

Nikola Popović, Srboljub Ranković
(Ministarstvo spoljnih poslova SCG)
Julijana Mirčevski

ANALIZA PROBLEMA DIGITALIZACIJE DOKUMENATA U DRŽAVNOJ UPRAVI

Sažetak: U tekstu se iznose lični stavovi autora o problemu digitalizacije dokumenata u državnoj upravi. Analizira se proces organizacije digitalizacije. Utvrđuju se mere koje je potrebno primeniti od procesa projektovanja dokumentacionih sistema do organizacije rada na radnom mestu. Analiziraju se problemi koji su nerešeni u procesu rada sa dokumentima uopšte odnosno u procesu pretraživanja i predlažu rešenja u skladu sa svetskim tendencijama u ovoj oblasti.

Ključne reči: državna uprava, kancelarijsko poslovanje, digitalizacija, informaciono društvo

1. Uvod

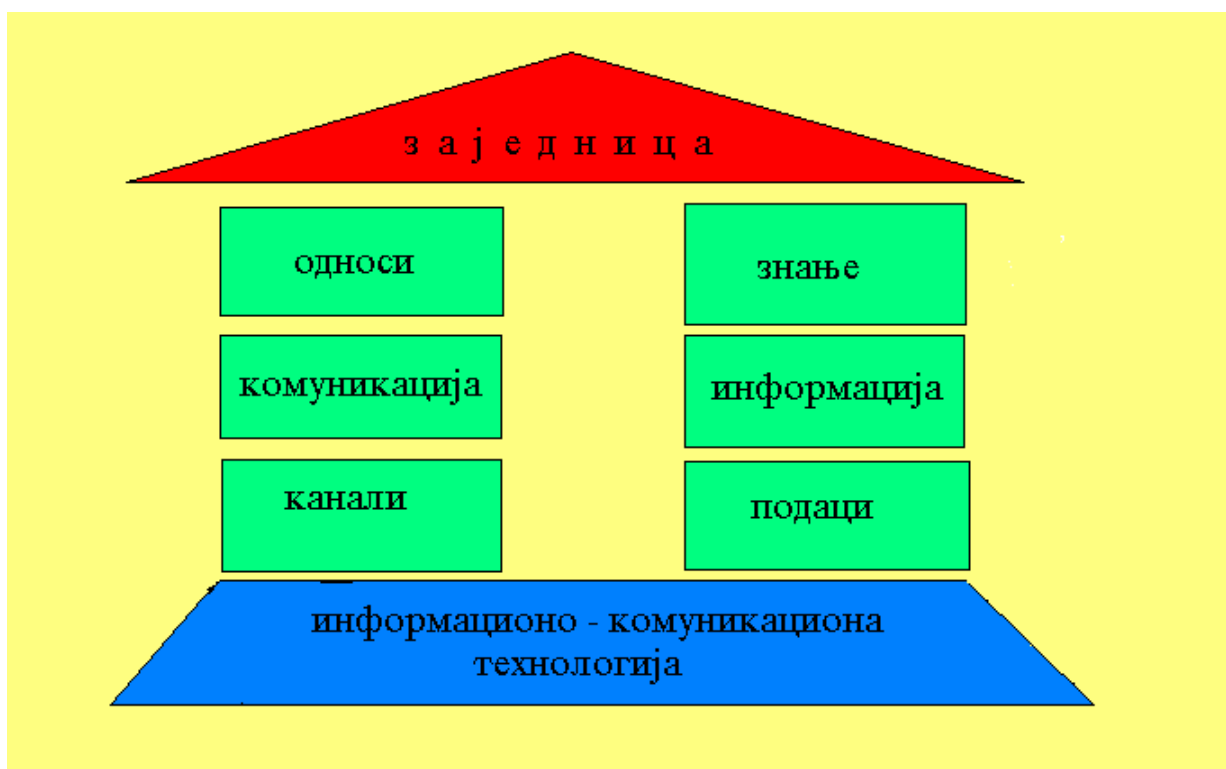
Posmatratrati informaciju kao nacionalno blago odraz je ozbiljnosti shvatanja njenog značaja za sveukupan razvoj društva u vremenu koje je bitno obeleženo kao informatičko doba. Koncept Informacionog društva uvodi racionalniji pristup - Informacija je nacionalni resurs. Informacija daje moć upravljanja odnosno orijentacije u prostoru stvarnih i mogućih događaja. Informacija može i treba da se posmatra i kao roba, odnosno ona ima svoju cenu, kako na tržištu tako i u procesu proizvodnje, što nikada ne smemo da izgubimo iz vida. Budućnost ne može da se projektuje bez fondova organizovanih informacija. U našoj sredini dešavanja u domenu dokumentacionih sistema državne uprave, koji čine okosnicu za procese generisanja informacija [1], objektivno su još uvek na početku. Iako je generisanje elektronskih dokumenata već postao redovan radni proces, još uvek se takvi dokumenti koriste samo za izradu štampanih, papirnih kopija, koje i dalje predstavljaju jedini originalni, zakonski priznati dokument. I to bi moglo i da se prihvati kao norma sadašnjeg trenutka da elektronske verzije dokumenata nakon toga gube značaj tj najčešće se brišu „da ne bi zauzimali mesto“. Posebno malo je učinjeno u procesu prevođenja (za redovan posao) značajnih papirnih dokumenata u elektronsku formu. Tehnologija digitalizacije je danas sa tehničkog ali i finansijskog aspekta u potpunosti prihvatljiva. Nažalost, još uvek se ne shvataju prave implikacije ovakvog pravca razvoja događaja. Stvaranje osnova „informacionog društva“ zahteva detaljno preispitivanje kompletnog poslovanja državne uprave na svim nivoima i usklađivanje sa kretanjima u razvijenom svetu. U pitanju je ne tako brz proces, pošto su mnoge stvari još uvek nedorečene, ali kašnjenje više od 5 godina za svetom, donosi višestruko kašnjenje u svim drugim oblastima funcionisanja društva pa i državnih institucija .

