

Centrálny dátový archív (CDA)

Univerzitná knižnica v Bratislave a TEMPEST vybudovali celonárodný digitálny archív. Slúži na záchranu a uchovávanie kultúrneho dedičstva. V archíve bude možné nájsť zdigitalizované historické listiny a dokumenty, digitálne kópie zbierok umeleckých či výtvarných diel i historické audiovizuálne záznamy.

25^{PB} / 75^{PB}

celková použiteľná kapacita archívu je 25 PB, celková hrubá úložná kapacita je 75 PB úložného priestoru (75 000 počítačov s kapacitou disku 1 000 GB)



systém kontroluje kvalitu dát i kvalitu metadát, overuje prítomnosť vírusov, prípadne škodlivých kódov v reálnom čase

381
miliónov

očakávaný objem uložených objektov je viac ako 381 miliónov

150^{TB}

pre spracovanie obrovského objemu dát poskytuje riešenie CDA celkovú hrubú priepustnosť na úrovni 150 TB týždenne

3x

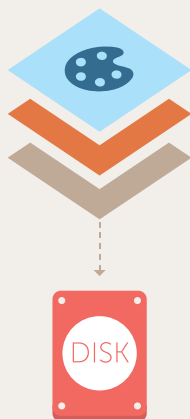
redundancia dát

archív realizuje trojnásobnú redundanciu dát z dôvodu ich zachovania v prípade katastrofy alebo výpadku

1^s

spracovanie, kontrola a uloženie štandardného dokumentu trvá menej ako 1 sekundu

BENEFITY



- jednotná, štandardizovaná metodika (OAIS) a technologické prostredie pre spracovanie a ukladanie digitálnych záznamov
- trojnásobná redundancia dát v geograficky oddelených lokalitách
- validácia, kontrola kvality a viacnásobná ochrana dát na všetkých úrovniach operácií
- flexibilita formátov ukladaných dát prispôsobená požiadavkám jednotlivých vkladateľov
- flexibilita vkladania údajov (on-line alebo prostredníctvom logistického riešenia)
- riešenie je pripravené na certifikáciu podľa normy ISO 16363:2012 a prevádzkujúca organizácia je pripravená získať certifikát ISO 27000

Hlavným prínosom projektu je dlhodobé uchovanie a ochrana kultúrneho dedičstva, v niektorých prípadoch nenahraditeľného rozsahu a významu, pre budúce generácie. V nemalej miere je prínosom aj inštitucionalizácia dlhodobého digitálneho archívu spĺňajúceho medzinárodné normy a štandardy v oblasti. Prínosom je definovanie legislatívneho, procesného a metodického rámca pre dlhodobé uchovanie digitálnych objektov, ich sprístupňovania a spracovania. V priebehu realizácie projektu sa tvorí komunita odborníkov pre oblasť digitalizácie a uchovania digitálnych objektov, ktorej know-how významne prispeje k rozvoju využitia danej oblasti pre širokú verejnosť. Realizáciou tohto projektu sa Slovensko dostáva na špičku v oblasti dlhodobého uchovania digitálnych dát v rámci EÚ.

Zapojené inštitúcie

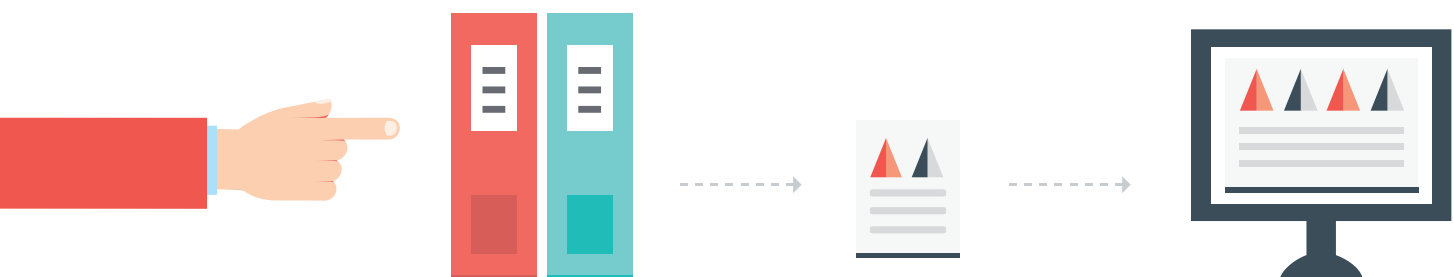
Súčasnými používateľmi služby sú významné inštitúcie v oblasti digitalizácie kultúrneho dedičstva:

