

# Digitalizace a ukládání

## Typizovaný projektový záměr

## OBSAH:

<b>1</b>	<b>ZÁKLADNÍ CHARAKTERISTIKA PROJEKTU</b>	<b>4</b>
<b>2</b>	<b>VÝCHODISKA</b>	<b>6</b>
2.1	Vize projektu	6
2.2	Cíle projektu	6
2.3	Návaznost na Technologická centra	6
2.4	Služby pro území	7
<b>3</b>	<b>ANALÝZA SOUČASNÉHO STAVU</b>	<b>8</b>
3.1	Problematika digitalizace dat	8
3.2	Problematika ukládání dat	9
<b>4</b>	<b>ANALÝZA POPTÁVKY A KONCEPCE MARKETINGU</b>	<b>11</b>
4.1	Analýza poptávky výstupů projektu	11
4.2	Marketing	11
<b>5</b>	<b>POŽADAVKY NA ŘEŠENÍ PROJEKTU</b>	<b>12</b>
5.1	Problematika digitalizace dat	Chyba! Záložka není definována.
5.2	Krajská digitalizační jednotka	13
5.3	Problematika ukládání dat	13
5.3.1	Dlouhodobé ukládání dokumentů (KDS, KDR)	14
5.3.2	Krajské digitální úložiště (KDU)	20
5.4	Architektura řešení	21
5.4.1	Softwarová architektura	21
5.4.1.1	Softwarová architektura KDS a KDR	22
5.4.1.2	Rozdíly ve funkčnosti KDS a KDR	25
5.4.1.3	Softwarová architektura KDÚ	26
5.4.1.4	Technologická architektura	27
5.4.1.5	Implementace subsystémů KDS a KDR	28
5.4.1.6	Implementace subsystému KDÚ	29
5.5	Organizační a procesní část	29
5.5.1	Studie proveditelnosti	29
5.5.2	Role kraje jako garanta projektu	30
5.5.3	Role organizací kraje	30
5.5.4	Role měst a obcí a jejich organizací	30
5.5.5	Role státu	30

<b>6</b>	<b>POSTUP REALIZACE .....</b>	<b>32</b>
<b>7</b>	<b>FINANČNÍ ANALÝZA A PLÁN .....</b>	<b>33</b>
<b>7.1</b>	<b>Přehled celkových nákladů v investiční fázi .....</b>	<b>33</b>
7.1.1	Náklady na lidské zdroje.....	33
7.1.2	Přehled celkových možných nákladů.....	33
<b>7.2</b>	<b>Příjmy provozní fáze.....</b>	<b>33</b>
<b>7.3</b>	<b>Způsobilé výdaje projektu .....</b>	<b>34</b>
<b>7.4</b>	<b>Udržitelnost projektu .....</b>	<b>34</b>

## Seznam zkratek a pojmů

<b>AIP</b>	Archive Information Package	Archivní informační balík, zahrnující ukládaný obsah a odpovídající popisné informace pro uchovávání (archivní a technické informace), ukládané uvnitř OAIS.
<b>Archiv</b>		Zařízení podle zákona č. 499/2004 Sb. ve znění pozdějších předpisů. o archivnictví a spisové službě, které slouží k ukládání archiválií a péči o ně.
<b>CAS</b>	Content addressable storage	Paměťové úložiště určené k dlouhodobému ukládání neměnného obsahu. Generuje globálně jedinečné jméno pro každý uložený objekt, které je nezávislé na běžné výpočetní infrastruktuře, organizaci, umístění nebo technologii. Objekt je vždy přístupný pod tímto jménem. Důsledkem toho je, že objekty se mohou v průběhu času libovolně přemísťovat, měníce své fyzické nebo technologické umístění, aniž by to mělo vliv na jejich přístupnost.
<b>DASTA</b>	Datový standard definovaný Ministerstvem zdravotnictví	Slouží k předávání dat mezi zdravotnickými informačními systémy, je využíván v každodenní praxi již více jak deset let a je zabudován do všech současných významných (českých) zdravotnických informačních systémů.
<b>DICOM</b>	Digital Imaging and Communications in Medicine	Standard pro manipulaci, ukládání, tisk a přenos informací uložených v lékařských obrazových materiálech. Zahrnuje definici formátů souboru a síťový komunikační protokol.
<b>DIP</b>	Dissemination Information Package	Informační balík odvozený z jednoho nebo více AIPů, posílaný uživateli (badatelům) jako odpověď na žádost o poskytnutí informace z OAIS.
<b>DRAMBORA</b>	Digital Repository Audit Method Based on Risk Assessment	Nástroj pro zpracování vnitřního auditu důvěryhodnosti digitálního úložiště.
<b>ERMS</b>	Electronic records management system	Elektronický systém spisové služby je informační systém určený ke správě dokumentů ve smyslu ustanovení § 2 písm. k) zákona, s použitím jeho § 63 odst. 3.
<b>ESS</b>	Elektronická spisová služba	Viz pojmy – elektronická spisová služba
<b>FTP</b>	File transfer protokol	Internetový protokol určený pro přenos souborů mezi počítači, na kterých mohou být provozovány rozdílné operační systémy.
<b>GÚ</b>	Garantované úložiště	
<b>HL7</b>	Health Level Seven	Nezisková organizace zabývající se vývojem standardů v oboru zdravotnické péče. Zde framework a s ním související standard pro výměnu, integraci, sdílení a získávání zdravotnických informací v elektronické podobě.
<b>HSM</b>	Hierarchical Storage Management	HSM je technika ukládání dat, která automaticky přenáší data mezi vysoko a nízko nákladovými ukládacími médii. Příkladem HSM scénáře může být např.: že datové soubory, které jsou často používány, zůstávají uchovávány na discích, ale pokud nejsou užívány po delší časové období (typicky několik měsíců), mohou být migrovány na pásky. Jestliže jsou tato data umístěna na

		páskách a uživatel je potřebuje znovu použít, data jsou automaticky přenesena zpátky na disky.
<b>KDR</b>	Krajský digitální repozitář	
<b>KDS</b>	Krajská digitální spisovna	
<b>KDÚ</b>	Krajské digitální úložiště	
<b>LIS</b>	Laboratorní informační systém	
<b>NAS</b>	Networked Attached Storage	Zařízení pro uložení dat a centrální sdílení dat mezi různými platformami. Skládá se ze dvou základních částí: bezpečného datového úložiště a zařízení (NAS hlava, nebo NAS box) se spec. operačním systémem, optimalizovaným pro rychlou a spolehlivou práci se soubory. Klientovi se celé NAS zařízení jeví jako jeden fyzický disk.
<b>NDA</b>	Národní digitální archiv	
<b>NDK</b>	Národní digitální knihovna	
<b>NFS</b>	Network file system	Internetový protokol pro vzdálený přístup k souborům přes počítačovou síť.
<b>NIS</b>	Nemocniční informační systém	
<b>OAIS</b>	Open Archival Information System	Základní norma, z níž vychází většina moderních návrhů na řešení archivů. Archiv zajišťuje dlouhodobé uložení informací od původců a jejich správu, zpřístupňuje obsah uživatelům (badatelům). Systém je „otevřený“ obsahující doporučení, která reagují na neustálý vývoj technologií, formátů, ...
<b>OCR</b>	Optical Character Recognition	Systémy optického rozpoznávání znaků
<b>ORP</b>	Obec s rozšířenou působností	ORP jsou mezičlánkem v přenesené působnosti státní správy mezi krajskými a obecními úřady. Obecní úřady obcí s rozšířenou působností mají oproti ostatním obecním úřadům některé oblasti působnosti navíc, a to nejen pro svůj vlastní, základní správní obvod, ale zpravidla i pro další obce v okolí.
<b>PACS</b>	Picture Archiving and Communication Systém	Síťová a archivační infrastruktura pro ukládání, přenos obrazových informací z medicínských oborů (EKG, X-RAY ....) a jejich zpřístupnění.
<b>PLATTER</b>	PLAnning Tool for Trusted Electronic Repositories	Metodika plánování a údržby důvěryhodných digitálních úložišť
<b>SIP</b>	Submission Information Package	Informační balík obdržený od původce určený do OAIS. Z něj je vytvořen jeden nebo více archivních informačních balíků (AIP).
<b>SW</b>	Software	
<b>TC</b>	Technologické centrum	Kumulovaná (centralizovaná) hardwarová a softwarová infrastruktura na úrovni kraje nebo ORP, která poskytuje své služby zřizovateli, obcím v jeho správním obvodu a zřízeným nebo založeným organizacím kraje nebo obcí.
<b>Tier</b>		V tomto významu: vrstvy resp. kategorie datových úložišť (médií), jejichž výběr a použití (vhodnost) pro ukládání různých typů dat se liší na základě

		úrovně požadované ochrany, výkonnostních požadavků, frekvenci použití dat atd.
<b>WORM</b>	Write once, read many	Technologie uložení dat, která umožní jednou zapsat informace na ukládací médium a následně zabránit jejich náhodnému nebo záměrnému smazání nebo změně.
<b>ZOOÚ</b>	Zákon o ochraně osobních údajů	

## 1 Základní charakteristika projektu

V agendách veřejné správy, a to na jednotlivých úrovních státní správy i samosprávy, vzniká již několik let **velké množství elektronických dokumentů a dat** nejrůznějších typů uchovávaných na nejrůznějších typech nosičů. Některé z nich mají **historický význam** a potřebují zvláštní péči, aby se dochovaly jako svědectví minulosti. Bohužel řada z nich se již nedochová z důvodu nedostatečné péče původce plynoucí i z nejasných pravidel pro jejich dlouhodobé uchovávání. A tak řada **těchto vzácných dokumentů, elektronických záznamů a dat** mizí zbytečně v propasti „digitálního temna“. V současnosti se díky legislativě elektronické dokumenty zrovnoprávňují s papírovými a to ve všech fázích jejich životního cyklu. Je nutné podpořit celý životní cyklus elektronických dokumentů a dat včetně **zajištění ukládání**. Kromě digitálních dokumentů je nutné bezpečně po neomezenou dobu zajistit také digitalizované dokumenty, vytvářené jednak pro umožnění snadnějšího přístupu badatelů k nim, ale také jako náhrady ohrožených nebo zanikajících fyzických podkladů. Projekt Digitalizace a ukládání, jako subprojekt typizovaného projektu Technologického centra kraje (dále TCK) zajistí proces **digitalizace a ukládání** na všech úrovních samosprávy, rozpracovává a vymezuje podmínky realizace, v souladu se strategickými záměry a paralelně běžícími aktivitami směřující k naplnění cílů Strategie Smart Administration. Základní HW komponenty vytváří projekt Technologická centra kraje.

