

VPLIV INFORMACIJSKE TEHNOLOGIJE V PROCESU PRETVORBE DOKUMENTACIJE V DIGITALNO OBLIKO IN ARHIVIRANJA

Aleš Kalan *

UDK: 930.25:004.3

Aleš Kalan: Vpliv informacijske tehnologije v procesu pretvorbe dokumentacije v digitalno obliko in arhiviranja. Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Zbornik referatov z dopolnilnega izobraževanja, Maribor 5/2006, št. 1, str. 399-404.

Izvirnik v slovenščini, izvleček v slovenščini in angleščini, povzetek v angleščini.

Prispevek obravnava strojno in programsko opremo, potrebno za prehod iz klasičnega v elektronsko arhiviranje.

UDC: 930.25:004.3

Aleš Kalan: The impact of information technology in the process of the conversion of records to digital form and in the process of archiving. Technical and Field Related Problems of Traditional and Electronic Archiving. Conference Proceedings, Maribor 5/2006, No. 1, pp. 399-404.

Original in Slovenian, abstract in Slovenian and English, summary in English.

This paper outlines the current hardware and software solutions in the field of electronic data storage that can be used for the transition from traditional to electronic methods of archiving paper documents.

ZAKAJ ELEKTRONSKI ARHIV

Prednosti elektronskega arhiva:

- hiter in enostaven dostop do dokumentov;
- nadzor nad dostopom in enostavno upravljanje z dokumenti;
- varnost pred uničenjem;
- zmanjšanje arhivskega prostora in nižji stroški poslovanja;
- centralen arhiv vseh dokumentov;
- lažja izmenjava dokumentov;
- drugo.

Za lažje razumevanje znižanja stroškov in povečanja uspešnosti poslovanja po prehodu iz klasičnega v elektronski arhiv v razmislek podajam nekaj vprašanj in ugotovitev:

- Koliko dokumentov, s katerimi imate opravka vsak dan, je pomešanih ali izgubljenih? Koliko časa porabite za iskanje teh dokumentov?
- Koliko časa porabite za razvrščanje in razpošiljanje dokumentov?

Aleš Kalan, univ. dipl. organizator dela, MFC&L d.o.o. Ljubljana, 1000 Ljubljana, Slovenija:

- S prehodom na elektronsko arhiviranje dokumentov se obseg klasičnega arhiva občutno zmanjša.
- Stroški hranjenja podatkov in informacij na papirju so cca. za faktor 75 višji kot hranjenje iste količine podatkov in informacij na optičnem disku.

ZAKONODAJA

Ravnanje z arhivskim gradivom (beri: gradivo kulturne in zgodovinske vrednosti) opredeljuje Zakon o arhivih in arhivskem gradivu (ZAGA), ki je hkrati tudi temeljno izhodišče za obvladovanje dokumentarnega gradiva. Obstoječo zakonodajo bo nadomestil, predvsem z vidika pretvorbe papirne dokumentacije v elektronsko, Zakon o varstvu dokumentarnega in arhivskega gradiva in arhivih (ZVDAGA).

Zakon o elektronskem poslovanju in elektronskem podpisu (ZEPEP) s pripadajočo Uredbo o pogojih za elektronsko poslovanje in elektronsko podpisovanje na splošno ureja področje elektronskega poslovanja in način uporabe elektronskega podpisa.

Zakon pravi, da se podatki lahko hranijo v elektronski obliki, če so dosegljivi in primerni za kasnejšo uporabo, če so shranjeni v obliki, v kateri so bili oblikovani, poslani ali prejeti ali v kakšni drugi obliki, ki verodostojno predstavlja oblikovane, poslane ali prejete podatke in če je iz shranjenega elektronskega sporočila možno ugotoviti, od kod izvira, komu je bilo poslano ter čas in kraj njegovega pošiljanja ali prejema in če uporabljena tehnologija in postopki v zadostni meri onemogočajo spremembo ali izbris podatkov oziroma obstaja zanesljivo jamstvo glede nespremenljivosti sporočila.

Delno se na arhiviranje v elektronski obliki navezujeta tudi Zakon o davčnem postopku in Zakon o davku na dodano vrednost, saj predpisujeta časovno hrambo gradiva.

Poleg domače zakonodaje je potrebno upoštevati tudi mednarodne standarde in priporočila, kot so: Splošni mednarodni standardi za arhivsko podpisovanje - ISAD/G, Mednarodni standardi za arhivski zapis o ustvarjalcih arhivskega gradiva pravnih oseb, fizičnih oseb in družinah - ISAAR/CPF, Mednarodni standard za pisarniško poslovanje - ISO 15489 in Priporočila za shranjevanje in razstavljanje arhivskega gradiva - BS 5454.

