

Digitalizacija naučne građe – metode i rešenja

Zoran Zdravković

Apstrakt — Digitalizacija tehničke i naučne građe je prvi korak digitalizacije u nauci uopšte. Biblioteke, opšte i specijalizovane, su centralna mesta gde se nalazi postojeća naučna građa i prikuplja nova. Digitalizacija je u ovim ustanovama uveliko otpočela i nalazi se u različitim fazama razvoja. Od posebnog interesa je digitalizacija građe iz domena tehničkih nauka zbog prisustva raznorodnih tipova dokumentacije: crteža, fotografija, planova, патената, opisa pronalazaka, video materijala i druge građe. Pored osnovnih pojmova i definicija, u radu će biti analiziran analogno / digitalni koncept, principi digitalizacije različitog tipa informacija i vrsta građe (analogni signal, tekst, slika, audio i video materijal), razmotrene posebnosti i limiti digitalizacije naučne građe. Kroz zaključna razmatranja biće prikazani svetski trendovi digitalizacije i digitalne transformacije.

Ključne reči — digitalizacija naučne građe, biblioteke, digitalizacija tehničke dokumentacije

I. UVOD

DIGITALIZACIJA - OSNOVNI POJMOVI

Analizom postojećih definicija koje na brojne načine pokušavaju da preciziraju fenomen digitalizacije, može se uočiti da termin *digitalizacija* obuhvata proces prevođenja informacije (podataka) iz originalno analognog oblika u digitalnu formu. Digitalizacija se najšire definiše kao proces prevođenja analognog u digitalni oblik, pri čemu izvor (*analogni objekat*) može biti veoma raznorodan: signal, zvuk, dokument, slika, trodimenzionalan objekat... a rezultat procesa je *digitalni objekat*, odnosno digitalna forma: hologram, digitalna fotografija, digitalni zvuk, digitalni signal..., visokog kvaliteta transformacije pri čemu se tokom procesa transformacije ne gube informacije. Najopštije, digitalizacija predstavlja beleženje analognog signala u digitalni oblik pri čemu se digitalizovana građa organizuje u digitalne kolekcije i može se pretraživati i koristiti od strane korisnika.

U digitalnom formatu informacije se organizuju u diskretne jedinice podataka – bitove, binarne podatke, koje računari i drugi digitalni uređaji (npr. digitalna kamera ili digitalni slušni aparat) mogu procesuirati.

Digitalizacija obuhvata proces snimanja analognih signala i čuvanje rezultata u digitalnom obliku. Započinje sensorima, koji registruju analogne signale poput svetla i zvuka, da bi se potom zabeležene analogne forme transformisale u ekvivalentne digitalne formate uz upotrebu čipa za analogno/digitalnu konverziju. Proces funkcioniše pretvaranjem neprekidnog toka signala i analognih tipova podataka u diskontinuirane vrednosti, koje se potom uzorkuju (sempliraju) u redovnim jednakim vremenskim intervalima kako bi rezultirali digitalizovanim izlazom. Primer su audio datoteke (najčešće frekvencije od 44,1 do 192 kHz) pri čemu se audio datoteka frekvence 48.1 kHz uzorkuje 48.000 puta u sekundi. Proces

digitalizacije je efikasniji i kvalitetniji ukoliko se obavlja na višim nivoima uzorkovanja.

Tekst i slike digitalizuju se po sličnom principu: skener snima sliku ili „sliku“ teksta i pretvara je u datoteku slika – *image file*, kao što je bitmap. Program optičkog prepoznavanja karaktera (OCR) analizira svetle i tamne oblasti tekstualne površine kako bi identifikovao slovo ili numerik (cifru) pretvarajući dalje svaki takav karakter u ASCII kod.

Digitalizacija audio i video materijala obavlja se konverzijom analognog u digitalni signal, pri čemu se u ovom procesu konstantno varijabilan analogni signal transformiše bez gubitka i promene bitnog sadržaja, u višeslojni (digitalni) signal.

Digitalizacija informacija olakšava njihovo čuvanje, pristup, prezentaciju i deljenje. Originalan (analogni) istorijski dokument, muzejski eksponat, rukopis u biblioteci, tehnički crtež i drugi tip naučne građe može biti dostupan jedino posećivanjem njegove fizičke lokacije, što za posledicu ima visok stepen lokalizovanosti, dok digitalizacijom postaje dostupan ljudima širom sveta, uz visok nivo demokratičnosti a nizak stepen lokalizovanosti. Digitalizacija naučne i tehničke građe, kao i naučno-istorijske baštine u celini, dobija sve više na zamahu i beleži trend porasta.