<b>Cíle projektu:</b>	Vytvořit a udržovat zejména SW nástroje digitalizace a ukládání dat a dokumentů na území kraje jako službu KrÚ, organizacím kraje, městům a obcím a jejich organizacím a v definovaném rozsahu jako veřejnou informační službu.
<b>Cílové skupiny:</b>	Krajské úřady, partnery projektu budou organizace kraje, města a obce, jejich organizace.
<b>Předpokládané výstupy:</b>	Krajská digitalizační jednotka, krajská digitální spisovna, krajský digitální repozitář, krajské digitální úložiště, digitalizovaná a bezpečně dlouhodobě uložená data a dokumenty, publikované digitální dokumenty.
<b>Očekávané přínosy:</b>	Vytvoření standardního systému digitalizace a ukládání dokumentů územních samospráv. Vytvoření základní báze uložených dat a dokumentů. Ochrana kulturního dědictví.
<b>Náležitosti žádosti:</b>	Studie proveditelnosti včetně detailní analýzy digitalizovaných a ukládaných fondů
<b>Členění projektu, indikátory:</b>	

Katalog služeb a výstupů typizovaných projektů výzvy IOP pro krajské projekty						
Téma	Výstup	Objektivně ověřitelné indikátory	Minimální hodnota naplnění indikátoru	Povinný výstup typizovaného projektu	Maximální doba realizace	Maximální uznatelné náklady
Digitalizace a ukládání	Krajské digitální spisovna (KDS)	plně elektronizovaná agenda místní veřejné správy - služba KDS	existuje	ano	2 roky	15mil Kč
Digitalizace a ukládání	Krajský digitální repozitář (KDR)	plně elektronizovaná agenda místní veřejné správy - služba KDR	existuje	ne	2 roky	20mil Kč
Digitalizace a ukládání	Krajské digitální úložiště (KDU)	plně elektronizovaná agenda místní veřejné správy - služba KDU	existuje	ne	2 roky	20mil Kč
Digitalizace a ukládání	Digitalizované a uložené dokumenty	nárůst podílu digitalizovaných dokumentů	min 20%	ne	3 roky	10mil Kč
Digitalizace a ukládání	Krajská digitalizační jednotka	Existence krajské digitalizační jednotky	existuje	ne	2 roky	10mil Kč
Cellkem MAX za téma						75mil Kč



## 2 Východiska

### 2.1 Vize projektu

Veřejná správa disponuje nástroji pro digitalizaci, dlouhodobou, důvěryhodnou a bezpečnou archivaci, ukládání a zpřístupňování hodnotných dokumentů a dat.

### 2.2 Cíle projektu

Cílem projektu je vytvoření a údržba nástrojů digitalizace a ukládání dat na území kraje, zejména pak:

- **Krajská digitalizační jednotka** - pořízení technologií pro digitalizaci - skenery a další SW/HW nástroje pro digitalizaci kulturního dědictví a úředních dokumentů,
- **Digitalizované a uložené dokumenty** - proces digitalizace, zpracování, popisu, ukládání a zpřístupnění dokumentů
- **Vytvoření krajské elektronické spisovny (KDS)** – nástroj pro uložení úředních dokumentů a spisů vzniklých jako produkt činnosti původců,
- **vytvoření krajského digitálního repozitáře (KDR)** – ukládá dokumenty převážně z oblasti kulturního dědictví regionu. To znamená vše, co lze považovat za dokumenty, má význam pro zachycení historie regionu a dokumenty přitom nevznikly jako produkt činnosti orgánů veřejné moci (knihy, filmy, fotografie, digitalizované umělecké předměty a sbírky),
- **vytvoření krajského digitálního úložiště (KDU)** – ukládá jiná data a dokumenty, která pocházejí z činnosti informačních systémů orgánů veřejné správy a je třeba je z nejrůznějších důvodů střednědobě až dlouhodobě ochránit proti ztrátě (zdravotní dokumentace, geodeta, záznamy z kamerových systémů, údaje z provozu informačních systémů důležité pro jejich audit, data síťového provozu apod.)

### 2.3 Návaznost na Technologická centra

Projekt technologických center (TC) je součástí projektu regionálních center, tzv. eGON center, která mají složku technologickou, vzdělávací a administrativní. Takto pojatá centra se stávají výrazným nositelem a šířitelem znalostí konceptu eGovernment. Z pohledu umístění v hierarchii veřejné správy, se eGON centra dělí na eGON centra na úrovni obecních úřadů obcí s rozšířenou působností (ORP) a na krajských úřadech. Ve vztahu k typizovanému projektu Digitalizace a ukládání budou Technologická centra určena zejména k provozu systémů:

- spisových služeb včetně potřebných datových úložišť a datových schránek ve vazbě na implementaci zákona 300/2008;
- vzorových projektových záměrů samospráv jako je projekt Digitalizace a ukládání;
- systémových služeb a dalších aplikací provozovaných pro potřeby samosprávy měst a obcí;

Předmětem projektu Digitalizace a ukládání je i rozšíření kapacit technologického centra kraje a to výhradně v prokazatelně nezbytně nutné míře.

## 2.4 Služby pro území

Z pohledu služeb do území musí projekt v oblasti digitalizace mimo jiné jako povinnou službu zajistit kapacity pro digitalizaci fondů obcí a jimi zřizovaných nebo zakládaných organizací specifikovaných v rámci studie proveditelnosti s doloženým stanoviskem všech ORP na území kraje.

Z pohledu služeb do území musí projekt v oblasti ukládání jako povinnou službu zajistit kapacity KDS a KDR (pokud je zřizován) pro obce a jimi zřizované nebo zakládané organizace specifikované v rámci studie proveditelnosti s doloženým stanoviskem všech ORP.

## 3 Analýza současného stavu

### 3.1 Problematika digitalizace dat

V problematice digitalizace a ukládání dat se na úrovni kraje v současné době prolínají minimálně dva projekty:

1. Projekt vytvoření Národní digitální knihovny,
2. Projekt vybudování Krajského technologického centra

Oba dva projekty mají být součástí širší koncepce s názvem Česká digitální knihovna, která je tvořena velkým množstvím dalších digitálních dokumentů oborového, regionálního, institucionálního i jiného charakteru. Pouze některé z těchto zdrojů se kvalifikují jako nejcennější součást - jádro národního kulturního dědictví umístěné v Národní digitální knihovně. Za jeho vytvoření a uchování nese odpovědnost Národní knihovna a Ministerstvo kultury. Za shromažďování, trvalé uchovávání (na centrálním datovém úložišti nebo lokálních datových úložištích) i zpřístupnění zdrojů mimo „jádro“ národního kulturního dědictví nesou odpovědnost resortně příslušná ministerstva, regiony, instituce atd.

Oba projekty by měly být koordinovány tak, aby se pokud možno nepřekrývaly ani v části digitalizace ani v části uložení (s výjimkou žádané redundance dat pro jejich ochranu). Pro koordinaci digitalizace knihovních fondů jsou připravovány dva nástroje: registr digitalizace a souborný katalog. Paralelně je řešena problematika jednoznačných identifikátorů digitalizovaných objektů.

Projekt Národní digitální knihovny je tvořen vybraným souborem publikovaných digitálních i digitalizovaných dokumentů kvalifikovaných jako základ – jádro národního kulturního dědictví určený k trvalému uchování a zpřístupnění současným i budoucím uživatelům.

Projekt Národní digitální knihovny se zabývá digitalizací, uložením a zpřístupněním podkladů celonárodního významu v následujícím členění:

- Digitalizace historických bohemikálních dokumentů v úplnosti
- Digitalizace bohemikálních dokumentů 19. stol. v úplnosti
- Digitalizace soudobých přírůstků bohemikální literatury
- Digitalizace bohemikálních dokumentů 20. a 21. stol.

Projekt zahrnuje podle odhadu zpracování cca 1,2 milionů dokumentů (svazků) což představuje cca 350 milionů stran.

