# ODPRTA VPRAŠANJA IN DILEME KVANTITATIVNEGA IN KVALITATIVEGA OCENJEVANJA V ZNANOSTI

**Franc Mali**
Fakulteta za družbene vede

Kontaktni naslov:
franc.mali@fdv.uni-lj.si

**Janez Jug**
Fakulteta za družbene vede

Kontaktni naslov:
janez.jug@fdv.uni-lj.si

URL:
http://www.fdv.uni-lj.si

**Izvleček**
Prispevek se uvodoma ukvarja z vprašanjem, kdaj in zakaj se je v razvitem znanstvenem okolju pojavila zahteva, da se poleg tradicionalnih kolegialnih ocenjevanj v znanosti uveljavi sistem kvantitativnega (bibliometričnega) spremljanja znanosti. Avtorja skušata prikazati stanje in trende razvoja na področju kvantitativnih in kvalitativnih ocenjevanj znanstvenih rezultatov danes in v prihodnje. Še posebej ju zanima delovanje evalvacijskega sistema znanosti v novih načinih produkcije znanstvenega vedenja. Zanima ju, kateri so ključni problemi zanesljivosti in veljavnosti kvantitativnih in kvalitativnih ocenjevanj v znanosti. V zadnjem delu so predstavljeni nekateri predlogi za dopolnitve veljavnega sistema vrednotenja rezultatov znanstvenega raziskovanja v Sloveniji, ki ga je v zadnjem času vzpostavila Javna agencija za raziskovalno dejavnost Republike Slovenije.

**Ključne besede**
bibliometrija, kolegialna kontrola v znanosti, indeks znanstvenih citatov, znanstvena publicistika, znanstvena politika, kriterij veljavnosti in zanesljivosti v kvantitativnih in kvalitativnih merjenjih znanosti, znanstveni razvoj

**Abstract**
The article is dealing with the issue, what are the reasons that bibliometrics has appeared in the second half of 20th century as additional element in R&D evaluations (along peer review which has longer historical tradition). The authors of this article try to analyse the state of the art and the development of quantitative and qualitative R&D evaluations in different part of the world. They give special attention to the functioning of R&D evaluation system in the new mode of knowledge production. In the article, the key interest is oriented to the criteria of validity and certainty in R&D evaluation procedures. In the last part of the article, there are elaborated some recommendations to improve the system of R&D evaluation established in the last few months by Slovenian Research Agency (ARRS).

**Keywords**
bibliometrics, peer review, science citation index, scientific publications, science policy, validity and reliability of quantitative and qualitative measures of science, scientific development

## UVOD

Pri obravnavanju odprtih vprašanj in dilem kvantitativnega in kvalitativnega ocenjevanja znanstvenih rezultatov ter z njimi povezanih dilem znanstvene uspešnosti moramo najprej poudariti, da lahko govorimo o nekakšni povsodnosti (ubikvitarnosti) ocenjevanja v moderni znanosti. Ocenjevanje je neizogiben element katere koli znanstvene dejavnosti, čeprav se včasih zdi, da postopke znanstvenega ocenjevanja omejujemo samo na njihovo vlogo pri vodenju konkretnih raziskovalno-razvojnih politik. Dejstvo je, da so raznovrstne oblike ocenjevanja v znanosti – če uporabimo izraz sodobne sociologije – del vsakdanjega življenjskega sveta znanosti. Naj naštejemo samo nekatere najbolj tipične primere:

1. ocenjevanje poteka že s samim citiranjem, saj s tem, ko v znanosti citiramo besedilo svojega profesionalnega kolega, že opravimo evalvacijo besedila;
2. ocenjevanje vključuje celotni recenzijski sistem v znanosti, saj v večini primerov uredniki znanstvenih revij ter založniki znanstvenih monografij šele na temelju ekspertnih ocen oziroma recenzij objavijo rezultate neke raziskave;
3. znanstveniki se na svoji karierni poti srečujejo z različnimi oblikami ocenjevanja njihovega dela (habilitacijski postopki, razpisi itd.).

Uvodoma je treba podati osnovno razliko med kategorijo kvalitativnega (ekspertnega) in kvantitativnega (biblio-

metričnega) ocenjevanja v znanosti. Pri kvalitativnem ocenjevanju gre za t. i. kolegialno kontrolo (angl. *peer review*) med znanstveniki samimi. Kolegi po isti stroki oziroma kompetencah znotraj svojega področja čim bolj nepristransko ocenjujejo primernost predlogov raziskovalnih projektov, prispevkov za objavo v znanstvenih revijah, kongresnih zbornikih itd. V nasprotju s kvantitativnim tipom ocenjevanja kvalitativni tip ocenjevanja posredno vključuje merjenje znanstvene kakovosti z različnimi vrstami bibliometričnih kazalcev. Takšni kvantitativni (bibliometrični) kazalci so: število znanstvenih objav, število citatov, število soavtorstev itd. Med kvantitativne (bibliometrične) vrste kazalcev štejemo tudi soobjave, socitiranje, indekse podobnosti, analize semantičnih omrežij itd. V zvezi s kvantitativnimi kazalci se v stroki večkrat postavlja vprašanje, ali predvsem druga skupina (kombiniranih) kazalcev lahko nadomesti ustaljene tipe kvantitativnih kazalcev. O tem vprašanju na tem mestu ne bomo obširno razpravljali. Bolj nas bodo zanimali odgovori na naslednja vprašanja:

1. Kdaj in zakaj se je v razvitem znanstvenem okolju pojavila zahteva, da se poleg tradicionalnih kolegialnih ocenjevanj v znanosti uveljavi sistem kvantitativnega (bibliometričnega) spremljanja znanosti?
2. Kakšno je stanje na področju kvantitativnega in kvalitativnega ocenjevanja znanstvenih rezultatov danes in v kolikšni meri novi načini produkcije znanstvenega vedenja vplivajo na spremembe v ocenjevanjih znanstvenih rezultatov?
3. Kateri so ključni problemi zanesljivosti in veljavnosti kvantitativnih in kvalitativnih ocenjevanj v znanosti?

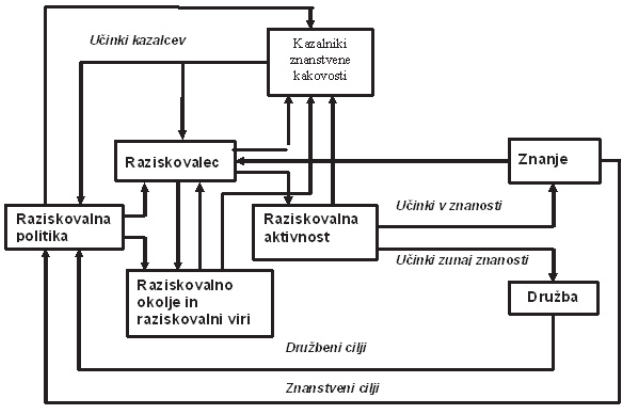
V zadnjem delu bomo predstavili nekaj predlogov za dopolnitev veljavnega sistema vrednotenja rezultatov znanstvenega raziskovanja v Sloveniji, ki ga je v zadnjem času vzpostavila Javna agencija za raziskovalno dejavnost Republike Slovenije (ARRS). Omenjeni evalvacijski sistem se ne glede na vse kritike s strani posameznih lobističnih skupin znanstvenikov odlikuje z vedno večjo stopnjo transparentnosti in objektivnosti. Je pa seveda še vedno daleč od idealnega modela, ki pa ga tako ali tako nikoli ne bo mogoče doseči.

## DRUŽBENO-ZGODOVINSKI KONTEKST KVANTITATIVNEGA IN KVALITATIVNEGA MERJENJA ZNANOSTI

Cilj kvantitativnega in kvalitativnega ocenjevanja v znanosti je ugotavljanje znanstvene kakovosti. Ni enoznačne opredelitve pojma kakovosti v znanosti. Ni nobene aprirorne kategorije znanstvene kakovosti, ki bi bila zunaj časa in prostora. Res je, da različni avtorji, ki se ukvarjajo z vprašanji kvantitativnega in kvalitativnega ocenjevanja v znanosti, uporabljajo različne vrste opredelitev znanstvene kakovosti, vendar gre pri teh opredelitvah vendarle bolj za t. i. operativne oziroma delovne opredelitve. Nizozemski scientometrik Anthony Van Raan, direktor Centra za znanstvene in tehnološke študije pri Univerzi Leiden in urednik obsežne monografije "Handbook of Quantitative Studies of Science and Technology", je podal na primer naslednjo opredelitev kategorije znanstvene kakovosti: "Kakovost je merilo, do katere stopnje skupina ali posamezni znanstveniki prispevajo k napredku našega znanja. Gre za zmožnost reševanja problemov, omogočanje novega pogleda na resničnost ali omogočanje nove tehnologije." (Van Raan, 1996, 398)

Kot smo že dejali, je takšna opredelitev znanstvene kakovosti samo ena izmed možnih opredelitev. Pri vrednotenju znanstvene kakovosti moramo predvsem upoštevati, da je raziskovalna dejavnost družbena dejavnost, ki deluje znotraj občega sistema kulturnih vrednot, ki vladajo v določenem zgodovinskem obdobju in družbenem okolju. Danes si razprav o različnih sistemih ocenjevanja znanstvene kakovosti ne moremo predstavljati kot nečesa, kar je vezano na družbeno izolirano kategorijo znanstvene kakovosti. Švedski teoretik znanosti Sven Hemlin ugotavlja, da "se ocene o tem, kaj je najprimernejši kazalec kakovosti v znanosti, kaj so zaželene lastnosti raziskovalnih dosežkov, in o tem, kakšne so posledice teh ocen za sam sistem znanosti, neprestano spreminjajo" (Hemlin, 1996, 217).

