

ARHIVIRANJE WEBA PRAVNI I ETIČKI ASPEKTI

Tanja Buzina

*Nacionalna i sveučilišna knjižnica
u Zagrebu, Zagreb, Hrvatska*

KLJUČNE RIJEČI:

*nacionalna knjižnica, digitalni
obvezni primjerak, arhiv weba,
pravni aspekti arhiviranja weba,
Opća uredba o zaštiti podataka*

SAŽETAK

U počecima arhiviranja weba aktivnosti nacionalnih knjižnica bile su usmjerene na izradu načela i kriterija za izgradnju zbirke online publikacija, programe za prikupljanje sadržaja i uspostavljanje sustava za trajno čuvanje podataka. Iako pitanja odnosa između prava na zaštitu osobnih podataka i prava na pristup informacijama i slobodu izražavanja nisu nova u knjižničarskoj i arhivističkoj zajednici, novim se zakonskim propisima, posebno u vezi sa zaštitom osobnih podataka, fokus prebacuje na etičke i pravne aspekte arhiviranja weba i pristupa arhiviranim podacima.

Uvod

U radu će biti riječi o svrsi i inicijativama za uspostavljanja arhiva *weba*, metodama i postupcima arhiviranja, pravnim temeljima procesa arhiviranja, utjecaju zakonskih propisa u vezi sa zaštitom osobnih podataka na postupanje s podacima u arhivima *weba*. Iako se arhiviranjem *weba* bave knjižnice, arhivi, muzeji, privatne tvrtke, neprofitne udruge, privatne osobe u različite svrhe, ovaj će se rad ograničiti na arhive *weba* nacionalnih knjižnica u kontekstu njihove zadaće prikupljanja i čuvanja obveznog primjerka.

Nacionalna i sveučilišna knjižnica u Zagrebu je u suradnji sa Sveučilišnim računskim centrom Sveučilišta u Zagrebu (Srce) već 2003. pokrenula *Projekt uspostave sustava za preuzimanje i arhiviranje obveznog primjerka hrvatskih mrežnih publikacija* (Willer, Milinović 2005; Buzina, Willer 2007; Willer 2007). Voditeljica projekta bila je Mirna Willer, a sustav se temeljio na konceptu selektivnog pobiranja statičnih i dinamičnih *online* publikacija (Willer i sur. 2008). Testna faza arhiviranja započela je u siječnju 2004., sustav je u punoj funkcionalnosti pušten u rad u studenome iste godine.

Svrha arhiviranja weba

Arhiviranje sadržaja s *weba*¹ traje skoro isto toliko koliko postoji i sam *World Wide Web*.² *World Wide Web* bio je zamišljen kao sustav za povezivanje i razmjenu informacija među znanstvenicima na sveučilištima i znanstvenim institutima. S vremenom je postao nezaobilazan izvor podataka i, između ostaloga, važan nakladnički medij. U početku je, kao i svaki novi medij, podržavao vrste sadržaja poznate iz prethodne paradigme – tiskanog medija. Druga generacija *weba* (web 2.0) proširuje mogućnosti stvaranja novih sadržaja, formata i žanrova koji nisu postojali u tiskanom izdavaštvu.

Knjižnice kao čuvari kulturnog nasljeđa potpuno su svojih zbirki oduvijek su nastojale održati uključivanjem sadržaja na novim medijima. Devedesetih godina 20. stoljeća bila je to nova multimedijaska platforma WWW kao sredstvo objavljivanja.³ Knjižnice su rano prepoznale sadržaje na *webu* vrijedne uključivanja u svoje zbirke za trajnu pohranu (Klarin, 2005).

- 1 Već od 1996. neprofitna organizacija Internet Archive periodički prikuplja i pohranjuje digitalne sadržaje s *weba*.
- 2 World Wide Web, skraćeno WWW, W3, ili samo *web*. Pojmovi World Wide Web i internet često se, pogrešno, poistovjećuju.
- 3 Sva se elektronička građa kojoj se daljinski pristupa smatra objavljenom; vidi ISBD objedinjeno izdanje (2011:A 2.2).

Da bi se osiguralo čuvanje i dostupnost publikacija s *weba* potrebno je bilo uspostaviti sustave (arhive *weba*)⁴ koji omogućuju upravljanje procesom arhiviranja i obuhvaćaju aktivnosti koje se odnose na prikupljanje dijelova *weba* prema određenim kriterijima, preuzimanje i pobiranje, pohranu i trajno čuvanje i omogućavanje pristupa arhiviranim sadržajima. Arhivi *weba* poseban su dio sustava digitalne knjižnice.

Karakteristike samog *weba*, kao što je dinamičnost kreiranja novih sadržaja, njihova nepostojanost, nestabilnost i kratkotrajnost⁵, razlog su više za uspostavu sustava za njihovu trajnu pohranu i čuvanje.

