

Evaluacija artefakata instanci prilikom provođenja istraživanja znanstvenim oblikovanjem u softverskom inženjerstvu

Marko Mijač

Fakultet organizacije i informatike

Sveučilište u Zagrebu

Pavlinska 2, 42000 Varaždin, Hrvatska

marko.mijac@foi.hr

Sažetak. *Oblikovanje je proces kreiranja primjenjivih rješenja za problem, i kao takvo je već dugo vremena prihvaćena istraživačka paradigma u tradicionalnim inženjerskim disciplinama. U skorije vrijeme, često se koristi u području informacijskih sustava i softverskog inženjerstva.*

Jedan od predloženih pristupa u provođenju sustavnog i metodološkog oblikovanja je znanstvenog oblikovanje. U suštini se radi o pragmatičnoj paradigmi usmjerenoj na rješavanje problema, koja rezultira izradom nekog od sljedećih artefakata: konstrukta, metoda, modela ili instance. Međutim, kako bi znanstveno oblikovanje zaista bilo znanstveno, izrađeni artefakti moraju biti evaluirani na odgovarajući način.

U ovom članku su prikazane smjernice za definiranje i provođenje evaluacije artefakata instanci prilikom provođenja istraživanja znanstvenim oblikovanjem u području softverskog inženjerstva.

Ključne riječi. Znanstveno oblikovanje, artefakti, evaluacija, softversko inženjerstvo

1 Uvod

Prema Merriam-Webster rječniku, oblikovanje (engl. Design) označava planiranje i izradu nečega za specifičnu uporabu ili svrhu. Kao proces izrade

primjenjivih rješenja za problem, oblikovanje je već dugo prihvaćena istraživačka paradigma u tradicionalnim inženjerskim disciplinama. U skorije vrijeme, često se koristi u području informacijskih sustava i softverskog inženjerstva.

Jedan od predloženih pristupa za provođenje sustavnog i temeljitog oblikovanja je znanstveno oblikovanje (engl. Design Science). U suštini se radi o pragmatičnoj paradigmi usmjerenoj na rješavanje stvarnih problema koja rezultira u izradi inovativnih artefakata: konstrukata, metoda, modela i instanci (engl. instantiations) [1].

Iako se svaki od ovih tipova artefakta može pojaviti na pojedinačan rezultat znanstvenog oblikovanja, rješenja se obično sastoje od nekoliko artefakata izgrađenih jedan na temelju drugog. Instance su uobičajeno na vrhu jedne takve hijerarhije, tj. koriste domenske konstrukte i implementiraju modele i metode. March i dr. [1] opisuju instancu kao realizacije artefakta u njegovoj okolini. U kontekstu softverskog inženjerstva, tipični predstavnici artefakata instanci su implementacije i prototipi informacijskih sustava, baza podataka, alata, komponenata, servisa, biblioteka koda, softverskih okvira, algoritama i sl.

Osim što artefakti moraju biti inovativni i relevantni za problemsku domenu, da bi znanstveno oblikovanje bilo zaista

znanstveno, izrađeni artefakti moraju biti evaluirani na odgovarajući način. Doista, evaluacijske aktivnosti su prisutne u svakoj metodi, okviru ili smjernicama za provođenje istraživanja znanstvenim oblikovanjem.

Zbog razlike u svrsi, obliku i karakteristikama, različiti tipovi artefakata mogu zahtijevati različit evaluacijski pristup. Iako nas ovdje primarno zanimaju artefakti instance, evaluacije instanci i evaluacije konstrukata, modela i metoda ugrađenih u instance su obostrano ovisne. S jedne strane, evaluacija konstrukata, modela i metoda kao artefakata niže razine apstrakcije zasigurno podiže cjelokupnu kvalitetu instance kao krajnjeg rezultata. S druge strane, March i dr. [1] tvrde kako kroz izgradnju instanci zapravo operacionaliziramo konstrukte, modele i metode sadržane u instanci, i time demonstriramo izvodljivost i djelotvornost. Nadalje, evaluacijom instanci također pružamo potvrdu vrijednosti artefakata ugrađenih u nju.

U ovom članku istražujemo postojeće metode, uzorke, okvire i smjernice za provođenje evaluacije u znanstvenom oblikovanju. Zatim nadograđujemo postojeće stanje sa vlastitim smjernicama za evaluaciju instanci u znanstvenom oblikovanju.

Članak je strukturiran na sljedeći način. Odjeljak 2 raspravlja općenito o evaluaciji u znanstvenom oblikovanju i njenom položaju u postojećim metodama znanstvenog oblikovanja. U odjeljku 3 raspravljamo o procesu definiranja same evaluacije u znanstvenom oblikovanju, s naglaskom na FEDS okvir [2] i njegovo moguće proširenje sa doprinosima iz drugih istraživanja. U odjeljku 4 spajamo postojeće pristupe i nudimo 7 smjernica za definiranje i provođenje evaluacije u znanstvenom oblikovanju. U konačnici, zaključujemo temu u odjeljku 5.

