

## Priložnosti izrabljanja tehnologij digitaliziranja : prispevek za Akademijo 2020

Janko Klasinc  
Narodna in univerzitetna knjižnica

Digitalizacija že dolgo ne predstavlja le transformacije zbirk podatkov in informacij iz fizične v digitalno obliko, temveč tudi preoblikovanje raznovrstnih procesov, ki se odvijajo v družbi. Tako govorimo o tehnologijah, kot so umetna inteligenca, internet stvari, pametne zgradbe in mesta. Digitalizacija prispeva k produktivnosti in učinkovitosti gospodarstva ter širšemu socialno-ekonomskemu razvoju. Nove storitve in rešitve na vseh področjih družbe nam omogočajo, da digitalizacijo uporabljamo za nove načine poslovanja, komuniciranja, izobraževanja, zabave in na drugih specializiranih področjih. Digitalizacija se ne odvija zgolj na ravni posameznih organizacij in podjetij, temveč lahko govorimo o digitalizaciji celotne družbe.

Eden od bistvenih pogojev za digitalno transformacijo procesov in družbenih odnosov je čim več kakovostno digitaliziranih podatkov. Z digitalizacijo gradiva, ki ga hranijo knjižnice, arhivi, muzeji in druge dediščinske ustanove ustvarjamo neprecenljivo zbirko digitalnih podatkov, ki so na voljo za raziskovanje in preučevanje. Podatki so na ta način enostavneje dostopni najširšemu krogu uporabnikov, hkrati pa omogočajo razvoj pametnih aplikacij in naprednih storitev, ki temeljijo na razvoju umetne inteligence ter obvladovanja masovnih podatkov (ang. *big data*). Vidnost, prepoznavnost in zmožnost odkrivanja informacij, ki so bile v fizičnem okolju pogosto težje najdljive in dostopne, pospešujejo in krepijo raziskovalno dejavnost ter pozitivno vplivajo na razvoj družbe znanja ter demokratizacijo družbe. Poleg tega digitalizirano gradivo predstavlja bogat vir za uporabo v medijih, založništvu, kreativnih industrijah in turizmu ter pozitivno vpliva na zagotavljanje enakih možnosti dostopa do gradiva ranljivim skupinam prebivalstva, kot so gibalno ovirani, slepi in slabovidni ter starejši.

Kako se je najbolj smiselno lotiti digitalizacije?

Ko se lotevamo kakršnega koli projekta digitalizacije gradiva, moramo najprej določiti naš namen. Običajno pri tem igrata glavno vlogo vidika trajnega ohranjanja gradiva in zagotavljanja boljše dostopnosti do gradiva, saj z digitalizacijo zmanjšamo uporabo fizičnih izvirkov in z objavo na spletu omogočimo enostavnejši dostop širšemu krogu uporabnikov. Poleg teh dveh vidikov so pri izboru gradiva za digitalizacijo bistvena tudi sredstva, ki jih imamo na voljo.

Praviloma digitaliziramo le gradivo, ki je v ustreznem stanju oziroma, za katerega ocenimo, da ga tekom postopka ne bomo poškodovali. Pri knjižničnem in sorodnem gradivu smo pozorni predvsem na potencialne poškodbe vezave ali drugih strukturnih elementov, nosilca (papir, lepenka, pergament, itd.) in nanosa. Pri zagotavljanju varne digitalizacije bistveno vlogo igra oprema, s katero opravimo postopek. Ta mora zagotavljati varno in celostno digitalizacijo, izberemo pa jo skladno s tipom gradiva, ki ga bomo digitalizirali. Na trgu obstaja veliko različnih naprav (optičnih čitalnikov), vse pa niso primerne za digitalizacijo gradiva, katerega smatramo za kulturno dediščino. Zelo pogosti so [ploski](#) in [pretočni](#) optični čitalniki, s katerimi lahko varno digitaliziramo le novejša neobčutljivo nevezano gradivo manjših dimenzij. Za gradivo na krhkem papirju, ki je prevladoval med prvo polovico 19. st. in drugo polovico 20. st., so veliko bolj primerne naprave, pri katerih gradivo ne pride v stik s skenirno površino (ang. [overhead scanners](#)). Mnogi modeli tovrstnih čitalnikov vključujejo tudi podperne elemente, ki omogočajo skeniranje vezanega gradiva pod kotom manjšim od 180°. To je še posebej pomembno pri občutljivejšem ali zelo trdo vezanem starejšem gradivu, pri katerem je zaradi

preprečevanja poškod hrbta priporočljivo odpiranje samo do kota med 120° in 90° ali celo manj. Če imamo opravka z zelo veliko količino manj občutljivega vezanega gradiva, je lahko ustrezna rešitev [robotski optični čitalnik](#), ki omogoča samodejno obračanje listov. Prednost je predvsem velika hitrost zajema in zmanjšanje potreb po človeških virih za izvedbo postopka.

