

Informatika a problematické hodnocení inovací





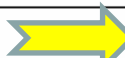

Jiří Zlatuška

Reforma výzkumu, vývoje a inovací v ČR (VaVal)







- „Reformní“ představa: peníze mezi malý počet „poskytovatelů“ se rozdělí podle bodů za výsledky v databázi RIV (metodika připravovaná v době dlouhé cca půl roku; v Británii byl podobný projekt REF nedávno odsunut a pilotní fáze jsou rozfázovány na cca pět let)
- U publikací je rozhodující, jak se na ně dívá Thomson Publishing (= ISI, WoS),
informatika je zde málo zastoupena, dominantní médium publikací jsou sborníky (podrobnou analýzu viz zejména F. Mattern: „Bibliometric Evaluations of Computer Science – Problems and Pitfalls“, ECSS 2008)
- Velký důraz reformy VaVal je na aplikační výstupy, zejména patenty, zde je potíž se softwarovými patenty (spor základní vs. aplikovaný výzkum)

Pro výsledky uplatněné do roku 2007 včetně:

(Převzato z Metodiky 2008)

Druh výsledku		I – obory NRRE ⁰⁾	II – ostatní obory
J _{imp}	článek v impaktovaném časopise		5 + 140 × Faktor ¹⁾
J _{neimp}	článek v recenzovaném časopise	světově uznávané databáze ²⁾	12
		seznam recenzovaných periodik ²⁾	10
B	odborná kniha	světový jazyk ³⁾	40
		ostatní jazyky	20
D	článek ve sborníku ⁴⁾		8
P	patent	evropský nebo mezinárodní patent (EPO, USPTO), patent USA a Japonska	 500
		český nebo národní patent s výjimkou patentu USA a Japonska, který je využíván na základě platné licenční smlouvy	 200
		ostatní patenty ⁵⁾	 40
Z (T)	poloprovoz, ověřená technologie, odrůda, plemeno		100 ⁶⁾
S, F, G, H, L, N, R	prototyp, uplatněná metodika, funkční vzorek, software, užitný a průmyslový vzor, specializované mapy, poskytovatelem realizované výsledky		40 ⁶⁾
V	výzkumná zpráva, která je výsledkem obsahujícím utajované informace		50 ⁷⁾

Pro výsledky uplatněné od roku 2008 včetně:

Druh výsledku		I – obory NRRE ⁰⁾	II – ostatní obory
J _{imp}	článek v impaktovaném časopise ¹⁾		10 až 305 ²⁾
	článek v prestižním impaktovaném časopise (<i>Nature</i> , <i>Science</i> , <i>Proc. Natl. Acad. Sci. USA</i>) ³⁾		500
J _{neimp}	článek v recenzovaném časopise	světově uznávané databáze ⁴⁾	12
		seznam recenzovaných periodik ⁴⁾	10
B	odborná kniha	světový jazyk ⁵⁾	40
		ostatní jazyky	20
D	článek ve sborníku ⁶⁾		8
P	patent	evropský nebo mezinárodní patent (EPO, USPTO), patent USA a Japonska	 500
		český nebo národní patent s výjimkou patentu USA a Japonska, který je využíván na základě platné licenční smlouvy	 200
		ostatní patenty ⁷⁾	 40
Z	poloprovoz, ověřená technologie, odrůda, plemeno		100
F	užitný vzor		40
	průmyslový vzor		40
G	prototyp, funkční vzorek		40
H	poskytovatelem realizované výsledky		40
N, L	certifikované metodiky a postupy, specializované mapy s odborným obsahem		40
R	software		40
V	výzkumná zpráva, která je výsledkem obsahujícím utajované informace		50

Hodnocení kvality výzkumu

- Kvantifikace výzkumu (počty publikací) s vazbou na kvalitu (citační ohlas) by měla být jedním z podkladů hodnocení, nikoli nahrazovat hodnocení jako takové
- Podstatná je normalizace výsledků s ohledem na oborové rozdíly
- Známá je nerovnoměrnost pokrytí vědních disciplin ve WoS
- Podstatně chybné je mechanické užívání IF časopisů
- **Český přístup: Kvantifikace publikací je jediným kritériem, tato kvantifikace lineárně odpovídá množství přidělených financí. Kvalitou nebo kalibrací dat se česká metodika vůbec nezabývá.**

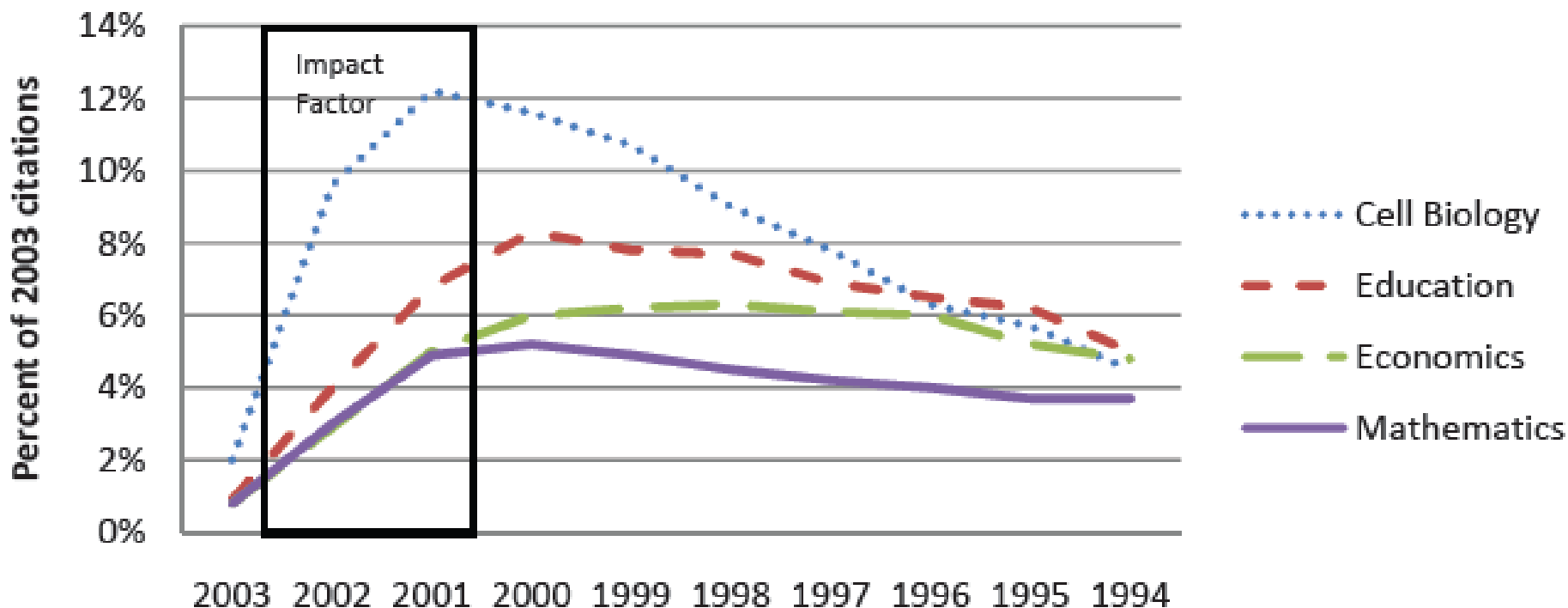
Chybné použití IF k hodnocení článků i autorů