II. ANALOGNO / DIGITALNO

Odnos analognog i digitalnog predstavljanja podataka (informacija i signala) može se definisati kroz određivanje pojmova i termina *analogni* i *digitalni* oblik informacije.

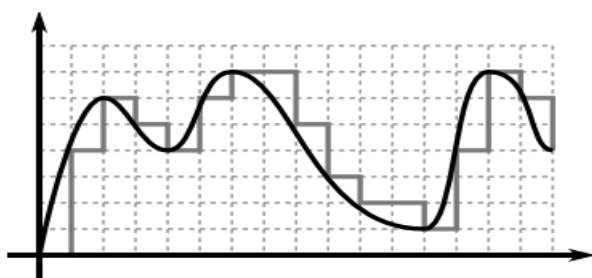
Analogni oblik izvora označava neprekidan signal iz okruženja: zvučni signal, svetlost, temperaturu, vlažnost vazduha, pritisak, elektromagnetno zračenje i druge fizičke fenomene i informacije iz sveta koji nas okružuje.

Digitalni oblik signala je brojčane, numeričke prirode – digitalne (od lat. *Digitus*: broj, cifra), i predstavljen je kontinualnim nizom nula i jedinica, binarno. Svaki digitalni podatak bez obzira na pojavni vid (tekst, slika, zvuk, animacija, film...) može se brojčano (digitalno) predstaviti celim brojevima dok se svaki pojedinačni broj može predstaviti nizom nula i jedinica, binarnim određivanjem prisustva (jedinica), odnosno odsustva (nula) signala od izvora informacija ka recipijentu.

Analogni mehanički fotoaparati, opremljeni analognom tehnologijom, koristili su analognu optiku da bi izvor svetlosti i refleksiju od fotografisanog objekta zabeležili kao sliku na fotografski film, a potom hemijske reakcije kako bi se film razvio i slika sa negativna prenela na fotografski papir. Digitalni fotoaparati koriste digitalnu tehnologiju da bi optikom, sa objektivom preneli sliku do CCD senzora (Charge-Coupled Device Sensors) sastavljenih od velikog broja fotoosetljivih elemenata pri čemu svaki mikroelement

predstavlja jednu tačkicu slike (Pixel), čime se svetlost koja ulazi u objektiv, reflektovana sa fotografisanog objekta, pretvara u niz tačaka na određenoj površini, broj tačaka po inču (dpi – dots per inch) predstavljenih brojevima, digitalno, odnosno binarno, prikazujući prisustvo/odsustvo osvetljenja.

Analogne signale karakterišu kontinuirani podaci i informacije, celi i u neprekidnom nizu, prikazani blagim sinusoidama, pri čemu amplituda može imati bilo koju vrednost i blage prelaze od pika do pika graničnih vrednosti, uz konstantnu fluktuaciju vrednosti sa protokom vremena, dok digitalni signali nisu kontinuirani: karakterišu ih oštri „stepenici“ – skokoviti prelazi, jasno određene diskretne vrednosti, respektabilno jasno izraženi pikovi, pri čemu amplituda digitalnog signala može imati samo ograničeni broj vrednosti. Za razliku od analognog signala koji može imati bilo koju vrednost, digitalni signal može imati jedno od dva stanja prekidača: uključeno ili isključeno, uz binarnu mogućnost prisustva ili odsustva signala od izvora ka recipijentu (postoji ili ne postoji napajanje). Tako binarni kod u računarskoj primeni ima vrednost 0 ili 1; Morzeova azbuka sastavljena je od kombinacija dve veličine: tačke (·) i crte (–).



Slika 1: Uporedni prikaz analognog signala (puna crna linija sa varijabilnim kvantitativnim vrednostima po protoku vremena) i digitalnog signala (stepeničasta linija sa diskretnim, određenim vrednostima) [1]

Suštinska razlika između analognog i digitalnog leži i u postojanosti kvaliteta i prednosti digitalnog signala jer se njegov kvalitet ne gubi prenosom.

