

# I

ARHIVSKI ČLANCI //  
ARCHIVAL ARTICLES

## **O MOGUĆNOSTIMA I POTREBAMA UMJETNE INTELIGENCIJE U ARHIVIMA**

**Apstrakt:** *U ovom radu su predstavljene osnovne karakteristike informacijskih rješenja, koja označavamo s pojmom “umjetna inteligencija”. Mnoge arhivske ustanove već upotrebljavaju usluge te vrste. Spomenimo samo društvene mreže, sintetizatore govora, kompleksnu digitalizaciju itd. Mogućnosti neposredne upotrebe rješenja te vrste pokazala su se dobro i kod svakodnevnog arhivskog stručnog rada, pri čemu se pojavljuju različite olakšice. One stvaraju mnoga stručna, etička, pravna, tehnološka i druga pitanja, na koja treba brzo i osnovano odgovoriti.*

*Poseban izazov za arhivsku stručnu zajednicu predstavlja adaptacija postojeće arhivske doktrine prema novim okolnostima, koje su određene u kombinaciji umjetne inteligencije s velikim količinama podataka (Big Data) i lančanom tehnologijom (Block Chain). Nova rješenja već danas na razne načine utiču na cjelokupno arhivsko gradivo, koje je u skladištima profesionalnih arhiva ili će tek nastati kod stvaralaca u njihovim okruženjima i na novim tehnološkim osnovama. Spomenimo samo oblikovanje različitih vrsta podataka, vrednovanje sadržaja i okolnosti obezbjeđenje i očuvanje elektronskog gradiva, upotreba arhivskog gradiva itd.*

**Ključne riječi:** *Umjetna inteligencija, inteligentna rješenja, arhivsko gradivo, odgovarajuće društvo, arhivska doktrina.*

### **ABOUT THE POSSIBILITIES AND NEEDS OF USING ARTIFICIAL INTELLIGENCE IN ARCHIVES**

**Abstract:** *This paper presents the basic features of information solutions, which we are referring to as “artificial intelligence”. Many archival institutions already use such services. We shall mention only social networks, speech synthesizers, complex digitization, etc. Possibilities of direct in-depth use of such solutions are evident in daily archival professional work. In this proces different deviations also occur. These raise many professional, ethical, legal, technological and other issues. Archivists will need to answer on it promptly and reasonably.*

*A special challenge for the archival professional community is the adaptation of the existing archival doctrine to new circumstances. These we are defining in a combination of artificial intelligence with big data and block chain technology. Today, however, new solutions affect the entirety of archives, which archivists are keeping in the repositories of professional archives, or the creators will create them in new environments and on new platforms. Just to mention designing different types of metadata, evaluating content and environments, protection and preservation of electronic material, use of archival material, etc.*

**Key words:** *Artificial intelligence, intelligent solution, archival material, modern society, archival doctrine.*

## Uvod

U odgovarajućem društvu postoji mnogo rasprava, kao i suprotnih stavova o neposrednoj upotrebi umjetne inteligencije. Čini se, kao da informacijsko-tehnološka rješenja u tom segmentu već prevazilaze granice trenutno dozvoljenog, odnosno, prihvatljivog za pojedince, ili bar ostavljaju takav dojam. Kod njih se pojavljuje strah, koji je inače osjećaj potrebe za samoodržanjem, kao nekakva vrsta korekcije na tom području među stvarnim i mogućim, odnosno između teoretskog i praktičnog. Upravo zbog toga se u periodu široke implementacije rješenja, na osnovu umjetne inteligencije, mnogo intenzivnije nego u drugim razvojnim periodima informacijske tehnologije, pojavljuju mnogi moralno-etički, socijalno-ekonomski, vojno-sigurnosni, stručno-razvojni, kao i religiozni i drugi vidici. Oni koji se odnose na posebna pitanja upotrebe umjetne inteligencije, združuju ili razdružuju odgovarajuće informacijsko društvo.

Dostignuti stepen razvoja umjetne inteligencije možemo odrediti na razne načine. Najčešće je to na osnovu brojnih patenata, određenih prije svega s visinom uloženog kapitala za tu vrstu istraživanja. S tim metodama ne možemo objektivno odrediti njene praktične mogućnosti. Zato je teško precizno odrediti njen stvarni domet u različitim područjima, a s tim također i moguće oblasti njene stvarne upotrebe. Razloge za to treba tražiti, prije svega, u brzom razvoju cjelokupnog sektora umjetne inteligencije, pri čemu važnu ulogu očito ima najšire obezbjeđivanje konkurentskih, sigurnosnih i drugih prednosti pojedinačnih rješenja.<sup>1</sup>