Problematika vlastní digitalizace je připravována, ověřována aktivitami Národní knihovny, Moravské zemské knihovny a krajských knihoven. Několik národních grantových projektů umožnilo odstartovat projekty digitalizace v knihovnách ČR již počátkem 90. let minulého století. Od samého počátku byly respektovány mezinárodní standardy a díky tomu je možné všechny výstupy integrovat do různých národních i nadnárodních portálů. Metodika digitalizace se průběžně upravuje podle získaných zkušeností. Metodika zahrnuje jak zřízení digitalizačního pracoviště (výběr skenerů, další potřebný hardware, software pro řízení, metodika zpracování naskenovaných předloh), tak jeho provoz a procesy, které ho podporují.

V současné době na úrovni regionů existují také samostatné aktivity v oblasti digitalizace zahrnující:

- Digitalizační zařízení pro lékařskou diagnostiku
- Nástroje pro 3D/4D skenování sbírkových fondů a modelování
- Digitalizaci a georeferencování mapových děl
- Digitalizaci starých tisků
- Digitalizaci tzv. šedé literatury
- Digitalizace novodobých fondů krajských knihoven
- a další.

### 3.2 Problematika ukládání dat

Díky nárůstu využívání moderních informačních technologií a stále častějšímu používání elektronických forem dokumentů, vzniká tlak na efektivní správu, ukládání a manipulaci s elektronickými dokumenty. Řada dokumentů v současné době začíná vznikat již v elektronické formě u původců (důležité evidence, zvukové a obrazové záznamy, fotografie apod.), také se provádí digitalizace stávajících fyzických dokumentů, aby se dochovalo svědectví o skutečnostech, které fyzické dokumenty obsahují, jelikož použitý materiál je citlivý a očekává se v dohledné době jeho dožití, nebo i za účelem zpřístupnění věrné podoby uložených archiválií uživatelské komunitě (badatelům) bez nutnosti používat původní dokument a tím snížit manipulaci s původním dokumentem.

Problematika střednědobého a dlouhodobého ukládání se zabývá daty a dokumenty, které byly v určitém okamžiku zafixovány jako neměnné. V případě dokumentů se jedná typicky o okamžik, kdy je připravovaný dokument prohlášen za hotový (např. po podpisu dokumentu není žádoucí provádět jakékoli jeho změny a měl by být v této podobě uchováván). V případě jiných dat je často tímto okamžikem myšlen okamžik vyexportování dat z jejich provozního systému. Systémově je třeba zajistit, aby ukládaný obsah zůstal neměnný.

Dlouhodobé ukládání digitálních dokumentů vyžaduje trvalou péči, která vyžaduje určité náklady (větší než na pouhé uložení). Také příprava digitálních dokumentů pro dlouhodobé uložení představuje určité náklady. Minimálně je třeba k dokumentům připojit i příslušná metadata (popisná, technická a administrativní) a dokumenty zabalit do balíčků vhodných pro archivaci.

Pro některá data mohou být náklady na jejich převedení do podoby vhodné pro dlouhodobé uložení velmi vysoké, protože dosud neexistují dostupné metody vhodné pro tento převod. Příkladem může být uložení složitějších databázových aplikací, která vyžaduje kromě uložení vlastních dat i dlouhodobé uložení aplikačního softwaru. Na výzkumu a vývoji obecných metod pro archivaci složitějších komplexnějších dat se ve světě dosud pracuje.

Z těchto důvodů jsme problematiku ukládání rozdělili na dvě části:

1. Dlouhodobé ukládání dokumentů,
2. Bezpečné dlouhodobé ukládání dat.



## 4 Analýza poptávky a koncepce marketingu

Analýza poptávky je dlouhodobě prováděna a reprezentována požadavky Svazu měst a obcí ČR a Asociace krajů ČR.

### 4.1 Analýza poptávky výstupů projektu

V rámci analýzy trhu byl proveden průzkum projektových záměrů, kterého se účastnily všechny kraje ČR. Šetření mimo jiné prokázalo, že panuje shoda o potřebnosti a enormní zájem o řešení problematiky digitalizace a ukládání, zejména pak ze strany knihoven.

Zároveň existuje obecná poptávka po uceleném řešení problematiky digitalizace elektronických písemností orgánů veřejné správy

### 4.2 Marketing

Garant projektu bude provádět a dohlížet na zajištění propagace projektu přidáním publicity podle pravidel výzvy, bude-li projekt spolufinancován ze strukturálních fondů.

## 5 Požadavky na řešení projektu

### 5.1 Digitalizované a uložené dokumenty

Zatímco zdroje dokumentů pro projekt Národní digitální knihovny jsou zmapovány celkem podrobně, zdroje (podklady) digitalizace regionálního charakteru v rámci Krajských technologických center zatím nejsou přesně specifikovány. V obecné poloze by digitalizace měla být podle zadání zaměřena na zpracování dokumentů pro potřebu fungování úřadů a dále na záchranu, ochranu a zpřístupnění dokumentů z oblasti knižních fondů, stavebních spisoven, zdravotnických spisoven nebo dokumentů významných svým obsahem či původem pro kulturní, politické, náboženské či jiné oblasti, kterým hrozí nebezpečí fyzického poškození či rozpadu v důsledku jejich častého používání.

Konkrétní seznam zdrojů v rámci kraje bude předmětem konkrétní studie proveditelnosti. Digitalizace je chápána jako jeden ze zdrojů digitálních dokumentů, který je nutno uchovávat. Vzhledem k celkovému značnému rozsahu a vysokým nákladům, bude vytvoření strategie postupu digitalizace regionálních zdrojů zcela jistě podléhat nejen odbornému, ale především politickému rozhodnutí.

Pro zdárný průběh procesu digitalizace doporučujeme rozčlenit digitalizaci regionálních dokumentů na jednotlivé menší subprojekty podle typů zdrojů, které mají být digitalizovány.

Na úrovni kraje tedy bude existovat jedno, nebo několik digitalizačních pracovišť, která budou předávat výstupy své činnosti k archivaci buď do Národní digitální knihovny (protože digitalizují podklady patřící do NDK) popř. do Národního digitálního archivu nebo do Krajského digitálního repozitáře (protože se jedná o dokumenty regionálního významu).

Všechny výstupy by měly být podchyceny v registru digitalizace, resp. v Souborném katalogu České republiky. V případě potřeby bude možno digitalizační pracoviště sdílet pro oba projekty.

Projekty digitalizace na krajské úrovni budou respektovat metodiku digitalizačních pracovišť zpracovanou Národní knihovnou, včetně standardů pro výstupní datové formáty.

Digitalizace dat úzce navazuje na následující specifikace technického vybavení, které bude pořízeno v rámci služby Krajská digitalizační jednotka. Předmětem tématu digitalizace dat v oblasti výdajů jsou pak především práce na digitalizaci dokumentů. Jde typicky o tyto činnosti:

- Vyhledávání a příprava fondů pro digitalizaci
- Příprava dokumentů
- Skenování
- Popis metadat a vyhledávání informačních zdrojů
- Práce s OCR
- Správa provozního úložiště a repozitáře

Výstupem projektu digitalizace je nárůst digitalizovaných dokumentů v kraji minimálně o 20% a naplněné úložiště Krajského digitálního repozitáře.

V oblasti uznatelných nákladů jde pak o pořízení SW na zpracování výstupů z krajské digitalizační jednotky:

- Sw na zpracování naskenovaných dat
- Nástroje na převody formátů a grafické úpravy
- Sw pro katalogizaci a tvorbu metadat
- Systémy OCR
- Systémy podpory workflow
- Integrace na primární evidenční systémy (např. knihovní a sbírkové)
- Systémy správy a konverze digitalizovaných 3D objektů
- personální náklady instituce a náklady na případnou dodávku digitalizace dokumentů (outsourcing).

Maximální přijatelné náklady na výstup Digitalizované a uložené dokumenty jsou 10mil Kč na kraj.

## 5.2 Krajská digitalizační jednotka

Předmětem projektu krajské digitalizační jednotky je pořízení SW a HW vybavení pro digitalizace dokumentů a dalších objektů. Přijatelným nákladem na vznik digitalizační jednotky je nákup příslušného HW a SW. Předpokládaná skladba prostředků pro realizaci digitalizační jednotky je následující:

- Robotický skener (formát až A2, min rozlišení 600 DPI)
- 1-2 ruční skenery (formát až 2A0, min rozlišení 600 DPI)
- 3D-4D skener
- Příprava vhodných prostor (eliminace negativních vlivů vnějšího světla, hluchost,...)
- Kapacity provozního úložiště (nejlépe rozšířením Technologického centra kraje)
- Základní Sw na zpracování naskenovaných dat

Výstupem projektu krajské digitalizační jednotky je realizace funkční jednotky jako služby pro následnou digitalizaci.

## 5.3 Problematika ukládání dat

Dokumenty obvyklých (běžných) formátů (textové, obrazové, audio, video), pro které jsou nástroje na převedení do podoby vhodné pro dlouhodobé uložení běžně k dispozici, navrhujeme ukládat a dále spravovat v systému dlouhodobého úložiště.

Data komplexnějšího charakteru, pro která by převod do archivních balíčků byl nákladný nebo v dané době nemožný, navrhujeme pouze ukládat do digitálního úložiště. Uložení takovýchto dat do úložiště znamená odložení řešení problému dlouhodobé ukládání na pozdější dobu. Přitom



předpokládáme, že čas potřebný pro vyřešení metod pro dlouhodobou archivaci těchto dat bude kratší, než je životnost příslušného softwaru či příslušného formátu.

Rozdělením problematiky na dvě části máme možnost vyřešit dlouhodobé uložení jen těch dokumentů, u kterých je tento problém v současné době rozumně řešitelný a zároveň neztratit dokumenty a data u kterých trvalé uložení zatím není uspokojivě vyřešeno.