- 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.

Impaktní faktor nezohledňuje odchylky v charakteru citací

- IF je počítán z dvouletého okna, charakteristiky citací se však výrazně liší mezi jednotlivými disciplinami
- Hodnocení významu rozdílu citační frekvence je proto závislé na uvažované disciplině

Citation Curves



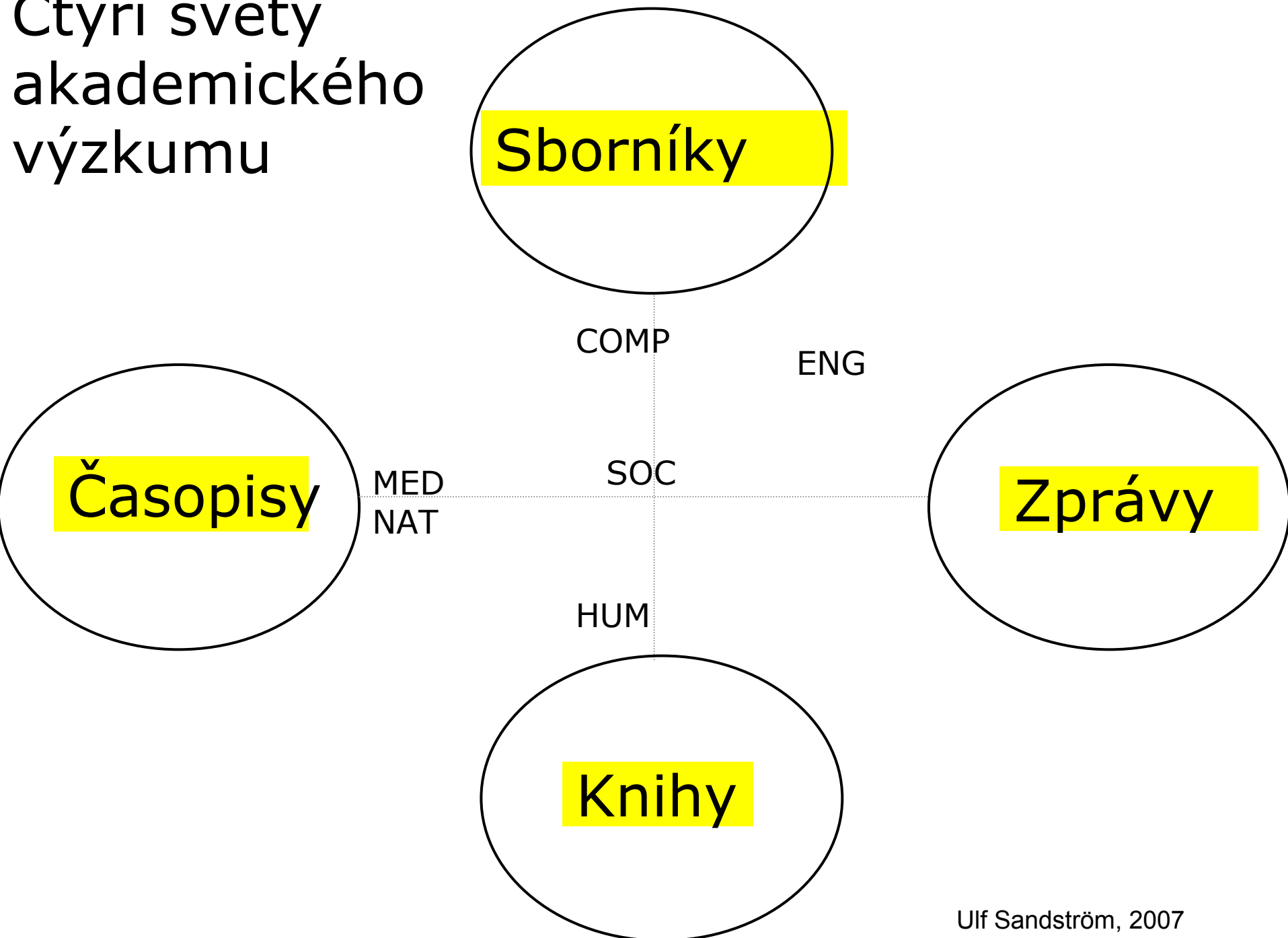
Eugen Garfield k chybnému užívání IF, citát z roku 2005:

- *Používání citačních indexů časopisů namísto skutečných citačních indexů článků pro hodnocení jednotlivých autorů je vysoce spornou záležitostí.*
- *Grantové agentury a jiné instituce zodpovědné za vědní politiku si často přejí vyhnout se práci, kterou představuje získání skutečných počtů citací pro jednotlivé články a autory.*
- *Typicky to vypadá tak, že se při posuzování bibliografie posuzovaného autora nahrazují skutečné počty citací jeho prací impaktními faktory časopisů, ve kterých byly publikovány. Impaktní faktor se tak používá k odhadu očekávaného dopadu jednotlivých článků, což je dost pochybné, uvážíme-li známé zkreslení pozorované u většiny časopisů.*

