

METODE VREDNOVANJA WEB UPOTREBLJIVOSTI

WEB USABILITY EVALUATION METHODS

mr.sc. Dijana Plantak Vukovac, Tihomir Orehovački

SAŽETAK

Dizajn orijentiran korisniku (eng. user-centered design, UCD) podrazumijeva dizajn sučelja imajući na umu potrebe, iskustvo i ograničenja krajnjih korisnika sustava. U kontekstu korištenja web sučelja, karakteristike korisnika web stranica samo su djelomično poznate. Stoga dizajn web stranica treba biti takav da i manje iskusnim korisnicima omogućuje jednostavnu, brzu i učinkovitu interakciju s web stranicom. Atribut kvalitete pomoću kojeg se procjenjuje u kojoj mjeri korisnik može uspješno, učinkovito i sa zadovoljstvom koristiti određeni proizvod (hardver, softver, web sučelje) da bi postigao specifične ciljeve u danom kontekstu naziva se upotrebljivost.

Dizajn web mjesta s aspekta upotrebljivosti podrazumijeva usklađivanje dizajna sa smjernicama web upotrebljivosti, kao i vrednovanje upotrebljivosti kako bi se identificirali i ispravili problemi upotrebljivosti tijekom razvojnog procesa web mjesta. U ovom radu dan je pregled najčešće korištenih analitičkih i empirijskih metoda vrednovanja web upotrebljivosti, te kriterija važnih za njihov odabir. Također su dane smjernice za jednostavnu integraciju nekih metoda u proces razvoja web mjesta.

ABSTRACT

User-centered design (UCD) refers to interface design where needs, experiences and limitations of a system's end users are taken into consideration. In the context of web interfaces characteristics of end users that visit web pages are only partially known. Web sites should therefore be designed in a manner that enables less experienced users an easy, quick and efficient interaction with web pages. The attribute of quality that allows for the assessment of the extent to which a product (hardware, software, web interface) can be used by a specified user effectively, efficiently and satisfactorily to achieve specified goals within a specified context of use is referred to as usability.

When web usability is considered, the design of web sites should be based on web usability design guidelines and include usability evaluation to identify and rectify usability problems during the web development process. This paper presents an overview of the most common analytic and empiric usability evaluation methods as well as the relevant criteria for methods selection. Guidelines for an easy integration of some of the methods in the website development process are also introduced.

1. UVOD

Tijekom posljednjih dvadesetak godina kroz istraživanja u području znanstvene discipline interakcija čovjeka i računala (eng. *human-computer interaction*, HCI) definirane su mnogobrojne smjernice i kriteriji za dizajn upotrebljivih korisničkih sučelja te su razvijene metode za njihovo vrednovanje. Koncept upotrebljivosti vuče svoje korijene iz općeg dizajna informacijskih sustava i potrebe da se sučelje sustava prilagodi korisnicima sustava.

Od sredine devedesetih godina 20. st., sa sve većim prodorom i korištenjem različitih web aplikacija, fokus istraživanja prebacio se na web upotrebljivost, zbog uočene tendencije da web dizajneri i projektanti razvijaju loša web mjesta i ne razmatraju pitanja upotrebljivosti web stranica [3].

Općenito, upotrebljivost (eng. *usability*) je svojstvo proizvoda ili sustava koje korisniku omogućuje jednostavno, učinkovito i ugodno izvršavanje željenih zadataka. Najjednostavnije, označava kvalitetu sustava ili proizvoda pri njegovoj upotrebi (eng. *quality in use*) [4]. Prema standardu ISO 9241-11 upotrebljivost se definira kao "opseg u kojem određeni korisnik može uspješno (*effectively*), učinkovito (*efficiently*) i sa zadovoljstvom (*satisfaction*) koristiti određeni proizvod da bi postigao specifične ciljeve u danom kontekstu" [15]. Jakob Nielsen [33], najpoznatiji svjetski stručnjak i promicatelj web upotrebljivosti (tzv. *usability guru*), pod

upotrebljivošću podrazumijeva atribut kvalitete pomoću kojeg se procjenjuje koliko je jednostavno koristiti sučelje, a također i metode kojima možemo poboljšati upotrebljivost sučelja prilikom njegovog dizajna. Navodi pet komponenti upotrebljivosti: jednostavnost učenja sučelja (*learnability*), učinkovitost postizanja cilja (*efficiency*), pamtljivost sučelja (*memorability*), prevencija grešaka pri izvođenju zadatka (*errors*) i zadovoljstvo u korištenju sučelja (*satisfaction*). Neki drugi autori [43] dodaju i parametre poput uspješnosti (*effectiveness*), tj. točnosti i potpunosti izvršenja zadatka, i korisnosti (*utility*), tj. mogućnosti sustava da pruži uslugu koja se očekuje, dok zadovoljstvo korisnika svrstavaju u kategoriju korisničkog iskustva (*user experience*), koja još uključuje parametre motivacije, zabave, ugone itd.

Definicije upotrebljivosti razlikuju se jer velikim dijelom ovise o atributima ili kriterijima pomoću kojih će se upotrebljivost mjeriti. Upotrebljivost se ne može direktno mjeriti, ali operacionalizacijom konstrukata upotrebljivosti, pojedini njeni aspekti mogu se izmjeriti objektivnim ili subjektivnim metrikama [14]. Na primjer, učinkovitost mjerimo pomoću vremena koje je korisnik utrošio na izvršenje određenog zadatka, dok uspješnost izražavamo pomoću postotka ukupno izvršenih zadataka [6, str. 240] i te mjere svrstavamo u objektivne metrike. S druge strane, subjektivne metrike identificiraju percepciju korisnika ili njihove stavove o sučelju, interakciji ili zadovoljstvu prilikom korištenja sučelja.

2. WEB UPOTREBLJIVOST

Većina korisnika weba dolazi na određeno web mjesto u potrazi za informacijama [31] i stoga im kroz različite aspekte web dizajna, a ponajprije dizajn informacija i navigacije, treba omogućiti jednostavno i brzo pronalaženje tražene informacije. Međutim, razna istraživanja pokazuju da ljudi teško pronalaze određenu web stranicu na nižoj hijerarhijskoj razini, a u još manjoj mjeri uspijevaju doći do web stranice i obaviti određeni zadatak (kupiti proizvod, ispuniti prijavu itd.) [31]. Stoga je za uspješnost nekog web mjesta i pozitivno korisničko iskustvo važno podići razinu upotrebljivosti web mjesta.

Razmatranje upotrebljivosti u fokusu je tzv. dizajna orijentiranog korisniku (*eng. user-centered design, UCD*), koji se temelji na planiranju, dizajnu, razvoju i vrednovanju proizvoda/sustava uz razmatranje potreba, iskustva i ograničenja korisnika te njihovo uključivanje u razvoj u svim fazama životnog ciklusa sustava. Standard *ISO 13407: Human-centred design process* opisuje generički proces razvoja sustava koji uvažava potrebe korisnika. Četiri aktivnosti koje je potrebno provesti kroz faze razvoja sustava su: definiranje konteksta upotrebe, definiranje zahtjeva (naručitelja, ali i korisnika), izrada većeg broja dizajnerskih rješenja, te vrednovanje dizajna (vrednovanje upotrebljivosti kao dijela cjelokupnog vrednovanja kvalitete sustava) [16].

Dizajn orijentiran korisniku također je nužno ostvariti i prilikom razvoja web stranica, bez obzira na razinu njihove složenosti (jednostavno web mjesto, web aplikacija ili složeni web sustav). Neophodno je identificirati ciljanu publiku (npr. studenti; roditelji; poslovna skupina ljudi; ili što šira populacija internet korisnika), njene karakteristike (npr. godine, stupanj obrazovanja, računalna pismenost, iskustvo na internetu itd.) i predvidjeti u kakvim uvjetima će pristupati web stranicama (npr. stolno računalo s *wide-screen* monitorom i širokopojasnom Internet vezom, *netbook* i bežični pristup ili mobitel i GPRS prijenos). Pristupom orijentiranom korisniku istovremeno se, uz zadovoljavanje ciljeva posjetitelja web stranica, nastoje uvažiti zahtjevi koje postavlja naručitelj, tj. nastoje se postići poslovni i marketinški ciljevi web mjesta. Iako ciljevi posjetitelja i naručitelja mogu biti u konfliktu, te ih je ponekad teško izbalansirati, njihova realizacija na web mjestu povećava cjelokupnu kvalitetu web mjesta.

Da bi se izradile upotrebljive web stranice, potrebno je poštovati smjernice dizajna koje povećavaju web upotrebljivost. Različiti istraživači, praktičari i institucije kreirali su smjernice web upotrebljivosti, među kojima su poznatije Nielsenove smjernice objavljene u više njegovih knjiga [31, 35], izvješća i portalu *www.useit.com*, zatim smjernice upotrebljivosti za web stranice američke javne uprave (*www.usability.gov*) [49] ili osnovni principi dizajna web stranica Lyncha i Hortonove [24]. Njihove smjernice [24, 31, 35, 49] fokusiraju se na sljedeće elemente dizajna web mjesta: strukturu web mjesta, navigaciju i pretraživanje, dizajn informacija, strukturu web stranice (naročito početne web stranice), grafički dizajn web stranice, tipografiju, korištenje grafike i multimedije, stil pisanja, označavanje poveznica, imenovanje elemenata web stranica, dostupnost, korisničko iskustvo itd. Također su definirali specifične smjernice za dizajn intraneta, *e-commerce* web mjesta, web mjesta s međunarodnim korisnicima i slično.

