

Migracija wordovih dokumentov

Tomaz Stensak¹, Tomaz Dogša²

¹ BSH, d. o. o., tomaz.stensak@bshg.com

² Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, tdogsa@uni-mb.si

Izvelek

Digitalni dokumenti bodo vsaj deloma nadomestili papir samo v primeru, ko bo zagotovljeno, da bodo berljivi prostorsko in časovno neodvisno. Berljivost starejših digitalnih dokumentov ohranjamo z ustrežno migracijo. V članku so predstavljeni rezultati testiranja migracije dokumentov, ki so bili tvorjeni s starejšimi različicami urejevalnika Word. Pri migraciji dokumentov, tvorjenih z urejevalnikom *Word 5.5* (različica za DOS), se je popolnoma ohranilo le povprečno 20 odstotkov obravnavanih lastnosti. Testiranje je pokazalo, da je za te dokumente najbolj primerna postopna migracija. Zelo uspešno je bilo mogoče migrirati dokumente, ki so bili tvorjeni z urejevalnikom *Word 2.0* (različica za Windows), pri katerih se je popolnoma ohranilo povprečno 89 odstotkov obravnavanih lastnosti.

Ključne besede: migracija, digitalni dokumenti, vzdrževanje dokumentov, ohranjanje dokumentov, formati dokumentov.

Abstract

Migration of Word Documents

Digital documents will, at least partially, replace paper documents only if they remain readable in the same way regardless of time and space, which means that old documents need to be transformed. Readability of older digital documents can be kept by proper migration. In this article are presented the results of old Word documents migration tests. On average, after migration of documents created in *Word 5.5* (version for DOS) only 20 % of document properties were completely preserved. The test shows that gradual migration is most suitable for these documents. The migration of documents created in *Word 2.0* (version for Windows) was very successful with on average 89 % of properties being completely preserved.

Key words: migration, digital documents, document maintenance, document preservation, document formats.

1 UVOD

Prvi digitalni dokumenti so bile preproste besedilne datoteke, ki so vsebovale samo golo besedilo. Z razvojem urejevalnikov se je kompleksnost strukture dokumentov neprestano večala. Zaradi vključevanja raznih objektov (slik, formul itd.) v dokument se format zapisa še danes neprestano spreminja. Word 2003 je poznal sedem različnih formatov, verzija 2007 pa jih ima že petnajst. Ker število dokumentov skokovito narašča, jih je treba ustrezno hraniti in omogočiti preprosto iskanje, branje in urejanje. Ali bo mogoča uporaba sedanjih oz. še starejših dokumentov tudi po daljšem obdobju (npr. čez 20 let)? Povprečnemu uporabniku takrat današnja programska in strojna oprema ne bosta več uporabni niti dostopni.

Načinov ohranjanja digitalne dokumentacije je več. Najbolj preprosta metoda je prenos na papir, ki pa od treh zahtev (branje, prostorsko in časovno neodvisno iskanje, urejanje) zagotavlja samo branje. Postopek, pri katerem dokumente, ki so bili tvorjeni v starem formatu, pretvorimo v novejši format, imenujemo migracija. Migracija je eden najbolj priljubljenih in razširjenih načinov ohranjanja digitalnih dokumentov [2].

Rezultat idealne migracije je dokument, ki je berljiv, ima ohranjeno integriteto v semantičnem in oblikovnem smislu in ga je mogoče ponovno urejati. Če je ciljni format kompatibilen z izvornim, se prenesejo vse pomembne lastnosti digitalnega dokumenta v novo različico. Tako je ohranjena verodostojnost in integriteta dokumenta.

Ker je treba učinkovitost vsake migracije predhodno preveriti, smo analizirali učinkovitost migracije dokumentov, ki so bili tvorjeni s starejšimi različicami urejevalnika besedil Word. Uspešnost migracije odvisna tudi od strukture dokumentov. Dokument, ki vsebuje samo preprosto besedilo, lahko v večini primerov brez težav prenesemo v novejši format. Zato smo za testne dokumente izbrali takšne, ki so bili sestavljeni iz treh segmentov: slogovno oblikovanega besedila, drugih objektov (slik, formul itd.) ter podatkov o dokumentu (avtor, datum nastanka itd.). V nadaljevanju so najprej predstavljeni primeri iz splošne prakse.