Prema istraživanju Ntoulas, Cho, Olston (2004) svakog tjedna nastaje 8 % novih *web*-stranica, nakon godinu dana oko 50 % sadržaja na *webu* novi su sadržaji, 80 % *web*-stranica nije dostupno na izvornoj lokaciji (URL-u) nakon godinu dana, 13 % poveznica na referentnu literaturu znanstvenih članaka nedostupno je nakon 27 mjeseci.

Inicijative za početak arhiviranja sadržaja s weba

Važan poticaj za arhiviranja sadržaja s *weba* bile su preporuke Konferencije ravnatelja nacionalnih knjižnica (CDNL, 1996). Preporuke naglašavaju potrebu za novim zakonskim okvirom u kojem će publikacije objavljene na *webu* biti uključene u korpus obveznog primjerka.

UNESCO-ova Povelja o zaštiti digitalne baštine (*Charter on the Preservation of Digital Heritage*, 2003) također naglašava potrebu za trajnim čuvanjem baštine u digitalnom obliku u svrhu zaštite i osiguranja pristupa javnosti. Navodi se da bi pristup javno dostupnim sadržajima trebao bi biti oslobođen pretjeranih ograničenja, ali uz zaštitu osjetljivih i osobnih podataka.

Preporuka Europske komisije o digitalizaciji i *online* dostupnosti kulturnih artefakata i digitalnoj zaštiti (*Commission Recommendation of 24 August 2006 on the digitisation and online accessibility of cultural material and digital preservation*, 2006) ponavlja potrebu za pravnom regulacijom trajnog čuvanja i pristupa građi u digitalnom obliku uz poštovanje prava intelektualnog vlasništva i autorskog prava.

UNESCO-ova konferencija *The Memory of the World in the Digital Age*:

4 Sredinom devedesetih godina 20. stoljeća nastali su prvi arhivi *weba*. Više na: List of Web archiving initiatives, https://en.wikipedia.org/wiki/List_of_Web_archiving_initiatives.

5 *Web*-središta često se ažuriraju, proširuju ili prenamjenjuju, premještaju na druge lokacije, gase i nestaju.

Digitization and Preservation (2012) zaključuje također da je trajno čuvanje digitalnih sadržaja neophodno kako bi svaka osoba mogla ostvariti temeljno ljudsko pravo na pristup informacijama, a bolje razumijevanje digitalne okoline ključno je za izradu modela digitalne pohrane koji poštuju osnovna pravna načela ustanovljena uredbenim okvirima institucija, usklađenost prava na privatnost i prava na znanje s ekonomskim pravima, poštovanjem vlasništva i nadzora autohtonog kulturnog nasljeđa i tradicionalnog znanja u digitalnom formatu.

Pravni temelj za arhiviranje weba

Nacionalni zakoni koji reguliraju postupanje s digitalnim obveznim primjerkom, uključujući prikupljanje i pohranu obveznog primjerka *online* građe, najbolja su pravna osnova za uspostavu arhiva *weba* (ISO/TR 14873:2013). Nacionalnim knjižnicama legitimitet za arhiviranje sadržaja s *weba* daje status ustanove koja obradu obavlja kao zadaću od javnog interesa.

Zakonom se može regulirati pravo na samu uspostavu arhiva *weba*, pravo ustanove na prikupljanje i trajno čuvanje obveznog primjerka s *weba*, metode arhiviranja, uvjeti pristupa arhiviranim sadržajima, definirati obuhvat nacionalne domene, učestalost pobiranja, uključivanje ili isključivanje određenih vrsta sadržaja. Bez obzira na razinu razrađenosti nacionalnih zakona o postupanju s *online* obveznim primjerkom, svaka knjižnica trebala bi imati javno dostupnu politiku izgradnje zbirke i kriterije odabira sadržaja koji se arhiviraju.

Devedesetih godina 20. stoljeća kada se osnivaju prvi arhivi *weba*, pravni temelj kojim se regulira obvezni primjerak *online* publikacija u većini zemalja uopće nije postojao ili je davao samo općeniti okvir.

Kako bi ipak započele s prikupljanjem *online* obveznog primjerka, nacionalne knjižnice koje su poslovale u okruženju u kojem to nije bilo pravno regulirano započele su otvarati arhive *weba* vodeći pritom računa o ključnim pravnim pitanjima kao što su poštovanje intelektualnog vlasništva i autorskog prava.

Jedan od prvih arhiva *weba* otvorila je Nacionalna knjižnica Australije (NLA) 1996. U tada aktualnom zakonu o obveznom primjerku nije bilo riječi o *online* publikacijama. Da bi osigurala legitimost svojeg poslovanja, NLA je od svakog nositelja autorskog prava pojedinačno tražila odobrenje za arhiviranje što je

dugotrajan i spor proces, pretpostavlja vrlo selektivan pristup, a rezultira često izostankom vrijedne građe.