2 Evaluacija u znanstvenom oblikovanju

2.1 Položaj evaluacije unutar procesa znanstvenog oblikovanja

Evaluacija je proces procjene kvalitete, važnosti ili vrijednosti nečega. Zajedno sa aktivnosti izgradnje, sačinjava interni *izgradi-evaluiraj* ciklus (engl. *build-evaluate*), koji je, prema Hevner [3], srce svakog istraživačkog projekta provedenog znanstvenim oblikovanjem.

Više autora je radilo na formaliziranju procesa evaluacije u znanstvenom oblikovanju. Npr. u metodološkom okviru predloženom od strane autora Johannesson i Perjons [4], dvije od ukupno pet aktivnosti su posvećene evaluaciji. Konkretno, radi se o aktivnostima *demonstriraj artefakt* i *evaluiraju artefakt*. Demonstriranje artefakta ovdje može biti promatrano kao oblik slabe evaluacije, i pokazuje da je artefakt izvodljiv i da radi. Aktivnost evaluacije, s druge strane, cilja provjeriti koliko dobro artefakt radi. Sličan prijedlog dolazi i od Peffers i dr [5], u čijem su modelu procesa znanstvenog oblikovanja također koraci nazvani *demonstracija* i *evaluacija*. Vaishnavi i dr. [6] predlažu opću metodologiju za provođenje znanstvenog oblikovanja, sa jednom od faza nazvanom *evaluacija*. Wieringa [7] gleda znanstveno oblikovanje iz perspektive tzv. inženjerskog ciklusa, te predlaže aktivnosti *validacije* i *evaluacije*. Opisuje validaciju kao način za predviđanje kako međudjelovanja artefakta i njegove okoline, prije nego je artefakt izgrađen. S druge strane, evaluacija istražuje kako se implementirani artefakt ponaša u stvarnosti. Evaluacija je također prisutna kao smjernica u dobro poznatim smjernicama za provođenje znanstvenog oblikovanja od strane Hevner i dr. [8]. Offermann i dr. [9] nude formalizaciju detaljnog procesa znanstvenog oblikovanja, pri čemu je jedna od tri faze *evaluacijska faza*. Sein i dr. [10] stavljaju evaluaciju u drugu fazu (izgradnja, intervencija i *evaluacija*) svoje metode akcijskog oblikovanja (engl. *Action Design Research*).

2.2 Ciklusi evaluacije

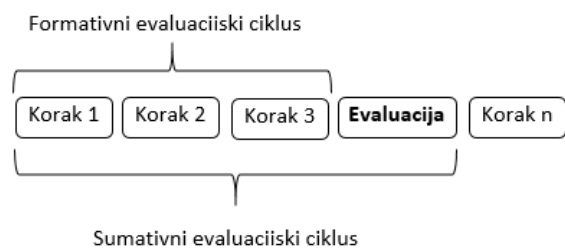
Kao što se može vidjeti, evaluacija je sastavni dio svakog formalnog procesa znanstvenog oblikovanja. Većina pristupa prikazuje evaluaciju kao odvojenu fazu ili korak koji se provodi nakon što je artefakt oblikovan i izgrađen. Međutim, proces znanstvenog oblikovanja se ne provodi nužno u stilu vodopadnog modela, nego može sadržavati iteracije i cikluse. Npr. rezultati iz evaluacijske aktivnosti mogu uzrokovati vraćanje na prethodne faze zbog otkrivanja nedostataka u oblikovanju artefakta, zbog promjena u shvaćanju samog problema, ili jednostavno zbog pojave novih ideja koje bi poboljšale artefakt. Iako njihov okvir izgleda kao slijedan skup aktivnosti, Johannesson i Perjons [4] ipak naglašavaju da se znanstveno oblikovanje uvijek izvodi na iterativan način, pomicanjem naprijed-nazad između svih aktivnosti. Offermann i dr. [9] također naglašavaju da je se, ovisno o rezultatu evaluacijske faze, moguće vratiti natrag na prethodne faze. Sein i dr. [10] idu još i dalje u njihovoj metodi akcijskog oblikovanja. Oni tvrde da je aktivnosti evaluacije svojstvena isprepletenost sa aktivnošću izgradnje artefakta i aktivnosti intervencije u organizaciju, zbog čega one moraju biti provedene istovremeno.

Kada raspravljamo o svrsi evaluacije, Venable i dr. [11] razlikuju formativnu i sumativnu evaluaciju. Formativna evaluacija se usredotočuje na pružanje povratne informacije i mjerenje napretka u toku razvoja artefakta. S druge strane, sumativna evaluacija podupire proces formiranja mišljenja i usporedbe artefakta, nakon što je razvoj završen.