Za nas je od posebnog značaja pažljivo izučavanje koncepta održivih sistema, da se ne bi ponovile greške iz 80-tih i 90-tih godina kada su se projektovali mamutski tzv.

Jedinstveni sistemi. Velika sredstva su ulagana u imitacije sistema bogatih zemalja i zato takvi sistemi nikada nisu ni proradili.

2. Struktura društvene zajednice – ICT aspekt

Šema strukture društvene zajednice [2], jasno ukazuje na značaj trijade podaci – informacija – znanje (tj. desnog „stuba zajednice“) koja se kod nas suviše dugo zanemaruje u odnosu na drugi, levi „stub zajednice“. Uspešno funkcionisanje društvene zajednice uslovljeno je jedinstvenim funkcionisanjem prikazanih modula. Njihova međusobna zavisnost i povezanost obezbeđuje kontinuitet upravljanja zajednicom kao zajedničkim



krovom društvene zgrade. Digitalizacija je upravo jedan od značajnih tehnoloških postupaka koji treba da snažno podrži realizaciju navedene trijade. Naime, već decenijama su procesi akumulacije, organizovanja i efikasnog korišćenja podataka najslabija karika državne uprave.

3. Vremenska dimenzija podataka, tehnološka nehomogenost i uticaj ovih faktora na projektovanje

Podatak sam po sebi zastareva, ali vremenske serije (istorija) podataka nose informaciju o procesima. Prelaskom na nove tehnologije često se gube izvorni podaci (ostavljaju se sintetičke informacije, izveštaji i sl., bez istorijskih, referentnih podataka). Poznavanje tih vidova problema kroz lično iskustvo podstaklo je autore na izradu ovog teksta. Autori teksta imaju iskustva sa najmanje tri tehnološke generacije elektronskih mašina za

obradu podataka i prosečno tri decenije iskustva u obradi podataka na mašinama. Sada znamo da su mnogi podaci mogli, i trebalo je, da budu sačuvani.

