

Hodnocení výzkumu v informatice přístupy a problémy

Jiří Zlatuška

Hodnocení kvality výzkumu

- **Český přístup: Kvantifikace publikací je jediným kritériem, tato kvantifikace lineárně odpovídá množství přidělených financí, metodika hodnocení založena primárně na IF časopisů navzdory faktu, že IF pro podobné posuzování není vhodný a že oborové deformace obsahu WoS podstatně poškozují některé obory, zejména informatiku (s pouze 40% interním pokrytím citací)**
- **Kritika mezinárodního auditu: Extrémní verze nemající v jinde obdoby, chybné principy v hodnocení i financování, chybné spojení hodnocení s financováním, sklon ke kvantitativním optimalizacím, nikoli potřebám výzkumného zaměření**

Chybné použití IF k hodnocení článků i autorů

- IF časopisů nebyl navržen nebo zamýšlen k užití jako měřítko nebo indikace kvality jednotlivých článků nebo autorů
- Výrazně asymetrické rozložení citací způsobuje, že ve většině případů IF nadhodnocuje impakt jednotlivých článků nebo autorů, kteří v časopise publikují
- IF je ovlivněn rozdíly mezi obory, mezioborová srovnání jsou zkreslená
- Sumace nebo kombinace IF to chyby ještě zvýrazňuje.

Kritika mechanických přístupů v informatice – Informatics Europe

1. Informatika je originální disciplína kombinující přírodní vědy a inženýrství. Potřebuje náležitě upravený přístup k hodnocení.
2. Významným rysem publikací v informatice je důležitost výběrových konferencí a knih. Časopisy nemají nutně vyšší prestiž.
3. V hodnocení přínosu jsou artefakty jako software stejně důležité jako publikace.

Kritika mechanických přístupů v informatice – Informatics Europe

4. Pořadí autorů není u infromatických publikací zpravidla důležité a bez dalších informací nemůže být užito k hodnocení.
5. Numerická měření jako např. počty citací nesmějí být nikdy užívána jako jediný nástroj hodnocení. Musí být filtrována lidskou interpretací a doplněna peer review a zhodnocením jiných výstupů, než jen publikací.
6. Počty publikací nejsou adekvátním měřítkem hodnoty výzkumu. Měří produktivitu, nikoli dopad nebo kvalitu.
7. Všechna hodnotící kritéria, zejména kvantitativním, musí být založena na jasných a zveřejněných kritériích.

Kritika mechanických přístupů v informatice – Informatics Europe

8. Numerické indikátory nesmí sloužit pro srovnání mezi disciplínami.
9. Pro hodnocení publikací a citací je ISI Web of Science nedostatečný a nesmí být užíván. Alternativy k němu zahrnují Google Scholar, CiteSeer a (potenciálně) Digitální knihovnu ACM
10. Hodnotící kritéria musí sama procházet pravidelným vyhodnocením a revizemi.

Problémy s neadekvátním pokrytím informatiky v ISI Web of Science

- Niklaus Wirth obdržel Turingovu cenu za report k pascalu z roku 1970, není obsažen
- Knuthova série The Art of Computer Programming má 15 tisíc citací v Google Scholar, ale není indexována, totéž se týká tří Knuthových článků nejcitovanějších podle Google
- Změna klasifikace LNCS (do 2006 počítány jako časopis)

Deformace informatiky komerčních databázích

- Padesát nejcitovanějších prací v informatice podle WoS obsahuje např. práci *Chemometrics in food science* z časopisu *Chemometrics and Intelligent Laboratory*
- „Horké“ články z informatiky ve Scopus z července 2011 obsahují mezi 25 pouze 5 relevantních přehledových článků z oblasti bezdrátových senzorů a bezpečnosti v cloud computing, zbytek je sociologie – tj. drtivá většina jsou aplikace nebo hraniční oblasti informatiky a jiných disciplin
- Turingovy ceny jsou téměř disjunktní s „250 highly cited“ Podle WoS

Deformace obsahu informatiky

"Scientist rankings in computer science"

1. HIGGINS, DG
2. FUCHS, R
3. BLEASBY, AJ
4. BILLETER, M
5. KORADI, R
6. WUTHRICH, K
7. SJOSTRAND, T
8. EVANS, SV
9. WAS, Z
10. SEYMOUR, MH
11. JADACH, S
12. OVERBEEK, R
13. WEBBER, BR
14. ABBIENDI, G
15. KNOWLES, IG
16. ...

