



METODIKA

Research Evaluation - Approaches and Methods

Daniel Münich
MŠMT

IPN Metodika:

10. ledna 2015

National Technical Library

www.metodika.reformy-msmt.cz

www.techlib.cz/cs/82734-kre-14#tab_program



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



1. Stručně o obecných principech hodnocení

2. Institucionální hodnocení

- Účely, struktury, principy, dimenze
- Česko a Evropa, Česko, Británie

3. Bibliometrická podpora

- Principy
- Ukázky
- Varování

Stanovení priorit
VaV

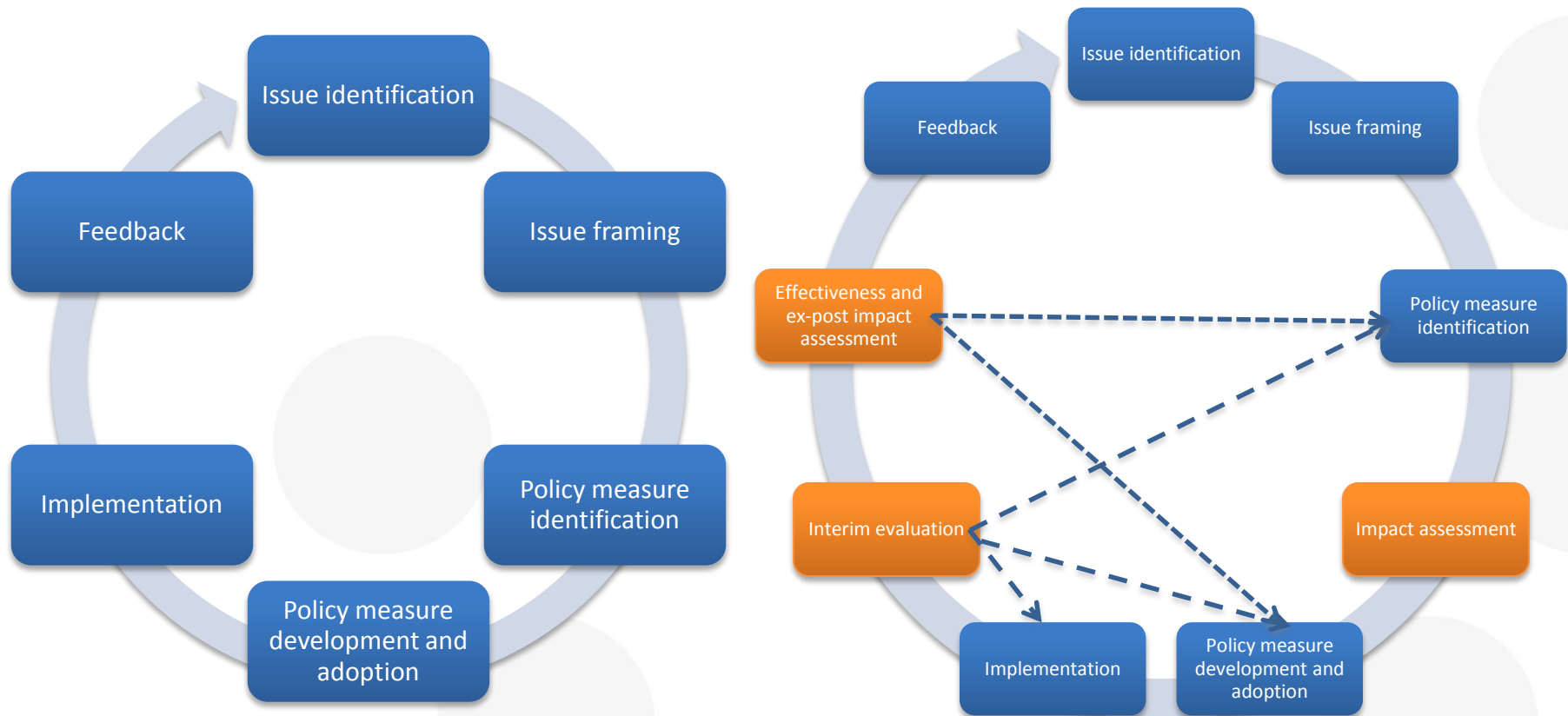
Řízení lidských
zdrojů VaV

**Bez kvalitního hodnocení
nebude kvalitního řízení**

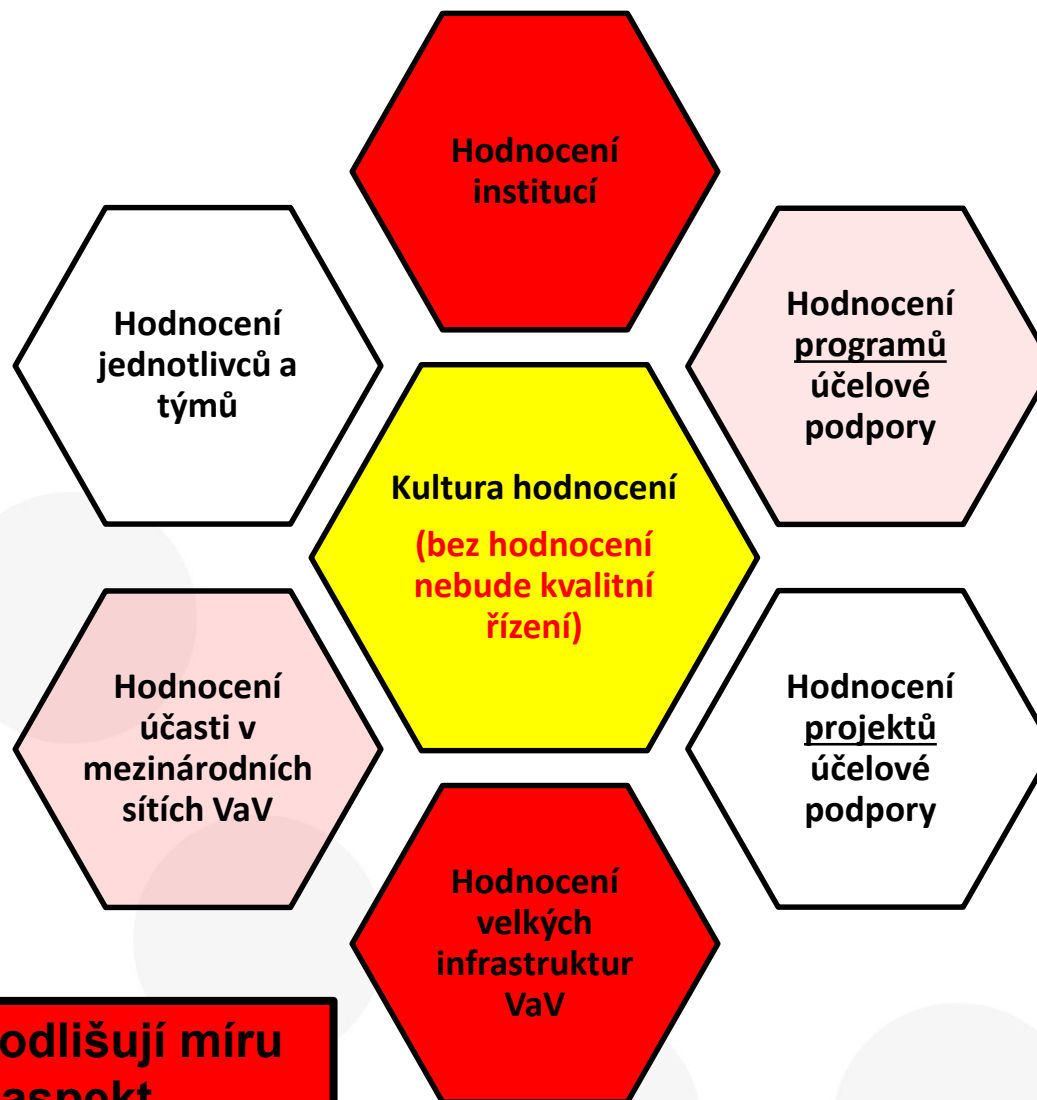
Efektivní pořizování
a využívání
infrastruktur VaV