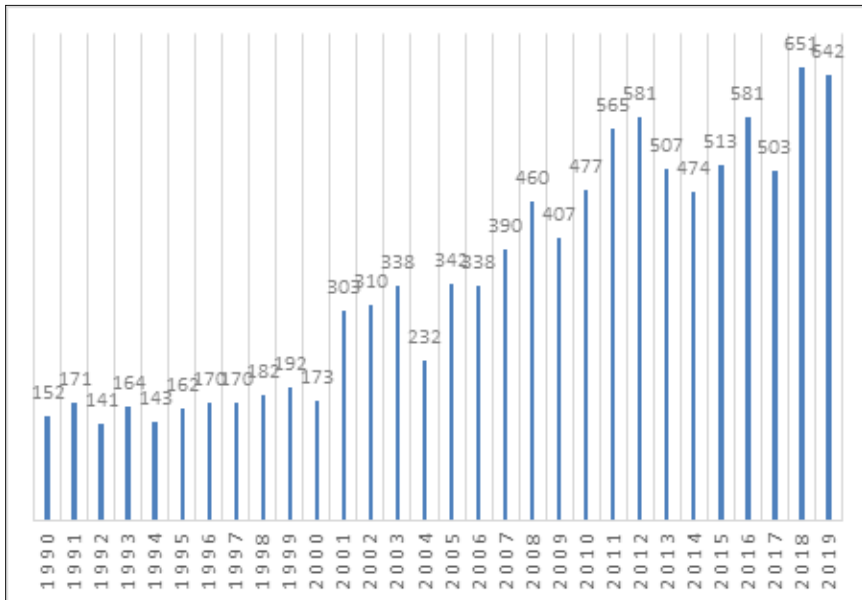
Poznato je, da pitanja u vezi sa obradom podataka sežu nekoliko stoljeća u prošlost. Često su povezana s izradom prve mehaničke računске mašine, koju je izradio francuski matematičar Blais Pascal sredinom 17. stoljeća.<sup>2</sup> Tehnološki napredak je u prvoj polovici 19. stoljeća donio prvi programirani stroj, koji su

1 *WIPO Technology Trends 2019 Artificial Intelligence*. Dobijeno 10.1.2020. s web strane [https://www.wipo.int/edocs/pubdocs/en/wipo\\_pub\\_1055.pdf](https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf).

2 *Pascaline*. Dobijeno 10. 1. 2020. s web strane <https://www.britannica.com/technology/Pascaline>.

napravili Charles Babbage i Ada Lovelace.<sup>3</sup> Slijedeća važna prekretnica, koja ide u smjeru implementacije umjetne inteligencije, bila je izvedena 1950. godine, kada je Alan Turing predstavio svoj test mašinske inteligencije.<sup>4</sup> Nakon toga su uslijedila još dva kraća intenzivna perioda razvoja iz oblasti razumijevanja i implementacije tehnologije.<sup>5</sup> U zadnjoj deceniji primijećen je intenzivan razvoj, što se ogleda u činjenici da univerziteti uvode studije sadržaja iz oblasti umjetne inteligencije, ustanovljavaju posebne razvojno-istraživačke grupe, laboratorije itd.

Stepen općeg interesa stručne i istraživačke zajednice za tu vrstu pitanja možemo odrediti kroz javno dostupne i objavljene, stručne i istraživačke radove. U bibliotekarskom informacionom sistemu COBISS+<sup>6</sup> dobivamo na osnovu sistema pretraživanja pojma “umjetna inteligencija” više od 11.000 odgovora. Pri tome možemo primjetiti, da se u periodu od 1990. godine do kraja 2019. godine pokazuje izuzetan porast stručnog interesa te vrste, a posebno od 2008. godine i dalje.



Prilog 1. Broj naslova s ključnom riječi “umjetna inteligencija” u sistemu COBBIS+ od 1990. do 2019. godine.

Pojam “umjetna inteligencija” kroz historiju je doživio evoluciju. U početku je to bio generički pojam za određenje kompleksnije mašinske (strojne) obrade

3 *Analytical Engine*. Dobijeno 10. 1. 2020. s web strane <https://www.britannica.com/technology/Analytical-Engine>.

4 *Turing Machines*. Dobijeno 10. 1. 2020. s web strane <https://plato.stanford.edu/entries/turing-machine/>.

5 *History of artificial intelligence*. Dobijeno 10. 1. 2020. s web strane [https://en.wikipedia.org/wiki/History\\_of\\_artificial\\_intelligence](https://en.wikipedia.org/wiki/History_of_artificial_intelligence).

6 *COBISS+ iskalnik – bralci*. Dobijeno 10. 1. 2020. s web strane <https://www.cobiss.si/>.

podataka. Danas je to kompleksan pojam, koji u svojoj definiciji obuhvata više obimnu podršku mašinske i programske opreme u različitim područjima ljudskog djelovanja.

Na sam razvoj umjetne inteligencije možemo gledati s više stanovišta, iz više uglova, pri čemu tražimo limitere, kao što su događaji, dostignuća ili karakteristike, koji su odredili njen razvoj. Drugi autori određuju razvoj umjetne inteligencije s načinom učenja različitih tehnoloških rješenja. Kao primjer navest ćemo petostepenu sistematizaciju razvoja inteligencije:

1. "Ponavljam" način učenja, koji se izvodi na osnovu naivnih algoritama – period traje od prvih začetaka izgradnje mašinske (strojne) obrade podataka pa do 1960. godine.
2. "Snimam", koji se izvodi na osnovu mašinskog (strojnog) učenja – period traje od 1960. godine do 2010. godine.
3. "Učim", način učenja, koji se izvodi na osnovu tzv. dubokog učenja – period traje od 2010. do 2018. godine.
4. "Učim se učiti", način učenja, koji se izvodi po osnovu tzv. duboko ojačanog učenja – taj period je određen do 2021. godine.
5. "Doprinosim i razmjenjujem", način učenja, koje će se izvoditi uz pomoć distribuiranih agenata, te masovnog duboko ojačanog učenja poslije 2021. godine.<sup>7</sup>

Sisteme, koji se zasnivaju na rješenjima umjetne inteligencije, možemo sistematizirati na slijedeće sisteme:

- s "uskom ali oštrom umjetnom inteligencijom"<sup>8</sup>
- s "opštom, ali jakom umjetnom inteligencijom"<sup>9</sup> i
- s "super umjetnom inteligencijom".<sup>10</sup>

Sva danas poznata rješenja spadaju u kategoriju uže umjetne inteligencije.<sup>11</sup> Pri tome važi saznanje, da se rješenja te vrste uvijek pojavljuju u odnosu prema nekom izazovu, zadatku, zahtjevu.

