



01 U P O R A B N A INFORMATIKA

2024 ◀ ŠTEVILKA 1 ◀ LETNIK XXXII ◀ ISSN 1318-1882

U P O R A B N A I N F O R M A T I K A

2024 ŠTEVILKA 1 JAN/FEB/MAR LETNIK XXXII ISSN 1318-1882

Znanstveni prispevki

Sandi Gec, Vlado Stankovski

Analiza nefunkcionalnih zahtev na primeru uporabe priporočilnega sistema pametnih pogodb

3

Lana Masnec, Marina Trkman

Analiza uporabe aplikacije za sledenje stikov med mladimi: študija primera Nemčije

13

Strokovni prispevki

Klara Žnideršič, Matija Marolt, Aleš Veršič, Matevž Pesek

Metodologije za kvalitativno vrednotenje kakovosti odprtih podatkov

20

Tjaša Šoltes, Jan Vasiljevič, Marko Bajec

Online notes: sistem za razpoznavo govora in strojno prevajanje v realnem času na ravni univerzitetnih predavanj

29

Informacije

Iz Islovarja

39

Ustanovitelj in izdajatelj

Slovensko društvo INFORMATIKA
Litostrojska cesta 54, 1000 Ljubljana

Predstavniki

Slavko Žitnik

Odgovorni urednik

Mirjana Kljajić Borštnar

Uredniški odbor

Andrej Kovačič, Anton Manfreda, Evelin Krnac, Jan Mendling, Jan von Knop, John Taylor, Lili Nemeč Zlatolas, Marko Hölbl, Miodrag Popović, Mirjana Kljajić Borštnar, Mirko Vintar, Pedro Simões Coelho, Saša Divjak, Sjaak Brinkkemper, Tatjana Welzer Družovec, Timotej Knez, Vesna Bosilj-Vukšić, Vida Groznik, Vladislav Rajkovič

Recenzentski odbor

Aleksander Sadikov, Alenka Baggia, Alenka Brezavšček, Aljaž Košmerlj, Andrej Brodnik, Andrej Kovačič, Andreja Pucihar, Anton Manfreda, Benjamin Urh, Blaž Rodič, Borut Batagelj, Borut Werber, Boštjan Šumak, Božidar Potočnik, Branko Kavšek, Branko Šter, Ciril Bohak, Damjan Fujs, Damjan Strnad, David Jelenc, Dejan Lavbič, Denis Trček, Domen Mongus, Drago Bokal, Eva Jereb, Evelin Krnac, Inna Novalija, Irena Nančovska Šerbec, Ivan Gerlič, Jernej Vičič, Jure Žabkar, Jurij Mihelič, Lovro Šubelj, Luka Pavlič, Luka Tomat, Maja Pušnik, Marina Trkman, Marjeta Marolt, Marko Bajec, Marko Hölbl, Marko Robnik Šikonja, Martin Šavc, Martina Šestak, Matej Klemen, Matjaž Divjak, Mirjam Sepesy Maučec, Mirjana Kljajić Borštnar, Mladen Borovič, Muhamed Turkanovič, Niko Schlamberger, Nikola Ljubešič, Patricio Bulić, Polona Rus, Robert Leskovar, Samed Bajrić, Sandi Gec, Saša Divjak, Slavko Žitnik, Tatjana Welzer Družovec, Tomaž Hovelja, Uroš Rajkovič, Vida Groznik, Vladislav Rajkovič, Živa Rant

Tehnični urednik

Timotej Knez

Lektoriranje angleških izvlečkov

Marvelingua (angl.)

Oblikovanje

KOFEIN DIZAJN, d. o. o.

Prelom in tisk

Boex DTP, d. o. o., Ljubljana

Naklada

110 izvodov

Naslov uredništva

Slovensko društvo INFORMATIKA
Uredništvo revije Uporabna informatika
Litostrojska cesta 54, 1000 Ljubljana
www.uporabna-informatika.si

Revija izhaja četrtletno. Cena posamezne številke je 20,00 EUR. Letna naročnina za podjetja 85,00 EUR, za vsak nadaljnji izvod 60,00 EUR, za posameznike 35,00 EUR, za študente in seniorje 15,00 EUR. V ceno je vključen DDV.

Revija Uporabna informatika je od številke 4/III vključena v mednarodno bazo INSPEC.

Revija Uporabna informatika je pod zaporedno številko 666 vpisana v razvid medijev, ki ga vodi Ministrstvo za kulturo RS.

Revija Uporabna informatika je vključena v Digitalno knjižnico Slovenije (dLib.si).

Izid publikacije je finančno podprla Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije.

© Slovensko društvo INFORMATIKA

Vabilo avtorjem

V reviji Uporabna informatika objavljamo kakovostne izvirne prispevke domačih in tujih avtorjev z najširšega področja informatike, ki se nanašajo tako na poslovanju podjetij, javno upravo, družbo in posameznika. Prispevki so lahko znanstvene, strokovne ali informativne narave, še posebno spodbujamo objavo interdisciplinarnih prispevkov. Zato vabimo avtorje, da prispevke, ki ustrezajo omenjenim usmeritvam, pošljejo uredništvu revije po elektronski pošti na naslov ui@društvo-informatika.si.

Avtorje prosimo, da pri pripravi prispevka upoštevajo navodila, ki so objavljena na naslovu <http://www.uporabna-informatika.si>.

Za kakovost prispevkov skrbi mednarodni uredniški odbor. Prispevki so anonimno recenzirani, o objavi pa na podlagi recenzij samostojno odloča uredniški odbor. Recenzenti lahko zahtevajo, da avtorji besedilo spremenijo v skladu s priporočili in da popravljeni prispevek ponovno prejmejo v pregled. Sprejeti prispevki so pred izidom revije objavljeni na spletni strani revije (predobjava), še prej pa končno verzijo prispevka avtorji dobijo v pregled in potrditev. Uredništvo lahko še pred recenzijo zavrne objavo prispevka, če njegova vsebina ne ustreza vsebinski usmeritvi revije ali če prispevek ne ustreza kriterijem za objavo v reviji.

Pred objavo prispevka mora avtor podpisati izjavo o avtorstvu, s katero potrjuje originalnost prispevka in dovoljuje prenos materialnih avtorskih pravic. Avtorji prejmejo enoletno naročnino na revijo Uporabna informatika, ki vključuje avtorski izvod revije in še nadaljnje tri zaporedne številke. S svojim prispevkom v reviji Uporabna informatika boste pomagali k širjenju znanja na področju informatike. Želimo si čim več prispevkov z raznoliko in zanimivo tematiko in se jih že vnaprej veselimo

Uredništvo revije

Navodila avtorjem člankov

Članke objavljamo praviloma v slovenščini, članke tujih avtorjev pa v angleščini. Besedilo naj bo jezikovno skrbno pripravljeno. Priporočamo zmernost pri uporabi tujk in, kjer je mogoče, njihovo zamenjavo s slovenskimi izrazi. V pomoč pri iskanju slovenskih ustreznih priporočamo uporabo spletnega terminološkega slovarja Slovenskega društva Informatika, Islovar (www.islovar.org).

Znanstveni prispevek naj obsega največ 40.000 znakov, kratki znanstveni prispevek do 10.000 znakov, strokovni članki do 30.000 znakov, obvestila in poročila pa do 8.000 znakov.

Prispevek naj bo predložen v urejevalniku besedil Word (*.doc ali *.docx) v enojnem razmaku, brez posebnih znakov ali poudarjenih črk. Za ločilom na koncu stavka napravite samo en presledek, pri odstavih ne uporabljajte zamika.

Naslovu prispevka naj sledi polno ime vsakega avtorja, ustanova, v kateri je zaposlen, naslov in elektronski naslov. Sledi naj povzetek v slovenščini v obsegu 8 do 10 vrstic in seznam od 5 do 8 ključnih besed, ki najbolje opredeljujejo vsebinski okvir prispevka. Sledi naj prevod naslova povzetka in ključnih besed v angleškem jeziku. V primeru, da oddajate prispevek v angleškem jeziku, velja obratno. Razdelki naj bodo naslovljeni in oštevilčeni z arabskimi številkami.

Slike in tabele vključite v besedilo. Opremite jih z naslovom in oštevilčite z arabskimi številkami. Na vsako sliko in tabelo se morate v besedilu prispevka sklicevati in jo pojasniti. Če v prispevku uporabljate slike ali tabele drugih avtorjev, navedite vir pod sliko oz. tabelo. Revijo tiskamo v črno-beli tehniki, zato barvne slike ali fotografije kot original niso primerne. Slikam zaslonov se v prispevku izogibajte, razen če so nujno potrebne za razumevanje besedila. Slike, grafikoni, organizacijske sheme ipd. naj imajo belo podlago. Enačbe oštevilčite v oklepajih desno od enačbe.

V besedilu se sklicujte na navedeno literaturo skladno s pravili sistema IEEE navajanja bibliografskih referenc, v besedilu to pomeni zaporedna številka navajenega vira v oglatem oklepaju (npr. [1]). Na koncu prispevka navedite samo v prispevku uporabljeno literaturo in vire v enotnem seznamu, urejeno po zaporedni številki vira, prav tako v skladu s pravili IEEE. Več o sistemu IEEE, katerega uporabo omogoča tudi urejevalnik besedil Word 2007, najdete na strani https://owl.purdue.edu/owl/research_and_citation/ieee_style/ieee_general_format.html.

Prispevku dodajte kratek življenjepis vsakega avtorja v obsegu do 8 vrstic, v katerem poudarite predvsem strokovne dosežke.

Analiza nefunkcionalnih zahtev na primeru uporabe priporočilnega sistema pametnih pogodb

Sandi Gec, Vlado Stankovski

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Večna pot 113, 1000 Ljubljana

sandi.gec@fri.uni-lj.si, vlado.stankovski@fri.uni-lj.si

Izveček

Nefunkcionalne zahteve so opredeljena kot ena izmed ključnih meril kakovosti programske opreme, ki vplivajo na zmogljivost, uporabnost, varnost in zanesljivost. Tehnologija veriženja blokov razširja nabor običajnih nefunkcionalnih zahtev z dodatnimi atributi, kot so preglednost, nespremenljivost in decentralizacija. Integracija tehnologije veriženja blokov običajno osredotoča na funkcionalne zahteve, pri čemer nefunkcionalne pogosto niso dovolj naslovljene. Zato se nefunkcionalne zahteve običajno analizirajo z namenski vmesniki uporabniškega programa, spletnimi pajki in drugimi orodji, ki zahtevajo nenehno vzdrževanje ter prilagajanje delovanja tovrstnih orodij. V tem delu predlagamo analizo nefunkcionalnih zahtev za rešitve tehnologije veriženja blokov, ki temelji na umetni inteligenci, kjer so uporabljena orodja, kot so Copilot, ChatGPT in druga. Takšna orodja uporabljamo za identifikacijo, vrednotenje in optimizacijo nefunkcionalnih zahtev za rešitve verig blokov na bolj celovit način. Rezultati naše raziskave so predstavljeni na primeru uporabe priporočljivega sistema, ki vključuje diskusijo o tem, kako izboljšati dolgoročno vzdržnost predlagane rešitve.

Ključne besede: umetna inteligenca, trilema, Ethereum virtualni stroj, nefunkcionalne zahteve

Analysis of non-functional requirements for smart contract recommender system

Abstract

Non-functional requirements (NFRs) are essential criteria of software quality that affect the performance, usability, security, and reliability of software. The blockchain, as a novel technology expands the set of common NFRs with additional attributes, such as transparency, immutability, and decentralization. However, blockchain integration usually focuses on functional requirements (FRs), while NFRs are not sufficiently addressed. Therefore, NFRs are usually analysed with dedicated APIs, crawlers, and other tools that require ongoing maintenance and manual intervention. This paper proposes an AI-based analysis of NFRs for blockchain solutions, where AI tools such as Copilot, ChatGPT, and others are leveraged. We use these tools to generate, evaluate, and optimize NFRs for blockchain solutions in a more comprehensive way. The results of our solution are presented on a recommender system use case, including explanation on how to improve the sustainability of the proposed solution.

Keywords: Artificial Intelligence, trilemma, Ethereum Virtual Machine, Non-Functional Requirements

1 UVOD

Na področju programskega inženiringa so nefunkcionalne zahteve ključnega pomena, še posebej pri oblikovanju programske opreme, analizi programske opreme in splošnem razumevanju programske opreme. Ker programska oprema običajno ni monolitna,

centralizirana ali homogena, temveč heterogena, sestavljena iz več komponent in v nekaterih primerih decentralizirana (npr. decentralizirane baze podatkov), je pozornost na nefunkcionalnih zahtevah bistvena. To velja še posebej na področju tehnologije

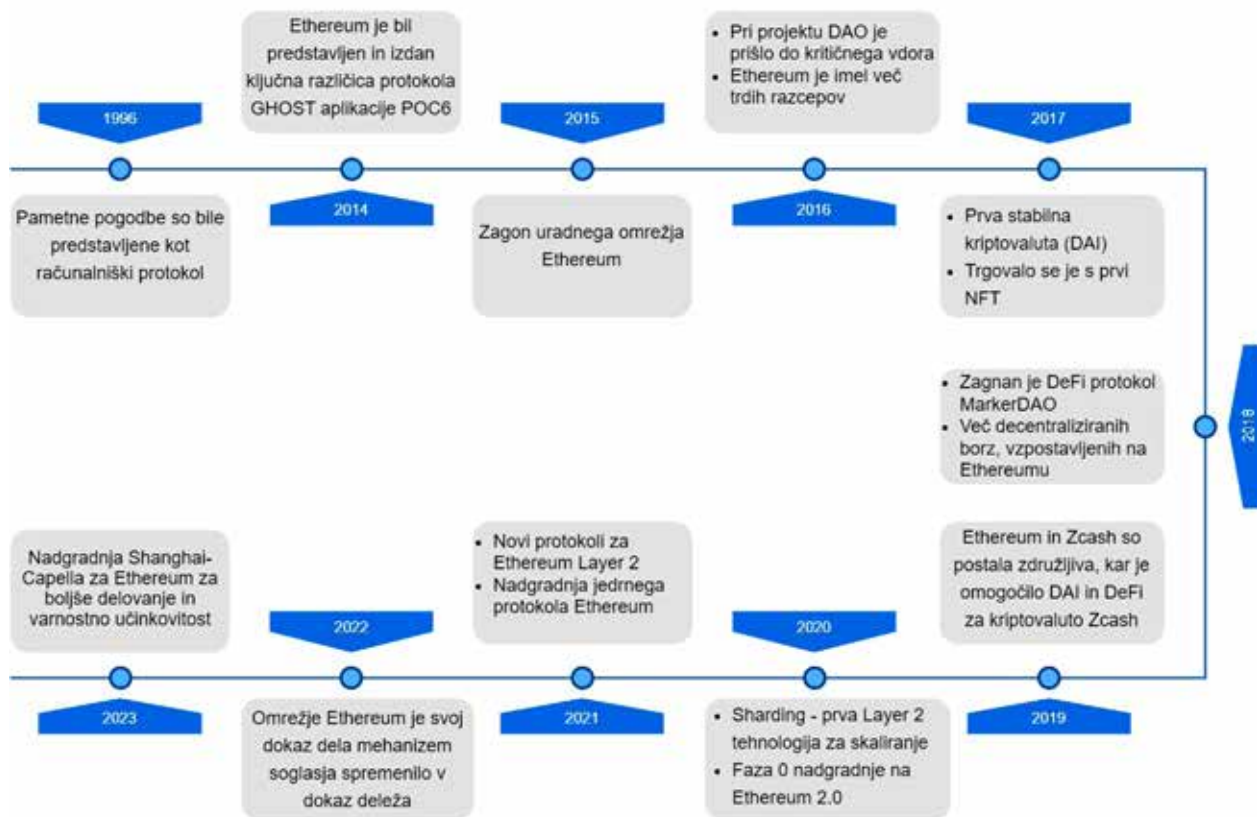
verženja blokov, kjer je kompleksnost programske opreme pogosto zelo visoka.

Pri razvijajočih se tehnologijah distribuiranih razpršenih knjig (angl. Distributed Ledger Technology, DLT) je pomembno temeljito razumevanje nefunkcionalnih zahtev za izboljšanje procesa izbire razpršene knjige. Naša osredotočenost je na DLT, ki temelji na Ethereumu, in vključujejo podporo za pametne pogodbe

Solidity prek komponente Ethereum virtualnega stroja (angl. Ethereum Virtual Machine, EVM), ki deluje v primerih vozlišč razpršene knjige. Čeprav primarna razpršena knjiga Ethereum mainnet morda ni primerna za vse primere uporabe, ki zahtevajo določene nefunkcionalne zahteve, je ključna za razvoj funkcionalnosti pametnih pogodb, saj se knjiga nenehno ter skrbno razvija [1], kot prikazuje Slika 1.

Zato postaja preučevanje razpoložljive razpršene knjige, ki omogočajo EVM, z vidika nefunkcionalnih zahtev. Čeprav se funkcionalne zahteve med različnimi platformami pogosto ujemajo, je razlika v ne-

funkcionalnih zahtevah znatna. V tej raziskavi bomo predstavili celovito analizo razpršenih knjig, ki temeljijo na EVM. Najprej se bomo lotili najbolj standardne analize nefunkcionalnih zahtev, imenovane trojček razpršene knjige, in preučili značilnosti ter razlike med vsemi razpoložljivimi plastmi EVM, vključno s Plastjo-0, ki ni EVM. Ker ni standardov za opredelitev zahtev pri EVM, bomo raziskali najbolj izstopajoče nefunkcionalne zahteve z uporabo storitev umetne inteligence. Razčlenili bomo razlike med storitvami umetne inteligence in analizirali rezultate umetne inteligence, kot so predlagane nefunkcionalne zahteve. Nadalje bomo predstavili integracijo nefunkcionalnih zahtev v primeru uporabe priporočilnega sistema za pametne pogodbe in opisali delovni tok nefunkcionalnih zahtev ter morebitne koristi v okviru primera uporabe. V eksperimentalni študiji bomo analizirali storitve umetne inteligence v kontekstu nefunkcionalnih zahtev, ki omogočajo EVM. Razčlenili bomo najbolj prevladujoče nefunkcionalne zahteve v kontekstu predstavljenega primera uporabe.



Slika 1: Časovni potek razvoja pametnih pogodb od teoretičnih konceptov [2] do zagona prve verige blokov s podporo pametnih pogodb Ethereum v letu 2015 ter vse nadaljnje pomembne mejnike. V letu 2024 je odmevna nadgradnja Ethereum Dencun, ki zmanjša stroške transakcij na verigah Plast-2.

Preostanek članka bo sledil naslednjim korakom. Razdelek 2 predstavi aktualne raziskovalne dosežke in identificira vrzel, ki jo obravnava ta študija. Razdelek 3 predstavi osnovne koncepte študije, vključno z Ethereum plasti in storitvami umetne inteligence. Razdelek 4 pojasni integracijo nefunkcionalnih zahtev v motivacijskem primeru uporabe. Razdelek 5 predstavi eksperimentalno študijo. Razdelek 6 razpravlja o izvedenih eksperimentih in podaja zaključke.

2 SORODNA DELA

Znanstvene raziskave nefunkcionalnih zahtev programskih rešitev so aktualne že desetletja, pri čemer se je v zgodnjih raziskavah definirala osnovna taksonomija z izhodiščnimi kategorijami, kot so: (i) funkcionalnosti obnašanja, (ii) performančne lastnosti (npr. časovne, prostorske, hitrostne in druge), (iii) kvalitativne lastnosti (npr. uporabnost, varnost, zanesljivost in druge) in (iv) druge lastnosti (npr. pravni vidik, okoljski in drugi) [3]. Temeljne ugotovitve raziskave leta 2007 so obsegale programske rešitve predvsem monolitnih ter centraliziranih aplikacij in niso vključevale računalništva v oblaku, ki je bil v takrat v povojih. Pomemben primer decentralizirane aplikacije (angl. decentralised Applications, dApps) je bil predstavljen leta 2009 z zagonom prve javne rešitve, ki temelji na tehnologiji veriženja blokov s strani neznanega avtorja s psevdonimom Satoshi Nakamoto [4]. Zanimanje za tehnologijo veriženja blokov je privedlo do raziskav ter razvoja, ki je privedla do izdaje javnih kriptovalut z različnimi nameni. Določene verige blokov so bile predlagane z namenom izboljšav verige blokov Bitcoin, druge so se pa osredotočale na nefunkcionalne zahteve, kot je zasebnost v korelaciji z anonimnostjo ter različnih raziskovalnih vej [5]. Precej raziskovalnega zanimanja na področju tehnologije veriženja blokov je bilo naklonjenega primeru uporabe samostojna suverena identiteta (angl. Self-Sovereign Identity, SSI) [6], ki je še dandanes izjemno aktualna tudi s širše uporabe (npr. na nivoju Evropske unije). Pri pregledu aktualnih raziskovalnih publikacij so avtorji identificirali ter opredelili 22 nefunkcionalnih zahtev aktualnih na področju SSI.

Zanimiv koncept je bil predstavljen z zagonom verige blokov Ethereum [7], kjer so avtorji predstavili možnost decentraliziranih skript imenovane pametne pogodbe. Z vidika nefunkcionalnih zahtev so M. Staderini idr. [8] predstavili problematiko izbire ustrezne verige blokov, kjer so predlagali metodolo-

gijo izbire na podlagi namenskih diagramov odločanja. Posebno pozornost so avtorji namenili identifikaciji nefunkcionalnih zahtev, pri čemer so se osredotočili na njihovo stopnjo pomembnosti v različnih tipih verig blokov: javne, zasebne, stopnje dovoljenj uporabnikov in odločanje po konceptu konzorcija. Raziskovalne ugotovitve avtorjev nudijo izhodišče za nadaljnjo analizo nefunkcionalnih zahtev, ki so obravnavane v našem delu na javnih verigah blokov. Analizo nefunkcionalnih zahtev praktičnih produkcijskih primerov uporabe podprte s tehnologijo veriženja blokov je predstavil M. Kassab [9] in sicer tako, da je najprej opredelil identificirane nefunkcionalne zahteve ter nato ocenil njihovo pomembnost glede na krovne funkcionalnosti v primerih uporabe. Na podlagi 1327 zahtev zastopanih v 7 različnih produkcijskih primerih uporabe je avtor identificiral ter opredelil 6 osnovnih razredov nefunkcionalnih zahtev: (i) zasebnost, (ii) skalabilnost/performančnost, (iii) interoperabilnost, (iv) uporabnost, (v) skladnost z zakonodajo (npr. regulatorji) in (vi) Operativne in finančne omejitve (npr. zadostno število vozlišč, velikost blokov idr.). V našem delu se poleg omenjenih zahtev osredotočamo predvsem na tiste, ki se pogosto pojavljajo kot implicitne (npr. v arhitekturni opredelitvi plasti EVM verig blokov) in eksplicitne (npr. kvantitativne) lastnosti.