Nielsen je kroz godine analiziranja web dizajna različitih web mjesta objavio više tisuća smjernica za web upotrebljivost. Njih 113, koje se odnose na upotrebljivost početne web stranice, objavio je u knjizi "Homepage Usability: 50 Websites Deconstructed", dok u deset najvažnijih smjernica spadaju sljedeće [31]:

1. logotip i ime organizacije trebaju biti na svakoj web stranici – logotip treba imati poveznicu na početnu web stranicu (osim na početnoj web stranici);
2. ako web mjesto ima više od 100 web stranica, treba omogućiti pretraživanje web mjesta;
3. naslovi stranica trebaju biti jednostavni i direktni tako da jasno objašnjavaju svrhu stranice i da imaju smisla kada se pojave u rezultatima pretraživanja web pretraživača;
4. struktura stranice treba biti takva da olakšava brzo pregledavanje – grupiranjem i korištenjem poglavlja rascjepkajte dugačak sadržaj stranice u nekoliko manjih cjelina;
5. umjesto širokog objašnjavanja neke teme ili proizvoda na jednoj stranici, koristite hipertekst za strukturiranje sadržaja kroz kratak pregled karakteristika na osnovnoj stranici, a na nekoliko sekundarnih stranica detaljno razradite određenu podtemu. Na ovaj način posjetitelji weba mogu zaobići podteme koje ih ne interesiraju;
6. koristite fotografije proizvoda, ali izbjegavajte stranice s mnogo fotografija iste grupe proizvoda. Umjesto toga stavite jedan umanjeni prikaz fotografije (*eng. thumbnail*) na svaku stranicu pojedinog proizvoda i poveznice na jednu ili više većih fotografija na sekundarnim stranicama (omogućite zumiranje, rotiranje i slično, već prema prirodi proizvoda). Time je osnovna stranica proizvoda ograničena na ikoničku fotografiju i brzo se učitava;
7. kod priprema umanjenih prikaza slika ograničiti se na neki važan detalj na slici, umjesto da se samo smanje velike slike jer će tako objekti na njoj postati neprepoznatljivi;
8. koristite jasne i kratke nazive poveznica, tako da posjetitelji weba znaju kuda će ih poveznica odvesti prije nego kliknu na nju;
9. osigurajte da su sve važnije web stranice pristupačne ljudima s invaliditetom, naročito slabovidnim i slijepim osobama;
10. radite ono što i drugi rade na svojim web mjestima – ako se nešto koristi na određeni način na većini web mjesta, posjetitelji weba će isto očekivati i za vaše web mjesto. "Jakobov zakon o iskustvu web korisnika" kaže da korisnici većinu vremena provode na nekim *drugim* web mjestima pa tamo formiraju svoje mišljenje o tome kako web stranice trebaju funkcionirati.

Kako bi se identificirali problemi upotrebljivosti i na vrijeme ispravili, vrednovanje upotrebljivosti poželjno je provesti tijekom svih faza razvoja sustava. Zašto je vrednovanje web upotrebljivosti važno? Postoji više razloga za to:

- korisnik koji ima teškoća u korištenju web stranica jer, npr. ne može pronaći traženi sadržaj na web mjestu, ili teško može pročitati sadržaj, ili se izgubi na web mjestu zbog loše navigacije, otići će na konkurentsko web mjesto [33];
- korisnik koji ostvaruje svoje ciljeve na web mjestu je zadovoljan korisnik, postaje lojalan korisnik koji se uvijek vraća te će lako preporučiti proizvode ili usluge web mjesta ostalima [51];
- dizajneri/projektanti nisu tipični korisnici sustava zbog poznavanja proizvoda/sustava kojeg razvijaju, ali i općenito zbog više razine računalne pismenosti

[36], te mogu zanemariti neke važne aspekte dizajna koji su važni s gledišta korisnika;

- iako je problem upotrebljivosti manje izražen nego ranijih godina, dizajneri još uvijek ponavljaju greške dizajna koje predstavljaju kritične probleme upotrebljivosti, npr. otvaranje pop-up prozora, vizualno nerazlikovanje neposjećenih i posjećenih poveznica, web dizajn koji sliči na oglas itd. [35];
- provedba vrednovanja upotrebljivosti u ranim fazama životnog ciklusa sustava smanjuje mogućnost potrebe većih izmjena u redizajnu na kraju razvojnog ciklusa, čime se smanjuju vrijeme i troškovi razvoja i održavanja [50];
- otkrivanjem problema upotrebljivosti i redizajnom stranica usklađenim sa smjernicama dizajna upotrebljivosti povećava se postotak prodaje proizvoda za 100%, posjećenost za 150% ili korisnička učinkovitost za 160% [32].

Ukupno gledajući, web dizajneri se slažu da izrada web mjesta prema smjernicama upotrebljivosti i dostupnosti povećava kvalitetu web mjesta te smanjuje troškove razvoja i olakšava održavanje web mjesta [41].

3. PREGLED METODA ZA VREDNOVANJE UPOTREBLJIVOSTI

Vrednovanje upotrebljivosti provodi se pomoću različitih metoda koje je moguće svrstati u tri različite kategorije [17]:

- metode pregledavanja (eng. *inspection methods*),
- metode testiranja (eng. *testing methods*),
- metode ispitivanja (eng. *inquiry methods*).

Prva kategorija metoda su tzv. analitičke metode koje koriste stručnjaci koji se bave procjenom upotrebljivosti, dok se druge dvije kategorije odnose na empirijske metode kojima se upotrebljivost testira pomoću korisnika. U svakoj će od spomenutih kategorija najkorištenije metode biti detaljno opisane u nastavku, dok su one manje važne ukratko opisane u tablici 1 koja se nalazi u Dodatku A na kraju članka.

3.1 Metode pregledavanja

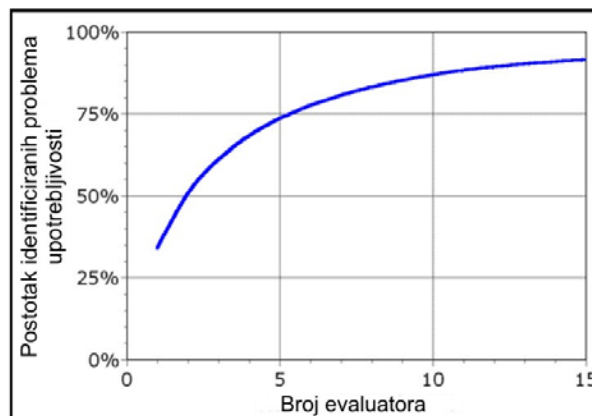
Korištenjem metoda pregledavanja (eng. *inspection methods*), procjenjuje se usklađenost web aplikacije sa važećim standardima i smjernicama koji osiguravaju upotrebljivost. Konačna ocjena web aplikacije ovisi isključivo o osobnom sudu pojedinca ili skupine ljudi koji su evaluaciju proveli.

Metoda procjene prema heuristikama (heurističko vrednovanje)

Heuristika je naziv za opće načelo ili smjernicu prema kojoj se tijekom procesa razvoja web aplikacije oblikuju novi ili ocjenjuju postojeći funkcionalni dijelovi njena sučelja. Metoda procjene prema heuristikama ili heurističko vrednovanje (eng. *heuristic evaluation*) je najčešće korištena neformalna metoda koju su prije gotovo dva desetljeća razvili Jacob Nielsen i Rolf Molich [25]. Sastoji se od toga da relativno mali broj stručnjaka analizira usklađenost svakog interaktivnog elementa sučelja web aplikacije sa popisom poznatih i usvojenih principa upotrebljivosti. U većini slučajeva, riječ je o heuristikama koje je predložio Nieleesen [29, 34]. Tijekom provođenja evaluacije, svaki stručnjak individualno prolazi kroz sve elemente sučelja najmanje dvaput. Ovo ograničenje je vrlo važno kako bi se osigurala nezavisne, nepristrane i visokokvalitetne procjene.

U prvom se prolasku evaluator upoznaje s aplikacijom i njenim mogućnostima, dok se u drugom fokusira na točno određene elemente sučelja i funkcionalnosti, te ih

procjenjuje prema popisu heuristika. Konačni rezultat procjene prema ovoj metodi je lista problema upotrebljivosti koji nisu ništa drugo nego popis nepridržavanja ili povreda heuristika. Heuristike moraju biti pažljivo odabrane tako da odražavaju specifičnosti sustava koji se vrednuje. Jednom kada je evaluacija dovršena, uspoređuju se pronalasci različitih evaluatora. Generirani popis problema olakšava njihovo ispravljanje obzirom da su smjernice za poboljšanja sastavni dio popisa heuristika. Ova metoda je posebice dragocjena kada je potrebno provesti procjenu sa ograničenom količinom raspoloživih resursa. Naime, ukoliko su u evaluaciju uključeni iskusni stručnjaci, oni mogu, bez potrebe za uključivanjem korisnika, generirati visoko kvalitetne rezultate u ograničenoj količini vremena [18]. Zapravo, evaluaciju po heuristikama bi mogao provesti i samo jedan evaluator. Međutim, istraživanje je pokazalo da jedan evaluator može pronaći samo 35% od ukupnog broja postojećih problema upotrebljivosti [25]. Prema tome, ako želimo dobiti kvalitetnije podatke, u evaluaciju je potrebno uključiti što veću i heterogeniju skupinu stručnjaka. Na slici 1. prikazan je omjer postotka utvrđenih problema upotrebljivosti sa brojem uključenih evaluatora. Iz položaja krivulje je evidentno da se relativno kvalitetni podaci mogu dobiti ukoliko u evaluaciji sudjeluje pet stručnjaka no njihov broj nikako ne bi smio biti manji od tri. Međutim, neki autori ovu činjenicu osporavaju obzirom da je njihovo istraživanje pokazalo da se i sa petoro evaluatora može otkriti tek oko 35% postojećih problema upotrebljivosti [46].