Ako je do nedavno važilo, da su različita inteligentna rješenja, kako u mašinskom (strojnom), tako i u programskom okruženju, određena kao "umjetna

7 *AI & academic history*. Dobijeno 15. 1. 2020. s web strane [https://atos.net/wp-content/uploads/2018/07/IA-infography\\_mini.pdf](https://atos.net/wp-content/uploads/2018/07/IA-infography_mini.pdf).

8 Pod pojmom "uska umjetna inteligencija" podrazumijevamo onu umjetnu inteligenciju koja je jednaka ili premašuje ljudske performanse određenih, visoko prilagođenih zadataka.

9 Pod pojmom "opšta umjetna inteligencija" podrazumijevamo onaj nivu umjetne inteligencije koji je jednak ljudskoj inteligenciji za bilo koji zadatak.

10 "Super umjetnu inteligenciju" definiramo kao onu umjetnu inteligenciju koja premašuje ljudsku inteligenciju za bilo koji zadatak.

11 Odbor za znanost in tehnologijo (STC), Pododbor za tehnološki razvoj in varnost (STCTTS), *Umetna inteligenca: posledice za oborožene sile nato*, Poročilo Matej Tonin (Slovenija) Poročevalec 088STC 19E, Izvor: engleski, 22. marec 2019. Dobijeno 15. 1. 2020. s web strane <https://imis.dz-rs.si/imis/1a16df40f0da165e8254.pdf>.

inteligencija”, pa je u zadnje vrijeme postalo jasno da je potrebno razlikovati pojam “inteligentna rješenja” od pojma “umjetna inteligencija”. Osnovna razlika između ova dva pojma je u tome, da su inteligentna rješenja rezultat čovjekovog uma i ona su sposobna da izvode vrlo zapletene logičke ili fizičke operacije, iako nisu sposobna donositi vlastite zaključke na osnovu dobivenih podataka. Na drugoj strani su rješenja, čija je osnovna karakteristika da se zasnivaju na umjetnoj inteligenciji, te da se rješenje uči i na osnovu vlastitih iskustava predlažu rješenja bez posredne ili neposredne čovjekove aktivnosti.<sup>12</sup>

Tehnologije i namjenski programi, koji se zasnivaju na umjetnoj inteligenciji, imaju već danas veliki uticaj. Trenutno najveći uticaj je poznat u:

- telekomunikacijama,
- transportu,
- biološkim i medicinskim naukama,
- ličnim napravama,
- kompjuterima i interakciji između kompjutera i čovjeka,
- bankarstvu,
- zabavnoj industriji,
- sigurnosti,
- industriji i proizvodnji,
- poljoprivredi,
- društvenim digitalnim mrežama itd.<sup>13</sup>

## Inteligentna rješenja u arhivskoj teoriji i praksi

O “inteligentnim”, odnosno “pametnim” rješenjima u arhivskoj teoriji i praksi ne nalazimo mnogo zapsanog. Razlog je očito u tome, da se ta terminološka rješenja još nisu ustalila u arhivskoj struci. Ipak se u arhivskoj stručnoj literaturi pojavljuju poveznice riječi, kao što su “kompleksna programska rješenja”, “racionalizacija postupka”, “paketna rješenja” i slično.

Spojmom “inteligentna/pametna rješenja” (engl. Smart Solution) određujemo sva rješenja, koja su zamišljena s optimiziranim postupcima i racionaliziranu upotrebu ljudskih i drugih izvora za postizanje željenog cilja na najširem području arhivske teorije i prakse.<sup>14</sup>

12 *Sindrom sjajnih novih igračaka ili kako umjetna inteligencija mijenja marketing.* (2018). Dobijeno 15. 1. 2020. s web strane <https://smk.si/sindrom-sijocih-novih-igrac-ali-kako-umetna-inteligencija-spreminja-marketing/>.

13 *WIPO Technology Trends 2019 Artificial Intelligence.* Dobijeno 10. 1. 2020. s web strane [https://www.wipo.int/edocs/pubdocs/en/wipo\\_pub\\_1055.pdf](https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf).

14 *What are Smart Solutions?* Dobijeno 27. 2. 2020. s web strane <https://apollogic.com/it-for-companies/internet-of-things-for-companies>.

Arhivisti u pravilu prihvataju i uvode inteligentna rješenja, mada ih tačno ne nazivaju (pojmovno). Pregled starije arhivske literature realno pokazuje, da je bilo na nekim područjima mnogo uvedenih različitih racionalizacija, odnosno inteligentnih rješenja, kako u arhivskim skladištima<sup>15</sup> tako i u oblasti uređivanja i popisivanja arhivskog gradiva<sup>16</sup> ili preuzimanja njegove upotrebe.

Mnogo inteligentnih rješenja se pojavilo s implementacijom različitih tehnoloških rješenja, koja su uticala na poboljšanje poslovanja pojedinačnih arhiva. Na području arhivskih informacionih sistema možemo među takva rješenja ubrojiti istraživanje sa upitima za punim tekstom, obuhvat koji će uzajamno povezati i obuhvatiti podatke, uspostavljanje zbirke podataka, koje se temelje na međusobnim relacijama i slično.