2 MIGRACIJA V PRAKSI

2.1 Problematika migracij digitalnih dokumentov

Migracija je periodično pretvarjanje digitalnih vsebin iz ene strojne ali programske konfiguracije v drugo ali iz predhodnih generacij računalniških tehnologij v kasnejše. [8] Glavni problem migracije je, da zahteva specializirano preverjanje parov formatov izvornih in končnih datotek dokumentov. Le tako jo je mogoče izvesti kontrolirano in zanesljivo. [2]

Za dokumente, ki jih je treba ohranjati dlje časa, je potrebnih več migracijskih ciklov. Pri tem se pojavi vprašanje, kdaj začeti z migracijskimi cikli in kako pogosto jih izvajati. Migracija dokumentov je potrebna vsakič, ko to zahteva sprememba tehnologije ne glede na to, ali je v tistem trenutku znano, ali bodo določeni dokumenti še uporabni ali ne. Poleg tega je treba zagotoviti tudi ustrezno migracijo metapodatkov, ki se navezujejo na dokumente. Vse to otežuje pripravo stroškovnega modela brez katerega migracija ni izvedljiva. Migracija je zahtevna tudi glede zagotavljanja kadrov, razen kadar jo je mogoče izvajati popolnoma avtomatizirano. [2]

Migracija digitalnih dokumentov je večkrat kritizirana zaradi neznanih učinkov, ki jih lahko ima proces na verodostojnost in integriteto dokumentov. Pri vsakem kopiranju lahko pride do napake zaradi hroščev v programu, nepravilnega ravnanja z datotekami ali zaradi mehanske okvare podatkovnih nosilcev. Kljub vsemu izsledki raziskav kažejo, da je migracija s pravilno oceno tveganja primerna metoda za dolgoročno hranjenje večine digitalnih dokumentov. [2]

2.2 Raziskave in projekti na temo migracije digitalnih dokumentov

V zadnjih desetih letih je bilo na področju migracij digitalnih dokumentov izvedenih veliko raziskav, kar je omogočilo identifikacijo glavnih problemov ter tveganj. Prihodnji razvoj je tako naravnani v iskanje najprimernejših rešitev.

Eden pomembnejših projektov je bil CAMiLEON. Začeli sta ga skupaj univerzi Michigan (ZDA) in Leeds (VB) leta 1999. Ena izmed nalog je bila priprava smernic uporabe emulacije in migracije za ohranjanje digitalnih dokumentov. [9]

Projekt InterPARES (The International Research on Permanent Authentic Records in Electronic Systems) je osnovan na univerzi British Columbia v

Kanadi. Med drugim se ukvarja s pripravo smernic kot podlago za standarde, strategije, politiko, ki zagotavljajo dolgoročnost digitalnih dokumentov in možnost, da lahko uporabniki zaupajo v njihovo verodostojnost. [10] Kljub temu da se njihovi dokumenti ne nanašajo neposredno na migracijo, je predlagane smernice in modele mogoče neposredno uporabiti za pripravo ustrezne strategije migracije digitalnih dokumentov.

Nizozemski narodni arhivi (National Archives of the Netherlands) so leta 2000 vzpostavili projektno skupino Digital Preservation Testbed. Ukvarja se z raziskovanjem različnih metod dolgoročnega ohranjanja digitalnih dokumentov. [14] V letih 2000 do 2003 so izvajali projekte testiranja migracij besedilnih dokumentov, baz podatkov in preglednic v različne formate. Rezultati so pokazali, da migracija s povratno združljivostjo ni primerna za dolgoročno ohranjanje digitalnih dokumentov (50 let in več). Migracija besedilnih dokumentov v XML se je izkazala kot zelo dobra, kadar so dokumenti tvorjeni na podlagi vnaprej pripravljenih predlog. Pri migraciji besedilnih dokumentov v PDF so ugotovili, da je to najboljša strategija, če pri tvorbi dokumentov niso bile uporabljene vnaprej pripravljene predloge. Migracija v XML je bila spoznana kot najboljša strategija za migracijo baz podatkov in preglednic. [11]