Konverzija analognog u digitalni signal praćena je gubicima. I pored razvoja digitalne tehnologije i postizanja sve veće preciznosti digitalnog signala, ostaju ograničenja tehnološkim mogućnostima očitavanja i zapisivanja, brzine prenosa, frekvencije centralne procesorske jedinice (CPU), brzine procesora, kompresije zapisa, memorijskim kapacitetom i drugim tehničkim limitima. Digitalna tehnologija je postigla visok nivo kvaliteta digitalnog zapisa i unapredila ga, u meri da se ljudskim čulima, inače analognim, ne može primetiti razlika u kvalitetu i sadržaju digitalizovane verzije ostvarene konverzijom analognog objekta ili signala.

Razvoj video-monitora i ekrana omogućio je unapređenje kvaliteta prikaza tehničke građe i porast rezolucije. Rezolucija kao mera razlučivosti grafičkog sadržaja, odnosno stepen vidljivosti detalja, izražava se u tačkama po inču (dpi), odnosno pikselima po inču (ppi). Veća rezolucija donosi bolji kvalitet slike/dokumenta ali i porast veličine digitalnog fajla i zahteva veći memorijski prostor potreban za smeštaj.

Digitalizacijom se pored fajlova visoke rezolucije (300-1.200 dpi) pri skeniranju materijala za master skenove koji se trajno čuvaju, kreiraju i snimaju i digitalne datoteke tipično niske rezolucije od 72 dpi (tzv. „ekranska rezolucija“), koje se koriste samo za pregled na

ekranu, uz tzv. *thumbnail* prikaz „malih sličica“, i datoteke rezolucije od 300 dpi za kreiranje digitalnih objekata u procesima digitalizacije tekstualne građe, fotografija, karata, tehničkih crteža, planova i sl. sa ciljem optimalnog smeštaja, distribucije i korišćenja digitalizovane tehničke građe.

Iako je postignut visok nivo digitalne konverzije ona ipak nije savršena i potpuno jednaka analognoj stvarnosti. Iako ljudsko oko nije u mogućnosti da razlikuje dve susedne nijanse jedne iste boje na ekranu monitora, računari moćnih performansi sa naprednim grafičkim kartama prikazuju mnogo više boja nego što se ljudskim okom može registrovati koristeći 32-bitni i viši kvalitet boja. Uobičajene vrednosti dubine boja kreću se od 1 do 64 uz bitnu osobinu da sa porastom broja informacija eksponentno raste i broj boja, ali respektivno rastu i veličina datoteke, dužina vremena za transfer i učitavanje, iscrtavanje i štampanje, ali je veći i kvalitet, izražajnost i vernost boja. Standard kvaliteta boja ekrana današnjih računara je 16-, odnosno 32-bitna dubina boje, a preporučeni parametar za snimanje pri digitalizaciji tekstualne građe, fotografija, karata, planova, tehničkih crteža i slične tehničke građe je dubina boje 8-16 bita za sivu skalu (Gray Scale) i 24-48 bita za pun kolor.

Specifični su limiti i analogne tehnologije. Oni ometaju proces komunikacije koga u slučaju jednosmernog komunikacionog kanala čine: izvor informacije (emiter) → medij prenosa podataka → prijemnik signala, informacije, tj. podataka. Faktori uticaja na komunikacioni kanal mogu biti 1) spoljni činioci: „informacioni šum“ ili „buka“ koji remete prenos signala, odnosno analogni prenos informacija, i 2) udaljenost izvora i recipijenta kada kvalitet signala i prenosa informacije opada sa udaljenošću.

Nedostaci digitalnog slanja, obrade i prijema signala potiču iz realnog okruženja koje je po prirodni analogno, ali osnovni nedostaci digitalne konverzije su:

1. potreba stalnog kodiranja i dekodiranja kako bi se informacija poslala, odnosno primila, moguća u dva smera konverzije: A/D konverzija (kodiranje analogne u digitalnu informaciju) D/A konverzija (dekodiranje digitalnog u analogni podatak, signal ili informaciju);
2. duže vreme potrebno za slanje digitalnog signala;
3. porast količine podataka u procesu digitalizacije radi izrade i postizanja kvaliteta digitalnog objekta (gustina, odnosno visoka rezolucija slike ili video zapisa, kvalitet audio zapisa i sl. – primer: istovetan audio zapis na LP gramofonskoj ploči u analognoj formi očitavan iglom staje na spiralu dugu 500 m, dok je spirala CD-a duga 6 km, uz neophodan laser kako bi se očitao audio zapis.