Neke se zemlje oslanjaju na dobrovoljnu dostavu obveznog primjerka (Nizozemska), a neke su, poput Francuske, ograničile pristup unutar ustanove što je jedna od mogućnosti kojom se koriste arhivi *weba* radi zaštite autorskih prava i prava privatnosti. Nedostatak restriktivnog pristupa arhiviranim sadržajima unutar ustanove jest slab doseg i korištenje arhiva.

Uporište za uspostavu sustava za arhiviranje *weba* Hrvatski arhiv *weba* imao je u Zakonu o knjižnicama iz 1997. Zakon ne regulira detaljnije obveze i prava Knjižnice te način pobiranja, pristup i čuvanje podataka. Navodi se tek da u opseg obveznog primjerka ulaze i *online* publikacije. Novi Zakon o knjižnicama i knjižničnoj djelatnosti (2019) nešto detaljnije regulira postupanje s obveznim primjekom *online* građe.

Svaki se element procesa arhiviranja dotiče pravnih pitanja; od prava na uspostavu arhiva *weba*, ako to nije zakonom regulirano, pitanja pobiranja (imaju li knjižnice uporište u zakonu ili traže dozvolu za pobiranje od svakog vlasnika *web*-sjedišta); što je pokriveno zakonom o obveznom primjerku (je li preopćenit ili je dovoljan); imaju li ustanove od javnog interesa koje posjeduju arhive *weba* pravo na javno objavljivanja arhiviranih sadržaja, čak i kada je pravo pobiranja pravno regulirano (javni interes).

Povijest i postupci arhiviranja *weba*

Svjesne važnosti čuvanja sadržaja objavljenih na *webu* i opasnosti njihova nepovratnog gubljenja, od devedestih godina 20. stoljeća sve veći broj nacionalnih knjižnica uspostavlja arhive *weba*. Arhivi *weba* i metode arhiviranja vrlo su različito koncipirani i razlikuju se s obzirom na nacionalni zakonski okvir, razvojne strategije ustanove, financijske, organizacijske, kadrovske i tehničke kapacitete, kriterije za odabir građe, tehnologije i alate kojima se koriste pri arhiviranju.

Unatoč tim razlikama proces arhiviranja *weba* sadrži sve poznate elemente upravljanja građom u knjižnicama: selekciju, prihvata, spremanje, trajnu zaštitu, korištenje i pristup građi.

U istraživanju Gomes, Miranda, Costa (2011) o inicijativama arhiviranja *weba* utvrđeno je da u pravilu postoji jedan arhiv *weba* na državnoj razini, da na tome rade mali timovi, a arhiviraju se sadržaji koji se odnose na određenu zemlju (na-

cionalna domena). Broj arhiva *weba* od 2010. u naglom je porastu.⁶ Cilj je istraživanja bio ustanoviti lokaciju arhiva, godinu nastanka, kriterije selekcije, formate koji se arhiviraju, broj osoblja koje radi na arhiviranju, veličinu arhiva, način pristupa i alate za pobiranje kojima se arhiv služi.⁷

U počecima uspostave prvih arhiva *weba* u središtu pozornosti bile su teme o unapređenju alata za pobiranje i spremanje *weba*, analizi arhiviranih podataka, uspostavi standarda za pohranu i trajno čuvanje podataka, uključivanju društvenih medija i društvenih mreža koji dolaze s *webom* 2.0 (Costa, Gomes, Silva, 2017), dok je pitanje zaštite osobnih podataka ostalo u drugom planu.

Selekcija

Selekcija je postupak u kojem se određuju kriteriji uključivanja i/ili isključivanja određenog sadržaja u opseg arhiviranja. Opseg arhiviranja može se odnositi na pojedinačnu *web*-stranicu, skup stranica okupljenih prema zajedničkoj temi ili na nacionalnu domenu. Selekcija se u pravilu odnosi na sadržaj, tehničke karakteristike publikacije, javnu dostupnost, jedinstvenost sadržaja na *webu*, učestalost promjena sadržaja, format objavljivanja.

U prvim su nacionalnim arhivima iz devedesetih godina politike odabira sadržaja za arhiviranje bile vrlo različite. Prema kriterijima Australске nacionalne knjižnice (*PANDORA Archive*) prednost su imale službene publikacije, zbornici radova s konferencija, svučilišni udžbenici, časopisi, službene stranice etničkih zajednica. Norveška nacionalna knjižnica opredijelila se za svakodnevno pobiranje *online* izdanja odabranih dnevnih novina. Francuska prikuplja sve što je objavljeno na francuskoj internetskoj domeni. Neki su arhivi davali prednost odabranim sadržajima bitnim s kulturnog i znanstvenog aspekta, ugledu nakladnika, *web*-stranicama događaja kao što su izborne kampanje i glazbeni festivali, tehnički inovativnim stranicama, zbornicima s konferencija, knjigama i časopisima, nepostojanju istog sadržaja na drugim medijima/formatima.⁸

Arhivi *weba* orijentirali su se na sadržaje koji su javno i slobodno dostupni i za koje nije potrebna prethodna registracija za pristup.

