



Dr. Tina Štemberger

Sekundarna analiza podatkov in masovni podatki v pedagoškem raziskovanju

Prejeto 10. 10. 2022 / Sprejeto 8. 5. 2023

Znanstveni članek

UDK 001.891+311.213.5

KLJUČNE BESEDE: sekundarna analiza podatkov, masovni podatki, rudarjenje podatkov, učna analitika

POVZETEK – Prispevek se usmerja v dva koncepta, ki se tudi na področju raziskovanja vzgoje in izobraževanja vse pogostejše pojavljata v strokovnih razpravah, nista pa bila znotraj tega področja pri nas še deležna dovoljšne pozornosti, zato sta manj poznana in razširjena ter skromno implementirana. V prvem delu prispevka se usmerjamo v sekundarno analizo podatkov, pri čemer najprej opredelimo sam koncept sekundarne analize podatkov, nato pa analiziramo potencial in zadržke pri rabi tovrstnih podatkov. V nadaljevanju predstavimo pojem masovni podatki, ki z vse večjo digitalizacijo, tudi v šolskem prostoru, postajajo vse pomembnejši vir podatkov, ki predstavlja pomembno osnovo za načrtovanje sprememb za višjo kakovost izobraževanja. Hkrati predstavimo tudi pomembne slabosti rabe masovnih podatkov, med katerimi se zlasti usmerimo v problem varnosti podatkov na spletu ter pomanjkanja kompetentnih strokovnjakov za analizo tovrstnih podatkov. Oba koncepta skušamo tudi smiselno umestiti v polje pedagoškega raziskovanja.

Received 10. 10. 2022 / Accepted 8. 5. 2023

Scientific paper

UDC 001.891+311.213.5

KEYWORDS: secondary data analysis, massive data, data mining, learning analytics

ABSTRACT – The paper focuses on two concepts that are increasingly emerging in professional debates in the field of educational research, but which have not yet received sufficient attention within the field and are therefore less known, less widely used and modestly implemented. The first part of the paper focuses on secondary data analysis; the concept itself is introduced, while its potential and drawbacks, such as being collected for other purposes, are analysed. We then introduce the Big Data concept which, with the shift of digitalization, has also become an important source of data in education, and a linchpin for planning changes to improve the quality of education. At the same time, we highlight the challenges of using Big Data, such as security and the lack of competent professionals for data analysis. The paper also aims to embed both concepts in the field of educational research.

1 Uvod

Če velja, da je raziskovanje načrten, sistematičen, kritičen in samokritičen proces pridobivanja novega znanja, ki izhaja iz raziskovalnih ciljev in poteka po vnaprej določenih fazah, za raziskovanje na pedagoškem področju trdimo, da je v osnovi njegov cilj pridobivanje oz. ustvarjanje znanja na pedagoškem področju (Cencič, 2009). Pedagoško raziskovanje je tudi zelo kompleksno področje, ki vključuje in združuje različne raziskovalne paradigme, številne tehnike zbiranja podatkov, raznolike udeležence raziskav in raznolike obdelave podatkov (Štemberger, 2020). Hkrati gre tudi za področje, ki mora slediti družbenim spremembam in potrebam ter se nanje tudi kar najhitreje odzivati; med njimi je tudi vse pogostejša vsakdanja raba digitalnih tehnologij za namene poučevanja in učenja, s svojim načinom delovanja pa vodi tudi v ogromne količine zbranih podatkov. Količina podatkov in dejstvo, da so ti podatki že zbrani, odpirata številne nove možnosti

raziskovanja, pri čemer pa je pomembno, da se zavedamo prednosti in tudi pasti tako zbranih oz. izbranih podatkov ter tudi pasti teh postopkov. V prispevku se zato osredinjamo na dva ključna koncepta, ki ju je v povezavi z velikimi količinami podatkov treba tudi na pedagoškem področju ustrezno opredeliti in utemeljiti – to sta sekundarna analiza podatkov (angl. secondary data analysis) in masovni podatki (angl. big data).

2 Sekundarna analiza podatkov

Omeniti je treba, da sekundarna analiza podatkov ni nov pojav. Tako so denimo Kozina idr. (2012, str. 131) na osnovi baz raziskav TIMSS, v katerih so zbrani predhodno zbrani podatki za dijake, učitelje in ravnatelje, v svojo raziskavo “vključili tiste dijake, ki so se pripravljali na izpitni preizkus iz fizike”. Kunčič idr. (2013, str. 45) so v prispevku analizirali zapise spletnih dnevnikov, ki so bili v osnovi “zapisani kot učno sredstvo za doseganje ciljev pouka”. Čepar idr. (2022) so za prispevek, ki naslavlja povezovanje srednjega šolstva in gospodarstva, uporabili podatke “iz Statističnega urada Republike Slovenije”, kar so tudi eksplicitno označili kot sekundarne podatke (prav tam, str. 132).

Vendarle pa so navedeni primeri dokaj osamljena praksa na področju pedagoškega raziskovanja, kar kaže na to, da je sekundarna analiza podatkov na tem področju še vedno premalo poznana in uporabljena. Smithova (2008) sicer ugotavlja, da se pojavljajo številne definicije analize sekundarnih podatkov, pri čemer prednjačijo tiste, ki poudarjajo uporabnost analize sekundarnih podatkov za raziskovanje s postavljanjem novih raziskovalnih vprašanj. Te definicije po njenem mnenju sicer zanemarjajo potencial sekundarnih podatkov za ponovno analizo istih podatkov z uporabo drugih statističnih ali teoretskih pristopov, da se torej ponovno išče odgovore na prvotno postavljena raziskovalna vprašanja, pri čemer se na primer uporabi naprednejše statistične metode.

V osnovi so sekundarni podatki definirani kot podatki, ki jih je nekdo predhodno že zbral in tudi uporabil (Smith, 2008), pri čemer pa Schutt (2007) navaja, da to velja, tudi če podatke ponovno uporabi ista oseba ali ista skupina oseb, ki je te podatke že zbrala in obdelala. Iz tega sledi, da ni vedno mogoče postaviti natančne ločnice med primarnimi in sekundarnimi podatki. Zato Jary in Jary (2000, str. 540) predlagata precej splošno definicijo analize sekundarnih podatkov, in sicer da gre za vsako raziskavo, ki temelji na ponovni analizi predhodno analiziranih podatkov.