Utuda je kompresija podataka potencijalno dobro rešenje naročito pri transportu na daljinu, deljenje na mreži i sl. Kompresija s gubitkom (engl. Lossy Compression) je tehnologija sažimanja analognih podataka i kreiranje digitalnih podataka s prihvatljivim gubicima u odnosu na analogni izvor. Obuhvata tehnike kompresije kod kojih se gubi prihvatljiva količina podataka pri čemu se eliminišu suvišne, redundantne ili nepotrebne informacije. Najrasprostranjenija primena i upotreba je kod audio, video i multimedijalnih datoteka (JPG i TIFF su najpoznatiji formati za

digitalne fotografije, MP3 i OGG za audio fajlove, a MPEG za video-snimke). Preporučeni formati za digitalizaciju fotografija su nekompresovan TIFF, kompresovan JPEG bez gubitka informacija i PDF format.[2]

TIFF (eng. Tagged Image File Format) nekompresovan format idealan je za smeštaj arhivskih datoteka i operisanje sa tekstualnim dokumentima i fotografijama s tim što je veličina TIFF datoteke znatno veća u odnosu na druge formate.

JPEG format (eng. Joint Photographic Experts Group Format) kao vrsta kompresovane datoteke idealno je rešenje za smeštaj kopija dokumenata i za smeštaj, transport i korišćenje dokumenata i fotografija na vebu. Veličina datoteke u JPG formatu znatno je manja od one u TIFF formatu. Kompresijom se dobija datoteka manjeg kvaliteta, ali značajno manje veličine i zauzima mnogo manje memorijskog prostora od originala (JPG fotografija tako nastala može biti manja i za više od 80% originala bez (ili sa malo) vidljivog efekta umanjjenja dok kompresovani MP3 fajl može biti jedna desetina veličine originalne audio datoteke uz neprimetnu razliku zvuka ljudskom uhu) što ovaj metod kvalifikuje kao idealan za multimedijalni sadržaj, ali isto tako i nemogućim za binarne datoteke, kod kojih ne sme doći do gubitka informacija, kada se primenjuje kompresija bez gubitka i kreiraju tzv. fajlovi bez gubitaka.

Odnos kvaliteta i veličine datoteka, odnosno kompromis umanjjenja podataka originala uz balans kvaliteta, od izuzetnog značaja je za skladištenje, prenos i korišćenje sadržaja. Dobro dizajnirana tehnologija kompresije uz gubitke značajno smanjuje veličine datoteka do mere da se degradacija ne uočava. Kompromis između veličine i kvaliteta datoteke ostaje značajna stavka procesa digitalizacije naučne i tehničke građe.

Kompresija bez gubitka (engl. Lossless Compression) obuhvata tehnologiju komprimovanja digitalnih podataka kojom se mogu povratiti sve informacije originalnog sadržaja kako je izgledao pre kompresije. Kompresija podataka bez gubitaka je reverzibilna kompresija podataka jer ne degradirajući podatke, svaki bit informacije koji je izvorno bio u datoteci ostaje i nakon kompresije, a sve informacije mogu u potpunosti biti obnovljene nakon raspakivanja i dekompresije, što kod kompresije sa gubitkom nije moguće.

Ova tehnika kompresije preporučuje se i koristi kao dobar izbor za digitalizaciju tekstualnih i tabelarnih datoteka, računskih podataka, ali i kod naučnih članaka i većih tekstualnih formi (tehnička literatura, udžbenici, tehničke enciklopedije, naučni radovi, disertacije, patentni, standardi, smernice, uputstva i prateća dokumentacija) gde gubitak pojedinih delova, reči ili cifara, i grafičkih podataka može predstavljati problem. GIF digitalni format slika (engl. Graphics Interchange File) koristi se na vebu kao grafički interfejs za digitalnu kompresiju bez gubitaka. Standardni format podataka kompresije bez gubitaka su „loosless“ fajlovi, poput datoteka FLAC kodek formata – audio formata sličnog MP3 formatu samo bez gubitaka. Ovaj tip formata je najpoborniji za digitalizaciju zvuka, jer je audio kompresovan u FLAC bez gubitka kvaliteta.