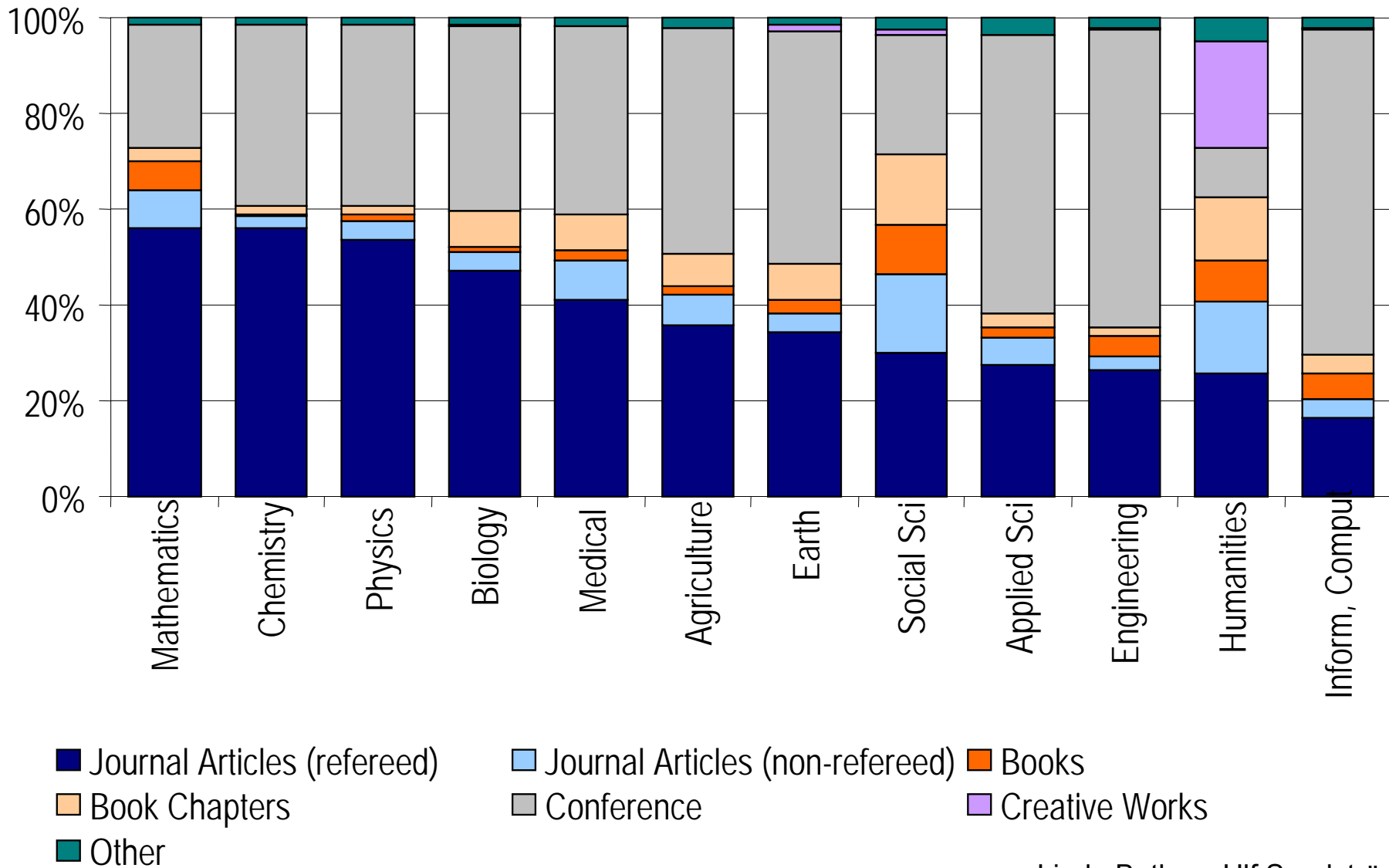
Celkové pokrytí hlavních oborů v ISI

EXCELLENT (> 80%)	VERY GOOD (60-80%)	GOOD(40-60%)
Biochem & Mol Biol	Appl Phys & Chem	Mathematics
Biol Sci – Humans	Biol Sci – Anim & Plants	Economics
Chemistry	Psychol & Psychiat	Engineering
Clin Medicine	Geosciences	MODERATE (<40 %)
Phys & Astron	Soc Sci ~ Medicine	Other Soc Sci
		Humanities & Arts
		(Computing)