- **Pamiatkový úrad SR**
- **Múzeum Slovenského národného povstania**
- **Slovenská národná galéria**
- **Slovenská národná knižnica**
- **Štátna vedecká knižnica v Prešove**
- **Slovenský filmový ústav**

Cieľovou skupinou je komunita inštitúcií, ktoré realizujú digitalizáciu kultúrneho dedičstva Slovenska. Archív je pripravený na otvorenie pre ďalšie inštitúcie, prípadne pre komerčné alebo nekomerčné subjekty aj mimo oblasti kultúry.

Univerzitná knižnica v Bratislave (UKB)

Univerzitná knižnica v Bratislave (UKB) je najstaršia a najväčšia vedecká knižnica v Slovenskej republike. Vznikla v roku 1919 ako Knižnica Univerzity Komenského. Súbežne so svojím akademickým poslaním plnila do roku 1954 funkciu národnej knižnice. Od roku 1954 je samostatnou vedeckou knižnicou slúžiacou širokej verejnosti. Koncom roka 2013 fond UKB obsahoval v súhrne 2 650 tis. knižničných jednotiek.



Situácia pred projektom

Do roku 2012 neexistovala spoľahlivá, primeraná a udržateľná infraštruktúra na dlhodobú archiváciu. Operačný program Informatizácia spoločnosti (OPIS) a jeho prioritná os 2 pokrývajú cca. 3,5 mil. kultúrnych objektov, na ktoré nebolo predchádzajúce riešenie pripravené. Ministerstvo kultúry SR charakterizovalo existujúce podmienky ako nevyhovujúce a nedostatočné – či už ide o parametre technického vybavenia, alebo nízku úroveň interoperability a šandardizácie, bezpečnosti i jednotného metodického riadenia príslušných pracovných postupov.

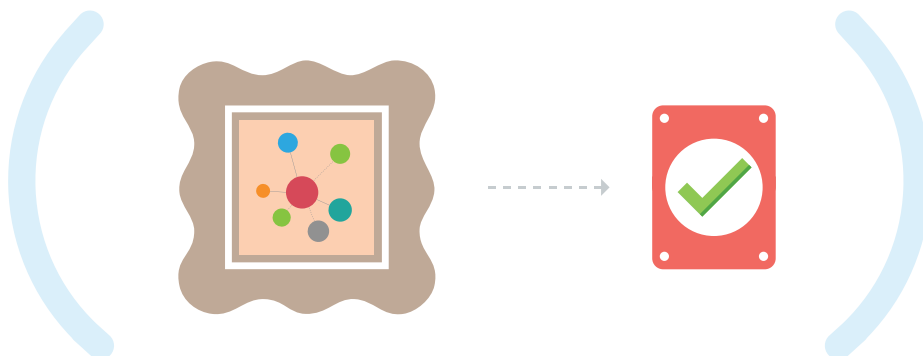
Fragmentované úložiská údajov, ktoré sú súčasťou informatizačných projektov jednotlivých pamäťových a fondových inštitúcií (PFI) z uplynulých rokov (KIS3G, CEDVU, CEMUZ a pod.), boli zastarané, výkonovo poddimenzované a kapacitne nedostatočné. Východiskovou situáciou pre CDA tak bola absencia dôveryhodného dátového úložiska pre nabiehajúce a pripravované digitalizačné projekty a funkčného riešenia komplexného a spoľahlivého zálohovania programových systémov a údajov.