Pro každý nový vstup (data či dokumenty), který bude nutno uložit, bude třeba nejprve specifikovat a vyhodnotit potřebnou délku uložení, formát dat a jeho potenciální ohrožení, existenci nástrojů pro převod dat do archivního formátu, hodnotu dat. Na základě těchto kritérií bude možno rozhodnout, zda má smysl data uložit do dlouhodobého úložiště, nebo zda postačuje uložení v digitálním úložišti. (Uložení dat mimo dlouhodobé úložiště přináší potencionální nebezpečí ztráty logické čitelnosti vlivem nedostupnosti software, kterým byla data pořízena, resp. kterým je možno jejich formát interpretovat).

Mezi uznatelné náklady projektů ukládání dat patří:

- Realizace popř. nákup systémů KDS, KDR a KDU
- Náklady na spuštění a integraci systémů
- Nezbytně nutné rozšíření kapacit krajského technologického centra

### 5.3.1 Dlouhodobé ukládání dokumentů (KDS, KDR)

Při návrhu digitálního úložiště je nutné mít na paměti neustálý vývoj technologií. Je třeba především volit řešení, které bude dostatečně otevřené, ověřené a podporované a tím oddálí potřebu vyměnit technologie. Na druhou stranu, pokud bude nutné technologii vyměnit (dlouhodobě se s nutností výměny musí počítat, viz životní cyklus úložiště), umožní použité řešení jednodušší přechod a migraci do nového prostředí počítačových technologií.

Pro budování digitálního úložiště zaměřených na dlouhodobé uchovávání digitálních dokumentů se v současnosti ve světě vychází především ze standardu OAIS (Open Archival Information System). Ten specifikuje základní funkční části otevřeného archivu, komunikaci s okolím, procesy a informační model ve formě informačních balíčků přijímaných, poskytovaných a především uložených v repozitáři. Z tohoto modelu vycházíme i při návrhu systému pro dlouhodobé uložení digitálních dokumentů v rámci Technologického centra kraje.

Technologie pro provoz systému digitálního úložiště, dlouhodobých úložišť a zpřístupňování informací se neustále mění a je nutné na tyto změny reagovat tak, aby systém zůstal životaschopný. V rámci životního cyklu systému je nezbytné provádět pravidelný dohled nad morálním zastaráváním technického řešení. Je nezbytné mít připravené takové postupy, aby bylo možno dostatečně včas spustit proces obnovy včetně přípravy a finančního zajištění. V modelu OAIS se touto činností zabývá část nazývaná „Preservation planning“. Tato služba

monitoruje změny vnějšího prostředí, které by mohly mít dopad na schopnost archivu chránit a udržet přístup k informacím v péči. Jako reakci na změny vytváří tato služba doporučení pro aktualizaci politik a procedur OAIS a pro přizpůsobení se těmto změnám (např. navrhuje změnit technický formát uložených souborů, který je ohrožen na jiný-nový formát. Připravuje a ověřuje i software pro migraci formátu.) Provozování této služby představuje jeden z podstatných nákladů na dlouhodobé uložení.

Model OAIS je všeobecně odbornou veřejností přijímán i pro projekty Národní digitální archiv a Národní digitální knihovna, jejichž realizace se připravuje na celonárodní úrovni (nositeli projektů jsou Národní archiv ČR a Národní knihovna ČR). Oba tyto subjekty se budou dlouhodobě zabývat i činnostmi souvisejícími s „Preservation planning“. Podle názoru autorů této studie by se výstupy z jejich činnosti mohly využít i pro úložiště na krajské úrovni. Obě instituce by tedy měly provádět sledování vývoje, přípravu SW, potřebného pro migraci doslouživších formátů, změny ukládacích technologií a vydávat pokyny a doporučení pro provedení migrací či změny technologie (např. typu média) pro dlouhodobé uložení svých dokumentů. Tyto pokyny by měly být obdobné i pro dlouhodobé uložení na krajské úrovni (KDS a KDR). Pokud by byl přijat tento postup „metodického řízení“ z NDA a NDK, mohly by se na krajské úrovni vydané pokyny a doporučení využít. To znamená například podle vydaného doporučení vybrat dokumenty s ohroženým formátem k migraci a tyto, přes migrační SW (získaný z NDA či NDK) převést do nového formátu a uložit je zpět do repozitáře. Tím by se náklady na digital preservation na krajské úrovni mohly podstatně zmenšit. Z tohoto důvodu doporučujeme najít či upravit pro vztah mezi Národním archivem, provozujícím Národní digitální archiv, Národní knihovnou provozující Národní digitální knihovnu a krajskými úřady provozujícími KDS a KDR rozumný právní resp. smluvní statut, který by výše zmíněné metodické řízení umožňoval.

Problematika dlouhodobého ukládání se dále věnuje dvěma kategoriím dokumentů:

- 1. Úřední dokumenty**
- 2. Ostatní dokumenty kulturního dědictví.**

### *Ad 1. Úřední dokumenty*

Správa úředních dokumentů je stanovena příslušnými legislativními předpisy především zákon č. 499/2004 Sb. ve znění pozdějších předpisů.

Novela Archivního zákona předepisuje původcům předávat digitální archiválie po uplynutí skartační lhůty do NDA.

Mezinárodní standard OAIS doporučuje vytváření informačních balíčků obsahujících data a metadata.

Nesporně nejjednodušším okamžikem pro vytvoření těchto balíčků je okamžik uzavření spisů či dokumentů a jejich výstup ze systému ERMS do KDS, protože v této době lze získat to nejširší množství metadat přímo ze spisové služby. Nelze racionálně předpokládat vůli, časové možnosti nebo dostatek podkladů pro vytváření těchto balíčků až po uplynutí skartačních lhůt, které se pohybují v rozmezí od 5 do 100 let.

Pokud vezmeme například dokument s archivační lhůtou A50, pak po 50ti letech od vzniku dokumentu se mají získat jeho metadata, sestavit balíček a předat do NDA. Otázky typu kde za 50 let bude možno metadata získat, v jakém formátu bude uchován originální dokument, zdabude ještě čitelný apod. možná nejsou nyní aktuální, ale systémový návrh řešení by měl nabídnout jejich řešení. V rámci skartačního řízení se balíčky obsahující spisy či dokumenty určené k archivaci pouze přesunou do NDA.

Použití standardu OAIS pro systém digitální spisovny se snaží uvedeným problémům předejít a sestavit informační balíčky s dokumenty ihned po uzavření dokumentu či spisu, kdy jsou veškeré potřebné údaje dostupné v aktuálních formátech a systémech. Takto připravený balíček se po 50 letech již jen do NDA přesune.

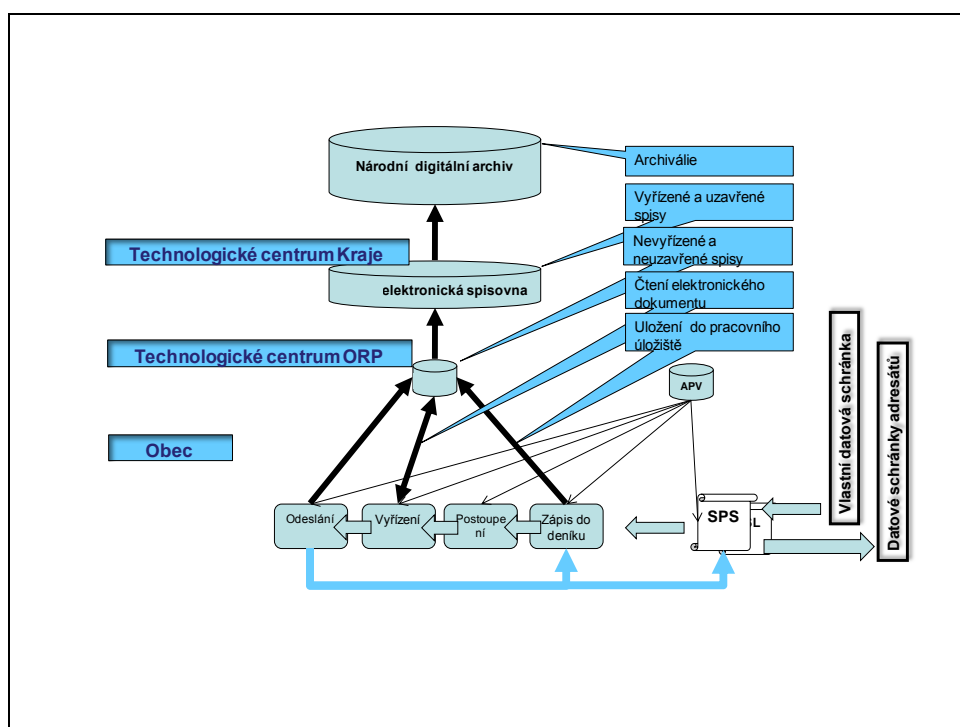
Úřední dokumenty a spisy vznikají a vyřizují se v různých IT systémech a aplikacích jako např. "výběrová řízení", "systém pro stavební řízení atd. (nebo přímo v systému spisové služby (ERMS)) a jako metadata jim musí být přiřazeny mimo jiné především věcná skupina a skartační režim dle spisového plánu ERMS původce. **Uzavřený dokument** se již nesmí měnit a pro jeho uchování je třeba s ním zacházet předepsaným způsobem. Listinné dokumenty se předávají do listinných spisoven. Elektronické dokumenty a spisy se po uzavření ukládají do **elektronické spisovny**. Životnost dokumentů a spisů uložených v elektronické spisovně je řízena spisovým plánem organizace. Uložené dokumenty a spisy čekají v elektronické spisovně na skartační řízení. Po uplynutí skartační lhůty dojde buď ke skartaci dokumentů, nebo dojde k výběru archiválií, které se předávají do nadřízeného digitálního archivu (např. Národní digitální archiv). Je třeba počítat s tím, že některé dokumenty mohou v elektronické spisovně zůstat po velmi dlouhou dobu, aniž by se skartovaly či předávaly. Analýzou spisového plánu Krajského úřadu zjistíme, že některé typy dokumentů (výběr je uveden v příloze studie) mají skartační lhůty 50, 70 i více let.