Narodni arhivi Avstralije (National Archives of Australia) se od leta 1994 aktivno ukvarjajo z raziskovanjem in razvojem rešitev za dolgoročno hranjenje digitalnih dokumentov. [12] Njihove raziskave na področju migracije so osredinjeni na besedilne datoteke in HTML-datoteke, ki so del projekta, ki se ukvarja s spletnimi dokumenti PANDORA (Preserving and Accessing Networked Documentary Resources of Australia). [13]

Našteti projekti in organizacije so eni pomembnejših in so pogosto omenjeni v literaturi, ki obravnava migracijo digitalnih dokumentov. V nadaljevanju bodo predstavljene vrste migracij, ki se pogosto uporabljajo.

3 VRSTE MIGRACIJ

Pojavlja se več vrst migracije. V literaturi, ki je trenutno na voljo, se najpogosteje pojavljajo:

- migracija s pomočjo povratne združljivosti (backward compatibility [3]),
- migracija s pomočjo medobratovalnosti (interoperability [3]) in

- migracija v standardne formate (migration to standard formats) [1].

Odločitev, katera oblika je najprimernejša, je odvisna od dolžine obdobja hranjenja dokumentov. Upoštevati je treba vse zahteve za ohranitev verodostojnosti in integritete, kar je ključna zahteva pri upravljanju z digitalnimi dokumenti [4].

3.1 Migracija z uporabo povratne združljivosti

Migracija z uporabo povratne združljivosti omogoča preslikavo in pravilno reprodukcijo dokumenta, ki je bil tvoren v starejši različici aplikacije, z uporabo njene kasnejše različice. Migracija je najlažja in najcenejša, kadar razvijalec originalne programske opreme zagotavlja pretvornike v okviru novejših različic programov, npr. Word 2007 lahko bere datoteke tvorjene v Word 95.

Bistvena nevarnost te vrste migracije je, da programi običajno omogočajo migracijo le za dve ali tri predhodne generacije in ne za vse predhodne različice. Poleg tega razvijalci niso dolžni zagotavljati povratne združljivosti in je to odvisno od njihove poslovne strategije. [2]

Običajno je treba izvesti migracijo na novo različico vsakih nekaj let, odvisno od tega, kako pogosto izhajajo nove različice programske opreme. Povratna združljivost je primerna strategija za kratkoročno hranjenje. Primeri takšnega načina migracije sicer ne kažejo večjih težav, če med različicami urejevalnikov ni preteklo preveč časa. S tem verodostojnost ter integriteta večinoma nista ogroženi. [1]

3.2 Migracija z uporabo medobratovalnosti

Migracija z uporabo medobratovalnosti pomeni, da so datoteke (dokumenti) prenosljive z ene platforme na drugo in jih je še vedno mogoče prikazati enako. Datoteko je mogoče brati in urejati z uporabo različnih verzij iste aplikacije, ki deluje na različnih operacijskih sistemih. Datoteke, pri katerih uporabljamo za opis besedila samo prvih 127 znakov v tabeli ASCII, so zelo prenosljive. Ker pri njih z migracijo skorajda ni težav, so podlaga za razne označevalne jezike (npr. HTML, RTF). Drugi primer medobratovalnosti je prenos med podobnimi programskimi aplikacijami. Tretja oblika medobratovalnosti temelji na uporabi programskega vmesnika za pretvorbo. To pomeni, da je dokument, tvoren v enem urejevalniku in shranjen v izmenljivem formatu, npr. RTF (Rich Text Format [5]), berljiv v drugem urejevalniku besedil.