Inteligentna rješenja su bila u prošlosti tijesno povezana s profesionalnim odnosom, ličnim angažovanjem i na kraju sa zdravom pameću. Često su se uvažavala zbog nedostatka ljudskih ili finansijskih izvora. Upravo zbog toga možemo reći, da u svakoj arhivskoj ustanovi sigurno možemo naći inteligentna, racionalna, odnosno pametna rješenja, koja nisu odgovarajuća, ili uopšte nisu promovisana, tako da ih okruženje ne prepoznaje i ne prihvata kao takva.

### **Neka već upotrijebljena rješenja, koja se zasnivaju na umjetnoj inteligenciji**

U arhivskoj teoriji i praksi prepoznajemo sporadičnu upotrebu rješenja na osnovu umjetne inteligencije. Zato te pojave treba razmatrati objektivno i sa cjelokupnom stručnom odlučnošću. Konačan rezultat zasnivamo na spoznaji, da je odgovarajuća informacijsko-tehnološka podrška nastanku, upravljanju i upotrebi arhivskog i dokumentarnog gradiva, jako zaljuljala tvrde temelje tradicionalne arhivske struke. Zato da spomenemo samo pojavu dematerijalizacije odgovarajuće dokumentacije, poteškoće s obezbjeđivanjem univerzalne vjerodostojnosti arhivskog gradiva, umnoženog očuvanog sadržaja, itd.

Zbog kompleksnosti izazova, s kojima se susreće odgovarajuća arhivska struka, kao i zbog ograničenih čovjekovih i drugih izvora, arhivska struka se koncentriše prije svega na teškoće i izazove, koji su povezani sa arhivskim gradivom. Opšteg i masovnog uvođenja usluga ili rješenja na području kompleksnih procesa u arhivima, a koji bi se temeljili na umjetnoj inteligenciji, još nije moguće primjetiti. Rješenja na osnovu umjetne inteligencije se pojavljuju kao podrška promociji arhivske djelatnosti. Zato možemo zaključiti, da umjetnu inteligenciju implementiramo u arhivskim ustanovama vrlo neprimjetno, s relativno pojednostavljenim stručnim razmišljanjima. O njenim pozitivnim i mogućim negativnim posljedicama za razvoj

15 Spomenimo samo mrežne planove skladišnih kapaciteta, ventilacione sisteme skladišta, osiguravanje selektivne dostupnosti u skladišnim prostorijama, uvođenje bar koda, itd.

16 Spomenimo samo uzajamno zahvatanje podataka, umjetno stvaranje inventara arhivske građe, semantičku integraciju sadržaja, itd.

same arhivske struke u globalnim ili lokalnim kontekstima, još nemamo izrađenih rezultata.

Danas skoro svaka arhivska ustanova svoju djelatnost promovira preko društvenih mreža. U tom kontekstu da spomenemo samo implementaciju odgovarajućih servisa, kao što su Google, Facebook, Twitter, YouTube itd.<sup>17</sup> Preko njih arhivske ustanove posreduju podatke i kao posljedicu, također dobivaju bolje usluge na osnovu rješenja s područja umjetne inteligencije.

Značajan skok u toj oblasti nalazimo i u servisima široke primjenjivosti. Da spomenemo samo Google Translate<sup>18</sup> i s tim povezana rješenja, kao što su npr. razni sintetički zvučnici (vi samo slušate).<sup>19</sup> Oni otvaraju ne samo invalidima, nego, i drugim korisnicima široke mogućnosti pristupa arhivskim sadržajima. Istovremeno s tim utiču na način predstavljanja arhivskog gradiva, npr. na oblikovanje popisa arhivskog gradiva u arhivskim informacionim sistemima.<sup>20</sup>

Isto važi za strojno raspoznavanje sadržaja, odnosno informacionih nivoa na kompleksnim arhivskim dokumentima, kao i za njihovu interpretaciju u elektronskim okruženjima.<sup>21</sup> Ovdje možemo ubrojiti i rješenja vezana za raspoznavanje i prevođenje audio-video sadržaja, stavljanje u podnaslove iste vrste sadržaja na različitim jezicima i slično.<sup>22</sup>

Za zahtjevnije korisnike informacionih sistema se pojavljuje vrlo zanimljivo rješenje semantičnog povezivanja različitih zbirki podataka za potrebe oblikovanja kontekstnih meta podataka u okviru sistema Wolfram Alpha.<sup>23</sup> Pored toga moramo dodati i rješenja namijenjena dugoročnom čuvanju jako velikih količina podataka u datotekama, u elektronskim skladištima, koji se zasnivaju na rješenjima iz oblasti umjetne inteligencije.<sup>24</sup>

Neke implementacije umjetne inteligencije su ukazale na pojavu neočekivanih rezultata obrade u obliku odstupanja. Pokazuju se na različite načine.

17 Nina Gostenčnik, (2013). *Regional Archives Maribor in Web 2.0 and the overall situation in Slovenian archives*. Dobijeno 10. 2. 2020. s web strane <https://archive20.hypotheses.org/518>.

18 *Google Translate*. Dobijeno 10. 2. 2020. s spletne strani [https://en.wikipedia.org/wiki/Google\\_Translate](https://en.wikipedia.org/wiki/Google_Translate).