Slika 1. Omjer postotka identificiranih problema upotrebljivosti i broja uključenih evaluatora [28]

Od prednosti ove metode potrebno je svakako izdvojiti mogućnost upotrebe u svim fazama razvoja aplikacije, intuitivna i jeftina primjena te brza i djelotvorna identifikacija glavnih i sporednih problema upotrebljivosti. S druge strane, ova metoda ima i nedostataka od kojih je najveći njena ovisnost o iskustvu i vještinama evaluatora [7]. Naime, prilikom odabira evaluatora treba voditi računa o tome da su podaci koje će generirati evaluatori početnici i do tri puta lošiji od onih koje će nam dati stručnjaci domene koji se bave istraživanjem web upotrebljivosti [27]. Osim toga, otežano je identificirati potrebe krajnjih korisnika obzirom da oni nisu uključeni u proces evaluacije. Nadalje, kako ne postoji mehanizam kontrole, evaluatori mogu zastraniti i procjenu usmjeriti samo na neke dijelove sučelja, a ostale koji su jednako važni, u potpunosti zanemariti. Konačno, provedena su mnoga istraživanja koja su preispitivala valjanost Nielsenovih heuristika [44]. To se posebice odnosi na web 2.0 aplikacije za koje se pokazalo da mogu ostvariti ogroman prihod te privući milijune korisnika, a da pri tome zadovoljavaju minimalan broj predloženih i usvojenih heuristika [12, 45]. Jednako tako, Nielsenove se heuristike ne mogu bez izmjena i nadogradnje

uspješno koristiti u vrednovanju upotrebljivosti tehnologija za potporu učenju i sličnih specifičnih sustava [14].

Kognitivna šetnja

Metoda kognitivne šetnje (eng. *cognitive walkthrough*) simulira ponašanje korisnika tijekom rješavanja zadatka pomoću sučelja web aplikacije [42]. Svaki evaluator dobiva predefiniran scenarij prema kojem, korak po korak, izvršava zadatak te komentira svaki problem upotrebljivosti na koji naiđe. Kroz identifikaciju odnosa između korisničkih ciljeva, korisničkih akcija i reakcija aplikacije, moguće je utvrditi što će i kako napraviti korisnik u određenoj problemskoj situaciji. Ova je metoda posebice prikladna za evaluaciju prezentacijskih aspekata sučelja te utvrđivanje problema upotrebljivosti koji su usko povezani s učenjem primjene web aplikacije (eng. *learnability*). Tipičan proces provođenja evaluacije započinje odabirom zadataka koje će evaluator izvršavati pomoću sučelja. Nakon svake aktivnosti koju provede, evaluator pokušava interpretirati reakciju aplikacije te procijeniti sljedeće korake koji vode krajnjem cilju. Pri tome mu mogu pomoći pitanja vezana uz standardan skup atributa upotrebljivosti. Svaki negativan odgovor na postavljena pitanja povećava količinu identificiranih problema. Na kraju evaluacije, kompletira se lista sa popisom problema te preporukama za poboljšanja koja se onda šalju razvojnom timu. Postoji nekoliko verzija ove metode, od kojih je potrebno izdvojiti multidisciplinarnu kognitivnu šetnju (eng. *pluralistic walkthrough*) gdje krajnji korisnik, programer i stručnjak koji se bavi proučavanjem web upotrebljivosti zajedno prolaze kroz cijelu web aplikaciju komentirajući svaki interaktivni element sučelja. Prednosti ove metode uključuju djelotvornu identifikaciju problema koji proizlaze iz interakcije sa aplikacijom te mogućnost utvrđivanja korisničkih ciljeva i ponašanja prilikom korištenja aplikacije. S druge strane, sporost i pretjerana detaljnost u provođenju, opasnost od odabira neprikladnih zadataka, te izoliranost krajnjih korisnika spadaju u glavne nedostatke ove metode.

3.2 Metode testiranja

Metode testiranja (eng. *testing methods*) daju nam informacije o tome kako korisnici upotrebljavaju web aplikacije te na koje probleme pritom nailaze. Tijekom testiranja korisnici izvršavaju određeni zadatak, a evaluatori prate njihov rad i bilježe rezultate. Analizom rezultata moguće je utvrditi koliko je pojedina web aplikacija prikladna za izvršavanje određene skupine zadataka, koliko često i zašto dolazi do određenih grešaka u njenom radu i slično.

Razmišljanje naglas

Razmišljanje naglas (eng. *thinking aloud*) ubraja se među najvažnije metode testiranja upotrebljivosti. Sastoji se od toga da korisnici cijelo vrijeme korištenja web aplikacije verbaliziraju svoje misli [30]. Na taj način korisnici pružaju mogućnost evaluatorima da shvate na koji način oni percipiraju pojedine elemente sučelja, te sa kojim se problemima susreću prilikom korištenja aplikacije. Promatrajući način na koji korisnici opisuju svaki element sučelja, evaluator može jednostavno identificirati one sa kojima korisnici imaju najviše problema. Postoji nekoliko varijanti ove metode. Prva je retrospektivni izvještaj (eng. *retrospective thinking aloud*) [11] prema kojem korisnik tijekom korištenja aplikacije ne razmišlja naglas već nakon dovršetka zadatka deskriptivno ili verbalno prenosi svoja zapažanja evaluatoru. Prednost ovakvog pristupa je u tome što se korisnik može bolje fokusirati na izvođenje zadatka. S druge strane, klasičnim razmišljanjem naglas se dobivaju

puno kvalitetniji podaci jer se verbaliziraju na licu mjesta za razliku od retrospektivnog izvještaja gdje se korisnik prisjeća nečega što je bilo prije nekog vremena te pritom postoji opasnost da zaboravi ili slučajno izostavi neke aspekte koji bi mogli biti jako važni za studiju upotrebljivosti. Druga varijanta ove metode je konstruktivna interakcija (eng. *codiscovery learning*) u kojoj dva korisnika istovremeno upotrebljavaju aplikaciju. Na taj je način testna situacija mnogo prirodnija te je izglednije da će kroz interakciju oba korisnika dati veću količinu komentara nego bi to bio slučaj kada bi svaki zasebno razmišljao naglas. Prednosti ove metode uključuju mogućnost prikupljanja velike količine jasnih i detaljnih podataka od malog broja korisnika, dobivanje predodžbe o tome kako korisnici upotrebljavaju sustav što olakšava predviđanje nastanka problema te mogućnost istovremenog prikupljanja podataka o performansama aplikacije i preferencama korisnika. Bez obzira na mnoštvo prednosti, ova metoda često oduzima jako puno vremena, posebno u fazi pripreme kada je ispitanicima potrebno dati detaljne upute o načinu njene provedbe. Osim toga, mnogim je korisnicima ovaj oblik evaluacije naporan, neprirodan, odvraća im pažnju te im je zbog toga nemoguće istovremeno obavljati zadatak i verbalizirati misli.

Metoda praćenja oka

Metoda praćenja oka (eng. *eye tracking*) se ubraja u metode testiranja upotrebljivosti jer se pokreti očiju smatraju indikatorom aktivnosti kognitivnog procesiranja [9], a time i uloženog mentalnog i fizičkog napora prilikom korištenja web aplikacije. Osim toga, pomoću ove je metode moguće dobiti uzorak načina pretraživanja i pregledavanja elemenata sučelja. Osnovni tipovi pokreta oka su [8]:

- fiksacija oka – usmjeravanje i zadržavanje pogleda na elementu sučelja (npr. gumb na izborniku). Fiksacija traje otprilike 150-600 ms što je dovoljno za realizaciju kognitivnog procesa pomoću kojeg će se shvatiti značenje pojedinog elementa sučelja ili izdvojiti tražena informacija iz teksta;
- trzaj oka (eng. *saccade*) – vrlo brz pokret oka između dvije fiksacije koji traje 10-100 ms. Za vrijeme trzaja, informacije se ne dohvaćaju.
- Pomoću modernih uređaja za praćenje oka, kao što je npr. Tobii, moguće je prikupiti sljedeće podatke: [21]:
- prosječno trajanje fiksacija oka izraženo u milisekundama predstavlja kognitivno opterećenje koje može biti interpretirano kao poteškoće u razumijevanju sadržaja na web sučelju ili može biti pokazatelj da je neki element sučelja jako interesantan;
- prosječna amplituda trzaja izražena u stupnjevima. Do određene granice, veća amplituda predstavlja manju količinu uloženog mentalnog napora;
- prosječan promjer zjenica izražen u milimetrima koji može biti mjerilo kognitivnog napora tijekom korištenja elemenata web sučelja;
- ukupan broj trzaja oka – veća količina trzaja oka predstavlja povećani fizički napor zbog opsežnog pretraživanja i nemogućnosti efikasnog rješavanja zadatka. Veliki broj trzaja oka označava poteškoće kod usmjeravanja na pojedini element korisničkog sučelja;
- ukupan broj fiksacija. Povećani broj fiksacija ukazuje na poteškoće u dohvaćanju informacije ili frustracije u interakciji sa korisničkim sučeljem;
- snaga ekstraokularnog mišića je količina energije izražena u gramima po stupnjevima u sekundi za izvođenje operacija ekstraokularnog mišića za vrijeme trzaja ili fiksacije oka.