Medobratovalnost zaradi številnih pomanjkljivosti ni zanesljiv način za hranjenje digitalnih dokumentov. Uporabna je lahko kot začasna rešitev do uvedbe dolgoročnejšega načina [1].

3.3 Migracija s pretvorbo v standardni format

Migracija s pretvorbo v standardni format je migracija lastniškega formata (običajno zaprt in v binarni obliki) v format, ki temelji na javnem standardu in je odprt. Digitalni dokumenti niso več odvisni od originalne programske opreme, s katero so bili tvorjeni. Posledično niso več izpostavljeni nevarnosti zastaranja programske opreme.

Formati so lahko formalni, odprte narave ali akreditirani pri organizacijah za standardizacijo (ISO, NEN, W3C) [2]. Eden takšnih standardnih formatov je XML (Extensible Markup Language [6]). XML je bil razvit z namenom povečanja medobratovalnosti jezikov SGML in HTML [6]. Določeni formati temeljijo na standardih, ki niso uradno priznani, so pa tako razširjeni, da jih uporablja veliko število uporabnikov [2]. Takšna standardna formata sta na primer PDF (Portable Document Format), preden je postal odprti standard, ter format RTF (Rich Text Format [5]).

Format PDF je leta 2008 uradno postal odprti standard, poznan kot ISO 32000. Datoteke v tem formatu so berljive in jih je mogoče natisniti na platformah, kot so Mac OS, Microsoft Windows, UNIX, in na mnogih mobilnih platformah. Datoteke PDF ohranjajo izgled kot izvorniki in ohranijo besedilo, risbe, video, 3D-slike, načrte, fotografije ipd. [7].

Format RTF je metoda kodiranja besedila in grafike za enostaven prenos med aplikacijami. Trenutno so na voljo prevajalniki, ki omogočajo prenos dokumentov med različnimi aplikacijami, ki delujejo v okoljih DOS, Microsoft Windows, OS/2, Macintosh in Power Macintosh [5].

Tudi pri migraciji v standardne formate se pojavljajo težave. Na izbiro je veliko standardov in težko je določiti, kateri je najprimernejši. Poleg tega se standardi spreminjajo in prihajajo nove različice. To pomeni, da je treba tudi v tem primeru migracijo izvajati večkrat. [2]

Kljub vsemu je pretvorba v standardne formate primeren pristop za hranjenje dokumentov. Takšen način omogoča tako povratno združljivost kot medobratovalnost [1].

4 TESTIRANJE MIGRACIJE STAREJŠIH WORDOVIH DOKUMENTOV

Glede na izkušnje uporabnikov je znano, da se določene lastnosti pri migraciji starejših dokumentov vedno ne ohranijo popolnoma. Namen testiranj je bil ugotoviti, kakšna migracija omogoča čim bolj popolno ohranitev verodostojnosti in integritete obravnavanih dokumentov.

Izvedena sta bila dva testa. V prvem smo opazovali uspešnost različnih migracij dokumentov, ki so bili tvorjeni z urejevalnikom Word 5.5, ki je deloval v okolju DOS. V drugem testu smo opazovali uspešnost različnih migracij dokumenta, ki je bil tvorjen z urejevalnikom Word 2.0, ki deluje v okolju Windows.

Glavne lastnosti testnih dokumentov

V okviru testiranj smo opazovali, kako uspešna je migracija enajstih najpomembnejših lastnosti in elementov testnih dokumentov:

- slike, kopirane iz grafičnega programa prek odložišča,
- slike, vstavljene kot hiperpovezave,
- polja s samodejnim posodabljanjem (številke strani, številke slik),
- tabele,
- standardne oblike pisav,
- šumniki,
- posebni znaki,
- enačbe,
- sprotne opombe,
- slogi besedil,
- povzetek dokumenta.

Testni dokumenti:

- dokumenti, tvorjeni z Word 5.5 (datoteke: testni_dokument_1.doc, testni_dokument_2.doc in testni_dokument_3.doc),
- dokument, tvorjen z Word 2.0 (datoteka: testni_dokument_4.doc).