Uzroci problema su: tehnološka nehomogenost, nekonzistentna distribucija poznavanja različitih računarskih tehnologija, komplikovan i skup proces konverzije, ne postoje sveobuhvatna pravila za postupak sa podacima.

Podaci se zapisuju na različite nosioce informacija: kamen, glinene pločice, papir, CD, magnetne medijume itd. I pored mnoštva savremenih memorijskih tehnologija ogromnog kapaciteta, trajnost i dostupnost nosioca informacije čini suštinu čuvanja informacija i posebno procesa digitalizacije. Pouzdanije informacije o trajnosti medijuma postoje samo za magnetne medijume – trake i diskove. Problem promene nosioca informacije odnosno tehnologije zapisivanja i čitanja informacije sve je veći, pošto se informaciono-komunikaciona tehnologija (ICT) možda i suviše brzo menja.

Problem tehnološke nehomogenosti javlja se usled različitih: generacija HW opreme, operativnih sistema, DBM dokumentalističkih baza podataka i sistema za obradu teksta.

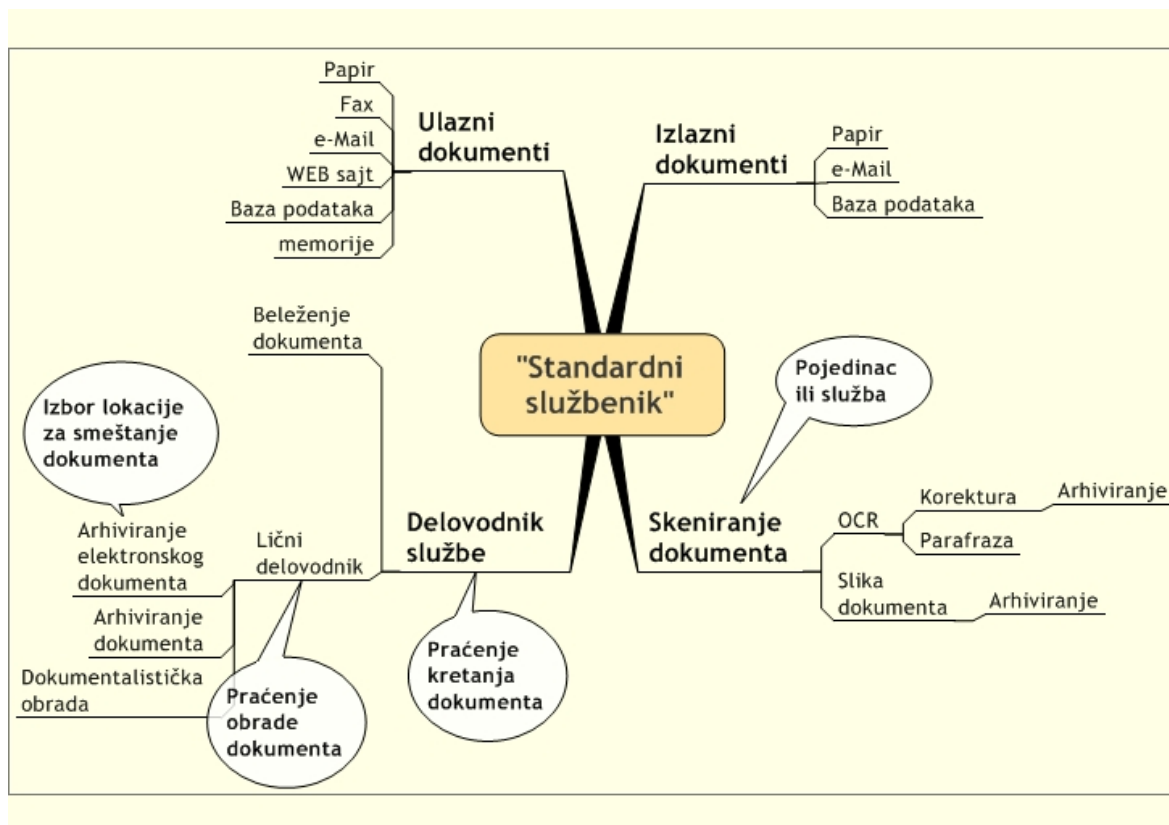
Ovde su evidentirani problemi u domenu nečitljivosti zapisa na računaru (operativnost, neadekvatan format i sl.), nemogućnosti konverzije podataka na savremeniji softverski sistem, nemogućnost aktiviranja SW sistema na novim HW platformama ili obrnuto, otežan reinženjering sistema zbog nedostatka dokumentacije ili manjkavo projektovanih baza podataka.