Temeljem prosječnog trajanja fiksacija, prosječne amplitude trzaja i prosječnog promjera zjenice moguće je izračunati indeks uloženog mentalnog napora dok je na osnovi ukupnih trzaja i fiksacija oka, te snage ekstraokularnih mišića moguće je izračunati količinu uloženog fizičkog napora kod izvođenja zadatka pomoću web aplikacije.

Osim za izračun uloženog napora, većina uređaja za praćenje oka ima implementirane tri tehnike analize prikupljenih podataka [40]: mogućnost ponovnog pregledavanja, nacrt pregledavanja i toplinske karte.

Mogućnost ponovnog pregledavanja (eng. *gaze replay*) sastavni je dio alata za praćenje oka koji nam omogućava da snimamo rad očiju korisnika za vrijeme testiranja. Nakon što testiranje završi, evaluator može pokrenuti snimku, prilagoditi brzinu te analizirati redoslijed pregledavanja pojedinih elemenata sučelja tijekom izvršavanja zadatka. Iako se ubraja u najtočnije tehnike prikupljanja informacija, njihova analiza oduzima jako puno vremena.

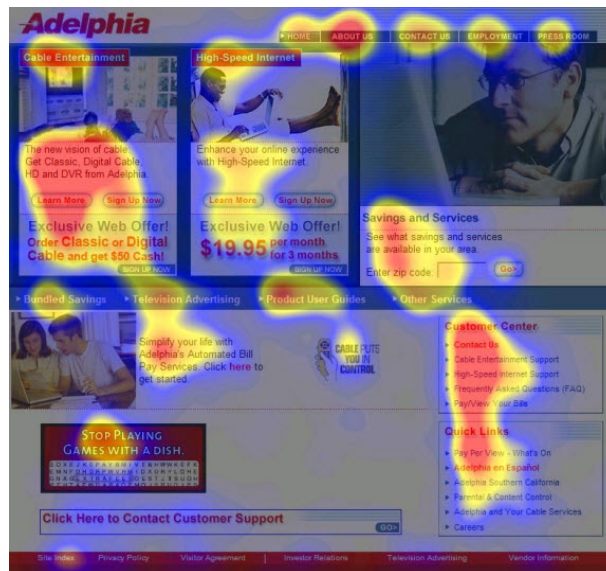
Nacrt pregledavanja (eng. *gazeplot*) je grafički prikaz redoslijeda pregledavanja elemenata web sučelja. Svaki nacrt kojeg generira uređaj za praćenje oka odnosi se na jednog korisnika te jednu web stranicu. Nacrt se sastoji od: 1) plavih točaka koje pokazuju gdje su se nalazile fiksacije, 2) brojeva uz točke koji predstavljaju redoslijed kojim su ispitanici pregledavali web sučelje. Ukupan broj točaka određuje količinu važnih dijelova sučelja dok veličina točke predstavlja duljinu trajanja fiksacije oka. Naravno, što je točka veća, fiksacija je dulje trajala. Premda je riječ o tehnici kojom je moguće prikupiti veliku količinu podataka, njihova analiza može biti otežana i dugotrajna. Naime, na jednom dijelu sučelja se može nalaziti veliki broj točaka te ih je stoga teško razlikovati, a gotovo nemoguće identificirati koji redni broj pripada kojoj točki. Primjer generiranog nacrta pregledavanja tijekom čitanja teksta i izvršavanja zadatka nalazi se na slici 2.



Slika 2. Primjer generiranog nacrta pregledavanja [40]

Toplinska karta (eng. *heatmap*) predstavlja grafički prikaz frekvencije pregledavanja pojedinih elemenata web sučelja. Dijelovi web sučelja označavaju se određenom bojom, ovisno o broju identificiranih fiksacija oka. Tako se crvenom bojom označavaju dijelovi koji su privukli najviše pažnje korisnika nakon koje slijede narančasta, žuta te plava boja kojom su označena dijelovi sučelja sa najmanjim brojem fiksacija. Siva boja označava da se na tom dijelu web aplikacije nije

realizirao značajan broj fiksacija oka. Primjer toplinske karte koju je izgenerirao uređaj za praćenje oka temeljem pregledavanja web sučelja od strane 60 korisnika nalazi se na slici 3. Toplinske karte vrlo jednostavno generirati te se iz njih mogu izlučiti korisni zaključci vezani uz pozicioniranje pojedinih funkcionalnosti, reklama i sl. na web sučelju. Međutim, toplinske karte ne predstavljaju kvalitativne već kvantitativne podatke bazirane na statistici te je za dobivanje relevantnih podataka potrebno uključiti više korisnika nego što je to slučaj kod ostalih metoda testiranja upotrebljivosti (najmanje 39 sudionika po toplinskoj karti [40]).



Slika 3. Toplinska karta pregledavanja web stranice [40]

Automatsko zapisivanje podataka

Automatskim zapisivanjem (eng. *logging actual use*) pohranjuju se svi podaci vezani uz interakciju korisnika i sučelja web aplikacije, uključujući pokretanje određenih akcija, vrijeme izvođenja zadatka, broj pritisaka tipaka na tipkovnici, broj klikova mišem, korištenje pomoći, broj grešaka u radu itd. Pomoću ove metode, evaluatori dobivaju informacije o tome na koji način različite skupine korisnika pristupaju rješavanju istog problemskog zadatka. Osim toga, s obzirom da je riječ o automatiziranom postupku, moguće je prikupiti podatke od velike količine korisnika koji zadatak rješavaju u različitim web sučeljima. Međutim, glavni nedostatak ove metode je što ona pokazuje što korisnici rade, ali ne i zašto to oni rade. Stoga se gotovo uvijek preporuča korištenje ove metode u kombinaciji sa drugim metodama, primjerice razmišljanje naglas, pomoću kojih korisnik može objasniti značenje poduzetih akcija.

3.3 Metode ispitivanja

Metode ispitivanja (eng. *inquiry methods*) koriste se za evaluaciju cjelokupnog zadovoljstva web aplikacijom na kraju njenog razvojnog ciklusa. Osim toga, mogu se koristiti i tijekom razvoja aplikacije kako bi se lakše utvrdile potrebe korisnika.

Upitnici

Upitnici (eng. *questionnaires*) se sastoje od predefiniranih pitanja i skupa otvorenih ili zatvorenih odgovora. Riječ je o indirektnoj metodi ispitivanja upotrebljivosti koja se ne bavi proučavanjem funkcionalnosti sučelja već prikupljanjem mišljenja korisnika. Subjektivne aspekte upotrebljivosti web aplikacije najlakše je izmjeriti Likertovom ljestvicom stavova. Zasnovana je na pretpostavci da svaka izjava

na ljestvici ima jednaku važnost u odnosu na to koliko odražava stav prema pitanju u upitniku. Iako je upitnik moguće primijeniti u svim razvojnim fazama, najčešće se upotrebljava tek nakon što je korisnik ostvario prvi kontakt sa web aplikacijom te isprobao funkcionalnosti njena sučelja. Na taj se način mogu prikupiti informacije o tome kako korisnici upotrebljavaju aplikaciju te njihovom zadovoljstvu i preferencama. U literaturi se može pronaći mnoštvo upitnika kao što su SUMI [19], SUS [5], ASQ [22], PSQ [22], CSUQ [22], PSSUQ [22], QUIS [37] i PUTQ [23] za procjenu upotrebljivosti programskih proizvoda bilo koje vrste te WAMMI [20], UWIS [39], WUS [26], EUCS [1] i PWQ [2] za procjenu zadovoljstva korisnika i upotrebljivosti web stranica. Iako je riječ o upitnicima koji se često koriste, većina njih omogućava vrednovanje relativno malog broja atributa upotrebljivosti. Osim toga, ne postoji dokaz da su ti upitnici prikladni za vrednovanje upotrebljivosti web 2.0 aplikacija. Stoga je provedeno istraživanje koje predlaže skup atributa koje treba uzeti u obzir prilikom izrade upitnika za vrednovanja kvalitete, a time i upotrebljivosti web 2.0 aplikacija [38]. Od prednosti ove metode potrebno je izdvojiti mogućnost identifikacije subjektivnih preferenci korisnika vezano uz upotrebu web aplikacije te što se temeljem prikupljenih podataka može napraviti kvalitetna statistička obrada te se iz nje polučiti važni zaključci. S druge strane, rezultati koji se dobivaju ovom metodom često se ne podudaraju s podacima koji se prikupljaju objektivnim metodama mjerenja upotrebljivosti. Nadalje, kako bi dobili signifikantne podatke, u istraživanju je potrebno sudjelovati najmanje trideset ispitanika. Osim toga, potrebno je napomenuti da korisničko mišljenje u trenutku ispunjavanja upitnika može varirati te često ovisi o jako puno faktora poput npr. raspoloženja ispitanika. Konačno, u odnosu na druge opisane metode, pomoću upitnika je moguće identificirati najmanju količinu problema upotrebljivosti.