U konačnici, neovisno o odabranog procesu znanstvenog oblikovanja, možemo identificirati dva evaluacijska ciklusa: formativni i sumativni. U formativnom evaluacijskom ciklusu evaluacija se provodi kontinuirano i usporedo sa oblikovanjem i izgradnjom artefakta. Pri tome imamo za cilj što ranije moguće dati povratnu informaciju, kako bi inkrementalno radili na usavršavanju artefakta. U ovom ciklusu može doći do

velikog broja iteracija sa implicitnim ili eksplicitnim mikro evaluacijama. Sumativni evaluacijski ciklus, s druge strane, podrazumijeva da je artefakt već izgrađen, te da eksplicitna formalna evaluacija artefakta kao krajnjeg rezultata znanstvenog oblikovanja može početi. U ovom evaluacijskom koraku, artefakt može biti ocijenjen kao nezadovoljavajući, što također može zahtijevati povratak na prethodne korake i rješavanje nedostataka. Međutim, broj iteracija u sumativnom evaluacijskom ciklusu je obično mnogo manji. Potrebno je pri tome naglasiti, da će evaluacija rijetko zaključiti da je artefakt savršen i da nema mjesta daljnjem poboljšanju. Zbog toga bi istraživač morao voditi računa o ciljevima i ograničenjima istraživačkog projekta, te procijeniti kada iteracije i poboljšanja treba u potpunosti zaustaviti, ili ih odgoditi za buduća istraživanja.

Slika 1 Evaluacijski ciklusi u znanstvenom oblikovanju



2.3 Instance

Gregor i Jones [12] opisuju instance kao materijalne artefakte koji fizički postoje u stvarnom svijetu, te se u temelju razlikuju od apstraktnih artefakata, tj. konstrukata, modela i metoda. March i dr. [1] ukazuju da je instanca zapravo realizacija artefakta u njegovom okruženju. Slično tome, Johannesson i Perjons [4] opisuju instancu kao funkcionirajući sustav koji može biti korišten u praksi.

Instance također mogu biti okarakterizirane u smislu razlike između *artefakata proizvoda* i *artefakata procesa* [11]. Dok artefakti procesi predstavljaju metode i procedure koje usmjeravaju ljude tokom izvršenja nekog zadatka, artefakti proizvodi predstavljaju

alate, dijagrame, softver i sl., koje ljudi koriste za izvršenje nekog zadatka. Očito je da će se u softverskom inženjerstvu artefakti instance većinski pojavljivati kao artefakti proizvodi.

Još jedan pogled na instance kao artefakte u softverskom inženjerstvu je iz perspektive *tehničkih* i *sociotehničkih* artefakata [11]. U tom smislu, većinu instanci u softverskom inženjerstvu možemo svrstati u sociotehničke artefakte, što podrazumijeva da se radi o tehničkim sustavima koji moraju biti u interakciji sa ljudima da bi bili korisni (npr. informacijski sustavi, ERP sustavi, računalne igre, CASE alati, i dr.). S druge strane, instance se također mogu pojaviti kao pretežito ili čak u potpunosti tehnički artefakti, što znači da je interakcija sa ljudima minimalna ili nije uopće potrebna (npr. softverske komponente ugrađene u veće, uglavnom sociotehničke artefakte).

Gledano iz perspektive evaluacije u znanstvenom oblikovanju, instance su osobito važne. Npr. March i dr. [1] tvrde da instance operacionaliziraju ugrađene konstrukte, modele i metode, i time demonstriraju njihovu izvodljivost i djelotvornost. Slično tome, prema Hevner i dr. [8] instance pokazuju kako konstrukti, modeli i metode mogu biti implementirani u sustav koji radi. Instance demonstriraju izvodljivost, te omogućavaju izravnu ocjenu prikladnosti artefakta za onu svrhu za koji se namjerava koristiti. Gregor i Jones [12] zaključuju da dok je se konceptualni rad na oblikovanju dokazao kao utjecajan, može se očekivati da će vjerodostojnost tog rada biti povećana pružanjem instance kao praktičnog primjera.

3 Oblikovanje evaluacije

Iako svi procesi znanstvenog oblikovanja uključuju evaluaciju i nude općenite savjete za njeno provođenje, nedostaje im jasna i detaljna procedura evaluacije. Johannesson i Perjons [4] navode kako je upravo uporaba znanstvenih istraživačkih metoda u provođenju evaluacije ključ za razlikovanje znanstvenog od rutinskog oblikovanja. Venable i dr. [2] dodaju da, ukoliko će

znanstveno oblikovanje zaslužiti svoj naziv, evaluacija mora biti relevantna, temeljita i znanstvena. Kako bi planirali, oblikovali i proveli tako rigorozne evaluacijske aktivnosti, potrebne su odgovarajuće procedure, okviri i smjernice. Ovisno o karakteristikama samog istraživačkog projekta koji se provodi znanstvenim oblikovanjem, te samog artefakta, takve smjernice bi nam trebale pomoći u odluci kada, što, zbog čega i kako evaluirati.

3.1 Postojeći pristupi

U skorije vrijeme, više autora se bavilo problemom oblikovanja evaluacije u okviru znanstvenog oblikovanja. Pries-Heje [13] je predložio strateški okvir za evaluaciju u znanstvenom oblikovanju, koji može biti korišten kao pomoć pri odabiru prikladne evaluacijske strategije za nova istraživanja, kao i za klasifikaciju evaluacijskih strategija u već objavljenim istraživanjima. Okvir propisuje dvije dimenzije: *ex-ante* nasuprot *ex-post* evaluacije kao prva, te *naturalistička* nasuprot *vještačka* evaluacija kao druga dimenzija. Cleven i dr. [14] su predstavili morfološko polje sa 12 varijabli i njihovih respektivnih vrijednosti, koje se može koristiti prilikom odlučivanja o različitim alternativama za evaluacijsku strategiju. Venable i dr. [2] su razvili okvir za evaluaciju u znanstvenom oblikovanju (FEDS), koji propisuje proceduru od 4 koraka za oblikovanje evaluacijske strategije. Sonnenberg i Brocke [15] predstavljaju opći evaluacijski uzorak, koji propisuje četiri evaluacijske aktivnosti provedene kroz cijeli proces znanstvenog oblikovanja.