Podobu uložení elektronického dokumentu musí systém elektronické spisovny zajistit ochranu uložených informací před ztrátou, důvěryhodnost uložených informací (nezměněnost a prokazatelnost vzniku v uvedeném čase) a čitelnost uložených informací v budoucnosti. Kromě toho musí elektronická spisovna zajistit i ochranu uložených informací proti neoprávněnému přístupu.

**Správu úředních dokumentů na úrovni kraje v době od uzavření (vyřízení) do skartace či vyřazení do Národního digitálního archivu zajistí Krajská digitální spisovna (KDS). Z právního**

pohledu bude KDS provozována především podle zákon č. 499/2004 Sb. ve znění pozdějších předpisů, Národního standardu pro elektronické systémy spisové služby a další platné legislativy.

Krajská digitální spisovna navazuje na projekty implementace (či rozšíření) elektronické spisové služby u různých subjektů kraje. Propojení elektronických spisových služeb, Krajské digitální spisovny a následně i Národního digitálního archivu ukazuje následující obrázek.



Budoucími původci KDS budou následující subjekty, u nichž vznikají digitální dokumenty:

- orgány Kraje (hejtman, Rada, Zastupitelstvo, Krajský úřad včetně jeho organizačních jednotek). Zdrojem dokumentů jsou členové volených orgánů či zaměstnanci krajského úřadu.
- Krajem zřízené nebo založené organizace
- zrušená nestátní zdravotnická zařízení, převážně v případě úmrtí soukromého lékaře, jehož zdravotní dokumentaci jednotlivých pacientů má kraj povinnost převzít (týká se převážně převzetí elektronické zdravotnické dokumentace soukromých lékařů v případě jejich náhlého úmrtí)

- obce Kraje
- organizace zřízené nebo založené obcemi Kraje

Do digitální spisovny Kraje se budou dostávat dokumenty převážně z elektronických systémů spisové služby (ERMS) původců. Nebudou to ovšem jediné zdroje elektronických dokumentů. Dalšími zdroji jsou samostatné elektronické agendy, produkující elektronické dokumenty různého typu podle metodik ministerstev. Dnes jsou to především ekonomické evidence (účetnictví), agendy životního prostředí, agendy živnostenských úřadů, agendy grantových agentur, matrika a v budoucnu možná i další

Odhadovaný roční přírůstek digitálních dokumentů do KDS činí 2-4 TB dat za kraj ročně.

Formáty elektronických dokumentů uložených v KDS musí odpovídat aktuální vyhlášce MV ČR, která určí výstupní formáty systémů spisových služeb. Současná vyhláška č. 191/2009 Sb., o podrobnostech výkonu spisové služby stanovuje tyto povolené typy:

- Jako výstupní datový formát statických textových dokumentů a statických kombinovaných textových a obrazových dokumentů se použije:
  - Formát Portable Document Format/Archive (PDF/A, ISO19005) se použije jako výstupní datový formát statických textových dokumentů a statických kombinovaných textových a obrazových dokumentů.
- Jako výstupní datové formáty statických obrazových dokumentů se použijí:
  - formát Portable Network Graphics (PNG, ISO/IEC 15948)
  - formát Tagged Image File Format (TIFF, revize 6 - nekomprimovaný)
  - formát JPEG File Interchange Format (JPEG/JFIF, ISO/IEC 10918)
- Jako výstupní datové formáty dynamických obrazových dokumentů se použijí:
  - video programový nástroj pro komprimaci dat (kodek) Moving Picture Experts Group Phase 2 (MPEG-2, ISO/IEC 13818)
  - video programový nástroj pro komprimaci dat (kodek) Moving Picture Experts Group Phase 1 (MPEG-1, ISO/IEC 11172)
  - formát Graphics Interchange Format (GIF)
- Jako výstupní datové formáty zvukových dokumentů se použijí:
  - zvukový programový nástroj pro komprimaci dat (kodek) MP2 (MPEG-1 Audio Layer 2)
  - zvukový programový nástroj pro komprimaci dat (kodek) MP3 (MPEG-1 Audio Layer 3)

- formát Waveform audio format (WAV), modulace Pulse-code modulation (PCM)

## **Ad 2. Dokumenty kulturního dědictví regionálního významu.**

Kromě dokumentů úředního významu, popsanych v předchozím odstavci, existuje na regionální úrovni celá řada dalších dokumentů, které nemají úřední charakter ani nevznikly činností nějakého úřadu. Jedná se například kulturní památky, cenné písemnosti a umělecká díla, fotografie, historické mapy, audio, video, časopisy a ostatní publikace vztahující se k regionu.

Podle způsobu vzniku se může jednat o výstupy z digitalizačních linek (to znamená, že po provedení digitalizace existuje ještě po nějakou dobu originální fyzický podklad) či o dokumenty vzniklé již primárně v digitální podobě. Vzhledem k tomu, že životnost původních fyzických originálů je často omezena (viz problematika tiskovin vytištěných na kyselém papíru) je třeba se k oběma těmto druhům chovat z pohledu dlouhodobého ukládání stejně. Ztráta digitálního dokumentu z digitálního repozitáře je nenahraditelná.

### **Na úrovni Kraje budou ukládány tyto dokumenty v Krajském digitálním repozitáři (KDR).**

Bude se jednat o tyto typy dokumentů:

- dokumenty, cenné písemnosti, umělecká díla a vybrané knihovní fondy spravované Krajskou knihovnou a dalšími krajem nebo obcemi zřizovanými paměťovými institucemi, zejména knihovnami, muzei, archivy.
- historické dokumenty a cenné písemnosti vzniklé z činnosti nebo spravované školami a vědeckými institucemi
- 3D digitalizované vybrané kulturní památky
- historické dokumenty a cenné písemnosti vzniklé z činnosti náboženských obcí a kongregací
- dokumenty vytvořené soukromými osobami
- webové stránky regionálního významu vytvořené libovolnými původci
- data uložená a pevných nosičích (typicky s problematickou životností a čitelností – diskety, pásky, CD/DVD apod.)

Do KDR se dostávají elektronické dokumenty ve formě vstupních informačních balíčků (SIP), které mohou vytvářet z dodaných datových souborů (popisná metadata, obrázky) archiváři KDR nebo tyto balíčky vytvářejí podle dohody příslušní původci, kteří mají k dispozici uživatelské rozhraní KDR, umožňující tvorbu těchto balíčků a zasílají je do KDR. V případě digitalizace na digitalizačním pracovišti budou balíčky SIP sestavovány v rámci exportu dat.

Protože jedním z nejvýznamnějších původců KDR bude Krajská knihovna příslušného kraje, která bude digitalizovat kulturní dědictví regionálního významu z knihoven v regionu kraje, bude mít mimo standardní možnosti vytvářet balíčky SIP i možnost **vzdáleně přistupovat do KDR** ve správcovských rolích a ze zaslaných dokumentů vytvářet fondy či sbírky a dle potřeby upravovat popisná metadata

Podobným zdrojem dokumentů bude i projekt digitalizace předmětů ze sbírek muzeí a galerií.

Odhadovaný průměrný roční přírůstek digitálních dokumentů do KDR, činí 3-5 TB dat za kraj.

Jednoznačnou identifikaci dokumentů – digitálních objektů – zajišťuje po jejich vstupním zpracování Systém správy dat KDR, který tento jednoznačný identifikátor vygeneruje a uloží jak do své databáze, tak současně i s ostatními metadaty a samotným dokumentem do archivního informačního balíčku AIP. V rámci projektu Národní digitální knihovny bude definován i způsob jednotné identifikace dokumentů. Jakmile bude způsob přidělování identifikátorů v tomto projektu stanoven, bylo by žádoucí jej převzít i pro KDR tak, aby identifikace byla jednotná. Pro uživatelské vyhledávání dokumentů se využijí popisná metadata, která mohou být koncipována podle následujících standardů či celostátních, nebo mezinárodních doporučení:

- standard metadat stanovený Národní knihovnou pro knihovní systémy,
- možné vazby na číselníky stanovené Národní knihovnou (např. Celostátní databáze národních autorit vedená Národní knihovnou),
- základní archivní metadata používaná při budování archivních fondů a sbírek,
- potencionálně může vzniknout požadavek i na využití standardů používaných v muzejnictví,

Povolené formáty elektronických dokumentů pro uložení v KDR budou převážně stejné jako formáty pro KDS. Některé formáty, např. pro 3D skenování, bude třeba doplnit. Seznam povolených formátů lze pro KDR v podstatě libovolně rozšiřovat.

Z právního pohledu se v případě KDR předpokládá zakotvení mimo působnost archivního zákona, tedy na smluvním základě. Jeho činnost a fungování je založeno především ve vztahu k externím subjektům na normách civilního závazkového práva.