Pregled najpomembnejših pristopov inženiringa programskih zahtev na področju tehnologij veriženja blokov so bili predstavljeni s strani M. S. Farooq idr. [10] pri čemer je bil pregled del med leti 2015 in 2021 kritično analiziran ter izpostavljene omejitve posameznih pristopov na izbranih domenah. Pomembno delo predstavlja tudi pregledni znanstveni prispevek, ki izpostavlja raziskovalne izzive, kot so integracija zahtev, identifikacija metrik in predstavitev zahtev v sistemih s podporo tehnologije veriženja blokov [11]. V primerjavi z našim delom, kjer sta najbolj zastopani domeni oblachnega računalništva ter tehnologije veriženja blokov s podporo pametnim pogodbam, v predhodnih raziskavah tovrstna domena ni bila zadostno obravnavana.

V našem delu se osredotočamo na verige blokov, ki omogočajo EVM in posledično omogočajo izvajanje Solidity pametnih pogodb. Zanimiv pristop z uporabo odločitvenih modelov za najpogostejše funkcionalnosti (npr. avtentikacija, avtorizacija, komunikacija izven verige idr.) s katerimi so pri modeliranju modelov upošteevane tudi nefunkcionalne

zahteve. Čeprav je največja omejitev predlaganih modelov rigidnost (npr. monetizacijski modeli so le minimalen del kapacitet, ki jih omogočajo pametne pogodbe), le-ti predstavljajo dobro izhodišče za razumevanje zahtev v vsakdanjih primerih uporabe. Podobno je predstavljena možnost izbire verige blokov v odvisnosti od različnih vidikov kot je možnost migracije na druge verige blokov ali najpogosteje funkcionalnosti, ki jih veriga blokov omogoča [12]. Avtorji so na teoretičnem modelu igre predlagali matematično analizo s katere so ugotovili, da uporabniki EVM verig blokov niso osredotočeni zgolj na natančno eno verigo blokov. Izbira ustrezne EVM verige blokov glede na primer uporabe pametnih pogodb je ključnega pomena, da lahko zagotavljamo pričakovano kakovost storitve ter kakovost izkušnje.

Omenjeni pristopi so nas motivirali, da smo v našem preteklem delu predlagali analizo EVM verig blokov za primere uporabe internet stvari z nefunkcionalnega vidika, bolj natančno smo se osredotočili na hitrost transakcije in ceno transakcij [13]. V analizi smo s pomočjo vmesnikov uporabniškega programa (angl. Application Programming Interface, API) in s pomočjo namenskih spletnih pajkov ter ostalih orodij predlagali primerjalno analizo med obetavnimi EVM verigami blokov.

3 NEFUNKCIONALNE ZAHTEVE ZA VERIGE BLOKOV S PODPORO EVM

V tem poglavju je najprej opredeljen pomen večplastnega razvoja, ki temelji na težavah osnovnega Ethereum omrežja. Omrežje je kot prvo izjemno decentra-

lizirano ter posledično varno, po drugi strani pa so transakcije počasnejše ter drage, kar v splošnem povzamemo kot manj skalabilno. Za boljše razumevanje nefunkcionalnih zahtev primerjamo različne storitve umetne inteligence, ki nam praviloma bolj podrobno opredelijo nefunkcionalne zahteve verig blokov.

3.1 Analiza plasti

Z razvojem tehnologije veriženja blokov, ki je bila prvotno pretežno decentralizirana, so se pojavile težave pri možnosti podpore različnim primerom uporabe. Pri višji stopnji decentralizacije je posledično skalabilnost nižja. Obenem so aplikacije z nižjo stopnjo decentralizacije manj varne zaradi večje izpostavljenosti vozlišč ter možnosti raznih napadov (npr. 51% napad). Problematiko kompromisa omenjenih nefunkcionalnih zahtev (varnost, skalabilnost in decentralizacija) je opredelil ustanovitelj Ethereum verige blokov Vitalik, in sicer kot trilemo (angl. trilemma) tehnologije veriženja blokov. Glede na različne namene uporabe Ethereum verige blokov ter podpor različnih primerov uporabe, so se razvijalci odločili razvit več plasten pristop, kjer so trilema lastnosti različno zastopani. Iz prvotne plasti Ethereum omrežja (Plast-1) se je razvilo skupno pet plasti, kjer so glavne lastnosti ponazorjene v Tabeli 1.

3.2 Pristopi umetne inteligence za analizo nefunkcionalnih zahtev

Javne storitve umetne inteligence so postale pomembna orodja, ki nam omogočajo učenje na podlagi interakcije v obliki klepeta s sistemom. Na

Tabela 1: Osnovne plasti, ki opredeljujejo EVM verige blokov.

Ime plasti	Varnostni mehanizem	Odvisnost	Stopnja skalabilnosti	Lastnosti	Primeri verig
Plast-0	Lastni (deli paraverigam)	Neodvisen	Omejena	Deluje kot zanašajoča se veriga, s poudarkom na varnosti in interoperabilnosti med glavnimi paraverigami	Polkadot [14]
Plast-1	Lastni	Neodvisen	Omejena	Zajema podatkovne, omrežne, konsenzne in aktivacijske podplasti v logični arhitekturi	Bitcoin, Ethereum
Plast-2	Se zanašajo na Plast-1	Zgrajen na temeljih Plasti-1	Višja	Neodvisen ali vzporedno, si prizadevajo za reševanje izzivov transakcijske hitrosti in razširljivosti	Polygon, Optimism
Plast-3	Se zanašajo na Plast-2	Zgrajen na temeljih Plasti-2	Višja	Aplikacijska plast, gosti dApps, interoperabilna, razdeljena na dve podplasti: aplikacijska in izvajalna	Uniswap, Axie Infinity
Stranska veriga (angl. Sidechain)	Lastni	Neodvisen	Višja	Povezuje z nadrejeno verigo preko dvosmernega mostu	Cardano Milkmeda V2

tak način sistem bolje razume kontekst iskanja kot običajni iskalniki, čeprav marsikateri brskalniki beležijo zgodovino našega iskanja in nam, zato lahko vračajo bolj relevantne rezultate. Storitve umetne inteligence je mogoče uporabljati v vsakdanjem življenju, kot orodje za izobraževanje, reševanje nalog in drugo [15]. V Tabeli 2 je predstavljen nabor primerov aktualnih javnih storitev umetne inteligence, pri čemer so, med drugimi, opredeljene tudi lastnosti, ki so pomembne v kontekstu integracije. Pri pedagoškem področju postajajo takšna orodja problematična, saj so storitve zmožne reševanja marsikaterega problema, tudi programerskih. Kljub temu je treba rezultate kritično preveriti, saj se mnogokrat zgodi, da rešitev v obliki programske kode ne prevede, narobe deluje ali nam podaja napačen rezultat. V našem delu uporabljamo storitve umetne inteligence kot del priporočilnega sistema, predvsem kot dopolnilo obstoječih informacij, predvsem kvantitativnih, o EVM verigah blokov. V našem delu se osredotočamo zgolj na brezplačne različice storitev umetne inteligence.

4 INTEGRACIJA ZAHTEV V OBLAČNE REŠITVE

V tem poglavju najprej predstavimo osnovne značilnosti priporočilnega sistema pametnih pogodb. Nato predstavimo potek integracije nefunkcionalnih zahtev kot del sistema.

4.1 Primer uporabe priporočilnega sistema pametnih pogodb

Priljubljenost oblačnih ter sorodnih arhitektur, kot so v megli (angl. Fog) in na robu (angl. Edge) je v nenehnem vzponu zaradi možnosti dodeljevanja namenskih nalog na različne nivoje ter možnost integracije različnih protokolov komunikacije. V tovrstnih arhitekturah so lahko vozlišča del podatkovnih centrov, pri čemer del infrastrukture lahko zagotavljajo končni uporabniki [16]. Postopek razbremenitve vključuje končne uporabnike, ki zagotavljajo storitve ali vire v infrastrukturi megle, kar običajno zahteva ročne operacije, kot so dogovor o ceni in politiki, registracija nove komponente zagotavljanja in drugo. Te ročne operacije se pogosto izvajajo prek centraliziranih komponent. Za poenostavitev takšnih procesov je mogoče ročne operacije opredeliti v namenskih pametnih pogodbah, ki upoštevajo sistemske zahteve in potrebe deležnikov. Razvoj tako zapletenih pametnih pogodb je izziv in zahteva izkušenega razvijalca pametnih pogodb. Zato smo predlagali priporočilni sistem, ki kot rezultat vrača predloge pametnih pogodb, ki predstavljajo preverjeno varno programsko kodo ter so namenjene razširitvam za podporo primerom uporabe. Prednost takšnega priporočilnega sistema je povzeta kot izboljšanje procesa integracije pametnih pogodb s predlogi pametnih pogodb, ki se vračajo na podlagi oblačne arhitekture sistema ter funkcionalne zahtev, ki jih izbere razvijalec [17].

Tabela 2: Primeri aktualnih javnih storitev umetne inteligence

Ime storitve	Kategorija	Primarni namen	API podpora	Možnosti integracije
ChatGPT	Klepetalnik z umetno inteligenco	Pogovor in pomoč uporabnikom	Da	Da
Microsoft Copilot	Klepetalnik z umetno inteligenco, asistent izvorne kode	Pomoč uporabnikom in razvijalcem programske opreme	Da	Da
Google Bard	Klepetalnik z umetno inteligenco	Pogovor in pomoč uporabnikom	Da	Da
Craiyon AI	AI generator slik	Ustvarjanje slik	Da	Da
HeyWire.ai	B2B SaaS	Novinarstvo in ustvarjanje korporativnih vsebin	Ne	Ne
Eightify	Izločanje glavnih točk iz dolgih videoposnetkov	Razširitev brskalnika	Ne	Ne
Feathery	Pomoč razvijalcem programske opreme	Izdelava prilagojenih obrazcev in delovnih tokov brez kodiranja	Da	Da
Azure AI Vision	Umetna inteligenca kot storitev (angl. Artificial Intelligence as a Service, AlaaS)	Računalniški vid	Da	Da
Cloud Vision API	AlaaS	Računalniški vid	Da	Da
Amazon Rekognition Image	AlaaS	Prepoznavanje slik in globoko učenje	Da	Da

Osnovni scenarij uporabe priporočilnega sistema obravnava pogoste izzive, s katerimi se soočajo mlajši razvijalci, vključno s pomanjkanjem izkušenj pri razvoju pametnih pogodb in omejenim znanjem o potencialnih predlogah pametnih pogodb za večkratno uporabo, oblčnih arhitektur ali oboje. Kot odgovor predstavljenim izzivom smo razvili nov priporočilni sistem pametnih pogodb, zasnovan za doseganje naslednjih ciljev:

- ustvarjanje predlog pametnih pogodb, usklajenih s specifikacijami aplikacije (funkcionalne zahteve), ob upoštevanju najboljših praks in vzorcev oblikovanja;
- podpreti implementacijo verige blokov tako, da razvijalcem interneta stvari ponudi zanesljive, standardizirane pametne pogodbe, prilagojene njihovim primerom uporabe;
- pospešiti splošni razvoj in izvajanje pametnih pogodb v oblčnih okoljih.

Pomemben prispevek takšnega sistema se ne odraža zgolj s splošnega vidika programskega inženiringa za namene razvoja programske opreme, ampak tudi kot pedagoško orodje za izboljšanje razumevanja možnosti razvoja pametnih pogodb. To poteka skozi izbiro različnih funkcionalnih zahtev na osnovni oblčne tri-nivojske arhitekture spletne aplikacije, kjer študentom poenostavi razvoj pametnih pogodb na projektnem delu razvoja spletne aplikacije [18]. Prvotna različica priporočilnega sistema je zasnovana, da vključuje kot vhodne podatke le funkcionalne tipe zahtev. V nadaljevanju je orisan predlog integracije nefunkcionalnih zahtev z namenom olajšanja izbiro ustrezne EVM verige blokov.

4.2 Integracija nefunkcionalnih zahtev

V osnovni različici priporočilnega sistema so bile naslovljene funkcionalne zahteve, saj vloga nefunkcio-

nalnih zahtev ne bi imela dodane vrednosti pri osnovnem delovanju in algoritmu sistema pri priporočilu pametnih pogodb. Kljub temu je vključitev nefunkcionalnih zahtev v takšen sistem smotrna, saj uspešna integracija pripomore k sledečim izboljšavam:

- boljše razumevanje raznolikosti EVM verig blokov,
- podatki EVM verig blokov v kontekstu zahtev so pridobljeni iz javnih storitev umetne inteligence,
- nov tip izhodnega podatka sistema, kjer so predlagane EVM verige blokov (izhod sistema), ki ustrezajo podanim nefunkcionalnim zahtevam (vhodni podatki) in
- celovit pristop razvoja in integracije pametnih pogodb.

Pri integraciji nefunkcionalnih zahtev osnovni algoritem ter obstoječi vhodni podatki sistema ostanejo nespremenjeni. Zahteve med sabo niso komplementarne, saj vračajo dva različna vhoda. Na Sliki 2 je orisan potek delovanja priporočilnega sistema z integracijo nefunkcionalnih zahtev, kjer so glavni gradniki sistema opisani v naslednjih korakih:

1. Nefunkcionalne zahteve so izbrane s strani razvijalca programske opreme, kjer razvijalec izbere nefunkcionalno zahtevo ter vrednost, ki je lahko:
 - kvantitativna vrednost (npr. povprečna cena transakcije, hitrost potrjevanja transakcij, vključno z interakcijami pametnih pogodb, hitrost generiranja bloka idr.),
 - kvalitativna vrednost, izbrana na podlagi mehko opredeljene lastnosti (npr. skalabilnost, varnostni mehanizmi idr.) in
 - želja po dodatnih informacijah pri zahtevah, kjer je definicija metrike bodisi preveč kompleksna bodisi se zahteva lahko nanaša na različne lastnosti sistema (npr. varnost).



Slika 2: Nadgrajena arhitektura priporočilnega sistema pametnih pogodb na podlagi nefunkcionalnih zahtev kot vhodnega podatka sistema vrne seznam predlaganih verig blokov.

2. Konvencionalni pristopi iskanja kvantitativnih nefunkcionalnih zahtev iz EVM verig blokov [13]. Pristopi lahko vključujejo REST API storitve, beleženje dejavnosti v realnem času na verigi blokov, spletne pajke ter druge vire. Takšen pristop zahteva vzdrževanje logike za pridobivanje podatkov (npr. protokoli, spremembe specifikacij, posodabljanje URL naslovi spletnih virov itd.). Poleg tega se nove EMV verige blokov na novo zaganjajo za pokrivanje namenskih primerov uporabe ter obstoječe EVM verige blokov se nadgrajujejo z novimi koncepti, specifikami in drugimi lastnostmi. Z uporabo storitev umetne inteligence, predvsem klepetalniki (angl. chatbot), je mogoče vzdrževanje minimizirati. V tem koraku je smotrno podatke hraniti v relacijski podatkovni bazi.
3. V koraku izvajanja procesiranja priporočilnega sistema se neodvisno izvajata algoritma funkcionalnih in nefunkcionalnih zahtev. Pri procesiranju nefunkcionalnih zahtev je algoritem opredeljen kot večkriterijsko odločanje (glej Algoritem 1) pri čemer je vsaka zahteva predstavlja kriterij, ki mora biti izpolnjen. Kriteriji se pregledujejo neodvisno na podlagi podatkov iz prejšnje točke in sicer pred procesiranimi podatki pridobljenimi s storitvami umetne inteligence in ostalimi viri s prejšnje točke. V tem koraku je pomembno, da se ponovno preverijo storitve umetne inteligence, saj se modeli podatkov za učenje pogosto posodabljaajo. Poleg tega je smotrno uporabiti več virov storitve umetne inteligence komplementarno, saj uporabljajo različne modele, metodologije iskanja podatkov ter metode strojnega učenja, posledično različno podrobne odgovore. Na primer, pri podanem vprašanju „Ali je bila v zadnjem času izvedena kakšna pomembna nadgradnja Ethereum omrežja?“ so odgovori različni. Pri uporabi brezplačne storitve *ChatGPT* pridobimo odgovor o nadgradnjah v preteklem letu 2023. Po drugi strani nam pa storitev *Microsoft Copilot* vrne odgovor o pomembni nadgradnji v letu 2024 in sicer *Dencun*, vključno s ključnimi lastnostmi nadgradnje (glej opis Slike 1).
4. Pri pripravi podatkov je pomembno, da uporabniku podamo seznam verig blokov, ki ustrezajo podanim nefunkcionalnim zahtevam. V primeru, da ni mogoče najti verige, ki bi ustrezala vsem nefunkcionalnim zahtevam, se uporabniku izpiše seznam kandidatov ter poda tudi seznam nefunkcionalnih zahtev, ki ne ustrezajo zahtevam

vhodnih podatkov. Na tak način ima razvijalec programske opreme možnost vpogleda v pogojno sprejemljive rezultate.

5 EVALVACIJA

V tem poglavju predstavimo eksperimentalno študijo simulacije priporočilnega sistema EVM verig blokov na podlagi uporabnikovih nefunkcionalnih zahtev. Študija obsega simulacijo treh pogostih primerov uporabe, in sicer:

- Internet stvari (angl. Internet of Things, IoT),
- decentralizirane finance (angl. Decentralized Finance, DeFi) in
- lastništvo na podlagi nezamenljivih žetonov (angl. Non-fungible tokens, NFT).

Najprej določimo relevantne nefunkcionalne zahteve, ki smo jih pridobili s povpraševanjem v brezplačnih različic orodij *ChatGPT* in *Microsoft Copilot*. Skupni predlogi relevantnih nefunkcionalnih zahtev so: (i) varnost, (ii) skalabilnost (npr. odzivni čas oz. hitrost transakcije), (iii) zmogljivost ter učinkovitost (npr. hitrost transakcij), (iv) zanesljivost ter dostopnost, (v) mehanizem soglasja, (vi) celovitost ter doslednost podatkov ali integriteta, in (vii) interoperabilnost.

Izmed nabora identificiranih zahtev se bomo osredotočili zgolj na takšne, kjer je mogoče določiti kvantitativno ali kvalitativno vrednost. Na zahteve, ki so zgolj informativne narave ali jih je v osnovi težje interpretirati (npr. varnost) se ne bomo osredotočali. Na podlagi obstoječih rešitev primerov uporabe ter značilnosti poslovnih modelov določimo pomembnost posameznih kategorij, ki so prikazane na Sliki 3. Kot je predstavljeno v Poglavju 3.1 na primeru problematike trilema, so si določene konfliktne in je zato treba vedno iskati kompromise.

Najbolj dominantne kvantitativne nefunkcionalne zahteve opredelimo kot vhodne podatke priporočilnega sistema. Iz osnovnih predstavljenih kategorij bomo izbrali kvantitativne lastnosti, ki so pomembne za posamezni primer uporabe. Pri simulaciji bomo podali realistične ocene, da pridobimo rezultate kateri ustrezajo javnim EVM verigam blokom. Rezultati simulacije so predstavljeni v Tabeli 3 kjer se osredotočamo na nefunkcionalne kvalitativne zahteve značilne za uporabljene primere uporabe. Predvsem nas zanimajo stopnja distributivnosti (npr. število vozlišč, ki potrjujejo transakcije), hitrost transakcij v

korelaciji s hitrostjo generiranja blokov in cena transakcije. Slednja je zelo pomembna z vidika dolgoročne vzdržnosti, saj pri IoT primerih uporabe je število transakcij praviloma višje.

Algorithm 1 Algoritem večkriterijskega odločanja nefunkcionalnih zahtev

```

procedure VEČKRITERIJSKO ODLOČANJE
    rezultati ← rezultat priporočilnega sistema nefunkcionalnih zahtev v obliki seznama EVM verig blokov
    pogojniRezultati ← rezultat v obliki seznama EVM verig blokov kjer niso vse zahteve izpolnjene
    neizpolnjeneZahteve ← seznam neizpolnjenih zahtev
    zahteveUporabnika ← seznam vhodnih podatkov nefunkcionalnih zahtev v JSON obliki

    while seznamzahteveUporabnikanipregledan do
        ▷ iteriramo skozi seznam nefunkcionalnih zahtev uporabnika

        if kvantitativna zahteva then
            ▷ pregled ujemajočih se pogojev na podlagi kvantitativnih vrednosti iz obstoječih podatkov sistema in storitev umetne inteligence

        else if kvalitativna zahteva then
            ▷ pregled ujemajočih se pogojev na podlagi kvalitativnih zahtev iz obstoječih podatkov sistema in storitev umetne inteligence

        else if informativna zahteva then
            ▷ pregled pogojev informativne narave iz obstoječih podatkov in storitev umetne inteligence

        end if
        if neizpolnjena zahteva then

            neizpolnjeneZahteve ← trenutna zahteva
            pogojniRezultati ← rezultati
            rezultati ← null
        else

            while seznamrezultatovnipregledanALiseznampogojniRezultatatinipregledan do
                ▷ pregled ustreznosti zahteve z elementom rezultata (EVM veriga blokov) ▷ neustrezne elemente se odstrani iz seznama rezultati ALI pogojniRezultati ▷ v primeru, da je elementov v rezultati enako 0 se vse rezultate shrani v pogojniRezultati ter zahteve doda v neizpolnjeneZahteve
            end while

        end if
    end while
end procedure
    
```

Tabela 1: Osnovne plasti, ki opredeljujejo EVM verige blokov.

Primer uporabe	Nefunkcionalne zahteve	Kvantitativne zahteve	Predlagane EVM verige blokov
IoT	Distributivnost, hitrost generiranja bloka, cena transakcije	nizka, minimalno \$2\$ sekund, nizka	Optimism, Arbitrum, ShimmerEVM
DeFi	Distributivnost, hitrost generiranja bloka, cena transakcije, št. vozlišč	srednja, minimalno \$17\$ sekund, srednja, visoko	Ethereum, Avalanche
NFT	Distributivnost, hitrost generiranja bloka, cena transakcije	visoka, minimalno \$30\$ sekund, visoka	Binance Smart Chain, Polygon, Avalanche, Fantom



Slika 3: Pomembnost nefunkcionalnih zahtev primerov uporabe.

6 DISKUSIJA IN ZAKLJUČEK

Delo predstavlja analizo nefunkcionalnih zahtev, ki so smotrne za verige blokov s podporo EVM pametnih pogodb. Najprej smo predstavili razvoj nefunkcionalnih zahtev, ki je doseglo razpon s pojavom oblačnih arhitektur konec prvega desetletja tega tisočletja. Nato smo podrobneje pregledali razvoj plasti EVM verig blokov, ki so nastali kot odgovor na zahteve po podpori novih primerov uporabe. Pomemben aspekt našega dela predstavljajo javne storitve umetne inteligence, ki nam kot dodano vrednost nudijo informacije o EVM verigah blokov, saj obstoječi pristopi so težavni za dolgoročno vzdrževanje (npr. analiza preko API storitev, lastni spletni pajki idr., spletni viri z namenskimi informacijami idr.). Nefunkcionalne zahteve smo naslovili kot nadgradnjo obstoječega priporočilnega sistema, ki primarno kot rezultat vrača seznam predlog pametnih pogodb glede na podane funkcionalne zahteve sistema ter oblačne arhitekture.