3.2 Odabrani pristup

Kako bi raspravili detalje oblikovanja evaluacije, osloniti ćemo se na FEDS okvir predložen od autora Venable i dr. [2]. Okvir propisuje sljedeću proceduru od četiri koraka za oblikovanje evaluacije: (1) *razjasni ciljeve evaluacije*, (2) *odaberi evaluacijsku strategiju*, (3) *utvrdi svojstva za evaluaciju*, (4) *oblikuj pojedinačne evaluacijske epizode*.

Međutim, s obzirom da se FEDS okvir ne evaluacijskim kriterijima na sustavan način, niti ih stavlja u odnos sa metodama evaluacije, upotpuniti ćemo korake (3) i (4) sa rezultatima iz drugih relevantnih istraživanja.

3.2.1 Razjasni ciljeve evaluacije

Venable i dr. [2] navode četiri konkurentna cilja koje moramo uzeti u obzir prilikom oblikovanja evaluacije: (1) rigoroznost, (2) nesigurnost i smanjenje rizika, (3) etika i (4) efikasnost.

Rigoroznost ovdje promatramo u smislu uspješnosti (engl. *efficacy*) – utvrđivanja da su poboljšanja zaista uzrokovana artefaktom, te efektivnosti (engl. *effectiveness*) – utvrđivanja da artefakt radi u realnim situacijama. Potrebno je pri tome napomenuti da je formativni evaluacijski ciklus primjereniji za evaluaciju uspješnosti, a sumativni za evaluaciju efektivnosti.

Nesigurnost i smanjenje rizika podrazumijeva napor da se eliminiraju ili umanje društveni i tehnički rizici što je prije moguće. Formativna evaluacija je po definicije osobito važna za ovaj cilj.

Na *etička* pitanja treba obratiti pažnju osobito kada evaluiramo artefakte koji uvode rizike za sigurnost, zdravlje ili privatnost.

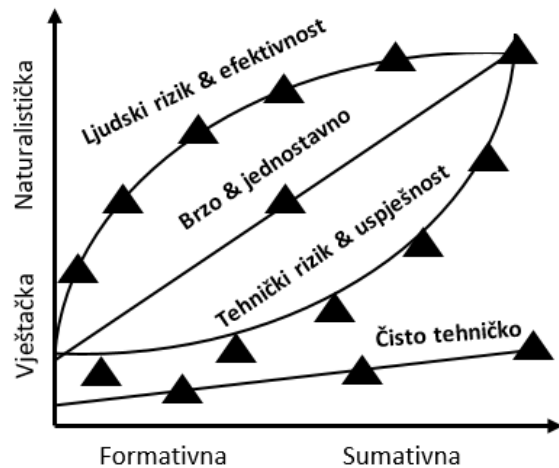
U konačnici, evaluacije bi trebala biti efikasna (engl. *efficient*) u smislu da bude izvodljiva unutar ograničenih istraživačkih resursa (vremena, novca, ljudi...).

3.2.2 Odaberi evaluacijsku strategiju

Da bi opisali i pozicionirali različite evaluacijske strategije, Venable i dr. [2] predlažu dvodimenzionalni prostor sa dimenzijama: (1) *funkcionalna svrha* i (2) *evaluacijska paradigma*. Funkcionalna svrha se bavi pitanjem zašto evaluirati i pozicionira evaluaciju prema već spomenutom kontinuumu formativna-sumativna evaluacija. S druge strane, evaluacijska paradigma se bavi pitanjem kako evaluirati i pozicionira se na kontinuumu vještačka-naturalistička evaluacija. Kao što i sami naziv ukazuje, vještačka evaluacija se izvodi u

vještačkom okruženju (npr. laboratorij ili simulator), dok se naturalistička evaluacija izvodi pod realističnim uvjetima, u stvarnom, ili najbliže moguće stvarnom okruženju.

Slika 2 Dvodimenzionalni prostor za evaluacijske strategije [2]



U ovom dvodimenzionalnom prostoru, evaluacijska strategija je predstavljena kao putanja formirana spajanjem pojedinačnih evaluacijskih epizoda. Pri čemu su evaluacijske epizode prethodno pozicionirane u skladu sa navedene dvije dimenzije.

Venable i dr. [2] predlažu četiri arhetipa evaluacijskih strategija, i to: *Ljudski rizik & efektivnost*, *Brzo & jednostavno*, *Tehnički rizik & uspješnost*, i *Čisto tehničko*. Uz ove arhetipe priložena je i jednostavna heuristika koja pomaže istraživačima pri odabiru najprikladnijeg arhetipa za konkretno istraživanje. Međutim, autori naglašavaju da je svako istraživanje provedeno znanstvenim oblikovanjem specifično, i potiču istraživače da po potrebi prilagode predložene arhetipe, ili čak da predlože u potpunosti nove evaluacijske strategije.