Najbitnija prednost digitalnog prenosa signala leži u mogućnosti prenosa svakog mogućeg pojavnog vida informacije obuhvatajući tekst, informaciju o boji, zvuk, animaciju, video signal itd. Za razliku od analognog signala koji je po svojoj prirodi podložan ometanju, informacionom šumu i smetnjama, digitalni podatak je postojan. Svaki, pa i bilo kakav, jednom digitalizovani podatak može se čuvati, smestiti, prenositi i koristiti, istovetnog stepena kvaliteta, upravo procesima kodiranja i dekodiranja: faks uređaji prenosili su tekst i grafički sadržaj između dva telefona; računarski mrežni modem omogućavao je digitalnu MODulaciju i DEModulaciju po principu kodiranja i dekodiranja svih pobrojanih vidova informacija između dva udaljena računara, konverzijom digitalnog signala iz računara u analogni koji može biti prenet na velike daljine postojećim analognim instalacijama, najčešće telefonskim vezama.

Motivi i razlozi za uvođenje digitalizacije između ostalih leže i u prednosti digitalizovanih podataka koji imaju mogućnost popravljavanja i oporavka narušenog ili oštećenog signala, time što se može ponovo poslati i sklopiti na određitu dopunom nedostajućih delova informacije i restaurirati.

III. SPECIFIČNOSTI DIGITALIZACIJE NAUČNE I TEHNIČKE GRAĐE

Prilikom digitalizacije naučnog materijala kao vrste građe potrebno je razmotriti materijal koji je potrebno digitalizovati, obaviti selekciju građe, racionalno sagledati neophodne hardverske i softverske resurse i planirati ljudske resurse.

Selekcija naučne građe, odnosno dokumenata i objekata koji bi bili digitalizovani, od presudnog je značaja i obavlja se prema prioritetima:

- vrsta građe (papir, mikrofilm, tehnički crtež, plan...)
- značaj građe, (istorijsko-naučni značaj, značaj za instituciju, za istraživače...)
- stanje u kome se nalazi
- stepen oštećenja
- uslovi smeštaja
- frekventnost korišćenja
- zakonska ili druga obavezujuća regulativa digitalizacije
- obaveza arhiviranja i čuvanja tehničke dokumentacije i sl.

Digitalizacija, proces konverzije dokumenata u digitalnu formu, zahteva definisanje ključnih elemenata koji će, pre svega, biti razmatrani pri određivanju vrste materijala – nosioca naučne informacije, odnosno tipa građe. Među najznačajnijim su veličina, oštećenost, količina, koloritet, vrsta dokumenta, tajnost, odnosno javnost dokumenta, atributi pretrage, rezolucija i tip izlaznog fajla (veličina, kompresija, broj boja, dubina boje), vid pregleda digitalnih fajlova, smeštaj i organizacija u digitalne kolekcije, a na kraju i način korišćenja digitalnih fajlova u okviru digitalne biblioteke institucije, mogućnost pristupa, dozvola za preuzimanje i autorska prava, što na kraju definiše i cenu, odnosno isplativost procesa, tj. projekta digitalizacije.

Iamjući sve ove parametre i posebnost naučne građe kao materijala za digitalizaciju iskazuje se potreba neophodnosti kreiranja, grupisanja i ažuriranja dokumentacije koja obuhvata čitav

spektar navedenih elemenata, dokumentovanih u vidu formulara, ili dnevnika digitalizacije, za svaki digitalizovani objekat. Ovaj dokument prati čitav tok kretanja naučne dokumentacije u procesu digitalizacije. S druge strane, definiše radni zadatak operateru na skeneru a ujedno je i dokument na osnovu koga će biti kreiran finalni proizvod na zahtev korisnika, istraživača ili institucije. Ključna faza u procesu skeniranja je **kontrola kvaliteta kreiranih digitalnih fajlova**.

Podela prema vrsti materijala koji se digitalizuje definiše i opremu za skeniranje. Najobimnija (i najčešća) vrsta materijala na kojoj se nalazi naučna dokumentacija je **papirna dokumentacija**, i to:

- **papirna dokumentacija do A3 formata koja nije povezana**
 - Dokumentacija u registratorima ili kutijama (A5, A4, A3+)
 - fakture, ugovori, dosijei, predmeti i ostala razna akta
- **papirna dokumentacija do A3 formata – osetljiva, uvezana i ne sme se razdvajati**
 - knjige, časopisi, novine, brošure, sveske...
- **papirna dokumentacija od A3 do A1 formata uvezana ili pojedinačna**
 - osetljiva građa, povelje, plakete, grafički crteži, knjige velikih formata, poster, plakati..
- **papirna dokumentacija A0 formata koja nije uvezana**
 - tehnički crteži, planovi, mape, geografske karte, poster, plakati...