6 Hrvatski arhiv *weba* prve je publikacije (novine i časopise) arhivirao 2004.

7 Kao rezultat istraživanja objavljen je na Wikipediji popis *web*-arhiva (List of Web archiving initiatives) s namjerom da se stalno dopunjuje, https://en.wikipedia.org/wiki/List_of_Web_archiving_initiatives (24-09-2018). Broj članica IIPC konzorcija vidjeti na: <http://netpreserve.org/about-us/members/>.

8 Kriteriji selekcije arhiva *weba*: Webarchiv: český webový archiv <https://www.webarchiv.cz/en/source-selection> (14-09-2018); UK Web Archive: preserving UK websites <http://www.webarchive.org.uk/ukwa/> (14-09-2018); New Zealand Web Archive <https://natlib.govt.nz/collections/a-z/new-zealand-web-archiv> (14-09-2018); Eesti Veebiarhiiv <http://veebiarhiiv.digar.ee/> (14-09-2018); Hrvatski arhiv *weba* http://haw.nsk.hr/kriteriji_odabira (14-09-2018).

S „drugom generacijom“ *weba* poznatijom kao *web 2.0* došlo je do ubrzanog nastanka novih žanrova i neki su arhivi svoje kriterije selekcije proširili uključivanjem sadržaja s društvenih mreža.⁹ Razvojem i promjenama samog *weba* u sadržajnom i tehničkom smislu, mijenjali su se i zakonski propisi prema većoj zaštiti osobnih podataka, što je opet izazvalo zaokret u kriterijima za selekciju, tako da neke knjižnice sadržaje s društvenih mreža (Twitter, Facebook) isključuju iz odabira za arhiviranje (Bibliothèque et Archives nationales du Québec).

Pobiranje

Pobiranje (harvestiranje) proces je koji pomoću računalnog programa pregledava *web*-stranice, slijedi poveznice i kopira ih u svrhu arhiviranja. Opseg pobiranja može se kretati od pojedinačne *web*-stranice (selektivno pobiranje) do cijele nacionalne domene. Selektivno pobiranje uvijek podrazumijeva veću kvalitetu arhiviranih kopija (Pigac, Buzina, 2006).

Kod prvih arhiva *weba* načini pobiranja znatno su se razlikovali i knjižnice su u načelu primjenjivale samo jednu metodu – ili selektivno prikupljanje pojedinačne *web*-stranice ili pobiranje (harvestiranje) nacionalne domene. Pojam prostora nacionalne domene nije jednoznačno definiran. Katkad se određuje prema dvoslovnoj oznaci države (nacionalna vršna domena), a za neke je opseg nacionalne domene proširen i na domene općeg tipa kao „com“, „org“, „net“, „info“ (Brügger, Schroeder, 2017) uz minimalan uvjet da je sadržaj pisan službenim jezikom zemlje (primjeri su Danska i Hrvatska).

U prvim inicijativama za arhiviranje *weba* devedesetih godina (Australaska nacionalna knjižnica) selektivno se pobiranje provodilo tako da se od svakog vlasnika *web*-stranice tražio pristanak za pobiranje ili je vlasnik sam davao pristanak popunjavanjem obrasca za prijavu svoje publikacije za arhiviranje (HAW, UK Web Archive). Kod masovnog pobiranja *web*-stranica to nije provedivo. Prava vlasnika u tom se slučaju mogu ostvariti poštovanjem zaštite metodom oduzimanja (*opt out*) (ISO/TR 14873:2013). Ovaj pristup podrazumijeva implicitno dopuštenje za arhiviranje *weba*. Urednicima *web*-stranica šalje se poruka u kojoj se navodi da knjižnica želi arhivirati i učiniti stranicu javno dostupnom radi prikupljanja obveznog primjerka. Ako u zadanom roku ne stigne odgovor o odbijanju, to se smatra implicitnim ili prešutnim pristankom. Takvom se praksom koriste hrvatski i nozozemski arhivi *weba*.

9 Kongresna je knjižnica na temelju sporazuma s kompanijom Twitter 2010. arhivirala sve javne *tweetove* objavljene između 2006. i 2010.

Kod sveobuhvatnog pobiranja nacionalne domene važna je odluka hoće li se standard za isključivanje robota (*robot.txt.protocol*) poštovati ili ignorirati. Pridržavanjem standarda za isključivanje robota znatno se smanjuje kvaliteta rezultata pobiranja, pa ih neki arhivi zaobilaze kako bi arhivirane kopije u što većoj mjeri zadržale izvorni oblik (Celjak, Milinović, 2012; Keksitalo, 2010).