3.2.3 Utvrdi svojstva za evaluaciju

Kako bi utvrdili koja točno svojstva artefakta instance evaluirati, potrebno je uzeti u obzir više kriterija, uključujući općeniti cilj evaluacije, odabranu strategiju, karakteristike artefakta i svrhu artefakta. Prema Venable i dr. [2] konačan odabir svojstava za evaluaciju je nužno jedinstven za artefakt.

Različiti autori su predložili različita svojstva/kriterije za evaluaciju artefakata instanci. March i dr. [1], na primjer, spominju *efikasnost, efektivnost, i utjecaj artefakta na okolinu i korisnike*. Hevner i dr. [8] u svojim smjernicama za evaluaciju navode da artefakt mora demonstrirati *korist, kvalitetu i uspješnost*.

Iako ne postoji u općeprihvaćen popis, Prat i dr. [16] su analizirali literaturu iz područja znanstvenog oblikovanja i formirali popis evaluacijskih svojstava uključujući učestalost njihove uporabe. Ovaj popis predstavlja dobru početnu točku za odabir svojstava za evaluaciju koja su prikladna za pojedini artefakt. Međutim, treba imati na umu i učestalost uporabe ovih svojstava u literaturi, jer veća učestalost može ukazivati na već ustaljene dobre prakse i moguću bolju prihvaćenost od strane recenzenata.

Sa originalnog popisa svojstava [16] smo izostavili svojstvo *manjak konstrukta*, (engl. *construct deficit*) jer se očito odnosi na evaluaciju artefakta konstrukta. Osim toga, možemo tvrditi da svojstva prikazana u **Tablica 2** mogu biti primijenjena prilikom evaluacije artefakata instanci.

3.2.4 Oblikuj pojedinačne evaluacijske epizode

Evaluacijska epizoda *Ep* može biti definirana kao konkretna evaluacija o okviru evaluacijske strategije, karakterizirana sa 4 dimenzije: svrha evaluacije (*Pu*), evaluacijska paradigma (*Pa*), metoda evaluacije (*M*) i jedno ili više evaluacijskih svojstava (*Pr*):

$$Ep = \{Pu, Pa, M, Pr(p1, p2, \dots)\}, \text{ gdje je}$$

$$Pu = (\text{formativna} \mid \text{sumativna}),$$

$$Pa = (\text{vještačka} \mid \text{naturalistička}),$$

$$Pr = (\text{uspješnost} \mid \text{korisnost} \mid \dots),$$

$$M = (\text{eksperiment} \mid \text{studija slučaja} \mid \dots),$$

Dok smo već raspravljali o svrsi evaluacije, paradigmi i svojstvima, još uvijek ostaje pitanje mogućih metoda evaluacije. **Tablica 1** prikazuje neke od najčešće spominjanih

metoda i uzoraka evaluacije u literaturi znanstvenog oblikovanja.

Neki od autora su također proveli pregled znanstvenih istraživanja provedenih znanstvenim oblikovanjem kako bi utvrdili koje se metode evaluacije zaista koriste. Na primjer, Peffers i dr. [17] su analizirali 148 članaka, te su izvijestili da su tehnički eksperiment, eksperiment temeljen na subjektu, prototipiranje, i demonstracija na ilustrativnim scenarijima dominantne metode evaluacije za artefakte instance. Prat i dr. [16] su, analizirajući 121 istraživanje provedeno znanstvenim oblikovanjem, razvili taksonomiju metoda za evaluaciju. Pri tome su identificirali demonstraciju (na ilustrativnim ili stvarnim primjerima), simulaciju, vrednovanje (engl. *benchmarking*), studiju slučaja i kontrolni eksperiment kao najzastupljenije tehnike evaluacije za artefakte instance.

Tablica 1 Metode i uzorci evaluacije u istraživanjima provedenim znanstvenim oblikovanjem

Metode/uzorci evaluacije	Spominje se u
Eksperiment	[18][2][8][17][6][9][7][16][14]
Studija slučaja	[18][2][8][17][9][7][16][14]
Simulacija	[18][2][8][6][9][7][16]
Informirani argument	[18][2][8][17][6][15]
Demonstracija / Scenariji	[18][8][17][6][15][16]
Terenska studija	[18][2][8][14]
Matematički dokazi	[18][2][6][14]
Anketa	[18][2][7][14]
Akcijsko istraživanje	[2][17][9][14]
Ekspertna evaluacija	[18][17][9]
Vrednovanje	[18][6][16]
Statička / Dinamička analiza	[18][17][16]
Prototipiranje	[17][15][14]
Testiranje	[8][7]
Metrike	[2][16]

Tablica 2 Frekvencija učestalosti (stupac f) evaluacijskih svojstava u istraživanjima provedenim znanstvenim oblikovanjem (prilagođeno iz Prat i dr. [16])