Terensko ispitivanje

Terensko ispitivanje (eng. *field observation*) uključuje posjećivanje jednog ili većeg broja korisnika u njihovoj radnoj okolini te praćenje izvršavanja radnih zadataka pomoću određene web aplikacije. Prikupljanje podataka i vođenje bilješki je potrebno provoditi što je moguće više nenametljivo kako se korisnike ne bi ometalo u njihovom radu. Naime, buka i ostali oblici uznemiravanja korisnika mogu dovesti do pogrešnih rezultata i zaključaka. Prema tome, idealna bi situacija bila kada bi evaluator prikupljao podatke neprimjetno te time osigurao normalne radne uvjete. Kako bi ispitivanje bilo što manje uznemirujuće, u nekim se situacijama ova metoda kombinira sa metodom odgođenog emitiranja (eng. *retrospective testing*). Na ovaj se način rad korisnika snima kamerom, a podaci prikupljaju analizom video zapisa. Međutim, kako je ova metoda usmjerena na prikupljanje podataka o problemima upotrebljivosti koji su očiti na prvi pogled, snimanje kamerom te kasnija detaljna analiza video zapisa bi oduzimala previše vremena.

4. KRITERIJI ZA ODABIR METODA ZA VREDNOVANJE UPOTREBLJIVOSTI

S obzirom da postoji velik broj metoda vrednovanja upotrebljivosti, njihov odabir nije jednostavan i ovisi o različitim kriterijima. Nekoliko znanstvenika [6, 13, 43] razmatralo je kriterije koji mogu utjecati na izbor jedne ili više metoda za vrednovanje upotrebljivosti. U nastavku se nalazi popis tih kriterija:

- primjena u fazi životnog ciklusa sustava [6, 13] – da li se metoda jednostavno može primijeniti u svim fazama razvoja sustava;
- oprema potrebna za provedbu vrednovanja upotrebljivosti [6, 13, 43] – uključuje opremu za provođenje inspekcije ili testiranja (npr. računalo, softver za snimanje ekrana, video kamera, upute/upitnici itd.);
- vrijeme i novac potreban za provedbu metode [6, 13, 43] – koliko je metoda vremenski zahtjevnija, te koliki je budžet nužan za njenu provedbu;
- stil vrednovanja upotrebljivosti [6] – vrednovanje se može obaviti u laboratorijskim uvjetima, na terenu ili oboje;
- razina objektivnosti/subjektivnosti metode [6] – npr. procjene koje obavlja čovjek su subjektivne, no subjektivnost se može smanjiti većim brojem evaluatora ili korisnika, dok eksperiment ili praćenje pokreta očiju daju objektivne podatke o problemima upotrebljivosti;
- vrste informacija i razina informacija dobivenih metodom [6] – da li metodom dobivamo kvalitativne informacije o upotrebljivosti ili kvantitativne podatke na koje možemo primijeniti statističke metode. Informacije više razine daju objašnjenje uočenog problema (npr. kod intervjua ili heurističkog vrednovanja) dok informacije niže razine daju konkretan podatak (npr. vrijeme, identificirani element itd.);
- da li je metoda ometajuća [6, 13] – u kojoj mjeri metoda ometa korisnike u njihovoj procjeni sučelja, npr. razmišljanje naglas stvara dodatan neprirodan mentalni napor;
- broj testiranih korisnika [13, 43] – minimalan potreban broj testiranih korisnika za dobivanje relevantnih podataka o problemima upotrebljivosti. Broj ovisi o samoj metodi, te kod primjene razmišljanja naglas iznosi najmanje 3 korisnika, preporučuje se 5 korisnika kod mjerenja korisničkih performansi [33], dok kod korištenja upitnika treba ispitati barem 30 korisnika.
- broj stručnjaka upotrebljivosti [13] – minimalan potreban broj evaluatora koji provode pregled sučelja. Kod terenskog ispitivanja dovoljni su jedan do dva evaluatora, dok kod kognitivne šetnje i heurističkog vrednovanja preporučuje se 3-5 evaluatora. Također je važna i ekspertiza evaluatora [6, 13, 43].

Općenito, metode pregledavanja su jeftine i brze metode za identificiranje problema upotrebljivosti. Ne zahtijevaju specijalnu opremu za vrednovanje i pomoću njih mali broj evaluatora s visokom ekspertizom može identificirati velik broj kritičnih i manje kritičnih problema upotrebljivosti. S druge strane, metode testiranja korisnika su skuplje i vremenski zahtjevnije, ali ako se koriste, prednost bi trebalo dati testiranju interakcije korisnika sa sučeljem, a ne ispitivanju percepcije korisnika pomoću upitnika [13]. Upitnici ne daju toliko pouzdane podatke jer se temelje na prisjećanju korisnika o tome kako su koristili sučelje i kojim njegovim komponentama jesu ili nisu zadovoljni.

Jedan od kriterija odabira i primjene metoda vrednovanja upotrebljivosti jest mogućnost njihova kombiniranja, radi dobivanja pouzdanijih rezultata vrednovanja upotrebljivosti [10, 13, 48]. Na taj se način postiže da do izražaja dođu prednosti pojedinih metoda, a nedostaci se smanje na najmanju moguću mjeru. Kombinacije primjene metoda moguće su unutar kategorija (npr. procjena prema heuristikama i kognitivna šetnja) ili među kategorijama (npr. upitnik i razmišljanje naglas). U slučaju da je potrebno odabrati samo jednu metodu, heurističko vrednovanje pokazalo se kao optimalna metoda vrednovanja upotrebljivosti [47].

5. PREPORUKE ZA VREDNOVANJE WEB UPOTREBLJIVOSTI

Prilikom vrednovanja web upotrebljivosti važno je utvrditi da li korisnici na web mjestu postižu svoje ciljeve i imaju li pozitivno iskustvo tijekom korištenja web mjesta. Ove parametre nije moguće utvrditi samo jednom metodom vrednovanja.

Posljednjih godina web upotrebljivost se vrednuje kroz dva pristupa: korištenjem heurističkog vrednovanja i vrednovanjem pomoću scenarija [48], te oba pristupa uključuju i stručnjake upotrebljivosti i korisnike.

Prvi pristup temelji se na korištenju heuristika pomoću kojih je moguće identificirati širok spektar problema upotrebljivosti, ali zbog nedostatka konkretnih smjernica (što točno treba pregledati na nekom web mjestu), moguće je previdjeti neke složenije scenarije interakcije. Na primjer, jedna od Nielsenovih heuristika koja se odnosi na konzistentnost i standarde definira da se korisnici ne bi trebali pitati da li različite riječi, situacije ili akcije označavaju jednu te istu stvar (*Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.* [34]). Heuristiku kao takvu je teško primijeniti bez detaljnijih smjernica koje opisuju kako konkretnu smjernicu provjeriti na web mjestu. Specifične smjernice koje bi pojašnjavale ovu osnovnu heuristiku u kontekstu vrednovanja web mjesta mogle bi biti slijedeće: provjeriti je li web mjesto konzistentno u navigaciji, korištenoj terminologiji i dizajnu (font, boja i slično); osnovna navigacija ne mijenja svoj položaj na web stranici; za svaku poveznicu jasno je kuda vodi; ne koriste se nestandardne tehnologije koje zahtijevaju dodatnu instalaciju itd. Evaluators početnik bez sličnog detaljnog opis smjernica teško može kvalitetno obaviti procjenu sučelja.

Stoga vrednovanje pomoću scenarija nadopunjuje vrednovanje pomoću heuristika jer omogućuje konkretizaciju zadataka koje treba provesti na web mjestu. Scenariji fokusiraju vrednovanje aplikacije kroz specifikiranje korisničkih profila, ciljeva i zadataka i mogu se primijeniti kod pregleda aplikacije ili tijekom testiranja korisnika. Primjer scenarija može biti slijedeći: američki turist želi posjetiti određeni muzej u nekom gradu, te ga zanima da li na točno određeni dan u muzeju postoji neka specijalna izložba ili vođeni obilazak [48]. Iz navedenog scenarija možemo definirati korisnički profil (turist), njegov cilj (posjetiti muzej na definiran dan) i zadatke koje bi korisnik obavio na web mjestu da bi postigao cilj (pronaći izložbu ili događaj na definiran dan, pronaći lokaciju muzeja, pronaći pravce dolaska do muzeja, radno vrijeme muzeja i sl.).

Dva primjera istraživanja koja su obuhvatila oba pristupa vrednovanja web upotrebljivosti dana su u nastavku:

1. istraživanje [10] – vrednovala se upotrebljivost novinskih portala u Hrvatskoj primjenom slijedećih metoda vrednovanja: mekše verzije heurističkog vrednovanja (tzv. *guideline inspection*) s 10 priučenih evaluators i korisničkog testiranja s grupom od 15 naprednih korisnika sa znanjem iz područja informacijskih tehnologija i grupom od 15 običnih korisnika različite dobi. Testiranje korisnika uključivalo je interakciju sa sučeljem prema definiranim scenarijima i zadacima, ispitivanje zadovoljstva korisnika primjenom SUS upitnika, te polustrukturirani intervju za pojašnjenje identificiranih problema. Analiza rezultata istraživanja pokazala je raskorak u identificiranim problemima između stručnjaka i korisnika, te da je

bolje koristiti manji broj evaluators s visokom ekspertizom (5) i manji broj korisnika (do 15).