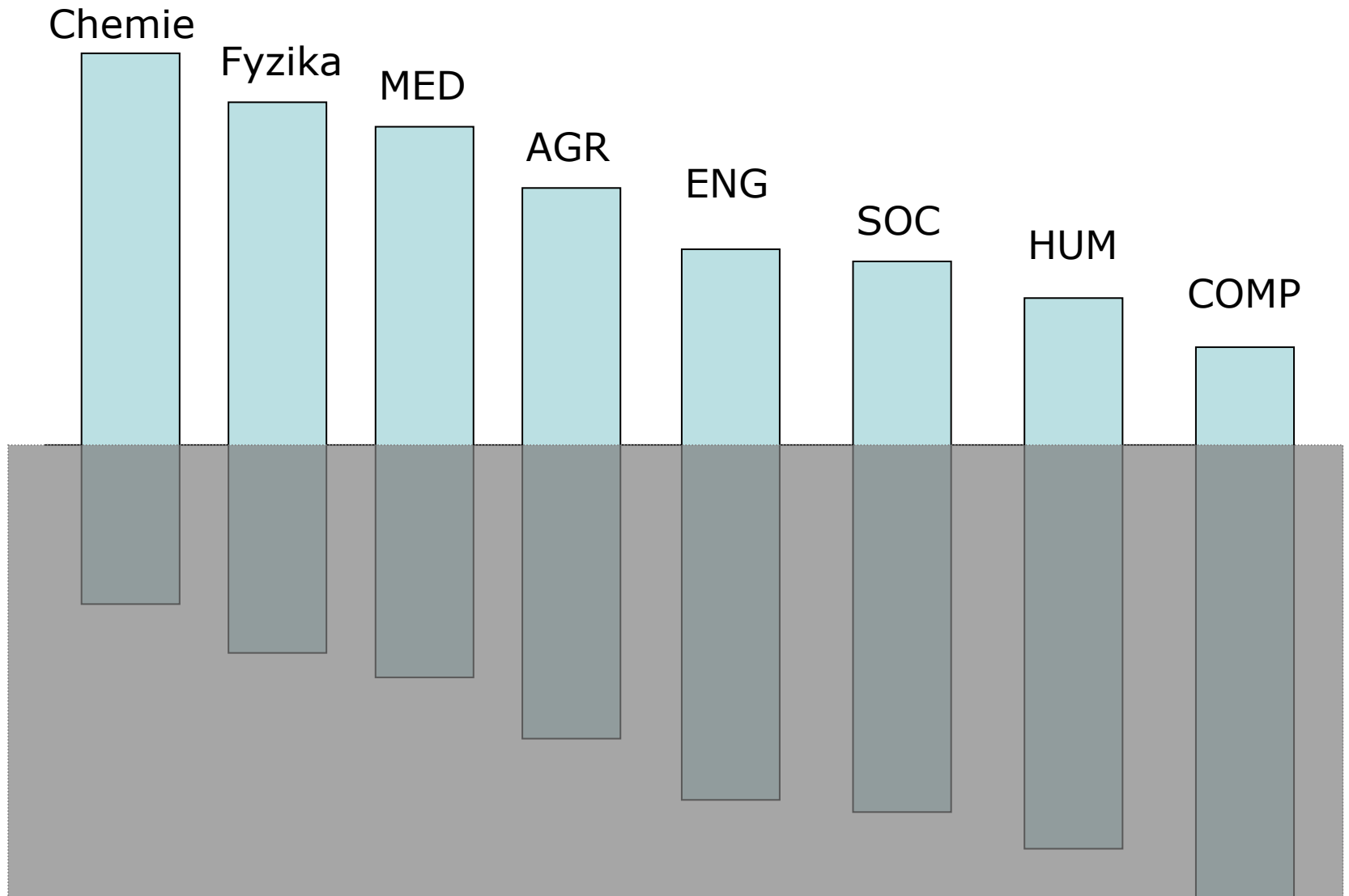
Čtyři světy akademického výzkumu



Rozdílná struktura typů publikací mezi různými obory



Metafora ledovce



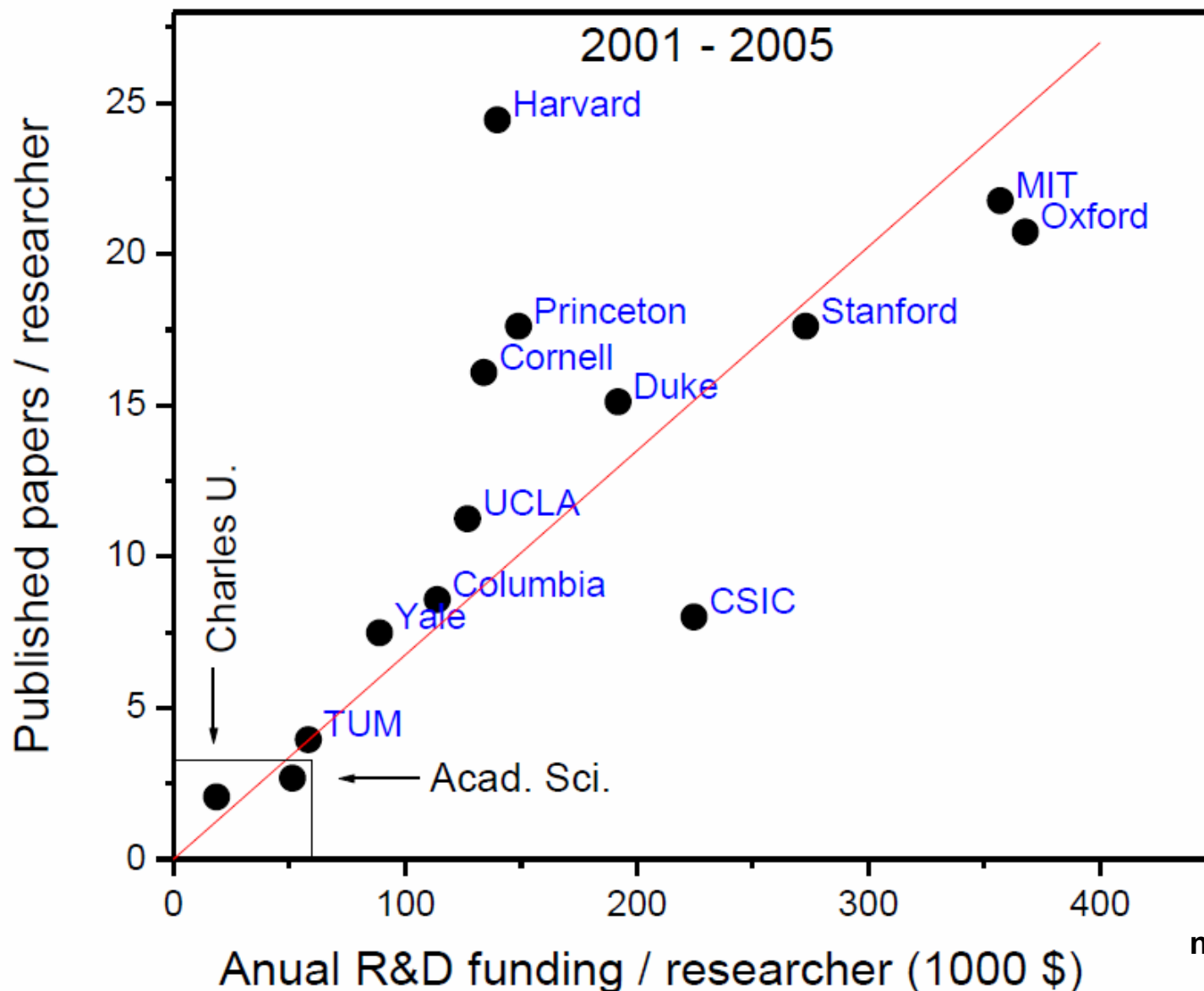
Bibliometrie zneužitá podle politických požadavků