V eksperimentalni študiji smo najprej izpostavili tri aktualne primere uporabe (IoT, DeFi in NFT). Najprej smo na podlagi orodij umetne inteligence ChatGPT in Microsoft Copilot identificirali najpomembnejše nefunkcionalne zahteve oziroma prisotne kot rezultat obeh orodij. Ocenili smo pomembnost nefunkcionalnih zahtev posameznih primerov uporabe ter ugotovili, da IoT primeru uporabe največjo težo dajejo hitrosti izvedbe pri čemer mehanizem soglasja igra najmanjšo vlogo. Pri primeru uporabe DeFi so najpomembnejša varnost, predvsem z vidika funkcionalnosti pametnih pogodb ter vsesplošna dovzetnost EVM verige blokov za napade (npr. premajhno število vozlišč verige blokov lahko vodi do t.i. 51% napada). Pri primeru uporabe NFT je varnost tudi

dominantna lastnost, interoperabilnost je najmanj pomembna lastnost, saj praviloma se elementi transakcij ne prenašajo preko različnih verig in skalabilnost tudi ne igra pomembne vloge zaradi pretežno determinističnega delovanja primera uporabe.

Glede na to, da je to prva različica sistema s podporo nefunkcionalnih zahtev, obstajajo precejšnje možnosti izboljšave. Kakovost informacij je mogoče izboljšati z bolj aktualnimi modeli za učenje orodij umetne inteligence, ki so prisotni v plačljivih različicah. Poleg tega bi bilo smiselno analizirati odgovore različnih storitev umetne inteligence z namenom izboljšave priporočilnega sistema. V prihodnje bi se lahko osredotočili na migracijo podatkov iz relacijske baze v obliko predstavitve podatkov ontologije, saj bi tako lahko vključili naprednejše pristope strojnega učenja ter pripravili naprednejša pravila sklepanja (angl. reasoning) na podlagi podatkov.

ZAHVALA

Raziskava je bila finančno podprta s sredstvi projekta Evropske unije Obzorje program za raziskave in inovacije na podlagi sporazumov št. 101093274 (TrustChain projekt: Fostering a Human-Centered, Trustworthy and Sustainable Internet) in št. 101092052 (BUILDCHAIN projekt: BUILDing knowledge book in the blockCHAIN distributed ledger. Trustworthy building life-cycle knowledge graph for sustainability and energy efficiency).

LITERATURA

- [1] F. Liu, S. He, Z. Li, P. Xiang, J. Qi, and Z. Li, "An overview of blockchain efficient interaction technologies,"
- [2] *Frontiers in Blockchain*, vol. 6, p. 996070, 2023.
- [3] N. Szabo et al., "Smart contracts: Building blocks for digital markets, extropy," *The Journal of Transhumanist Thought*, vol. 16, no. 18, pp. 2–20, 1996.
- [4] M. Glinz, "On non-functional requirements," in *15th IEEE international requirements engineering conference (RE 2007)*, pp. 21–26, IEEE, 2007.
- [5] S. Nakamoto, "Re: Bitcoin p2p e-cash paper," *The Cryptography Mailing List*, pp. 1–2, 2008.
- [6] J. Yli-Huomo, D. Ko, S. Choi, S. Park, and K. Smolander, "Where is current research on blockchain technology?—a systematic review," *PLOS ONE*, vol. 11, pp. 1–27, 10 2016.
- [7] R. Nokhbeh Zaeem, K. C. Chang, T.-C. Huang, D. Liao, W. Song, A. Tyagi, M. Khalil, M. Lamison,
- [8] S. Pandey, and K. S. Barber, "Blockchain-based self-sovereign identity: Survey, requirements, use-cases, and comparative study," in *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pp. 128–135, 2021.
- [9] V. Buterin et al., "Ethereum white paper," *GitHub repository*, vol. 1, pp. 22–23, 2013.

- [9] M. Staderini, E. Schiavone, and A. Bondavalli, "A requirements-driven methodology for the proper selection and configuration of blockchains," in *2018 IEEE 37th Symposium on Reliable Distributed Systems (SRDS)*, pp. 201–206, 2018.
- [10] M. Kassab, "Exploring non-functional requirements for blockchain-oriented systems," in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*, pp. 216–219, 2021.
- M. S. Farooq, M. Ahmed, and M. Emran, "A survey on blockchain acquainted software requirements engineering: Model, opportunities, challenges, and future directions," *IEEE Access*, vol. 10, pp. 48193–48228, 2022.
- [11] T. A. Tisha and M. M. A. Shibly, "Non-functional requirements for blockchain: Challenges and new directions," in *IOP Conference Series: Materials Science and Engineering*, vol. 1110, p. 012016, IOP Publishing, 2021.
- [12] R. Jia and S. Yin, "To evm or not to evm: Blockchain compatibility and network effects," in *Proceedings of the 2022 ACM CCS Workshop on Decentralized Finance and Security, DeFi'22*, (New York, NY, USA), p. 23–29, Association for Computing Machinery, 2022.
- S. Gec, D. Lavbič, V. Stankovski, and P. Kochovski, "Towards a smarter iot environment with ethereum virtual machine enabling ledgers," in *2022 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 1229–1232, 2022.
- G. Wood, "Polkadot: Vision for a heterogeneous multi-chain framework," *White Paper*, vol. 21, pp. 2327–4662, 2016.
- [13] S. K. Singh, S. Kumar, and P. S. Mehra, "Chat gpt google bard ai: A review," in *2023 International Conference on IoT, Communication and Automation Technology (ICICAT)*, pp. 1–6, 2023.
- [14] Z. A. Mann, "Notions of architecture in fog computing," *Computing*, vol. 103, p. 51–73, jan 2021.
- [15] S. Gec, V. Stankovski, D. Lavbič, and P. Kochovski, "A recommender system for robust smart contract template classification," *Sensors*, vol. 23, no. 2, 2023.
- [15] S. Gec, P. Kochovski, V. Stankovski, and D. Lavbic, "Project oriented teaching approach of decentralised applications for undergraduate students," in *The 15th International Conference on Education Technology and Computers, ICETC 2023*, (New York, NY, USA), p. 207–215, Association for Computing Machinery, 2024.

■

Sandi Gec je zaposlen kot asistent na Fakulteti za računalništvo in informatiko Univerze v Ljubljani, kjer poučuje predmete s področij informacijskih sistemov, spletnih tehnologij in decentraliziranih aplikacij. Dejavno je bil na številnih projektih, največ na mednarodnih v okviru Obzorje 2020 (npr. SWITCH, ENTICE, DECENTER, ONTOCHAIN idr.). Trenutno sodeluje na projektih TRUSTCHAIN in BUILDCHAIN v okviru Obzorje Evropa ter mednarodnem projektu ESSA, kjer se ukvarja z različnimi področji programskega inženirstva, predvsem z novimi pristopi tehnologije veriženja blokov s poudarkom na pametnih pogodbah ter interoperabilnost med različnimi verigami.

■

Vlado Stankovski je redni profesor računalništva in informatike. Ima bogate izkušnje na področju programskega inženirstva, računalništva v oblaku, na robu in v megli, porazdeljenih sistemov, semantike ter tehnologij umetne inteligence (strojno, globoko učenje). Sodeloval je pri načrtovanju, razvoju in integraciji tehnoloških vmesnih programske opreme. Sodeloval je pri več nacionalnih in mednarodnih projektih, v konzorciju Superračunalniški center Slovenije, na projektu pametne specializacije IQ DOM ter v gruči za programsko inženirstvo projektov Obzorje 2020 kot predstavnik projektov ENTICE, SWITCH in DECENTER. Trenutno je v okviru Obzorje Evropa predstavnik projektov TRUSTCHAIN, BUILDCHAIN, ExtremeXP, EBSI-VECTOR in Swarmchestrare.

Analiza uporabe aplikacije za sledenje stikov med mladimi: študija primera Nemčije

Lana Masnec², Marina Trkman¹

¹ Univerza v Ljubljani, Fakulteta za upravo, Gosarjeva 5, 1000 Ljubljana

² Univerza v Ljubljani, Fakulteta za družbene vede, Kardeljeva ploščad 5, 1000 Ljubljana
masneclana@gmail.com, marina.trkman@fu.uni-lj.si

Izveček

Pandemijo koronavirusne bolezni covid-19, so mnogi zaznali kot prelomno točko za množično vpeljevanje in uporabo storitev e-zdravja. Del storitev e-zdravja, ki lahko pomembno prispevajo k obvladovanju širjenja bolezni, so aplikacije za sledenje stikom. Te so učinkovite, ko jih dejansko uporablja zadostni delež prebivalstva (56-60%), ki pa v evropskih državah, v času pandemije covid-19, ni bil dosežen. Ker lahko v prihodnosti pričakujemo širjenje nevarnih virusov, je potrebno razumeti dejavnike aktivne uporabe in zavračanja aplikacij za sledenje stikom. V članku preučujemo ali prej znani dejavniki vpliva na namero uporabe, vplivajo tudi na dejansko uporabo aplikacije. V empirični študiji se osredotočamo na mlade odrasle, stare med 18 in 35 let. Z rezultati binarne logistične regresije lahko potrdimo, da na dejansko uporabo aplikacije statistično značilno vplivajo: zaznana enostavnost uporabe, družbeni vpliv, zaznana uporabnost na ravni posameznika in zaskrbljenost glede zasebnosti. Vpliva pa nismo potrdili za zaznana uporabnost na ravni družbe in za zaupanje v tehnologije.

Ključne besede: aplikacija za sledenje stikom, privzemanje tehnologij, dejanska uporaba, mladi odrasli, pandemija covid-19.

Analysis of factors of actual use of contact tracing application

Abstract

The COVID-19 coronavirus pandemic was seen by many as a major milestone for the mass uptake and use of eHealth services. Contact tracing applications are part of the eHealth services that can make an important contribution to controlling the spread of the disease. These are only effective when a sufficient proportion of the population (56-60%) actually uses them, which was not achieved in European countries at the time of the COVID-19 pandemic. As we can expect the spread of dangerous viruses in the future, it is necessary to understand holistically the drivers of active use and rejection of contact tracing applications. In this article, we examined whether previously known factors influencing the intention to use also influence the actual use of the application. We focus on young adults aged between 18 and 35. We find that the key factors influencing the actual use of a contact tracing application are perceived ease of use, social influence, perceived usefulness at the personal level, and privacy concerns. However, we have not confirmed the impact on perceived usefulness at the societal level and on trust in technologies.

Keywords: contact tracing application, technology acceptance, actual use, young adults, COVID-19 pandemic.

1 UVOD

E-zdravje velja za pomemben del javnega zdravstva in že več let lajša dostop do zdravstvenih informacij posameznikov. Pandemijo koronavirusne bolezni covid-19 so mnogi zaznali kot prelomno točko za množično vpeljevanje in uporabo storitev e-zdravja [27]. Te so omogočile upravljanje in obvladovanje več različnih vidikov bolezni covid-19 [13], med drugim preprečevanje in nadzor širjenja bolezni [10]. Del

storitev e-zdravja, ki so pripomogle k obvladovanju pandemije koronavirusne bolezni covid-19, so aplikacije za sledenje stikom. Te omogočajo hitro dostopne in natančne informacije o stanju okužb, kar pripomore k ustreznemu ukrepanju, zaustavitvi širjenja bolezni in hkrati omejitvi škode gospodarstvu [8].

Aplikacije za sledenje stikom lahko pomembno prispevajo k obvladovanju pandemije, le ob zadostni uporabi [24]. Tovrstne aplikacije so učinkovite, ko jih

aktivno uporablja vsaj 56 % do 60 % prebivalstva [2,5]. Stopnja dejanske uporabe aplikacij za sledenje stikom je v evropskih državah bila nizka. Delež prebivalcev, ki so dejansko uporabljali aplikacije za sledenje stikom, je bil najvišji na Finskem (45 %), medtem ko je v Sloveniji aplikacijo dejansko uporabljalo le 6% državljanov [11]. Mnoge države so poročale o slabem sodelovanju državljanov pri zaježitvi virusa. Ker je v prihodnosti pričakovati širjenje nevarnih virusov, je nujno celostno razumeti dejavnike, ki vplivajo na sprejemanje in zavračanje aplikacij za sledenje stikom [25].

Mladi odrasli ljudje so bili v literaturi večkrat označeni kot skupina ljudi nagnjena k neupoštevanju priporočil za ščitenje javnega zdravja [20]. Prav tako je znano, da se mladi odrasli vedejo drugače kot starejši odrasli, saj so v povprečju bolj tesnobni, depresivni in nesigurni [14]. V članku preučujemo vedenjske zakonitosti mladih odraslih oseb starih med 18 in 35 let. Naše raziskovalno vprašanje je: kateri dejavniki vplivajo na mlade ljudi, da dejansko uporabljajo aplikacijo za sledenje stikom? Nabor dejavnikov smo poiskali v obstajajoči literaturi privzemanja aplikacij za sledenje stikom, ki pa so v večini preučevale vplive dejavnikov na namero uporabe in ne na dejansko uporabo. Analizirali smo sekundarne podatke pridobljene v Nemčiji leta 2022.

Članek sledi sledeči strukturi. V drugem poglavju, opredelimo področje e-zdravja. V tretjem poglavju predstavimo aplikacijo za sledenje stikom ter v četrtem zanimive dejavnike, ki vplivajo na namero uporabe. V petem poglavju predstavimo anketo v okviru empirične študije ter v šestem diskutiramo rezultate. Sledijo zaključne besede.

2 TEORETIČNA IZHODIŠČA

2.1 E-zdravje med pandemijo covid-19

E-zdravje že več let lajša dostop do zdravstvenih informacij posameznikov. Pandemijo koronavirusne bolezni covid-19 so mnogi zaznali kot prelomno točko za množično vpeljevanje in uporabo storitev e-zdravja [27]. Številne evropske vlade so v tem obdobju začele vlagati v razvoj in posodobitev sistema ter aplikacij e-zdravja, kar je omogočilo upravljanje in obvladovanje različnih vidikov bolezni covid-19 [13], in sicer preprečevanje, pregledovanje, diagnosticiranje, zdravljenje, izobraževanje, upravljanje in nadzor [10].

Uporaba informacijske tehnologije se je izkazala kot pomemben del preventivnega zdravstva, katerega

cilj je zmanjšati verjetnost pojava bolezni z zaščito in zgodnjim odkrivanjem [18]. Pametni telefoni in podobne nosljive naprave so omogočile vpeljevanje novih mobilnih aplikacij za sledenje stikom, ki so pripomogle k izvedbi preventivnih ukrepov samoizolacije prebivalstva in s tem zmanjšanju tveganja za nadaljnje širjenje okužbe [13]. Pandemija covid-19 je pospešila privzemanje storitev e-zdravja na več področjih [10].

Pri omejevanju širjenja nalezljive bolezni je sledenje preteklim stikom okužene osebe z drugimi osebami ključno [6]. Tradicionalen način sledenja stikom je proces, ki se opravlja ročno s pomočjo vprašalnikov in papirnatih zemljevidov, za kar so potrebni usposobljeni izvajalci intervjujev in visoka natančnosti pridobljenih podatkov [7]. Do podatkov o preteklih stikih posameznika, okuženega z boleznijo, vodijo dolgotrajni razgovori, ki pa so lahko nesmiselni, saj se zanašamo le na človeški spomin [6]. Poleg tega ni mogoče identificirati oseb, ki jih posameznik ne pozna [26].

V času pandemije so se v večini držav po svetu uveljavile digitalne tehnike sledenja stikom [6]. Namen aplikacije za sledenje stikom je izboljšanje časovne učinkovitosti in natančnosti pridobivanja informacij o stikih ter opozarjanja ljudi, ki so bili v neželenem kontaktu [6].

Aplikacije za sledenje stikom (angl. Contact tracing applications – CTA; angl. Proximity tracing applications – PTA) so del omenjenih digitalnih orodij, ki s pomočjo lokacijske tehnologije GPS ali tehnologije Bluetooth omogočajo anonimno sledenje stikom [34]. Veljajo za eno izmed ključnih strategij, ki so jo vlade in organi javnega zdravstva v večini držav po svetu, uporabili za obvladovanje pandemije koronavirusne bolezni covid-19, saj omogočajo hitro dostopne in natančne informacije o stanju okužb.

2.1 Dejavniki privzemanja aplikacij za sledenje stikom

Za pojasnjevanje dejavnikov privzemanja aplikacij za sledenje stikom so avtorji v preteklih študijah uporabili že uveljavljene modele in teorije in jih prilagodili kontekstu dotične aplikacije. Enotna teorija sprejemanja in uporabe tehnologije (angl. Unified Theory of Acceptance and Use of Technology – UTAUT) je široko uporabljen teoretični model, ki sprejemanje in uporabo nove tehnologije napoveduje glede na posameznikovo pričakovano uspešnost, pričakovani napor in družbeni vpliv [31]. Ker privzemanje aplikacij za sledenje stikom predstavlja edinstveni kon-

tekst delovanja, je poleg univerzalnih dejavnikov iz teorij privzemanja kot je UTAUT, smiselno preučevati dodatne kontekstualne dejavnike. Predhodno so koristnost aplikacije proučevali tako iz vidika posameznika kot iz vidika družbe [30].

Med zaznamimi koristmi na ravni posameznika so identificirali občutek varnosti in občutek pravočasnega informiranja [1]. Megnin-Viggars in drugi [19] poleg pridobivanja informacij o tveganju okužbe in vpogleda v širjenje virusa, kot pomembno korist na ravni posameznika, prepoznavajo zaščito lastnega zdravja in zdravja družinskih članov ter prijateljev.

Zetterholm in drugi [35] poudarjajo, da je pomembnost zaznanih koristi na ravni posameznika vprašljiva, zaradi narave aplikacij za sledenje stikom. Koristi od posameznikove vestne uporabe aplikacije in ustreznega vedenja so deležni predvsem drugi uporabniki aplikacije oziroma sohržavljeni. Zato je smiselno kot pomemben dejavnik privzemanja analizirati vpliv zaznane koristi uporabe na ravni družbe [9,19,35], ki pa so: obvladovanje pandemije, vrnitev v normalno življenje in zmanjšanje obremenitev zdravstvenega sistema [1].

Zaupanje se je prav tako pokazalo kot pomemben dejavnik privzemanja [35]. V kontekstu aplikacije za sledenje stikom je relevantnih več vrst zaupanj. Ker so aplikacije za sledenje stikom v večini držav bile izdane s strani vladnih organov, zaupanje v vlado pomembno vpliva na posameznikovo privzemanje aplikacije. V azijskih državah so avtorji prepoznali visoko stopnjo zaupanja v vlado in visoko stopnjo sprejemanja aplikacij za sledenje stikom, kljub temu, da aplikacije v teh državah v višji meri posegajo v zasebnost državljanov [35]. Da je zaupanje v vlado

in zdravstvene organe povezano z večjo verjetnostjo dejanske uporabe aplikacije potrjujejo tudi von Wyl in drugi [32], Albrechtove in drugi [3], [15], Huang in drugi [17] ter Bannister-Tyrrell in drugi [4].

Zaupanje v tehnologijo pozitivno vpliva na posameznikovo namero uporabe aplikacije za sledenje stikom [21]. Uspešna uvedba mobilnih aplikacij za boj proti pandemiji covid-19 je v demokratičnih državah odvisna od sposobnosti zagotavljanja zaupanja ljudi v tehnologijo [21, 22]. Pomemben dejavnik sprejemanja aplikacije za sledenje stikom je tudi zaupanje v samo aplikacijo, ki se kaže v dojemanju aplikacije kot učinkovite, zanesljive in varne [19].

Med ključnimi dejavniki zavračanja in neuporabe aplikacij za sledenje stikom, avtorji prepoznavajo skrb za zasebnost [4, 16, 17, 28, 29, 30, 33]. Skrb za zasebnost lahko prepoznamo kot pomemben etični pomislek, ki se kaže v strahu pred zlorabo osebnih podatkov, strahu pred hekerji ali pred posredovanjem osebnih podatkov tretjim osebam [35]. Posamezniki niso pripravljeni sprejeti aplikacije za sledenje stikom, v kolikor menijo, da njihovi osebni podatki niso dobro zaščiteni [9]. Študije Prakasha in Dasa [23], Rossa [25], Foxa in drugih [12], von Wyla in drugih [32] ter Megnin-Viggars in drugih [19] potrjujejo, da posameznikova skrb za zasebnost informacij vpliva na odpor do uporabe aplikacije za sledenje stikom.

3 EMPIRIČNA ŠTUDIJA

3.1 Spremenljivke

V anketi so anketiranci odgovarjali na vprašanja predstavljena v Tabeli 1 na lestvici od 1 (sploh se ne strinjam) do 7 (zelo se strinjam).

Tabela 1: Spremenljivke uporabljene v binarni logistični regresiji

Spremenljivka	Vprašanje v anketi
Starost	Katerega leta ste rojeni?
Izobrazba	Izberite ustrezno skupino izobrazbe.
Spol	Kaj je vaš spol?
Zaznana enostavnost uporabe	Uporaba aplikacije je zame jasna in razumljiva.
Družbeni vpliv	Ljudje, ki so zame pomembni, menijo, da bi moral/a uporabljati aplikacijo.
Zaznana uporabnost na ravni posameznika	Skrbi me, da lahko razvijalci oziroma ponudniki aplikacije, preko nje pridobijo moje osebne podatke.
Zaznana uporabnost na ravni družbe	Z uporabo aplikacije je življenje v družbi varneje.
Zaupanje v vlado	Menim, da država deluje v interesu državljanov.
Zaupanje v tehnologijo	Zaupam novim tehnologijam, dokler mi ne dokažejo, da jim ne bi smel zaupati.
Skrbi glede zasebnosti	Skrbi me, da lahko razvijalci oziroma ponudniki aplikacije, preko nje pridobijo moje osebne podatke.

3.2 Pridobivanje podatkov

V okviru empiričnega dela smo opravili analizo sekundarnih podatkov pridobljenih v anketi, ki se je navezovala na uporabo nemške mobilne aplikacije Corona-Warn-App. Anketa je bila pripravljena v okolju Qualtrics. Izvedena je bila v nekaj dneh začetku junija 2022. Spletni panel je priskrbelo agencija Prolific. Podatke smo za namen preučevanja vedenja mladih filtrirali in v vzorec za analizo vključili le mlade odrasle, stare med 18. in 35. letom. Z orodjem SPSS smo pripravili tabelo z deskriptivno statistiko in rezultate binarne logistične regresije.

3.3 Deskriptivna statistika

Kot je razvidno iz Tabele 2 je bilo v raziskavo vključenih 290 posameznikov, ki so bili v letu 2022 stari med 18 in 35 let. Vzorec predstavlja enako število moških (145) in žensk (145). Največ sodelujočih je v času ankete imelo opravljeno višješolsko ali univerzitetno izobrazbo (183 oziroma 63,1%). Srednješolsko ali nižjo stopnjo izobrazbe je dokončalo 107 (36,9%) anketirancev.