### 5.3.2 Krajské digitální úložiště (KDU)

Zdrojem dat pro Krajské digitální úložiště mohou být různé informační systémy provozované u různých původců.

Na rozdíl od problematiky ukládání dokumentů uvedené v předchozích dvou kapitolách se KDU zabývá uložením předem obecně nespecifikovaných dat.

Předmětem projektu bude uložení dat v KDÚ ve formě souborů resp. datových bloků ukládaných původci prostřednictvím určených standardních souborových systémů nebo prostřednictvím aplikací umístěných v TC kraje (typicky využitím úložiště vrstvy Tier 3 TC kraje). Úložiště neslouží pro přímé ukládání provozních dat původců/uživatelů ale pro dlouhodobé uložení výstupních dat typicky IS.

Příkladem jsou soubory vzniklé záznamem z kamer, nebo logy z provozních systémů. Ty se pomocí některého standardního souborového protokolu (ftp, nfs, cifs, apod.) dostupného v operačním systému přenesou po síti a uloží se do adresářové struktury na disk.

Vnitřní struktura uložených dat a manipulace s nimi (včetně např. skartace) je zodpovědností uživatele. Provozovatel KDU řeší pouze samotné uložení dat, jejich dostupnost a zálohování dle parametrů domluvených s uživatelem (typicky RTO, RPO, životnost, SLA, počet kopií,...).

Dalším příkladem je ukládání zdravotnické dokumentace kompatibilní s PACS (Picture Archiving and Communication System) systémy. Tento typ zdravotnické dokumentace je výstupem z jednotlivých digitálních lékařských přístrojů (modalit) schopných předávat obrazová data v různých formátech, z nichž nejvýznamnější jsou DICOM, HL7 (případně i národní standard DASTA).

Odhadovaný průměrný roční přírůstek digitálních dat do KDU je 10-20 TB dat.

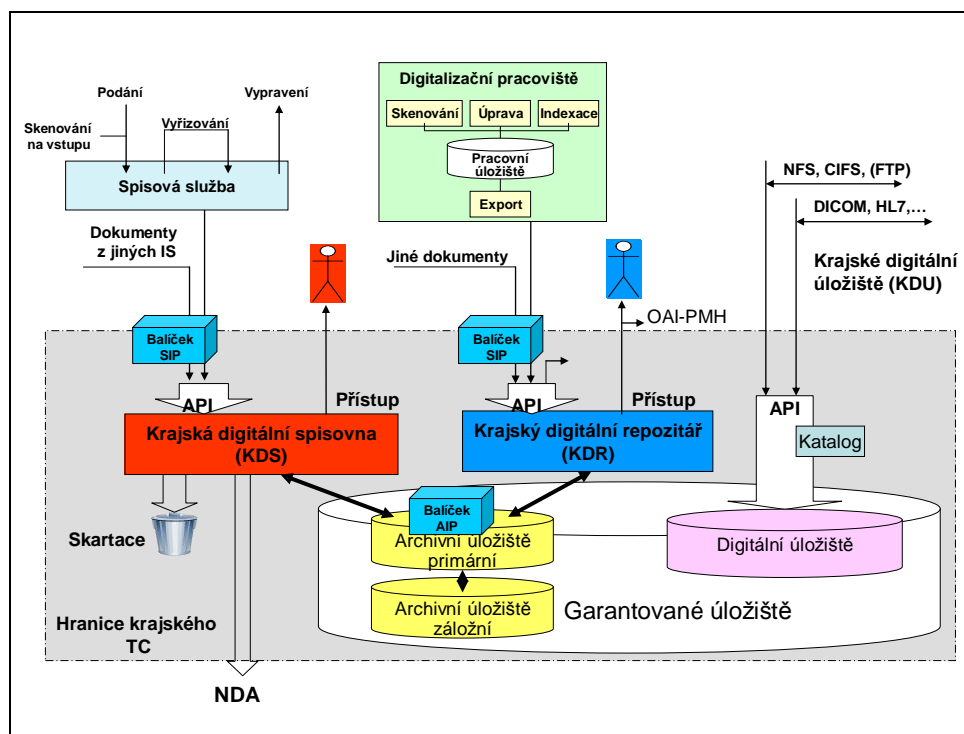
## 5.4 Architektura řešení

### 5.4.1 Softwarová architektura

Návrh architektury aplikací použitých při implementaci KDS a KDR vychází ze standardního modelu OAIS (viz. předchozí kapitola). Struktura navržených komponent těchto dvou subsystémů je tedy podobná. Podstatné rozdíly jsou v konfiguraci funkcí, typu interních procesů, použití formátů a metadat, realizaci přístupu uživatelů a definici přístupových oprávnění.

KDÚ je budováno jako systém pro zabezpečené uložení souborových dat bez evidence i kontroly metadat a s nižšími požadavky na zajištění dlouhodobého důvěryhodného uložení dokumentů při nižších nákladech na implementaci a provoz úložiště. Vzhledem k těmto cílům je softwarová architektura KDÚ podstatně jednodušší. Předpokládá se implementace i rozhraní využívající možností standardních souborových systémů.





### 5.4.1.1 Softwarová architektura KDS a KDR

Subsystémy KDS a KDR založené na principech OAIS přistupují k ukládaným dokumentům a spisům jako k balíčkům, obsahujícím předmětná data a současně jejich metadata za účelem dlouhodobého uložení. Podle fáze jejich životního cyklu se jedná o vstupní (SIP), archivní (AIP) a výstupní (DIP) balíčky.

Rozhraní pro přístup k těmto systémům je specificky navrženo pro příjem a výdej balíčků v příslušném formátu definovaném na základě standardů. Vzhledem k zajištění bezpečnosti a konzistence uložených dat probíhá příjem dat do úložiště asynchronně v rámci procesu, který se skládá z několika kontrolních a transformačních procedur.

Systém digitálního repozitáře i digitální spisovny se skládá z těchto softwarových komponent:

- **Vstupní modul**

Příjem dat

Zajišťuje komunikaci s původcem, autentizaci, autorizaci a uložení přijatých balíčků SIP do pracovního úložiště.

Kontrola kvality vstupních dat (kontrola datové struktury, kontrola na obsah škodlivého kódu).

Kontroluje formální strukturu balíčků a přítomnost virů a jiného škodlivého obsahu balíčků. V rámci tohoto modulu je zřízena i tzv. karanténna zóna pro zajištění spolehlivosti kontrol. Struktura vstupních SIP balíčků může být doplněna dle příslušné metodiky např. o detailnější data typicky ze životního cyklu dokumentu (např. ve spisové službě).

#### Řízení příjmu

Kontrola popisných a technických metadat, kontrola přípustnosti souborových formátů, kontrola struktury balíčku SIP a vzájemného provázání balíčků.

#### Generování balíčků AIP

Automatické doplnění zejména technických metadat, konverze formátů metadat, možnost manuálního doplnění metadat, vstupní migrace formátů včetně generování náhledů pro prezentaci dat archivu v určeném formátu.

#### Řízení ukládání

Zajišťuje konzistentní uložení metadat a obsahu archivních balíčků současně do archivního systému, systému správy dat a systému pro přístup.

- **Modul správy dat**

#### Evidence číselníků

Zajišťuje ukládání a přístup k číselníkům používaným v rámci vstupní kontroly a vyhledávání. Jedná se zejména o tyto číselníky - původci, klasifikace, povolené souborové formáty, kategorizace dokumentů podle kritérií přístupnosti, požadavků na zachování důvěryhodnosti, doby uložení.

#### Evidence přijímaných a uložených balíčků.

Zajišťuje vedení a přístup ke katalogu uložených dokumentů včetně stavu příjmu a uložení.

#### Evidence periodické obnovy časových razítek.

Zajišťuje evidenci historie obnovy časových razítek pro jednotlivé balíčky pro trvalé zajištění důvěryhodnosti uloženého obsahu.

#### Evidence kontroly konzistence.

Uložení kontrolních součtů jednotlivých uložených balíčků AIP na aplikační úrovni pro účely periodické kontroly konzistence uloženého obsahu nezávisle na vlastnostech použitého archivního úložiště (CAS/NAS).

#### Evidence procesů skartace a ukládání.

Informace o stavu skartace a informace o stavu jednotlivých balíčků AIP zařazených do skartačního řízení.

- **Archivní systém**

Zajišťuje vlastní důvěryhodné uložení obsahu balíčků AIP

Je implementováno primárně prostřednictvím technologie CAS (v případě vybraných klasifikací dokumentů v subsystému KDS i technologie NAS).

- **Modul administrace**

Řízení procesu příjmu

Pro administrátora zajišťuje přehled o stavu příjmu balíčků SIP, umožňuje řešení problémů se strukturou a obsahem balíčků při příjmu.

Řízení procesů migrace

Spouštění migrace souborových formátů v uložených balíčcích a přehled o provedených migracích.

Řízení procesu časového razítkování.

Kontrola periodické obnovy časových razítek u uložených balíčků, případně i manuální spouštění obnovy razítek.

Skartační řízení

Příprava návrhu a jeho schvalování, provedení skartace, případně exportu do jiného archivu v definovaném formátu.

Správa kontroly konzistence

Přehled o průběhu ověřování kontrolních součtů a o nalezených problémech s uložením balíčků AIP.

Správa číselníků.

Zajišťuje pro administrátory původce a archivu aktualizaci a čtení číselníků používaných v rámci vstupní kontroly a vyhledávání.