V okviru znanstvene dejavnosti in širšega družbenega okolja, znotraj katerega ta dejavnost poteka, deluje vrsta med sabo povezanih dejavnikov. Ti na takšen ali drugačen način vplivajo na to, kako bodo opredeljeni kazalniki znanstvene kakovosti. V shemi skušamo ponazoriti povezanost teh raznovrstnih dejavnikov.




Slika 1: Kontekstualni dejavniki in njihov vpliv na delovanje evalvacijskih sistemov

Slika 1 kaže, da je kategorija znanstvene kakovosti določena s celo vrsto dejavnikov, ki določajo kontekst znan-

stvenega delovanja znanstvenikov. Ti tvorijo več vzročno povezanih verig, ki vplivajo na oblikovanje kazalnikov znanstvene kakovosti. Na primer, eno verigo sestavljajo povezave med cilji raziskovalno-razvojne politike in postopki ocenjevanja. Druga tipična veriga se nanaša na povezanost med spremembo kriterijev ocenjevanja in raziskovalno dejavnostjo znanstvenikov. Znanstveniki si v okviru svoje raziskovalne dejavnosti vedno prizadevajo, da bi se prilagodili obstoječim kriterijem ocenjevanja kakovosti znanstvenega dela. Končno ne smemo pozabiti tudi na tiste dejavnike, ki so zunaj domene znanosti in ki v zadnjem času posebno močno opredeljujejo naravo postopkov ocenjevanja v sistemu znanosti.

V zvezi z vprašanjem konteksta kazalnikov znanstvene kakovosti je treba izpostaviti še eno vrsto povezave, ki se zdi včasih zunanjim opazovalcem zakrita, ker na analitični ravni vse prevečkrat ločujejo obe ravni. Gre za to, da se neredko želi prikazovati kvantitativne kazalce znanosti kot odslikavo znanstvene dejavnosti, kot povsem neodvisne od subjektivnih presoj in odločitev znanstvenikov. Takšne predstave so napačne. Namreč, vsi podatki, iz katerih izhaja bibliometrija, so na takšen ali drugačen način "poddeterminirani" s kvalitativnimi (subjektivnimi) sodbami znanstvenikov. Kvantitativni kazalci ne morejo uiti svoji kontekstualni pogojenosti. Podobno kot govorimo v hermenevtiki o nujni medsebojni povezanosti med delom in celoto, obstaja tudi v tem primeru neka nujna soodvisnost: ni kvantitativnih meril brez kvalitativnih sodb in obratno: kvalitativne sodbe se ravno tako opirajo na kvantitativne parametre.

Vzemimo preprost, a hkrati najbolj očiten primer: na znanstveno publicistiko (objave, citiranost) posredno ali neposredno vplivajo kvalitativne ekspertne ocene. Za objavljeni članek v ugledni znanstveni reviji samoumevno predpostavljamo, da je šel skozi natančno recenzijsko sito. Posredno je torej tako objava članka v ugledni znanstveni reviji kot tudi njegova morebitna kasnejša citiranost vedno odvisna od predhodne kvalitativne ekspertne ocene oziroma recenzentskih postopkov. Velja seveda tudi obratno: kvalitativne (subjektivne) ekspertne ocene se ravno tako opirajo na različne postopke kvantifikacije. Pomislimo samo na postopek rangiranja, ki mu morajo slediti eksperti oziroma recenzenti, ko za nacionalne raziskovalne agencije ocenjujejo predloge raziskovalnih projektov.

Kvantitativni kazalci torej niso nastali in se razvili neodvisno od kvalitativnih ekspertnih sodb. Res pa je, kar bomo skušali pokazati v nadaljevanju, da je na temelju kritike kvalitativnih ekspertnih ocen nastal močan zagon za nastanek kvantitativnega merjenja znanosti.

## KDAJ IN ZAKAJ SE JE V RAZVITIH ZNANSTVENIH OKOLJIH NAJPREJ POJAVILA ZAHTEVA PO KVANTITATIVNEM VREDNOTENJU ZNANOSTI

Kvalitativne ekspertne ocene oziroma – če uporabimo v strokovni terminologiji bolj razširjen izraz – kolegialne ocene (angl. *peer review*) so še vedno najbolj pogosta oblika ocenjevanja znanstvenih dosežkov v skupnosti znanstvenikov. Znanost je vedno temeljila na komunikaciji med t. i. profesionalnimi kolegi. Sociologi znanosti so takšen tip komunikacije označevali na različne načine. Robert Merton ga je imenoval prestižna hierarhija (angl. *prestige hierarchy*) (Merton, 1973, 36), Diana Crane nevidno omrežje (angl. *invisible colleges*) (Crane, 1969, 5). Takšen tip komunikacije in z njim povezani postopki znanstvenega ocenjevanja imajo daljšo zgodovinsko tradicijo, saj so raznovrstni mehanizmi kolegialne kontrole v znanosti nastali že s pojavom moderne znanosti konec sedemnajstega stoletja, torej takrat, ko začnejo delovati prve akademije znanosti (Royal Society v Angliji, Academie des Sciences v Franciji), čeprav v tistem času še ne moremo govoriti o zelo formaliziranih tipih znanstvenih komunikacij (na primer o široki razvejanosti znanstvenih revij itd.). Z vidika napredka znanosti kot delnega družbenega sistema, ki je v naraščajočih procesih funkcionalne družbene diferenciacije sledil lastnim pravilom delovanja, je bila vzpostavitev in kontinuirano delovanje mehanizma profesionalne kolegialne kontrole v znanosti izredno pomembna.

1. Omenjeni mehanizem je omogočal znanstvenikom dostop do raziskav oziroma objav, s tem pa – v sferi akademske znanosti – pridobivanje znanstvenih nagrad in znanstvenega ugleda. Sistem nagrajevanja in pridobivanja znanstvenega ugleda je eden glavnih motivacijskih dejavnikov delovanja družbenega sistema znanosti skozi ves njegov zgodovinski razvoj.
2. Z vidika širše družbene vloge znanosti je sistem kolegialnega ocenjevanja in kontrole kakovosti v znanosti vedno bil sinonim za avtonomijo in svobodo znanstvenega raziskovanja. S to predstavo je močno povezana predpostavka, da profesionalne vrednote, ki jih člani znanstvene skupnosti usvojijo, usmerja njihovo obnašanje v različnih vlogah znotraj znanstvene skupnosti. Specifični profesionalni etos znanosti, ki ga po ameriškem sociologu znanosti Robertu Mertonu utemeljujejo vrednote univerzalnosti, javnosti, organiziranega skepticizma, nepristranosti, skromnosti in originalnosti, vplivajo na objektivnost kolegialnih ocen v znanosti.[1]

Kvantitativna merjenja kakovosti v znanosti so rezultat novejših znanstvenih dogajanj. Pravi pomen začnejo pridobivati šele v sedemdesetih letih 20. stoletja. Spodbuda

za uporabo kvantitativnih kazalcev merjenja znanstvene kakovosti je prihajala iz dveh strani. Na tako imenovani uporabniški strani so zahtevo po "objektivnih podatkih" začele forsirati nacionalne raziskovalno-razvojne politike. To se je seveda zgodilo najprej v državah z razvitimi znanstvenimi sistemi (Van Raan, 1988; Weingart, 2001). Ključni akterji raziskovalno-razvojne politike so začeli postavljati ostrejše zahteve po uporabnosti in večji družbeni odgovornosti. Ko se je državni proračun za znanost povečeval s 5-odstotne letne stopnje rasti na 10-odstotno, politika ni izražala večjih pomislekov do takrat prevladujočih kvalitativnih ekspertnih ocen. Nadzor znanstvene kakovosti je bil predvsem stvar zaupanja: v znanstveno kompetenco, v motivacijo raziskovalcev, da skrbijo za kakovost svojih rezultatov, v institucije akademske znanosti, ki so v okviru svoje tradicije vzpostavile celo vrsto implicitnih in eksplicitnih nadzornih mehanizmov kakovosti raziskovalnih rezultatov. Zadeve v znanstveno in tehnološko razvitih državah so se začele spreminjati, ko je začela javna politika postavljati znanosti večje zahteve.