3.1. Uticaj na projektovanje. Realizacija savremenih dokumentalističkih sistema se uglavnom zasniva na gotovim konfekcioniranim rešenjima. Tržište nudi rešenja sa kompromisima i/ili uslovima koji su često teško prihvatljivi (specifični formati, posebni programski jezici itd.). Prethodno navedeni problemi nehomogenosti sugerišu korišćenje jednostavnijih rešenja. Ideologija tržišne privrede uslovljava da se tehnologija suviše često menja ili firme isporučioici opreme nestaju sa tržišta. Dobitak od trenutno tehnološki optimalnog rešenja najčešće se gubi pri konverziji na novi “savršeniji” sistem. Jednostavnije i “čitljivije” strukture podataka, po cenu povećane redundantnosti, pa i manje efikasnosti, imaju prednost [1]. Već u procesu projektovanja novog sistema veoma je korisno predvideti “izlazne tačke” za konverziju podataka na budući sistem. Takođe je važno pripremiti dokumentaciju za budući prelazak na novu generaciju sistema ili potpuno drugi tehnološki sistem. Dokumentacija mora da sadrži veoma detaljne meta-podatke.

Autori predlažu obaveznu procenu dužine trajanja i obavezni opis uslova migracije sa tekuće na sledeću generaciju ICT opreme. Ovakva analiza trebalo bi neizostavno da sadrži i procenu ekonomske vrednosti procesa migracije, a da se realizuju samo ekonomski isplative migracije informatičkih sistema.

4. Digitalizacija – značaj i problemi primene u državnoj upravi

Digitalizacija je proces konvertovanja postojećeg sadržaja u formi teksta ili slike u mašinski čitljivu formu. Digitalizacija omogućava manipulaciju podacima u smislu: njihove izmene, ponovnog kopiranja, vizuelizacije i podvrgavanja procesima analize i sinteze. Problemi digitalizacije su: stepen vernosti reprodukcije (modela), u slučaju dokumentacije aspekt pravne valjanosti kopije odnosno autentifikacije, korišćenje i čuvanje digitalizovanih dokumenata.



Digitalizacija naglašava pitanje: šta je originalni dokument. Pitanje je doslovno biblijsko, skoro dve hiljade godina se proverava vreme nastanka i autentičnost jevanđelja, primenom krajnje složenih lingvističkih metoda. Danas, pitanje postojanja nesumnjivog originala papirnih dokumenata je manje izraženo u odnosu na dokumenta u video zapisu kao i dokumenta čije je originalno poreklo Web sajt gde ovaj problem dobija izuzetno složenu pozadinu [3].

Šta je karakteristično za Ministarstvo:

Ministarstvo je objekat odnosno proces koji se odvija u realnom vremenu (tokom 24 sata) (ovo znači da ne može da se doslovno posmatra preko danas modernog koncepta – projekta sa definisanim početkom i krajem).

Veliki obim dokumentacije (najmanje 3.000.000 listova papira godišnje u lokalnu).

Izuzetna nehomogenost dokumentacije po sadržaju – od pitanja rata i mira pa do rešavanja penzija i ostavina tj. od nacionalnog do ličnog nivoa.

Nehomogenost dokumentacije po formi – papir, karte, e-mail, fax, video/audio zapisi, pdf, word, CD, DVD, diskovi itd.