Potencial sekundarne analize podatkov

Glaser (1963) je že pred desetletji poudaril, da z rabo sekundarnih podatkov prihranimo čas in denar, a tudi ohranjamo kariere, raziskovalne interese, samopodobo ter tako na neki način “rešimo” ogromne količine podatkov pred njihovo izgubo.

Smithova (2008) povzema naslednje prednosti sekundarne analize podatkov:

- Raziskovalcem omogoča dostop do velikih vzorcev, ki bi jih z lastnim zbiranjem podatkov bistveno težje dosegli.
- Če raziskovalec uporabi sekundarne podatke, mu ni treba vložiti časa in npora za pripravo instrumenta zbiranja podatkov, kar je tudi prednost za raziskovalce z manj metodološkega pa tudi tehničnega znanja. Še več, kot poudarja Yorke (2011), so

sekundarni podatki na voljo takoj, tako da na rezultate ni treba dolgo čakati, kar v hitro spreminjajoči se družbi omogoča tudi sorazmerno hitro prilagajanje trenutnim razmeram.

Navaja pa tudi, da sekundarni podatki lahko predstavljajo priložnost za generiranje novih raziskovalnih idej, npr. za odkrivanje drugačnih razmerij, povezanosti ali razlik, kot so bile zasnovane v primarni raziskavi. Podobno navaja tudi Yorke (2011), saj meni, da zbrani podatki omogočajo tudi uresničevanje raziskovalnih ciljev, ki jih primarna analiza ni naslovila. Poudarja, da celo organizacije, ki imajo dostop do ogromnih baz podatkov in razpolagajo s kompetencami za statistično obdelavo podatkov (pri čemer navaja npr. National Center for Educational Statistics (ZDA), Higher Education Funding Council (Velika Britanija)), ne uspejo izkoristiti vsega potenciala teh podatkov, kar odpira pot sekundarni analizi podatkov. Smithova (2011) dodaja, da sekundarni podatki ponujajo številne priložnosti za ponavljanje analiz in ponovno interpretacijo. Omogočajo izpeljavo longitudinalnih raziskav, raziskovanje preteklih dogodkov ter eksploratorne raziskave za testiranje novih idej, teorij in modelov. Z uporabo podatkov drugih virov je omogočena triangulacija. Že leta 1974 je Cook (1974) poudaril, da zastavljanje drugačnih raziskovalnih vprašanj, uporaba drugih statističnih metod za sekundarne podatke lahko pripelje tudi do tega, da se razkrijejo morebitne napake prvotnih analiz.

Nadalje se poudarja tudi (Hakim, 1982; Glaser, 1963), da so sekundarni podatki običajno visokokakovostni in s finančnega vidika (Hakim, 1982; Smith, 2008; Yorke, 2011) visoko dostopni. Ta vidika sta pomembna, ker je v številnih primerih sicer zbiranje podatkov lahko zelo drag proces in je zato dostopnost do visokokakovostnih podatkov na eni strani pomembna za raziskovalce začetnike, pa tudi za izkušenejše raziskovalce, saj jim omogoča visoko stopnjo raziskovalne neodvisnosti. Z izpustitvijo faze zbiranja podatkov lahko raziskovalci začetniki in tudi študentje pridobijo dragoceno izkušnjo izvedbe neodvisne raziskave na področju zanimanja in možnosti objave rezultatov. V tem smislu imajo sekundarni podatki pomembno vlogo pri izgrajevanju raziskovalnih kompetenc, pa tudi razvoju raziskovalnih interesov (Smith, 2021). Na pomen sekundarnih podatkov za izgrajevanje raziskovalnih kompetenc opozarja tudi Sobal (1981), ki meni, da imajo ti pomembno vlogo pri poučevanju, še posebej pri poučevanju raziskovalne metodologije. Ob sekundarnih podatkih imajo lahko zainteresirani namreč tudi vpogled v vprašalnike (s katerimi so bili podatki zbrani), ki so lahko izhodišče za diskusijo o pozitivnih in slabih praksah pri oblikovanju vprašanj, konstrukciji lestvic, vrstnem redu vprašanj in pri uporabljenih metodah statistične obdelave (Sobal, 1981). Prav tako so sekundarni podatki zelo uporabni pri učenju in poučevanju statistike, od postavljanja hipotez, vzorčenja do izbire ustreznih statističnih metod za obdelavo podatkov (Smith, 2008).

Bulmer (1980) poudarja, da raba sekundarnih podatkov predstavlja zelo nevsiljiv način zbiranja podatkov (pravzaprav le obdelave podatkov), saj ne posega v življenje posameznika ali skupine, ščiti njihovo zasebnost, saj ne sprašuje o nekom in njegovih dejavnostih. Ta značilnost je še toliko pomembnejša v raziskavah, ki naslavljajo občutljive teme ali skupine ljudi, še posebej v primeru, ko so skupine ali posamezniki težje dostopni. Kot navaja Gorard (2002), so sekundarni podatki v današnjih družbenih razmerah lahko še kako pomembni tudi z vidika opisovanja neenakosti, saj omogočajo opisovanje stanja brez poseganja v ranljive skupine oz. izpostavljanja ranljivih skupin zaradi namenskega zbiranja podatkov. Ti podatki lahko služijo politiki za postavljanje mehanizmov, ki bi omogočili premagovanje neenakosti, tudi v šolstvu. Kot opozarja

Petlák (2021, str. 42), so zato potrebne reforme izobraževalnih sistemov, ki vnašajo inovacije v vsebine izobraževanja, pa tudi take, ki poudarjajo kakovost izobraževanja učiteljev. Gorard (2002) predlaga dvostopenjski pristop: najprej se na osnovi obstoječih sekundarnih podatkov identificira neki trend, vzorec ali situacijo, nato pa se to preuči bolj poglobljeno, predvsem z rabo kvalitativnih pristopov.

Pomembno je poudariti tudi pomen analize sekundarnih podatkov v deskriptivnih raziskavah, in to kljub dejstvu, da imajo v akademskih krogih deskriptivne študije običajno nižji status kot študije, ki preverjajo povezanost, razlike ali napovedi, nižje pa je tudi zaupanje v možnost posploševanja, pri čemer se raziskovalci pogosto usmerjajo v pojasnjevanje nekega pojava, še preden se dokaže njegov dejanski obstoj. Zadnje vodi v to, da so številni pojavi prešibko opisani in slabo izmerjeni, saj se prehitro preide v preverjanje napovedi itn. (Campbell idr., 1982).