S obzirom na metodu pobiranja (selektivno ili sveobuhvatno), odbirali su se i alati za pobiranje.¹⁰ S promjenama pravnih propisa o obveznom primjerku, većim iskustvom arhiviranja *weba* i razvojem alata za pobiranje, pristupi arhiviranju sve su ujednačeniji i kombiniraju se metode selektivnog pobiranja pojedinačnih stranica, tematski okupljeni sadržaji i pobiranje nacionalne domene.

Radi legitimnosti poslovanja važno je da opseg i načini pobiranja budu definirani na razini zakona, ali i aktima ustanove koja provodi pobiranje.

Pohrana i trajno čuvanje

Primarna svrha arhiva *weba* jest pohrana i trajno čuvanje podataka u digitalnom obliku i održavanje integriteta izvornika (Celjak, Milinović, 2012).¹¹ Nedostatak standardiziranih deskriptivnih i tehničkih metapodatka i metapodataka za zaštitu najveći je izazov koji se odnosi na arhiviranje i trajnu pohranu *web*-stranica (Klarin, Holub, 2018).¹²

Kao jedna od najvećih prednosti arhiva *weba* naglašava se upravo funkcija trajnog čuvanja podataka i kada oni iz različitih razloga više ne postoje na izvornoj internetskoj adresi. Objavljivanjem na *webu* autori/vlasnici svjesno omogućuju javnu dostupnost podataka, istodobno imajući u vidu i mogućnosti njihova jednostavnog gašenja bez najave ili upozorenja. Međutim, nisu u dovoljnoj mjeri svjesni da jednom objavljen sadržaj može ostati negdje trajno sačuvan. Većina nacionalnih knjižnica internetske izvore prikuplja kao obvezni primjerak i po analogiji na tiskanu građu imaju obvezu čuvati ih trajno (Zakon o knjižnicama i knjižničnoj djelatnosti 2019; ISO/TR 14873:2013).

10 Npr. Archive-It, Heritrix, HTTrack, Memento, Netarchive Suite, SiteStory, Social Feed Manager, Wayback Machine, Web Archive Discovery, Web Curator Tool, Webrecorder. Za selektivno je pobiranje Hrvatskog arhiva *weba* Srce (Sveučilišni računski centar Sveučilišta u Zagrebu) razvilo vlastiti alat DAMP.

11 Potpunu cjelovitost izvornika teško je postići zbog tehničkih ograničenja na razini pobiranja; programi za pobiranje (roboti-pobirači) unaprijed ograničavaju dubinu pobiranja, maksimalan broj datoteka po poslužitelju, maksimalnu veličina datoteke.

12 Trajna zaštita arhiviranih sadržaja promatra se kroz širi kontekst digitalne knjižnice.

Pristup

Pristupom se definiraju uvjeti korištenja i prava pristupa *web*-arhivu. Uvjeti pristupa određuju se s obzirom na politiku privatnosti, metodu pobiranja i širi zakonski okvir.¹³ U načelu javni je pristup legitiman ako od nositelja prava/vlasnika stranica postoji dopuštenje ili kada je zakonskim propisima regulirano pobiranje nacionalne domene kao načina prikupljanja obveznog primjerka. U drugim je slučajevima vjerojatnije da će pristup arhivu biti na neki način ograničen; arhiv je dostupan samo osoblju knjižnice (norveški Norsk nettarkiv); arhiv je dostupan samo u istraživačke svrhe (BnF- Archives de l'Internet, danski Netarchive); arhiv je dostupan korisnicima samo u prostoru knjižnice (Deutsche Nationalbibliothek, Latvijski arhiv weba); korisnicima je javno dostupan samo dio arhiva (UK Web Archive, češki Webarchiv). U većini knjižnica dostupnost je potpunom sadržaju arhiva ograničena, čak i kada se od vlasnika stranica unaprijed dobilo dopuštenje za arhiviranje (nizozemski arhiv weba – KB Webarchieff). Kada je riječ o ograničenjima pristupa arhivima *weba* Gomes, Miranda i Costa naglašavaju interes korisnika:

Da bi arhivi weba bili korisni, pristup arhiviranim sadržajima trebao bi biti obveza. Ako je sadržaj javno dostupan na webu, trebao bi i dalje biti javno dostupan i kada postane povijesni sadržaj. Međutim, takva se politika sukobljava s nacionalnim zakonodavstvima koja ograničavaju pristup ili čak sprečavaju proaktivno arhiviranje weba (Gomes, Miranda, Costa, 2011).