Tabela 2: Demografske značilnosti sodelujočih v raziskavi

Spremenljivka	n	%
Spol		
Moški	145	50,0
Ženski	145	50,0
Starost		
18-24 let	95	32,8
25-35 let	195	67,2
Izobrazba		
Srednješolska izobrazba ali manj	107	36,9
Višješolska in univerzitetna izobrazba	183	63,1
	290	100,0

4 REZULTATI

Z binarno logistično regresijo smo preverili povezanost med odvisno spremenljivko – dejanska uporaba aplikacije za sledenje stikom in neodvisnimi spremenljivkami – spol, starost, izobrazba, zaupanje v vlado, zaupanje v tehnologijo, zaznana enostavnost uporabe, družbeni vpliv, zaznana uporabnost na ravni posameznika, zaznana uporabnost na ravni družbe in skrbi glede zasebnosti.

Omnibus test (hi kvadrat = 97,336, $p < 0,001$) nakazuje, da je regresijski model statistično značilen in dobro pojasnjuje del variance odvisne spremenljivke.

Nagelkerke R2 (0,384) kaže, da regresijski model pojasnjuje približno 38% variance odvisne spremenljivke. Na podlagi Hosmer-Lemeshow testa (hi-kvadrat = 2,404; $p = 0,966$) pa lahko zaključimo, da je model statistično neznačilen, kar kaže, da se ta ustrezno prilega danim podatkom. Glede na dobljene vrednosti, je model primeren za nadaljnjo vsebinsko interpretacijo.

Rezultati binarne logistične regresije (glej Tabelo 3) kažejo, da vpliv starosti ($B = -0,323$; $\text{Exp}(B) = 0,724$; $p = 0,358$), izobrazbe ($B = -0,378$; $\text{Exp}(B) = 0,685$; $p = 0,271$), spola ($B = 0,189$; $\text{Exp}(B) = 1,208$; $p = 0,518$), stopnje zaupanja v vlado ($B = 0,119$; $\text{Exp}(B) = 1,126$; $p = 0,331$) in zaznane uporabnosti na ravni družbe ($B = 0,213$; $\text{Exp}(B) = 1,238$; $p = 0,101$) na dejansko uporabo aplikacije za sledenje stikom, pri 5% stopnji značilnosti, ni statistično značilen. Posamezniki, ki se popolnoma strinjajo, da tehnologiji zaupajo, dokler jim ta ne dokaže nasprotno ($B = -0,277$; $\text{Exp}(B) = 0,758$; $p = 0,014$), bodo manj verjetno aktivno uporabljali aplikacijo za sledenje stikom. Posamezniki, ki se popolnoma strinjajo, da je uporaba aplikacije jasna in razumljiva ($B = 0,337$; $\text{Exp}(B) = 1,400$; $p = 0,044$), bodo bolj verjetno aktivno uporabljali aplikacijo za sledenje stikom. Posamezniki, ki se popolnoma strinjajo, da njihovi pomembni drugi menijo, da bi ti morali uporabljati aplikacijo za sledenje stikom ($B = 0,325$; $\text{Exp}(B) = 1,384$; $p = 0,002$), jo bodo tudi bolj verjetno aktivno uporabljali. Tisti, ki se popolnoma strinjajo, da bi se z uporabo aplikacije počutili varneje ($B = 0,280$; $\text{Exp}(B) = 1,323$; $p = 0,015$), bodo bolj verjetno aktivno uporabljali aplikacijo. Posamezniki, ki se popolnoma strinjajo s trditvijo »Skrbi me, da bi lahko razvijalci oziroma ponudniki aplikacije, preko nje, pridobili moje osebne podatke ($B = -0,282$; $\text{Exp}(B) = 0,754$; $p = 0,001$), bodo manj verjetno aktivno uporabljali aplikacijo za sledenje stikom.

Dobljeni rezultati kažejo, da lahko pri 5% stopnji značilnosti oziroma s tveganjem, ki je manjše od 5%, trdimo, da zaupanje v tehnologijo, zaznana enostavnost uporabe, družbeni vpliv, zaznana uporabnost na ravni posameznika in skrbi glede zasebnosti pomembno vplivajo na dejansko uporabo aplikacije za sledenje stikom.

5 DISKUSIJA

V članku smo raziskovali vpliv zaupanja v vlado, zaupanja v tehnologijo, zaznane enostavnosti uporabe, družbenega vpliva, zaznane uporabnosti na osebni ter družbeni ravni in zaskrbljenosti glede zasebnosti, na dejansko uporabo aplikacije za sledenje stikom.

Tabela 3: Rezultati binarne logistične regresije

	B	S.E.	Exp(B)	Sig.
Starost	-0,323	0,352	0,724	0,358
Izobrazba	-0,378	0,344	0,685	0,271
Spol	0,189	0,292	1,208	0,518
Zaznana enostavnost uporabe	0,337	0,167	1,400	0,044
Družbeni vpliv	0,325	0,104	1,384	0,002
Zaznana uporabnost na ravni posameznika	0,280	0,115	1,323	0,015
Zaznana uporabnost na ravni družbe	0,213	0,130	1,238	0,101
Zaupanje v vlado	0,119	0,122	1,126	0,331
Zaupanje v tehnologijo	-0,277	0,113	0,758	0,014
Skrbi glede zasebnosti	-0,282	0,088	0,754	0,001

Opombe: S sivo so označene kontrolne spremenljivke. Omnibus test: hi-kvadrat = 97,336; df = 10; p < 0,001; Nagelkerke R² = 0,384; Hosmer-Lemeshow test: hi-kvadrat = 2,404; df = 8; p = 0,966

Osredotočili smo se na populacijo mladih odraslih oseb. Rezultati empirične študije so pokazali, da zaupanje v vlado nima statistično značilnega vpliva na posameznikovo dejansko uporabo aplikacije za sledenje stikom. Drugi avtorji naši ugotovitvi nasprotujejo in zaključujejo, da zaupanje v vlado pomembno vpliva na posameznikovo privzemanje aplikacije [3,15,17,32,35]. Zetterholm in drugi [35] odkrivajo, da se je v nekaterih študijah o privzemanju aplikacij za sledenje stikom, zaupanje v vladne organe izkazalo za pomembnejši dejavnik od skrbi za zasebnost, vprašanj varnosti in tveganj za posameznikovo svobodo. Bannister-Tyrrell in drugi [4] so v sistematičnem pregledu literature, zaupanje v vlado prepoznali kot najpogosteje naveden dejavnik sprejemanja aplikacij za sledenje stikom. Na drugi strani Ross [25], Touzani in drugi [29], Walrave in drugi [33] in Megnin-Viggars in drugi [19] nezaupanje v vladne organe navajajo kot enega ključnih dejavnikov zavračanja uporabe aplikacije.

V empirični študiji smo ugotovili, da posameznikovo zaupanje v tehnologijo statistično značilno vpliva na neuporabo aplikacije za sledenje stikom. Naši ugotovitvi sicer nasprotuje raziskava Oldeweme in drugih [36], ki ugotavljajo, da zaupanje v tehnologijo pozitivno vpliva na posameznikovo namero uporabe aplikacije za sledenje stikom. Ne glede na to, pa je raziskava Trkman in drugih [30] pokazala, da ima zaupanje v tehnologijo posreden vpliv na namero upo-

rabe preko skrbi za zasebnost. Naši rezultati regresije kažejo, da ima skrb za zasebnost statistično značilen vpliv na dejansko uporabo, kar je v skladu z mnogimi drugimi raziskavami [9,16,23,30].

Naša empirična študija je pokazala statistično značilen vpliv dveh spremenljivk iz modela UTAUT. Ena predstavlja vpliv zaznane enostavnosti uporabe in druga družbenega vpliva na uporabo aplikacije. Posamezniki, ki uporabo aplikacije za sledenje stikom zaznavajo kot enostavno, bodo bolj verjetno aktivno uporabljali samo aplikacijo. Zaznana enostavnost uporabe je povezana s tehnološko samoučinkovitostjo [35]. Oldeweme in drugi [36] družbeni vpliv opredeljujejo kot pomembno strategijo iskanja informacij, pri čemer se posameznik o tehnologiji posvetuje s pomembnimi drugimi. V raziskavi Guazzinija in drugih [15] se je izkazalo, da je poznavanje pomembnih drugih, posebej družinskih članov in prijateljev, pomembno vplivalo na posameznikovo uporabo aplikacije za sledenje stikom. Sodelovanje v družbenih skupinah na vedenje posameznika ustvarja pritisk, saj se skuša obnašati v skladu s postavljenimi pričakovanji [36]. Družbeni vpliv se je izkazal za posebej pomembnega v okoliščinah, ki za uspeh zahtevajo množično sprejetje nove tehnologije, kar velja za tehnologije vladnega nadzora v času pandemije covid-19 [12].

Rezultati empirične študije so pokazali, da zaznane koristi na ravni posameznika statistično značilno vplivajo na dejansko uporabo, medtem ko zaznana uporabnost na ravni družbe nima statistično značilnega vpliva. Abramova in drugi (2022) kot ključne koristi na ravni posameznika prepoznajo občutek varnosti in občutek pravočasnega informiranja [1]. Naši rezultati o vplivu zaznane uporabnosti na ravni družbe na dejansko uporabo niso v skladu s Trkman in drugimi [30] in Fox in drugi [12], ki so sicer ugotovili, da le-ta ima vpliv na namero uporabe aplikacije. Torej, direktnega vpliva zaznane uporabnosti na ravni družbe na dejansko uporabo ni, obstaja pa posreden vpliv. Seveda ljudje v svetovni zdravstveni krizi pri sprejemanju odločitev upoštevajo tudi koristi na ravni družbe [1,9,19,35]). V kolikor posameznik verjame, da lahko uporaba aplikacije koristi njemu in širši skupnosti, bo aplikacijo bolj verjetno privzel, tudi če bo ta zahtevala razkritje občutljivih podatkov o lokaciji in zdravju [12,19]. Na drugi strani, bodo ljudje, ki zaznavajo pomanjkanje koristi pri aplikaciji za sledenje stikom, bolj verjetno razvili odpor do njene uporabe [19,23,32,33].

6 ZAKLJUČEK

V članku smo preučevali aplikacije za sledenje stikom in ključne dejavnike, ki vplivajo na njihovo privzemanje in dejansko uporabo. Raziskovali smo dejavnike sprejemanja in zavračanja decentraliziranih aplikacij za sledenje stikom med mladimi na primeru Nemčije. Naša raziskava je prispevala k dodatnemu razumevanju dejavnikov dejanske uporabe aplikacij za sledenje stikom med mladimi odraslimi. V kolikor želimo v prihodnosti bolje razumeti privzemanje tovrstnih tehnologij, bo potrebno celostno upoštevati okoliščine, v katerih so bile vpeljane.

Namen članka je bil odgovoriti na raziskovalno vprašanje: kateri dejavniki vplivajo na mlade ljudi, da dejansko uporabljajo aplikacijo za sledenje stikom? Preverili smo vpliv več spremenljivk na dejansko uporabo aplikacije za sledenje stikom. V empirični študiji smo izvedli binarno logistično regresijo in tako analizirali sekundarno pridobljene podatke. Naša empirična študija kaže, da tako družbeni vpliv kot zaznana enostavnost uporabe vplivata na dejansko uporabo aplikacije. Zanimiva ugotovitev je, da medtem ko zaznana uporabnost na ravni posameznika statistično vpliva na dejansko uporabo, pa zaznana uporabnost na ravni družbe ne. Statistično neznačilen vpliv ima tudi zaupanje v vlado. Nazadnje, zaupanje v tehnologijo in skrb za zasebnost pomembno vplivata na dejansko uporabo. Zanimivo je, da statistično značilen vpliv zaupanja v tehnologijo na dejansko uporabo negativen. To pa v bistvu pomeni, da je v bodoče potrebno vključiti v analizo tudi vmesni člen med njima. Bodoče raziskave bi lahko bolje raziskale vlogo skrbi za zasebnost [30].

Kot omejitev analize je potrebno izpostaviti, da so dejavniki uporabe aplikacije bili merjeni z eno samo trditvijo. V kolikor bi spremenljivke bile merjene z več trditvami in bi naredili sestavljene spremenljivke, bi lahko dobili bolj zanesljive rezultate in model prilagodili teoretičnim konceptom iz literature. Morebiti bi bilo v bodoče smiselno raziskati razlike pri privzemanju aplikacije med moškimi in ženskami, med različnimi izobrazbenimi skupinami in med različnimi starostnimi skupinami.

ZAHVALA

Raziskovalno delo soavtorice Marine Trkman je financirala Agencija za raziskovalno dejavnost Republike Slovenije (ARIS) v okviru raziskovalnega projekta J7-50185. Omenjena soavtorica deluje v okviru programske skupine P5-0399.

LITERATURA

- [1] Abramova, O., Wagner, A., Olt, C. M. in Buxmann, P. (2022). One for all, all for one: Social considerations in user acceptance of contact tracing apps using longitudinal evidence from Germany and Switzerland. *International Journal of Information Management*, 64. <https://dx.doi.org/10.1016/j.ijinfomgt.2022.102473>
- [2] Afroogh, S., Esmalian, A., Mostafavi, A., Akbari, A., Rasoulkhani, K., Esmaeili, S. in Hajiramezanali, E. (2022). Tracing app technology: an ethical review in the COVID 19 era and directions for post COVID 19. *Ethics and Information Technology*, 24(30). <https://dx.doi.org/10.1007/s10676-022-09659-6>
- [3] Albrecht, R., Jarecki, J. B., Meier, D. S. in Rieskamp, J. (2021). Risk preferences and risk perception affect the acceptance of digital contact tracing. *Humanities and Social Sciences Communications*, 8(195). <https://doi.org/10.1057/s41599-021-00856-0>
- [4] Bannister-Tyrrell, M., Chen, M., Choi, V., Miglietta, A. in Galeac, G. (2023). Systematic scoping review of the implementation, adoption, use, and effectiveness of digital contact tracing interventions for COVID-19 in the Western Pacific Region. *The Lancet Regional Health – Western Pacific*, 2023(34). <https://doi.org/10.1016/j.lanwpc.2022.100647>
- [5] Blom, A. G., Wenz, A., Cornesse, C., Rettig, T., Fikel, M., Friedel, S., Möhring, K., Naumann, E., Reifenscheid, M. in Krieger, U. (2021). Barriers to the Large-Scale Adoption of a COVID-19 Contact Tracing App in Germany: Survey Study. *Journal of Medical Internet Research*, 23(3). <https://dx.doi.org/10.2196/23362>
- [6] Chen, A. T.-Y. in Thio, K. W. (2021). Exploring the drivers and barriers to uptake for digital contact tracing. *Social Sciences & Humanities Open*, 4(100212). <https://dx.doi.org/10.1016/j.ssho.2021.100212>
- [7] Dar, A. B., Lone, A. H., Zahoor, S., Khan, A. A. in Naaz, R. (2020). Applicability of mobile contact tracing in fighting pandemic (COVID-19): Issues, challenges and solutions. *Computer Science Review*, 38(100307). <https://dx.doi.org/10.1016/j.cosrev.2020.100307>
- [8] de Montjoye, Y.-A., Ramadorai, T., Valletti, T. in Walther, A. (2020). Privacy, adoption, and truthful reporting: A simple theory of contact tracing applications. *Economics Letters*, 198(109676). <https://dx.doi.org/10.1016/j.econlet.2020.109676>
- [9] Duan, S. X. in Deng, H. (2022). Exploring privacy paradox in contact tracing apps adoption. *Internet Research*, 32(5), str. 1725-1750. <https://doi.org/10.1108/INTR-03-2021-0160>
- [10] Eslami, P., Niakan Kalhori, S. R., in Taheriyani, M. (2021). eHealth solutions to fight against COVID-19: A scoping review of applications. *Medical Journal of the Islamic Republic of Iran (MJIRI)*, 35(43). <https://dx.doi.org/10.47176/mjiri.35.43>
- [11] European Commission (2021). Digital Contact Tracing Study: Study on lessons learned, best practices and epidemiological impact of the common European approach on digital contact tracing to combat and exit the COVID-19 pandemic.
- [12] Fox, G., Clohessy, T., van der Werff, L., Rosati, P. in Lynn, T. (2021). Exploring the competing influences of privacy concerns and positive beliefs on citizen acceptance of contact tracing mobile applications. *Computers in Human Behavior*, 121. <https://doi.org/10.1016/j.chb.2021.106806>
- [13] Gerli, P., Arakpogun, E. O., Elsahn, Z., Olan, F. in Prime, K. S. (2021). Beyond contact-tracing: The public value of eHealth application in a pandemic. *Government Information Quarterly*, 38(101581). <https://dx.doi.org/10.1016/j.giq.2021.101581>
- [14] Glowacz, F. in Schmits, E. (2020). Psychological distress during the COVID-19 lockdown: The young adults most at risk. *Psychiatry Research*, 293. <https://doi.org/10.1016/j.psychres.2020.113486>

- [15] Guazzini, A., Fiorenza, M., Panerai, G. in Duradoni, M. (2021). What Went Wrong? Predictors of Contact Tracing Adoption in Italy during COVID-19 Pandemic. *Future Internet*, 13(286). <https://doi.org/10.3390/fi13110286>
- [16] Harborth, D., Pape, S. in McKenzie, L. T. (2023). Why Individuals Do (Not) Use Contact Tracing Apps: A Health Belief Model Perspective on the German Corona-Warn-App. *Healthcare*, 11(583). <https://dx.doi.org/10.3390/healthcare11040583>
- [17] Huang Z., Guo H., Lim H.Y.-F., Chow A. (2022). Determinants of the acceptance and adoption of a digital contact tracing tool during the COVID-19 pandemic in Singapore. *Epidemiology and Infection* 150(54). <https://doi.org/10.1017/S0950268822000401>
- [18] Koivumäki, T., Pekkarinen, S., Lappi, M., Väisänen, J., Juntunen, J. in Pikkarainen, M. (2017). Consumer Adoption of Future MyData-Based Preventive eHealth Services: An Acceptance Model and Survey Study. *Journal of medical internet research*, 19(12). <https://dx.doi.org/10.2196/jmir.7821>
- [19] Megnin-Viggars, O., Carter, P., Melendez-Torres, G. J., Weston, D. in Rubin, G. J. (2020). Facilitators and barriers to engagement with contact tracing during infectious disease outbreaks: A rapid review of the evidence. *PLOS ONE*, 15(10). <https://doi.org/10.1371/journal.pone.0241473>
- [20] Nivette, A., Ribeaud, D., Murray, A., Steinhoff, A., Bechtiger, L., Hepp, U., Shanahan, L., Eisner, M. (2020). Non-compliance with COVID-19-related public health measures among young adults in Switzerland: Insights from a longitudinal cohort study. *Social Science & Medicine*, 268. <https://doi.org/10.1016/j.socscimed.2020.113370>
- [21] Oldeweme, A., Märtings, J., Westmattelmann, D. in Schewe, G. (2021). The Role of Transparency, Trust, and Social Influence on Uncertainty Reduction in Times of Pandemics: Empirical Study on the Adoption of COVID-19 Tracing App. *Journal of Medical Internet Research*, 23(2). <https://doi.org/10.2196/25893>
- [22] Parker, M. J., Fraser, C., Abeler-Dörner, L. in Bonsall, D. (2020). Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic. *J Med Ethics*, 46, str. 427-431. <https://dx.doi.org/10.1136/medethics-2020-106314>
- [23] Prakash, A. V. in Das, S. (2022). Explaining citizens' resistance to use digital contact tracing apps: A mixed-methods study. *International Journal of Information Management*, 63. <https://dx.doi.org/10.1016/j.ijinfomgt.2021.102468>
- [24] Ranisch, R., Nijsingh, N., Ballantyne, A., van Bergen, A., Buyx, A., Friedrich, O., Hendl, T., Marckmann, G., Munthe, C. in Wild, V. (2021). Digital contact tracing and exposure notification: ethical guidance for trustworthy pandemic management. *Ethics and Information Technology*, 23, str. 285-294. <https://dx.doi.org/10.1007/s10676-020-09566-8>
- [25] Ross, G. M. (2021). I use a COVID-19 contact-tracing app. Do you? Regulatory focus and the intention to engage with contact-tracing technology. *International Journal of Information Management Data Insights*, 1. <https://dx.doi.org/10.1016/j.ijimedi.2021.100045>
- [26] Shahroz, M., Ahmad, F., Younis, M. S., Ahmad, N., Boulos, M. N. K., Vinuesa, R. in Qadir, J. (2021). COVID-19 digital contact tracing applications and techniques: A review post initial deployments. *Transportation Engineering*, 5(10072). <https://dx.doi.org/10.1016/j.treng.2021.100072>
- [27] Stanimirović, D. in Rant, Ž. (2022). Vloga digitalizacije pri obvladovanju javnozdravstvenih groženj: skokovita rast uporabe rešitev e-Zdravja v obdobju pandemije COVID-19 v Sloveniji. V Gabrovec, B., Eržen, I., Trop Skaza, A., Fafangel, M., Vrdelja, M. in Selak, Š. (ur.) *Javno zdravje in COVID-19: Zbornik povzetkov in recenziranih prispevkov*, 2. znanstvena in strokovna konferenca (str. 151-158). NIJZ.
- [28] Tomczyk, S., Barth, S., Schmidt, S. in Muehlan, H. (2021). Utilizing Health Behavior Change and Technology Acceptance Models to Predict the Adoption of COVID-19 Contact Tracing Apps: Cross-sectional Survey Study. *Journal of Medical Internet Research*, 23(5). <https://dx.doi.org/10.2196/25447>
- [29] Touzani, R., Schultz, E., Vandentorren, S., Arwidson, P., Guillemain, F., Bouhnik, A.-D., Rouquette, A. in Mancini, J. (2023). Digital contact tracing during the COVID-19 pandemic in France: Associated factors and reasons for non-use. *International Journal of Medical Informatics*, 171 (2023). <https://doi.org/10.1016/j.ijmedinf.2023.104994>
- [30] Trkman, M., Popovič, A. in Trkman P. (2023). The roles of privacy concerns and trust in voluntary use of governmental proximity tracing applications. *Government Information Quarterly* 40(101787). <https://dx.doi.org/10.1016/j.giq.2022.101787>
- [31] van der Waal, N. E., de Wit, J., Bol, N., Ebbers, W., Hooff, L., Metting, E., van der Laan, L. N. (2022). Predictors of contact tracing app adoption: Integrating the UTAUT, HBM and contextual factors. *Technology in Society*, 71. <https://doi.org/10.1016/j.techsoc.2022.102101>
- [32] von Wyl, V., Höglinger, M., Sieber, C., Kaufmann, M., Moser, A., Serra-Burriel, M., Ballouz, T., Menges, D., Frei, A. in Puhan, M. A. (2021). Drivers of Acceptance of COVID-19 Proximity Tracing Apps in Switzerland: Panel Survey Analysis. *JMIR Public Health Surveillance*, 7(1). <https://dx.doi.org/10.2196/25701>
- [33] Walrave, M., Waeterloos, C. in Ponnet, K. (2022). Reasons for Nonuse, Discontinuation of Use, and Acceptance of Additional Functionalities of a COVID-19 Contact Tracing App: Cross-sectional Survey Study. *JMIR Public Health and Surveillance*, 8(1). <https://doi.org/10.2196/22113>
- [34] World Health Organization (WHO) (2020). Digital tools for COVID-19 contact tracing. Annex: Contact tracing in the context of COVID-19.
- [35] Zetterholm, M. V., Lin, Y. in Jokela, P. (2021). Digital Contact Tracing Applications during COVID-19: A Scoping Review about Public Acceptance. *Informatics*, 8(48). <https://dx.doi.org/10.3390/informatics8030048>
- [36] Oldeweme, A., Märtings, J., Westmattelmann, D. in Schewe, G. (2021). The Role of Transparency, Trust, and Social Influence on Uncertainty Reduction in Times of Pandemics: Empirical Study on the Adoption of COVID-19 Tracing App. *Journal of Medical Internet Research*, 23(2). <https://doi.org/10.2196/25893>

Lana Masnec je magistrica družboslovne informatike. Magistrirala je na Fakulteti za družbene vede na Univerzi v Ljubljani, z delom Dejavniki vpliva na aktivno uporabo aplikacije za sledenje stikov med mladimi: študija primera Nemčije. Osredotoča se na področje privzemanja in oblikovanja digitalnih storitev in je zaposlena kot freelance oblikovalka uporabniške izkušnje.