19 *eBralec bere, kar je zapisano. Vi samo poslušate*. Dobijeno 10. 2. 2020. s web strane <https://ebralec.si/>.

20 Mojca Kosi, Ocena dostopnosti digitaliziranega gradiva v podatkovni bazi Arhiva Republike Slovenije za invalide, *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Digitalno in digitalizirano. Arhivsko gradivo včeraj, danes in jutri*, Pokrajinski arhiv Maribor, Maribor 2017, 212-230. Dobijeno 10.2.2020. s web strane [http://www.pokarh-mb.si/uploaded/datoteke/Radenci/radenci\\_2017/16\\_kosi\\_2017.pdf](http://www.pokarh-mb.si/uploaded/datoteke/Radenci/radenci_2017/16_kosi_2017.pdf).

21 *High-Performance Digitisation. Project duration 2018-2020*. Dobijeno 10. 2. 2020. s web strane <https://www.csc.fi/en/-/high-performance-digitisation>.

22 *Kung-Hsiang, Huang (Steeve). (2018). Automatic Speech Recognition Data Collection with Youtube V3 API, Mask-RCNN and Google Vision API*. Dobijeno 21. 2. 2020 s web strane <https://towardsdatascience.com/automatic-speech-recognition-data-collection-with-youtube-v3-api-mask-rcnn-and-google-vision-api-2370d6776109>.

23 *What Is Wolfram|Alpha?* Dobijeno 10. 2. 2020 s web strane <https://www.wolframalpha.com/>.

24 *ARCHANGEL - Trusted Archives of Digital Public Records*. Dobijeno 21. 2. 2020. s web strane <https://www.archangel.ac.uk/>.



S arhivskog stručnog stanovišta, posebno su osjetljiva pravna i moralno-etička pitanja. Slijedeća je postavio Microsoftov projekat umjetne inteligencije Tay.AI bot s rasističkim homofobnim i drugim neprimjerenim tvitovima (Tweets). Istovremeno moramo istaći saznanja toga pokusa, koji je vezano s njegovom brzinom i obimom širenja diskutabilno moralno-etičkih sadržaja po cjelokupnoj svjetskoj mreži.<sup>25</sup>

Druga grupa arhivskih stručnih pitanja se odnosi na široku oblast, koja je poznata pod imenom “deep fake” ili duboke laži (prevare, neistine, nepravilnosti). Izvedene su uz pomoć programskih rješenja, zasnovanih na umjetnoj inteligenciji i izvode se na posebnoj mehaničkoj opremi.

Arhivska teorija i praksa se suočava s nepravilnostima još od srednjovjekovnih dokumenata. Nepravilnosti arhivisti dobro raspoznaju uz pomoć raznih metoda, kao na primjer historijska kritika izvora ili analiza sadržaja itd.<sup>26</sup> Značaj određenih nepravilnosti je u tome, da utiču ne samo na potencijalno arhivsko gradivo uskog spektra, kao što su fotografije ili video snimci i s tradicionalnim arhivskim pristupima nije ih moguće sigurno odrediti, nego imaju mnogo širi uticaj na odgovarajuće društvo, a s tim i na nastalo arhivsko gradivo u najširem značenju riječi.<sup>27</sup>

## O perspektivama upotrebe umjetne inteligencije u arhivima

Rezultat rasprave o mogućnostima upotrebe umjetne inteligencije u arhivima zasniva se na dokumentu s nazivom Etičke smjernice za umjetnu inteligenciju, vrijednu povjerenja, koji je napravila posebna radna grupa u okviru EU.<sup>28</sup> Arhivske stručne ambicije na tom području treba ograničiti još na realnu mogućnost implementacije tzv. uske ili slabe umjetne inteligencije. Pri tome moramo uvažavati poznata ograničenja, kao što su moguće teškoće s podacima i pristrasnost djelovanja sistema, pojavu crne kutije umjetne inteligencije i problem njene raspoloživosti.<sup>29</sup>

Teoretski se rješenja na osnovu umjetne inteligencije mogu implementirati prije svega za velike količine arhivskih sadržaja. Ti sadržaji su u pravilu u različitim

25 Hope Heese, *Why Microsoft's 'Tay' AI bot went wrong*, (2016). Dobijeno 20. 2. 2020. s web strane <https://www.techrepublic.com/article/why-microsofts-tay-ai-bot-went-wrong/>.

26 Hope Heese, *Why Microsoft's 'Tay' AI bot went wrong*, (2016). Dobijeno 20. 2. 2020. s web strane <https://www.techrepublic.com/article/why-microsofts-tay-ai-bot-went-wrong/>.

27 Paul Kari, *California makes 'deepfake' videos illegal, but law may be hard to enforce*, (2019). Dobijeno 20. 2. 2020. s web strane <https://www.theguardian.com/us-news/2019/oct/07/california-makes-deepfake-videos-illegal-but-law-may-be-hard-to-enforce>.

28 Radna grupa osnovana je sredinom 2018. godine i u osnovi se sastojala od 52 stručnjaka iz područja razvoja i upotrebe umjetne inteligencije. Službeno se grupa naziva Stručna grupa za umjetnu inteligenciju na visokom nivou (High-Level Expert Group on Artificial Intelligence). O njenom radu detaljnije na web strani *High-Level Expert Group on Artificial Intelligence*. Dobijeno 27. 2. 2020. s web strane <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>.