Uporabljena programska in strojna oprema:

- Word 5.5 (angleška različica za DOS),
- Word 2.0 (angleška različica za Windows),
- Word 95 (angleška različica 7),
- Word 2007 SP2 (slovenska različica 12.0.6535.5002),
- Adobe reader 9, različica 9.3.1,
- program za pretvorbo dokumentov v format pdf: FreePDF XP 3.26,

- operacijski sistem: MS Windows Xp; Home Edition; različica 2002; servisni paket 3; slovenska različica,
- procesor: Intel Core™ 2; 1,66GHz, 1GB RAM.

Merila za ovrednotenje uspešnosti rezultatov migracije:

- **popolnoma uspešna migracija:** vsi elementi in lastnosti v ciljnem formatu so popolnoma enaki kot v izvornem;
- **delno uspešna migracija:** določen element ali lastnost v ciljnem formatu nista popolnoma enaka kot v izvornem (npr. vrsta pisave), vendar s tem ni ogrožena njegova integriteta;
- **neuspešna migracija:** določen element ali lastnost v ciljnem formatu se tako razlikuje od izvornega, da je ogrožena njegova integriteta.

Programi za preverjanje uspešnosti migracij

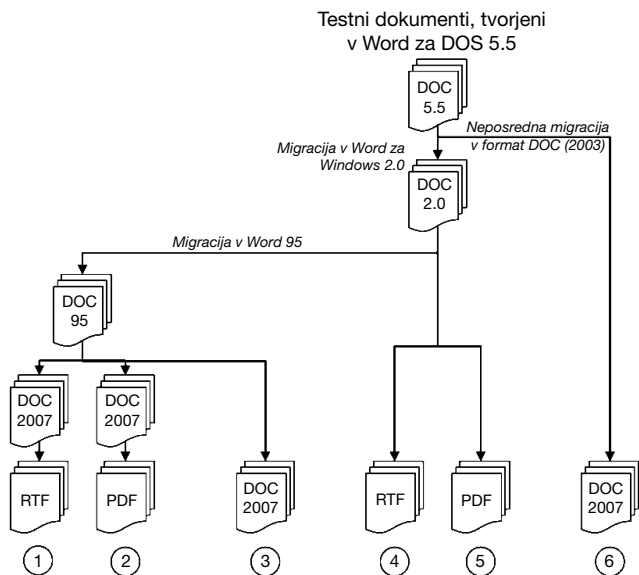
Uspešnost migracije smo spremljali na podlagi ohranjenih lastnosti v ciljnem formatu glede na izvorni dokument. Dokumente v novem formatu smo odpirali s programi:

- format DOC; Word 2007,
- format RTF; Word 2007,
- format PDF; Adobe reader 9.

4.1 Testiranje migracije dokumentov, tvorjenih z urejevalnikom Word 5.5

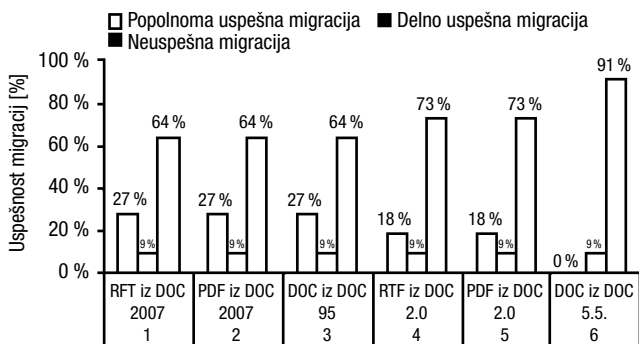
V prvem testu smo opazovali uspešnost neposredne in postopne migracije dokumentov, tvorjenih v urejevalniku Word 5.5. Dokumente smo v obeh primerih pretvorili v formate DOC (različica 2007), PDF (različica 1.3) in RTF. Glede na to, da urejevalniki v osnovi ne omogočajo pretvorbe v format PDF, je bil uporabljen brezplačni program FreePDF XP, ki deluje kot navidezni tiskalnik. Različica formata PDF po migraciji je 1.3. Neposredna migracija v formata RTF in PDF neposredno iz Word 5.5 z razpoložljivo programsko opremo ni bila mogoča, zato je bila izvedena postopna migracija najprej v format Word 2.0. Potek prvega testa prikazuje slika 1.