Doc. dr. **Marina Trkman** je zaposlena na Fakulteti za upravo in Fakulteti za družbene vede Univerze v Ljubljani.. Doktorat je opravila na Fakulteti za računalništvo in informatiko UL. Svoj podoktorski projekt je vodila na Institutu Jožefa Štefana. Raziskovalno se ukvarja: 1) s preiskovanjem dejavnikov, ki vplivajo na privzemanje informacijsko komunikacijskih tehnologij in 2) z analiziranjem poslovnih procesov za optimizacijo poslovanja. Kot prvi avtor je objavila članke v prestižnih revijah kot sta *International journal of information management* in pa *Government information quarterly*.

Premikamo meje za bolnike.

Smo Sandoz,
vodilno farmacevtsko
podjetje v svetu za generična
in podobna biološka zdravila.
In smo Lek, pionirji farmacevtske industrije
v Sloveniji.

Naša strast so odličnost in vrhunska kakovost zdravil.
Navdušujejo nas biotehnoški postopki za razvoj in
proizvodnjo podobnih bioloških zdravil ter najvišji standardi
farmacevtske proizvodnje.

SANDOZ



Lek farmacevtska družba d. d.
Verovškova ulica 57
1526 Ljubljana, Slovenija
www.lek.si

Metodologije za kvalitativno vrednotenje kakovosti odprtih podatkov

Klara Žnideršič¹, Matija Marolt¹, Aleš Veršič², Matevž Pesek¹

¹Fakulteta za računalništvo in informatiko Univerze v Ljubljani, Večna pot 113, Ljubljana

²Ministrstvo za digitalno preobrazbo

klara.znidersic@fri.uni-lj.si, matija.marolt@fri.uni-lj.si, aversic@gov.si, matevz.pesek@fri.uni-lj.si

Izveček

Članek predstavlja pregled metodologij za kvalitativno vrednotenje odprtih podatkov. Različni pristopi so bili razviti tako za ocenjevanje metapodatkov, ki opisujejo posamezne zbirke podatkov, kot tudi za ocenjevanje samih podatkov. Metodologije pogosto temeljijo na standardih in širše sprejetih smernicah za zagotavljanje kakovosti podatkov, posamezne vidike (meta)podatkov pa obravnavajo kot dimenzije, združene v kategorije. Slednje med različnimi metodologijami niso poenotene, kar otežuje primerjavo ali združevanje rezultatov, vendar poznavanje različnih pristopov pomembno pripomore pri razvoju novih, učinkovitejših in bolj interoperabilnih rešitev. V primeru razvoja novih orodij za vrednotenje odprtih vladnih podatkov v Sloveniji je potrebno upoštevati še dodatna priporočila. Trg odprtih podatkov v državah Evropske unije namreč regulirajo zakonodaje in smernice tako na nacionalni, kot na ravni Unije. V tem prispevku predstavljamo pregled smernic in standardov ter izpostavljammo izzive za izdelavo poenotene metodologije, ki bi pripomogla k dvigu ravni kakovosti in dostopnosti odprtih podatkov.

Ključne besede: metodologije kvalitativnega vrednotenja, odprti podatki, pregled področja

Methodologies for qualitative assessment of open data

Abstract

The article provides an insight into the methodologies for the qualitative assessment of open data. Various approaches have been developed both for the assessment of metadata and for the data itself. The methods are often based on standards and generally accepted guidelines for ensuring data quality, whereby individual aspects of the (meta)data are treated as dimensions grouped into categories. However, these categories are not standardized across different methodologies, making it difficult to compare or integrate the results. Nevertheless, understanding different approaches contributes significantly to the development of more effective and interoperable solutions. When developing new tools for open government data in Slovenia, additional recommendations need to be considered. Laws and directives regulate the open data market in European Union countries at both national and EU level. We provide an overview of guidelines and standards, and highlight the challenges in creating a uniform methodology to improve the quality and accessibility of open data.

Keywords: quality assessment methodologies, open data, overview of the field

1 UVOD

Pojem odprti podatki predstavlja podatke, do katerih lahko vsakdo svobodno dostopa, jih uporablja, modificira in deli za kakršenkoli namen (»Open Definition 2.1«, 2015). Ocena skupne vrednosti trga odprtih podatkov v Evropski uniji z vsakim letom izrazito narašča, o čemer poroča tudi najnovejša uradna študija evropskega podatkovnega trga, ki napoveduje,

da bo trg do leta 2030 presegel vrednost 118 milijard €. To rast lahko razumemo kot posledico kopičenja podatkov na različnih evropskih podatkovnih portalih, s tem pa se hkrati dvigujejo tudi pričakovanja in zahteve potencialnih uporabnikov ponujenih podatkov. Podjetja, ki bi z rabo odprtih podatkov spodbudila inovacije in ekonomsko rast, se za to pogosto ne odločajo, saj se kljub navidez preprosti uporabi

brezplačno dostopnih odprtih podatkov neredko srečujejo z znatnimi ovirami pri identificiranju primernih nizov podatkov in pripravi le-teh za uporabo, ovire pa se pogosto navezujejo na kakovost podatkov (Krasikov & Legner, 2023). Raziskave preteklih let, ki naslavljajo to problematiko, so predstavile že več različnih pristopov k bolj in manj celostnem vrednotenju kakovosti zbirk, vendar se med področjem metapodatkov in podatkov pojavlja vrzel, ki ni zanemarljiva. Medtem ko je za vrednotenje metapodatkov, ki opisujejo objavljene zbirke podatkov, pripravljenih več uspešno evalviranih metodologij, je vrednotenje podatkov kompleksnejše narave in so zato metode pogosteje prilagojene specifičnim podatkovnim zbirkam.

V pričujočem članku so predstavljeni glavni izsledki pregleda področja kvalitativnega vrednotenja (meta)podatkov, ki so pomembni za razvoj novih metodologij ter vzpostavitev čim bolj celostnega in prilagodljivega pristopa k podajanju vrednostne ocene podatkovnih zbirk, predvsem v kontekstu odprtih podatkov. Cilj prispevka je pregled metodologij, ki bodo služile kot osnova za izdelavo celostne metodologije kot rezultat projekta *Izdelava metodologije za določanje kakovosti podatkov ter ocena kakovosti posameznih podatkovnih zbirk na nacionalnem portalu odprtih podatkov Slovenije - portalu OPSI*.

2 TEMELJI ZA ZAGOTAVLJANJE KAKOVOSTI PODATKOV

Dolgoletna zgodovina zbiranja podatkov in vse bolj zgoščenega kopičenja informacij izpostavlja velik pomen zagotavljanja in vzdrževanja kar najvišje stopnje kakovosti podatkov. S tem namenom so bili razviti standardi in druga pravila, ki ponudnikom podatkov predstavljajo smernice za doseganje kakovosti in so usmerjena v doseganje optimalne stopnje urejenosti na danem področju.

2.1 Standard in aplikacijski profil evropskih podatkovnih portalov

Besednjak podatkovnih katalogov DCAT («Data Catalog Vocabulary - Version 2», 2020) za podatkovni model RDF (Resource Description Framework oz. ogrodje za opis virov), razvit v okviru organizacije W3C, ponuja smernice in specifikacije za opisovanje naborov podatkov v katalogih podatkov z namenom izboljšanja prepoznavnosti in interoperabilnosti naborov podatkov na spletu. DCAT je standard, ki se lahko uporablja globalno in ni povezan z nobenim regulativnim okvirjem, a je za specifične potrebe različnih podatkovnih portalov lahko dopolnjen z aplikacijskimi profili. Kot specifikacija za opis povezanih javnih podatkov v Evropi je v uporabi aplikacijski profil za evropske podatkovne portale DCAT-AP (Slika 1). Aplikacijski profil DCAT-AP, ki je bil zasnovan za rabo v evropskih javnih upravah, ponuja posebne smernice za zajemanje pravnih in licenčnih informacij v zvezi z nabori podatkov, ki so v skladu z evropskimi pravnimi zahtevami za odprte podatke ter zaradi svoje specializacije pripomore k večji interoperabilnosti med državami članicami EU.

Podatkovni model RDF za predstavitev informacije uporablja trojice subjekt–predikat–objekt, pri čemer se za opis podatkovnih katalogov uporablja nabor razredov in lastnosti, ki jih definira DCAT.

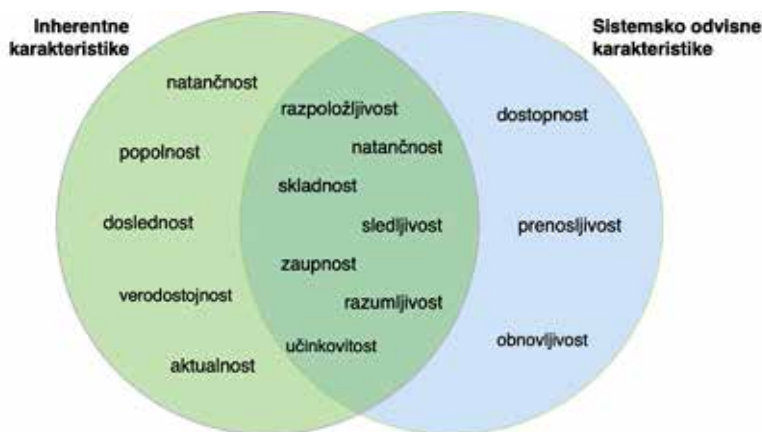
2.2 Standard ISO 25012

Serijski standardov Mednarodne organizacije za standardizacijo (ISO), katerih glavni cilj je usmerjati razvoj programskih izdelkov s specifikacijo kakovostnih zahtev in kriterijev za evalvacijo, je združena pod kratico SQuaRE (System and Software Requirements and Evaluation). Za podatke, shranjene v strukturirani obliki v računalniškem sistemu, je splošni model kakovosti podatkov opredeljen s standardom ISO/IEC 25012:2008. Uporablja se lahko za

```
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix example-ds: <https://data.gov.gr/id/dataset/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
```

```
example-ds:BeePopulation a dcat:Dataset;
dct:title "Bee population"@en;
dct:description "A dataset about bee population in Greece"@en;
dct:publisher <https://agencies.gov.gr/id/GreekEnvironmentAgency> .
```

Slika 1: Primer DCAT-AP opisa podatkovne zbirke o populaciji čebel (Van Nuffelen, 2023)



Slika 2: Tabela karakteristik kvalitete podatkov ISO/IEC 25012 (»ISO 25012«, 2022)

opredeljevanje kvalitativnih meril ali načrtovanje in izvajanje ocenjevanja kakovosti podatkov. Kakovost podatkovnega izdelka je lahko interpretirana kot stopnja, do katere podatki izpolnjujejo zahteve petnajstih karakteristik modela, te pa so razvrščene v dve glavni kategoriji - inherentna in sistemsko odvisna kakovost podatkov (Slika 2). Inherentna kakovost podatkov se nanaša zlasti na podatkovne vrednosti, razmerja med temi vrednostmi ter metapodatke, kategorija sistemsko odvisnih kakovosti podatkov pa določa stopnjo, do katere je kakovost dosežena in ohranjena v računalniškem sistemu. V tej kategoriji na kakovost vplivajo zmogljivosti komponent računalniških sistemov (strojna in programska oprema).

2.3 5-zvezdični sistem za ocenjevanje

Leta 2006 je ustanovitelj svetovnega spleta Tim Berners-Lee predstavil sistem za ocenjevanje povezanih podatkov (Berners-Lee, b.d.), s katerim je želel spodbuditi predvsem upravljavce vladnih podatkov k dvigu kakovosti podatkov in boljšo povezanost. Sistem temelji na standardnih spletnih tehnologijah, kot so HTTP, RDF in URI, ki jih uporablja za deljenje informacij na strojno berljivi način in s tem omogoča povezovanje podatkov iz različnih virov ter razširja možnosti semantičnih poizvedb. Za vpeljavo odprtih podatkov je Berners-Lee predlagal petzvezdično shemo, pri kateri je zadostovanje pogojem vsake predhodne stopnje predpogoj za izpolnjevanje zahtev naslednje stopnje:

1. Podatki so dostopni na spletu v formatu z odprto licenco.

2. Podatki so na voljo v strojno berljivi obliki (npr. tekstovna datoteka namesto digitalno preslikane razpredelnice).
3. Podatki so dostopni v nelastniškem formatu (npr. CSV namesto Excel).
4. Za identificiranje so uporabljeni odprti standardi organizacije W3C (RDF in SPARQL) ter URI.
5. Podatki so za namen kontekstualizacije povezani z ostalimi dostopnimi podatki.

2.4 Načela FAIR

Potreba po izboljšavi infrastrukture, ki podpira ponovno uporabo znanstvenih podatkov, je bila naslovljena z načeli FAIR (»GO FAIR: FAIR Principles«, b. d.), ki poseben poudarek usmerjajo v izboljšanje strojnih zmogljivosti za samodejno iskanje in uporabo podatkov. Utemeljena so na štirih osrednjih kriterijih – najdljivosti (F - findability), dosegljivosti (A - accessibility), interoperabilnosti (I - interoperability) in ponovni uporabnosti (R - reusability). FAIR opisuje jedrnata, od področja neodvisna načela, ki jih je mogoče uporabiti za širok spekter znanstvenih rezultatov.

3 PRISTOPI H KVALITATIVNEMU VREDNOTENJU (META)PODATKOV

V začetku je potrebno izpostaviti pomen razlikovanja med izrazoma profiliranje podatkov in ocenjevanje kakovosti podatkov. Medtem ko je prvo zgolj sistematizirano zbiranje informacij o podatkih, drugo predstavlja preverjanje ustreznosti podatkov določenim kakovostnim merilom. Profiliranje torej lahko razumemo kot komplementarno zgodnjo fazo vrednotenja podatkov (Debattista in sod., 2016).

3.1 Vrednotenje metapodatkov

Velika večina raziskav o kvalitativnem vrednotenju podatkovnih zbirk je bila osredotočena na vrednotenje metapodatkov. Različne metode in kazalniki za oceno kakovosti metapodatkov so bili preizkušeni na številnih zbirkah podatkov, med drugim tudi na vladnih portalih odprtih podatkov. Metode so pogosto razčlenjene tako, da obravnavajo posamezne aspekte podatkov - dimenzije, ki so razdeljene v kategorije. V posamezno kategorijo so zbrane kvalitativne dimenzije, v katerih je kot kazalnik uporabljena skupna vrsta informacije (npr. dostopnost, ki vključuje dimenzije razpoložljivosti, varnosti in delovanja), saj takšno združevanje pri velikem številu dimenzij omogoča boljši pregled vseh aspektov kvalitete. Posamezna dimenzija ima lahko eno ali več metrik, t.j. konkretnih meril kakovosti za posamezni kazalnik, ki so običajno povezani z merilnim postopkom in vrnejo numerično ali boolean oceno.

Ena obsežnejših metapodatkovnih evalvacij preteklih let je bila izvedena v okviru projekta *Open Data Portal Watch* (Neumaier in sod., 2016), kjer so predstavili orodje za avtomatizirano vrednotenje kvalitete metapodatkov s pretvorbo v standardizirano shemo DCAT. Na podlagi izbranih 29 dimenzij iz besednjaka DCAT in izračunu metrik za posamezno dimenzijo je bilo ovrednotenih preko milijon podatkovnih zbirk, kar je služilo tudi nadzoru in razvrščanju portalov. Nadgradnjo projekta je predstavljala implementacija analitičnega hierarhičnega procesa, ki je omogočil primerjavo preko 250 portalov, na katerih so bile objavljene ovrednotene zbirke (Kubler in sod., 2018).

Po številu primerljiv obseg zbirk dosega tudi Uradni portal za evropske podatke, ki za preučevanje kakovosti metapodatkov objavljenih zbirk uporablja orodje Metadata Quality Assessment (MQA). Metrike tega orodja so predstavljene v Tabeli 1. MQA naslavlja osnovni vprašnji o kakovosti metapodatkov za podatke iz javnega sektorja v vseevropski regiji in identificira največje ovire pri doseganju boljše kakovosti. Poleg ocene kakovosti metapodatkov zbirk javnega sektorja v vseevropski regiji, MQA prav tako pomaga pri iskanju preprek pri doseganju višje kakovosti. Metodologija MQA preverja:

- skladnost z aplikacijskim profilom DCAT in derivati aplikacijskega profila DCAT;
- razkritje informacij, ki niso obvezne za aplikacijski profil DCAT;

- dostopnost podatkov, na katere se sklicujejo metapodatki prek URL povezave za dostop in URL povezave za prenos;
- strojno berljivost referenčnih podatkov;
- uporabo licenc.

Proces je omejen z metapodatki, ki jih orodje lahko pregleda in ki so dostopni na portalu data.europa.eu. V izogib prenosu napak iz izvirnih metapodatkov v proces vrednotenja, MQA že pred objavo zbirk zagotavlja storitev, ki jo lahko uporabijo ponudniki podatkov za potrjevanje veljavnih formatov in skladnost z DCAT-AP, preden svoje metapodatke vključijo v postopek vrednotenja. Integriran je tudi mehanizem za ponovno preverjanje dostopnosti objavljenih spletnih povezav po določenem časovnem obdobju. V postopku vrednotenja je upoštevanih pet dimenzij – štiri iz principov FAIR ter dimenzija kontekstualnosti – z vnaprej določenimi metriki in največjim možnim številom točk, ki so pretvorjene v opisne ocene (odlično, dobro, zadostno ali slabo). V večini primerov metodologija ocenjuje zgolj Boolovo vrednost metrike (da/ne).

Pogosteje so bile izvedene manj obsežne raziskave, ki so podlago za izbiro dimenzij poiskale v drugih smernicah, a so kljub temu sorodne zgoraj opisanim. Primer je Bogdanović-Dinić in sod. (2014), ki je 7 portalov odprtih podatkov ovrednotila na podlagi 8 principov (»The 8 Principles of Open Government Data«, b.d.): odprtost (opis, možnost prenosa, strojna berljivost, povezljivost), primarnost (podatki v izvorni obliki), pravočasnost (časovno obdobje, frekvenca posodobitev, zadnja posodobitev), dostopnost, možnost strojne obdelave (PDF/XLS, CSV/HTML/XTT, XML/RDF), nediskriminatornost dostopa, objava v nelastniških formatih, brez licence oz. licenca za odprti dostop.

3.2 Vrednotenje kakovosti podatkov v podatkovnih zbirkah

Izrazito manj zastopane so študije kakovosti dejanske vsebine podatkovnih zbirk. Teoretično izhodišče tem raziskavam je v več primerih predstavljalo znanstveno delo *Data Quality: Concepts, Methodologies and Techniques* (Batini & Scannapieca, 2006). Principi, ki jih obravnava delo, so formulirani neodvisno od konteksta uporabe in med raziskovalci zanimivi zaradi svoje univerzalnosti, predvsem zato, ker je v novejši literaturi pogosteje izpostavljeno vrednotenje kako-

Tabela 1: Metrike v posameznih kategorijah MQA dimenzij

Dimenzija	Metrika
Najdljivost	ključne besede kategorije geografsko iskanje časovno iskanje
Dostopnost	dostopnost URL-ja za dostop URL za prenos dostopnost URL-ja za prenos
Interoperabilnost	format vrsta medija vrsta medija/formata iz slovarja strojna berljivost skladnost z DCAT-AP
Ponovna uporabnost	licenca slovar licence omejitve dostopa slovar omejitev dostopa kontaktna točka izdajatelj
Kontekstualnost	pravice velikost datoteke datum izdaje datum spremembe

vosti, prilagojeno specifičnim kontekstom in nadaljnji rabi. Dimenzije kakovosti, definirane v knjigi, so povzele številne metodologije, ki so nabor dopolnile glede na potrebe raziskav in razvoj tehnologij. Knjiga ob tem ponuja tudi praktične rešitve in metodologije za lokalizacijo in odpravljanje podatkovnih napak, identifikacijo objektov ter integracijo podatkov.

Med raziskavami velja izpostaviti nabor dimenzij kakovosti, ki je nastal na podlagi pregleda 30 različnih pristopov in 12 orodij (Zaveri in sod., 2016). Za vsako izmed 18 dimenzij je ponujena natančna definicija, več metrik (ki so opredeljene kot kakovostne ali kvantitativne) in primer uporabe. Kvalitativna oz. kvantitativna narava metrike je pomembna predvsem za proces avtomatizacije vrednotenja, saj lahko slednja zagotavlja zanesljivejše rezultate pri objektivnih kvantitativnih metrikih (npr. št. sintaktično pravih vrednosti), medtem ko so kvalitativne ocene lahko odvisne od subjektivne ocenjevalčeve percepcije (npr. ugled podatkovne zbirke).

Na podlagi sorodne metodologije LANG (Zhang in sod., 2019) je bil zasnovan dvostopenjski model za vrednotenje sintaktičnih (enolična razpoznavnost podatkov, konsistentnost formata, referenčna integriteta, skladnost metapodatkov in skladnost s poslovnimi pravili) in semantičnih (popolnost atributov, semantična konsistentnost, konsistentnost vrednosti, natančnost in neredundanca/nepodvojenost) vidikov kvalitete. Rezultati evalvacije na več kot 20 podatkovnih zbirkah so prikazali potencial, ki ga ima

LANG za pospeševanje in razširjanje procesov raziskovanja podatkov, študija možnosti za avtomatizacijo tega procesa pa je dala rezultate, skoraj povsem enake rezultatom ročnega vrednotenja.

sSQuaRE-Aligned Portal Data Quality Model (SPDQM) (Moraga in sod., 2009) je bil utemeljen na 30 karakteristikah starejšega modela PDQM in dopolnjen s petimi karakteristikami, izbranimi na podlagi dodatnega sistematičnega pregleda literature ter sedmimi ISO standarda, ki še niso bili del modela. Navezuje se tako na podatke same kot tudi na kakovost sistema, saj obravnava vlogo sistema, kontekst in reprezentacijo podatkov. Za vrednotenje posamezne kategorije so bile določene naslednje karakteristike:

- Inherentna kakovost: natančnost, kredibilnost (objektivnost, ugled), sledljivost, aktualnost, zastaranost, popolnost, doslednost, dostopnost, skladnost, zaupnost, učinkovitost, natančnost, razumljivost;
- Operativna kakovost: razpoložljivost, dostopnost (interaktivnost, enostavnost uporabe in pomoč strankam), preverljivost, zaupnost, prenosljivost, popravljivost;
- Kontekstualna kakovost: veljavnost (zanesljivost, obseg), dodana vrednost (aplikabilnost, prilagodljivost, novost), relevantnost (novost, pravočasnost), specializacija, uporabnost, sledljivost, skladnost, preciznost;
- Kakovost predstavitev: jedrnatost, doslednost, razumljivost (interpretabilnost, količina podatkov, dokumentacija, organizacija), privlačnost, berljivost, učinkovitost, uspešnost.