2. istraživanje [48] – vrednovala su se muzejska web mjesta u Italiji primjenom specifične metode za web upotrebljivost, tzv. MiLE metode [*Milano-Lugano Evaluation Method*]. Ova metoda temelji se na primjeni heurističkog vrednovanja sa specifičnim heuristikama, inspekciji pomoću scenarija i testiranju korisnika (koje nije obavezno). MiLE+ metoda sadrži biblioteke alata za vrednovanje upotrebljivosti koje se sastoje od: a) biblioteke tehničkih heuristika – evaluators vode kroz vrednovanje tehničkih karakteristika aplikacije: navigacije, sadržaja, tehnologiju i dizajn sučelja, b) biblioteke indikatora korisničkog iskustva – evaluators omogućuju stavljanje u ulogu korisnika kako bi anticipirao probleme s kojima se korisnik aplikacije može susresti, c) biblioteke scenarija – kreiraju se za pojedinu domenu vrednovanja kroz specifikiranje korisničkih profila, ciljeva i zadataka. Nakon što se provede vrednovanje prema tehničkim heuristikama i indikatorima korisničkog iskustva, te se svaka heuristika i zadatak kvalitativno i kvantitativno ocijene, prelazi se na testiranje korisnika ako je inspekcija sučelja pomoću indikatora korisničkog iskustva otkrila kritične probleme sa sučeljem web aplikacije.

S obzirom na brojne navedene metode i pristupe vrednovanju upotrebljivosti, postavlja se pitanje kako iskustva znanstvenika i iskusnih praktičara najjednostavnije primijeniti u praksi? Prije svega, potrebno je isplanirati vrednovanje upotrebljivosti, te ga provesti unutar životnog ciklusa web aplikacije. Uvjet za vrednovanje je da su faze identifikacije i analize potreba korisnika (i naručitelja) završene.

Badre u [3] preporučuje slijedeći proces vrednovanja web upotrebljivosti:

- postavljanje ciljeva vrednovanja upotrebljivosti
Ciljevi upotrebljivosti web mjesta trebaju biti što konkretnije definirani i mjerljivi, a ne apstraktni. Cilj poput "web mjesto treba biti jednostavno i intuitivno za korištenje" je loše definiran cilj, dok je cilj "korisnik treba napraviti (određeni zadatak) bez postavljanja pitanja" konkretan i mjerljiv. Ciljeve možemo izraziti kao apsolutne, npr. pronaći stranicu za kupovinu proizvoda u 2 ili manje klikova mišem, ili relativne ciljeve, npr. obaviti kupovinu proizvoda u najviše 85% vremena u odnosu na konkurentsko web mjesto.
- rano prototipiranje na papiru
Rano prototipiranje omogućuje da ispitamo u kojoj mjeri inicijalni web dizajn reflektira svrhu web mjesta, te da li je dizajn jasan i estetski prihvatljiv. Vrednuje se jedan ili više dizajna jedne web stranice koja je skicirana na papiru. Za testiranje su dovoljna 2-4 korisnika koji odgovaraju na nekoliko pitanja o predloženim skicama dizajna budućih web stranica, kako bi ih međusobno usporedili. Tipično se postavljaju slijedeća pitanja: prema predloženom dizajnu ove web stranice odredite svrhu web mjesta; pronađite elemente na web stranici pomoću kojih ćete saznati adresu poduzeća; koji od prijedloga dizajna vam se više sviđa (bojama, rasporedom elemenata, odabirom fonta itd.). Korisnicima se mjeri vrijeme koje im je potrebno da odgovore na pitanje.
- testiranje strukture web mjesta na papiru
Testira se dizajn i struktura web mjesta koje je prikazano u obliku *storyboarda*, tj. mape web mjesta na papiru. Također, na zasebnim papirima prikazan

je dizajn pojedinih web stranica. U ovoj fazi vrednovanje upotrebljivosti provode 2-3 stručnjaka upotrebljivosti i 2-3 korisnika, tj. ukupno oko 5 osoba. Stručnjaci trebaju predložiti zadatke kojima se testiraju ciljevi i funkcionalnost web mjesta, a korisnici izvedu zadatke i komentiraju pitanja poput "možete li odrediti gdje se točno nalazite u strukturi web mjesta", "da li su tekstualne poveznice jasno definirane", itd. Evaluatori bilježe elemente (riječi, poveznice, ikone) koje korisnici ne razumiju ili ih frustriraju.

➤ testiranje interaktivnog prototipa

U ovoj fazi provodi se heurističko vrednovanje i testiranje korisnika na prototipu web stranica. Najbolji rezultati kod heurističkog vrednovanja postižu se ako vrednovanje provodi više evaluatora koji imaju ekspertizu iz web upotrebljivosti i domene vrednovanja jer odabrane heuristike treba promatrati u kontekstu domene (npr. e-poslovanja, web mjesta za osnovnoškolce i sl.). Korisnici se testiraju u laboratorijima gdje izvedu konkretne zadatke na web mjestu, a evaluatori bilježe uočene probleme, frustracije i druga ponašanja korisnika, te dodatno može se bilježiti vrijeme izvršenja zadataka, broj grešaka, broj klikova i druge mjerljive varijable da bi se utvrdilo u kojoj mjeri korisnici ostvaruju zadatke.

Iako se preporučuje izvedba sve četiri faze vrednovanja upotrebljivosti, radi smanjenja troškova i trajanja vrednovanja, dovoljno je izvesti prvu fazu (postavljanje ciljeva vrednovanja) i posljednju fazu (testiranje interaktivnog prototipa) bez testiranja korisnika [3].

Najjeftiniji način vrednovanja upotrebljivosti je tzv. "discount usability" vrednovanje od Nielsena [3]. Kod tog pristupa vrednovanju dovoljno je imati jednog evaluatora koji će vrednovati web stranice s manjim brojem heuristika, dok se korisničko testiranje obavlja s tri korisnika koji verbaliziraju interakciju s web mjestom pomoću metode razmišljanja naglas i pritom obavljaju nekoliko zadataka koji su vezani uz jedan scenarij. Kako bi se korisnici osjećali što ugodnije tijekom testiranja, važno im je prije početka testiranja naglasiti da se testira sustav, a ne oni kao individue.

6. ZAKLJUČAK

Metode vrednovanja upotrebljivosti koje su potekle iz discipline interakcija čovjeka i računala radi identificiranja problema upotrebljivosti tradicionalnih grafičkih korisničkih sučelja, mogu se jednako uspješno primijeniti i na razne web aplikacije. Kod weba je, međutim, dodatno naglašena potreba pozitivnog korisničkog iskustva u interakciji s aplikacijom, tj. zadovoljstvo i ugodna korisnika pri ostvarivanju ciljeva na web mjestu, kao vrlo bitna komponenta za zadržavanje korisnika upravo na našem web mjestu.

Kroz dizajn orijentiran korisniku i vrednovanje različitih aspekata dizajna korištenjem metoda vrednovanja upotrebljivosti, moguće je izraditi kvalitetna web mjesta koja su usklađena s potrebama korisnika.

Metode koje se mogu primijeniti kod vrednovanja web upotrebljivosti svrstavaju se u tri različite kategorije: metode pregledavanja, metode testiranja i metode

ispitivanja. Za svaku od spomenutih kategorija u ovom radu detaljnije su opisane one metode koje se najčešće koriste u praksi. Metode pregledavanja koriste se da bi se utvrdilo da li je web sučelje usklađeno sa poznatim i usvojenim standardima i smjernicama upotrebljivosti, a provode ih stručnjaci upotrebljivosti. Od metoda pregledavanja izdvojene su heurističko vrednovanje kod kojeg se u većini slučajeva provjerava usklađenost elemenata sučelja s Nielsenovim heuristikama, te kognitivna šetnja koja ima za cilj utvrditi na koji će način krajnji korisnici upotrebljavati web aplikaciju. Metode testiranja omogućavaju evaluatorima da tijekom korištenja web aplikacije promatraju interakciju ispitanika sa sučeljem te bilježe rezultate. Najčešće korištene metode testiranja su razmišljanje naglas kod koje ispitanici za vrijeme izvršavanja zadatka pomoću elemenata web sučelja verbaliziraju svoje misli, zatim metoda praćenja oka koja pomoću specijaliziranih alata omogućava identifikaciju najfrekventnijih elemenata web sučelja kao i izračun uloženog fizičkog i mentalnog napora u korištenje web aplikacije, te automatsko zapisivanje podataka koje olakšava prikupljanje podataka od veće količine korisnika. Konačno, metode ispitivanja služe prikupljanju podataka o zadovoljstvu korisnika web aplikacijom nakon što su istu koristili neko određeno vrijeme. U tu kategoriju ubrajamo upitnike pomoću kojih korisnici, koristeći ljestvice stavova, izjašnjavaju svoje mišljenje povezano sa upotrebljivošću web sučelja, te terensko ispitivanje gdje evaluatori posjećuju korisnike u njihovom okruženju te bez narušavanja radnog ritma ispitanika prikupljaju podatke. Za svaku od spomenutih metoda navedene su prednosti i nedostaci.

Kako bi se nedostaci pojedinih metoda upotrebljivosti smanjili na najmanju moguću mjeru, preporuča se da se tijekom vrednovanja upotrebljivosti, metode kombiniraju, kako unutar kategorija, tako i između njih. Pri tome izbor metoda uvjetuju razni kriteriji, od kojih su među najvažnijima resursi potrebni za izvođenje metoda (vrijeme, budžet, broj potrebnih evaluatora i njihova ekspertiza, broj korisnika za testiranje, oprema za testiranje), potrebna razina objektivnosti, te mogućnost primjene u raznim fazama izrade web aplikacije.