Dostupnost sadržaja arhiva samo u prostoru knjižnice smanjuje njegovo korištenje. Arhivi *weba* dostupnost sadržaju reguliraju na način da ne dopuštaju pristup tražilicama, ne daju javni pristup arhivu ili nekim njegovim dijelovima, a na uspješno dokazivanje autora/nakladnika da njegovim interesima šteti pristup arhiviranoj kopiji, neke knjižnice pobiranje brišu iz arhiva (Keskitalo, 2010). S obzirom na načela prema kojima se arhivi i sve baštinske ustanove vode, brisanje sadržaja u suprotnosti je s profesionalom etikom i dovodi u pitanje vjerodostojnost njihova djelovanja.

13 Sadržaj arhiva *weba* nije dostupan internetskim tražilicama.

Arhiv *weba* i zaštita osobnih podataka

Iz današnjih rasprava u zajednici koja se bavi arhiviranjem *weba* vidljivo je da se težište s tehničkih pitanja o alatima za pobiranje i formatima za pohranu podataka premješta na pitanja pristupa arhiviranim sadržajima i na društvene, etičke i pravne implikacije koje arhivi *weba* donose. U uvodu programa Konferencije o arhiviranju *weba* održanoj u Nacionalnoj knjižnici Novog Zelanda, Koerbin i Hutař (2018) navode da arhiviranje *weba* nije samo stvar tehnologije, već se tiče etičkih i pravnih pitanja. Tema zaštite osobnih podataka sadržanih u arhivima *weba* posebno je aktualizirana Općom uredbom o zaštiti podataka (Uredba. 2016), odnosno s početkom njezine primjene 2018. koja je u cijelosti obvezujuća u svim državama članicama Europske unije. U uvodnom dijelu Uredbe navodi se:

Zbog brzog tehnološkog razvoja i globalizacije pojavili su se novi izazovi u zaštiti osobnih podataka. Opseg prikupljanja i razmjene osobnih podataka značajno se povećava. Tehnologijom se privatnim društvima i tijelima javne vlasti omogućuje uporaba osobnih podataka u dosada nedosegnutom opsegu radi ostvarenja njihovih djelatnosti. Pojedinci svoje osobne informacije sve više čine dostupnima javno i globalno. Tehnologija je preobrazila i gospodarstvo i društveni život te bi trebala dalje olakšavati slobodan protok osobnih podataka u Uniji i prijenos trećim zemljama i međunarodnim organizacijama, osiguravajući pri tome visoku razinu zaštite osobnih podataka (Uredba. Recital 6).

Prema Povelji Europske unije o temeljnim pravima (2007) svatko ima pravo na zaštitu osobnih podataka. Povreda osobnih podataka povreda je sigurnosti koja dovodi do slučajnog ili nezakonitog uništenja, gubitka, izmjene, neovlaštenog otkrivanja ili pristupa osobnim podacima koji su preneseni, pohranjeni ili na drugi način obrađivani (Uredba. Recital 12).

U odnosu na prethodni pravni okvir Uredba propisuje više i strože obveze za fizičke ili pravne osobe, tijela javne vlasti, agencije ili druga tijela koja sama ili s drugima određuju svrhe i sredstva obrade osobnih podataka.¹⁴

Poseban interes izazvale su točke Uredbe koje se odnose na pravo na brisa-

14 Osobni su podaci svi podaci koji se odnose na pojedinca čiji je identitet utvrđen ili se može utvrditi; to može biti ime, adresa, adresa e-pošte, IP adresa, telefonski broj itsl.

nje, pravo na ispravak podataka i pravo na ograničenje obrade. IFLA je objavila dokument o utjecaju Opće uredbe o zaštiti podataka na rad knjižnica (IFLA, 2018). U tom se području prepoznaje potencijalna kolizija s temeljnim etičkim načelima knjižničarske struke kao što je poštovanje načela intelektualne slobode, neometanog pristupa informacijama i slobode izražavanja (IFLA, 1999). Iz perspektive arhiva i knjižnica postavlja se pitanje ne ugrožavaju li se tim odredbama njihove profesionalne vrijednosti, poput odgovornosti za točnost podataka i odgovornosti za njihovo trajno čuvanje (Vavra, 2018).

U Izjavi o pravu na zaborav (IFLA Statement on the Right to be Forgotten, 2016) IFLA ne podupire trajno zatvaranje ili uništavanje zapisa koji sadrže osobne podatke, čak ni u ime privatnosti, poslovne povjerljivosti ili nacionalne sigurnosti.

Prema Uredbi (2016) „pravo na zaštitu osobnih podataka nije apsolutno pravo; mora ga se razmatrati u vezi s njegovom funkcijom u društvu te ga treba ujednačiti s drugim temeljnim pravima u skladu s načelom proporcionalnosti“ (Uredba. Recital 4).