29 *Etične smernice za zaupanja vredno umetno inteligenco*. Dobijeno 27. 2. 2020. s web strane <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>.

međusobnim relacijama, koje su ovisne od vremena i prostora, te su predstavljene kao zamršene strukture podataka.<sup>30</sup> Za to su potrebni prilagođeni programski i drugi alati, koji su poznati pod imenom “topic modeling“.<sup>31</sup>

### *Vrednovanje stvaralaca i nadzor nad registrom stvaralaca*

Pitanja vezana za vrednovanje i vođenje odgovarajućih registara stvaralaca arhivskog gradiva, zahtijevaju od javne arhivske službe znatne napore. Problem je u tome, što se u praksi kod stvaralaca događaju brze promjene vlasništva i druge statusne izmjene, koje utiču i dinamički mijenjaju obaveze prema nadležnoj arhivskoj ustanovi. S tim se dinamično mijenjaju i obaveze i postupci nadležnih arhivskih ustanova u vezi ovlaštenja i djelovanja vanjske službe, te načina kako će se izvesti preuzimanje arhivskog gradiva od stvaralaca.

Uporedo s tim pojavljuju se pitanja mogućnosti posudbe registra stvaralaca javnog arhivskog gradiva. Dinamično mijenjanje ovlaštenih lica za upravljanje arhivskim gradivom, a time i njihovih statusa, predstavlja dosta obimne zadatke iz oblasti održavanja prilagođenih podataka. Te podatke bi mogli u perspektivi održavati sa sistemima i rješenjima, koji bi se zasnivali na algoritmima umjetne inteligencije.

### *Vrednovanje i preuzimanje velikih količina arhivskog gradiva*

U oblasti vrednovanja arhivskog gradiva, u arhivima se suočavaju s jako velikim količinama dokumentacije, koja je u fizičkom obliku. Obim dokumentacije, koja je nastala i koja se čuva u elektronskoj formi kod stvaralaca, nije moguće odgovarajuće odrediti. Problem je toliko veći jer se kod elektronskih oblika radi o različito (ne)uređenom arhivskom gradivu, koje je često i djelimično ili u cijelosti nestrukturiranih oblika.

Arhivski stručni izazovi na području upravljanja s tom dokumentacijom sežu od formiranja i vrednovanja formata, pa sve do uspješnog čitanja i razumijevanja arhivskih sadržaja i njihovih konteksta. Prilikom obrade gradiva te vrste, pojavljuju se i različite druge aktivnosti, kao što su: deduplikacija arhivskih sadržaja, provjera njihove cjelovitosti i vjerodostojnosti, izračunavanje i provjeravanje kontrolnih suma<sup>32</sup>, na kraju i izrada odgovarajućeg prijemnog informacionog paketa (SIP) u skladu s važećim odredbama, koji se može preuzeti u arhivskoj ustanovi i ubaciti u

30 Hrvoje Stančić, Computational Archival Science, *Moderna arhivistika*, 2018, št. 2, 323-330. Maribor: Pokrajinski arhiv Maribor. Dobijeno 27. 2. 2020. s web strane [http://www.pokarh-mb.si/uploaded/datoteke/2\\_2018\\_323-330\\_stancic.pdf](http://www.pokarh-mb.si/uploaded/datoteke/2_2018_323-330_stancic.pdf).

31 *Introduction to Topic Modeling*. Dobijeno 27. 2. 2020. s web strane <https://monkeylearn.com/blog/introduction-to-topic-modeling/>.

32 Više o ovim postupcima u: Tatjana Hajtnik, Aida Škoro Babić, Ali nam lahko pri vrednotenju in odbiranju elektronskega gradiva pomaga tehnologija?, *Moderna arhivistika*, 2018, št. 1, 169-196. Maribor: Pokrajinski arhiv Maribor. Dobijeno 27. 2. 2020. s spletne strani [http://www.pokarh-mb.si/uploaded/datoteke/radenci\\_20181/1\\_2018\\_169-196\\_%C5%A0koro.pdf](http://www.pokarh-mb.si/uploaded/datoteke/radenci_20181/1_2018_169-196_%C5%A0koro.pdf).

sistem dugoročnog čuvanja arhivskog gradiva. Kada sumiramo, kod tih postupaka se radi o zapetljanim tehničko-tehnološkim rješenjima, uključujući i postupke ovladavanja sadržajem. Radi obima tih sadržaja, treba tražiti primjereno tolikom obimu, odgovarajuća rješenja iz oblasti umjetne inteligencije.

### *Digitalizacija i pretvaranje oblika u kojima se pojavljuje arhivsko gradivo*

Masovna digitalizacija arhivskog gradiva i s tim strojno raspoznavanje tih sadržaja se u laičkoj javnosti doima, kao da je to već dugo riješen problem. Ipak je za precizno prepoznavanje arhivskog gradiva potrebno obezbijediti odgovarajuće kvalitetno izvedene postupke digitalizacije. Pri tome se arhivski stručni radnici susreću s problemima obezbjeđivanja zahtjevnih znanja spretnosti savladavanja velikih količina podataka, ne samo na nivou datoteka, nego i na nivou sadržaja, te na nivou svih slojeva, svih vrsta i pojavnih oblika arhivskog gradiva. Posebne izazove na tom području čine dokumenti, koji postavljaju zahtjeve za raspoznavanjem međusobno djelimično, ili u cjelosti sakrivenih znakova i zapleta u različitim bojama. Kompleksnija rješenja na osnovu umjetne inteligencije bi u tom segmentu povećala produktivnost postupka digitalizacije i njihovu precizniju kontekstualizaciju. Na taj način bi povećali i opšti stepen dostupnosti cjelini arhivskog gradiva.