Spodbude za večjo rabo kvantitativnih kazalcev znanosti so prihajale tudi od same bibliometrične stroke. Ko govorimo o razvoju in uporabi kvantitativnih kazalcev znanosti v drugi polovici 20. stoletja, moramo poudariti, da se je njihova uporaba najprej začela v akademskih krogih znanosti. Scientometrija kot posebna znanstvena disciplina se je najprej razvila pod okriljem sociologije znanosti. Njene kvantitativne analize so šele kasneje postale priročno sredstvo nacionalnih raziskovalnih in razvojnih politik. Lep primer uporabe kvantitativnih podatkov za sociološko analizo razvoja znanosti v tistem času je še danes znana Priceova logistična krivulja rasti znanosti (Price, 1963 in1975). Osnovna ideja Priceovega statističnega modela rasti znanosti je bila naslednja: statistično preučevanje znanstvenih publikacij, znanstvenih kadrov, sredstev za znanost in še nekaterih drugih kazalcev nas vodi k sklepu, da je moderna znanost od svojega nastanka naprej (sredina 17. stoletja) približno tristo let naraščala eksponentno. V drugi polovici dvajsetega stoletja naj bi prišlo do linearne, torej bolj umerjene rasti znanosti in – kar je v Priceovem statističnem modelu rasti znanosti za večino teoretikov znanosti najbolj sporno – tej rasti naj bi se obetalo, da doseže stanje nasičenosti. Priceova ocena dolgoročne krivulje kvantitativne rasti znanosti je ostala do danes predmet kontroverznih ocen in razprav.

Končno ne smemo pozabiti še na en dejavnik v sedemdesetih letih prejšnjega stoletja, ki je spodbujal večjo uporabo kvantitativnega merjenja znanosti. Vedno bolj intenzivni procesi diferenciacije in specializacije v znanosti ob hkratno vedno večji tekmovalnosti za pridobitev znanstvenega ugleda so v nekaterih primerih privedli tudi do negativnih pojavov. V znanstvenih in tudi zunajznanstvenih krogih so začeli v večji meri opozarjati na goljufije in prevare v znanosti. To je bil čas, ko se je v javnosti nasploh povečal dvom v objektivnost in nevtralnost kvalitativnih kolegialnih ocen v znanosti. Naraščajočo tekmovalnost na področju znanosti, ki sama po sebi seveda ni sporna, najbolje označuje krilatica "publiciraj ali pogini" (angl. *publish or perish*).

V celoti gledano se je ravno v sedemdesetih letih prejšnjega stoletja pojavil pretirani optimizem v moč kvantitativnega merjenja znanosti, tako med strokovnjaki na področju bibliometrije kot tudi med glavnimi akterji raziskovalno-razvojne politike. Derek de Solla Price je v tistem času celo izražal prepričanje, da bodo kvantitativna merjenja lahko zelo kmalu v nadomestila kvalitativne ekspertne ocene v znanosti. To se seveda ni nikoli zgodilo. Je pa res, kot ugotavlja nemški teoretik znanosti Peter Weingart, da je redkokateri instrument znanstvene politike imel toliko povratnih učinkov na posamezne zakonitosti delovanja znanstvenih skupnosti kot ravno uporaba kvantitativnih merjenj kakovosti v znanosti (Weingart, 2005).

## STANJE NA PODROČJU KVANTITATIVNEGA IN KVALITATIVNEGA OCENJEVANJA V ZNANOSTI DANES IN IZZIVI ZA PRIHODNOST

Kljub nekaterim optimističnim pričakovanjem glede rabe kvantitativnih znanstvenih kazalcev za namene znanstvene politike, ki naj bi odpravili vse druge oblike ocenjevanja, kvalitativne ekspertne ocene za namene raziskovalno-razvojne politike še zdaleč niso izgubile veljave. Nasprotno, uveljavilo se je prepričanje, da je za merjenje kakovosti v znanosti treba uporabljati kombinirani pristop. Kombinacija ekspertnih kvalitativnih in bibliometričnih kvantitativnih postopkov ocenjevanja je še vedno najbolj primerna. Tej poti sledi tudi veljavni evalvacijski sistem znanosti, ki ga je vzpostavila ARRS.

Če pogledamo situacijo danes v nekaterih znanstveno in tehnološko najbolj razvitih državah, potem bomo glede vprašanja večje ali manjše rabe kvantitativnega ocenjevanja v znanosti naleteli na velike razlike. Ne sledijo vse nacionalne raziskovalno-razvojne politike enakim modelom ocenjevanja. Pogosto se razlike pojavljajo že znotraj posameznih držav. Na primer, večina vladnih agencij v ZDA, ki financirajo znanost, v postopkih ex ante ali ex post ocenjevanja znanstvene uspešnosti raziskovalcev in raziskovalnih skupin še vedno ne uporablja zelo na široko bibliometričnih kazalcev znanosti (Roessner, 2002). To velja še posebej, če primerjamo situacijo v ZDA s situacijo v Evropi (Hicks idr., 2004). Tudi v državah EU smo priča različnim pristopom k ocenjevanju rezultatov znanstvenega raziskovanja. Po ocenah različnih strokovnjakov je uporaba bibliometrije, in to v njeni najbolj ekstremni

obliki, za ocenjevanje rezultatov znanosti na univerzah in inštitutih, ki se ukvarjajo s temeljno znanostjo, posebno razširjena na Finskem (Adam, 2002; Weingart, 2005). Po navedbah Davida Adama, ki se sklicuje na izjave Karija Raivioja, nekdanjega rektorja Univerze v Helsinkih, je Finska edina država, v kateri je bil faktor vpliva znanstvene revije uzakonjen. Raivio je med drugim naredil izračun, po katerem naj bi v letu 2000 vsak posamezni članek finskega avtorja, ki je bil objavljen v znanstveni reviji s faktorjem vpliva večjim od 2, pripomogel, da je vsaka znanstvena institucija, katere član je avtor objave, financirana z dodatnimi 7.000 dolarji iz državnega proračuna (Adam, 2002, 727).

Finska je primer države, v kateri se ne ozirajo na velikost znanstvene skupnosti kot morebitnega omejujočega dejavnika pri uporabi podatkov za namene ocenjevanja kakovosti v znanosti, ki jih producira korporacija Thomson (število citatov, razvrstitev revij s faktorjem vpliva itd.). V zvezi s tem so zanimivi nekateri podatki, ki jih v eni novejših bibliometričnih analiz navaja Nancy K. Bayers. Po tej njeni analizi iz leta 2005 so države EU, vzete skupaj, prehitele ZDA, kar zadeva število člankov, indeksiranih v Web of Science. Od 3,6 milijona člankov, ki so bili indeksirani v obdobju 2000–2005, jih je 37 odstotkov pripadlo državam EU in 34 odstotkov ZDA. Na tretjem mestu je Japonska z 9,6 odstotka. Glede na posamezne države članice EU se je v letu 2005 na vrhu seznama z največ indeksiranimi članki v Web of Science znašlo veliko majhnih oziroma srednje velikih držav. Velika Britanija z 8,8 odstotka, Nemčija s 6,3 odstotka, Francija s 4,2 odstotka, Italija s 3 odstotki in Španija s 2,6 odstotka, sledijo Švedska (1,9 odstotka), Švica (1,4 odstotka) itd. (Bayers, 2005, 155)

Po mnenju številnih teoretikov znanosti smo v zadnjem poldrugem desetletju priča radikalnim spremembam v družbenih načinih produkcije znanstvenega vedenja. Te spremembe gredo zlasti v smeri večjega povezovanja še nedavno precej zaprtega sistema znanosti z njegovim družbeno-ekonomskim okoljem. Ta odprtost navzven povratno vpliva na notranje razvojne zakonitosti znanosti. Tudi glede na svoj disciplinarni in siceršnji teoretski pedigre posamezni avtorji spremembe v družbeni in epistemološki strukturi znanosti opisujejo z različnimi oznakami: eni govorijo o prehodu iz klasičnega načina produkcije znanstvenega vedenja (Mode 1) v nov način produkcije znanstvenega vedenja (Mode 2) (gl. Nowotny, Scott in Gibbons, 2001), drugi spet govorijo o pojavu trojne spirale, kjer prihaja do vzpostavitve močnih posredniških struktur med tremi družbenimi podsistemi (znanostjo, ekonomijo, politiko) (Etzkowitz, Leydesdorff, 2001). Ponekod se je uveljavila tudi oznaka "postakademska znanost" (Ziman, 2000). Ne glede na različno poimenovanje teh znanstvenih sprememb, ki smo jim priča na začetku 21. stoletja, je zanje značilna predvsem vedno večja usmerjenost v t. i. kontekst aplikacije znanstvenega vedenja (kontekst aplikacije je več kot samo enostaven proces komercializacije rezultatov znanstvenega raziskovanja), naraščajoča transdisciplinarnost znanstvenega vedenja (ne samo interdisciplinarnost), nastop novih (heterarhičnih) oblik organiziranosti znanosti, ki presegajo klasične hierarhične odnose v sistemu znanosti, ter, nenazadnje, zahteva po večji družbeni odgovornosti znanstvenikov za rezultate svojega znanstvenega dela.