Ciele projektu



- **vybudovanie dátových úložísk v stanovených lokalitách**
- **záznam, zbieranie, dlhodobé archivovanie a ochrana digitálneho obsahu**
- **realizácia webharvestingu a webarchivingu**
- **održanie definovaných parametrov, ukazovateľov, kritérií a štandardov**

Cieľom projektu bolo vybudovať komplexný integrovaný systém dlhodobého uchovávanía a ochrany digitálneho obsahu, jeho získavania, spracovania, ochrany a využitia. V rámci projektu sa vybudoval centrálny dátový archív, ktorý v zmysle predpísaných štandardov zabezpečuje uloženie kópií digitálnych objektov v najmenej dvoch geograficky oddelených lokalitách vzdialených minimálne 50 km. Lokality sú vybavené príslušným HW i SW vybavením a prostredníctvom dátových telekomunikačných sietí sú vzájomne prepojené. Súčasťou dátového archívu je aj pasívny sklad pamäťových médií, ktorý je umiestnený v objekte UKB. Tieto ciele vytvárajú bázu pre systematickú podporu digitalizácie kultúrneho, vedeckého a intelektuálneho dedičstva, ktoré budú v rámci digitalizačných projektov a kampaní riešiť príslušné pamäťové a fondové inštitúcie (PFI).



