
Prevádzkové hodiny	12 hodín	od 6:00 hod. - do 18:00 hod. počas pracovných dní
Servisné okno	10 hodín	od 19:00 hod. - do 5:00 hod. počas pracovných dní
	24 hodín	od 00:00 hod. - 23:59 hod. počas dní pracovného pokoja a štátnych sviatkov Servis a údržba sa bude realizovať mimo pracovného času.
Dostupnosť produkčného prostredia IS	98,5%	<ol style="list-style-type: none"> <li>98,5% z 24/7/365 t.j. max ročný výpadok je 66 hod.</li> <li>Maximálny mesačný výpadok je 5,5 hodiny.</li> <li>Vždy sa za takúto dobu považuje čas od 0.00 hod. do 23.59 hod. počas pracovných dní v týždni.</li> <li>Nedostupnosť IS sa počíta od nahlásenia incidentu Zákazníkom v čase dostupnosti podpory Poskytovateľa (t.j. nahlásenie incidentu na L3 v čase od 6:00 hod. - do 18:00 hod. počas pracovných dní). Do dostupnosti IS nie sú započítavané servisné okná a plánované odstávky IS.</li> <li>V prípade nedodržania dostupnosti IS bude každý ďalší začatý pracovný deň nedostupnosti braný ako deň omeškania bez odstránenia vady alebo incidentu.</li> </ol>

## 8. POŽIADAVKY NA PERSONÁL

Objednávateľ zostaví Riadiaci výbor (RV), a to v zložení:

1. Predseda RV
2. Podpredseda RV – zástupca vlastníka procesov objednávateľa
3. Zástupca kľúčových používateľov
4. Projektový manažér za ŠA MV SR
5. Projektový manažér - zástupca dodávateľa (bude doplnený až po verejnom obstarávaní).

ID	Meno a Priezvisko	Pozícia	Oddelenie	Rola v projekte
1.		Generálny riaditeľ SITB MV SR	Sekcia informatiky, telekomunikácií a bezpečnosti MV SR	Predseda RV
2.		Riaditeľ OA, SITB MV SR	Odbor aplikácií, sekcie informatiky, telekomunikácií a bezpečnosti MV SR	Podpredseda RV – zástupca vlastníkov procesov
3.		Riaditeľka OAR, SVS MV SR	Odbor archívov a registratúr, sekcie verejnej správy MV SR	Zástupca kľúčových používateľov
4.		Vedúca oddelenia RAaR, OAR, SVS MV SR	Oddelenie riadenia archívov a registratúr, odboru archívov a registratúr, sekcie verejnej správy MV SR	Zástupca kľúčových používateľov / portfólio ŠA
5.		Starší referent špecialista, OŠpIS, OA, SITB MV SR	Oddelenie štandardov pre informačné systémy, odboru aplikácií, sekcie informatiky, telekomunikácií a bezpečnosti MV SR	Projektový manažér MV SR (bez hlasovacieho práva)
6.	Bude definované	Bude definované	Bude definované	Projektový manažér za Dodávateľa (bez hlasovacieho práva)

Po spustení realizačnej fázy bude zostavený pracovný Projektový tím, pričom budú definované nasledovné pozície:

Aktivita	ID	Rola v projekte	Organizácia
Hlavné aktivity	1.	Finančný manažér	Bude definované
	2.	IT Analytik	Bude definované
	3.	IT Architekt	Bude definované
	4.	IT programátor / vývojár	Bude definované
	5.	IT tester	Bude definované
	6.	Manažér kvality	Bude definované



	7.	Dátový kurátor	Bude definované
	8.	Kľúčový používateľ	Bude definované
	9.	Vlastník procesov	Bude definované
	10.	Dátový špecialista	Bude definované
	11.	Špecialista pre bezpečnosť IT	Bude definované
	12.	IT/IS konzultant	Bude definované
	13.	Špecialista pre infraštruktúru	Bude definované
	14.	Špecialista pre databázy	Bude definované
	15.	Odborník pre IT dohľad, QA	Bude definované
	16.	Školiteľ pre IT systémy	Bude definované
	17.	Špecialista na publicitu	Bude definované
	18.	Pracovník digitalizácie	Bude definované
	19.	Podporné odborné služby	Bude definované
Podporné aktivity	1.	Projektový manažér	MV SR
	2.	Finančný manažér	Bude definované
	3.	Administratívny pracovník	Bude definované
	4.	Špecialista na publicitu	Bude definované

Mená členov projektového tímu budú doplnené pred zahájením realizačnej fázy projektu, na základe rozhodnutia riadiaceho výboru a návrhu predloženého projektovým manažérom.