Podjetniški vidik uporabe odprtih podatkov je naslavljala novejša študija (Krasikov & Legner, 2023), ki se je osredotočila na vprašanje, kako podjetjem pomagati pri sistematičnem pregledu, oceni in pripravi odprtih podatkov za uporabo. Pri iterativnem razvoju metode so avtorji sodelovali s podjetji, končni izdelek pa omogoča sistematično analizo ter integracijo odprtih podatkov in s tem pojmuje pripravo odprtih podatkov kot smiselni proces ustvarjanja dodane vrednosti. Tristopenjski pristop k ocenjevanju kakovosti je zajemal

1. pregled metapodatkov po predlogi starejših raziskav;
2. oceno popolnosti sheme (prisotnost potrebnih atributov);

Dimenzija	Definicija
Aktualnost	Kdaj so bili podatki vneseni v vir in/ali podatkovno skladišče.
Dodana vrednost	V kolikšni meri so podatki koristni in zagotavljajo prednosti zaradi njihove uporabe.
Doslednost	V kolikšni meri so podatki predstavljeni v enaki obliki.
Dosegljivost	V kolikšni meri so podatki na voljo oz. jih je mogoče enostavno in hitro pridobiti.
Enostavnost uporabe in vzdrževanja	V kolikšni meri je mogoče podatke uporabljati, posodobljati, vzdrževati in upravljati.
Jedrnatost	V kolikšni meri so podatki jedrnato predstavljeni.
Kredibilnost	V kolikšni meri se podatki štejejo za resnične in verodostojne.
Natančnost	V kolikšni meri so podatki pravilni, zanesljivi in zagotovljeno brez napak.
Objektivnost	V kolikšni meri so podatki nepristranski.
Popolnost	V kolikšni meri podatki ne manjkajo ter so dovolj obsežni in poglobljeni za zadevno nalogo.
Pravilnost	V kolikšni meri so podatki pravilni in zanesljivi.
Pravočasnost	V kolikšni meri so podatki dovolj posodobljeni za zadevno nalogo.
Razločljivost	V kolikšni meri so podatki v ustreznih jezikih, simbolih in enotah, opredelitve pa so jasne.
Razumljivost	V kolikšni meri so podatki jasni, brez dvoumnosti in zlahka razumljivi.
Relevantnost	V kolikšni meri so podatki uporabni in koristni za zadevno nalogo.
Sledljivost	V kolikšni meri so podatki dobro dokumentirani, preverljivi in zlahka pripisani viru.
Spremenljivost	Časovno obdobje, v katerem so informacije veljavne v resničnem svetu.
Učinkovitost	V kolikšni meri lahko podatki hitro zadovoljijo potrebe po informacijah za zadevno nalogo.
Ugled	V kolikšni meri so podatki visoko cenjeni glede na njihov vir ali vsebino.
Uporabnost	V kolikšni meri so podatki jasni in enostavni za uporabo.
Ustreznost količina podatkov	V kolikšni meri je obseg podatkov primeren za zadevno nalogo.
Varnost	V kolikšni meri je dostop do podatkov ustrezno omejen, da je zagotovljena njihova varnost.

Tabela 2: Definicije pogosto uporabljenih dimenzij

3 na nivoju vsebine uporabo tradicionalnih meril za kvaliteto podatkov (popolnost, edinstvenost, veljavnost).

Kvalitativne dimenzije in merila za metapodatke, shemo in podatke so bile utemeljene na starejših raziskavah (Neumaier in sod., 2016; Vetrò in sod., 2016; Zhang in sod., 2019), definicije nekaterih pogosto uporabljenih dimenzij so zbrane v Tabeli 2, kot nezamisljiva faza procesa pa je izpostavljena izdelava dokumentacije podatkovnih zbirk in integracija odprtih podatkov v podatkovne zbirke znotraj podjetja. Proces natančne dokumentacije je poenostavljen z vse bolj razširjeno rabo standardiziranih RDF slovarjev, kar pozitivno vpliva tudi na nadaljnje povezovanje zbirk.

Posebno področje kvalitativnega vrednotenja podatkov predstavlja ocena primernosti za uporabo (ang. *fitness for use*), ki so ga obravnavali avtorji odprtokodnega orodja Luzzu (Debattista in sod., 2016). Pristop za obteženo razvrščanje naborov podatkov temelji na vrednotenju podatkov, pri čemer so dimenzije za vrednotenje večinsko povzete po Zaveri in sod. (2016), lasten domenskospesificni jezik pa omogoča tudi definiranje metrik brez poznavanja programskih jezikov. Utež ocen posamezne metrike (ali dimenzije ali kategorije) lahko uporabnik določi glede na potrebe, saj je pripomoček primarno name-

njen nadaljnjim uporabnikom, ki lahko s filtriranjem in razvrščanjem podatkov, ovrednotenih po izbranih merilih, poiščejo najprimernejši nabor podatkov. Pri procesiranju obsežnih zbirk podatkov so uporabljene verjetnostne metode za aproksimacijo, kar omogoča, da orodje potencialnim uporabnikom v doglednem času ponudi rezultate, ki so dovolj natančni za večino primerov uporabe.

3.3 Izzivi pri vrednotenju kakovosti podatkov

Pogosta opazka v raziskavah, ki obravnavajo kvaliteto podatkov ali metapodatkov, je problematika definicije *dobrih* podatkov. Presoja o kvaliteti je namreč pogosto odvisna od namena njihove uporabe ali zahtev uporabnikov in ponudnikov podatkov. Pri meritvah kakovosti dodatno nastajajo ovire zaradi pomanjkljivih, nenatančnih in nepopolnih informacij, zastarelih ali neveljavnih podatkov, nejasnih vrednosti, prevelike količine informacij za procesiranje, odsotnosti bistvenih podatkov in enake vsebine, ki v različnih sistemih ponuja različne rezultate (Jansen & Kronenburg, 2012; Krasikov & Legner, 2023). Tehnične ovire se pojavljajo pri nejasnih licenčnih pogojih, neprimerno opredeljenem ali težko dostopnem formatu, odsotnosti standardov, zastarelih sistemih za objavo podatkov, pomanjkanju standardne programske opreme za obdelavo odprtih podatkov ter razdrobljenosti programske opreme in aplikacij.

Na področju metapodatkov so se težave pojavljale tudi pri avtomatskih DCAT pretvorbah, nedelujočih HTTP zahtevah, razlikovanju med splošno definicijo in dejanskim stanjem odprtih podatkov ter nedostopnosti strojno berljivih kontaktnih podatkov skrbnikov zbirk.

Malo pristopov k vrednotenju podatkovnih zbirk je bilo dejansko podprtih s konkretnim orodjem, toda v večini teh primerov izvorna koda, ki bi bila uporabna na drugih podatkovnih zbirkah, ni dostopna. Izjema, ki jo velja izpostaviti, je Luzzu, dostopen v odprtokodni različici, vključno z dodatnimi zunanjimi metrikami.

Tudi priljubljena petzvezdična lestvica kakovosti Tima Bernersa-Leeja predstavlja težavo, saj pokriva zgolj specifičen aspekt kvalitete podatkov, kar je format oz. kodiranje, ki je v rabi za objavo podatkov. Zbirka podatkov na lestvici lahko doseže najvišjo oceno in kljub temu z drugih vidikov ne vključuje kakovostnih podatkov. Kot primer navajajo natančnost oz. nenatančnost pri ročnem vnašanju podatkov ali napake v programski opremi, pa tudi težave z vidika popolnosti, konsistentnosti podatkov in pravočasnosti objave.

Neenotnost definicij različnih kvalitativnih dimenzij se je izkazala za problematično v večih raziskavah metodologij vrednotenja kakovosti podatkov (Vetrò in sod., 2016; Zuiderwijk in sod., 2021). Zaradi različnih definicij karakteristik podatkov so bila posledično različno določena tudi merila, kar je pri uporabi metodologij na istih zbirkah podatkov dalo različne rezultate. Zuiderwijk in sod. (2021) so pri raziskovanju meril kakovosti, na podlagi katerih so ustvarjene lestvice kakovosti podatkov različnih ponudnikov (tj. držav), odkrili, da se pri uporabi raznih metodologij enake zbirke podatkov na lestvicah različno uvrščajo.

4 VLADNI ODPRTI PODATKI

Trg vladnih odprtih podatkov na območju Evropske unije urejajo evropske direktive in državne zakonodaje, EU pa je zaradi strme rasti količine podatkov v okviru Uradnega portala za evropske podatke podprla tudi izvedbo številnih študij na več nivojih podatkovne krajine. Njihovi rezultati sodelujočim državam predstavljajo spodbudo in zagotavljajo osnovo za izboljšavo, popravilo in razvoj nacionalnih portalov in zbirk odprtih podatkov. Smernice v splošnem spodbujajo čim širšo dostopnost podatkov

in nadaljnjo rabo, evropski podatkovni portal pa visoko kvaliteto pripisuje podatkom, ki so človeško in strojno enostavno berljivi ter dosegaajo čim višjo vrednost na petzvezdični ocenjevalni lestvici. Direktiva (EU) 2019/1024 predstavlja ključni dokument Evropske unije, ki obravnava področje odprtih podatkov in ponovne uporabe informacij javnega sektorja. S tem dokumentom so bile uvedene smernice in pravila, ki spodbujajo prost dostop do podatkov, ki jih proizvaja javni sektor, ter njihovo ponovno uporabo v gospodarske in družbene namene. Hkrati spodbuja izmenjavo informacij med državami članicami EU, združljivost z načeli FAIR ter prispeva k razvoju in inovacijam v digitalni družbi. Od leta 2015 izhaja tudi vsakoletno poročilo o zrelosti podatkov na nacionalnih portalih odprtih podatkov, ki je osredotočeno na štiri glavna področja – državno zakonodajo, meritve vpliva ponovne uporabe podatkov, oceno nacionalnih portalov in mehanizme za kvaliteto (meta)podatkov.

V slovenski zakonodaji se na temo odprtih podatkov nanašata Zakon o dostopu do informacij javnega značaja in Uredba o posredovanju in ponovni uporabi informacij javnega značaja, ki prenašata evropsko direktivo na pravni red Republike Slovenije. Z zakonom je urejena definicija relevantnih pojmov in postopek, ki omogoča prost dostop in ponovno uporabo informacij javnega značaja ter obvezuje organe k objavi informacij na namenskem portalu.

5 ZAKLJUČEK

Vpogled v stanje na področju kvalitativnega vrednotenja podatkov predstavlja najvidnejše standarde, definicije in načela za delo z (odprtimi) podatki, ki so pomembni gradniki pri vzpostavitvi metodologij vrednotenja, saj uporaba poenotenih pravil predstavlja predpogoj za zanesljive in primerljive rezultate. Kljub temu pa zgolj uporaba opisanih osnov ne zadoštuje za kakovosten proces vrednotenja in je za to potreben tudi podrobnejši razmislek o kriterijih, ki naj bi jih upoštevala generalizirana metodologija za vrednotenje odprtih podatkov. Pretekle raziskave so v večjem delu obravnavale vrednotenje metapodatkov, saj je to področje bolj standardizirano (npr. uporaba RDF slovarjev in različnih lestvic kriterijev), med tem ko je zaradi raznolikosti vsebine podatkovnih zbirk, raziskav, usmerjenih na polje samih podatkov, manj. Slednje pri metodologijah pogosto izpostavljajo velik pomen določanja cilja vrednotenja – prilagajanje kriterijev potrebam želene nadaljnje uporabe podatkov.

Zgolj dostop do podatkov ni dovolj za učinkovito nadaljnjo rabo, temveč je zanj potrebno upoštevati tudi večdimenzionalni aspekt kakovosti objavljenih podatkov. Predstavljene metodologije zato pogosto temeljijo na dimenzijah podatkov in njim prilagojenim metrikah, vendar se nabor dimenzij kot tudi njihove definicije med metodologijami razlikujejo. Posledično prihaja do odstopanj v kakovostnih lestvicah podatkovnih zbirk, raziskovalci pa navajajo tudi številne druge težave pri razvoju metodologij, predstavljenih v poglavju 3.3.

V tem pregledu stanja metodologij smo identificirali nekatere izzive, ki jih bomo v prihodnosti naslavljali v kontekstu odprtih podatkov. Z ozirom na slovenski portal odprtih podatkov - OPSI - je takšna metodologija nujno potrebna zaradi raznovrstnosti objavljenih podatkov, dostopnosti in povečanih potreb deležnikov po dvigu kakovosti in količini podatkov. Kljub upoštevanju evropskih in državnih smernic ter zakonodaj, navedenih v tem članku, so realne zmožnosti državnih agencij napram vedno večjim zahtevam uporabnikov odprtih podatkov pomemben omejevalni dejavnik.

ZAHVALA

Projekt Izdelava metodologije za določanje kakovosti podatkov ter ocena kakovosti posameznih podatkovnih zbirk na nacionalnem portalu odprtih podatkov Slovenije – portalu OPSI (št. V2-2388) je sofinancirala Javna agencija za raziskovalno dejavnost Republike Slovenije iz državnega proračuna.

LITERATURA

- [1] The 8 Principles of Open Government Data [Dostop dne: 27. 3. 2024]. (b.d.). <https://opengovdata.org/> Batini, C., & Scarnapiccia, M. (2006). *Data Quality*. Springer Berlin Heidelberg. <https://doi.org/10.1007/3-540-33173-5>
- [2] Berners-Lee, T. (b.d.). 5-star Open Data — 5stardata.info [Dostop dne: 29. 05, 2024]. <https://5stardata.info/en/> Bogdanović-Dinić, S., Veljković, N., & Stoimenov, L. (2014). How open are public government data? an assessment of seven open data portals. *Public Administration and Information Technology*, 5, 25–44. https://doi.org/10.1007/978-1-4614-9982-4_3
- [3] Data Catalog Vocabulary - Version 2. (2020). <https://www.w3.org/TR/vocab-dcat-2/>
- [4] Debattista, J., Auer, S., & Lange, C. (2016). Luzzu-A Methodology and Framework for Linked Data Quality Assessment. *Journal of Data and Information Quality (JDIQ)*, 8. <https://doi.org/10.1145/2992786>
- [5] GO FAIR: FAIR Principles [Dostop dne: 29. 05, 2024]. (b.d.). <https://www.go-fair.org/fair-principles/> ISO 25012. (2022). <https://iso25000.com/index.php/en/iso-25000-standards/iso-25012>
- [6] Janssen, K., & Kronenburg, T. (2012). EPSIP: Open Data Standardization before publication?
- [7] Krasikov, P., & Legner, C. (2023). A Method to Screen, Assess, and Prepare Open Data for Use. *Journal of Data and Information Quality*, 15. <https://doi.org/10.1145/3603708>
- [8] Kubler, S., Robert, J., Neumaier, S., Umbrich, J., & Traon, Y. L. (2018). Comparison of metadata quality in open data portals using the Analytic Hierarchy Process. *Government Information Quarterly*, 35, 13–29. <https://doi.org/10.1016/J.GIQ.2017.11.003>
- [9] Moraga, C., Moraga, M. Á., Calero, C., & Caro, A. (2009). SQuaRE-aligned data quality model for web portals. *Proceedings - International Conference on Quality Software*, 117–122. <https://doi.org/10.1109/QSIC.2009.23>
- [10] Neumaier, S., Umbrich, J., & Polleres, A. (2016). Automated Quality Assessment of Metadata across Open Data Portals. *Journal of Data and Information Quality (JDIQ)*, 8. <https://doi.org/10.1145/2964909>
- [11] Open Definition 2.1 [Version 2.1]. (2015). <https://opendefinition.org/od/2.1/en/>
- [12] Van Nuffelen, B. (2023). DCAT-AP 3.0. <https://semiceu.github.io/DCAT-AP/releases/3.0.0/>
- [13] Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, 33, 325–337. <https://doi.org/10.1016/J.GIQ.2016.02.001>
- [14] Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., & Auer, S. (2016). Quality assessment for Linked Data: A Survey. *Semantic Web*, 7, 63–93. <https://doi.org/10.3233/SW-150175>
- [15] Zhang, R., Indulska, M., & Sadiq, S. (2019). Discovering Data Quality Problems: The Case of Repurposed Data. *Business and Information Systems Engineering*, 61, 575–593. <https://doi.org/10.1007/S12599-019-00608-0>
- [16] Zuiderwijk, A., Pirannejad, A., & Susha, I. (2021). Comparing open data benchmarks: Which metrics and methodologies determine countries' positions in the ranking lists? *Telematics and Informatics*, 62, 101634. <https://doi.org/10.1016/J.TELE.2021.101634>

■

Klara Žnideršič je od leta 2023 članica Laboratorija za računalniško grafiko in multimedije na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Ukvarja se z raziskovanjem odprtih podatkov in metodologij vrednotenja ter z digitalnimi pristopi k učenju glasbe.

■

Matija Marolt je redni profesor in vodja Laboratorija za računalniško grafiko in multimedije na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Raziskovalno se ukvarja z analizo in prepoznavanjem multimedijskih signalov in z vizualizacijo podatkov.

■

Aleš Veršič je sekretar na Ministrstvu za digitalno preobrazbo, kjer se ukvarja s pripravo strategij in politik za področje podatkovnega gospodarstva ter odprtih podatkov. Je tudi vodja omrežja skrbnikov podatkov (Data Stewards) v državni upravi. Kot doktorski študent je vpisan na Fakulteto za organizacijske vede, Univerze v Mariboru, kjer raziskuje področje podatkovnih prostorov (Data Spaces).

■

Matevž Pesek je docent in raziskovalec na Fakulteti za računalništvo in informatiko Univerze v Ljubljani, kjer je diplomiral (2012) in doktoriral (2018). Od leta 2009 je član Laboratorija za računalniško grafiko in multimedije. Njegovi raziskovalni interesi so iskanje glasbenih informacij, vključno z glasbenim e-učenjem, biološko navdihljenimi modeli in globoke arhitekture. Od leta 2014 raziskuje odprte podatke in je sorazvil več storitev, med drugim Avtolog.si, Tocen.si in Vodostaji.si.

The logo for MODRA, featuring three white circles of varying sizes stacked vertically above the word "MODRA" in a bold, white, sans-serif font.

Zavarovalnica za dodatno
pokojninsko zavarovanje



MANJ DOHODNINE. **VEČ POKOJNINE.**

ZAKORAKAJ Z MODRO V PRIHODNOST.

Z varčevanjem v dodatnem pokojninskem zavarovanju ste upravičeni do posebne davčne olajšave. Vplačila v posameznem letu vam znižajo osnovo za odmero dohodnine in država vam del dohodnine vrne ali pa se vam zniža morebitno doplačilo dohodnine.

IZRAČUNAJTE
DAVČNO OLAJŠAVO



ONLINE NOTES: sistem za razpoznavo govora in strojno prevajanje v realnem času na ravni univerzitetnih predavanj

Tjaša Šoltes, Jan Vasiljević, Marko Bajec

Fakulteta za računalništvo in informatiko Univerze v Ljubljani, Večna pot 113, Ljubljana

tjasa.soltes@fri.uni-lj.si, jv1721@student.uni-lj.si, marko.bajec@fri.uni-lj.si

Izvleček

Online Notes je sistem za avtomatsko razpoznavo govora in prevajanje v realnem času na ravni univerzitetnih predavanj, ki ga razvijamo na Univerzi v Ljubljani. Sistem je primarno namenjen študentom Univerze v Ljubljani, ki ne razumejo ali slabše razumejo slovensko, ter študentom s senzornimi oviranostmi. V sklopu projekta smo razvili dva modela za razpoznavo govora in spletno aplikacijo s štirimi portali. Do sedaj je bilo izvedenih 29 pilotnih testiranj sistema, med drugim eno semestrsko, v okviru katerega smo spremljali dve študentki na izmenjavi programa Erasmus+, ki sta tekom celotnega semestra obiskovali predmet Prostorska statistika na Fakulteti za gradbeništvo in geodezijo UL v slovenščini z uporabo sistema Online Notes. Poleg spremljanja v živo sta imeli kot učni material na voljo tudi avtomatsko generirane podnapise v angleščini, ki so nastali na podlagi popravljenih slovenskih transkriptov. Študentki sta uspešno opravili vse študijske obveznosti. Izpostavili sta, da je sistem zelo uporaben za študente, ki ne razumejo slovensko, saj omogoča sledenje predavanjem, poleg tega pa sta izpostavili tudi določene pomanjkljivosti, kot je kognitivna zahtevnost pri spremljanju sistema, predavatelja in table hkrati in mestoma slabša kvaliteta prevodov pri predavanjih v živo.

Ključne besede: avtomatska razpoznavo govora, strojno prevajanje, univerzitetna predavanja

Online notes: a real-time speech recognition and machine translation system in university lectures

Abstract

Online Notes is a system for automatic speech recognition and real-time translation at the level of university lectures which has been developed at the University of Ljubljana. The system is primarily intended for students of the University of Ljubljana who do not understand Slovenian well or do not understand it at all, as well as students with sensory impairments. Within the project, we have developed two speech recognition models and a web application consisting of four portals. So far, 29 pilots have been carried out, among them a semester pilot, where we followed two exchange students of the Erasmus+ program who, during the entire semester, attended the Spatial Statistics course at the Faculty of Civil Engineering and Geodesy at the University of Ljubljana in Slovenian using the Online Notes system. In addition to the real time use, they also had automatically generated subtitles in English available as learning material, which were created on the basis of corrected Slovenian transcripts. The two students successfully completed all their study obligations. They pointed out that the system is very useful for students who do not understand Slovene, as it makes it possible to follow lectures, and they also pointed out certain shortcomings, such as the cognitive complexity of monitoring the system, the lecturer in the table at the same time, and the poor quality of translations in lectures in alive.

Keywords: automatic speech recognition, machine translation, university lectures

1 UVOD

V letu 2022 je bilo na Univerzi v Ljubljani vpisanih 3.508 tujih študentov od skupno 37.509 (torej 9,4 %),

na študijsko izmenjavo pa jih je prišlo 2.256 [7]. Ker predavanja večinoma potekajo v slovenščini, so študentje, ki slovensko ne razumejo, omejeni le na

peščico predavanj, ki se izvajajo v angleščini. Da bi tovrstnim študentom omogočili oziroma olajšali spremljanje vseh predavanj na Univerzi v Ljubljani, je bil razvit sistem Online Notes. Gre za sistem za avtomatsko razpoznavo govora in prevajanje v realnem času. V prvi vrsti je namenjen študentom, ki ne razumejo ali slabo razumejo slovensko, predvsem so to študentje na študijskih izmenjavah ali redno vpisani študentje, katerih materni jezik ni slovenščina, poleg tega pa je sistem namenjen tudi študentom s senzornimi oviranostmi (npr. za gluhe in naglušne). Razvili smo aplikacijo s štirimi portali ter dva modela razpoznavalnika govora (en namenjen naravoslovnim in drugi namenjen družboslovnim predavanjem).