Základné technické parametre projektu a schémy

ŽIVOTNÝ CYKLUS DOKUMENTU, ŠTANDARDY A OCHRANA

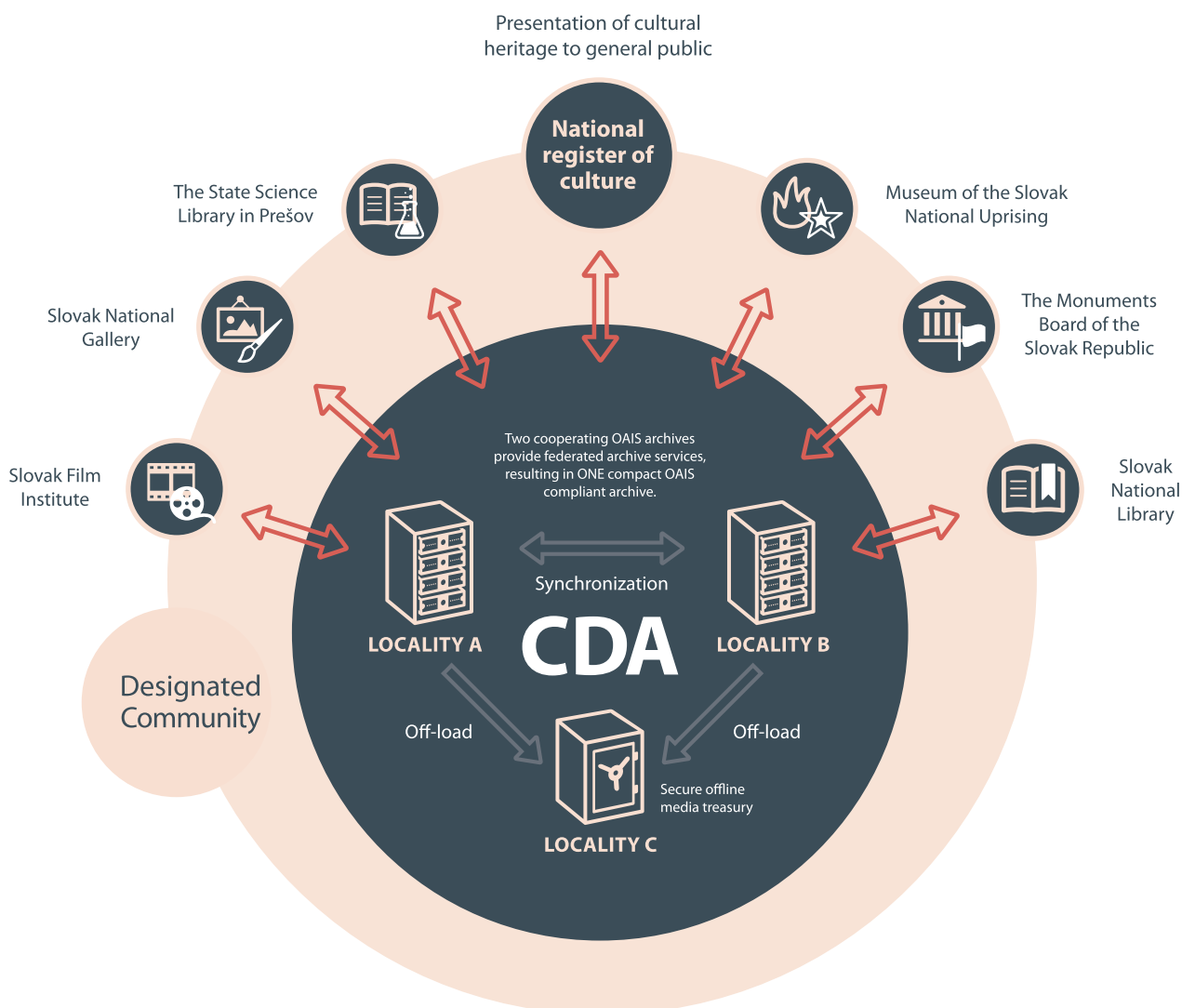
Riešenie poskytuje dlhodobé archivačné služby digitalizovaného alebo digitálneho obsahu v zmysle medzinárodného štandardu OAIS (Open Archival Information System). Hlavnou myšlienkou dlhodobého archívu nie je len uchovanie dát, ale najmä zachovanie ich čitateľnosti a zrozumiteľnosti. CDA sa preto sústreďuje na sledovanie životného cyklu uložených formátov, ich konverziu alebo emuláciu a zároveň na bitovú ochranu. Súčasťou riešenia CDA je aj podpora vkladateľov v rámci výberu a identifikácie formátov vhodných na archiváciu, zabezpečenie legislatívnej podpory pre uchovanie dát a samozrejme variability pri výbere dát. Súčasťou CDA je aj ochrana duševného vlastníctva v zmysle platnej legislatívy a bezpečnosť uložených objektov.

REDUNDANCIA

Každý údaj je uložený na troch nezávislých médiách, ktoré sú umiestnené v troch geograficky oddelených lokalitách. Dve z týchto lokalít sú živé systémy, ktoré môžu plniť požiadavky komunity kladené na služby. Tretia lokalita je bezpečné úložisko vytvorené na dlhodobú ochranu médií.

KAPACITA

Hrubá úložná kapacita: 75 PB * Využitelná úložná kapacita: 25 PB * Očakávané objekty: 381 346 900 * Škálovateľnosť systému: neobmedzená



KOMPATIBILITA S OAIS (OTVORENÝ ARCHÍVNY INFORMAČNÝ SYSTÉM)

CDA je vybudovaný ako riešenie, ktoré plne vyhovuje požiadavkám OAIS. Pozostáva z dvoch samostatných archívov, pričom každý z nich vyhovuje OAIS. Tieto archívy fungujú ako spolupracujúce a združené archívy, ktoré ponúkajú plne kompatibilné a uzavreté riešenie. Údaje sú pravidelne synchronizované, čo zabezpečuje stabilnú redundanciu a dostupnosť. Riešenie je pripravené na certifikáciu podľa normy ISO 16363:2012 a prevádzkujúca organizácia je pripravená získať certifikát ISO 27000. Všetky interné procesy sú vytvorené tak, aby rešpektovali a dodržiavali súlad s odporúčaniami a požiadavkami rámca ITIL (knížnica infraštruktúry informačných technológií).