U tom kontekstu upotreba umjetne inteligencije možemo dodati i raspoznavanje i odgovarajuće pretvaranje različitih formata zapisa u takve formate, koji obezbjeđuju dugoročnu čitljivost, upotrebljivost i povjerenje u očuvane arhivske sadržaje.

### *Čuvanje elektronskih i fizičkih oblika arhivskog gradiva*

Na području odgovarajućeg čuvanja arhivskog gradiva po sadašnjim naznakama umjetna inteligencija se može implementirati, kako u oblasti očuvanja elektronskih, tako i fizičkih oblika arhivskog gradiva.

Kod očuvanja elektronskih formi arhivskog gradiva, radi se o obezbjeđenju dugoročne konzistencije i garantovanju javnog povjerenja i dostupnosti elektronskim sadržajima na osnovu robotizacije upravljanja medijima, ili upravljanju na primjer konzistentnosti velikog broja datoteka u elektronskim skladištima.

Odgovor na pitanje, da li prilikom garantovanja očuvanja fizičkog arhivskog gradiva, ima smisla razmišljati o robotizaciji arhivskih skladišta, nije tako jednostavan. Taj odgovor treba tražiti u širokom spektru socijalnih, ekonomskih, kao i tehnoloških uticaja u pojedinačnom arhivu. Neupitno je da bi kompleksnija robotizacija arhivskih skladišta mogla pripomoći pri brznoj i preciznijoj manipulaciji tehničkih jedinica u arhivskim skladištima, istovremeno bi pripomogla i lakšem otkrivanju grešaka kod nosilaca ili zapisa arhivskog gradiva, te cjelovitijem nadzoru mikroklimе u arhivskim prostorijama.

Za arhivska skladišta fizičkog arhivskog gradiva, mogu biti zanimljiva rješenja širokog spektra od različitih robotskih ruku i sistema horizontalnog i verikalnog transporta arhivskog gradiva, pa sve do humanoidnih robota, koji su već razvijeni za obavljanje nekih tipično ljudskih poslova i djelatnosti, prije svega na području zdravstvene zaštite.

### *Popisivanje i kontekstualizacija arhivskih sadržaja*

Popisivanje arhivskog gradiva i postavljanje tih opisa u šire sadržajne, historijske, geografske i druge kontekste postaje sa stručnog aspekta sve zahtjevnije i kompleksnije. To nam govori i o poređenju zahtjeva standarda, gdje se broj predviđenih relacija među entitetima popisa jako povećava.<sup>33</sup> Istovremeno praksa pokazuje, da opise arhivskog gradiva treba dobro opremiti s tzv. kontekstnim meta podacima. Tako opstaju zapisi u bazama, koji imaju sto i više meta podatkovnih poveznica.<sup>34</sup>

	<i>ISAD(g)</i>	<i>ISAAR (CPF)</i>	<i>ISDIAH</i>	<i>ISDF</i>	<i>RiC</i>
<i>Broj gornjih entiteta</i>	1	3	1	1	14
<i>Broj sadržanih entiteta</i>	26	31	35	26	67
<i>Broj predviđenih relacija</i>	Nije određeno	4 niza	Nije određeno	2 niza	792

Prilog 2. *Broj entiteta i njihovih relacija, koje su predviđene u arhivskim stručnim standardima.*

Ako na sve to dodamo još zahtjeve za očuvanje npr. ličnih i drugih zakonom određenih podataka, možemo ustanoviti, da se u arhivskim zbirkama podataka pojavljuje zanimljiv fenomen, da popisi na nižim nivoima zahtijevaju više pažnje i implementacije, anonimizacije i nadzora nad dostupnošću podacima te vrste, dok na višim nivoima dolazi do njihove izrazite kompleksne kontekstualizacije. Upravo te dvije arhivske stručne djelatnosti predstavljaju veliku potrebu ljudskih izvora, što bi se s implementacijom rješenja na osnovu umjetne inteligencije, moglo sadržajno bitno racionalizovati.

33 Miroslav Novak, Teorijski i praktični aspekti upravljanja relacijama u savremenoj arhivskoj teoriji i praksi, *Arhivski vjesnik*, 62/2019, Hrvatski državni arhiv, Zagreb 2019, 47-60.

34 Kao primjer spomenimo popis arhivskog fonda Okružne komisije za ratne štete Gornja Radgona, koji ima 144 veze do lokalnih deskriptora. Dobijeno 27. 2. 2020. s spletne strani <http://www.siranet.si/detail.aspx?ID=42726>.

## *Upotreba i interpretacija arhivskih sadržaja*

Upotreba arhivskog gradiva je uvijek određena s relacijom korisnika i arhivskog informativnog pomagala, tj. nadležne osobe u arhivskoj čitaonici. Postojeća arhivska informativna pomagala, osim rijetkih izuzetaka još nisu standardno pripremljena za alternativne oblike upotrebe kao na primjer za potrebe invalida.<sup>35</sup> Još uvijek su ograničeni prije svega na pisani pojavni oblik. Na drugoj strani je razvoj sintetiziranog govora tako brz, da je pitanje vremena, kada ćemo standardno glasovno komunicirati s arhivskim informacionim sistemima. U tom slučaju se pojavljuju novi zahtjevi komunikacije. Jedna od vrsta komunikacije jeste, da krajnji korisnik mora znati s kim komunicira, jer je moguće da je na drugoj strani čovjek ili stroj. Isto tako mora znati, da li upotrebljava arhivske sadržaje s javnim povjerenjem (vjerodostojne) ili su to interpretacije sadržaja, umjetničko predstavljanje i slično, koje su isto tako važne i sačuvane, kao i opšte dostupne. Ti zahtjevi predstavljaju ozbiljan izazov za buduće arhivske sisteme umjetne inteligencije, kako za cjelokupnu arhivsku struku, tako i izvedena stručna rješenja predstavljanja arhivskog gradiva javnosti.