## 8.1 Pracovné náplne

Pre potreby riadenia projektu bude vytvorený riadiaci výbor projektu a vytvorený projektový tím prijímateľa. Projekt bude realizovaný ako internými, tak aj externými kapacitami dodávateľa. Štátny archív s podporou rezortu disponuje dostatočnými personálnymi kapacitami pre obsadenie nasledovných rolí v riadiacom výbore, resp. projektových rolí:

**Predseda RV** - Hlavným záujmom a zodpovednosťou predsedu Riadiaceho výboru projektu je:

1. zastupovať záujmy prijímateľa v projekte,
2. kontrolovať súlad projektu a projektových cieľov so strategickými cieľmi,
3. zabezpečiť a udržať finančné krytie (rozpočet) realizácie projektu,
4. zabezpečiť nákladovo prijateľný prístup v projekte.

**Podpredseda RV** – zástupca vlastníkov procesov - Hlavným záujmom a zodpovednosťou zástupcu vlastníkov procesov (biznis vlastník) je:

1. schválenie funkčných a technických požiadaviek, potrieb, obsahu, kvalitatívnych a kvantitatívnych prínosov projektu,
2. definovanie očakávaní na kvalitu projektu, kritérií kvality projektových produktov, prínosov pre koncových používateľov a požiadaviek na bezpečnosť,
3. definovanie merateľných výkonnostných prvkov a ukazovateľov projektu,
4. schválenie akceptačných kritérií,
5. akceptáciu rozsahu a kvality dodávaných projektových výstupov pri dosiahnutí platobných míľnikov,
6. odsúhlasenie spustenia výstupov projektu do produkčnej prevádzky,
7. dostupnosť ľudských zdrojov alokovaných na realizáciu projektu.

**Zástupca kľúčových používateľov** - hlavným záujmom a zodpovednosťou zástupcu kľúčových používateľov objednávateľa (end user), ktorí reprezentuje záujmy budúcich používateľov projektových produktov alebo projektových výstupov je:

1. návrh a špecifikácia funkčných a technických požiadaviek, potreby, obsah, kvalitatívnych a kvantitatívnych prínosov projektu, požiadaviek koncových používateľov na prínos systému a požiadaviek na bezpečnosť,
2. návrh a definovanie akceptačných kritérií,
3. akceptačné testovanie a návrh na akceptáciu projektových produktov alebo projektových výstupov a návrh na spustenie do produkčnej prevádzky,
4. predkladanie požiadaviek na zmenu funkcionalít produktov.

**IT analytík** - IT analytík na projekte zabezpečuje analyzovanie procesných a ďalších požiadaviek a špecifikácií budúceho používateľa riešenia, následne navrhuje dizajn a programátorské riešenie. Aktívne sa zúčastňuje analytických stretnutí s kľúčovými používateľmi Štátneho archívu k detailnej špecifikácii požiadaviek. Participuje na vývoji nových, ale i na vylepšovaní existujúcich funkčností v rámci celého vývojového cyklu. Úzko spolupracuje s IT architektom a vykonáva tiež ďalšie činnosti vyplývajúce z požiadaviek projektu (Vzhľadom na charakter projektu bude pravdepodobne potrebný IT Analytík so znalosťou využívaného ISEA). V rámci projektu má zabezpečovať definovanie činností zabezpečujúcich ochranu IS a jeho zložiek proti bezpečnostným hrozbám a nepriateľským aktivitám a zároveň zabezpečovať technickú podporu pre databázové systémy, chod databáz alebo databázového systému.