Správně nastavené
motivace ve VaV

Evaluation in general should inform the whole policy cycle



Rozdílné typy hodnocení ve VaV



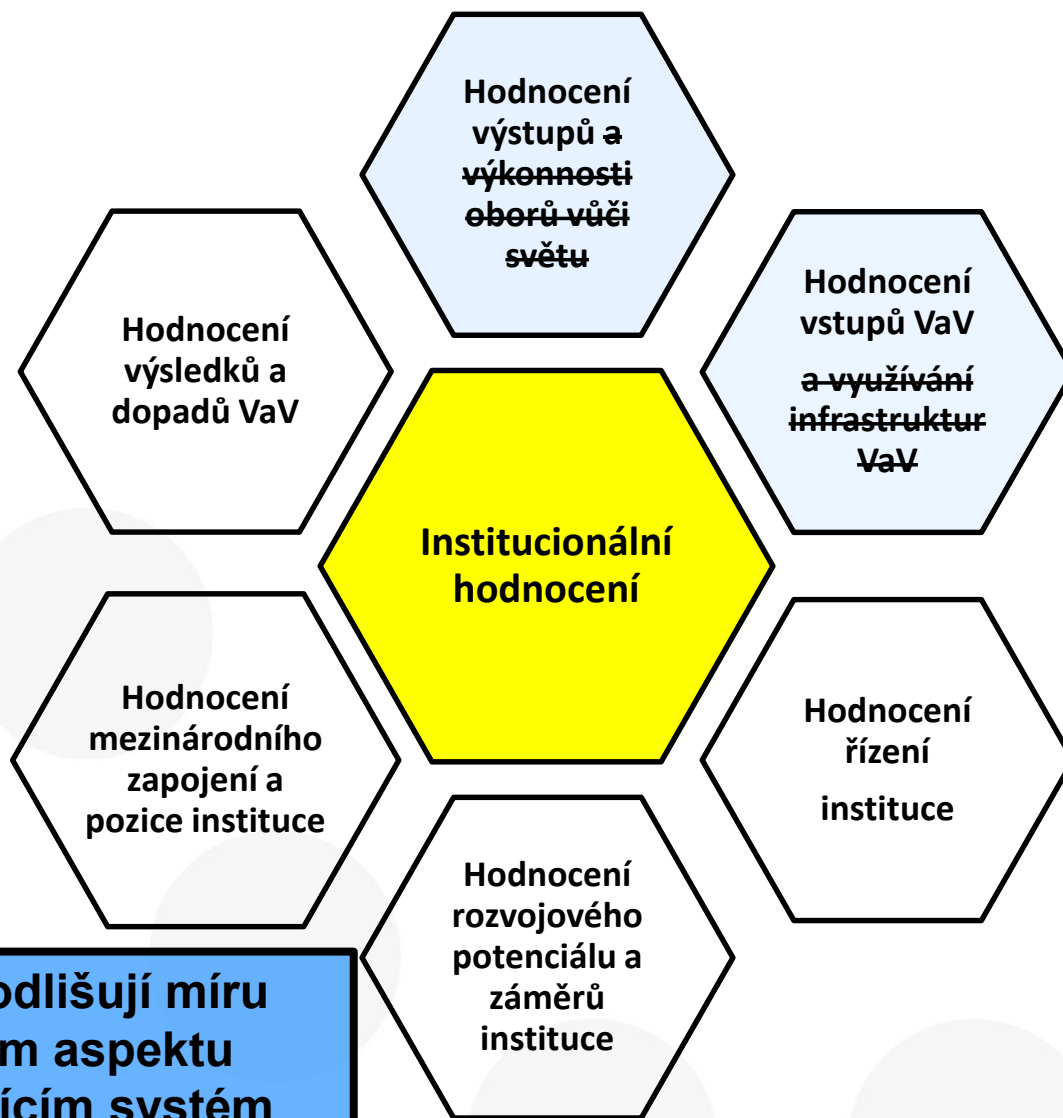
Odstíny červené odlišují míru toho, jak je daný aspekt obsažen v IPN Metodika.