Zadržki pri sekundarni analizi podatkov

Yorke (2011) opozarja, da je sekundarne podatke treba natančno pregledati, da se razjasnijo morebitne šibkosti in da se, če je treba, izpostavijo omejitve teh podatkov, ki jih je treba upoštevati pri analizi in predstavitvi rezultatov teh analiz. Smithova (2008) kot glavna zadržka pri rabi sekundarnih podatkov navaja dejstvo, da je lahko v obstoječih zbirkah podatkov veliko napak in da so družboslovni podatki "zreducirani" na številke, ki ne morejo predstaviti kompleksnosti tovrstnih pojavov. Na obstoj napak v zbirkah podatkov kažejo napake v podatkih, ki se uporabljajo za uradne statistike. Ob tem je veliko podatkov zreduciranih na določene kategorije, kot na primer na etnično pripadnost ali socialni sloj, kar v nekaterih krogih vodi v zadržke o rabi numeričnih podatkov v družboslovnem raziskovanju. Smithova (2008) pa opozarja, da kompleksnost podatkov ne bi smela biti razlog, da sekundarnih podatkov ne bi uporabljali. Meni namreč, da lahko sekundarni podatki omogočijo še drug pogled na svet, lahko pripomorejo k identifikaciji trendov in neenakosti, kar je lahko izhodišče za nadaljnje, bolj poglobljene raziskave. Uradni statistični podatki so po njenem mnenju preveč pomembni, da jih ne bi vključili v raziskave oz. da bi njihovo interpretacijo prepustili politiki ali medijem. Poudarja, da se prav tu kaže ta pomen in vloga raziskovalcev, saj pomagajo povezati in osmisliti empirične podatke, družbeni kontekst in obstoječe teoretične modele.

Nadalje avtorji Yorke (2011), Smith (2008) ter Gorard in Smith (2004) kot pomembno šibkost sekundarne analize podatkov izpostavljajo dejstvo, da podatki mogoče ne predstavljajo točno tistih značilnosti, ki jih je raziskovalec želel raziskati: gre torej za kompromis med podatki, ki so na voljo, in podatki, ki bi jih želeli imeti, torej "idealnimi podatki". Raziskovalec se mora v zgodnji fazi raziskovalnega procesa odločiti, ali so podatki dovolj dobri za namen raziskave.

Yorke (2011) nadaljuje, da je treba ob vsaki raziskavi vzeti v obzir dejstvo, da so bili podatki zbrani za druge namene in da se je za vsakokratno situacijo treba odločiti, ali je tveganje upravičeno. Pri tem se mora raziskovalec zavedati, da podatki niso vedno najvišje kakovosti, saj se na primer lahko zgodijo napake pri vnosu in tudi v povezavi s hranjenjem podatkov. Pri zadnjem gre npr. za princip, ali in kako neki novejši podatki izpodrinejo stare (npr. višanje ocen). Baze podatkov mogoče niso kompatibilne, kar vodi v kompromis o kategorizaciji podatkov.

3 Masovni podatki

V zadnjem času je znotraj osrednjega koncepta sekundarnih podatkov opazen trend rabe masovnih podatkov, katerega poimenovanje izhaja iz angleške besedne zveze “big data” in je na pedagoškem področju pri nas res šibko poznano, kar pa sicer velja tudi za druge države (Smith, 2021). Masovni podatki se lahko različno definirajo, v osnovi pa je to poimenovanje za velike količine strukturiranih in nestrukturiranih podatkov (Smith, 2021). Brands (2014) opredeljuje masovne podatke kot proces združevanja enormnih količin na različne načine zbranih podatkov, ki se jih analizira z rabo kompleksnih algoritmov z namenom informiranja v procesu odločevanja.

Dejstvo je, da digitalizacija – tudi kot posledica izrednih razmer v povezavi s šolanjem na daljavo – vse bolj vstopa tudi v izobraževanje ter s tem omogoča zbiranje in množenje digitalnih podatkov, ki se lahko uporabijo za namene raziskovanja učenja in poučevanja ter v končni fazi za zviševanje učinkovitosti izobraževalnega procesa. Razširjena in, kot ugotavljata Müller in Kuprešak (2018, str. 95), tudi priljubljena raba e-okolij, forumov, spletnih pogovorov, spletnih učilnic odprtih izobraževalnih virov (angl. OER – open educational resources), raba pametnih telefonov za učenje izjemno povečujejo količino že zbranih podatkov, pri čemer pa pristopi k raziskovanju ostajajo še precej tradicionalni. Kljub temu pa se že kažejo nekatere rabe masovnih podatkov, ki uspešno izkoriščajo prednosti že zbranih podatkov (Anirban, 2014). Keržič (2022) podobno izpostavlja, da se med izobraževanjem na daljavo v različnih informacijskih sistemih hranijo podatki o študentu, prav tako pa izobraževalne ustanove na terciarni ravni s spletnimi vprašalniki zbirajo mnenja o vsebinah, ki se navezujejo npr. na kakovost študijskega procesa. V spletnih učnih okoljih se shrani vsaka interakcija, vsak klik v e-okolju, kar se zapiše v dnevniških datotekah (npr. vstop v učilnico, ogled vira, objava na forumu, oddaja naloge, odgovarjanje na vprašanja v kvizu). Vse te velike količine raznovrstnih in nestrukturiranih podatkov, shranjenih v različnih oblikah, lahko ob ustrezni interpretaciji predstavljajo pomemben vir informacij o procesu učenja posameznika. Bamiah idr. (2018) poudarjajo, da je z razmahom digitalnih tehnologij in povečano rabo MOOC-ov prišlo do spremenjenega načina poučevanja, učenja in ocenjevanja, pa tudi do spremenjenega dostopa do izobraževanja. Izobraževanje po njihovem mnenju prehaja v okolje pametnega kombiniranega učenja z vključevanjem najrazličnejših digitalnih orodij, vendar pa izobraževalne ustanove običajno nimajo resursov za upravljanje in analiziranje kompleksnega korpusa tako zbranih podatkov. Prav na osnovi teh podatkov pa je mogoče učinkovito napovedati npr. dosežke učencev. Računalniški napredek omogoča samodejno in v realnem času (ne naknadno) izvedeno zbiranje podatkov, njihovo hranjenje, obdelavo in analizo.