V povezavi s predhodnimi spremembami naj bi se dogajal še en ključni premik. Šlo naj bi za spremembo v postopkih znanstvene evalvacije. Po mnenju nekaterih analitikov to spremembo lepo označuje prehod od "kontrole kakovosti" (angl. *quality control*) k "*nadzorovanju kakovosti*" (angl. *quality monitoring*) (Hemlin, 2006, 181). Pod temi sicer precej ohlapnimi oznakami naj bi se nezadržno dogajale spremembe, ki so vezane na:

1. večjo vlogo zunanjih družbenih akterjev v postopkih znanstvenega ocenjevanja;[2]
2. na ocenjevanja, ki se izvajajo na višjem nivoju "agregiranosti" (premik od ocenjevanja posameznih raziskovalcev k ocenjevanju večjih projektnih skupin raziskovalcev in znanstvenih institucij, vključno z ocenjevanjem raziskovalnih omrežij);
3. na ocenjevanja v vseh fazah znanstvenega raziskovanja, ne samo v fazi, ko so rezultati znanstvenega dela že doseženi;
4. na ocenjevanja, ki vključujejo širši krog ekspertov kot samo tiste eksperte, ki delujejo v ožjem disciplinarnem okviru.

Širitev "baze" ekspertov z vidika preseganja ozkih disciplinarnih meja ne škoduje nadaljnjemu razvoju znanosti, še posebej zato ne, ker je vloga ekspertnih ocen velika v "ex ante" in ne v "ex post" fazah vrednotenja znanosti. Razen tega se tudi na področju naravoslovnih in tehničnih znanosti monoparadigmatski model razvoja znanosti umika multiparadigmatskemu razvojnemu modelu. Ta zadnji izhaja iz predpostavke o obstoju množice lokalnih, med seboj tekmujočih in hitro spreminjajočih se paradigem, tako da ni več mogoče izhajati iz sistemov ocenjevanja znanosti, ki so ukrojeni po vzoru klasičnih znanstvenih disciplin.

V različnih delih Evrope so v zadnjem času zrasli številni centri odličnosti v znanosti. Za te centre odličnosti je značilno, da se njihov celotni akademski raziskovalni potencial usmerja v reševanje praktičnih problemov. Ti centri odličnosti v znanosti se ocenjujejo na temelju kriterijev kakovosti in družbene relevantnosti. Tradicionalna delitev na temeljno in uporabno znanost, iz katere navsezadnje izhaja celo OECD-jeva klasifikacija

znanstvenih ved (npr. Priročnik Frascati), izgublja pomen. Meje med bazično in uporabno znanostjo postajajo vedno bolj fleksibilne oziroma se ukinjajo. Pomislimo samo na najbolj propulzivna znanstvena področja danes, kot so na primer bioznanost, nanoznanost, informacijske in komunikacijske znanosti, bioinformatika itd. Na primer, za epistemološki in družbeni status sodobnih biotehnoloških raziskovanj je bolj kot iskanje njenih disciplinarnih izvorov in povezav pomembna njihova izredna aplikativna in transdisciplinarna naravnanost. Revolucionarna znanstvena odkritja v biotehnologiji se uporabljajo v zdravstvu, farmaciji, kmetijstvu in pridelavi hrane, pri varstvu okolja in še v številnih drugih industrijskih sektorjih. Hkrati njen transdisciplinarni značaj vodi k nastanku novih in izredno dinamično razvijajočih se raziskovalnih "niš", kot so genomika, bioinformatika itd. Z nastankom teh novih raziskovalnih niš je povezana vrsta aplikacij: genski testi, reprodukcija človeških organov, organskih tkiv itd.

Glede na brisanje meja med različnimi tipi znanstvenega vedenja (temeljno, uporabno, razvojno) smo v zadnjem času tudi na področju uporabe bibliometričnih kazalcev priča nekaterim zanimivim spremembam. Ti se vedno bolj spreminjajo tudi v kazalce vrednotenja praktične (ekonomske) vrednosti znanosti. Ameriški scientometriki v eni svojih analiz ugotovljajo, da so članki iz revij "Nature" in "Science" najpogosteje citirani viri v patentih na področju bioznanosti in nanoznanosti (Hicks idr., 2004). Bibliometrija, ki se ukvarja z vprašanjem medsebojne citiranosti znanstvenih člankov in znanstvenih patentov, postaja za nekatere agencije, ki financirajo znanost, tako vedno bolj zanimiva.

## PROBLEM ZANESLJIVOSTI IN VELJAVNOSTI BIBLIOMETRIČNIH KAZALCEV

Publicistična dejavnost je še vedno najpomembnejši element znanstvenega komuniciranja, komuniciranje pa je bistvo znanosti. Informacija je osnovna sestavina znanosti in njen glavni proizvod. Po mnenju ekonomistov, ki so se zelo natančno ukvarjali z družbeno funkcijo znanstvenih informacij, kot na primer Fritz Machlup, je pojem družbene produktivnosti v največji meri odvisen od takšnih informacij, ki prispevajo k izboljšanju kakovosti rezultatov, odpravi podvojitev in napak. Publicistika skrbi za ohranjanje in kontinuirano rast svetovne zakladnice znanstvenega vedenja. Širjenje lastnih in spremljanje oziroma priznavanje tujih znanstvenih dosežkov je temeljna predpostavka vsakega znanstvenega dela, kar izraža znani Newtonov aforizem: če sem videl dlje, sem zato, ker sem stal na ramenih velikanov.

Čeprav različne sociološke analize znanosti razkrivajo v zadnjih nekaj desetletjih povečano vlogo neformalnih oblik znanstvenega komuniciranja, te niso bistveno vplivale na mehanizme znanstvenega ocenjevanja. Z vidika bibliometričnih kazalcev je v delitvi med formalnim in neformalnim tipom znanstvenega komuniciranja morda še najbolj zanimivo vprašanje, kako ovrednotiti t. i. "sivo literaturo". Zanjo velja, da je nekje "vmes". Na eni strani ni javno dostopna, na drugi strani pa presega zgolj ad hoc vzpostavljena osebna omrežja znanstvenikov.

Bibliometrični kazalci dobijo pravi pomen oziroma smisel (niso samo podatki) šele z jasno teoretsko definicijo pojava (pojma oziroma konstrukta), ki ga merimo oziroma želimo meriti. Nasploh velja, da kolikor želimo ustrezno uporabljati katero koli vrsto kvantitativnega kazalca znanosti za merjenje kakovosti v znanosti, naj gre za "input" ali "output" kazalce znanosti, potem je ključnega pomena, da imamo jasna konceptualna izhodišča, kaj je cilj naših meritev. Šele glede na ta osnovni cilj oziroma kriterij lahko potem uporabljamo ustrezne metode oziroma postopke merjenja. Skratka, operacionalizacija kvantitativnih kazalcev znanosti je odvisna od jasnih teoretskih izhodišč.

Najbrž ni naključje, da je ravno Eugen Garfield, ustanovitelj indeksa znanstvenih citatov, ob številnih priložnostih poudarjal, da je glavni pogoj za učinkovito in nepristransko uporabo znanstvenih citatov, objav v znanstvenih revijah z visokim faktorjem vpliva itd., njihova skrbna teoretska interpretacija (Garfield, 1979). Gole številke še niso podatki in podatki še ne kazalci. Tu bi lahko uporabili naslednjo analogijo iz medicine: izmerjeni človekov pritisk je podatek brez pomena, dokler nekdo, ki v zadostni meri pozna fiziologijo, ne spozna tega podatka kot pokazatelja človekovega zdravja. Za splošnimi definicijami (npr. kaj je kakovost, kaj je produktivnost, kaj je vpliv, kaj je odzivnost v znanosti itd.) se skrivajo različni teoretski pristopi oziroma razumevanja produkcije in rezultatov znanosti.