Ideal: sve digitalizovano i pretraživo. To je ustvari san o kancelariji bez papira „paperless office“ [3] [4]. – stvarnost je drukčija. Nažalost, ova paradigma može suviše često da motiviše snage ljudi koji donose odluke da prihvate rešenja za računarska pretraživanja izvan potreba i izvan realnosti.

Stvarnost: potrebno je doneti odluke:

Šta da se digitalizuje (ko vrši trijažu dokumenata i po kojim kriterijumima?)

Šta da se pretvori u tekst primenom OCR-a (neki dokumenti treba da idu u full-text baze podataka [5], a neki samo da imaju referencu radi uvida u original)

Kako da se obavlja digitalizacija (individualno ili da se organizuje posebna služba?)

Šta da se radi sa digitalizovanim dokumentima (lokalne arhive, centralno skladište i sl.)

4.1 Koncept “Standardni službenik”. Upotreba računara u državnoj upravi obično se modelira prema postojećim procesima rada što najčešće ne daje očekivane rezultate [6] [7]. Digitalizacija menja strukturu modela mnogo više nego što se to uobičajeno misli i očekuje – pojavljuju se složeni poslovi promene forme dokumenta. Iskustvo, odnosno zahtevi za digitalizacijom dokumenata ukazuju da je bilo potrebno razraditi koncept “Standardnog službenika”. “Standardni službenik” prima dokumenta, zavodi ih u delovodnik [8], analizira predmet i određuje dalji postupak obrade: po upoznavanju sadržaja odlaže dokument u arhivu, ili ga prosleđuje kompetentnoj osobi, ili generiše nove dokumente analizirajući druge već postojeće dokumente, pravi potrebne kopije dokumenata, prosleđuje dalje nove dokumente, ili čeka instrukcije.

Organizacione jedinice državne uprave mogu da se posmatraju kao proizvodne celine koje prerađuju informacije (“fabrike papira”). Ideja za koncept “Standardnog službenika” ima neposrednu vezu sa određenim konceptima iz oblasti Virtual Enterprise i Flexible Manufacturing Systems, pošto postoji veoma velika sličnost sa materijalnom proizvodnjom, naime “standardni službenik” mora da obrađuje različite predmete koristeći istu opremu [9,10].

4.2 Organizacija procesa digitalizacije. Postoji izuzetno značajna proliferacija papirnih (ali i elektronskih) kopija što sve više opterećuje redovan rad, zahteva veliki broj uređaja i značajne troškove. Na kraju životnog veka dokumenti idu na trijažu pre odlaska u arhivu što sada predstavlja dodatni problem (najčešće se u projektima „aproximira“) i dodatni značajan trošak Digitalizacija je sada već aktuelan način smanjenja obima redundantnih kopija dokumenata. Izlazni dokumenti (radni) sada su već u ogromnoj većini u digitalnoj formi (osim potpisanog originala).

Praktična pitanja digitalizacije se mogu razdvojiti na sledeća potpitanja:

Da li je potrebno sve digitalizovati (trijaža po zadatim kriterijumima)

Šta je potrebno čuvati kao digitalni dokument – ulazni i / ili izlazni

Kako i ko će obaviti dokumentalističku obradu digitalizovanih dokumenata

Gde se čuvaju digitalizovani dokumenti i kako se povezuju sa papirnim kopijama

Kako se koriste digitalizovani dokumenti (prava kopiranja, uvida i sl.)

Registracija dokumenata

Pravni aspekt

Problem autentifikacije elektronske kopije

Ko potvrđuje valjanost kopije

Kako se verifikuje kopija (elektronski potpis)

Gde se čuvaju elektronski dokumenti (centralno skladište ili lokalne elektronske arhive). Procedure čuvanja i autentifikacije backup-a?

Ko vrši digitalizaciju dokumenta:

Službenik na radnom mestu (pitanje kvaliteta kopije, ispravnosti kopije, obuke za kopiranje „loših“ dokumenata itd.)