Whom do you recognize?

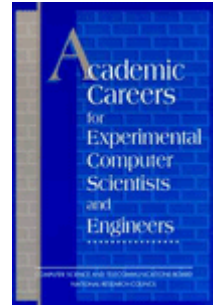
1. Desmond G. Higgins: Empirical estimation of the reliability of ribosomal RNA alignments
2. Rainer Fuchs: Predicting protein function: a versatile tool for the Apple Macintosh
3. Alan J. Bleasby: Information Resources for the Bioinformatician

This is CS in the ISI sense!

Malá reprezentativnost komerčních citačních dat

- Vnitřní pokrytí WoS u fyziky, chemie nebo biomedicinských oborů je 80 % a více, u „informatiky“ pouze 38 %
- Přidání ACM, IEEE-CS a LNCS by pokrytí zvýšilo na pouhých 51 procent (Moed&Visser, 2007)
- WoS je obzvláště špatný pro praktickou informatiku s pokrytím pod 28 % (CWTS 2007)

Akademická kariéra pro experimentální informatiky



- National Research Council, 1995
- Hardwarové nebo softwarové systémy
- Prototypy umožňující studiem složitých konceptů, které nejsou studovatelné čistě teoreticky
- Neuposuzovat experimentální informatiku závislou na instalované infrastruktuře podle měřítek teoretického oboru



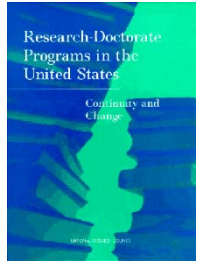
Best Practices Memo

- Computing Research Association, 1999
- Důraz na experimentální a inženýrské přístupy v informatice
- Sociální kontext impaktu prací v informatice
- Posuzování artefaktů jinými odborníky, nikoli jen citační výstupy publikací
- Odmítnutí deformace disciplíny podřízením se publikačním zvyklostem „zavedených“ disciplin

Research Excellence Framework

- Návrh projednávaný ve Velké Británii jako nástupce RAE
- Hodnocení založené na peer review, 12 z 36 panelů bude používat citační data jako podkladů; informatika bere citace i z Google Scholar, matematika nepoužívá citace vůbec
- Individuálně zkoumané publikační výstupy – originalita, preciznost, význam pro obor a širší dopad (akademický i aplikovaný)
- Výzkumné prostředí, doktorandi (důraz na hotové diplomy), výzkumná strategie, indikátory uznání, aplikovaný a interdisciplinární výzkum
- Hodnocení na škále 4*-1* (významný dopad na obor – přínos k oboru bez naděje na větší dopad)

Hodnocení výzkumných doktorských programů v USA



- Poslední v roce 1994, před tím 1982 a v 60. a 50. letech
- 41 skupin oborů, pouze výzkum
- Cca 300 univerzit, pro každý program hodnoceno 200 dotazníků v peer review
- Výsledek prezentován po čtvrtinách
- Publikační výsledky za 10 let
- Bez hodnocení průmyslu, bez impakt faktoru z komerčních zdroj; (na rozdíl od 1982)

Department Evaluation Initiative Informatics Europe



- Záměr vytvoření agentury vhodné specificky pro informatiku
- Vícedimenzionální pohled posuzující kombinaci výzkumných výsledků, managementu výzkumu, výzkumnou politikou a vybavením
- Integrovanou složkou posuzování kvality je výchova PhD studentů a společenské relevance prováděného výzkumu
- Výsledkem analýza indikující slabé a silné stránky, příležitosti a potřebná zlepšení

Department Evaluation Initiative

řídící výbor

- Fausto Giunchiglia - University of Trento, Italy,
- Manuel Hermenegildo - Technical University of Madrid, Spain,
- Jeff Magee - Imperial College, UK,
- John Mylopoulos - University of Toronto, Canada,
- Manfred Nagl - Aachen University, Germany,
- Joseph Sifakis - CNRS, France,
- Letizia Tanca - Politecnico di Milano, Italy,
- Jan van Leeuwen - University of Utrecht, The Netherlands,
- Peter Widmayer - ETH Zurich, Switzerland.