Vzemimo primer uporabe indeksa znanstvenih citatov in faktorjev vpliva znanstvenih revij, ki so danes predmet kontroverznih razprav v znanstveni javnosti. Najbrž ni treba posebej poudarjati, da je od zornega kota interpretacije odvisno, kako bomo presojali domnevne pristranosti oziroma napake pri indeksu znanstvenih citatov kot meri kakovosti znanosti. Eric Archambault s soavtorji opozarja v eni svojih novejših bibliometričnih analiz, kjer se sicer ukvarja predvsem z vprašanjem jezikovne pristranosti publikacij, ki so indeksirane v SCI-ju in SSCI-ju, da je največ problemov z veljavnostjo podatkov korporacije Thomson povezanih še z naslednjimi težavami: obseg vključenih znanstvenih revij, izključitvijo monografij in še nekaterih drugih vrst znanstvenih publikacij, disciplinarno klasifikacijo znanstvenih revij, nezmožnostjo zaznavanja različno vloženega dela med avtorji indeksi-

ranih člankov, negativni in napačni citati, kartelne oblike citiranja, izbrisani citati, sekundarni citati, itd. (Archambault idr., 2006, 332). Do podobnih sklepov, ki so temeljili na vsebinski analizi stališč uglednih strokovnjakov s področja bibliometrije, smo pred časom prišli tudi sami (Mali in Jug, 1995).

Kriterij veljavnosti znanstvenim citatom kot meri kakovosti v znanosti daje šele ustrezna teoretska interpretacija njegove kognitivne in socialne funkcije v sistemu znanstvenega komuniciranja. Že omenjeni nizozemski teoretik znanosti Loet Leydesdorff, ki se ne ukvarja samo z vprašanji posredniških znanstvenih struktur in trojne spirale, temveč tudi z kompleksnimi teoretskimi in empiričnimi analizami citiranj v znanosti, še zlasti opozarja na ta multidimenzionalni značaj citiranj v modernem znanstvenem komuniciranju. Ali kot pravi sam: "Število citatov kot kazalec vpliva citiranega avtorja omogoča izpeljavo iz kognitivne uporabe citatov v besedilu k delovanju družbenega sistema nagrad v znanstveni skupnosti. Poleg tega je v areni znanstvene politike argument citata pomagal znanstvenikom do večjega zavedanja družbenega in komunikacijskega konteksta njegovih trditev in tako pripomogel k legitimnosti sociooloških analiz znanosti" (Leydesdorff, 1998, 19).

Tudi ko gre za vprašanje zanesljivosti bibliometričnih kazalcev, se je treba zavedati, da ni možna popolna (absolutna) standardizacija podatkov. Bibliometrični kazalci niso enake vrste kot strogo kontrolirane spremenljivke v laboratorijskih eksperimentalnih preučevanjih. Opirajo se na podatkovne baze, katerih prvotni namen ni bil podpora evalvacijskim postopkom. To velja navsezadnje tudi za indekse citatov. Prvotni namen teh baz ni bila podpora raziskovalnim politikam pri vrednotenju znanstvene kakovosti. Prvotni namen je bil iskanje znanstvene literature. Drugi namen je bil vzpostavitev baze podatkov za kvantitativne sociološke in zgodovinske analize razvoja znanosti. Nekatere scientometrične raziskave, ki so bile opravljene za posamezna področja znanosti, so še konec osemdesetih let 20. stoletja odkrivala visoko stopnjo nezanesljivosti podatkovnih baz. Znana je statistična analiza R.E. Rica in soavtorjev, v okviru katere so v bazah Social Science Citation Index in Journal Citation Reports odkrili 25-odstotno mersko napako pri citiranju, ki se nanašajo na revije za informacijsko in bibliotekarsko vedo. Šlo je za tip napak, ki so nastajale kot posledica opustitve predhodnega naslova znanstvene revije, združevanje podatkov o citatih in citiranosti v nerazpoznavno celoto, uporaba nestandardiziranih okrajšav naslovov znanstvenih revij, manjkajoči citati od "polno pokritih" znanstvenih revij, nepravilno in nepopolno razvrščanje revij v specifične znanstvene discipline. (Rice et al., 1989)

Ne smemo pozabiti, da so vodilni predstavniki ISI-ja, še najbolj njen ustanovitelj Eugen Garfield, kar naprej opozarjali, kako pomemben je ustrezen interpretacijski okvir in uporaba citatov kot mera kakovosti v znanosti. V zadnjem času prihajajo na površje strahovi, da ISI postopoma izgublja to kritično avtorefleksijo glede dosega in omejitev njegovih podatkovnih baz za namene ocenjevanja znanstvene kakovosti in vpliva v znanosti (Adam, 2002; Weingart, 2005). Potem ko je Eugen Garfiled, ustanovitelj ISI-ja, omenjeno informacijsko servisno podjetje prodal korporaciji Thomson, je ta družba, ki sedaj poseduje ISI, sicer razvila vrsto programov, ki naj bi pomagali pri pregledovanju podatkovnih baz (npr. Essential Science Indicators). Hkrati pa je zaostrila pogoje uporabe njenih podatkovnih baz za neodvisne in kritične raziskovalce, ki se ukvarjajo z vprašanji scientometrije za namene znanstvene politike. Thomson Scientific se je v zadnjem času do te mere podredil komercialnim interesom, zato številni bibliometriki izražajo strahove, da je zaradi visokih stroškov uporabe njihovih podatkovnih baz in drugih omejitev, ki jih postavlja Thomson Corporation, vedno težje opravljati neodvisne in tudi kritične raziskave o citiranosti v znanosti kot meri znanstvene kakovosti.[3]

Delni dokaz, da moramo biti glede zanesljivosti podatkov, ki jih producira Thomson Scientific, vseskozi zelo pozorni, saj ne gre za bazo podatkov brez napak, so navsezadnje tudi izkušnje IZUM-a. Slovenski raziskovalci imajo trenutno v vzajemni bazi podatkov COBIB.SI več kot 32.000 bibliografskih enot, objavljenih v revijah, ki jih indeksira Thomson Scientific. Strokovnjaki IZUM-a so do konca leta 2006 v Web of Science našli le nekaj nad 75 odstotkov bibliografskih enot, pod katere so podpisani tudi raziskovalci iz Slovenije. Neskladje deloma izvira iz t. i. tehničnih napak v bazah (napačno zapisani priimki, pomanjkljivi podatki o institucionalni pripadnosti znanstvenikov, itd.).

Če skušamo podati nekaj splošnih sklepov o problemu zanesljivosti in veljavnosti bibliometričnih kazalcev, potem se ne moremo izogniti omembi naslednjih opozoril:

1. Noben kvantitativni kazalec v evalvacijskih postopkih znanosti ne more v celoti zaobseči pojava, ki ga želi izmeriti. Bibliometrija nima statusa eksperimentalnih laboratorijskih raziskav. Zato bi zanjo lahko rekli, da se giblje bolj v okviru hevristike kot natančno operacionaliziranih definicij pojavov, ki jih želi meriti. Ravno iz tega razloga je previdnost pri uporabi bibliometrije kot "orodja" za raziskovalno-razvojno politiko še toliko bolj na mestu.
2. Bibliometrija, ki se prvenstveno ukvarja z merjenjem znanstvenega outputa, se praviloma srečuje z bolj kompleksno situacijo, kot je pri merjenju inputa v znanost. V nasprotju z golimi številskimi vrednostmi

se indikatorji znanstvenega "outputa" nanašajo na nekaj, kar ni število kot tako, temveč nekaj, kar je definirano preko predhodnih konceptov: na primer na kategorijo znanstvene kakovosti, znanstvenega vpliva, znanstvene učinkovitosti, itd. Pri kazalcu znanstvenega osebja je v postopku operacionalizacije edini problem uporaba ustrezne klasifikacije. Pri bibliometriji kot meri znanstvenega "outputa" se srečujemo s precej bolj kompleksnimi konstrukti.

3. V bibliometriji pogosto prihaja do naslednjega paradoksa: pretirana težnja po visoki zanesljivosti (ki jo omogoča – ali vsaj omogoča preverjati – prav uporaba kvantitativnih metod zbiranja in analize podatkov) vodi v manjšo veljavnost raziskovanja. Četudi veljavnost načeloma predpostavlja zanesljivost.