Slika 2. Skeniranje papirne nepovezane građe skenerima sa ulagačima za dokumentaciju (tzv. document feeder), automatizovanim transportom i opcijom masovne digitalizacije

Masovna digitalizacija se najčešće koristi pri skeniranju i obradi poslovne dokumentacije ali se može primeniti i na naučni tip dokumenata (tzv. *protočni system obrade*). Prilikom određivanja grupe dokumenata u okviru naučne dokumentacije za masovnu digitalizaciju od značaja je **stepen automatizacije digitalizacije** koji se može primeniti. Veći stepen automatizacije direktno je srazmeran količini digitalizovane naučne dokumentacije, kvantumu i nivou obrade, respektivno i smanjenju cene, efektivnosti i ekonomičnosti/isplativosti procesa digitalizacije. Step automatizacije može biti poprilično visok a proces digitalizacije sofisticiran uz primenu tehnologije bar kodova, još i više uz primenu OCR tehnologije.

Naučna dokumentacija koja se ne može digitalizovati i obrađivati protočnom tehnologijom (nije moguće njeno propuštanje kroz document feeder), skenira se na **skenerima sa staklenom pločom** (ravno staklo – flatbed, ili V staklo). Ovakav vid digitalizacije naziva se **koračni postupak**, svaki dokument se ponaosob skenira, bilo ručno ili automatizovano, kao kod robot-skenera sa **robotizovanom rukom** koja automatizovano okreće strane.

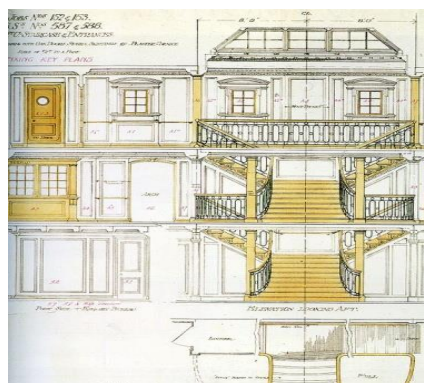


Slika 3. Skeniranje papirne građe (povezane) skenerom sa staklenom pločom – flatbed skener sa kamerom i lampom odozgo (overhead)

Ova grupa naučne dokumentacije deli se prema veličini:

- dokumentacija do veličine A4 formata
- dokumentacija do veličine A3 formata
- dokumentacija do veličine A2 formata
- dokumentacija do veličine A1 formata
- plakati i tehnički crteži do veličine A0 formata.

Posebno specifičan detalj vezan za tehničku dokumentaciju je da je ona **koncipirana kao jedinstven oblik**: jedan predmet koji se sastoji od tehničkih crteža, ali i od poslovne prateće dokumentacije A4 i A3 formata koja opisuje detalje tehničkog crteža, razne dozvole i ostalu pravnu dokumentaciju kao prateći material. Specifičnost ove grupe dokumenata je da završni digitalizovan fajl imora biti jedinstven fajl i ne može se razdvajati. Za ovakav tip tehničke i naučne građe koristi se hibridna tehnologija digitalizacije, skeniranja i softvera koji omogućavaju praćenje hronologije dokumentacije u predmetima prilikom skeniranja i jednostavne izmene skenera kojima će se obrađivati dokumentacija. Za tehničke crteže koriste se najsavremenija oprema i uređaji, a kada se u predmetu naiđe na manju dokumentaciju u softveru se izvršava odabir drugog skenera, nastavlja skeniranje i obrađuje, odnosno snima u jedinstven fajl. Ovakvim konceptom digitalizacije zastupljene su i prisutne sve vrste zapisa (papir, negativ, mikrofilmovi), objedinjene kao sastavni deo jedinstvenog dosijea.



Slika 4. Tehnički crtež sa pratećom dokumentacijom kao celina

OCR - OPTIČKO PREPOZNAVANJE KARAKTERA

Osnovni zadatak OCR softvera je pretvaranje digitalnih slika (JPEG, TIFF, PNG...) na kojima su skenirani tekstovi sa matričnih štampača, kućaćih mašina, iz knjiga, novina, časopisa ili tehničke dokumentacije, u promenljive digitalne tekstualne forme, tako što se iz rastera tačaka sa slika, softverski prepoznaju slova, reči i čitavi tekstovi.