Od leta 2021 dalje je bilo na osmih različnih fakultetah Univerze v Ljubljani izvedenih 29 pilotnih predavanj z uporabo sistema Online Notes, od tega tudi dve semestrski testiranji [1], eno bomo v prispevku tudi podrobneje predstavili. Pri enem od semestrskih testiranj smo spremljali dve študentki na izmenjavi projekta Erasmus+, ki sta tekom celotnega zimskega semestra študijskega leta 2022/2023 spremljali predavanja v slovenščini z uporabo sistema Online Notes in naknadno z videoposnetki, podnaslovljeni z avtomatsko generiranimi podnapisi v angleščini, pripravljene na podlagi popravljenih slovenskih transkriptov.

Prispevek je strukturiran sledeče: v poglavju 2 na kratko predstavimo sorodne raziskave, v poglavju 3 predstavimo sam sistem in aplikacijo, in sicer arhitekturo, komponento za razpoznavo govora ter uporabniški vmesnik, v četrtem poglavju pa podrobneje predstavimo rezultate semestrskega testiranja sistema.

2 PREGLED LITERATURE

Številne univerze v državah, kjer učni proces ne poteka v angleščini, se soočajo s podobnimi težavami. Tujim študentom namreč niso na voljo predavanja, ki bi potekala v angleščini, ali pa ali pa je takšnih vsebin le omejeno število. Zaradi jezikovnih ovir so takšne univerze posledično za tuje študente manj zanimive.

Sorodni sistemi za razpoznavo govora in prevajanje v realnem času so že bili razviti in preizkušeni tudi v drugih univerzitetnih okoljih. Zelo podoben sistem je denimo Karlsruhe Institute of Technology (KIT) Lecture Translation System [2, 3], ki ponuja transkripcije in strojne prevode nemških predavanj v obliki podnapisov. Pri evalvaciji sistema so poročali, da je večina študentov, ki je uporabljala sistem, oce-

nila sistem kot koristen, posebno samo transkripcijo govora. Med negativnimi vidiki sistema so izpostavili latenco pri prevodih, poleg tega pa tudi kognitivno zahtevnost pri sočasnemu sledenju materialom (preglednicam) in prevodom transkriptov predavanja. Eden od udeležencev raziskave je še poudaril, da je bil sistem uporaben predvsem za študente, ki se učijo nemško.

Shadiev in Huang [5] sta raziskovala uporabo razpoznave govora in strojnega prevajanja na primeru tujih predavateljev, ki v angleščini predavajo študentom, ki jim je angleščina tuji jezik. Raziskovala sta kognitivno obremenitev, pozornost in stres ob poslušanju predavanja v tujem jeziku. Ugotovila sta, da so imeli študentje, ki so predavanja spremljali s strojnimi prevodi v svojem prvem jeziku, najmanjšo kognitivno obremenitev in najvišjo stopnjo zadovoljstva v primerjavi s študenti, ki so predavanje spremljali samo s prepoznavanjem govora v angleščini, in študenti, ki sploh niso imeli podpore govorne tehnologije.

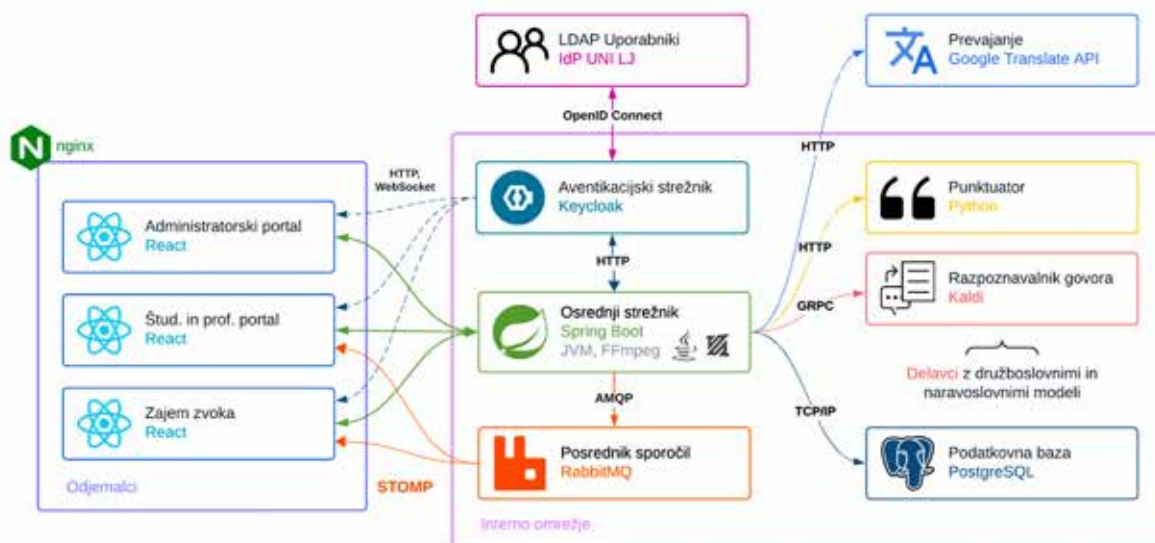
3 APLIKACIJA

Sistem Online Notes je sistem za avtomatsko razpoznavo govora in prevajanje v realnem času. V nadaljevanju predstavimo samo aplikacijo in pa komponente za razpoznavo govora in strojno prevajanje.

3.1 Arhitektura

Sistem je sestavljen iz več komponent, ki jih delimo na dva glavna dela. Čelni del (ang. frontend) sestavljajo tri spletne aplikacije, ki temeljijo na ogrodju React, za njihovo delovanje pa skrbi nginx strežnik. Odjemalske aplikacije komunicirajo z zalednim delom (ang. backend) preko več komunikacijskih kanalov, ki zahtevajo avtentikacijo. Med te kanale spadajo protokoli HTTP, Websockets in STOMP.

Osrčje zalednega dela sistema predstavlja spletni strežnik, implementiran v ogrodju Spring Boot in programskem jeziku Java 21, ki opravlja več nalog. Strežnik odjemalcem ponuja vmesnik, podoben REST-u. Prejete zahteve preverja preko avtentikacijskega strežnika Keycloak, ki pridobi tudi podatke o LDAP uporabnikih Univerze v Ljubljani. Strežnik preko protokola gRPC pošilja avdio zapis, prejet preko WebSocketov, v storitev za prepoznavanje govora. Le-ta glede na attribute predavanja dinamično alocira delo Kaldi delavcu. Prejeti transkripti so nato poslani punktualorju, ki dodaja ločila, in storitvi za prevajanje.



Slika 1: Shema arhitekture sistema

Med pretakanjem predavanja strežnik komunicira tudi s posrednikom sporočil RabbitMQ, ki skrbi za razpošiljanje transkriptov in prevodov odjemalcem. Avdio posnetki so dodatno obdelani z ogrodjem FFmpeg in shranjeni na datotečni sistem. Vsi preostali podatki kot so vsebina predavanj, podatki o predmetih in podobno, se hranijo v podatkovni bazi PostgreSQL.

3.2 Razpoznavo govora in strojno prevajanje

Ena od ključnih komponent sistema je tudi razpoznavnik govora. Za razpoznavo govora smo razvili dva ločena modela, in sicer enega za tehnična in drugega za družboslovna predavanja. Razpoznavnik govora temelji na ogrodju Kaldi [4], ki za usposabljanje in dekodiranje uporablja utežene pretvornike končnega stanja (*WFST*; ang. *weighted finite state transducers*).

Prvi korak v procesu prepoznavanja govora je branje valovne oblike signala in ekstrakcija značilnosti iz zvoka, ki je predstavljen v obliki Melovega spektrograma. Te podatke nato obdela akustični model, ki vrne verjetnostno matriko podbesednih enot skozi čas. Ta verjetnostna matrika se dekodira skupaj z jezikovnim modelom, ki vrne besedilo brez ločil. Besedilo procesira puntuator (ki je implementiran kot ločena storitev), ki v besedilo vstavi ločila.

Trenutno se znotraj sistema Online Notes uporabljata dva ločena modela za razpoznavanje govora, in sicer model za družboslovna predavanja in model za tehnične vede. Prednost ogrodja Kaldi je, da sta

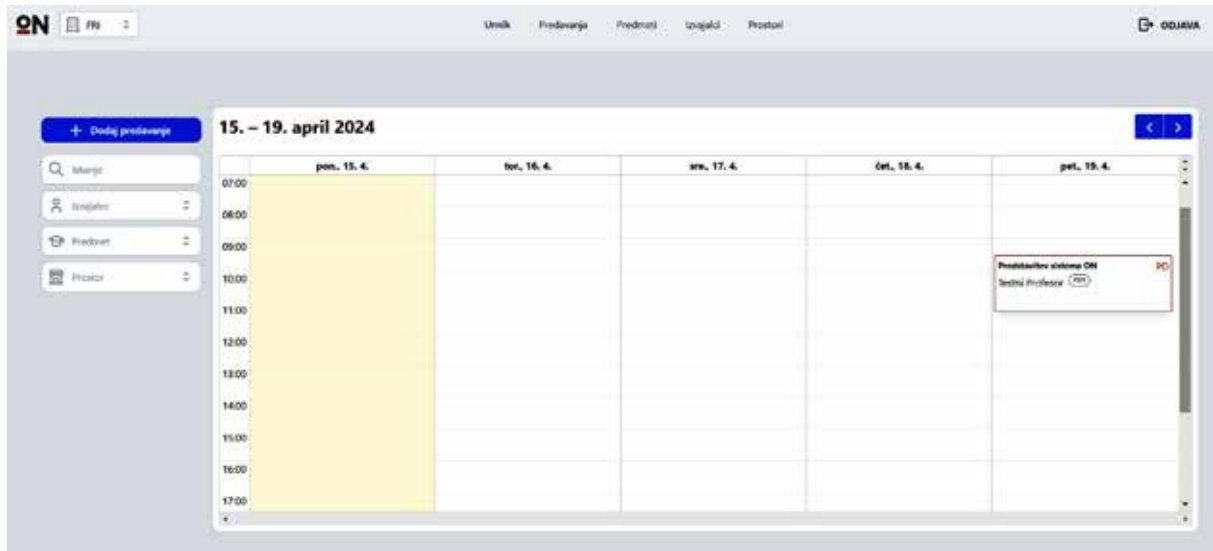
lahko akustični model in jezikovni model neodvisna, zato si oba modela za prepoznavanje govora delita skupen akustični model.

Osnova za jezikovni model je besedilni korpus. Vsebina predavanj je običajno specializirana, zato je nujno, da je jezikovni model prilagojen predavanju. Predavatelje zato prosimo, da nam pošljejo morebitne avdio ali video posnetke svojih predavanj iz preteklih let. Ti posnetki so samodejno transkribirani z uporabo razpoznavnika govora in ročno popravljani. Če gradiva ni na voljo, se za nadgradnjo modela uporabijo druga gradiva, kot so članki, revije, diplomske naloge ipd. Ko je podkorpus predavanja pripravljen, ga dodamo v enega od dveh glavnih korpusov, torej v tehnični ali družboslovni korpus. Ko je besedilni korpus sestavljen, med učenjem pripravimo leksikon besed s pripadajočimi izgovorjavami.

Akustični model je bil učen na približno 200 urah ortografsko transkribiranega besedila, jezikovna modela pa se ves čas posodabljata s transkripti preteklih predavanj. Trenutna učna množica je sestavljena iz okvirno 2 milijonov povedi in 1,3 milijona besed.

Razpoznavnik govora vrača razmeroma kratke besedilne segmente, ki se nato prevedejo v angleščino. Vsak besedilni segment se prevede samostojno, da se minimizira zamuda pri prevodu. Prevod je nato na študentskem (in kasneje na profesorskem) portalu prikazan vzporedno s slovenskim besedilom.

Za prevode se trenutno uporablja zunanji komercialni nevronske strojni prevajalnik.



Slika 2: Urnik na administratorskem portalu

3.3 Uporaba sistema in opis portalov

Sistem je sestavljen iz štirih portalov, in sicer: profesorski portal, študentski portal, portal za zajem zvoka in administratorski portal. Študentski in profesorski portal sta združena v enotno spletno aplikacijo, ki uporabnika glede na njegove pravice avtomatično preusmeri na ustrezní portal.

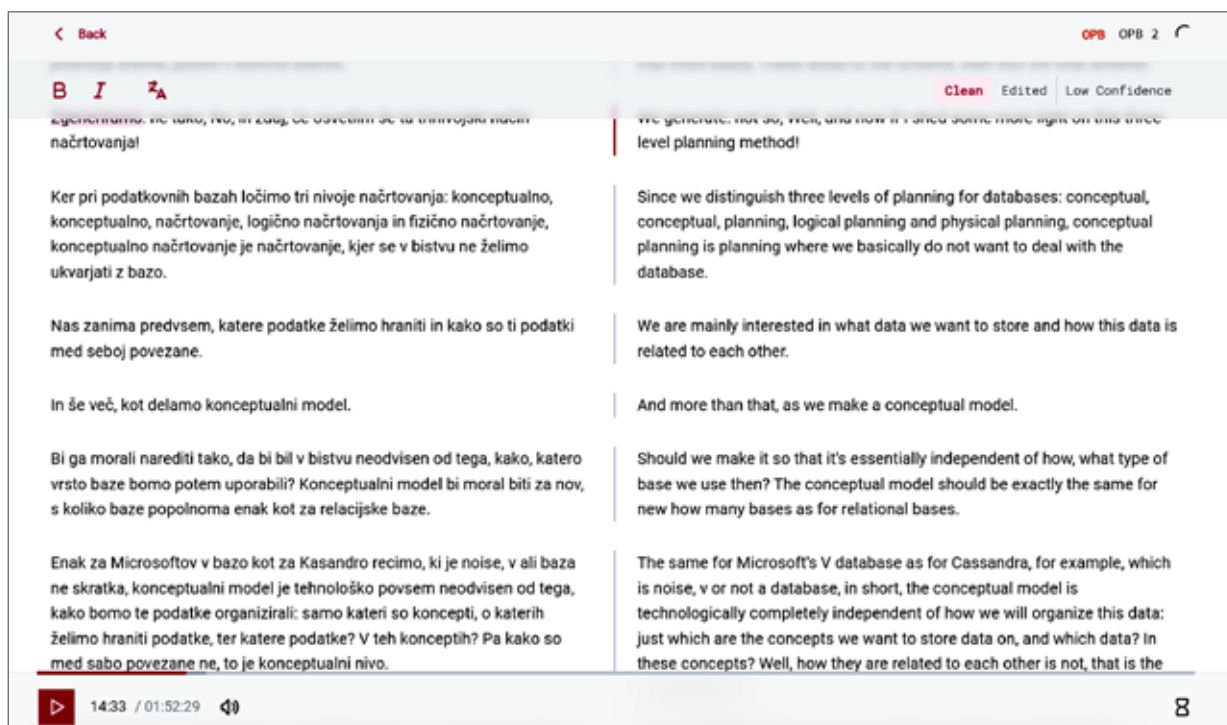
Na administratorskem portalu je možno v sistem vnesti predavatelje, predmete, posamezna predava-

nja in prostore oziroma predavalnice. Pred začetkom uporabe sistema je treba ustvariti posamezno predavanje, torej posamezno sejo, v katero se bo zapisovala vsebina in ki jo bodo študentje lahko spremljali. Ob ustvarjanju seje se določi predmet, predavatelja, prostor, prav tako pa se izbere tudi primerni model razpoznave govora (naravoslovni ali družboslovni).

Ko je seja ustvarjena, se lahko predavatelj s številko PIN vpiše v svoj račun na portalu za zajem zvoka.



Slika 3: Portal za zajem zvoka



Slika 4: Vsebina v profesorskem portalu po končanem predavanju

Pred snemanjem lahko preveri stanje vseh storitev in stanje povezave ter testira razpoznavnik govora, ne da bi se vsebina testiranja prikazovala študentom. Ko predavatelj začne s snemanjem, lahko vidi zadnje transkripte. Predavanje lahko začasno ustavi in kasneje z njim nadaljuje, ali pa ga zaključi. Za zagotavljanje dovoljšne kvalitete zvoka smo sodelujočim v pilotnih predavanjih namestili prenosne mikrofone, saj je kvaliteta opreme po različnih prostorih in članicah UL variirala.

Ko je predavanje aktivno, ga lahko študentje v živo spremljajo na študentskem portalu, kamor se vpišejo s svojimi študentskimi računi digitalne identitete UL. Ob govoru predavatelja se jim prikazujejo transkripti v slovenščini in prevodi v angleščino. Izberejo si lahko, ali naj se jim prikazuje le slovenski transkript, le angleški prevod ali oboje naenkrat, razen na manjših napravah, kot so mobilni telefoni, kjer je zaradi omejitve s prostorom možno predavanje spremljati le v enem jeziku.

Po končanem predavanju lahko predavatelj do vsebine dostopa na profesorskem portalu, kamor se vpiše s svojim računom. Tam si lahko ogleda slovenski transkript in prevod, ter popravi oziroma uredi vsebino predavanja. Predavatelj se lahko za vsako predavanje v sistemu odloči, ali ga želi deliti z na-

ročniki (tj. vsi študentje, ki so se v sistemu naročili oziroma prijavi na predavanje) ali ne. Če se odloči predavanje objaviti, postane vidno tudi na študentskem portalu.

4 SEMESTRSKO TESTIRANJE

V sklopu semestrskega testiranja smo spremljali dve študentki na študijski izmenjavi programa Erasmus+, ki sta v študijskem letu 2022/2023 cel semester spremljali predavanja pri predmetu Prostorska statistika na Fakulteti za gradbeništvo in geodezijo Univerze v Ljubljani pri prof. dr. Goranu Turku. Predavanja pri predmetu so potekala v slovenščini, študentki pa sta spremljali predavanje z uporabo sistema Online Notes s prevodi v angleščino. Poleg tega sta pri predmetu obiskovali tudi vaje, ki so potekale v angleščini.

Ena od študentk (v nadaljevanju Študentka A) je bila vpisana v tretji letnik dodiplomskega študija in prihaja iz Estonije, prav tako je estonsčina njen materni jezik. Svoje razumevanje angleščine je ocenila kot »tekoče« (C1/C2). Druga študentka (v nadaljevanju Študentka B) je bila vpisana v prvi letnik bolonjskega magistrskega študija in prihaja iz Francije. Njen materni jezik je francoščina. Svoje razumevanje angleščine je ocenila kot »dobro« (B1/B2).

Študentki smo prosili, da redno obiskujeta predavanja in jih spremljata z uporabo sistema Online Notes ter nam periodično podajata mnenje, po koncu semestra pa smo z njima opravili tudi končni intervju.

Tekom celotnega semestra sta obe študentki spremljali predavanja v živo z uporabo orodja Online Notes s svojimi študentskimi računi. Po vsakem predavanju so študentje, ki sodelujejo pri projektu, popravili transkript, da je ustrezal temu, kar je povedal predavatelj. Popravljeni transkripti so bili osnova za pripravo avtomatsko generiranih podnapisov v angleščini, profesor je namreč samostojno (brez uporabe sistema) posnel vsa predavanja in nato video-posnetke opremil z angleškimi podnapisi, ki so bili izvoženi iz sistema. Študentki sta torej imeli na voljo tudi avtomatsko generirane angleške podnapise za učenje po končanih predavanjih.

Na tem mestu je pomembno poudariti, da se kvaliteta prevodov v živo v primerjavi s prevodi v podnapisih zelo razlikuje. Za razliko od prevodov v živo, ki so prevodi posameznega končnega besedilnega segmenta v realnem času, gre tu za prevod popravljenega besedila. Vsaka napaka v slovenskem transkriptu se namreč v angleščini še potencira. Druga razlika pa je v tem, da se zaradi zagotavljanja hitre storitve v načinu v živo prevajajo posamezni odseki – segment po segment. Pri generiranju podnapisov po samem predavanju pa lahko ponovno prevedemo celotno vsebino predavanja naenkrat, kar seveda posledično zaradi več konteksta izboljša tudi kvaliteto prevoda.

4.1 Povratne informacije študentk

Študentki sta sistem uporabljali tekom enega semestra. Njuna mnenja smo zbrali ob prvi uporabi, na sredini semestra in ob koncu semestra.

Po prvem predavanju in s tem po prvi uporabi sistema Online Notes smo študentki prosili za povratne informacije v obliki odgovorov na anketni vprašalnik. Zanimala nas je splošna uporabnost sistema in kaj so glavni moteči elementi.

Študentki smo vprašali, ali se jima zdi sistem uporaben, ali bi ga priporočili svojim kolegom in ali bi ga uporabljali tudi za učenje po končanih predavanjih, pri čemer sta obe na vsa tri vprašanja odgovorili pritrdilno.

Poleg tega smo izpostavili 8 potencialnih pomanjkljivosti sistema in ju prosili, da ocenita, kako moteče so. Njuni odgovori so prikazani v Tabeli 1.

Na koncu je študentka A v razdelku splošno mnenje zapisala še, da je sistem zelo uporaben, a intenziven, študentka B pa je omenila, da je bilo težko spremljati sistem in obenem, kaj predavatelj v slovenščini piše po tabli. Poleg tega je omenila tudi, da mestoma zaradi napak v prevodu zelo težko razume, kaj je bilo mišljeno.

Ponovno smo ju prosili za mnenje v začetku decembra, tj. po dveh mesecih aktivne uporabe sistema. Zanimivo je, da sta obe študentki poročali, da se je kvaliteta angleških prevodov izboljšala, čeprav sami nismo zaznali bistvenega odstopanja v kvaliteti transkriptov ali prevodov tekom semestra, prav tako se jezikovni model v tem času ni dopolnjeval z novimi materiali s področja vsebine predmeta.

Študentka A je zapisala, da je spremljala skoraj vsa predavanja in da se ji zdi, da so se prevodi sčasoma izboljšali,¹ medtem ko je študentka B zapisala, da je besedilo bolj razumljivo kot na začetku semestra².

Študentka A je izpostavila slabše delovanje razpoznavalnika (in posledično tudi slabše prevode), kadar predavatelj piše na tablo. Izpostavila je tudi občasne neobičajne prevode, denimo »*God is uniformly distributed.*« ali »*We also go to our funeral.*«

Tabela 1: **Ocene motečih elementov v prevodih v živo**

	Ni moteče	Nekoliko moteče	Zelo moteče	Ne morem določiti
Besede manjkajo ali so odveč		A	B	
Napačen besedni red	A, B			
Napačno kapitalizirane besede	A, B			
Napačno postavljena ločila	A	B		
Napačno črkovanje		A	B	
Popolnoma napačne besede ali fraze		A	B	
Nerodno oblikovane povedi	A		B	
Zamik pri podajanju angleškega prevoda		A	B	

¹ »I have been following almost all the lectures and I do feel that the translation has improved over that time.«

² »I think that the software improved because things seem more understandable now than compare to the beginning of the semester.«

Študentka B je izpostavila težave pri razpoznavi spremenljivk, na primer „*here I have both Anič and B have different values*“, kjer je napačno razpoznana spremenljivka v prevodu posledica napačno prepoznane spremenljivke v slovenščini, gre namreč za spremenljivko *A0*.