- **Politické požadavky jednoduchého hodnocení jedním číslem**
- Vědecká praxe zdánlivě legitimizuje užití IF
- Zdánlivost „objektivity“ čísel dělá věc lákavou pro politiky
- **Pro „odborníky“, kteří se podíleli na KHV, je to etická diskreditace**
- F. Mattern (ETH Zurich) mj. demonstroval na ECSS 2008, jak Harzingův „Publish or perish“ vyrobí citovaného autora ze slova „Bullet“ – ani v nejnovější verzi programu to není odstraněno, přesto se tento nástroj často velmi naivně a zcela mechanicky používá

Problémy amatérské bibliometrie

- Užívání bibliometrie **na úrovni jednotlivců v institucích**
- Grantové agentury vynucují uvádět bibliometrické parametry a **nutí uchazeče do amatérské bibliometrie**
- Přes konstrukce žebříčků se produkty amatérské bibliometrie šíří dál
- Amatérská bibliometrie zvyšuje poptávku po hodnocení; **evaluace pochybné kvality jsou artiklem marketingu**; úspěšní vědci jsou zainteresováni „objektivně“ ukazovat, že jsou lepší, než jejich kolegové
- Amatérská bibliometrie **snižuje standardy bibliometrie obecně**
- **Peer review může pozbýt charakter alternativního hodnocení kvality**, pokud se posuzovatelé coby amatérští bibliometristé budou pouze „věřit číslům“

Výsledek? Redukce AV ČR destruuje nejvýkonnější instituci české vědy



Zdroj: J. Málek
(člen RVV),
prezentace ČR
na semináři OECD,
Praha 16.10:2009

Problémy hodnocení aplikovaného výzkumu

- Patenty hodnoceny analogicky hrubým počtům nerecenzovaných a necitovaných publikací
- Krátká doba na hodnocení – udělení Evropského patentu může trvat i 5-10 let
- Žádné hodnocení výnosu patentu z licencí versus veřejné i soukromé zdroje na jeho vývoj, poplatky za udělení patentu – náklady se stávají uznatelnými náklady projektů, po výnosech se nikdo nepídí
- Realizace patentů a hodnocení přínosu se nevejde do časového „okna“ hodnocení
- U zahraničních profesorů vykazujících patenty jich většina zanikne pro neplacení udržovacích poplatků
- Neřešeny jsou „investice“ v podobě soudních sporů po udělení patentu – mohou být o několik řádů vyšší, než přímé náklady patentování

Patenty jako podpora inovací?

- **Patenty samy nestimulují invenci**, ale jsou nástrojem boje s konkurencí na trhu
- **Patenty rychlost výzkumu nestimulují, ale brzdí**

[Joseph Stiglitz (Nobelova cena za ekonomii 2001) a John Sulston (Nobelova cena za medicínu 2003), *Who Owns Science*, Institute for Science, Ethics and Innovation, University of Manchester, červenec 2008]

Patenty jsou z hlediska výzkumné produkce jen smetí

- Přes 90 procent patentů není nikdy použito, protože nemají žádnou komerční hodnotu – analogie nerecenzovaných a nikým nečtených publikací
- Uplatňované patenty slouží především k potlačení konkurence, u komplexních průmyslových odvětvích se jedná také o přístup k technologiím konkurence (také tzv. „patent pools“)
- Pouze v situaci, kdy trh s produkty nebo službami, ke kterým se vynález vztahuje, nemohou reagovat dostatečně rychle, mají patenty smysl
- Z hlediska uplatnění práv z patentu je důležitá schopnost vymezit přesně hranice – funguje dobře u farmacie nebo chemie, zejména pak ve vazbě na léčiva; **zcela selhává u abstraktních pojmů, což je zejména případ informatiky**

Problémy s patenty v souvislosti s informačním prostředím ?

- Otevřené problémy v souvislosti s patentovými nároky na abstraktní logické konstrukce
- Kritérium „obtížnosti dosažení vynálezu“ lze použít k vyloučení triviálních patentů nebo vynálezů
- „Post-industriální“ inovace spočívají na využití abstraktních kalkulů, nikoli experimentů nebo nákladného vývoje prototypů, výrobních protor, atd.

Dilema

- Patenty jako zdroj inovací a nových vynálezů
- Patenty spojené se stroji, aparáty nebo systémy
- Patenty nejsou uzpůsobeny na zvládnutí abstrakcí založených na logice a bitech
- Patenty vycházejí z konkrétního modelu vynalézání, ten se všach mění s přechodem od energie k informaci

V oblasti softwaru jsou patenty obzvláště problematické

- Americká soudkyně P. Sarrisová: *„Duševní procesy a abstraktní intelektuální koncepty nejsou patentovatelné, protože jsou to základní nástroje vědecké a technické práce.“*
- Srvn. též úvahu J. Grusky „A Perception of Informatics“ o metodologické důležitosti užití informatiky, Academia Europæana, květen 2010

Příklad patentované softwarové triviality

- Scalable virtual world chat client-server system

Abstract

- The present invention provides a highly scalable architecture for a three-dimensional graphical, multi-user, interactive virtual world system. A plurality of users can interact in the three-dimensional, computer-generated graphical space where each user executes a client process to view a virtual world from the perspective of that user. The virtual world shows avatars representing the other users who are neighbors of the user viewing the virtual world. In order that the view can be updated to reflect the motion of the remote user's avatars, motion information is transmitted to a central server which provides position updates to client processes for neighbors of the user at that client process. The client process also uses an environment database to determine which background objects to render as well as to limit the movement of the user's avatar.
- Inventors: **Leahy; Dave** (Oakland, CA), **Challinger; Judith** (Santa Cruz, CA), **Adler; B. Thomas** (San Francisco, CA), **Ardon; S. Mitra** (San Francisco, CA) Assignee: **Worlds, Inc.** (San Francisco, CA)
Appl. No.: **08/747,420** Filed: **November 12, 1996**

Z tiskové zprávy Worlds.com: “[The patent was awarded] for its scalable 3D server technology ... [by] the United States Patent Office. The Company believes the patent may apply to currently, in use, multi-user games, e-Commerce, web design, advertising and entertainment areas of the Internet.”