S obzirom da se web aplikacije razvijaju u znatno kraćem vremenskom periodu nego klasični informacijski sustavi [41], vrednovanje upotrebljivosti često se preskače, a dodatan razlog za to leži u činjenici da je za primjenu pojedinih metoda potrebna sofisticirana oprema i visoka ekspertiza evaluatora. Međutim, web dizajneri svjesni su dobrobiti koje vrednovanje upotrebljivosti donosi, ponajprije kroz skraćivanje razvoja web aplikacije ako se problemi upotrebljivosti identificiraju u ranim fazama životnog ciklusa i kroz povećanje njene kvalitete, te je osnovno pitanje praktičara kako na najučinkovitiji način vrednovanje web upotrebljivosti integrirati u svakodnevni rad.

Korištenje kombiniranog pristupa koje podrazumijeva heurističko vrednovanje s nekoliko osnovnih heuristika i 1-2 priučena eksperta, te testiranje troje korisnika pomoću definiranih scenarija, čini se kao dobar kompromis između potrebe provođenja kvalitetnog vrednovanja web upotrebljivosti, te potrebnog vremena i troškova provedbe.

Literatura:

- 1 Abdinnour-Helm, S.F., Chaparro, B.S., Farmer, S.M. (2005) Using the End-User Computing Satisfaction (EUCS) Instrument to Measure Satisfaction with a Web Site. *Decision Sciences*, Vol. 36, No. 2, pp. 341-364.
- 2 Aladwani, A.M., Palvia, P.C. (2002) Developing and validating an instrument for measuring user-perceived web quality. *Information & Management*, Vol. 39, No. 6, pp. 467-476.
- 3 Badre, A. N. (2002.) *Shaping Web Usability: Interaction Design in Context*. Pearson Education, Inc., Boston, SAD

- 4 Bevan, N. (1995) Measuring usability as quality of use. *Software Quality Journal* 4, 115–150. Dostupno na <http://www.nigelbevan.com/papers/qusab95.pdf> (pristupano 05.04.2010.)
- 5 Brooke, J. (1996) *SUS: a 'quick and dirty' usability scale*. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability evaluation in industry. London: Taylor & Francis Ltd.
- 6 Dix, A., Finlay, J., Abowd, G., Beale, R. (2004) *Human-Computer Interaction*. (Third Edition). Harlow Assex: Pearson Education Limited.
- 7 Doubleday, A., Ryan, M., Springett, M., Sutcliffe, A. (1997) A Comparison of Usability Techniques for Evaluating Design. *Symposium on Designing Interactive Systems: Processes, Practices, Methods and Techniques (ACM DIS)*, Amsterdam, NL, pp 101-110.
- 8 Duchowski, A.T. (2007) *Eye Tracking Methodology: Theory and Practice*, Second edition. London: Springer-Verlag.
- 9 Goldberg, J.H., Wichansky, A.M. (2003) Eye tracking in usability evaluation: A practitioner's guide. In: *The Mind's Eye: Cognitive and Applied Aspects of Eye Movements*. Hyona J, Radach R, Duebel H, editors. Elsevier, Amsterdam, NL, pp. 493-516.
- 10 Granić, A., Mitrović, I., Marangunić, N. (2009) Web Portal Design: an Employment of a Range of Assessment Methods. In Papadopoulos, G. A.; Wojtkowski, W.; Wojtkowski, W. G.; Wrycza, S.; Zupancic, J. (Eds.): *Information Systems Development: Towards a Service Provision Society*. Springer-Verlag: New York, pp.
- 11 Guan, Z., Lee, S., Cuddihy, E. and Ramey J. (2006) The Validity of the Stimulated Retrospective Think-Aloud Method as Measured by Eye Tracking. *CHI 2006 Proceedings*, pp. 1253-1262.
- 12 Hart, J., Ridley, C., Taher, F., Sas, C., Dix, A. (2008) Exploring the Facebook Experience: A New Approach to Usability. *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*; October 20-22, Lund, Sweden, ACM, pp. 471-474.
- 13 Holzinger, A. (2005) Usability engineering methods for software developers. *Communications of the ACM*, Vol. 48, No. 1 (Jan. 2005), pp. 71-74.
- 14 Hornbæk, K. (2006) Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, Volume 64, Issue 2, pp. 79-102.
- 15 ISO 9241-11 (1998) *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability*
- 16 ISO 13407 (1999) *Human-centred design process for interactive systems*
- 17 Ivory M.Y., Hearst, M.A. (2001). The State of the Art in Automating Usability Evaluation of User Interfaces, *ACM Computing Surveys*, Vol. 33, No. 4, pp. 470-516.
- 18 Kantner, L., Rosenbaum, S. (1997) Usability Studies of WWW Sites: Heuristic Evaluation vs. Laboratory Testing. *Proceedings of International Conference on Computer Documentation (ACM SIGDOC)*, Snowbird, USA, pp 153-160.
- 19 Kirakowski, J., Corbett, M. (1993) *SUMI: the software measurement inventory*. British Journal of Educational Technology, Vol. 24, No. 3, pp. 210-212.
- 20 Kirakowski, J., Claridge, N., Whitehand, R. (1998) Human Centered Measures of Success in Web Site Design. *Proceedings of the 4th Conference on Human Factors & the Web*, June 5; Basking Ridge, New York, USA. AT&T Labs. Dostupno na <http://research.microsoft.com/en-us/um/people/marycz/hfweb98/kirakowski/> (pristupano 19.01.2010)
- 21 Komogortsev O.V., Mueller, C.J., Tamir, D., Feldman, L. (2009) An Effort Based Model of Software Usability. *Proceedings of the International Conference on Software Engineering Theory and Practice (SETP-09)*, July 13-16; Orlando, Florida, USA. pp. 75-83.
- 22 Lewis, J.R. (1993) *IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use*. Technical Report 54.786. IBM Corporation.
- 23 Lin, H.X., Choong, Y., Salvendy, G. (1997) A proposed index of usability: a method for comparing the relative usability of different software systems. *Behavior and Information Technology*, Vol. 16, No. 4–5, pp. 267–278.
- 24 Lynch, P.J., Horton, S. (2009) *Web Style Guide: Basic Design Principles for Creating Web Sites*, 3rd Edition, Yale University Press. Dostupno na <http://webstyleguide.com/wsg3/index.html> (pristupano 10.04.2010.)
- 25 Molich, R., Nielsen, J. (1990) Improving a Human-Computer Dialogue. *Communications of the ACM*, Vol. 33, No. 3, pp. 338-348.
- 26 Muylle, S., Moenaert, R., Despontin, M. (2004) The conceptualization and empirical validation of web site user satisfaction. *Information & Management*, Vol. 41, No. 5, pp. 543-560.
- 27 Nielsen, J. (1992) The Usability Engineering Lifecycle. *IEEE Computer*, Vol. 25, No. 3, pp. 12-22.
- 28 Nielsen, J., Landauer, T.K. (1993). A Mathematical Model of the Finding of Usability Problems. *ACM International Conference on Human Factors in Computing Systems*, Amsterdam, NL, April 1993, pp 296-213.
- 29 Nielsen, J. (1994) Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds) *Usability Inspection Methods*. Wiley & Sons, New York.
- 30 Nielsen, J. (1994). *Usability Engineering*. Morgan Kaufmann, San Francisco.
- 31 Nielsen, J. (1999) *Designing Web usability: The practice of simplicity*, New Riders Publishing, Indianapolis.
- 32 Nielsen, J. (2003) Return on Investment for Usability, *Jakob Nielsen's Alertbox*, January 7, 2003. Dostupno na <http://www.useit.com/alertbox/roi-first-study.html> (pristupano 10.04.2010.)

- 33 Nielsen, J. (2003) Usability 101: Introduction to Usability. *Jakob Nielsen's Alertbox*, August 25, 2003. Dostupno na <http://www.useit.com/alertbox/20030825.html> (pristupano 15.03.2010.)
- 34 Nielsen, J. (2005) *Ten Usability Heuristics*. Dostupno na http://www.useit.com/papers/heuristic/heuristic_list.html (pristupano 10.04.2010.)
- 35 Nielsen, J., Loranger, H. (2006) *Prioritizing Web usability*, New Riders Press, Indianapolis.
- 36 Nielsen, J. (2008) Bridging the Designer–User Gap, *Jakob Nielsen's Alertbox*, March 17, 2008. Dostupno na <http://www.useit.com/alertbox/designer-user-differences.html> (pristupano 10.04.2010.)
- 37 Norman, K.L., Shneiderman, B.A., Harper, B.D., Slaughter, L.A. (1998) *Questionnaire for User Interaction Satisfaction*, Version 7.0: Users' Guide. College Park: University of Maryland.
- 38 Orehovački, T. (2010) Proposal for a Set of Quality Attributes Relevant for Web 2.0 Application Success. *Proceedings of the 32nd International Conference on Information Technology Interfaces*, June 21-24, Cavtat, Croatia.
- 39 Oztekin, A., Nikov, A., Zaim, S. (2009) UWIS: An assessment methodology for usability of web-based information systems, *The Journal of Systems and Software*, Vol. 82, No. 12, pp. 2038-2050.
- 40 Pernice, K., Nielsen, J. (2009) *Eyetracking Methodology: How to Conduct and Evaluate Usability Studies Using Eyetracking*, Nielsen Norman Group. Dostupno na <http://www.useit.com/eyetracking/methodology> (pristupano 12.05.2010.)
- 41 Plantak Vukovac, D., Kliček, B. (2007) Web and Multimedia Systems Development – State of the Practice. In Aurer, B., Bača, M. (eds): *Proceedings of the 18th International Conference on Information and Intelligent Systems*, September, 12-14, Varazdin, Croatia, pp. 396-403.
- 42 Polson, P., Lewis, C., Rieman, J., Wharton, C. (1992) Cognitive Walkthrough: A Method for Theory-based Evaluation of User Interfaces. *International Journal of Man-Machine Studies*, Vol. 36, pp. 741-773.
- 43 Preece, J., Rogers, Y. & Sharp, H. (2002) *Interaction Design: Beyond Human-Computer Interaction*. New York: Addison-Wesley.
- 44 Sears, A.L. (1997) Heuristic walkthroughs: Finding problems without the noise. *International Journal of Human-Computer Interaction*, Vol. 9, No. 3, pp. 213–234.
- 45 Silva, P.A., Dix, A. (2007) Usability - Not as we know it! *Proceedings of the 21st British HCI Group Annual Conference on HCI 2007: People and Computers XXI: HCI...but not as we know it - Volume 2*, September 03 – 07; University of Lancaster, United Kingdom, ACM, pp. 103 – 106.
- 46 Spool, J., Schroeder, W. (2001) Testing Web Sites: Five Users Is Nowhere Near Enough, *Proceedings of ACM CHI Conference on Human Factors in Computing*, Seattle, Washington pp. 285- 286.
- 47 Ssemugabi, S., de Villiers, R. (2009) A comparative study of two usability evaluation methods using a web-based e-learning application. In *Proceedings of the 2007 Annual Research Conference of the South African institute of Computer Scientists and information Technologists on IT Research in Developing Countries*, SAICSIT '07, Vol. 226. ACM, New York, 132-142.
- 48 Triacca, L., Inversini, A., Bolchini, D. (2005) Evaluating Web Usability with MiLE+. In *Proceedings of Seventh International Symposium on Web Site Evolution (WSE 2005)*, Budapest, September 22-29.
- 49 U.S. Department of Health & Human Services. *Usability.gov*. Dostupno na <http://www.usability.gov/> (pristupano 11.04.2010.)
- 50 Usability Professionals' Association. *What is User-Centered Design: About Usability: UPA Resources* Dostupno na http://www.upassoc.org/usability_resources/about_usability/what_is_ucd.html (pristupano 10.04.2010.)
- 51 Webnauts Net. *Why usability is important to you*. Dostupno na <http://www.webnauts.net/usability.html> (pristupano 11.04.2010.)