Urejevalnik Word 2007 ne omogoča neposrednega odpiranja datotek, ki so bile tvorjene v urejevalniku Word 2.0 brez poseganja v register Windows. Glede na to, da za povprečne uporabnike poseg v register ni priporočljiv, je bilo treba te dokumente predhodno pretvoriti še v format DOC MS Word 95.



Slika 1: **Potek testiranj migracije dokumentov, tvorjenih v Word 5.5**

Rezultati testiranja so pokazali, da je bila migracija slik v prvem testu neuspešna. Ohranil se je le del slik, kar bistveno vpliva na integriteto dokumentov. Polja s samodejnim posodabljanjem so se ohranila le pri postopni migraciji najprej v Word 2.0 in nato še v Word 95. V ostalih primerih je bila migracija neuspešna. V vseh primerih je bila neuspešna migracija tabel, šumnikov, posebnih znakov in povzetka dokumenta. Pri tem je treba upoštevati, da je bila tabela v testnem dokumentu sestavljena iz posebnih znakov in razmikov med vrednostmi s pomočjo tabulatorjev. Edina lastnost, ki je bila delno uspešna v vseh primerih, je ohranitev standardne oblike pisav brez upoštevanja šumnikov. V vseh primerih razen pri neposredni migraciji v format DOC 2007 je bila popolnoma uspešna migracija slogov besedil in sprotnih opomb.



Slika 2: **Uspešnost migracij dokumentov, tvorjenih v Word 5.5, v nekatere trenutno aktualne formate**

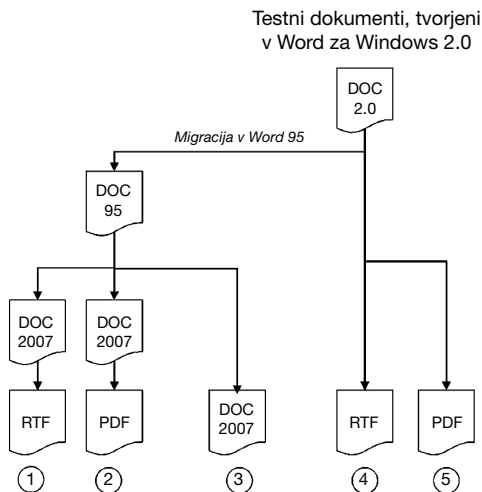
Kot prikazuje slika 2, delež neuspešnih migracij prvega testa v nobenem primeru ni manjši od 64 odstotkov.

Največ neuspešnih pretvorb posameznih lastnosti je bilo v primeru neposredne migracije izvornih dokumentov v Word 2007, in sicer 91 odstotkov. Popolnoma uspešna migracija je bila le pri 27 odstotkih obravnavanih lastnosti pri postopni migraciji originalnih dokumentov v treh korakih, tj. najprej v Word 2.0, nato v Word 95 in nazadnje v ciljni format. Pri postopni migraciji izvornih dokumentov v dveh korakih, tj. najprej v Word 2.0 in nato v končni format, je bil delež lastnosti, ki so prestali migracijo popolnoma uspešno, 18 odstotkov.

Zgornji rezultati kažejo, da v primerih, ko preteče daljše časovno obdobje med posameznimi različicami programske opreme, lahko pričakujemo večji delež neuspešnih migracij.