Svojstvo	Opis	f (%)
Uspješnosti (engl. <i>Efficacy</i>)	Stupanj do kojeg artefakt postiže svoj cilj, gledano usko, ne uzimajući u obzir situacijske aspekte.	80%
Korisnost (engl. <i>Usefulness</i>)	Stupanje do kojeg artefakt pozitivno utječe na učinak pojedinaca.	35%
Tehnička izvodljivost (engl. <i>Technical feasibility</i>)	Ocjenjuje lakoću korištenja i rukovođenja predloženim artefaktom s tehničkog gledišta.	32%
Točnost (engl. <i>Accuracy</i>)	Stupanj podudaranja između rezultata artefakta i očekivanih rezultata.	28%
Učinak (engl. <i>Performance</i>)	Stupanj u kojem artefakt izvršava svoje funkcije unutar zadatih ograničenja vremena ili prostora.	23%
Efektivnost (engl. <i>Effectiveness</i>)	Stupanj do kojeg artefakt postiže svoj cilj u stvarnoj situaciji.	18%
Lakoća korištenja (engl. <i>Ease of use</i>)	Stupanj truda koji korištenje artefakta iziskuje od pojedinca.	10%
Robusnost (engl. <i>Robustness</i>)	Sposobnost artefakta da se nosi sa nevaljanim ulazima ili stresnim okolnim uvjetima.	10%
Skalabilnost (engl. <i>Scalability</i>)	Sposobnost artefakta da podnese povećanje obima posla, ili da ga je moguće nadograditi.	10%
Izvodljivost u radu (engl. <i>Operational feasibility</i>)	Stupaj do kojeg će menadžment, zaposlenici, i drugi dionici poduprijeti predloženi artefakt, koristiti ga i integrirati u svoje svakodnevne poslove.	10%
Korist (engl. <i>Utility</i>)	Mjera postizanja cilja artefakta, tj. razlika između vrijednosti postizanja tog cilja i cijene koja je za to plaćena.	7%
Valjanost (engl. <i>Validity</i>)	Valjanost znači da artefakt radi ispravno, tj. da ispravno postiže svoj cilj.	6%
Potpunost (engl. <i>Completeness</i>)	Stupanj do kojeg aktivnost artefakta sadrži sve potrebne dijelove i veze između tih dijelova.	3%
Prilagodljivost (engl. <i>Adaptability</i>)	Lakoća s kojom artefakt može raditi u kontekstima za koje nije primarno namijenjen. Sinonim: fleksibilnost.	2%
Pouzdanost (engl. <i>Reliability</i>)	Sposobnost artefakta da ispravno radi u danom okruženju tijekom određenog vremenskog perioda.	2%
Sposobnost učenja (engl. <i>Learning capability</i>)	Sposobnost artefakta da uči iz iskustva.	2%
Jednostavnost (engl. <i>Simplicity</i>)	Stupanj do kojeg struktura artefakta sadrži minimalan broj dijelova i veza između tih dijelova.	1%
Ekonomska izvodljivost (engl. <i>Economic feasibility</i>)	Evaluiira hoće li prednosti predloženog artefakta prevagnuti troškove izgradnje i rada sa artefaktom.	1%
Općenitost (engl. <i>Generality</i>)	Odnosi se na opseg cilja artefakta. Što je širi opseg cilja, to je artefakt općenitiji.	1%

Prat i dr. [16] su također identificirali 6 često korištenih stilova kompozicije za evaluaciju artefakata instanci, i to: (1) *demonstracija*, (2) *simulacija i vrednovanje temeljno na metrikama*, (3) *evaluacija efektivnosti u praksi*, (4) *simulacija i apsolutna evaluacija temeljena na metrikama*, (5) *evaluacija korisnosti ili lakoće korištenja u praksi*, (6) *laboratorijska evaluacija korisnosti na studentima*. Prilikom definiranja evaluacijske strategije, ovi stilovi kompozicije mogu biti korišteni kao već provjerene evaluacijske epizode.

4 Smjernice za oblikovanje i provođenje evaluacije instanci

Evaluacija se smatra jednom od ključnih aktivnosti u istraživanju znanstvenim oblikovanjem. To je vidljivo iz istraživanja koje se bave teorijom znanstvenog oblikovanja ali i iz istraživanja koje se provode znanstvenim oblikovanjem. U cilju provođenja evaluacije u znanstvenom oblikovanju na rigorozan i sustavan način, predstavljamo nekoliko smjernica sintetiziranih iz literature o znanstvenom oblikovanju.

Smjernica 1 – *Koristi prihvaćene okvire za provođenje znanstvenog oblikovanja*

Ako zanemarimo za trenutak važnost odabira relevantnog istraživačkog problema, prvi korak koji trebamo napraviti u kontekstu evaluacije u znanstvenom oblikovanju je odabrati odgovarajuću metodu/proces provedbe znanstvenog oblikovanja. Iako odabir i primjena dobre metode/procesa nije nužno garancija da će i artefakt biti uspješan, zasigurno povećava šanse da se to dogodi. Uz to, svaka metoda znanstvenog oblikovanja uključuje i evaluacijski korak, koji zatim pozicionira u odnosu na druge aktivnosti znanstvenog oblikovanja. Primjeri formaliziranih metoda za provođenje znanstvenog oblikovanja se mogu naći u [4], [5], [6], [7], [8], [9], [10].