### **Zaključak**

Uvođenje rješenja, koja su zasnovana na slaboj ili uskoj umjetnoj inteligenciji na području arhivske teorije i prakse, otvara potpuno nove i druge dileme. One se ne odnose samo na uži stručni rad u arhivu, nego i na arhivsko gradivo koje će tek nastati pri stvaraočima u okruženjima te vrste. U tom kontekstu su najaktualnije moralno-etičke dileme. Pojavljuju se i pravni, organizacijski, socijalni, sigurnosni i drugi vidici te upotrebe. Sigurno će se pojaviti i takve dileme, koje danas još ne možemo ili ne znamo odrediti.

Arhivska stručna javnost se suočava s činjenicom, da je dosadašnje uvođenje informacijske tehnologije na području poslovanja s dokumentacijom jako promijenilo arhivsku doktrinu. Osnovna arhivska stručna načela treba interpretirati u kontekstima nematerijalnih oblika odgovarajuće dokumentacije, u prenošenju konteksta realnog svijeta u digitalna okruženja i slično.

Sa uvođenjem rješenja na osnovu umjetne inteligencije možemo očekivati neke drastične promjene u arhivskom stručnom radu. Umjesto konkretnog arhivskog stručnog rada s arhivskim gradivom bit će potrebno razviti metode i postupke provjere i garancije javnog povjerenja u arhivske sadržaje. Trebat će odrediti i garantovati vjerodostojne podatke za učenje sistema.

35 Tatjana Hajtnik, Mojca Kosi, Robert Hrovat Merič, Prilagoditve različnih tipov arhivskih gradiv za dostopnost ranljivim skupinam, *Moderna arhivistika*, 2018, št. 1, 83-93. Maribor: Pokrajinski arhiv Maribor. Dobiljeno 27. 2. 2020. s web strane [http://www.pokarh-mb.si/uploaded/datoteke/radenci\\_20181/1\\_2018\\_083-094\\_kosi.pdf](http://www.pokarh-mb.si/uploaded/datoteke/radenci_20181/1_2018_083-094_kosi.pdf).

Poseban izazov se pojavljuje u garantovanju kontrole nad izmjenama programskih kodova, koji će izvoditi djelatnosti na području upravljanja s dokumentacijom. Ti zahtjevi će biti jednako važni kod stvaralaca, kao i kod arhivskih ustanova.

Predviđena implementacija slabe umjetne inteligencije u arhivske stručne ustanove pokazuje, da ona podržava cjelokupni životni ciklus dokumenta. Isto tako se može implementirati na sve položaje modela otvorenog arhivskog informacionog sistema.

Predviđena arhivska struka se mora baviti s teškoćama i izazovima upotrebe slabijih oblika umjetne inteligencije prije svega zbog:

- uspostavljenih postupaka digitalizacije, koji sa sadašnjim metodama rada ne daju očekivane rezultate na području brze i precizne obrade podataka, te s tim povezane opšte dostupnosti digitalizatora,
- velikih količina obuhvaćenih i ostvarenih elektronskih zapisa, koji međusobno nisu dovoljno semantički povezani, da bi u budućnosti mogle ostati informacije.
- kompleksnosti pretraživačkih postupaka i mehanizama u arhivskim informacionim sistemima, koji spominju neposrednu upotrebu arhivskih sadržaja.

## Summary

The implementation of solutions based on already weak or narrow artificial intelligence in the archival theory and practice opens completely new archival professional and other dilemmas. These not only relate to narrower professional work in the archive, but also to the archival material that creators produce in these environments. In this context, moral and ethical considerations exist. There are also legal, organizational, social, security and other aspects of this use. We can assume that there will be some other aspects, we cannot yet identify.

The archival professional community knows that the introduction of information technology for supporting the whole documentation life cycle has dramatically changed traditional archival doctrine. Archivists must interpret the basic archival professional principles in the contexts of intangible forms of contemporary documentation, or they must map of real-world contexts into digital environments etc.

However, with the introduction of artificial intelligence solutions, we can expect some drastic changes in archival professional work. Instead of concrete archival professional work with archival material, it will be necessary to develop methods and procedures for verifying and ensuring the public faith of archival content. It will be necessary to identify and provide credible data for system learning.

A particular challenge is to provide control over the modification of the program code, which will carry out activities in the field of documentation management. This will be equally important for creators as well as archival institutions.

The implementation of weak artificial intelligence into archival professional institutions shows us that this can support the entire life cycle of the document. We can also implement it in to all positions of the model OAIS (open archive information system).

The contemporary archival profession must address the problems and challenges of using a weak form of artificial intelligence due to:

- Used digitization processes cannot produce the expected results in the field of rapid and accurate processing digitalized documents and the related general availability of digitalized archival material with current working methods.
- Large amounts of captured and created electronic records are semantically not interconnected enough to become and stay stable information in the future.
- The complexity of search procedures and mechanisms in archival information systems that limit the direct use of archival content.