Cilj OCR-ovanja je dematerijalizacija papirnih dokumenata, pri čemu se primenjuju sledeće discipline:

- masovna digitalizacija, kada se postiže velika brzina na velikom zahvatu
- automatizovana separacija dokumenata, auto-sortiranje, strukturiranje dokumenata i sadržaja
- Optical Character Recognition, Intelligent Document Recognition, indexiranje
- kompresija dokumenata i digitalnih slika.

Vektorizacija tehničkih crteža

Pri digitalizaciji tekstualne građe i poslovne dokumentacije primenjuje se tehnologija optičkog prepoznavanja karaktera (OCR procesuiranje) u procesu transformacije "slika teksta" u editabilan tekst, dok se kod digitalizacije tehničke dokumentacije koristi proces vektorizacije za transformaciju linija iz rastera tačaka u vektorski oblik. Sam proces OCR-a smanjuje potrebu za prekućavanjem tekstova sa digitalnih slika, a vektorizacija respektivno smanjuje potrebu za precrtavanjem tehničkih crteža.



Slika 5. Vektorski oblik tehničkog crteža

(izvor: <https://www.dass.rs/images/1/vektorski%20oblik.png>)

IV. ZAKLJUČNA RAZMATRANJA – GLOBALNI TRENDOVI U DIGITALNOJ TRANSFORMACIJI

Iako je digitalizacija u podmaklom stadijumu još nije kasno da se uvede minimum opštih zahteva i zajedničkih kriterijuma koji bi bili poštovani u daljem toku procesa digitalizacije naučne građe i u nauci uopšte. Imajući u vidu savremena kretanja na polju digitalizacije navodimo 10 ključnih globalnih trendova u digitalnoj transformaciji, kao budućim stremljenjima naučnih institucija i društva u celini: [3]

1. Internet of Things (IoT)

IoT je najbrže rastući koncept na Internetu (procene su najavljujvale više od 8 milijardi „Things (stvari)“ na Internetu u prvim godinama od pojavljivanja, da bi do 2020. godine broj dostigao i više od 30 milijardi), imajući u vidu da nisu samo stvari na Internetu u pitanju, već način na koji se koriste, povezuju, analiziraju i unapređuju digitalnu transformaciju. Pametni gradovi, kuće i zdravstvo su primeri oblasti koje najbolje demonstriraju značaj i primenu IoT-a. Pored neminovnosti daljeg razvoja i napretka u ovoj oblasti, dinamika rasta i razvoja će zavisiti i od IoT sigurnosti i ograničavanja rizika koji se javljaju u primeni. IoT je danas ključni digitalni trend, u osnovi je sledeća 3 globalna trenda, i biće njihov katalizator u budućnosti.

2. Analitika – Big Data

Ogromna količina informacija kreirana kroz IoT postaje snaga promena koje iz korena menjaju društvo, od proizvodnje do zdravstva, čak i načina funkcionisanja gradova, omogućavajući im da povećaju svoju efikasnost i profitabilnost. Slični rezultati se postižu i očekuju i u drugim oblastima, naročito kada su podaci naučnih istraživanja u pitanju.

3. Edge Computing

Trend koji postepeno preuzima primat od Cloud Computing-a: obrada podataka u realnom vremenu (real-time) postaje sve važnija sa drastičnim porastom količine podataka koje proizvodi rastući IoT. Pametni dronovi, autonomna vozila i AI pametni uređaji zahtevaju trenutno povezivanje i komunikaciju kroz IoT, što prevazilazi sadašnje mogućnosti Cloud Computing-a. I pored toga Cloud neće biti zamenjen u svim, bar ne u doglednom narednom periodu – još uvek će biti opravdana opcija u većini procesa.

4. 5G

Usled sve većeg pritiska narastajućeg kvantuma broja podataka kroz opet rastući IoT, mobilni provajderi se odlučuju za uvođenje 5G mreže i usluga. Prelazak na 5G se neće desiti preko noći: do 2025. godine 5G mreža će pokriti trećinu svetske populacije i omogućiti suštinsku i realnu digitalizaciju brojnih oblasti poslovanja i života.

5. Blockchain

Tehnologija poznatija po svojoj primeni u kriptovalutama i najčešće spominjana kao alat u finansijskom sektoru, postaje sve šira u raznorodnim oblastima: zdravstvena industrija i industrija zabave su samo neke od oblasti koje se priključuju primeni ove tehnologije, uz očekivanja su da će se posle 2020. godine više od 20% svetske finansijske trgovine odvijati na ovoj platformi.