Smjernica 2 – *Koristi postojeće okvire za oblikovanje evaluacije*

Nakon što je odabrana općenita metoda za provođenje znanstvenog oblikovanja, sljedeći korak vezan uz evaluaciju je oblikovanje evaluacije. Oblikovanje evaluacije je složen zadatak, i, kao i sam proces znanstvenog oblikovanja, treba biti proveden sustavno. Da bi to napravili, možemo koristiti jedan ili kombinaciju od više postojećih pristupa navedenih u odjeljku 3.1., uključujući: [2], [13], [14], [15]. Međutim, mišljenja smo da FEDS okvir [2], sa svojom procedurom od četiri koraka, nudi najbliže sveobuhvatnom vođenju.

Smjernica 3 – *Prilikom oblikovanja evaluacije razmisli o evaluaciji uobičajenih svojstava artefakta*

Prilikom određivanja svojstava artefakta koja će biti evaluirana, potrebno je uzeti u obzir istraživanja navedena u smjernici 2. Npr., FEDS okvir [2] upravo nudi heuristiku za to. Međutim, svakako bi trebalo uzeti u obzir i popis često evaluiranih svojstava u znanstvenom oblikovanju, naveden u **Tablica 2**. Često evaluirana svojstva mogu ukazivati

na ustaljene dobre prakse i moguću bolju prihvaćenost od recenzenata.

Smjernica 4 – *Uzmi u obzir često korištene evaluacijske metode prilikom oblikovanja evaluacije*

Iako ništa ne sprječava istraživača da odabere koju god metodu smatra prikladnom za evaluaciju pojedinog svojstva, korisno je uzeti u obzir evaluacijske metode koje se često koriste. **Tablica 1** sadrži popis evaluacijskih metoda koje se često spominju u literaturi o znanstvenom oblikovanju. Također, članci [16] i [17] spominju evaluacijske metode često primijenjene u istraživanjima znanstvenim oblikovanjem.

Smjernica 5 – *Prilikom oblikovanja evaluacije uzmi u obzir često korištene stilove kompozicije za evaluaciju*

Oblikovanje konkretnih evaluacijskih epizoda unutar cjelokupne evaluacijske strategije uključuje određivanje koja će metoda biti korištena za evaluaciju kojeg svojstva artefakta. Iako se velik broj *svojstvo-metoda* kombinacija može formirati uparivanjem svakog svojstva sa svakom metodom, neke od ovih kombinacija su više uobičajene od drugih. Česti stilovi kompozicije navedeni u [16] se mogu uzeti u obzir prilikom odlučivanja o evaluacijskog strategiji i pojedinačnim epizodama.

Smjernica 6 – *Koristi prikladne okvire za provođenje određene metode evaluacije*

Korištenje određene metode evaluacije unutar evaluacijske epizode je često cijelo istraživanje unutar istraživanja. Istraživačke metode koje se koriste kao metode evaluacije u znanstvenom oblikovanju u pravilu imaju jasno definirane korake za provođenje, npr. planiranje, prikupljanje podataka, analiziranje podataka i izvještavanje. Preporučljivo je pokušati pronaći okvire za provođenje pojedine metode evaluacije u samom znanstvenom oblikovanju i, ukoliko postoje, koristiti njih. Na primjer, sljedeće metode se

spominju u kontekstu evaluacije u znanstvenom oblikovanju: Fokus grupe [19], evaluacija ugrađena u softver (engl. *Software-embedded evaluation*) [20], Tehničko-akcijsko istraživanje [21], i eksperimentiranje [22][23]. Alternativno, okviri i metode korištene u kontekstu softverskog inženjerstva ili drugih srodnih područja mogu biti također u potpunosti prikladne, npr.: studija slučaja [24], eksperimentiranje [25], akcijsko istraživanje [26] itd.

Smjernica 7 – Uzmi u obzir korištenje postojećih modela kvalitete softvera i metrika prilikom evaluacije instanci

U slučaju kada je evaluirani artefakt proizvod, prema Pries-Heje i dr. [13] možemo koristiti postojeće modele kvalitete softvera za evaluaciju. Postoje različiti modeli kvalitete koji su predloženi za ocjenu kvalitete softverskih proizvoda. Jedan od njih je opisan i u ISO/IEC 25010:2011 standardu [27]. Ovaj model kvalitete propisuje osam karakteristika kvalitete (dalje podijeljenih u podkarakteristike) uključujući i odgovarajuće mjere i funkcije korištene za kvantifikaciju ovih karakteristika.

5 Zaključak

U ovom članku raspravljali smo o evaluaciji artefakata instanci u okviru znanstvenog oblikovanja. Evaluacija predstavlja značajan poduhvat i često cijelo novo istraživanje unutar istraživanja znanstvenim oblikovanjem. Kako bi pomogli u oblikovanju i provođenju sustavne i rigorozne evaluacije, nudimo 7 smjernica. Radi se o smjernicama više razine apstrakcije u smislu da se ne bave specifičnim metodama i kriterijima evaluacije. Umjesto toga, usmjeravaju istraživača prema postojećim okvirima i metodama za pozicioniranje evaluacije unutar znanstvenog oblikovanja, oblikovanje evaluacije, i odabir svojstava i metoda evaluacije. Iako je članak usmjeren na evaluaciju artefakata instanci, smjernice se u većoj mjeri mogu primijeniti i na druge tipove artefakata.