Náklady na patentové spory jsou násobkem agregátního přínosu patentů

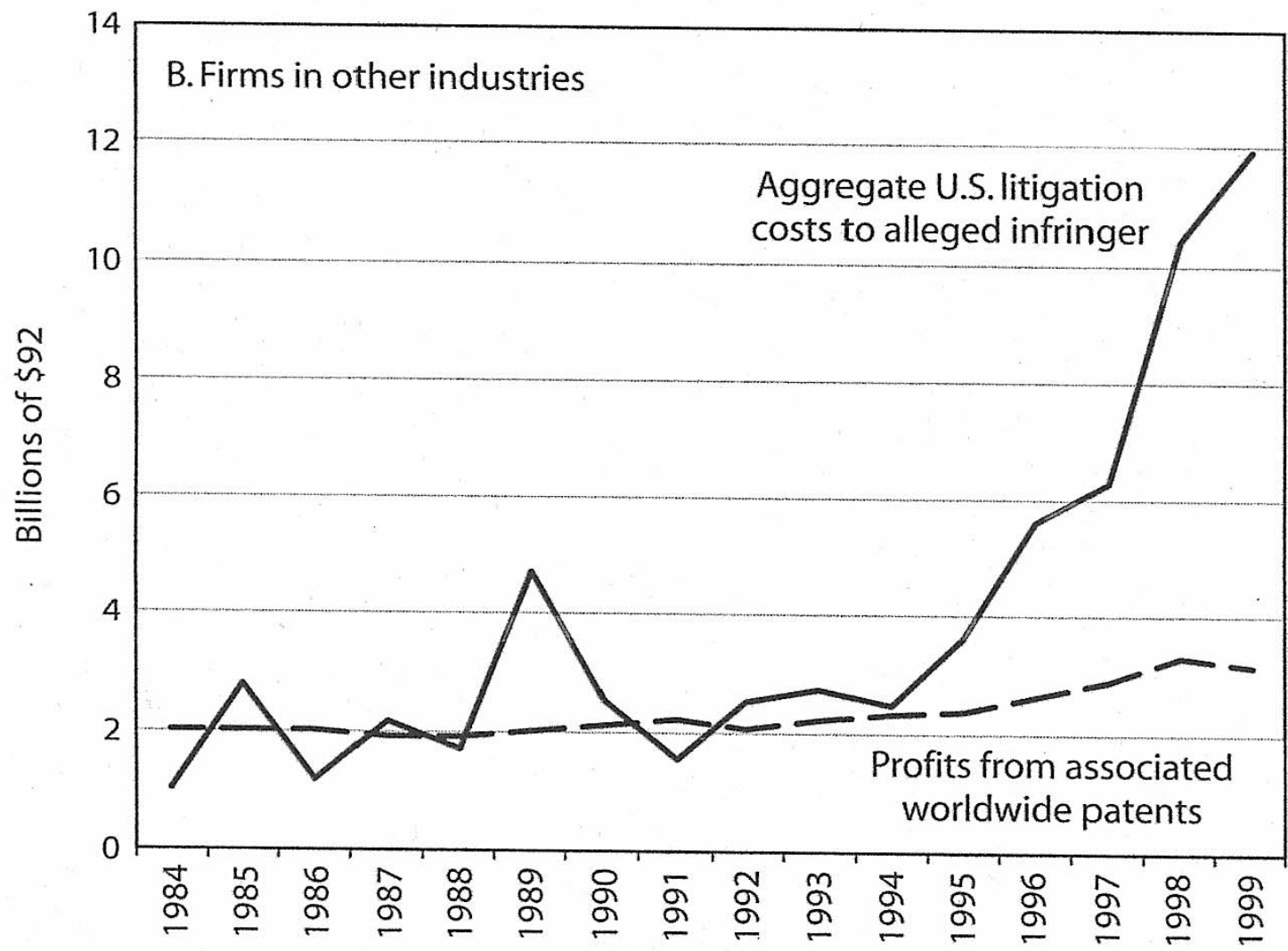
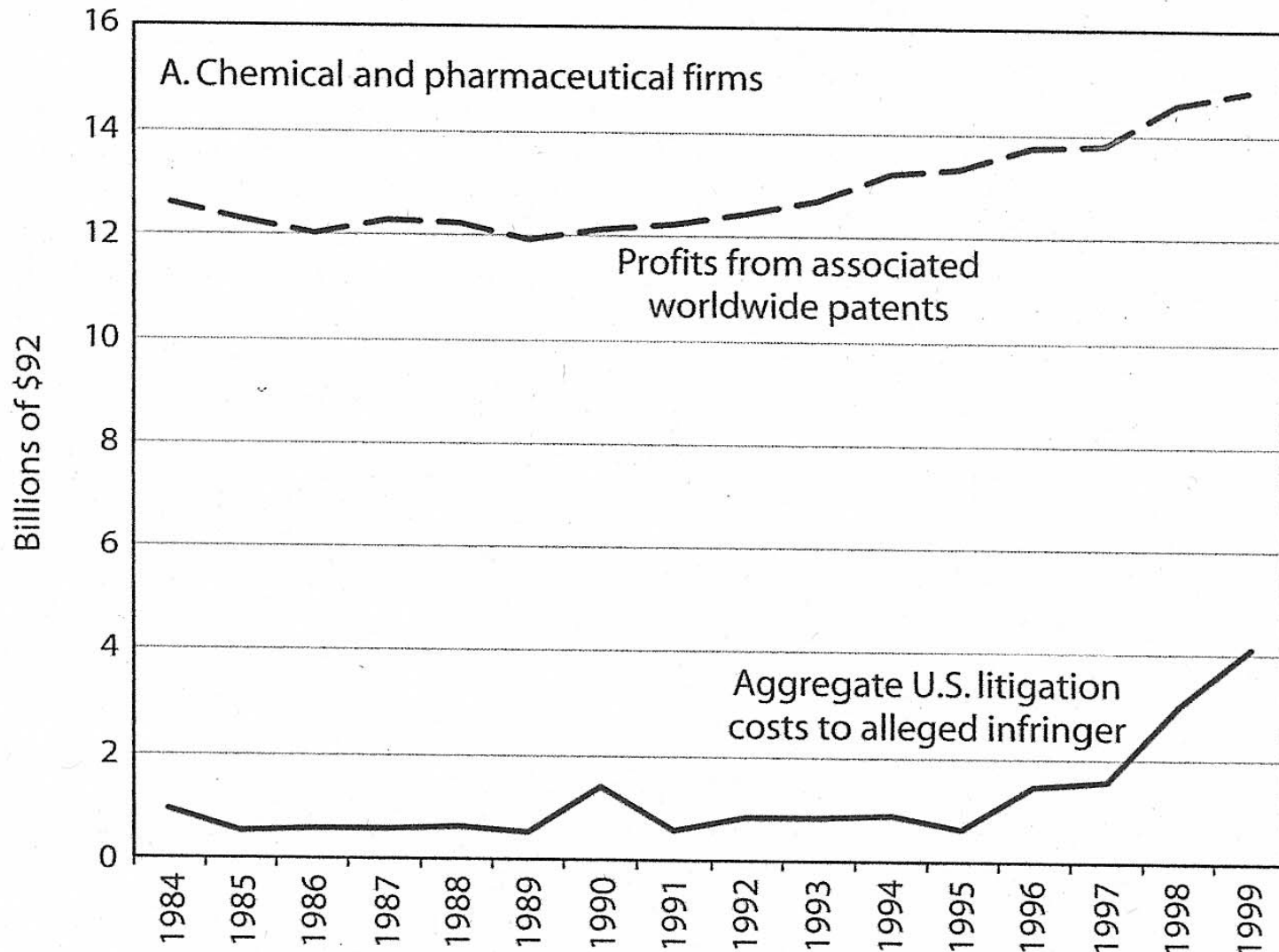