4.2 Testiranje migracije dokumentov, tvorjenih z urejevalnikom Word 2.0

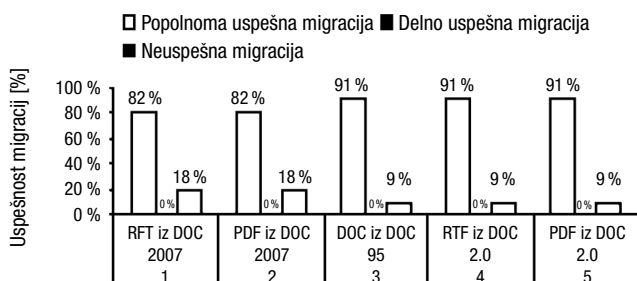
V drugem testu smo opazovali uspešnost neposredne migracije dokumentov, ki so bili tvorjeni v urejevalniku Word 2.0. Tudi v tem primeru so bili dokumenti pretvorjeni v iste formate kot v prvem testu. Potek drugega testa prikazuje slika 3.



Slika 3: **Potek testiranj migracije dokumentov, tvorjenih v Word 2.0**

Migracija v drugem testu je bila neuspešna le pri pretvorbi šumnikov, slogih besedila in povzetku dokumenta. V vseh drugih primerih je bila popolnoma uspešna.

Šumniki se niso ohranili v primeru migracije iz formata DOC Word 2.0 v Word 95 in nato v RTF, PDF ter format DOC 2007. Slogi besedil se niso ohranili pri migraciji testnega dokumenta iz formata DOC 2.0 v format RTF. Naslovi posameznih poglavij na isti strani so se pomaknili na začetek strani, kar je povzročilo prekrivanje besedila. S tem je ogrožena integriteta dokumenta. Migracija povzetkov dokumenta je bila neuspešna v primerih migracije v format PDF, saj se vpisane lastnosti ne ohranijo. S tem je ogrožena verodostojnost dokumenta.



Slika 4: Uspešnost migracij dokumenta, tvorjenega v Word 2.0, v nekatere trenutno aktualne formate

Rezultate uspešnosti migracije v različne formate v drugem testu prikazuje slika 4. Najnižji delež popolnoma uspešnih migracij obravnavanih lastnosti v tem testu je bil 82 odstotkov pri migraciji iz formata Word 2.0 v Word 95, nato v Word 2007 in nato v format RTF ali PDF. V drugih primerih je delež popolnoma uspešnih migracij 91 odstotkov.

Največjo težavo pri migraciji so predstavljali šumniki, in sicer v primerih, da je postopna migracija vključevala tudi pretvorbo v format Word 95. V teh treh primerih je bila migracija šumnikov neuspešna.

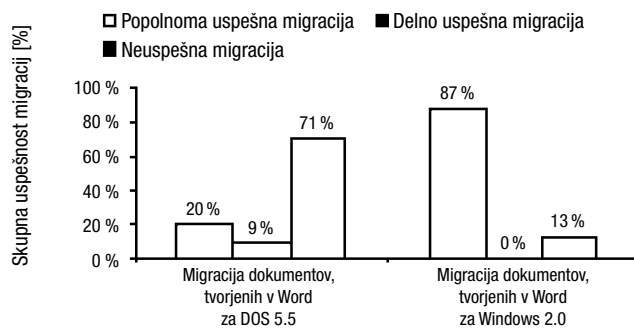
Rezultati drugega testa kažejo, da je neposredna migracija dokumenta Word 2.0 v formata RTF in PDF bolj uspešna kot v primeru postopne migracije.

Pri migraciji v format Word 2007 je bila potrebna postopna migracija v format Word 95, saj urejevalnik Word 2007 ne omogoča neposredne migracije dokumentov iz formata Word 2.0.

4.3 Primerjava uspešnosti migracij digitalnih dokumentov, tvorjenih v Word 5.5 in v Word 2.0

Rezultati uspešnosti migracij obeh testov, ki jih prikazuje slika 5, kažejo, da je bila migracija obravnavanih lastnosti bistveno bolj uspešna v primeru, ko je bil dokument tvorjen v urejevalniku Word 2.0, kot

v primeru dokumentov, tvorjenih v Word 5.5. Pri migraciji dokumentov, tvorjenih v Word 5.5, se je popolnoma ohranilo le povprečno 20 odstotkov obravnavanih lastnosti. Pri migraciji dokumenta, tvorjenega v Word 2.0, se je popolnoma ohranilo povprečno 87 odstotkov obravnavanih lastnosti.