**IT architekt** - IT architekt na projekte zabezpečuje činnosti vychádzajúce z požiadaviek organizácie, transformuje ich do konkrétnej koncepcie architektúry IS/IT. Zodpovedá za návrh a implementáciu technológií predovšetkým z pohľadu udržateľnosti, kvality a nákladov. Jeho úlohou je vytvoriť návrh technologického riešenia a odporúčanej schémy infraštruktúry, má analyzovať a navrhnúť vytvorenia služieb, definovať výstupy a postupy pre prípady použitia, má navrhnúť zmeny procesov a využitie analytických metód pre lepšie rozhodovanie, analyzovať dátové potreby, definovať požiadavky na dátové zdroje, ako aj ďalšie činnosti vyplývajúce z požiadaviek projektu.

**Špecialista pre bezpečnosť IT** - zodpovedá za dodržanie princípov a štandardov na kybernetickú a IT bezpečnosť, za kontrolu a audit správnosti riešenia v oblasti bezpečnosti. Koordinuje a riadi činnosť v oblasti bezpečnosti prevádzky IT, spolupracuje na projektoch, na rozvoji nástrojov a postupov k optimalizácii bezpečnostných systémov a opatrení. Stanovuje základné požiadavky, podmienky a štandardy pre oblasť bezpečnosti programov, systémov, databázy či siete. Spracováva a kontroluje príslušné interné predpisy a dohľada nad plnením týchto štandardov a predpisov. Kontroluje a riadi činnosť nad bezpečnostnými testami, bezpečnostnými incidentmi v prevádzke IT. Poskytuje inštrukcie a poradenstvo používateľom počítačov a informačných systémov pre oblasť bezpečnosti.

**IT programátor/vývojár** - IT programátor/vývojár transformuje návrh technického riešenia, na základe jeho detailnej špecifikácie, vývojových diagramov a návrhu dátovej integrácie, do podoby fyzického, funkčného a overeného zdrojového kódu. Zabezpečuje alebo priamo vykonáva jednotkové a funkčné testovanie a asistuje IT testerom pri vyšších úrovniach testovania. Takisto je zodpovedný za dokumentáciu zdrojového kódu tak, aby tento mohol byť ďalej využívaný a rozvíjaný nezávisle od autora kódu (tzn. od konkrétnej osoby IT programátora/vývojára, ktorý kód vytvoril), ako i za ďalšie činnosti vyplývajúce z požiadaviek na SW riešenia projektu.

**IT tester** - IT tester na projekte hľadá chyby/ nedostatky/ nezhody oproti zadaniu, v zrealizovanom technickom riešení / softwarovej aplikácii, vyhľadáva prípadné chyby v kóde s cieľom dosiahnuť čo najvyššiu kvalitu dodávaného riešenia. Testovanie prebieha podľa prípadov použitia v analýze, testovacích prípadov a scenárov odsúhlasených Riadiacim výborom projektu.

**Manažér kvality** - Špecialista kvality, resp. odborník pre IT dohľad/Quality Assurance navrhuje a implementuje do praxe také postupy, techniky, pravidlá, ktoré maximalizujú efektivitu práce a kvalitatívne parametre vývoja platformy. Zároveň definuje postupy, navrhuje a vyjadruje sa k plánom testov a testovacích scenárov. Analyzuje výsledky testovania. Komplexný prístup ku kvalite si vyžaduje jeho účasť vo všetkých fázach vývoja platformy.

**Projektový manažér** - projektový manažér riadi projekt v súlade so Zmluvou o poskytnutí NFP, usmerneniami a pokynmi poskytovateľa súvisiacimi s čerpaním fondov EÚ a tiež komunikáciu s RO. Zabezpečuje publicitu projektu a informovanosť v súlade s podmienkami vyhlášky 85/2020 Z.z. o riadení projektov, so zmluvou o poskytnutí NFP a ďalšími relevantnými riadiacimi dokumentami, a usmerneniami riadiaceho orgánu alebo sprostredkovateľského orgánu, rovnako má zodpovednosť za činnosti súvisiace s podpornými aktivitami publicity projektu.