DODATAK A.

Tablica 1. Pregled metoda vrednovanja upotrebljivosti (prošireno i prilagođeno prema [17])

Naziv metode	Kratak opis metode
1. Metode testiranja	evaluatori tijekom testiranja promatraju interakciju ispitanika sa sučeljem te bilježe rezultate
Razmišljanje naglas (<i>Thinking-Aloud Protocol</i>)	korisnik verbalizira svoje misli tijekom testiranja
Postavljanje pitanja (<i>Question-Asking Protocol</i>)	evaluator postavlja pitanja korisniku vezano uz web aplikaciju i njenu uporabu
Izravni prijenos (<i>Shadowing Method</i>)	stručnjak objašnjava korisničke akcije evaluatoru
Stručno podučavanje (<i>Coaching Method</i>)	korisnik može evaluatoru postavljati pitanja vezano uz funkcioniranje web aplikacije
Učenje (<i>Teaching Method</i>)	iskusniji korisnik podučava novog korisnika kako riješiti zadatak pomoću web sučelja
Konstruktivna interakcija (<i>Codiscovery Learning</i>)	kolaboracija među korisnicima tijekom rješavanja zadatka
Mjerenje performansi (<i>Performance Measurement</i>)	ispitanik bilježi podatke o korištenju tijekom testiranja
Analiza dnevnika (<i>Log File Analysis</i>)	evaluator analizira podatke o korištenju aplikacije
Odgođeno emitiranje (<i>Retrospective Testing</i>)	evaluator pregledava video zapis o korištenju sučelja zajedno sa korisnikom te mu postavlja pitanja
Udaljeno testiranje (<i>Remote Testing</i>)	evaluator i ispitanik su fizički razdvojeni ili se testiranje provodi sa vremenskim razmakom
Metoda praćenja oka (<i>Eye Tracking</i>)	pomoću specijaliziranih alata prikupljaju se podaci o tome kako korisnik pregledava elemente web sučelja
Automatsko zapisivanje postupaka (<i>Logging Actual Use</i>)	automatsko prikupljanje detaljnih podataka o upotrebi web aplikacije
2. Metode pregledavanja	stručnjaci provjeravaju usklađenost web sučelja s poznatim i usvojenim standardima i smjernicama
Usklađenost sa smjernicama (<i>Guideline Review</i>)	stručnjak provjerava usklađenost web sučelja sa smjernicama upotrebljivosti
Kognitivna šetnja (<i>Cognitive Walkthrough</i>)	stručnjak simulira način na koji korisnik rješava zadatak pomoću web sučelja
Multidisciplinarna kognitivna šetnja (<i>Pluralistic Walkthrough</i>)	korisnik, programer i stručnjak zajedno provode kognitivnu šetnju
Heurističko vrednovanje (<i>Heuristic Evaluation</i>)	stručnjak identificira moguću neusklađenost sa heuristikama
Perspektivan pregled (<i>Perspective-Based Inspection</i>)	stručnjak provodi usko usmjereno heurističko vrednovanje
Pregled funkcionalnih dijelova (<i>Feature Inspection</i>)	stručnjak vrednuje funkcionalne mogućnosti web sučelja
Formalni pregled (<i>Formal Usability Inspection</i>)	stručnjak provodi formalno heurističko vrednovanje
Pregled konzistentnosti (<i>Consistency Inspection</i>)	stručnjak provjerava konzistentnost kroz sve elemente web sučelja
Usklađenost sa standardima (<i>Standards Inspection</i>)	stručnjak provjerava poštivanje standarda upotrebljivosti
3. Metode ispitivanja	evaluatori prikupljaju podatke o tome koliko su korisnici zadovoljni web aplikacijom
Kontekstualno ispitivanje (<i>Contextual Inquiry</i>)	evaluator ispituje korisnike u njihovom okruženju
Terensko ispitivanje (<i>Field Observation</i>)	evaluator promatra korištenje sustava u radnom okruženju korisnika
Ciljane skupine (<i>Focus Groups</i>)	veći broj korisnika sudjeluje u raspravi o upotrebljivosti web sučelja, a evaluator je moderator rasprave
Razgovori (<i>Interviews</i>)	evaluator razgovara sa korisnikom postavljajući mu (unaprijed definirana) pitanja
Ankete (<i>Surveys</i>)	evaluator postavlja korisniku specifična pitanja
Upitnici (<i>Questionnaires</i>)	korisnik izražava svoj stav odgovarajući na specifična pitanja
Vođenje dnevnika (<i>Self-Reporting Logs</i>)	korisnik snima operacije korisničkog sučelja
Spremanje izgleda ekrana (<i>Screen Snapshots</i>)	korisnik sprema izgled ekrana web sučelja
Povratna informacija od korisnika (<i>User Feedback</i>)	korisnik predaje svoje bilješke

Podaci o autorima:**Mr.sc. Dijana Plantak Vukovac**

Sveučilište u Zagrebu

Fakultet organizacije i informatike

Pavlinska 2, 42 000 Varaždin

tel: 042 390 850; fax: 042 213 413;

e-mail: dijana.plantak@foi.hr

Dijana Plantak Vukovac diplomirala je na Fakultetu organizacije i informatike (FOI) Sveučilišta u Zagrebu 1999. godine te magistrirala na istom fakultetu 2007. godine s temom „Istraživanje paradigmi razvoja web mjesta“. Radno iskustvo stjecala je na nekoliko radnih mjesta (fotograf, sistem administrator, webmaster), a od 2001. godine radi na FOI u svojstvu znanstvene novakinje i asistentice, trenutno na kolegijima „Obrada teksta i slike“, „Ured i uredsko poslovanje“, „Multimedijski sustavi“ i „Sustavi za elektroničko učenje“. Svoje znanje i vještine usavršavala je radom na nekoliko domaćih znanstvenih i stručnih projekata, međunarodnim Tempus, EUREKA i FP6 projektima, te na različitim radionicama i tečajevima koji uključuju prezentacijske vještine i vještine podučavanja. Područja interesa Dijane Plantak Vukovac su različiti aspekti primjene i razvoja multimedije, naročito razvoj multimedijskih edukacijskih materijala, metode razvoja hipermedijskih sustava i web upotrebljivost.

Tihomir Orehovački, mag. inf.

Sveučilište u Zagrebu

Fakultet organizacije i informatike

Pavlinska 2, 42 000 Varaždin

tel: 042 390 840; fax: 042 213 413;

e-mail: tihomir.orehovacki@foi.hr

Tihomir Orehovački diplomirao je 2005. godine na Fakultetu organizacije i informatike u Varaždinu, smjer Informacijski sustavi. Nakon diplomiranja, kraće je vrijeme radio kao nastavnik u osnovnoj i srednjoj školi. Od 2006. godine zaposlen je na Fakultetu organizacije i informatike u zvanju asistenta te je dosad sudjelovao u izvođenju nastave na predmetima „Programiranje I“, „Programiranje 2“, „Strukture podataka“ i „Računalom posredovana komunikacija“. Praktično znanje i vještine usavršavao je nizom seminara i radionica te radom na međunarodnom Tempus i domaćim projektima. Područja od posebnog interesa su mu upotrebljivost i pristupačnost programskih proizvoda, web 2.0 tehnologije, sigurnost i vjerodostojnost informacija na webu, generativno programiranje i metodika podučavanja programiranja te je iz tih područja objavio dvadesetak znanstvenih i stručnih radova.