Slika 5: Primerjava uspešnosti migracij testnih dokumentov, tvorjenih v Word 5.5, in testnega dokumenta, tvorjenega v Word 2.0

5 SKLEP

Rezultati prvega testa kažejo, da je za dokumente, tvorjene v formatu Word 5.5, najbolj učinkovita postopna migracija v format Word 2007.

V drugem testu je bil obravnavani dokument tvorjen v urejevalniku Word 2.0, ki deluje v operacijskem sistemu Windows. Najbolj uspešna je bila neposredna migracija v standardna formata RTF in PDF. Enak rezultat je pokazala posredna migracija v format DOC 2007. Najmanj učinkovita je bila posredna migracija v formata RTF in PDF.

Rezultati testov kažejo, da je uspešnost migracije dokumentov, tvorjenih v Word 5.5, slabša kot uspešnost migracije dokumentov, tvorjenih v Word 2.0.

6 LITERATURA IN VIRI

- [1] ICTU: From digital volatility to digital permanence, Preserving text documents (version 1.0); ICTU, Hag, december 2003; <http://www.digitaleduurzaamheid.nl> (dec. 2009).
- [2] ICTU: Migration: Context and Current Status, ICTU, Hag, december 2001; <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/Migration.pdf> (dec. 2009).
- [3] Slovensko društvo Informatika, Slovar informatike (iskanje ustreznih prevodov); <http://www.islovar.org> (dec. 2009).
- [4] William E. Neale, IBM FileNet Records Manager: Assuring Records Integrity, IBM corporation, 2008; <ftp://ftp.software.ibm.com/software/data/ECM/compliance/fn-records-integrity-wp.pdf> (dec. 2009).
- [5] Microsoft Corporation, Rich Text Format (RTF) Specification, version 1.6, maj 1999; <http://msdn.microsoft.com/en-us/library/aa140277%28office.10%29.aspx> (dec. 2009).

- [6] World Wide Web Consortium, Extensible Markup Language (XML) Version 1.0, december 1997; www.w3.org/TR/PR-xml-971208.rtf (dec. 2009).
- [7] Adobe Systems, <http://www.adobe.com/products/acrobat/adobepdf.html> (dec. 2009).
- [8] CPA/RLG, Preserving Digital Information: Report of the Task Force on Archiving of Digital Information, maj 1996; <http://www.clir.org/pubs/reports/pub63watersgarrett.pdf> (sept. 2010).
- [9] <http://www2.si.umich.edu/CAMILEON/reports/mor/index.html> (sept. 2010).
- [10] <http://www.interpares.org/> (sept. 2010).
- [11] Caroline van Wijk; Migration Research; Koninklijke Bibliotheek; april 2006; http://www.kb.nl/hrd/dd/dd_projecten/Starting_Point_Migration_Research.pdf (sept. 2010).
- [12] NLA, Digital Preservation: The Australian Experience; oktober 2000; <http://www.nla.gov.au/nla/staffpaper/dw001004.html> (sept. 2010).
- [13] <http://pandora.nla.gov.au/> (sept. 2010).
- [14] <http://en.archief.nl/> (sept. 2010).

■

Tomaž Stenšak je diplomiral leta 2002 na mariborski Fakulteti za elektrotehniko, računalništvo in informatiko na univerzitetnem študiju Gospodarsko inženirstvo, smer Močnostna elektrotehnika. V okviru magistrskega študija raziskuje področje vzdrževanja tehniške dokumentacije v elektroenergetskih podjetjih.

■

Tomaž Dogša je izredni profesor na Fakulteti za elektrotehniko, računalništvo in informatiko Univerze v Mariboru, kjer predava na dodiplomski in podiplomski stopnji in vodi Center za verifikacijo in validacijo sistemov. Na raziskovalnem področju se ukvarja predvsem s tehnologijo V&V oz. s testirnimi orodji.