Pri zajemu v digitalno obliko skušamo čim bolj popolno zajeti vse vsebinske elemente gradiva ne glede na to ali gre za besedilno ali slikovno vsebino. Na končnem izdelku naj bi bili razpoznavni tudi najdrobnejši in najmanj poudarjeni deli vsebine. Pri nekaterih vrstah gradiva poleg vsebine informacijsko vrednost predstavljajo tudi oblika ali druge fizične značilnosti izvornika. V teh primerih je smiselno poleg vsebinskih zajeti tudi nevsebinske elemente, kot so platnice, hrbet, obreza, prazne strani itd. Pri tem lahko uporabimo fotografski aparat ali se odločimo za bolj napreden postopek [3D digitalizacije](#) oziroma [fotogrametrije](#). Postopek je vedno bolj uveljavljen na področjih [arheologije](#) in [muzealistike](#), uporabimo pa ga lahko tudi za digitalizacijo pisne kulturne dediščine kot so srednjeveški kodeksi in drugo gradivo, pri katerem je tudi fizični nosilec lahko predmet raziskovanja. S [postopkom fotogrametrije](#)<sup>1</sup> izdelamo večje število posnetkov določenega objekta z različnih zornih kotov. S pomočjo ustrezne programske opreme potem iz fotografij izdelamo [3D model](#)<sup>2</sup>, ki uporabniku prikaže, kakšen je predmet v vseh svojih fizičnih razsežnostih.

Ko po uspešno opravljeni digitalizaciji poskrbimo za dostop do gradiva, moramo poskrbeti tudi za ustrezne metapodatke, s katerimi zagotovimo, da je gradivo najdljivo. Uporabnikom je v veliko pomoč tehnologija optične prepoznavne besedila (ang. [Optical Character Recognition](#)), s pomočjo katere programska oprema prebere besedilno vsebino skenogramov (digitalnih slik) in jo shrani v obliki surovega besedila. Na ta način je prek iskanikov, ki so na voljo v spletnih digitalnih zbirkah, omogočeno iskanje po celotni besedilni vsebini gradiva. Kljub temu so za učinkovito in hitro identifikacijo digitaliziranega gradiva pomembni tudi čim bolj bogati deskriptivni (bibliografski) metapodatki, kot so avtor, naslov dela, leto izida ali nastanka, založnik, itd. Ker na spletu obstajajo številne digitalne zbirke in repozitoriji, ki temeljijo na interoperabilnosti in izmenjavi podatkov, je pomembno, da so metapodatki pripravljene v enem od standardiziranih formatov, ki zagotavlja avtomatizirano izmenjavo podatkov med različnimi platformami. Poznamo različne metapodatkovne formate, med katerimi sta na področju klasične knjižničarske katalogizacije najbolj znana [MARC21](#) in [UNIMARC](#) (v Sloveniji [COMARC](#)), ter formate, ki so uporabljeni v digitalnem okolju (npr. [Dublin Core](#), [XML](#), [MODS](#), [METS](#) itd). Končni digitalni objekt je tako sestavljen iz kakovostne digitalne reprodukcije fizičnega izvornika in urejenih, čim bolj podrobnih metapodatkov.

Digitalizirano gradivo resnično zaživi šele, ko je prosto dostopno na svetovnem spletu. Za ta namen je poleg tehnoloških rešitev, ki so na voljo, ključno predvsem področje avtorske pravice, kjer si pogosto stojita nasproti interes najširše javnosti po dostopu do znanja in informacij ter finančni interes imetnikov pravic. Zato je poleg razvoja tehnologije pomembno, da tudi avtorsko-pravna zakonodaja sledi napredku in se neprestano prilagaja novim možnostim, ki so na voljo za odprto in prosto razširjanje znanja. Le tako bomo lahko v največji meri izkoristili obstoječe tehnološke rešitve in tiste, ki še prihajajo.

---

<sup>1,2</sup> Avtor video posnetka in 3D modela je Uroš Grčar (Narodna in univerzitetna knjižnica, Oddelek za digitalne vsebine).