ELEKTRONSKI ZAJEM IN ARHIVIRANJE DOKUMENTOV

Elektronski zajem in arhiviranje dokumentov (dobesedni prevod angleškega izraza imaging and archiving je računalniško upodabljanje in arhiviranje dokumentov) obsega ponavljajoče se faze skeniranja, indeksiranja in uvažanja dokumentov, ter dalje shranjevanja, iskanja in nadzora dostopa do dokumentov. Poleg strojne potrebujemo tudi ustrezno programsko opremo, ki mora podpirati temeljne funkcije za kvalitetno pretvorbo papirnih izvornikov v digitalne slike dokumentov, prepoznavanje posameznih znakov s slike in s tem pretvorbo podatkov v tekstovno obliko ter prenos le-teh v podatkovne baze in sisteme za upravljanje z dokumenti:

Za skener:

- kontrola skeniranja in nadzor delovanja skenerja;

- kontrola kakovosti oziroma vidnosti dokumenta;
- izboljšanje vidnosti dokumenta in poravnava strani;
- možnost izbire načina skeniranja (določitev resolucije, enobarvno/barvno, multi/singel page tiff itd.).

Vnos meta podatkov:

- možnost vpisa dodatnih opomb kot sestavni del skeniranega dokumenta;
- optično prepoznavanje znakov (OCR, ICR, DCR, črtne kode itd.), iskanje po polnem besedilu (Full Text Search).

Dokumentacijski sistemi za upravljanje dokumentov in procesov:

- iskanje in pregledovanje dokumentov;
- arhiviranje oziroma povezljivost s podsistemom arhiviranja.

SKENIRANJE

Prvi korak v sistemu računalniškega upodabljanja je pretvorba papirnih dokumentov v digitalne slike papirnih izvirnikov (skeniranje). Ko je dokument ustrezno shranjen, je pripravljen za nadaljnjo obdelavo.

PESTRA PONUDBA STROJNE OPREME ZA ZAJEM PODATKOV

Na tržišču obstajajo različni ponudniki skenerjev, namenjenih pisarniškemu ali masovnemu zajemu dokumentov, npr. Canon, Kodak, Bell&Howell, Fujitsu, ki ponujajo skenerje različnih zmogljivosti glede na potrebe.

Nizko produkcijski skenerji, katerih zmogljivost je do 1200 dokumentov dnevno so namenjeni za pisarniško poslovanje. Produkcijski skenerji, katerih zmogljivost je do 7500 dokumentov dnevno, so namenjeni večjim sprejemnim pisarnam in rešitvam za zajem podatkov. Visoko produkcijski skenerji, katerih zmogljivost je nad 10.000 dokumentov dnevno, so namenjeni masovnemu zajemu podatkov in imajo vgrajene varovalne mehanizme za nadzor pravilnosti transporta dokumentov ter napredne procesne enote za izboljšanje kakovosti zajetih dokumentov.

Mikrofilm je še zmeraj edini medij, ki zagotavlja, da bodo dokumenti, kljub vsem spremembam tehnologije in standardov, ostali varni, logično berljivi in shranjeni za dolgo dobo, poleg tega pa je strošek vzdrževanja in hrambe samega medija minimalen.

Hibridna rešitev: zajemanje dokumenta in pretvorba papirnega dokumenta v digitalno sliko ter vzporedna preslikava na mikrofilm, posredujeta uporabniku vse prednosti in udobnost računalniške obdelave dokumentov shranjevanja na magnetnih, optičnih diskih in arhiviranja na mikrofilmu.

VNOS META PODATKOV

Ročno vnašanje meta podatkov (uporablja se tudi izraz indeksiranje) je označevanje slike dokumenta s pomočjo ključnih besed (indeksni atributi). OCR (Optical Character Recognition) je postopek, ki nadomesti ročno vnašanje podatkov.

Gre za proces prepoznavanja posameznih znakov, s katerim iz slike dokumenta v grafičnem načinu dobimo podatke v tekstovni obliki.

Sistemi ICR (Intelligent Character Recognition) so poleg različnih strojnih pisav zmožni prepoznavati tudi rokopise, tehnologija OMR (Optical Mark Recognition) prepozna označbe na obrazcih, medtem ko so sistemi DCR (Document Content Recognition), sposobni identificirati in razumeti vsebino posameznega dokumenta in podatke preusmeriti v ustrezno obdelavo.

Najzanesljivejša pa še vedno ostaja tehnologija za prepoznavanje črtnih kod (Barcode Recognition). V bazo podatkov vnesemo poljubne podatke z dokumenta, dalje tem meta podatkom določimo nek šifrant, ki ga zapišemo na črtno kodo. Črtna koda nam tako omogoča hitrejše in natančnejše zajemanje podatkov.