Pri uresničevanju potenciala masovnih podatkov v izobraževanju sta zlasti pomembna koncepta rudarjenje izobraževalnih podatkov (angl. EDM – Educational Data Mining) in učna analitika (angl. LA – Learning Analytics), na osnovi katerih se lahko preučuje vedenje in dosežke učečih se. Omogočata analizo procesa učenja skozi preučevanje interakcije v učnem okolju prek avtomatizirane interaktivne platforme, kot so npr. LMS, ITS Learning Management Systems (LMS), Intelligent Tutoring Systems (ITSs) (Prakash idr., 2014). Rudarjenje izobraževalnih podatkov se uporabi za iskanje novih vzorcev, algoritmov in modelov, kar poteka na podlagi principov statističnih metod in na podlagi metod rudarjenja podatkov. Te podatke se lahko raziskuje za oblikova-

nje prilagodljivih sistemov učenja, ki lahko podprejo proces učenja. Učna analitika se nanaša na merjenje in analizo ogromnih baz podatkov in njihovih kontekstov s ciljem izboljšanja učnih dosežkov (Prakash idr., 2014). Učna analitika vključuje interpretacijo teh baz podatkov in predstavljanje podatkov z orodij (npr. vizualnimi), ki omogočajo napovedovanje prihodnjih dosežkov in zaznavanje morebitnih šibkih področij (Yu in Wu, 2015) oz. vodijo k zbiranju in vzorčenju podatkov o zainteresiranosti učencev, njihovem ravnanju, rabi učnih pripomočkov, navadah, rabi jezika, trajanju pozornosti, o akademski uspešnosti itn. (Beneito-Montagur, 2017; Jarke in Breiter, 2019). Keržič (2022) navaja, da učna analitika predstavlja preplet računalniških in statističnih znanosti ter učnih in vedenjskih teorij ter vključuje merjenje in zbiranje podatkov, ki nastajajo med izobraževanjem, ter njihovo analizo, ki je prilagojena tako, da se na njeni osnovi lahko opredeli procese učenja in poučevanja. Siemens (2011) pa učno analitiko deli na učno in akademsko, pri čemer obema pripisuje pomen pri izboljševanju učnega procesa in učnih dosežkov. Meni, da se učna analitika lahko izpelje na osebni ravni, na ravni nekega predmeta ali programa, akademska pa na institucionalni, regionalni, nacionalni ali na mednarodni ravni.

Ko izobraževanje v celoti ali delno poteka z rabo digitalnih tehnologij in spletnih platform, se o njem s pomočjo učne analitike generirajo številni podatki (Oi idr., 2017), kar učitelju omogoča, da takoj dostopa do učenčevih učnih dosežkov in vzorcev sodelovanja ter da učencu takoj da povratno informacijo (Black in William, 2018). Pravočasna in seveda konstruktivna povratna informacija lahko pomembno vpliva na učenčevo nadaljnje ravnanje, saj ga dodatno spodbudi in usmeri (Zheng in Bender, 2019). Podatki učitelju omogočijo tudi analizo lastnih pristopov k poučevanju in ga tako spodbudijo za spremembe lastne poučevalne prakse v smislu prilagajanja procesa potrebam in zahtevam učencev.

Anirban (2014), Russom (2011) in Siemens (2011) poudarjajo, da analitika masovnih podatkov vključuje rabo naprednih metod analitike na velikih količinah podatkov ter tako omogoča odkrivanje vzorcev in osmišljeno rabo informacij. Avtomatizirana poročila analitike povečujejo učinkovitost, omogočajo boljši vpogled in zavedanje, kar lahko vodi v dvig kakovosti izobraževanja z vidika posameznika in ustanove. Vendarle pa se, kot poudarja Daniel (2014), analitika masovnih podatkov, tudi zaradi kompleksnosti identifikacije relevantnih podatkov, ki terjajo znanje, resurse in čas, na področju vzgoje in izobraževanja uvaja le počasi. Ta vključuje rudarjenje podatkov, analitiko podatkov in spletne prikazovalnike. Kot navajata Cope in Kalantzis (2015), analitika omogoča poglobljen vpogled v področje. Na pedagoškem področju se torej masovni podatki uporabljajo za izboljšanje izobraževanja, predvsem za spodbujanje k boljšim učnim dosežkom, vztrajanje v izobraževanju, preučevanje načinov učenja in poučevanja ter ocenjevanja in prilagajanje procesa učenja in tudi poučevanja glede na potrebe učencev in njihove sposobnosti ter tudi institucionalno odgovornost.

Prednosti masovnih podatkov

Raba masovnih podatkov po navedbah nekaterih avtorjev (Riffai idr., 2016) lahko pomembno pripomore k izboljšanju učne izkušnje. Podatki s področja vzgoje in izobraževanja se namreč lahko agregirajo za veliko število učečih se in se do njih lahko dostopa prek algoritmov za rudarjenje s ciljem izgradnje modela. Modeli se uporabijo za izgradnjo prilagodljivih sistemov učenja, v katerih se na osnovi modelskih napovedi

prilagodi/spremeni učenčevu naslednjo učno izkušnjo ali se priporoča neko drugo učno izkušnjo. Na osnovi analize masovnih podatkov se tako lahko poveča učiteljeva učinkovitost, ki na osnovi povratne informacije lažje poskrbi za zagotavljanje ustreznega, učinkovitega okolja za učenje in poučevanje ter se lažje prilagaja potrebam učečih se in virom, ki jih ima na voljo. Podatki in množična vključenost v vzgojno-izobraževalni sistem pomembno prispevajo k oblikovanju pričakovanih dosežkov, odgovornosti in dostopa, kar se lahko uresniči z zagotavljanjem infrastrukture in uvajanjem trajnostnih sprememb, ki vodijo v institucionalizacijo ustvarjanja, in izmenjave znanja (Amershi in Conati, 2009). Liñán in Pérez (2015) navajata, da masovni podatki omogočijo učitelju, da nemudoma pridobi povratne informacije, na osnovi katerih lahko opravi evalvacijo nekega predmeta ter tudi lastnega poučevanja in ocenjevanja. Učitelj lahko nadzoruje proces učenja in tako zelo hitro odkrije šibka področja, dejavnike tveganja glede na sposobnosti učencev in raven njihovega znanja. Russom (2011) dodaja, da masovni podatki omogočajo tudi individualizacijo, formativno ocenjevanje, participacijo, aktivno sodelovanje učečih se in sodelovalno učenje. Omogočajo tudi napovedovanje učnih izidov in sledenje napredku znotraj nekega modula. V slovenskem prostoru dobro poznan Moodle na primer zbira podatke o nekem predmetu, sledi aktivnosti učečega, njegovim interakcijam, oddaji nalog, prezentacijam ali kvizom.