Prav tako je izpostavila, da ji je sistem v pomoč, ko profesor ne piše, pač pa na primer razlaga koncepte, več težav pa je pri izračunih.³ Tu gre najverjetneje za več dejavnikov, ki vplivajo na razliko v kvaliteti: v korpusu besedil, ki je vključen v jezikovni model, so enačbe, pa tudi same spremenljivke zastopane v manjšem številu, prav tako se način govora (npr. dolžina premorov) ob sočasnem pisanju lahko spremeni.

Izpostavila je tudi kognitivno kompleksnost sistema, saj hkrati spremljata predavanja in predstavitve v slovenščini v živo, poleg tega pa morata ves čas spremljati še vsebino na svojem prenosniku v angleščini, kar je zahtevno.⁴

Po končanem semestru smo s študentkama opravili tudi končni intervju, kjer nas je zanimala njuna splošna izkušnja, prednosti in slabosti sistema.

Obe sta se strinjali, da jima je na splošno sistem olajšal sledenje vsebini predmeta. Moteči so bili mestoma napačni prevodi ob spremljanju v živo. Ob razlaganju teorije oziroma ob samem govoru (brez eksternih dejavnikov) so bili prevodi po njunem mnenju razumljivi in točni in sta lahko s pomočjo sistema razumeli predavatelja, medtem ko je bilo ob pisanju po tabli več napak. Po njunih izkušnjah je bilo največ težav pri razpoznavi spremenljivk in enačb. Nekoliko ju je motila tudi neobičajna struktura prevodov, kot so krajše povedi in ponavljanje besed, ki pa sicer izvira iz narave nebranega govora – pri nebranem govoru namreč govorci pogosto ponovijo besede, premori niso le na koncu povedi in podobno. Zanimivo je, da sta se obe strinjali, da so se prevodi tekom semestra izrazito izboljšali, medtem ko naša analiza ne kaže bistvenega odstopanja v kvaliteti transkriptov in prevodov tekom semestra. Zamik pri prevodih ju v nasprotju z našimi pričakovanji ni pretirano motil, študentko B je motil proti koncu semestra, študentki A pa se ni zdel pretirano moteč. V avtomatsko generiranih podnapisih na popravljenih

slovenskih transkriptih so se obema podnapisi zdeli precej bolj pravilni in razumljivi.

Drugi problem, ki ga je izpostavila študentka B, je kognitivna zahtevnost pri uporabi – spremljati je namreč treba predavatelja in tablo v enem jeziku, poleg tega pa še prenosnik s prevodi v drugem jeziku, hkrati pa noben od teh dveh jezikov ni materni, torej je vpet še tretji jezik. Kot alternativo je predlagala, da bi se prevodi generirali v avdio obliki, kjer bi bilo treba manj vsebine spremljati vizualno. Študentke A spremljanje programa in table ni motilo in ji je v določenih primerih olajšalo razumevanje, predvsem ko so bile v prevodu napačno razpoznane spremenljivke.

Glede materialov, ki so bili na voljo (sistem za spremljanje v živo in podnaslovljeni videoposnetki), sta se strinjali, da zadostujejo za samo učenje, da pa so jima pri učenju zelo pomagali podnapisi in bi jima bilo brez tega precej težje. Med samim predavanjem sta občasno dodatne informacije poiskali na spletu, vendar sta se strinjali, da to ni nujno posledica kvalitete in točnosti prevoda, pač pa same zahtevnosti vsebine predmeta.

Na splošno sta bili s sistemom zadovoljni in bi ga priporočili tudi svojim kolegom na domači univerzi.

4.2 Intervju s predavateljem

Po koncu semestra smo opravili tudi intervju s profesorjem. Profesor je bil s sistemom zadovoljen, predvsem so se mu zdeli koristni podnapisi. Svoja predavanja snema že od študijskega leta 2017/2018 naprej, vendar bi priprava podnapisov v angleščini vzela preveč časa.

Kvaliteto podnapisov (na podlagi popravljenih transkriptov) je ocenil kot boljše od pričakovanega. Terminološko je sicer zaznal nekaj pomanjkljivosti, zato si je pripravil seznam besed, ki so bile običajno napačno prevedene, in jih popravil. Po njegovi oceni mu je priprava podnapisov vzela okvirno 10 minut na predavanje. Ocenil je, da bi sicer sam marsikateri podnapis ubesedil drugače, da pa je vsebina razumljiva in da sta se študentki lahko naučili iz videov, da sta tudi izpit naredili brez težav, iz česar je sklepal, da sta iz videov razumeli, kaj želi povedati.

Vprašali smo ga, ali bi uporabljal sistem Online Notes tudi na ostalih predavanjih, na kar je odgovoril:

³ »I would say that it is helpful when the professor is not writing on the board, just talking (explaining one concept for example). But when calculations parts arrive, it is difficult to follow because the translation of mathematic expressions is chaotic. Furthermore, we have to read at the same time the board in slovene and our computer in english so it is a lot to process at once»

⁴ »Furthermore, we have to read at the same time the board in slovene and our computer in english so it is a lot to process at once.«

ril, da bi sistem uporabljal, če bi bili v razredu tuji, saj se mu zdi to dobra rešitev, posebno v kombinaciji s podnaslovljenimi videoposnetki.

Na vprašanje, ali je opazil, da njegov način govora kakor koli vpliva na samo kvaliteto razpoznavne, je odgovoril, da je pri spontanem oziroma nebranim govoru način govora drugačen, govor denimo vsebuje več premorov sredi povedi, besede se lahko ponavljajo in podobno, kar seveda vpliva na samo kvaliteto razpoznavne in posledično razumljivosti prevoda, da pa vendar ni namen predavanja, da je v naprej pripravljeno in brano.

Ta predmet je sicer v celoti v preteklosti že opravil študent na izmenjavi Erasmus+ iz Francije, vendar je predmet opravljal konsultacijsko. Konsenz je namreč, da se izvede ločeno predavanje za tuje študente, če je v skupini vsaj 5 tujih študentov. Zanimalo nas je, ali bi v prihodnje izvedel ločen predmet v primeru, da bi bilo nanj prijavljenih dovolj tujih študentov, na kar je odgovoril, da ne bi, saj sistem ON v kombinaciji s podnapisi omogoča sledenje pouku.

5 DISKUSIJA

V študiji smo spremljali dve tuji študentki, Francozinjo in Estonko, ki sta s pomočjo Online Notes, sistema za avtomatsko razpoznavo govora in strojno prevajanje v realnem času, v zimskem semestru 2022/2023 spremljali predavanje Prostorska statistika na Fakulteti za gradbeništvo in geodezijo Univerze v Ljubljani pri prof. dr. Goranu Turku v slovenščini. Študentki sta imeli poleg izpisov v živo med predavanji na voljo tudi videoposnetke predavanj, ki jih je predavatelj posnel samostojno (izven okvira sistema Online Notes), opremljeni pa so bili z avtomatsko generiranimi podnapisi v angleščini, ki jih je možno generirati v okviru sistema Online Notes. Osnova za podnapise so bili popravljeni slovenski transkripti. Poleg predavanj v slovenščini sta obiskovali še vaje, ki so potekale v angleščini.

Študentki sta bili na splošno s sistemom zadovoljni in sta mnenja, da jima je omogočil spremljanje predavanj. Bili pa sta enotni pri tem, da so bili pri učenju in samem razumevanju vsebine ključni tudi videoposnetki s podnapisi.

Ena od študentk je v več fazah uporabe omenila, da je spremljanje različnih enot (sistema na računalniku, table, profesorja) miselno zelo zahtevno, po-

sebnost tudi, ker gre za dva tuja jezika poleg materne jezika, o čemer so sicer poročali tudi v drugih študijah, denimo pri evalvaciji sorodnega sistema KIT v Nemčiji [3]. Druga študentka tega ni opazila. Možno je, da na percepcijo zahtevnosti vpliva tudi stopnja razumevanja angleščine (prva študentka je namreč svoje razumevanje angleščine ocenila z B1/B2, druga pa na C1/C2), vendar bi bilo treba za takšne ugotovitve študijo razširiti še na večje število udeležencev.

Poleg tega sta študentki v vseh fazah poročanja omenili slabšo kvaliteto in natančnost prevodov v živo v primerjavi s prevodi pri podnapisih. Razlika v kvaliteti oziroma točnosti prevodov izvira iz dveh pomembnih razlik v postopku priprave. Glavni razlog izboljšave je ta, da so podnapisi (prevodi po predavanju) nastali iz popravljenih in urejenih slovenskih transkriptov, medtem ko prevodi v živo seveda ohranjajo in celo potencirajo vse napake razpoznavnika govora. Na tem mestu omenimo, da je bila sicer razpoznavna govora s takratno verzijo modela razpoznavnika govora za tehnična predavanja pri tem predavanju nekoliko slabša od povprečja, verjetno ravno zaradi velike vključenosti eksternih elementov v govor in velike zastopanosti spremenljivk in enačb, ki so v jezikovnem modelu slabše zastopane. S standardno metriko WER (ang. *word error rate*), ki izraža razmerje dodanih, manjkajočih in napačnih besed v razmerju do vseh pojavnici⁵ v primerjavi z referenčnim transkriptom, ki je pripravljen ročno, smo ocenili vsa pilotna predavanja. Prvo predavanje v sklopu semestrskega testiranja je imelo stopnjo napake 20 %, kar je za 2 % več kot povprečna stopnja napake razpoznavnika govora na pilotnih predavanjih v študijskem letu 2022/2023 [6].

Druga razlika v kvaliteti prevodov pa je v tem, da se pri pripravi prevodov po predavanju ponovno prevede celotno vsebino predavanja naenkrat, kvaliteta prevodov pa se običajno boljša s količino konteksta. Prevodi v živo se generirajo na ravni enega besedilnega segmenta, torej je v prevajalnik poslano minimalno konteksta – s tem se želimo izogniti pretiranemu časovnemu zamiku. Tekom predavanja, torej spremljanja v živo, želimo tujim študentom zagotoviti čim hitrejše prevode, zamik je namreč lahko zelo moteč, posebno v kombinaciji z drugimi vizualnimi informacijami (informacije na tabli oziroma predstavitvi).

⁵ WER se izračuna po naslednji enačbi: (substituti + manjkajoče besede + odvečne besede)/število vseh pojavnici.

Predavatelj je bil s samim sistemom zadovoljen. Iz njegovih odgovorov, ki smo jih prejeli tekom intervjuja, lahko sklepamo, da se mu je sistem zdel uporaben in mu je olajšal delo. Ločena predavanja običajno v angleščini izvajajo, če je v skupini več kot 5 tujih študentov, sicer pa se srečanja izvajajo konsultacijsko. Profesor vsa svoja predavanja snema že od študijskega leta 2017/2018 dalje, zato mu je posebno prav prišla opcija izvoza podnapisov iz sistema.

Prof. dr. Goran Turk je po izvedbi te študije sistem semestrsko uporabljal še pri dveh predmetih v letnem semestru študijskega leta 2022/2023 in ga trenutno uporablja v letnem semestru v študijskem letu 2023/2024.

6 NADALJNJE DELO

Ena izmed nadgradenj, ki jo želimo ponuditi v sklopu projekta in se je izkazala kot dobrodošlo tudi tekom semestrskega testiranja, opisanega v poglavju 4, je izbira ciljnega jezika prevodov. Sledenje vsebini v slovenščini v živo in branje v angleščini, poleg procesiranja v tretjem, maternem jeziku, je lahko zahtevno, zato je smiselno študentom ponuditi opcijo prevoda v njihov materni jezik. Posebno tudi zato, ker je na Univerzi v Ljubljani veliko študentov s področja Balkana, katerih materni jezik je bližje slovenščini kot angleščini, zato se zdi angleščina v takšnih primerih uporabe odvečen korak. Za prevode se trenutno uporablja zunanji strojni prevajalnik, v kratkem pa bo znotraj sistema na voljo tudi strojni prevajalnik iz slovenščine v angleščino, ki je bil razvit na Univerzi v Ljubljani. V primerih prevajanja v druge tuje jezike (poleg angleščine) pa bi bil za prevajanje še vedno na voljo eksterni prevajalnik.

Poleg uporabnosti za tuje študente je velik poudarek pri nadaljnjem razvoju same aplikacije tudi dostopnost sistema za osebe s senzornimi oviranostmi. Za izboljšanje dostopnosti in s tem večjo uporabnost sistema sodelujemo s strokovnjaki s področja dostopnosti.

Ves čas nadgrajujemo sistem tudi z vidika enostavnosti uporabe in uporabniške izkušnje. Cilj je, da je urejanju transkriptov po predavanju čim enostav-

nejše, zato poskušamo upoštevati čim več predlogov uporabnikov, ki smo jih zbrali v času pilotne faze projekta (na primer hitrejše pomikanje po vsebini predavanja, iskanje po celotni vsebini predavanja in podobno).

ZAHVALA

Raziskava je bila finančno podprta s sredstvi projekta Online Notes, ki ga financira Univerza v Ljubljani.

Hvala tudi prof. dr. Goranu Turku s Fakultete za gradbeništvo in Geodezijo na Univerzi v Ljubljani za sodelovanje v študiji in aktivno uporabo sistema ves čas pilotne faze projekta.

LITERATURA

- [1] Bajec, M., Lebar Bajec, I., Šoltes, T., Cvek, J., Čibej, J., Gantar, K., Sever, S., & Krek, S. (2023). Online Notes - a real-time speech recognition and machine translation system for Slovene university lectures. 7–10. https://is.ijs.si/wp-content/uploads/2023/11/IS2023_Volume-H.pdf
- [2] Dessloch, F., Ha, T.L., Müller, M., Niehues, J., Nguyen, T.S., Pham, N.Q., Salesky, E., Sperber, M., Stüker, S., Zenkel, T., Waibel, A.: KIT lecture translator: Multilingual speech translation with one-shot learning. In: Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations. pp. 89–93. Association for Computational Linguistics, Santa Fe, New Mexico (Aug 2018), <https://aclanthology.org/C18-2020>
- [3] Markus Müller, Sarah Fünfer, Sebastian Stüker, and Alex Waibel. 2016. Evaluation of the KIT Lecture Translation System. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), str. 1856–1861, Portorož, Slovenia. European Language Resources Association (ELRA).
- [4] Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hannemann, M., Motlicek, P., Qian, Y., Schwarz, P., Silovsky, J., Stemmer, G., Vesely, K.: The kaldi speech recognition toolkit. In: IEEE 2011 Workshop on Automatic Speech Recognition and Understanding. IEEE Signal Processing Society (Dec 2011), iEEE Catalog No.: CFP11SRW-USB
- [5] Shadiev, R., Huang, Y.M.: Investigating student attention, meditation, cognitive load, and satisfaction during lectures in a foreign language supported by speecheenabled language translation. Computer Assisted Language Learning 33(3), 301–326 (2020). <https://doi.org/10.1080/09588221.2018.1559863>
- [6] Šoltes, T., Bajec, M., Lebar Bajec, I., Gantar, K., & Žitnik, S. (2023). Online-notes system: real-time speech recognition and translation of lectures. 485–492. https://link.springer.com/chapter/10.1007/978-3-031-33080-3_29
- [7] Univerza v številkah. <https://www.uni-lj.si/univerza/o-nas/univerza-v-stevilkah>. Dostop 30. 3. 2024.

■

Tjaša Šoltes je raziskovalka v Laboratoriju za podatkovne tehnologije na Fakulteti za računalništvo in informatiko UL in doktorska študentka programa Digitalno jezikoslovje. Ukvarja se z obdelavo naravnega jezika, posebno za namene govornih tehnologij. Sodelovala je pri raznih projektih na temo digitalizacije slovenskega jezika in uporabe govornih tehnologij za slovenščino, denimo RSDO (Razvoj slovenščine v digitalnem okolju), Online Notes, razvoj sintetizatorja govora za slovenski jezik, pilotni projekt uporabe govornih tehnologij v Državnem zboru ipd.

■

Jan Vasiljević je študent prvega letnika magistrskega programa Računalništvo in informatika na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. V svoji diplomski nalogi je preučeval vpliv psihometričnih lastnosti v skupinskih okoljih in napisal članek, objavljen v reviji Electronics. Kot študent dela v Laboratoriju za podatkovne tehnologije in trenutno sodeluje pri projektu ON. V sklopu projekta razvija zaledne storitve in integracijo z govornimi tehnologijami.

■

Marko Bajec je redni profesor ter vodja Laboratorija za podatkovne tehnologije na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Raziskovalno in aplikativno se ukvarja z razvojem podatkovno intenzivnih sistemov. V zadnjih letih večino časa posveča govornim tehnologijam ter digitalizaciji slovenskega jezika. Na tem področju je vodil številne projekte kot npr. Samodejno podnaslavljanje TV programov na RTV SLO, Uvedba govornih tehnologij v Telekom Slovenije ipd.

Poenostavite upravljanje vašega IT-okolja z rešitvijo NIL Cloud Management Platform

Preoblikujte vaš podatkovni center v sodobno storitveno platformo. Zagotovite si preglednost stroškov in učinkovito dostavo storitev IT.

Prednosti NIL Cloud Management Platform



Ena platforma za celovito upravljanje okolja skozi storitveno tržnico



Izboljšanje odzivnosti in učinkovitosti IT-službe skozi avtomatizacijo in orkestracijo



Procesna in stroškovna preglednost vedno bolj kompleksnih IT-okolij z možnostjo integracije z zunanjimi sistemi (SIEM, XDR, EDR, ITSM...)

Kontaktirajte nas za demo:

consulting@conscia.com

www.nil.com



Iz Islovarja

Islovar je spletni terminološki slovar informatike, ki ga že več kot 20 let ureja jezikovna sekcija Slovenskega društva INFORMATIKA. Slovar je javno dostopen za vpoglede in vnašanje novih izrazov na naslovu <http://www.islovar.org>

diferencialna varnostna kópija -e -- -- ž (*angl. differential backup*)

varnostna kopija sprememb, nastalih po zadnjem popolnem varnostnem kopiranju; prim. popolna varnostna kopija, prirastna varnostna kopija

dískovno varnostno kopíranje -ega -a -a s (*angl. disk backup, backup-to-disk*)

varnostno kopiranje, pri katerem je nosilec varnostne kopije disk; prim. tračno varnostno kopiranje

dnévná varnostná kópíja -e -e -e ž (*angl. daily backup*)

prirastna varnostna kopija podatkov, ki so se spremenili na izbrani dan

hládnó varnostno kopíranje -ega -ega -a s (*angl. offline backup, cold backup*)

varnostno kopiranje, pri katerem je onemogočen dostop do kopiranih podatkov; prim. vroče varnostno kopiranje

infrastruktúra mréžné varnosti -e -- -- ž (*angl. grid security infrastructure*)

celovit nabor ukrepov in postopkov za zagotavljanje varnosti v okolju mrežnega računalništva

prirástno varnostno kopíranje -ega -ega -a s (*angl. incremental backup*)

varnostno kopiranje podatkov, nastalih po zadnjem diferencialnem ali popolnem varnostnem kopiranju

progámje za varnostno kopíranje -a -- -- -- s (*angl. backup software*)

namensko programje za izdelavo varnostnih kopij

ravén varnosti -vní -- ž (*angl. security level*)

merilo vrednotenja varnostnih ukrepov in postopkov v sistemu

sístém za uprávljanje informacijske varnosti -a -- -- -- -- m (*angl. information security management system, ISMS*)

upravljalni sistem, ki z ukrepi in postopki nadzira in zagotavlja informacijsko varnost; sin. SUIV

uprávljanje varnosti -a -- m (*angl. security management*)

načrtovanje, uporaba in nadzor varovalnih ukrepov, ki zagotavljajo zahtevano stopnjo varnosti

varnost informacijskega sistéma -i -- -- ž (*angl. information system security*)

stanje informacijskega sistema, ki kljub varnostnim grožnjam zadovoljivo opravlja zahtevane funkcije; prim. informacijska varnost

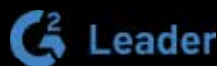
SOPHOS

Cybersecurity delivered.



Sophos Managed Detections and Response

Sophos MDR je najbolj razširjena MDR storitev na svetu. Zaupa nam že več kot **18.000** podjetij!



Distributer: Sophos d.o.o., www.sophos.si, slovenija@sophos.si, T: 07/39 35 600

Izpitni centri ECDL

ECDL (European Computer Driving License), ki ga v Sloveniji imenujemo evropsko računalniško spričevalo, je standardni program usposabljanja uporabnikov, ki da zaposlenim potrebno znanje za delo s standardnimi računalniškimi programi na informatiziranem delovnem mestu, delodajalcem pa pomeni dokazilo o usposobljenosti. V Evropi je za uvajanje, usposabljanje in nadzor izvajanja ECDL pooblaščen ustanova ECDL Foundation, v Sloveniji pa je kot član CEPIS (Council of European Professional Informatics) to pravico pridobilo Slovensko društvo INFORMATIKA. V državah Evropske unije so pri uvajanju ECDL močno angažirane srednje in visoke šole, aktivni pa so tudi različni vladni resorji. Posebno pomembno je, da velja spričevalo v 148 državah, ki so vključene v program ECDL. Doslej je bilo v svetu v program certificiranja ECDL vključenih že preko 16 milijonov oseb, ki so uspešno opravile preko 80 milijonov izpitov in pridobile ustrezne certificate. V Sloveniji je bilo doslej v program certificiranja ECDL vključenih več kot 18.000 oseb in opravljenih več kot 92.000 izpitov. V Sloveniji sta akreditirana dva izpitna centra ECDL, ki imata izpostave po vsej državi.



Znanstveni prispevki

Sandi Gec, Vlado Stankovski

ANALIZA NEFUNKCIONALNIH ZAHTEV NA PRIMERU UPORABE
PRIPOROČILNEGA SISTEMA PAMETNIH POGODB

Lana Masnec, Marina Trkman

ANALIZA UPORABE APLIKACIJE ZA SLEDENJE STIKOV MED MLADIMI:
ŠTUDIJA PRIMERA NEMČIJE

Strokovni prispevki

Klara Žnideršič, Matija Marolt, Aleš Veršič, Matevž Pesek

METODOLOGIJE ZA KVALITATIVNO VREDNOTENJE KAKOVOSTI
ODPRTIH PODATKOV

Tjaša Šoltés, Jan Vasiljevič, Marko Bajec

ONLINE NOTES: SISTEM ZA RAZPOZNAVNO GOVORA IN STROJNO
PREVAJANJE V REALNEM ČASU NA RAVNI UNIVERZITETNIH PREDAVANJ

Informacije

IZ ISLOVARJA

ISSN 1318-1882



9 771318 188001