Kao iznimku Uredba navodi da se pravo na zaštitu osobnih podataka ne primjenjuje radi ostvarivanja prava na slobodu izražavanja i informiranja, u svrhe arhiviranja u javnom interesu, u svrhu znanstvenog ili povijesnog istraživanja ili u statističke svrhe (Uredba. Čl. 17, stavak 3). U slučaju kada se osobni podaci skupljaju u navedene svrhe potrebno je primijeniti odgovarajuće tehničke i organizacijske mjere zaštite (Uredba. Čl. 89).

Primjena tehničkih mjera zaštite podataka, kao što su anonimizacija, pseudonimizacija i enkripcija ovise o arhitekturi programskog sustava za pobiranje i teško su provedive u arhivima weba. Kod planiranja novih sustava za arhiviranje ili razvoja postojećih svakako treba imati u vidu i način upravljanja osobnim podacima koje arhivi sadrže. U budućem razvoju informacijskih sustava knjižnica može se očekivati njihova bolja prilagođenost zahtjevima za zaštitu osobnih podataka.

Pitanje hoće li budućnost arhiva *weba* biti u sve većem ograničenju prava na pristup sadržaju i hoće li pristup biti ograničen samo za znanstvena istraživanja užem skupu korisnika ostaje otvoreno.

Zaključak

Od prvih arhiva *weba* nacionalne knjižnice bile su svjesne pitanja pravne utemeljenosti postupaka pri arhiviranju. Općom uredbom o zaštiti podataka naglašene su veće i strože obveze ustanova koje izvršavaju zadaće od javnog interesa, kao što su knjižnice, i imaju pravnu obvezu prikupljati, čuvati, uređivati, opisivati, komunicirati, promovirati, diseminirati i omogućiti pristup dokumentima od trajne vrijednosti za opći javni interes.

Knjižnice kao dio svoje djelatnosti moraju pratiti promjene zakonskih propisa i razmotriti njihove dugoročne implikacije na podatke koje prikupljaju i čuvaju te pronaći ravnotežu između ispunjavanja svojih profesionalnih obveza, kao što je pravo na pristup informacijama i trajno čuvanje podataka, i strože zakonske regulative u vezi sa zaštitom osobnih podataka. U *online* okruženju pravni propisi zadobivaju novu dimenziju i pitanje je uspijevaju li pravni instrumenti pratiti brz tehnološki razvoj i proces tranzicije u informacijsko društvo (Katulić, 2018).

LITERATURA

- BRÜGGER, N. i SCHROEDER, R. (2017). *The Web as History*, <https://doi.org/10.14324/111.9781911307563> (11-12-2018).
- BUZINA, T. i WILLER M. (2007). Croatian Digital Web Archive from project to service of the National and University Library in Zagreb. *DigItalia* 2 (2007), str. 20–34, <http://digitalia.sbn.it/article/view/437/277> (07-09-2018).
- CELJAK, D. i MILINOVIĆ M. (2012). Harvestiranje hrvatskog weba: arhitektura programskog sustava za harvestiranje i iskustva stečena njegovom uporabom. U: Hasenay, D. i Krtalić, M. (ur.). *15. Seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture: zbornik radova* (str. 144–160). Zagreb: Hrvatsko knjižničarsko društvo.
- CDNL Working Group (1996). *The legal deposit of electronic publications*, <http://unesdoc.unesco.org/images/0010/001055/105504e.pdf> (07-09-2018).
- Charter on the Preservation of Digital Heritage. (2003). U: *Resolution adopted on the report of Commission V at the 18th plenary meeting, on 15 October 2003*, http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=-

- DO_TOPIC&URL_SECTION=201.html (07-09-2018).
- The Commission of the European communities. (2006). *Commission recommendation of 24 August 2006 on the digitisation and online accessibility of cultural material and digital preservation*, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006H0585&from=ES> (07-09-2018).
- COSTA, M., GOMES D. i SILVA, M. (2017). The evolution of web archiving. *International Journal on Digital Libraries* 18 (3), 191–205. DOI 10.1007/s00799-016-0171-9.
- GOMES D., MIRANDA, J. i COSTA, M. (2011). A Survey on Web Archiving Initiatives. U: Gradmann, S., Borri, F., Meghini, C., Schuldt, H. (ur.) *Research and Advanced Technology for Digital Libraries*. (str. 408–420.). Berlin: Springer. DOI, https://doi.org/10.1007/978-3-642-24469-8_41.
- IFLA. (2018). *Briefing: Impact of the General Data Protection Regulation 2018*, https://www.ifla.org/files/assets/clm/publications/briefing_general_data_protection_regulation_2018.pdf (11-12-2018).
- IFLA Statement on Libraries and Intellectual Freedom. (1999), <https://www.ifla.org/publications/ifla-statement-on-libraries-and-intellectual-freedom> (11-12-2018).
- IFLA Statement on the Right to be Forgotten. (2016), <https://www.ifla.org/node/10272> (11-12-2018).
- ISBD: objedinjeno izdanje. (2011). Zagreb: Hrvatsko knjižničarsko društvo.
- ISO/TR 14873:2013 Information and documentation – Statistics and quality issues for web archiving. (2013). Geneva: ISO.
- KATULIĆ, T. (2018). Creative Commons ugovori: pravni temelj otvorene kulture. U: Hebrang, I. (ur.) *Otvorenost u znanosti i visokom obrazovanju*, str. 220–233. Zagreb: Školska knjiga.
- KESKITALO, E.-P. (2010). Web archiving in Finland, https://www.doria.fi/bit-stream/handle/10024/67051/webarchivingfinland_cndl.pdf?sequence=1 (11-12-2018).
- KLARIN, S. (2005). Predmet, motivi i metode arhiviranja sadržaja weba. U: Katić, T. (ur.) *8. seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture: zbornik radova* (str. 22–35). Zagreb: Hrvatsko knjižničarsko društvo.
- KLARIN ZADRAVEC, S. i HOLUB, K. (2018). Upravljanje digitalnim sadržajem: mogućnosti implementacije PREMIS-a u dokumentiranju postupaka zaštite. *Seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju*