Figure 6.5. Aggregate Profits from Patents and Aggregate Litigation Costs for U.S. Public firms

Pouze farmaceutické a chemické firmy mají pozitivní bilanci



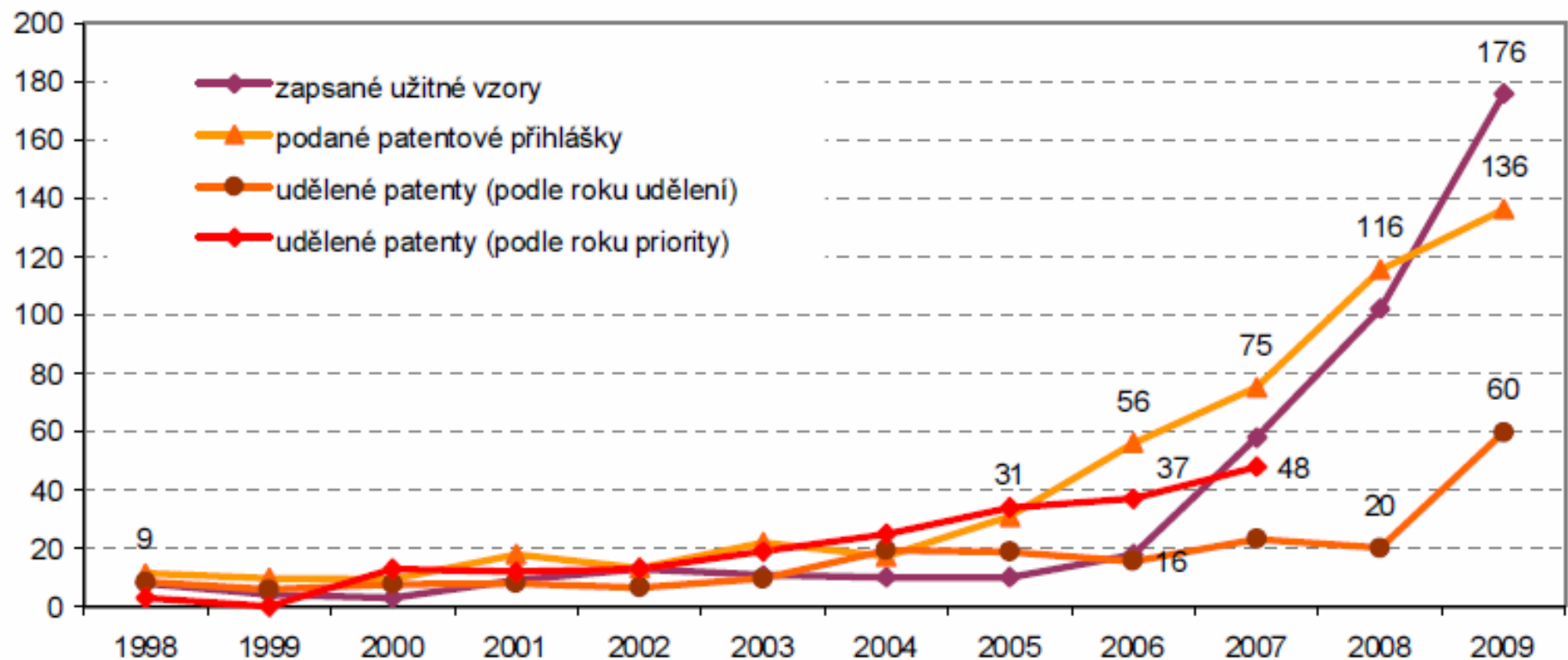
Patenty jako loterie

- Velká výhra pro pár úspěšných
- Většina ztrácí

- Podobnost se sázením?
Výhra přichází jen zřídka.

- Empirické výsledky z USA z 90.let: Celkové náklady na patenty jsou vyšší, než celkové přínosy, patenty navíc nekorelují s inovacemi, protože zahrnují i oportunistické chování (patent trolls, atd.)