## PROBLEM ZANESLJIVOSTI IN VELJAVNOSTI KVALITATIVNIH EKSPERTNIH OCEN V ZNANOSTI

Kvalitativne ekspertne ocene oziroma kolegialna kontrola ostaja centralni element v organizaciji in delovanju vsake znanstvene skupnosti, pa naj gre za zagotavljanje resursov za znanost, ohranjanje standardov znanstvenega raziskovanja, spodbujanje tekmovalnosti v socialnem sistemu znanosti z namenom doseči večjo učinkovitost in kakovost znanstvenega dela. Kvalitativna ekspertna ocenjevanja v znanosti segajo na različna področja znanstvene dejavnosti. Omenjeni mehanizem kontrole znanstvene kakovosti ne služi samo kot sredstvo za razdelitev finančnih sredstev, temveč se nanaša na raznovrstne oblike medsebojnega ocenjevanja znanstvenikov, ki so povezani z recenzijskimi, habilitacijskimi in drugimi postopki v znanosti. Z vidika delovanja družbenega sistema znanosti je dobro delovanje ekspertnih ocenjevanj pomembno iz več razlogov. V razmerah delovanja moderne znanosti kvalitativni ekspertni sistem ocenjevanja po eni strani omogoča dostop do dela v okviru različnih raziskovalnih projektov in do objav. Objave so še vedno ena ključnih predpostavk za pridobivanje znanstvenega ugleda. Težnja po pridobivanju znanstvenega ugleda pa je eden glavnih motivacijskih dejavnikov za delovanje družbenega sistema znanosti. Po drugi strani sistem kvalitativnih ekspertnih ocen nastopa kot nekakšen sinonim za avtonomijo in svobodo znanstvenega raziskovanja.

Čeprav so kvalitativne ekspertne ocene konstitutivni element delovanja znanstvenih skupnosti, je še vedno malo empiričnih analiz o mehanizmih njihovega delovanja. To je po svoje presenetljivo, saj lahko pridemo šele po empirični poti, na primer z analizo števila sprejetih in zavrnjenih člankov v postopkih znanstvenega recenziranja ali pa z analizo razlogov za zavrnitev predlogov raziskovalnih projektov, do odgovora na vprašanje o veljavnosti in zanesljivosti kvalitativnih ekspertnih ocen v znanosti.[4]

Namreč, tako kot kvantitativne ocene znanosti, ki se opirajo na bibliometrične kazalce, se tudi kvalitativne ekspertne ocene srečujejo s problemi zanesljivosti in veljavnosti. Zanesljivost in veljavnost ekspertnih ocen sta odvisni od različnih dejavnikov, ki delujejo na mikro- in makroravni in so med seboj povezani. Tudi v primeru kvalitativnih ekspertnih ocen se znajdemo pred vprašanjem, ki smo se ga dotaknili že na začetku in se glasi: Ali je v znanosti mogoče doseči soglasje o kategoriji znanstvene kakovosti?

Dobra stran kvalitativnih ekspertnih ocen je, da znanstveniki, ki se znajdejo v vlogi ocenjevalca, pri presoji znanstvene kakovosti upoštevajo celoten kontekst. Slaba stran je subjektivnost njihovih ocen. To prinaša nevarnost, da zaradi nezadostnih informacij ali preveč poudarjenega "ega" ocenjevalca pride do usodnih napak. Kvantitativne ocene namreč dosežejo visoko stopnjo zanesljivosti z možnostjo ponovitve merjenja (npr. načeloma napako lahko odpravimo s ponovnim štetjem publikacij, citatov itd.), kar seveda ni mogoče tako enostavno doseči pri kvalitativnih ekspertnih ocenah. Seveda imajo kvantitativni kazalci več težav z določitvijo kriterija veljavnosti. Namreč, ali res merijo tisto, kar želijo izmeriti?

Najbolj razvpiti primeri, ki kažejo, kako nezanesljive so lahko kvalitativne ekspertne ocene, so zavrnitve prispevkov znanstvenikov, ki so čez nekaj let dobili Nobelovo nagrado. Eden od primerov, ki je bil razkrit v zadnjem času, je usoda članka "The market for 'lemons': qualitative uncertainty and the market mechanism" Georga Akerlofa, Nobelovega nagrajenca za ekonomijo izpred treh let (Akerlof, 2003). Georg Akerlof z njim v 70. letih 20. stoletja kar pri treh različnih ekonomskih znanstvenih revijah ni uspel prodreti skozi njihova recenzijska sita. Uspel je šele v četrtem poskusu, in sicer v reviji "Quarterly Journal of Economics". Na temelju rezultatov, ki jih je predstavil v tem članku, je 34 let kasneje dobil Nobelovo nagrado. Primer Georga Akerlofa ni osamljen, saj je Juan Miguel Campanario v svojih študijah predstavil podobne slabe izkušnje sedmih Nobelovih nagrajencev z recenzentskimi postopki njihovih prispevkov, ki so jim kasneje, ko so se prebili čez vse prepreke in bili končno objavljeni, prinesli najvišjo znanstveno slavo in ugled (Campanario, 1998). Ti primeri so najbolj prepričljiv dokaz, kako se v znanosti bije stalen boj med tradicijo in inovacijo.

Kvalitativna ocena naj bi se štela kot zanesljiva, če se vsaj dva ocenjevalca strinjata o predloženem predlogu projekta ali predloženem rokopisu. Vendar raziskave, ki so se ukvarjale s to problematiko, opozarjajo, da visoka stopnja soglasja med eksperti sama po sebi dokazuje malo. Dva ocenjevalca lahko prideta do enako napačnih sklepov. Ali kot je to težavo v zvezi z zanesljivostjo

kvalitativnih znanstvenih ocen lepo povzel Stefan Hornbostel: "Ocene, ki v evalvacijskih postopkih temeljijo na soglasju, so lahko kolektivno napačne ali kolektivno pravilne sodbe, lahko pa so tudi povsem različne ocene z enakimi posledicami. Za soglasje v postopkih ekspertnih ocen torej ni nujno, da so tisti, ki nastopajo v vlogi ocenjevalcev, poenoteni glede tega, kaj je z vidika vsebinske presoje relevantno. V strogem 'tehničnem' oziru konsenzualno sprejete ocene, v okviru katerih eksperti iz medsebojno različnih ali celo nasprotujočih si zornih kotov tematizirajo različne vsebinske aspekte 'predmeta' ocenjevanja, nikakor niso zanesljive ocene". (Hornbostel, 1997, 205.)

Res pa je, da se lahko takoj vprašamo, ali takšno ozko pojmovanje zanesljivosti kvalitativnega ekspertnega ocenjevanja, kot ga v zgornjem citatu navaja Hornbostel, ne zgreši bistva problema. Določena stopnja odsotnosti ozko pojmovane zanesljivosti v ekspertnih ocenah je celo nujna, kolikor teh ocen ne razumemo v smislu mehaničnega algoritma, temveč v smislu fleksibilnih presoj, ki v dani situaciji upoštevajo celotni kontekst.

Problem veljavnosti in zanesljivosti glede na kvalitativni oziroma kvantitativni tip ocenjevanj v znanosti je opredelil Francis Narin na naslednji način: evalvacijske analize, ki so najbolj relevantne za merjenje dejanske stopnje prispevka k znanosti, izkazujejo največjo stopnjo nezanesljivosti glede na objektivnost, tiste, ki so najbolj objektivne, pa največjo stopnjo nezanesljivosti glede na relevantnost (Narin, 1978, 36).

Ne glede na vse morebitne slabosti, ki jih izkazujejo kvalitativne ekspertne ocene in smo jih tu nakazali samo nekaj, je treba znova poudariti, da so bile skozi ves razvoj moderne znanosti izredno pomembne. Prispevale so k ohranjanju kvalitativnih standardov raziskovanja. Razlog je nenazadnje v tem, da gre v sistemu znanosti bolj kot v katerem koli drugem družbenem podsistemu za specifično izmenjavo "znanstvenih dobrin". V tej specifični izmenjavi se vsak posamezen znanstvenik lahko hipotetično znajde enkrat v tej, drugič v drugi vlogi. Se pravi, da je enkrat lahko v vlogi tistega, katerega rezultati dela se ocenjujejo, drugič spet v vlogi tistega, ki ocenjuje rezultate dela svojih profesionalnih kolegov.

Čeprav se na tem mestu ne bomo obširno ukvarjali s tem vprašanjem, je treba na koncu reči, da se zlasti ex ante ekspertne ocene, ki so povezane z razdelitvijo sredstev za raziskovanje, torej se nanašajo na ocene predlogov raziskovalno-razvojnih projektov, v praksi izvajajo še vedno na zelo različne načine. Razlike v načinih uporabe kvalitativnih ekspertnih ocen kot pomembnih instrumentov raziskovalno-razvojne politike niso samo med posameznimi državami, temveč pogosto tudi med posameznimi znanstvenimi agencijami znotraj istih držav. Ena izmed študij, ki se je ukvarjala z vprašanjem kvalitativnih ekspertnih ocen v različnih raziskovalnih agencijah Evrope in ZDA, je sicer ugotovila, da so vse zavezane nekim osnovnim načelom, na katerih je utemeljena kolegialna kontrola v znanosti, se pa razlikujejo v vrsti praktičnih postopkov (Langfeldt, 2006). Po teh analizah se največje razlike pojavljajo glede števila faz ocenjevanja (ex ante, medium, ex post), glede števila ekspertov, ki so vključeni v evalvacijske panele (število varira od 3 do 30), in glede samih postopkov izbire ekspertov, ki ocenjujejo predloge raziskovalno-razvojnih projektov.