- globalne informacijske infrastrukture*, 21, 247–267.
- KOERBIN, P. i HUTAŘ, J. (2018). IIPC General Assembly and Web Archiving Conference Programme, <http://netpreserve.org/ga2018/programme> (11-12-2018).
- NTOULAS, A., CHO, J. i OLSTON, C. (2004). What's new on the web?: the evolution of the web from a search engine perspective. U: *Proceedings of the 13th International Conference on World Wide Web*, str. 1–12, <http://www2004.org/proceedings/docs/1p1.pdf> (07-09-2018).
- PIGAC, S. i BUZINA, T. (2006). Selektivno arhiviranje hrvatskog weba: rezultati i otvorena pitanja. U: Willer, M., Zenić, I. (ur.) *9. seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture: zbornik radova* (str. 28–39). Zagreb: Hrvatsko knjižničarsko društvo.
- Povelja Europske unije o temeljnim pravima. (2007), <https://eur-lex.europa.eu/legal-content/HR/TXT/PDF/?uri=CELEX:12007P/TXT&from=EN> (11-12-2018).
- UNESCO's Director General international conference. (2012). *The Memory of the World in the Digital Age: Digitization and Preservation*, 26 to 28 September 2012, Vancouver, British Columbia, Canada, http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco_ubc_vancouver_declaration_en.pdf (07-09-2018).
- Uredba (EU) 2016/679 Europskog parlamenta i Vijeća od 27. travnja 2016. o zaštiti pojedinaca u vezi s obradom osobnih podataka i o slobodnom kretanju takvih podataka te o stavljanju izvan snage Direktive 95/46/EZ (Opća uredba o zaštiti podataka). (2016), <https://eur-lex.europa.eu/legal-content/HR/TXT/PDF/?uri=CELEX:32016R0679&from=HR> (07-09-2018).
- VAVRA, A. N. (2018). The Right to be Forgotten: an archival perspective. *The American Archivist*, 81 (1), 100–111.
- WILLER, M. i MILINOVIĆ, M. (2005). Prema trećoj generaciji knjižnično-informacijskih sustava: hibridna knjižnica za hibridne usluge. U: Katić, T. (ur.) *8. seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture zbornik radova* (str. 36–53). Zagreb: Hrvatsko knjižničarsko društvo.
- WILLER, M. (2007). Arhiviranje hrvatskih mrežnih publikacija: od projekta do programa arhiviranja u Nacionalnoj i sveučilišnoj knjižnici u Zagrebu (NSK). *Bosniaca*, 12 (12), 36–42.

WILLER M. i sur. (2008). Selective archiving of web resources: a study of processing costs. *Program*, 42 (4), 341–364. DOI, <https://doi.org/10.1108/00330330810912043>.

Zakon o knjižnicama i knjižničnoj djelatnosti. (2019), https://narodne-novine.nn.hr/clanci/sluzbeni/full/2019_02_17_356.html (07-09-2018).

WEB ARCHIVING

LEGAL AND ETHICAL ASPECTS

KEYWORDS:

national library, digital legal deposit, web archive, legal aspects in web archiving, General Data Protection Regulation

ABSTRACT

In the beginnings of web archiving, the activities of national libraries were focused on the development of principles and criteria for building up a collection of online publications, selecting a web crawler software and establishing a digital data preservation system. Although the issues regarding the relation between the right to the protection of personal data and the right of access to information and freedom of expression are not new in the library and archive community, the new legislation, particularly concerning the protection of personal data, has shifted the focus onto the ethical and legal aspects of web archiving as well as access to the archive.