Reference

- [1] S. T. March and G. F. Smith, "Design and natural science research on information technology," *Decis. Support Syst.*, vol. 15, no. 4, pp. 251–266, Dec. 1995.
- [2] J. Venable, J. Pries-Heje, and R. Baskerville, "FEDS: a Framework for Evaluation in Design Science Research," *Eur. J. Inf. Syst.*, Studeni 2014.
- [3] A. Hevner, "A Three Cycle View of Design Science Research," *Scand. J. Inf. Syst.*, vol. 19, no. 2, Jan. 2007.
- [4] P. Johannesson and E. Perjons, *An introduction to design science*. 2014.
- [5] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A Design Science Research Methodology for Information Systems Research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, Dec. 2007.
- [6] V. Vaishnavi, *Design science research methods and patterns: innovating information and communication technology*. Boca Raton: Auerbach Publications, 2008.
- [7] R. J. Wieringa, *Design Science Methodology for Information Systems and Software Engineering*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.
- [8] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004.
- [9] P. Offermann, O. Levina, M. Schönherr, and U. Bub, "Outline of a design science research process," 2009, p. 1.
- [10] M. K. Sein, O. Henfridsson, S. Purao, M. Rossi, and R. Lindgren, "Action Design Research," *MIS Q.*, vol. 35, no. 1, pp. 37–56, Ožujak 2011.
- [11] J. Venable, J. Pries-Heje, and R. Baskerville, "A Comprehensive Framework for Evaluation in Design Science Research," in *Design Science Research in Information Systems. Advances in Theory and Practice*, K. Peffers, M. Rothenberger, and B. Kuechler, Eds. Springer Berlin Heidelberg, 2012, pp. 423–438.
- [12] S. Gregor and D. Jones, "The Anatomy of a Design Theory," *J. Assoc. Inf. Syst. Atlanta*, vol. 8, no. 5, pp. 312–323, 325–335, May 2007.
- [13] J. Pries-Heje, R. Baskerville, and J. Venable, "Strategies for Design Science Research Evaluation," *ECIS 2008 Proc.*, Jan. 2008.
- [14] A. Cleven, P. Gubler, and K. M. Hüner, "Design Alternatives for the Evaluation of Design Science Research Artifacts," in *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, New York, NY, USA, 2009, pp. 19:1–19:8.
- [15] C. Sonnenberg and J. vom Brocke, "Evaluation patterns for design science research artefacts," in

Practical Aspects of Design Science, Springer, 2011, pp. 71–83.

- [16] N. Prat, I. Comyn-Wattiau, and J. Akoka, “A Taxonomy of Evaluation Methods for Information Systems Artifacts,” *J. Manag. Inf. Syst.*, vol. 32, no. 3, pp. 229–267, Jul. 2015.
- [17] K. Peffers, M. Rothenberger, T. Tuunanen, and R. Vaezi, “Design Science Research Evaluation,” in *Design Science Research in Information Systems. Advances in Theory and Practice*, K. Peffers, M. Rothenberger, and B. Kuechler, Eds. Springer Berlin Heidelberg, 2012, pp. 398–410.
- [18] C. Sonnenberg and J. vom Brocke, “Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research,” in *Design Science Research in Information Systems. Advances in Theory and Practice*, vol. 7286, K. Peffers, M. Rothenberger, and B. Kuechler, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 381–397.
- [19] M. Tremblay, A. Hevner, and D. Berndt, “Focus Groups for Artifact Refinement and Evaluation in Design Research,” *Commun. Assoc. Inf. Syst.*, vol. 26, no. 1, Jun. 2010.
- [20] L. Chandra Kruse and et al., “Software Embedded Evaluation Support in Design Science Research,” presented at the Pre-ICIS workshop on Practice-based Design and Innovation of Digital Artifacts, 2016.
- [21] R. Wieringa and A. Morali, “Technical Action Research as a Validation Method in Information Systems Design Science,” in *Design Science Research in Information Systems. Advances in Theory and Practice*, K. Peffers, M. Rothenberger, and B. Kuechler, Eds. Springer Berlin Heidelberg, 2012, pp. 220–238.
- [22] L. Ostrowski and M. Helfert, *Design Science Evaluation – Example of Experimental Design*. .
- [23] T. Mettler, M. Eurich, and R. Winter, “On the Use of Experiments in Design Science Research: A Proposition of an Evaluation Framework,” *Commun. Assoc. Inf. Syst.*, vol. 34, no. 1, Jan. 2014.
- [24] B. Kitchenham, L. Pickard, and S. L. Pfleeger, “Case studies for method and tool evaluation,” *IEEE Softw.*, vol. 12, no. 4, pp. 52–62, Jul. 1995.
- [25] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.
- [26] D. E. Avison, F. Lau, M. D. Myers, and P. A. Nielsen, “Action research,” *Commun. ACM*, vol. 42, no. 1, pp. 94–97, Jan. 1999.
- [27] ISO, “ISO/IEC 25010:2011 - Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models.” 2011.