Ukládání transakčních záznamů.

Pro účely auditu zaznamenává veškeré provedené operace nad uloženými balíčky (příjem, kontrola, transformace, ukládání, čtení). Zaznamenané záznamy jsou zároveň ukládány do úložiště ve formě AIP.

Přístup k transakčním záznamům

Zobrazení transakčních záznamů pro účely auditu.

- **Přístupový modul**

Zabezpečení přístupu a autentizace uživatelů.

Zajištění přístupu uživatelů k uloženým metadatům a dokumentům.

Autorizace - omezení přístupů na základě klasifikace dokumentu, původce, uživatelských skupin a rolí uživatelů.

Modul povolí přístup ke čtení obsahu nebo metadat podle rolí přihlášeného uživatele a oprávnění příslušného balíčku.

Vyhledání uložených balíčků na základě zvolených metadat.

Zobrazení náhledů a distribuce uložených dokumentů ve formě DIP

Systém umožní výběr dokumentů a jejich zaslání oprávněnému uživateli ve standardizované podobě.

Provádění transakčních záznamů o přístupu k jednotlivým uloženým balíčkům

Programové rozhraní API na externí portál pro přístup

Systém eviduje veškeré přístupy k uloženým dokumentům a archivuje je.

### 5.4.1.2 Rozdíly ve funkčnosti KDS a KDR

- **Vstupní modul**

Jiné standardy použité při kontrole a doplňování metadat.

V subsystému KDR se použijí standardy a metadata definovaná Národní knihovnou, případně další, která budou dohodnuta s původci archivních balíčků (knihovní sbírky, muzejní exponáty atd.).

V subsystému KDS se použije „Národní standard pro elektronické systémy spisové služby“ definovaný MV ČR a schéma pro předávání dokumentů a jejich metadat do archivu, které je definováno v rámci tohoto standardu.

Obecně jiná konfigurace vstupních souborových formátů.

V subsystému KDS i KDR se využijí souborové formáty definované ve vyhlášce MV (vyhláška č. 191/2009 Sb.). V KDR navíc takové, které budou dohodnuty s původci archivních balíčků (knihovní sbírky, muzejní exponáty atd.).

Odlišné nastavení použití archivního úložiště (volba způsobu uložení CAS/NAS).

V subsystému KDS může být v některých případech (krátká skartační lhůta) použito úložiště typu NAS.

- **Modul správy dat a modul administrace**

V KDS probíhá zákonné skartační řízení. V subsystému KDR se provádí pouze interní skartační řízení.

V KDS probíhá opakovaná obnova časových razítek, v KDR probíhat nemusí.

Odlišné nastavení přístupových oprávnění – u KDS jsou operace týkající se správy obsahu uložených dat delegována na pověřené správce původců (každý původce si spravuje svoji spisovnu), podobně přístup k uloženým dokumentům je pro uživatele původce. V KDR tyto operace typicky nevykonává přímo pracovník původce. Pro přístup k dokumentům v KDR je nutno mít nastavitelné politiky (obecně přístupné, omezení autorskými právy, osobních údajů a jiná omezení).

- **Přístupový modul**

Datové balíčky KDR budou zpřístupněny prostřednictvím snadno ovladatelného přístupového modulu větší množině uživatelů nebo veřejnosti.

Datové balíčky KDS budou zpřístupněny pouze omezené množině autorizovaných uživatelů původce.

### 5.4.1.3 Softwarová architektura KDÚ

Subsystém KDÚ slouží k přímému, rychlému ukládání dat vybraných kategorií po dohodě s původci. Data jsou ukládána přímo ve formě souborů a povinně neobsahují popisná metadata v jednotně stanoveném formátu. Mimo zajištění spolehlivého uložení a zálohy obsahu datových souborů tento subsystém nezajišťuje žádné další obslužné operace sloužící k zajištění konzistence, důvěryhodnosti a přístupnosti obsahu.

Na úrovni Technologického centra kraje může být v technicky a ekonomicky odůvodněných případech, kromě vlastního úložiště nainstalován i příslušný software (např. archivní server pro ukládání zdravotní dokumentace ve formátu DICOM), který řídí jak komunikaci s archivující protistranou (původcem), tak i ukládání přenesených dat do fyzického úložiště. V tomto případě je komunikace po síti vedena na úrovni příslušného aplikačního protokolu a není využíván přímý přístup do souborového systému. Výhodou takového řešení může být snížení nákladů díky centralizaci archivačního subsystému a jeho obsluhy i možnost přímého využití specifických funkcí pro práci s archivovanými daty (vyhledávání, čtení, bezpečnost) informačním systémem původce.

Systém KDÚ je rozdělen na jednotlivé logické segmenty úložiště. Tyto logické segmenty jsou definovány v katalogu KDÚ a na jejich základě je vytvořena logická (adresářová) struktura úložiště. Pro každý segment je definován typ ukládaných dat, formát datových souborů, ukládací politika (pro řízení HSM), původce datových souborů a přístupová pravidla.

Systém digitálního archivu a digitální spisovny se skládá z těchto softwarových komponent:

- **Katalog KDÚ**

V katalogu jsou evidovány jednotlivé logické segmenty těmito parametry:

- Název logického segmentu a textový popis významu uložených dat
- Původce dat v logickém segmentu, jeho kontaktní osoby
- Definice typu ukládaných dat a formátu datových souborů v rámci logického segmentu
- Definice přístupového protokolu
- Způsob řízení životnosti dat v logickém segmentu
- Definice ukládací politiky požadovaného způsobu uložení s ohledem na rychlost přístupu (má vliv na konfiguraci HSM).
- Definice skupin uživatelů oprávněných k přístupu k souborům daného logického segmentu.
- Podrobný popis souborových formátů (dokumentace, standard), kdo standard vydal a udržuje, kdo jiný standard ještě používá.
- Předpisy/normy podle kterých je třeba zajistit bezpečnost dat (osobní data, data chráněná autorským zákonem) v jednotlivých logických segmentech.
- Způsob kryptování, periodicitu obměny kryptovacích klíčů, dostupnost a způsob zajištění dostupnosti klíčů pro vybrané logické segmenty úložiště.

- **Úložiště**

- Obsahuje adresářovou strukturu vytvořenou na základě logických segmentů definovaných v katalogu KDÚ.
- Oprávnění přístupu k adresářové struktuře a uloženým datům jsou nastavena na základě definice oprávnění a kategorií podle údajů katalogu KDÚ.
- Předpokládá se realizace primárně s použitím technologie NAS s možností připojení CAS pro vybrané kategorie dat prostřednictvím HSM (viz. technologická architektura).

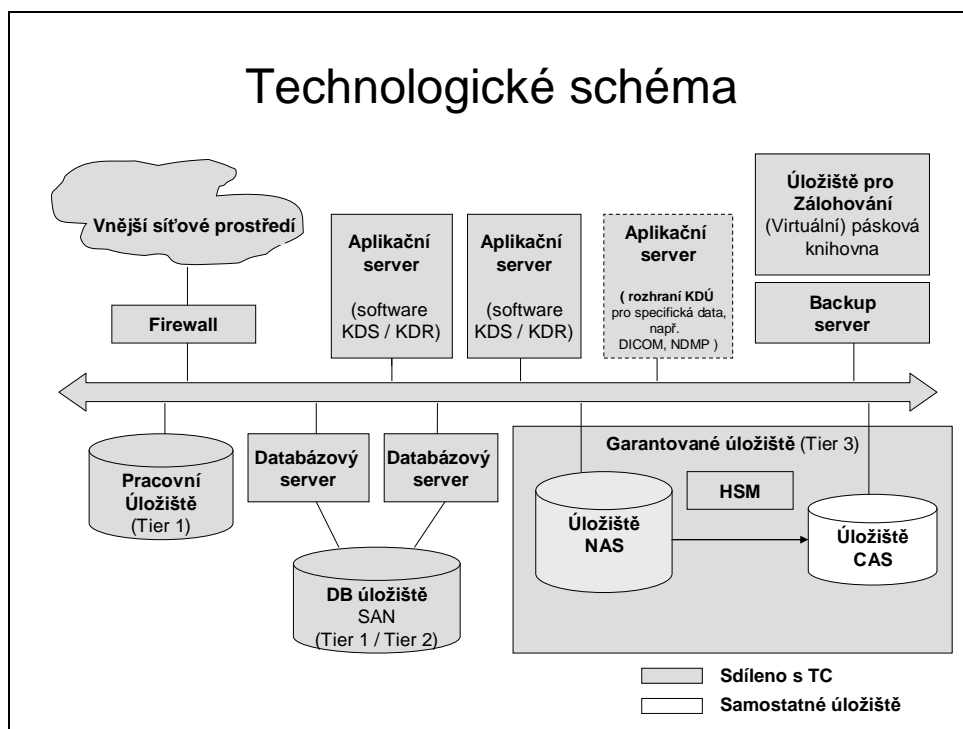
#### 5.4.1.4 Technologická architektura

Systém by měl v maximální možné míře využívat hw a sw vybavení nasazené v rámci projektu Technologického centra kraje (TC K).

V rámci tohoto projektu budou využity následující technologické části TC K:

- Aplikační servery pro obslužný SW subsystémů KDS a KDR
- Databázové servery využívané aplikacemi subsystému KDS a KDR
- Úložiště Tier 1 nebo 2 – pracovní prostory serverů a databáze subsystémů KDS a KDR
- Úložiště Tier 3 - technologie NAS a CAS (s použitím HSM) pro subsystémy KDS a KDR a KDÚ. Předpokládáme zajištění HA replikací úložiště Tier 3 na dvě geografické lokality
- Systém zálohování TC K pro systémy a pracovní prostory serverů a databáze.