6. Artificial Intelligence (AI)

Koncept koji predstavlja sposobnost mašina da učeći imitiraju inteligentno ponašanje ljudi opravdavaju očekivanja da će u budućnosti transport, proizvodnja i druge oblasti života biti digitalizovane i automatizovane uz pomoć veštačke inteligencije (AI). Poslovanje sve većeg broja kompanija oslanja se na podršku AI-a u marketingu, analitici, odnosu sa korisnicima, robotici i drugim oblastima. Ključne softverske kompanije ubrzano integrišu AI u svoje platforme sa jasnim ciljevima: ubrzavanje, ekonomičnija i pametnija automatizacija i digitalizacija procesa (u širokom dujapazonu primena od e-mailova, preko kreiranja sadržaja do industrijske proizvodnje).

7. Virtual (VR) i Augmented Reality (AR)

Uvećana stvarnost (AR) u profesionalnom kontekstu pokazuje bolje rezultate od virtuelne realnosti (VR), naročito usled povoljnije i jednostavnije primene, pogotovi u oblastima marketinga, edukaciji i dizajniranju novih proizvoda, gde se prednost daje AR-u u odnosu na VR. Puna komercijalna primena u globalno najlukrativnijoj kreativnoj industriji video igara potvrdiće se u narednim godinama.

8. Upravljanje glasom

Pretraživanje, obrada i upravljanje glasom rezultat je velikog napretka u nauci, naročito na polju analitike i hardverskim rešenjima pri čemu se ostvaruje novi nivo korisničkog iskustva kroz interakciju sa digitalnim platformama: pametne kancelarije, kuće i digitalni

asistenti su neki od primera u ovoj oblasti koja se bazira na napretku tehnologije prepoznavanja glasa i veštačke inteligencije.

9. Učenje na greškama

Failure-as-a-Service koncept po kome neuspeh postaje sam po sebi nova usluga, a učenje na greškama postaje alat u brzom prilagođavanju digitalnim uslovima poslovanja i omogućava dalji razvoj društva. Brzi neuspeh pomaže efikasnom prilagođavanju novom poslovnom okruženju, pre svega pri izradi prototipova novih proizvoda. Prepoznavanje grešaka i njihovo brzo ispravljanje postaju esencijalni deo uspešne digitalne transformacije, kao i inovativnog pristupa kreiranju novih proizvoda i usluga.

10. Digitalna transformacija kao preduslov opstanka

Porastom broja novih tehnologija koje se uvode u sve više oblasti ljudskog društva respektivno raste i brzina potreba za promenama. Naučne institucije i kompanije koje su ih prihvatile, i pored prisutnih ograničenja sa kojima se suočavaju, ostvaruju tržišnu prednost i povećavaju šanse za dalji razvoj. Dolazak digitalne transformacije i uvođenje nove poslovne kulture u brojnim institucijama postaje ključna prepreka: institucije koje su sposobne da se brzo i uspešno prilagode novim okolnostima imaju veće šanse za opstanak i uspeh.

LITERATURA

- [1] *Analog Signal*, URL: <http://www.chegg.com/homework-help/definitions/analog-signal-4> (12.3.2022.)
- [2] *Смернице за дигитализацију културног наслеђа у Републици Србији*, Министарство културе и информисања Републике Србије, 2017, стр. 18, URL: <http://www.kultura.gov.rs/docs/dokumenti/propisi-iz-oblasti-kulture/smernice-za-digitalizaciju-kulturnog-nasledja-u-republici-srbiji.pdf> (12.3.2022.)
- [3] „10 *svetskih trendova u digitalizaciji*“, <https://cdt.org.rs/index.php/2018/08/08/10-svetskih-trendova-u-digitalizaciji/>, (21.3.2022.)

ABSTRACT

Digitization of technical and scientific material is the first step of digitalization in science in general. Libraries, general and specialized, are central places where existing scientific material is located and new ones are collected. In these institutions, digitization has largely begun and is in various stages of development. Of special interest is the digitization of material from the domain of technical sciences due to the presence of various types of documentation: drawings, photographs, plans, patents, descriptions of inventions, videos and other materials. After basic terms and definitions of digitization, the paper will analyze the analog / digital concept, principles of digitization of different types of information and types of material (analog signal, text, images, audio and video material) and consider features, also limits, of scientific material digitization. Concluding remarks will present global trends in digitization and digital transformation.

Key words - digitization of scientific materials; libraries; digitization of technical documentation

Digitization of scientific material - methods and solutions

Zoran Zdravković