Način prepoznavanja znakov je odvisen od strukturiranosti dokumentov. Ti so lahko strukturirani - podatki, ki jih želimo zajeti, so vedno na istem mestu (npr. obrazci), delno strukturirani dokumenti vsebujejo potrebne podatke, a se ti nahajajo na različnih mestih (npr. prejeti računi dobaviteljev), medtem ko pri nestrukturiranih dokumentih ne vemo, kateri podatki so pomembni za zajem in kje se nahajajo (npr. tehnična dokumentacija). Največ dokumentarnega gradiva ima delno ali nestrukturirano obliko dokumenta, kar pomeni, da je potrebno v bazo podatkov »ročno« vnašati meta podatke dokumenta. Tako vnašanje meta podatkov je tudi najzanesljivejše.

ARHIVIRANJE DOKUMENTOV V ELEKTRONSKI OBLIKI

To je zadnja in najpomembnejša faza v sistemu za celovito upravljanje z dokumenti. Dokumentacijski sistemi za upravljanje dokumentov in procesov, omogočajo učinkovito ter uspešno obvladovanje velikih količin dokumentov ne glede na njihov izvor (papirna ali elektronska oblika). Namen dokumentacijskih sistemov je shranjevanje dokumentov v digitalni obliki za potrebe arhiva, omogočajo lahek in varen dostop do dokumentov, vodijo evidenco dokumentov v klasični in elektronski obliki, omogočajo izdelavo kopij dokumentov, workflow (pretok dokumentov) itd.

NOSILCI ELEKTRONSKIH ZAPISOV

Za zdaj še nimamo elektronskega nosilca, ki bi bil trajno ali dolgo časa uporaben oziroma berljiv. Za vsako vrsto nosilcev zapisa so pogojeni specifični načini materialnega varstva nosilca ter različno dolga uporabnost in življenjska doba. Zaradi krhkosti postane mikrofilm po 40 letih neuporaben za kopiranje, za magnetne in optične zapise je značilno, s stališča hitro spreminjajoče se strojne in programske opreme, razmeroma kratko obdobje obstojnosti in uporabnosti, medtem ko papir po trajnosti vzdrži pri normalnih klimatskih pogojih več kot 500 let, a ima druge praktične pomanjkljivosti. Poleg tehnoloških zahtev je potrebno upoštevati tudi zakonske zahteve in priporočila, ki predvidevajo varno hrambo in dostopnost do podatkov ne glede na nosilec zapisa.

Časovna in krajevna neodvisnost za dostop do dokumentov, poenostavljen nadzor nad uporabo, zanemarljivi stroški posredovanja, razmnoževanja in razpošiljanja dokumentov, zagotavljanje aktualnosti dokumentov, možnost prenosa na druge nosilce zapisov in hibridne rešitve (kombinacija hranjenja) so le nekateri od poglobitnih razlogov za hranjenje dokumentarnega gradiva na sodobnih nosilcih.

Magnetno-optični disk

Gre za kombinacijo magnetne in optične tehnologije. Laser segreje disketo, kar omogoča magnetenje površine za zapisovanje, za branje pa se uporablja laser manjše intenzitete. Število brisanj in sprememb zapisov je neomejeno kar za dolgoročno hrambo gradiva ni primerno.

Compact disk (CD) - zgoščenska

Podatki na zgoščenci so zapisani v obliki mikroskopskih utorov na reflektivno preplasteni površini polikarbonatskega diska. Zapis se bere s pomočjo infrardečega laserja, ki prepozna utore in ravnine na površini diska.

Digital Versatile Disc (DVD)

DVD je na videz zelo podoben starejšemu sorodniku, CD-ju. Pomembna razlika med njima je v gostoti zapisa. Izboljšan je tudi način strnjevanja podatkov, kar je privedlo do boljše izrabe in optimizacije zapisov. Obstaja možnost dvostranskega in dvostranskega zapisa. CD-ji in DVD-ji so cenovno najugodnejši in najprimernejši predvsem za distribucijo, za aktivno arhiviranje pa so prepočasni.

Write Once Read Many (WORM) disk

Trenutno najbolj priporočljiv nosilec zapisa. Ker omogoča le enkratni zapis in neomejeno branje, je primeren predvsem za dolgoročno oziroma trajno hrambo gradiva. Za branje zapisov je potrebna posebna programska in strojna oprema.