Nadalje Russom (2011) poudarja vlogo masovnih podatkov za višjo kakovost izobraževanja, pri čemer se usmerja predvsem na visoko šolstvo. Skrb za kakovost v tem sektorju je predvsem posledica konkurence med visokošolskimi zavodi, postopkov akreditacije, evalvacije in različnih drugih regulacij, pa tudi, kot navaja Blažič (2021, str. 94), konkurenčnega trga dela. Hmelak idr. (2020, str. 147) dodajajo, da je kakovost študija odvisna tudi od časa in prostora, ne nazadnje pa se morajo visokošolske institucije odzivati tudi na nenehne spremembe v družbi. Presoja kakovosti temelji na analizi podatkov, ki se zbirajo prav v vseh vzgojno-izobraževalnih ustanovah, prav tako pa se med zaposlenimi dviguje zavedanje, da so vzgojno-izobraževalne ustanove odgovorne za zadovoljstvo in uspeh študentov ali drugih učečih se.

Masovni podatki po avtorjih Amershi in Conati (2009) vodijo v učinkovito odločanje, saj omogočajo zaznavanje, razumevanje, analizo in napovedovanje ravnanj učečih se, njihovega napredka in učnih izidov. Med procesom učenja in poučevanja v virtualnem okolju se lahko analizira vedenje učečih se med opravljanjem neke aktivnosti, lahko se analizira participacijo, diskusije na forumih in prav s tem se lahko pridobijo informacije o močnih področjih in šibkih točkah učečih se. Še več, Bichsel (2012) trdi, da morajo vzgojno-izobraževalne ustanove nujno investirati v profesionalne analitike, ki bodo pomagali definirati ključna vprašanja v celotnem procesu do točke, ko bodo razvili podatkovne modele za oblikovanje in informiranje o nevrvalgičnih točkah, priporočilih in poročilih.

Slabosti masovnih podatkov

Med najpogostejšimi izzivi v povezavi z masovnimi podatki se omenjajo: varnost, zasebnost, hranjenje podatkov in vizualizacija (Bamiah idr., 2018).

Z vidika varnosti se pri masovnih podatkih pojavljajo varnostni zadržki, ki se tudi sicer omenjajo v povezavi s spletno varnostjo, tudi zato, ker se kaže, da tradicionalne varovalne strategije (gesla, požarni zidovi, protivirusni programi) niso več dovolj. Z vi-

dika masovnih podatkov in izobraževalnih ustanov pa se dodatno poudarja še pomanjkljiva določila glede intelektualne lastnine, pa tudi dostopa do podatkov ter njihovega hranjenja, obdelave in analize (Daniel, 2014).

Nadalje se pojavi še vprašanje zasebnosti. Masovni podatki v izobraževanju so namreč transparentni, saj pretežno razkrivajo identiteto učečega se. Tako se v okviru MOOC-ov (angl. Massively Open Online Courses) zbirajo, centralizirajo in analizirajo podatki učečih se udeležencev, ki pa lahko hitro postanejo predmet zlorabe (Young, 2015). Raba masovnih podatkov se poleg iskanja načinov za ohranjanje zasebnosti dotika tudi drugih etičnih izzivov, npr. vprašanja posameznikove privolitve, lastništva podatkov in transparentnosti. Pojavijo pa se tudi vprašanja o avtentičnosti uporabnikov, pooblaščenosti oseb za dostop do podatkov in namenu uporabe, npr. ali gre na primer za sisteme napovedovanja, ki povezujejo demografske podatke s preteklimi in z zdajšnjimi učnimi dosežki, za angažiranost za delo s spletnimi materiali ali prisotnostjo, končnimi učnimi izidi itn. (Wang, 2016).

Pomemben izziv za rabo masovnih podatkov je tudi pomanjkanje ustrezno kompetentnih strokovnjakov (Bamiah idr., 2018). Russom (2011) celo meni, da je prav to največja ovira pri rabi in nadaljnjem razvoju masovnih podatkov. Pri tem je treba omeniti, da morajo strokovnjaki za rabo masovnih podatkov na področju izobraževanja znati pridobivati, analizirati in smiselno uporabiti podatke, hkrati pa morajo tudi znati identificirati konstrukte, ki se vsebinsko navezujejo na področje vzgoje in izobraževanja.

Z masovnimi podatki se seveda odpira tudi vprašanje hranjenja in analize podatkov ter ravnanja z njimi. Izobraževanje ustvarja ogromne količine podatkov v sklopu LMS, spletnih repozitorijev, digitalnih knjižnic, informacijskih sistemov, družabnih medijev, posameznih računalnikov in administrativnih sistemov, ki vključujejo informacije o predmetih in programih, učnih poteh itn. Ti podatki so zbrani v različnih formatih, kar je izziv za integracijo in hranjenje, pojavi pa se tudi vprašanje kakovosti teh podatkov, predvsem v povezavi z robustnostjo v oblikovanju meril za njihovo izbiro (Bamiah idr., 2018).

Baig idr. (2020) so pripravili sistematični pregled literature na področju rabe masovnih podatkov v izobraževanju. Ugotavljajo, da se je v pedagoškem raziskovanju trend uporabe masovnih podatkov pojavil leta 2014 in se znatno povečal v letu 2017. Identificirali so štiri najpogostejše teme, ki so bile naslovljene z rabo masovnih podatkov. Prva taka tema je bila vedenje in dosežki učencev, znotraj katere so avtorji ugotavljali odnos učencev, njihovo zadovoljstvo, strategije, vedenje, pri čemer so se naslanjali na učno analitiko. Druga najpogostejša tema je bila modeliranje in hranjenje izobraževalnih podatkov. Gre za študije, ki so uvedle modeliranje z masovnimi podatki in analizirale orodja za analizo teh podatkov, skupaj s "skladišči" podatkov. Raziskovale so okolje oblaka za shranjevanje podatkov ter uporabile klastersko analizo za ugotavljanje dostopnosti in obdelovanje podatkov s področja vzgoje in izobraževanja. Tretja tema je bila izboljšanje izobraževalnega sistema. Te študije so se usmerjale v analizo statističnih orodij in meritev ter z njimi povezanih izzivov, pa tudi v učinkovitost informacijsko-komunikacijske tehnologije in rabo spletnih strani kot vira za izboljšanje izobraževalnega sistema. Zadnje področje, ki so ga avtorji poudarili, je vključitev masovnih podatkov v kurikulum. Gre za študije, ki vključujejo tematike masovnih podatkov v različne predmete in poudarjajo vlogo masovnih podatkov za izobraževanje.