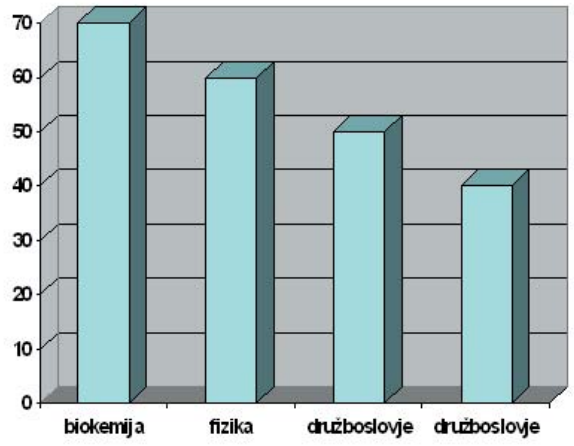
Nekateri predlogi sprememb ekspertnih ocenjevanj gredo predvsem v tej smeri, da bi prispevali k doseganju večje veljavnosti in zanesljivosti kvalitativnih ekspertnih ocen. Omenjeni predlogi se nanašajo predvsem na naslednje dejavnike (tu jih samo naštevamo, ne da bi se lotili njihove podrobne analize): zahteva po vzpostavitvi bolj formaliziranih pravil v postopkih ocenjevanja, ocenjevanje z vključevanjem načela dvojne prikritosti, kjer niti avtor predloga niti recenzent ne poznata drug drugega, pravica avtorja predloga raziskovalno-razvojnega projekta do ugovora na recenzijo in še nekateri drugi predlogi.

## NEKATERE POMANJKLJIVOSTI KVANTITATIVNEGA OCENJEVANJA ZNANSTVENE IN STROKOVNE USPEŠNOSTI RAZISKOVALCEV V SLOVENIJI

V zaključnem delu naše razprave bi radi našteli nekaj pomanjkljivosti, ki spremljajo zdaj veljavni sistem kvantitativnega merjenja uspešnosti dela znanstvenikov v Sloveniji. Omenjeni sistem kvantitativnega merjenja je del celotnega evalvacijskega sistema znanosti ARRS (gl. Pravilnik o kazalcih in merilih znanstvene in strokovne uspešnosti, 2006). Omenjeni sistem ARRS-ja, ki se uporablja v okviru ex ante ocenjevanja (ocena predlogov raziskovalno-razvojnih projektov) namreč temelji na kombinaciji kvalitativnih ekspertnih in kvantitativnih bibliometričnih pristopov. Gre za sistem ocenjevanja, ki je v zadnjem letu s pomočjo COBISS-ovih podatkovnih baz postal izredno izpopolnjen in predvsem vedno bolj transparenten. Seveda pa tako kot noben evalvacijski sistem znanosti v svetu tudi sistem v Sloveniji ni dovršen. Predlogi dopolnitev oziroma sprememb, ki jih podajamo v nadaljevanju, z izjemo problema citatnih oken seveda, se v veliki meri povezani tudi z težavami, s katerimi se strokovni delavci OSIC-ev srečujejo pri preverjanju različnih tipov del v sistemu COBISS. Podajamo nekaj najbolj očitnih pomanjkljivosti:

## Problem citatnih oken

Menimo, da enako dolgo obdobje za ugotavljanje števila citatov v javnih razpisih ARRS za vse vede ni ustrezno. V luči opozoril scientometrikov, ki govorijo o različni hitrosti teoretskega razvoja v različnih vedah, bi bilo najbrž smiselno upoštevati te razlike. Teoretski razvoj vpliva na prakso citiranja v posameznih znanstvenih vedah in disciplinah. Na področju družbenih in humanističnih ved citiranost člankov, ki so vključeni v Web of Science začne naraščati bistveno kasneje kot na področju naravoslovnih ved. Govorimo o različnem razponu citatnih oken (angl. *citation windows*). Izračuni t. i. Priceovega indeksa, ki ga na tem mestu ne bomo natančno predstavljali, so pokazali zanimivo sliko. V biokemiji je 70 odstotkov, v fiziki 60 odstotkov in v družbenih znanosti od 40 do 50 odstotkov citatov mlajših od pet let. Po Glänzelovem in Schöpflinovem izračunu, ki pomeni določeno prireditev "Priceovega indeksa", so razlike glede povprečne starosti uporabljenih virov v člankih, objavljenih v revijah, ki so indeksirane v SCI-ju ali SSCI-ju, naslednje: povprečje za biomedicino je 7–8 let, za sociologijo 12,5 let, za zgodovino in filozofijo znanosti 39 let itd. Ni bistvene razlike med naslednjimi znanstvenimi področji: fizika trdnih snovi, matematika, psihologija, psihiatrija, ekonomija in poslovne vede. Starost citiranih virov na teh področjih je v povprečju od 10 do 11 let in pol.




Slika 1: Citatna okna po posameznih področjih znanosti

## Zapostavljanje nekaterih objav

Pravilnik o kazalcih in merilih znanstvene in strokovne uspešnosti zapostavlja naslednje vrste objav: obsežna poglavja v monografijah, daljše prispevke v enciklopedijah, prevode rokopisov in večino recenziranih objavljenih prispevkov na konferencah.

Pravilnik v 3. točki 5. člena predvideva, da lahko "obsežni znanstveni sestavek ali poglavje (nad 50 strani)" izide samo "v obsežni znanstveni monografiji (2.01A nad 500 strani)", ne pa v monografiji do 500 strani. čeprav je takšnih v petletnem obdobju 2002–2006 kar 85 odstotkov.

Tudi poglavja v enciklopedijah vrednoti slabše kot poglavja v znanstvenih monografijah. Poglavje v enciklopediji prinese avtorju 10 točk, tudi če jo je izdala ugledna mednarodna založba s seznama ARRS. Če bi bilo objavljeno v monografiji, pa bi zanj dobil od 20 do 30 točk. Takšnih poglavij v 24 v tujini objavljenih enciklopedijah, katerih bibliografski zapisi so verificirani, je 44.

Veliko večji problem je neenako obravnavanje recenziranih objavljenih znanstvenih prispevkov na konferencah. Samo prispevki, ki so objavljeni v rednih številkah znanstvenih revij, vključenih v Web of Science, se vrednotijo enako kot izvirni znanstveni članki. Glede na to, da se upoštevajo samo objave recenziranih prispevkov na konferencah, kar pomeni, da se zanje uporablja enak recenzijski postopek kot za izvirne znanstvene članke in samostojne prispevke v monografijah, jih je treba enako obravnavati in enako vrednotiti.

Pravilnik v 2. (predzadnjem odstavku 5. člena določa: "Prevod ali nova izdaja lastnega znanstvenega dela se točkuje s 50 odstotkov točk ustrezne kategorije izvirnika." Beseda prevod tu pomeni prevod monografskega dela, ki je že bilo objavljeno v izvirnem jeziku in je avtor za izvirno objavo v ocenjevalnem obdobju že dobil 100 odstotkov točk. Kadar ne gre za prevod objavljenega dela, ampak za prevod rokopisa, avtor dobi za prvo objavo samo 50 odstotkov točk.

Po Pravilniku se znanstveni članki, objavljeni v revijah, ki niso vključene v mednarodne baze podatkov ali na seznam revij, ki se upoštevajo pri kategorizaciji, ne kategorizirajo. Pravilnik v 1. odstavku 6. člena določa: "Kot bibliografski kazalci strokovne uspešnosti se po tem pravilniku štejejo nekategorizirana znanstvena dela iz 5. člena tega pravilnika", zato dobijo 5 točk namesto 10. Pravilnik torej znanstveni članek, čeprav ima tip dela določen v skladu s Tipologijo dokumentov/del za vodenje bibliografij v sistemu COBISS in ga je verificiral OSIC, proglasi za strokovnega. To je v nasprotju z veljavno tipologijo dokumentov in v nasprotju s smislom verifikacije bibliografskih zapisov.