- Síťová infrastruktura TC K a zabezpečení přístupu z Internetu.
- Autentizace uživatelů.



#### 5.4.1.5 Implementace subsystémů KDS a KDR

Digitální repozitář a digitální spisovna budou využívat specifické softwarové vybavení implementující příjem a správu dokumentů v intencích modelu OAIS. Tento software bude instalovaný na aplikačních serverech TC K, přičemž bude možné využít i virtualizace těchto serverů.

K obsluhým aplikacím KDS a KDR umístěným na aplikačních serverech bude možný vnější přístup pro jednotlivé původce dokumentů, kteří budou komunikovat prostřednictvím zabezpečeného kanálu (https) v rámci klientských aplikací a poskytovaných webových služeb.

Pro správu obsluhých dat a metadat uložených balíčků subsystémů KDS a KDR bude využit databázový server a záložní databázový server podle potřeb těchto aplikací.

Pro účely důvěryhodného uložení balíčků AIP subsystémů KDS a KDR, obsahujících obsah dokumentů a jejich metadata, bude obslužnou aplikací použito přímo úložiště typu CAS, jehož obsah bude kompletně replikován v záložní lokalitě. Obslužná aplikace pracuje s daty uloženými

v CAS i po jejich uložení (čtení obsahu po vyžádání, procesy pro zajištění důvěryhodnosti a konzistence).

Pro vybrané dokumenty subsystému KDS (s krátkou skartační lhůtou - určené na základě konfigurace obslužného software KDS) je vhodné alternativně použít i úložiště NAS (Tier 3).

### 5.4.1.6 Implementace subsystému KDÚ

Pro vybrané původce bude ve formě digitálního úložiště (KDÚ) zřízen zabezpečený přístup k jim přiděleným adresářům úložiště.

Pro tento účel bude využito primárně úložiště NAS (Tier 3) současně s použitím technologie HSM. Pro vybrané kategorie uložených dat (například s požadavkem na uložení typu WORM) může být prostřednictvím HSM obsah přesunut do úložiště CAS, které může být takto využíváno současně se subsystémy KDS a KDR.

Na základě popisu jednotlivých typů logických segmentů v katalogu KDÚ budou pro jednotlivé adresáře úložiště definovány tyto vlastnosti:

- autorizace přístupu do úložiště
- uzamykání souborů (využití WORM vlastností NAS a CAS)
- pravidla pro přesun souborů mezi úrovněmi uložení prostřednictvím HSM v rámci NAS, případně uložení v CAS.

## 5.5 Organizační a procesní část

### 5.5.1 Studie proveditelnosti

- Projekt bude v rámci studie proveditelnosti mimo povinných kapitol také řešit:
- Z pohledu služeb do území musí projekt v oblasti ukládání jako povinnou službu zajistit kapacity KDS a KDR pro obce a jimi zřizované nebo zakládané organizace specifikované v rámci studie proveditelnosti s doloženým stanoviskem všech ORP.
- 
- Analýzu potřeby digitalizace a ukládání datových fondů na území kraje,
  - prioritně s ohledem na potřeby měst a obcí v oblasti KDS a KDR.
- Časový a kapacitní plán digitalizace a ukládání
- Kapitola popisující partnerství v projektu – bude vycházet z doporučených rolí partnerů
  - Cíl partnerství v projektu
  - Identifikace partnerů v projektu



- Role v partnerství,
- Práva a povinnosti partnerů a žadatele
- Udržitelnost partnerství - délka a systém zajištění udržitelnosti partnerství (smlouva, vyjádření zájmu apod.)
- Stanoviska partnerů k projektu

### 5.5.2 Role kraje jako garanta projektu

- Zpracuje studii proveditelnosti projektu
- Realizuje zpracování plánovaných prací v oblasti digitalizace a ukládání dat (vlastními silami, organizací zřízenou nebo zakládanou krajem, nákupem služby)
- Zajistí funkci správce systému v rámci kraje
  - Zajistí organizační vazbu na partnery projektu
  - Zajistí realizaci automatizovaných vazeb v systému (zejména KDS)

### 5.5.3 Role organizací kraje

- Poskytují svá data a fondy ke zpracování podle harmonogramu a obsahu definovaného ve studii proveditelnosti projektu.
- Realizuje zpracování plánovaných prací v oblasti digitalizace a ukládání dat dle dohody s krajem.
- Užívají výstupů systému.

### 5.5.4 Role měst a obcí a jejich organizací

- Poskytují svá data a fondy ke zpracování podle harmonogramu a obsahu definovaného ve studii proveditelnosti projektu.
- Užívají výstupů systému – zpracování je v rámci udržitelnosti projektu.
  - v oblasti KDS zdarma.
  - v oblasti KDR a případně KDU za cenu, vykalkulovanou v rámci studie proveditelnosti, odsouhlasenou v rámci partnerské smlouvy.

### 5.5.5 Role stát

- Zajistí potřebné standardy systému zejména pro předávání dat mezi instancemi jednotlivých úložišť.
- Přebírá vstupy z KDS, dle dohodnutých procesů.
- Zpřístupňuje data a fondy zpracované na národní úrovni

- Využívá služeb projektu dle dalších dohod.

## 6 Postup realizace

Orientační harmonogram:

Činnost	Harmonogram
<b>Přípravná fáze</b>	<b>2009-2010</b>
Vyhlášení výzvy	Prosinec 2009
Zpracování studie proveditelnosti	Prosinec - Leden 2010
Zpracování a administrace žádosti o dotaci včetně podpisu partnerských smluv, zpracování studie proveditelnosti včetně regionálního rozměru (partneři)	Leden 2010
Termín pro podání projektu	31.1.2010
Výběrové řízení na realizaci včetně zpracování zadávací dokumentace	Leden 2010
<b>Realizace – investiční část</b>	<b>Únor – Říjen 2010</b>
<b>Provozní fáze</b>	<b>2010 - 2020</b>
Zajištění udržitelnosti projektu	5 let

## 7 Finanční analýza a plán

Projekt Digitalizace a ukládání bude řešen v rámci finanční podpory IOP oblasti intervence 2.1 - Zavádění ICT v územní veřejné správě, kdy výše podpory je 85% pro investiční část, finanční spoluúčast garanta projektu (kraj) je 15%. Provozní náklady jsou hrazeny garantem projektu po dobu jeho udržitelnosti.

### 7.1 Přehled celkových nákladů v investiční fázi

Náklady v investiční fázi jsou dvojího druhu - náklady na lidské zdroje a náklady na HW/SW investice.

#### 7.1.1 Náklady na lidské zdroje

V souvislosti s realizací projektu a rolí žadatele je nutné zajistit odpovídající lidské zdroje, zejména s ohledem na kvalifikaci a kompetence. Zástupci kraje budou odpovědní za:

- koordinaci projektu
- zpracování zadávací dokumentace a výběr zpracovatele na základě výběrového řízení,
- přípravu dokumentů k digitalizaci (výběr dokumentů, čištění, tvorba popisných metadat)
- digitalizace dat dle obsahu a harmonogramu projektu
- postprocesing (tvorba strukturálních, administrativních a technických metadat)
- zpřístupnění dat
- nezbytné úpravy ICT prostředí (SW/HW)

S ohledem na výše uvedené činnosti se jedná o kapacitní požadavky v rozsahu maximálně 4 úvazky ročně na kraj po dobu investiční fáze projektu, maximálně však 3 roky.

#### 7.1.2 Přehled celkových možných nákladů

Výstup	Maximální doba realizace	Maximální uзнatelné náklady
<b>Krajské digitální spisovna (KDS)</b>	<b>2 roky</b>	<b>15mil Kč</b>
Krajský digitální repozitář (KDR)	2 roky	20mil Kč
Krajské digitální úložiště (KDU)	2 roky	20mil Kč
Digitalizované a uložené dokumenty	3 roky	10mil Kč
Krajská digitalizační jednotka	2 roky	10mil Kč
		<b>75mil Kč</b>

### 7.2 Příjmy provozní fáze

Projekt bude jako celek neziskový.

### 7.3 Způsobilé výdaje projektu

- Nákup nebo upgrade HW a SW pro digitalizaci a dovybavení TC kraje
- Nákup nebo upgrade HW a SW pro ukládání a archivaci
- Implementace SW
- Náklady na digitalizaci

### 7.4 Udržitelnost projektu

Udržitelnost projektu je 5 let po ukončení investiční fáze.

Celý projekt je směřován k dlouhodobému uchovávání dokumentů. Dlouhodobost v případě některých dokumentů uložených v KDS může znamenat 50 - 100 let. V případě dokumentů ukládaných v KDR znamená dlouhodobost časově neomezený interval. Z tohoto pohledu je třeba chápat pojem "udržitelnost projektu" jako pojem omezený na konkrétní technologické vybavení, pořízené v rámci tohoto projektu.

V dlouhodobém výhledu je třeba počítat s tím, že po uplynutí fyzické a morální životnosti pořízené technologie bude třeba technologii vyměnit za novou, aktuální v budoucnosti. Vzhledem k vývoji informačních technologií, lze okamžik záměny technologie za novou těžko předpovídat předem. Podle současných zkušeností je možno odhadovat, že životnost technologií (morální) je 5 až 15 let.