4 Sklep

V prispevku smo se najprej osredinili na sekundarno analizo podatkov, za katero ugotavljamo, da ima številne prednosti, med katerimi gre poudariti predvsem dejstvo, da so podatki že zbrani in da raziskovalec tako prihrani čas za pripravo instrumenta za zbiranje podatkov kot tudi za zbiranje podatkov, kar je dragoceno z vidika raziskovalcev začetnikov in tudi zelo koristno za študente, ki začenjajo raziskovanje. Sekundarno analizo podatkov odlikuje tudi možnost, da se podatke analizira z drugega zornega kota, da se torej postavi drugačna raziskovalna vprašanja, pa tudi, da se analizo lahko nadgradi z naprednejšimi načini obdelave podatkov, kar seveda pripomore h kompleksnejšim znanstvenim spoznanjem. Na tem mestu bi poudarili, da so tudi v Sloveniji na voljo podatki nacionalnih študij (npr. PISA, TIMMS), ki jih hrani Pedagoški inštitut, nekateri podatki drugih študij pa so na voljo tudi na spletni strani Arhiva družboslovnih podatkov. Velja pa seveda izpostaviti tudi zadržke pri sekundarni analizi podatkov, predvsem da se mora raziskovalec zavedati, da so lahko podatki tudi napačni in da se mora vsakič znova vprašati, ali podatki, ki primarno niso bili zbrani za cilje njegove raziskave, dejansko lahko odgovorijo na zastavljena raziskovalna vprašanja.

V drugem delu smo več prostora namenili masovnim podatkom, za katere ugotavljamo, da se v pedagoškem raziskovanju kljub pomembnim prednostim njihova vloga in vrednost uveljavljata le počasi. V praksi se masovne podatke največkrat upravlja v smislu rudarjenja podatkov za t. i. institucionalne evalvacije, npr. reakreditacije, samo-evalvacije itn., torej predvsem v terciarnem izobraževanju. Uporabljajo pa se tudi za napoved dosežkov, raziskovanje vedenja, sledenje prisotnosti/odsotnosti, priporočila za izvedbo predmetov, takojšnjo pomoč in ocenjevanje, za vizualno analitiko učenčeve interakcije v diskusiji znotraj foruma, povečano dostopnost izobraževanja, raziskovanje in razvoj. Predvsem vidimo prednosti tega pristopa v:

- rabi učne analitike za prilagajanje poučevanja sposobnostim učencev in razpoložljivim virom;
- napovedovanju prihodnjih dosežkov učečih se in iskanju vzorcev za identifikacijo dejavnikov tveganja za neuspeh ali opuščanje šolanja;
- evalvaciji;
- v višji dostopnosti izobraževanja, pripomore pa lahko tudi k izboljšanju raziskovanja, evalvaciji in odgovornosti.

Tina Štemberger, PhD

Secondary Data Analysis and Big Data in Educational Research

The paper addresses two concepts which appear more and more often in general professional discussions, but that have not yet received sufficient attention in educational research in our country, which is why they are less known, less used and modestly implemented. The first part of the paper focuses on secondary data analysis, which is not a new phenomenon (see for example Kozina et al., 2012, p. 136; Kunčič et al., 2013, p. 48;

Čepar et al., 2022, p. 132). Basically, secondary data are defined as data that someone has previously collected and used (Smith, 2008), with Schutt (2007) stating that this applies even if the data are reused by the same person or group of people that collected and processed the data. It follows that it is not always possible to draw a clear line between primary and secondary data. Therefore, Jary and Jary (2000, p. 540) propose a more general definition of secondary data analysis, namely that it is any research based on the re-analysis of previously analysed data. As Glaser (1963) pointed out decades ago, by using secondary data we not only save time and money, but also preserve careers, research interests and self-image, and thus in a sense “save” huge amounts of data from being lost.

Smith (2008) also states that secondary data can provide opportunities to develop new research ideas, such as discovering different relationships, associations, or differences than were hypothesized in the primary research. It is further emphasized that secondary data are usually of high quality and highly available from a financial perspective (Hakim, 1982; Smith, 2008; Yorke, 2011). These two aspects are important because data collection can be a very costly process in many cases and therefore access to high-quality data is important for both novice and more experienced researchers, because it allows them a high degree of research independence. By omitting the data collection phase, both novice researchers and students can gain valuable experience conducting independent research in an area of interest and have the opportunity to publish the results. Smith's (2008) main objection to the use of secondary data is that existing databases may contain many errors and that social science data are “reduced” to numbers that cannot reflect the complexity of such phenomena.

In recent years, there is a clear trend towards the use of Big Data, which is hardly known in the educational field in our country, but this is also the case in many other countries (Smith, 2021). Big Data can be defined in different ways, but basically it is a term for large amounts of structured and unstructured data (Smith, 2021). Brands (2014) defines Big Data as the process of combining enormous amounts of differently collected data, which are analysed using complex algorithms, with the goal of providing information for the decision-making process.

The fact is that digitization – also as a result of COVID-related emergency distance learning – is increasingly entering education, and thus enabling the collection and multiplication of digital data that can be used to research learning and teaching and, ultimately, to increase the efficiency of the educational process. The widespread use of e-environments, forums, online conversations, online classrooms with open educational resources (OER), and the use of smartphones for learning greatly increase the amount of data already collected, while research approaches are still quite traditional.

In realizing the potential of Big Data in education, the concepts of Educational Data Mining (EDM) and Learning Analytics (LA) are particularly important, based on which the behaviour and performance of learners can be studied. They enable the analysis of the learning process by examining interaction in the learning environment through an automated interactive platform, such as Learning Management Systems (LMS), Intelligent Tutoring Systems (ITSs) (Prakash et al., 2014). Educational data mining is used to find new patterns, algorithms and models, and is based on both the principles of statistical methods and the data mining methods. Data can be explored to design adaptive learning systems that can support the learning process. Learning analytics refers to the measurement and analysis of large databases and their contexts

with the goal of improving learning outcomes (Prakash et al., 2014). Learning analytics involves interpreting these databases and presenting data with tools (e.g., visual) that can predict future performance and identify potential weaknesses (Yu and Wu, 2015) or lead to the collection and sampling of data on students' interest, their behaviour, use of teaching aids, habits, use of language, attention span, academic performance, etc. (Beneito-Montagur, 2017; Jarke and Breiter, 2019). Keržič (2022) states that learning analytics is an intertwining of computer and statistical sciences with learning and behavioural theories, and involves the measurement and collection of data generated during instruction as well as their analysis, which is adapted to define the learning processes based on them. When education partly or in whole takes place with the use of digital technologies and online platforms, learning analytics generates a lot of data about it (Oi et al., 2017), which enables the teacher to immediately access the data on students' learning performance and participation patterns, so that he or she can provide immediate feedback to the learner (Black and William, 2018). Timely and naturally constructive feedback can have a significant impact on the learner's further behaviour, providing additional encouragement and guidance (Zheng and Bender, 2019). The data also allows the teachers to analyse their own teaching methods and thus encourage them to make changes to their own teaching practice in terms of adapting the process to the needs and demands of the students.