VYSOKÁ FLEXIBILITA

Každý člen komunity podpísal s CDA formálnu dohodu o zadávaní údajov. Táto dohoda obsahuje formáty údajov, validácie, konverzie a podrobnosti o spracovaní údajov, ktoré sú šité na mieru konkrétnym potrebám každého člena. Tieto podrobné údaje sa neskôr premieňajú na konkrétny archivačný profil.

LOGISTIKA

Údaje je možné zadávať cez internet (online) alebo prostredníctvom špecializovaného logistického systému, ktorý sa poskytuje ako súčasť riešenia CDA. Zásielka získava záruku vhodných podmienok prostredia pre prepravované médiá a správnej manipulácie a ochrany médií. Logistický systém ďalej poskytuje vhodné balenie médií a správne označenie balenia. Logistický systém sa orientuje na manipuláciu s magnetickými páskami, no riešenie je dostatočne otvorené na zavedenie nových dátových nosičov podľa potrieb v budúcnosti.

BEZPEČNOSŤ

Všetka online komunikácia externých subjektov je chránená šifrovaním a všetky citlivé operácie sú chránené dvojfaktorovou autentizáciou. CDA prevádzkuje svoj vlastný certifikačný orgán, ktorý ponúka služby PKI (infraštruktúry verejných kľúčov) všetkým členom určenej komunity. Riešenie CDA ďalej implementuje stabilné mechanizmy pre podpisovanie digitálneho obsahu a tak zaručuje vysokú konzistentnosť a autentickosť údajov.

OCHRANA ÚDAJOV

Okrem niekoľkonásobnej redundancie údajov na fyzickej úrovni CDA aplikuje moderné a technologicky vyspelé zásady logickej ochrany údajov, vďaka ktorým zabezpečuje dlhodobé zachovanie údajov z hľadiska dostupnosti, kvality a spoľahlivosti. Nastaviteľné postupy spracovávania založené na plug-in architektúre ponúkajú rozšíriteľnosť a budúce zlepšenie využiteľnosti. Rovnaké zásady sa aplikujú na periodické LTP procesy (procesy dlhodobej potenciácie), riadenie formátu, plánovanie uchovania a zhodnotenie rizík. Vďaka moderným nástrojom na riadenie plánov na uchovávanie ponúkajú riešenia svoj vlastný systém pre riadenie formátov údajov. Tento nástroj zhromažďuje informácie o formátoch spracovávaných v archívoch z rôznych zdrojov a externých databáz a umožňuje manuálne vyhodnocovanie a spracovávanie formátov, ktoré sú obzvlášť nové a nie sú všeobecne známe. Systém takisto umožňuje riadenie celého životného cyklu formátu od fázy plánovania až po koniec životnosti daného formátu a jeho konverziu.

VYSPELÉ TECHNOLOGIE

Riešenie CDA je vybudované na vyspelých technológiách s využitím všetkých moderných zásad a architektúr, ktoré sa využívajú na spracovávanie objemných údajov. Hlavným zámerom tohto prístupu je zaručiť budúce zlepšovanie riešenia, ktoré bude schopné naplniť budúce potreby rastu, rozšírenia a modernizácie. Hlavnými cieľmi počas budovania architektúry bolo zaručenie modularity, škálovateľnosti, robustnosti a rozšíriteľnosti.

OTVORENÝ SYSTÉM

Riešenie CDA je vybudované ako otvorený systém a je kombináciou open source (otvoreného zdroja) a nástrojov na podnikovej úrovni.

Partneri



 **Tempest**

I T m a k e s s e n s e

TEMPEST a.s.
GBC IV
Galvaniho 17/B
821 04 Bratislava 2
Slovenská republika

Tel.: +421 2 502 67 111
Fax: +421 2 502 67 100

info@tempest.sk
www.tempest.technology