Služba specijalizovana za digitalizaciju (pripremljena za efikasnu digitalizaciju i obradu starih i/ili oštećenih dokumenata)

Pošto je u pitanju velika masa dokumenata, postavlja se dodatni problem organizacije autentifikacije. Kada se menja tehnologija elektronske arhive, ko i kako će da izvrši autentifikaciju novih kopija dokumenata (baza podataka i svih drugih formi elektronskih dokumenata)? Koja procedura treba da se primeni prilikom promene tehnologije na kojoj se zasniva elektronski potpis? Kako se ta promena odražava na autentifikaciju postojećih (elektronskih) arhiva?

5. Pretraživanje – funkcionalnost, problemi pisma i jezika, podrška za različite tipove fajlova i podataka

Digitalizacija dokumenata u slučaju primene OCR procesa je komplikovana, spora i često zahteva rad u fazama. OCR je sada već toliko razvijen da jednostavno registruje istovremeno prisustvo napr. dva pisma u dokumentu. Međutim, i dalje se zahteva pažljiva korektura i redaktura dokumenta nakon digitalizacije. Istovremeno korišćenje dva (ili više) jezika u istom dokumentu otežava korišćenje spelling checker-a (na primer kada analizira reference). Obuka službenika za digitalizaciju ovakvih dokumenata je veoma složena. Dokumenti koji sadrže tekst sa različitim pismima i jezicima značajno otežavaju njihovo implementiranje u sisteme za pretraživanje full text-a

Digitalizovani dokumenti mogu da se pretražuju:

Po punom tekstu, ako su podvrgnuti OCR procesu i procesu lekture

Po referenci, ako su u vidu slike, sa dodatim dokumentalističkim opisom (struktura opisa u okviru ili van standarda), odnosno ukoliko ne postoji digitalizovana verzija dokumenta

U oba slučaja javlja se problem potpunosti odziva pretraživanja usled problema sa OCR i/ili nepostojanja spelling checker-a. Digitalizacija dokumentacije u državnoj upravi u sprezi sa OCR tehnologijom uslovljava primenu spelling checker-a za različite jezike (važan aspekt ulaska u EU). Još uvek ne postoji elektronski “državni” rečnik i spelling checker (u smislu obuhvata i nivoa verifikacije) što suštinski otežava razvoj valjanih dokumentacionih sistema [5] [6]. Treba imati u vidu regularno prisustvo dva pisma. Dokumentalistička obrada elektronskih kopija dokumenta (odnosno grupa dokumenata) složena je i obuhvata:

Dokumentalistički opis slike dokumenta

Dokumentalistički opis dokumenta koji je čitljiv tekst procesorom

Naglašavamo da elektronski dokument mora da se obradi, u dokumentalističkom smislu, mnogo detaljnije nego klasični papirni dokument [9], [11]. Naime, papir ide na relativno preciznu lokaciju u arhivi, a elektronski dokument ide u elektronsku arhivu gde je lokacija određena atributima – deksriptorima dokumenta. Proliferacija elektronskih dokumenata je danas već tolika da je teško obavljati linearno pretraživanje, odnosno pregled dokumenata, dok se ne nađe ono što tražimo. U pitanju je složen proces i pored toga što raspoložemo računarima.

Ugrađivanje nekorigovanog dokumenta u sistem za pretraživanje teksta ne može da se dopusti [5], (imamo veoma loša praktična iskustva – kada se ne pronade relevantan dokument nastaju veliki problemi) – pogrešno ispisani tekst čini u mnogim slučajevima dokument nevidljivim u odnosu na sistem za pretraživanje. Kvalitet dokumentalističke obrade u mnogome zavisi od raspoloživosti rečnika (listi deskriptora) specifičnih za oblasti kojima se bavi Ministarstvo ili bilo koji drugi državni organ. Na primer, jedan od velikih problema je često menjanje naziva organizacionih jedinica ili institucija koji-

ma pripada dokumentacija (problem hronoloških sinonima i sl.). Problemi mogu da često nastanu i usled karakteristika S/W sistema za pretraživanje, koji nije opremljen za istovremeni rad sa dokumentima na više jezika i pisama.