In addition, Russom (2011) emphasizes the role of Big Data for higher quality education, which is mainly discussed at the level of higher education. Concern about quality in this sector is primarily a result of competition among higher education institutions, accreditation procedures, evaluation and other regulations. Quality assessment is based on an analysis of data collected across educational institutions, and the employees' awareness that higher education institutions are responsible for student satisfaction and success.

Some of the most common challenges associated with Big Data are: security, privacy, data storage and visualization (Bamiah et al., 2018). The security concerns with Big Data are mainly related to the general online security issues, also because it appears that traditional security strategies (passwords, firewalls, antivirus programs) are no longer sufficient. From the perspective of Big Data and educational institutions, the insufficient regulations on intellectual property, as well as access to data and its storage, processing and analysis are additionally emphasized (Daniel, 2014).

Furthermore, the issue of privacy also arises; Big Data in education is transparent, as it mainly reveals the identity of learners. For example, in the context of MOOCs (Massively Open Online Courses), learners' data is collected, centralized and analysed, but can quickly become an object of misuse (Young, 2015). In addition to the question of how to preserve privacy, the use of Big Data also raises other ethical challenges, such as issues of individual consent, data ownership and transparency. There are also questions about the authenticity of the users and those authorized to access the data, as well as questions about the intended use, such as whether there are predictive systems that link demographic data to past and present learning performance, engagement to work with online materials or presence, final learning outcomes, etc. (Wang, 2016).

An important challenge in using Big Data is also the lack of sufficiently competent experts (Bamiah et al., 2018). Russom (2011) even believes that this is the biggest obstacle in the use and further development of Big Data. Experts in the use of Big Data

in education need to be able to collect, analyse and meaningfully use data, while also being able to identify constructs that are substantively related to education.

For secondary data analysis, we can sum up that it has many advantages, including the fact that the data has already been collected, which saves the researcher time not having to prepare the instrument for data collection and perform the data collection itself, which is valuable from the point of view of novice researchers and also very useful for students beginning their research. Secondary data analysis is also distinguished by the possibility of analysing the data from another point of view, so that different research questions can be proposed, as well as by the fact that the analysis can be extended with more advanced methods of data processing, which naturally contributes to more complex scientific findings. At this point, we would like to point out that in Slovenia data from national studies (e.g., PISA, TIMMS) are also available, which are kept by the Slovenian Educational Research Institute, while some data from other studies are also available on the website of the Social Science Data Archive. However, it is also necessary to point out caveats in secondary data analysis, especially that the researchers must be aware that the data might be wrong and that they must ask themselves every time whether the data, which were not primarily collected for the purposes of their research, can actually answer the research questions posed.

Regarding Big Data, we conclude that their role and value in educational research has been slow to take hold, despite the significant benefits. In practice, Big Data are most often managed in terms of data mining for so-called institutional evaluations (e.g., reaccreditation, self-evaluation, etc.), i.e., mainly in tertiary education. They are also used for performance prediction, behavioural research, attendance/absence tracking, course recommendation, instant help and assessment, visual analytics of student interaction, forum discussions, improved access to education, research and development. Above all, we see the benefit of this approach in:

- using learning analytics to match instruction to student abilities and available resources;
- predicting future learner performance and looking for patterns to identify risk factors for failure or dropping out of school;
- evaluation;
- its contribution to greater educational accessibility and improved research assessment and accountability.

LITERATURA

1. Amershi, S. in Conati, C. (2009). Combining Unsupervised and Supervised Classification to Build User Models for Exploratory Learning Environments. *Journal of Educational Data Mining*, 1(1), 18–71.
2. Anirban, S. (2014). Big Data Analytics in the Education Sector: Needs, Opportunities, and Challenges. *International Journal of Research in Computer and Communication Technology (IJRCCT)*, 3(11), 2278–5841.
3. Baig, I. M., Shuib, L. in Yadegaridehkordi, E. (2020). Big Data in Education: a State of the Art, Limitations, and Future Research Directions. *International Journal of Educational Technology in Higher Education*, 17(44), 1–23. Dostopno na: <https://doi.org/10.1186/s41239-020-00223-0> (pridobljeno 7. 7. 2022).