**Finančný manažér** - pripravuje podklady pre žiadosti o platbu, pripravuje a nahráva samotné žiadosti o platbu, pripravuje podklady pre monitorovanie projektu, pripravuje a nahráva samotné monitorovacie správy, pripravuje podklady pre verejné obstarávania k projektu, pripravuje kontroly verejných obstarávaní a práce súvisiace so zmenami na projekte, komunikuje s RO a pod.

**Pracovník digitalizácie (skenerista)** – obsluhuje technické vybavenie (skenery) na digitalizačných pracoviskách, zabezpečuje praktickú digitalizáciu dokumentov, tvorbu kvalitných digitalizátov a ich základnú evidenciu. Pracuje pod odborným dohľadom archivára. Podrobnejší popis role je obsiahnutý v *Prílohe č.2 – Zdroje cien*.

**Podporné odborné služby** – Zabezpečujú dohľad nad činnosťou digitalizačných pracovísk, resp. pracovníkov digitalizácie; poradenstvo pri vývoji, implementácii, nasadení a prevádzke nových aplikácií; tvorbu metodík a štandardov týkajúcich sa digitálnych dokumentov a metadát; všestranné školenia vlastníkov a používateľov HW, SW a aplikácií; manažment zmien a kontrolu kvality procesov. Podrobnejší popis role je obsiahnutý v *Prílohe č.2 – Zdroje cien*.

## 9. IMPLEMENTÁCIA A PREBERANIE VÝSTUPOV PROJEKTU

**Rámcový harmonogram implementácie projektu:**

**Inkrement 1 (pilot):**

1. Fázy	NÁZOV FÁZY / ETAPA PROJEKTU	ODHADOVANÝ ZAČIATOK	ODHADOVANÝ KONIEC	ODHADOVANÉ TRVANIE V MESIACHOCH
Fáza #1	Prípravná fáza	04/2023	06/2023	3
Fáza #2	Iniciačná fáza	06/2023	10/2023	5
Fáza #3	Realizačná fáza	11/2023	08/2024	10
Fáza #3a	Analýza a dizajn	11/2023	03/2024	4
Fáza #3b	Nákup HW, SW a služieb	03/2024	06/2024	3
Fáza #3c	Implementácia a testovanie	03/2024	06/2024	3
Fáza #3d	Nasadenie a podpora	06/2024	08/2024	2
Fáza #4	Dokončovacia fáza	08/2024	08/2023	1

**Inkrement 2 (plné riešenie):**

1. Fázy	NÁZOV FÁZY / ETAPA PROJEKTU	ODHADOVANÝ ZAČIATOK	ODHADOVANÝ KONIEC	ODHADOVANÉ TRVANIE V MESIACHOCH
Fáza #1	Prípravná fáza	08/2024	09/2024	1
Fáza #2	Iničiačná fáza	09/2024	10/2024	1
Fáza #3	Realizačná fáza	10/2024	09/2025	12
Fáza #3a	Analýza a dizajn	10/2024	01/2025	3
Fáza #3b	Nákup HW, SW a služieb	01/2025	03/2025	3
Fáza #3c	Implementácia a testovanie	01/2025	06/2025	5
Fáza #3d	Nasadenie a podpora	06/2025	09/2025	3
Fáza #4	Dokončovacia fáza	09/2025	09/2025	1

**Preberanie výstupov projektu:**

ID	Výstup	Odhad termínu
1.	Funkčná špecifikácia riešenia	t+2
2.	Obstarané riešenie – vývoj	t+6
3.	Riešenie nasadené na testovanie	t+7
4.	Vyškolení operátori	t+9
5.	Riešenie nasadené do produkcie	t+9

Pričom t= 1.1.2023

Projekt bude realizovaný Waterfall metódou so zohľadnením spôsobu implementácie prostredníctvom vývoja na mieru. Vzhľadom na uvedené bude realizácia projektu prebiehať v rámci niekoľkých.

## 10. PRÍLOHY

- - *Príloha č.1 – Kalkulácia – Digitalizačné pracoviská*
  - *Príloha č.2 – Zdroje cien*
  - *Príloha č.3 – Východiská prínosov.*
  - *Príloha č.4 – Referencie cenotvorby*
  - *Príloha č.5 – Dodatočné referencie*
  - *Príloha č.6 – Rozpočet NFP (inkrementy)*

Koniec dokumentu