Ponovo naglašavamo da je formalna valjanost dokumenta koji se dobija kao rezultat pretraživanja, u državnoj upravi od veoma velikog značaja, pošto na osnovu sadržaja dokumenta službenik donosi odluke. Bez obzira da li tangiraju državu ili građanina.

6. Zaključak

Za ulazak u Evropsku Uniju neophodno je imati razrađenu i proverenu metodiku rada sa elektronskim dokumentima na nivou državne uprave. Za ulazak u Evropsku Uniju potrebno je imati verifikovan rečnik(!) i program za proveru sintakse (spelling checker) kao i druge sisteme za dokumentalističku obradu podataka. Digitalizacija svih mogućih izvora podataka i pojava koji čine nacionalnu baštinu je prvorazredni instrument zaštite nacionalnih interesa. Nacionalni centar za digitalizaciju predstavlja dobru organizacionu osnovu (u ovom trenutku i jedinu) za rad na opisanim problemima.

References

- [1] N. Popović, B. Vratonjić, *EDI u državnoj upravi*, Zbornik radova 4. YUEDI konferencije 'Stanje i tendencije razvoja EDI tehnologije u svetu i kod nas', Budva, 1997.
- [2] M. Gordillo, *Access Networks for the Communication and Information Society*, Alcatel Telecommunications Review – 3rd Quarter 1996
- [3] Darlene Fichter, *It is time to convert? Migrating Static Pages to Database-Driven Web Sites*, Intranet Librarian, Web site management, ONLINE January/February 2005, pp. 51–53
- [4] Kim Guenther, *Choosing a Search Engine*, Web site management, ONLINE January/February 2005, pp. 48–50
- [5] Z. Jovanović, B. Lazić, N. Popović, B. Marović, M. Vitorović, V. Dikanović, D. Pisić, B. Vratonjić: *Projekat: Baze podataka u punom tekstu za SMIP*; Elektrotehnički fakultet Univerziteta u Beogradu – RCUB, oktobar 1996.
- [6] Nikola Popović, Branislav Vratonjić, Mirjana Milenković, *Some aspects of the model of the FMFA & DCM documentation system*, YU INFO 2002, 11–15.3.2002, Kopaonik, Zbornik apstrakta
- [7] Branislav Vratonjić, Nikola Popović, Jasmina Novaković, *The concept of the e-MFOA model based on the e-government model*, 7. Kongres JISA, e-Balkan, 3–8 jun 2002, Herceg Novi, Zbornik radova
- [8] N. Popović, M. Mihajlović: *Automatizovana pisarnica*, Savetovanje "Informacioni sistemi državnih organa uprave", Beograd, Zbornik radova s. 68, 1996.
- [9] N. Popović, J. Mirčevski, S. Jovanović, G. Putnik: *Baza podataka o novim tehnologijama*; Zbornik radova, Konferencija zajednice JUPITER, Cavtat, 1989.
- [10] J.P. Silva, G.D. Putnik, M.M. Cunha, *Technologies for virtual enterprise integration*, published in *Business Excellence: Performance measures, Benchmarking and Best Practices in New Economy*, Univ of Minho Press, 2003, pp. 706–712
- [11] Časlav Očić i Aleksandra Pošarac (priređivači): *BEOGRADIKA, Bibliografija radova o Beogradu*, Institut ekonomskih nauka, Beograd 1988.

**ANALYSIS OF THE PROBLEM
OF DOCUMENT DIGITIZATION
IN GOVERNMENTAL ADMINISTRATION**

Abstract: The text contains personal views of the authors regarding the problem of the digitalization of documents in the state government. The process of organizing the digitalization is analyzed. The actions to be taken, from the designing of documentation systems to the organization of work in a working environment, are justified. There is also an analysis of problem that remains unsolved during document processing in general and in the field of Information retrieval.

Key words: government administration, record keeping, digitization, information society

nikola.popovic@smip.sv.gov.yu

julijana@afrodita.rcub.bg.ac.yu