4. Bamiah, M. A., Brohi, S. N. in Bashari, B. R. (2018). Big Data Technology in Education: Advantages, Implementations, and Challenges. *Journal of Engineering Science and Technology, Special Issue on ICCSIT*, 229–241.
5. Beneito-Montagut, R. (2017). Big Data and Educational Reserach. V: Wyse, D., Selwyn, N., Smith, E. idr. (ur.). *The BERA. Sage Handbook of Educationa Research* (str. 913–914). London: Sage.
6. Bichsel, J. (2012). *Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations*. Louisville: EDUCAUSE Center for Applied Research.
7. Black, P. in Wiliam, D. (2018). Classroom Assessment and Pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575.
8. Blažič, M. (2021). Prispevek visokošolskega učnega okolja h kariernemu razvoju študentov. *Didactica Slovenica – Pedagoška obzorja*, 36(1), 93–113.
9. Brands, K. (2014). *TECH Practices. Big Data and Business Intelligence for Management Accountants. Strategic Finance*. Dostopno na: <http://sfmagazine.com/wpcontent/uploads/sfarchive/2014/06/TECH-PRACTICES-Big-Data-and-Business-Intelligence-for-Management-Accountants.pdf> (pridobljeno 3. 4. 2022).
10. Bulmer, M. (1980). Why Don't Sociologists Make More Use of Official Statistics? *Sociology*, 14(4), 505–524.
11. Campbell, J. P., Daft, R. L. in Hulin, C. L. (1982). *What to Study: Generating and Developing Research Questions*. Beverly Hills: Sage.
12. Cencič, M. (2009). *Kako poteka pedagoško raziskovanje: primer kvantitativne empirične neeksperimentalne raziskave*. Ljubljana: Zavod Republike Slovenije za šolstvo.
13. Cook, T. D. (1974). The Potential and Limitations of Secondary Evaluations. V: Apple, M. W., Subkoviak, M. J. in Luffer, H. S. (ur.). *Educational Evaluation: Analysis and Responsibility* (str. 15–48). California: McCutchan.
14. Cope, B. in Kalantzis, M. (2015). Sources of Evidence-of-Learning: Learning and Assessment in the Era of Big Data. *Open Review of Educational Research*, 2(1), 194–217.
15. Čepar, Ž., Likar, B. in Kunc, P. (2022). Povezovanje srednjega šolstva in gospodarstva ter kadrovske štipendije. *Didactica Slovenica – Pedagoška obzorja*, 37(1), 125–140.
16. Daniel, B. (2014). Big Data and Analytics in Higher Education: Opportunities and Challenges. *British Journal of Educational Technology*, 46(5), 904–920.
17. Glaser, B. G. (1963). Retrading Research Materials: the Use of Secondary Analysis by the Independent Researcher. *American Behavioral Scientist*, 6(10), 11–14.
18. Gorard, S. (2002). The Role of Secondary Data in Combining Methodological Approaches. *Educational Review*, 54(3), 231–237.
19. Gorard, S. in Smith, E. (2004). An International Comparison of Equity in Education Systems. *Comparative Education*, 40(1), 15–28.
20. Gorard, S., See, B. H., Smith, E. idr. (2006) *Teacher Supply: the Key Issues Workforce*. London: Continuum.
21. Hakim, C. (1982). *Secondary Analysis in Social Research: A Guide to Data Sources and Methods with Examples*. London: Allen & Unwin.
22. Hmelak, M., Rudaš, A. in Lepičnik Vodopivec, J. (2020). Vključevanje študentov v razvoj inovativnih izobraževalnih modelov. *Didactica Slovenica – Pedagoška obzorja*, 35(3–4), 147–163.
23. Jarke, J. in Breiter, A. (2019). Editorial: the Datafication of Education. *Learning, Media and Technology*, 44(1), 1–6.
24. Jary, D. in Jary, J. (2000). *Collins Dictionary of Sociology*. Glasgow: Harper Collins.
25. Keržič, D. (2022). *Učna analitika kombiniranega učenja v visokem šolstvu*. [Doktorska disertacija]. Ljubljana: Fakulteta za družbene vede.
26. Kozina, A., Rožman, M., Vršnik Perše, T. idr. (2012). Napovedna vrednost različnih ocen šolske klime za dosežke v raziskavah TIMSS. *Didactica Slovenica – Pedagoška obzorja*, 27(1–2), 127–144.
27. Kunčič, Š., Razdevšek-Pučko, C. in Rugelj, J. (2013). Spletni dnevnik v prvem obdobju osnovne šole. *Didactica Slovenica – Pedagoška obzorja*, 28(3–4), 43–56.

28. Liñán, C. in Pérez, J. (2015). Educational Data Mining and Learning Analytics: Differences, Similarities, and Time Evolution. *RUSC. International Journal of Educational Technology in Higher Education*, 12(3), 98–112.
29. Müller, M. in Kuprešak, I. (2018). Perceptions of High School Students of the Use of ICT in the Process of a Foreign Language. *Didactica Slovenica – Pedagoška obzorja*, 33(1), 95–103
30. Oi, M., Yamada, M., Okubo, F. idr. (2017). Reproducibility of Findings from Educational Big Data. V: *Proceedings of the Seventh International Learning Analytics Knowledge Conference* (str. 536–537). New York: Association for Computing Machinery.
31. Petlák, E. (2021). Self-Reflecionas as Basis of a Teacher's Work. *Didactica Slovenica – Pedagoška obzorja*, 36(3–4), 41–54.
32. Prakash, B., Hanumanthappa, M. in Kavitha, V. (2014). Big Data in Educational Data Mining and Learning Analytics. *International Journal of Innovative Research in Computer and Communication Engineering*, 2(12), 2320–2320.
33. Riffai, M., Edgar, D., Duncan, P. idr. (2016). The Potential for Big Data to Enhance the Higher Education Sector in Oman. 3rd MEC International Conference on Big Data and Smart City. *IEEE Xplore*.
34. Russom, P. (2011). Big Data Analytics. TDWI Best Practices Report. Dostopno na: <https://vivomente.com/wp-content/uploads/2016/04/big-data-analytics-white-paper.pdf> (pridobljeno 7. 6. 2022).
35. Siemens, G. (2011). How Data and Analytics can Improve Education. Dostopno na: <https://www.oreilly.com/ideas/education-data-analytics-learning> (pridobljeno 16. 4. 2022).
36. Smith, E. (2021). Secondary Data. V: Coe, R., Waring, C. in Hedges, L. (ur.). *Research Methods & Methodologies in Education* (str. 144–152). London: Sage.
37. Smith, E. (2008). Pitfalls And Promises: The Use of Secondary Data Analysis In Educational Research. *British Journal of Educational Studies*, 56(3), 323–339.
38. Sobal, J. (1981). Teaching With Secondary Data. *Teaching Sociology*, 8(2), 149–170.
39. Štemberger, T. (2020). Uvod v pedagoško raziskovanje. Koper: Založba Univerze na Primorskem.
40. Wang, Y. (2016). Big Opportunities and Big Concerns of Big Data in Education. *TechTrends*, 60(4), 381–384.
41. Yorke, M. (2011). Analysing Existing Datasets: Some Considerations Arising from Practical Experience. *International Journal of Research & Method in Education*, 34(3), 255–267.
42. Young, E. (2015). Educational Privacy in the Online Classroom: FERPA, MOOCs, and the Big Data Conundrum. *Harvard Journal of Law and Technology*, 28(2), 1–44.
43. Yu, X. in Wu, S. (2015). Typical Applications of Big Data in Education. *International Conference of Educational Innovation through Technology (EITT)* (str. 103–105). Wuhan: Conference Publishing Services.
44. Zheng, M. in Bender, D. (2019). Evaluating Outcomes of Computer-Based Classroom Testing: Student Acceptance and Impact on Learning and Exam Performance. *Medical Teacher*, 41(1), 75–82.

Dr. Tina Štemberger (1980), izredna profesorica za področje pedagoške metodologije na Pedagoški fakulteti Univerze v na Primorskem.

Naslov/Address: Podgraje 27, 6250 Ilirska Bistrica, Slovenija

Telefon/Telephone: (+386) 041 782 741

E-mail: tina.stemberger@pef.upr.si