Fizik Kurt Gerecke, ki je v nemškem oddelku IBM-a zadolžen za preizkušanje medijev, namenjenih shranjevanju in arhiviranju podatkov, je pri testiranju CD-R in CD-RW medijev ugotovil, da se podatki, ki so bili na ploščke shranjeni z običajnim CD-zapisovalnikom, ohranijo v povprečju le dveh do pet let (odvisno od kakovosti medija). Za kratko življenjsko dobo CD medijev so po mnenju nemškega fizika "krivi" ravno CD-zapisovalniki, ki pri zapisovanju podatkov zaradi visokih temperatur poškodujejo oziroma deformirajo kovinsko plast zgoščenk.

Vse težave nosilcev elektronskih zapisov rešuje prostorsko naslovljivi diskovni sistem (CAS - Content Addressed Storage), načrtovan za posebne zahteve nespremenljivih digitalnih dokumentov, ki jih je potrebno dolgotrajno hraniti zaradi različnih zakonskih predpisov. Je prva WORM naprava zasnovana na magnetnih diskih, ki je zgrajena glede na najstrožje zakonske zahteve. Je optimalna rešitev za končne dokumente in zagotavlja funkcionalnost, ki do sedaj ni bila na voljo pri trakovih, optičnih, oziroma tradicionalnih diskovnih rešitvah.

MOŽNOST IZVAJANJA STORITVE S POGODBENIM PARTNERJEM

V Sloveniji je v zadnjem času vse več podjetij, ki pogodbeno opravljajo skeniranje in mikrofilmanje dokumentacije. Za organizacije je ustrežnejše in ceneje, če za elektronsko zajemanje in arhiviranje podatkov najamejo zunanje izvajalce, ki na tem področju nudijo rešitve po dogovorjenih tehničnih standardih in v skladu s predpisanimi pogoji bolj kakovostno in ceneje, kot to lahko opravlja sama organizacija, kateri tako ni potrebno vlagati v lastno tehnologijo in usposabljanje

kader, ampak lahko vse svoje znanje usmeri v razvoj svoje osnovne dejavnosti, ker je skrb za izvedbo prepuščena usposobljenim strokovnjakom, poleg tega je čas od določitve zahtev do uvedbe rešitve zelo kratek, saj usposobljen zunanji izvajalec razpolaga s tehnologijo, znanjem in izkušnjami za izvedbo storitve.

LITERATURA IN VIRI

- *Zakon o arhivskem gradivu in arhivih /ZAGA/, UL RS 20/97, 32/97.*
- *Zakon o elektronskem poslovanju in elektronskem podpisu /ZEPEP/, UL RS, 57/2000, 30/2001.*
- *ISAD/G (General International Standard Archival Description - Splošni mednarodni standardi za arhivsko poslovanje).*
- *ISAAR/CPF (International Standard Archival Authority Record for Corporate Bodies, Persons and Families - Mednarodni standardi za arhivski zapis o ustvarjalcih arhivskega gradiva - pravnih osebah, fizičnih osebah in družinah).*
- *ISO 15489 (Information and Documentation - Records Management - Navodila in dokumentacija - pisarniško poslovanje).*
- *BS 5454 (Recommendation for the Storage and Exhibition of Archival Documents - Priporočila za shranjevanje in razstavljanje arhivskega gradiva).*
- *Mag. Aljoša Jerman Blažič: Dolgoročno varno arhiviranje poslovnega gradiva; DOK_SIS 2005 - Kranjska gora, 14. do 16. september 2005.*
- *Bine Žerko: Računalniško upodabljanje in arhiviranje dokumentov; Sodobni arhivi 2002.*
- <http://www.mfc-l.si/>
- <http://www.kodak.com>
- <http://www.src.si/>
- <http://www.centera.si>
- <http://www.wikipedia.org>

SUMMARY

THE IMPACT OF INFORMATION TECHNOLOGY ON THE PROCES OF CONVERSION OF RECORDS TO DIGITAL FORM AND ON THE PROCESS OF ARCHIVING

Trough the implementation of an electronic system of data storage, companies can, among other things, achieve lower business costs, a quick and easy access to important information, as well as gain a more secure method of storing their written documents. The existing Information Technology in the field of electronic data storage offers a variety of hardware and software solutions for imaging and archiving paper documents, the process of which is closely governed by legislature that must be followed in the transforming of written documents into electronic form, and in their subsequent storage. In the process of the transformation we use scanners to make a digital image of the paper document, which is then stored in the form of meta data in a database. The document in its electronic form can then easily and quickly be retrieved from the database when it is needed.

Aleš Kalan, univ. dipl. organizator dela, je zaposlen v podjetju MFC&L d. o. o. kot pomočnik tehničnega direktorja. Na področju elektronskega zajemanja in arhiviranja dokumentarnega gradiva ima 6-mesečne izkušnje.