Efekty státní podpory ochrany duševního vlastnictví na vysokých školách v ČR

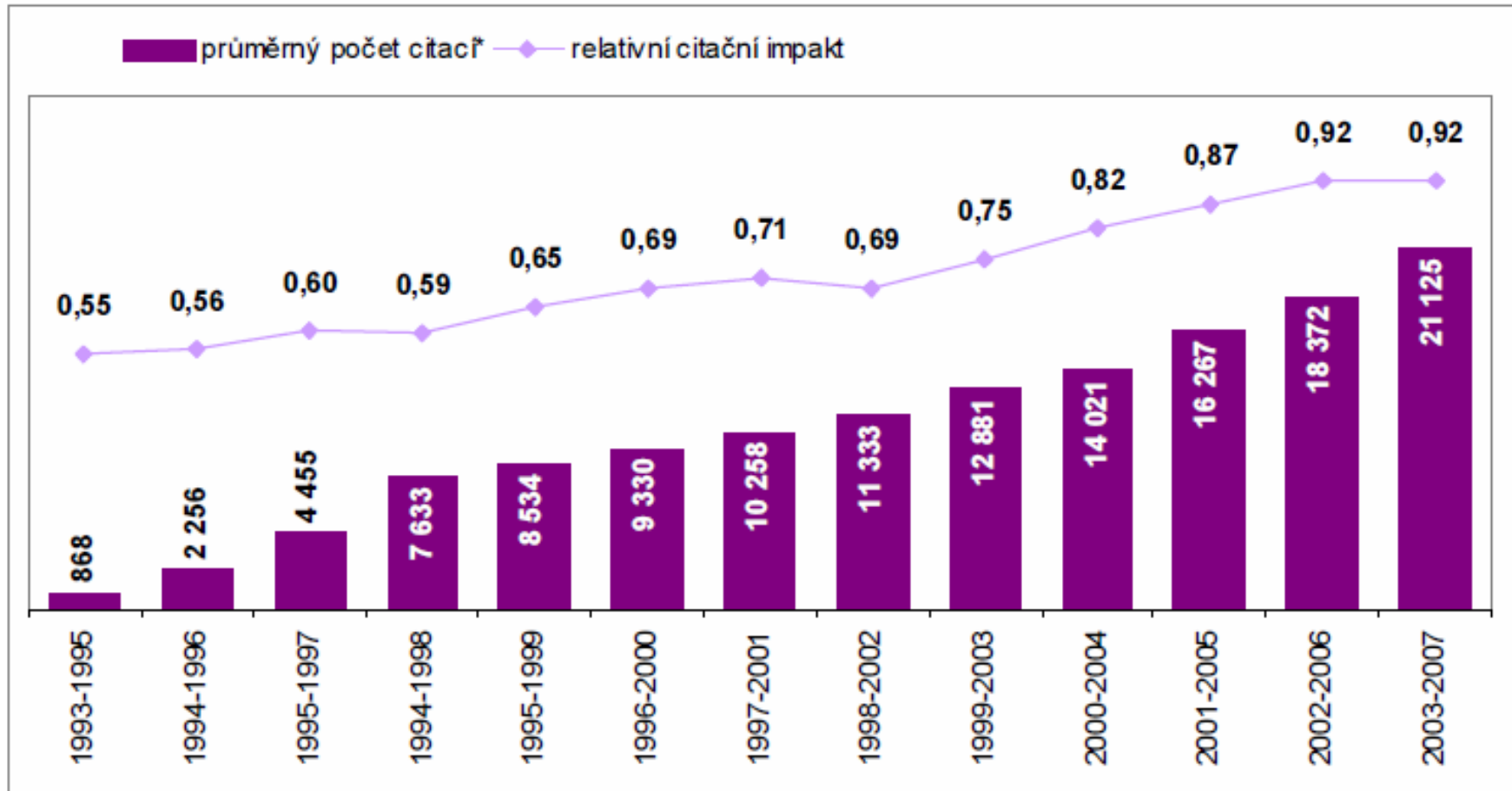


Zdroj: Úřad průmyslového vlastnictví ČR

Zdroj: Výzkum a vývoj na vysokých školách, ČSÚ, květen 2010

Relativní citační impakt rostl v ČR naopak velmi rychlým tempem *před* realizací reformy

Citace odborných článků vzniklých v ČR a jejich relativní citační impakt:



* roční průměry za dané pětileté období

Zdroj: Thomson ISI National Science Indicators, 1981-2007, Rada pro výzkum, vývoj a inovace, dopočty ČSÚ

Zdroj: Výzkum a vývoj na vysokých školách, ČSÚ, květen 2010

Chybějící koncepty

- Služby v IT – náhrada prodeje programů nebo licencí poskytováním služeb na nich založených
- Experimentální CS&E – užívání artefaktů na pomezí výzkumu a vývoje, prototypy uvádějící nové koncepty schopné praktického nasazení, které je testuje

Z vypořádání připomínek k aktuálně platné metodice hodnocení

<p>(RVŠ-4) Zásadní připomínka Původní znění: <i>Příloha č.1 Výsledky typu R, software, 40 bodů</i> Návrh na změnu: Vyřadit z bodového hodnocení. Zdůvodnění: Současná definice umožňuje bodování triviálního softwaru tvořeného jen za účelem bodového zisku. Chybí jakýkoliv způsob objektivního rozlišení mezi užitečným, netriviálním softwarem a softwarem potěmkinovským, tvořeným pro body</p>	<p>Neakceptováno Nelze tvůrcům takového výsledku upřít právo na hodnocení tohoto výsledku. Úroveň na posouzení, zda se jedná či nejedná o relevantní výsledek je spíše apelem na poskytovatele, aby důsledně plnili povinnosti uložené jim zákonem č. 10/2002 Sb. pro ohodnocení řešení výzkumných aktivit a správnost a pravdivost údajů předávaných do RIV. I zde byla možnost škálování diskutována, zatím však bez akceptovatelného výsledku. Připomínka bude znovu diskutována při přípravě Metodiky 2010.</p>
<p>(RVŠ-6) Zásadní připomínka Původní znění: <i>Příloha č. 2, písmeno D: D – článek ve sborníku – 8</i> Návrh na změnu: D – článek ve sborníku – 12 (informatické obory) / 8 (ostatní obory) Zdůvodnění: Prestižní publikace v oblasti informatiky a IT jsou publikovány především na špičkových konferencích – na rozdíl od jiných oborů, kde prestižní práce jsou publikovány především ve špičkových časopisech. Viz závěr č. 2 na str. 3 následujícího článku: http://www.informatics-europe.org/ECSS08/papers/Research_evaluation_CACM.pdf</p>	<p>Neakceptováno Zdůvodnění pro diferencované hodnocení příspěvků ve sbornících pro obory informatické a neinformatické nelze akceptovat. V opačném případě by bylo relevantní naopak např. článkům v časopisech v informatických oborech dát bodové ohodnocení nižší než u oborů jiných.</p>

- Informatika je v extrémně špatné situaci vzhledem ke zvoleným způsobům hodnocení (a přiřazování financí)
- Hesla o podpoře informační společnosti jsou tváří tvář této anomálii zvláště problematická
- Není divu, že finanční podpora VaV v informatice je malá, i když výdaje za konkrétní systémy nasazované státem jsou zpravidla více než štědré...