## SKLEP

V prispevku smo opozorili na nekatera odprta vprašanja in dileme kvantitativnega in kvalitativnega vrednotenja znanstvenih rezultatov v znanosti. Predstavili smo družbene in zgodovinske vidike razvoja različnih sistemov ocenjevanja znanstvene kakovosti. Današnji razvoj znanosti in njen odnos do družbenega okolja bolj kot kadar

koli prej postavlja v ospredje vprašanje kriterijev veljavnosti in zanesljivosti kvantitativnih in kvalitativnih vrednotenj znanosti. Zato smo velik del pozornosti namenili ravno tej problematiki. Vse dosedanje študije dokazujejo, da je uporaba bibliometričnih kazalcev za namene vrednotenja raziskovanja in razvoja nesmiselna ali celo škodljiva za razvoj znanosti, če ni utemeljena v ustrezni teoretski refleksiji o dosegu in omejitvah kvantitativnih merjenj znanosti. Enako seveda velja tudi za kvalitativne ekspertne ocene v znanosti, četudi se vprašanja kriterijev veljavnosti in zanesljivosti v tem primeru zastavljajo na drugačen način. Predvsem pa je treba v vsakem primeru iskati soglasje o tem, s katerimi instrumenti meriti znanost, med vsemi vpletenimi akterji (znanstveniki, akterji znanstvene politike, eksperti na področju bibliometrije itd.). V Sloveniji veljavni evalvacijski sistem znanosti sledi sodobnim trendom razvoja na tem področju. Upošteva se kombinacija kvalitativnega in kvantitativnega pristopa k vrednotenju kakovosti v znanosti. Zaradi dobro urejenih podatkovnih baz (COBISS) je zlasti stopnja veljavnosti in zanesljivosti kvantitativnih ocenjevanj pri nas vedno večja. Se pa pojavljajo tudi nekatere konceptualne in metodološke težave, na katere smo skušali vsaj deloma opozoriti.

## Opombe

1 Načelo univerzalnosti pravi, da ne obstajajo privilegirani viri znanstvenega vedenja, kar z drugimi besedami pomeni, da naj bi se znanstveni rezultati posameznega raziskovalca ocenjevali zgolj glede na interna znanstvena merila, ne glede na njegovo nacionalno, ideološko, rasno, spolno ali katero koli drugo družbeno opredeljenost. Načelo komunalnosti pravi, da so rezultati znanstvenega raziskovanja javni, načelo nepristranosti pravi, da znanost služi sama sebi, kar z drugimi besedami pomeni, da naj bi znanstveniki raziskovali in posredovali rezultate raziskovanja samo z enim namenom, namreč da prispevajo k napredku znanstvenega vedenja. Načelo sistematičnega (organiziranega) skepticizma pravi, da naj ne bi znanstveniki ničesar vzeli v dobri veri (lat. bona fide), temveč vseskozi skrbijo za kritično presojo in nadzor predloženega znanstvenega vedenja. Merton je v svojih kasnejših delih pripisal poseben pomen načelu skromnosti in originalnosti v znanosti (Merton, 1973).

2 Na Nizozemskem so že v začetku devetdesetih let prejšnjega stoletja začeli vključevati vse ključne elemente, ki označujejo premik od "kontrole" kakovosti znanosti k njenemu "monitoringu". Izredno velika teža je bila dana ravno širjenju tradicionalnega sistema kolegialnega ocenjevanja v smeri večjega vključevanja različnih družbenih akterjev v postopke ocenjevanja znanstvenih rezultatov. (Van der Meulen in Rip, 2000.)

3 Tibor Braun, urednik revije Scientometrics, je čas pred nakupom ISI-ja s strani korporacije Thomson označil za "romantično obdobje bibliometrije". Garfieldov močan osebni interes za različna konceptualna vprašanja kvantitativnega merjenja znanosti je vplival na to, da so imeli scientometriki več svobode glede ukvarjanja z osnovnimi podatkovnimi bazami ISI-ja. (gl. Adam, 2002: 729)

4 Presenečenje se seveda ne nanaša toliko na razmere pri nas (empirične sociološke analize o delovanju znanstvenega sistema pri nas so še vedno precej nezaželene in to predvsem med znanstveniki samimi), temveč na situacijo v zadnjih nekaj desetletjih nasploh v razvitem znanstvenem svetu. Torej tam, kjer izhaja množica uglednih mednarodnih znanstvenih revij. Po ugotovitvah Juana Miguela Campanaria naj bi bile pred letom 1975 objavljene samo 3 relevantne empirične študije o delovanju mehanizmov recenziranja v znanstvenih revijah (Campanario, 1998: 182). Razen tega naj bi se empirične raziskave v glavnem osredotočale na vprašanja zanesljivosti in nepristranosti ekspertnih kvalitativnih ocen, premalo pozornosti pa naj bi namenile vprašanju njihove veljavnosti (Daniel, 1993: 71).

## Reference


[1] Adam, David (2002). The counting house. *Nature,* 415 (6873), 726–729.

[2] Akerlof, George (2003). Writing the "Market for lemons'": a personal and interpretative essay. Dostopno na: http://nobelprize.org/economics/articles/akerlof/index.html, zadnji dostop: 12. januar 2007.

[3] Archambault, Eric, Vignola-Gagne Etienne, Grégoire Côté, Larivière, Vincent (2006): Benchmarking scientific output in the social sciences and humanities: the limits of existing databases. *Scientometrics*, 68 (3): 329–342.

[4] Bayers, Nancy K. (2005). Using ISI data in the analysis of German national and institutional research output. *Scientometrics*, 62 (1): 155–163.

[5] Campanario, Juan Miguel (1998). Peer review for journals as it stands today – Part 1. *Science communication*, 19 (3): 181–211.

[6] Campanario, Juan Miguel (1998). Peer review for journals as it stands today – Part 2. *Science communication*, 19 (4): 277–306.

[7] Crane, Diana (1969). Social structure in a group of scientists: a test of the "invisible college" hypothesis. *American sociological review*, 34 (2): 335–352.

[8] Daniel, Hans-Dietrich (1993). Guardians of science: fairness and reliability of peer review. Weinheim & New York: VCH.

[9] Etzkowitz, H., Leydesdorff, L. (2001). Universities and the global knowledge economy: a triple helix of university – industry – government relations. New York: Continuum.

[10] Garfield, Eugen (1979). Citation indexing: its theory and application in science, technology, and humanities. New York: John Wiley & Sons.

[11] Hemlin, Sven (1996). Research on research evaluation. *Social Epistemology*, 10 (2): 209–250.

[12] Hemlin, Sven (2006). The shift in academic quality control. *Science, technology & human value*s, 31 (2): 173–198.

[13] Hicks, Diana, Hiroyuki, Tomizawa, Yoshiko, Saitoh, Shinichi, Kobayashi (2004). Bibliometric techniques in the evaluation of federally funded research in the United States. *Research evaluation*, 13 (2): 78–86.

[14] Hornbostel, Stefan (1997). Wissenschaftsindikatoren: Bewertungen in der Wissenschaft. Opladen: Westdeutscher Verlag.

[15] Mali, Franc, Jug, Janez (1995). Scientometrija in analiza citatov. *Teorija in praksa*, 32 (9/10): 812–820.

[16] Langfeldt, Liv (2006). The policy challenges of peer review: managing bias, conflict of interests and interdisciplinary assessment. *Research evaluation*, 15 (1): 31–41.

[17] Leydesdorff, Loet (1998). Theories of citations? *Scientometrics*, 43 (1): 5–25.

[18] Merton, (1973). Sociology of science. Chicago: Chicago University Press.

[19] Narin, Francis (1978). Objectivity versus relevance in studies of scientific advance. *Scientometrics*, 1 (1): 35–41.

[20 Nowotny, Helga, Scott, Peter in Gibbons Michael (2001). Re-thinking science: knowledge and the public in an age of uncertainty. Oxford: Blackwell.

[21] Pravilnik o kazalcih in merilih znanstvene in strokovne uspešnosti (2006). Dostopno na: http://www.arrs.gov.si/sl/akti/prav-znan-strok-uspesn-okt06.asp, zadnji dostop: 17. januar 2007.

[22] Price, Derek de Solla (1963). Little science, big science. New York: Columbia University Press.

[23] Price, Derek de Solla (1975). Science since Babylon. Haven, CT: Yale University Press.

[24] Rice, R. E., Borgman, C.L., Bednarski, D., Hart, P. (1989). Journal to journal citation data: issues of validity and reliability. *Scientometrics*, 15 (3/4): 257–282.

[25] Roessner, J. D. (2002). Outcome measurement in the USA: state of the art. *Research evaluation*, 11 (2): 85–93.

[26] Van der Meulen, Barend in Rip, Arie (2000). Evaluation of societal quality of public sector research in the Netherlands. *Reserach evaluation*, 8 (1): 11–25.

[27] Van Raan, Anthony (ur.) (1988). Handbook of quantitative studies of science and technology. Amsterdam: Elsevier Science Publishers.

[28] Van Raan, Anthony (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics*, 36 (3): 397–420.

[29] Weingart, Peter (2001). Die Stunde der Wahrheit?: zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft. Weilerwist: Velbrueck Wissenschaft.

[30] Weingart, Peter (2005). Impact of bibliometrics upon the science system: inadvertent consequences?. *Scientometrics*, 62 (1): 117–131.

[31] Ziman, John (2000). Real science: what it is, and what it mean*s.* Cambridge: University Press.