Different evaluations and purposes



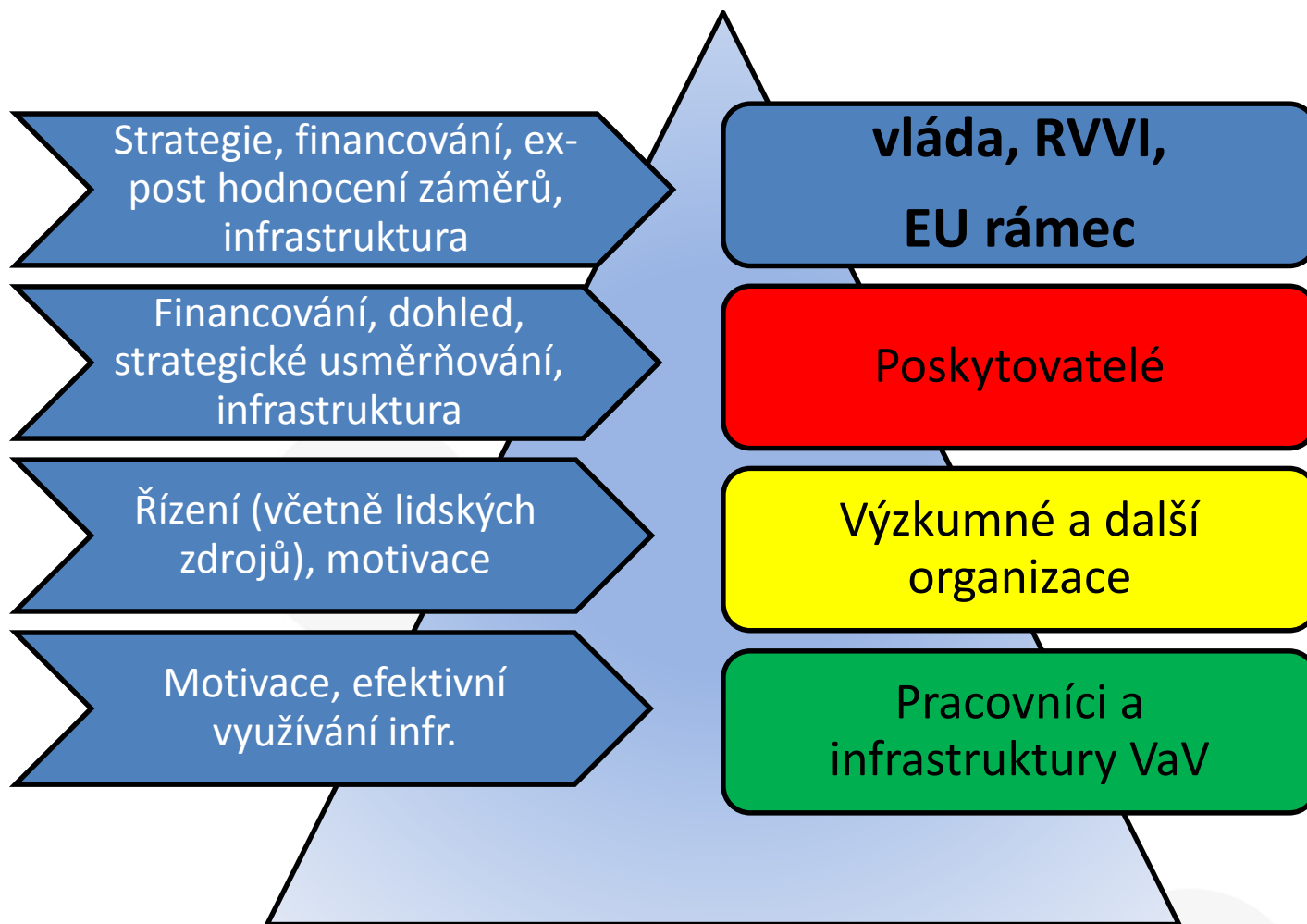
- **Hodnocení projektů a grantů**
 - Ex-ante, ex-post
- **Hodnocení programů podpory VaV**
 - Ex-ante, interim, ex-post
- **Hodnocení velkých (sdílených) infrastruktur**
 - Ex-ante, interim, ex-post
- **Hodnocení zapojení země do mezinárodních sítí VaV**
- **Hodnocení v institucích VaV (v.v.i.)**
 - průběžné / ad-hoc
 - VaV nebo i vzdělávání?
 - externí / **interní** (*internal governance*)

Komponenty institucionálního hodnocení

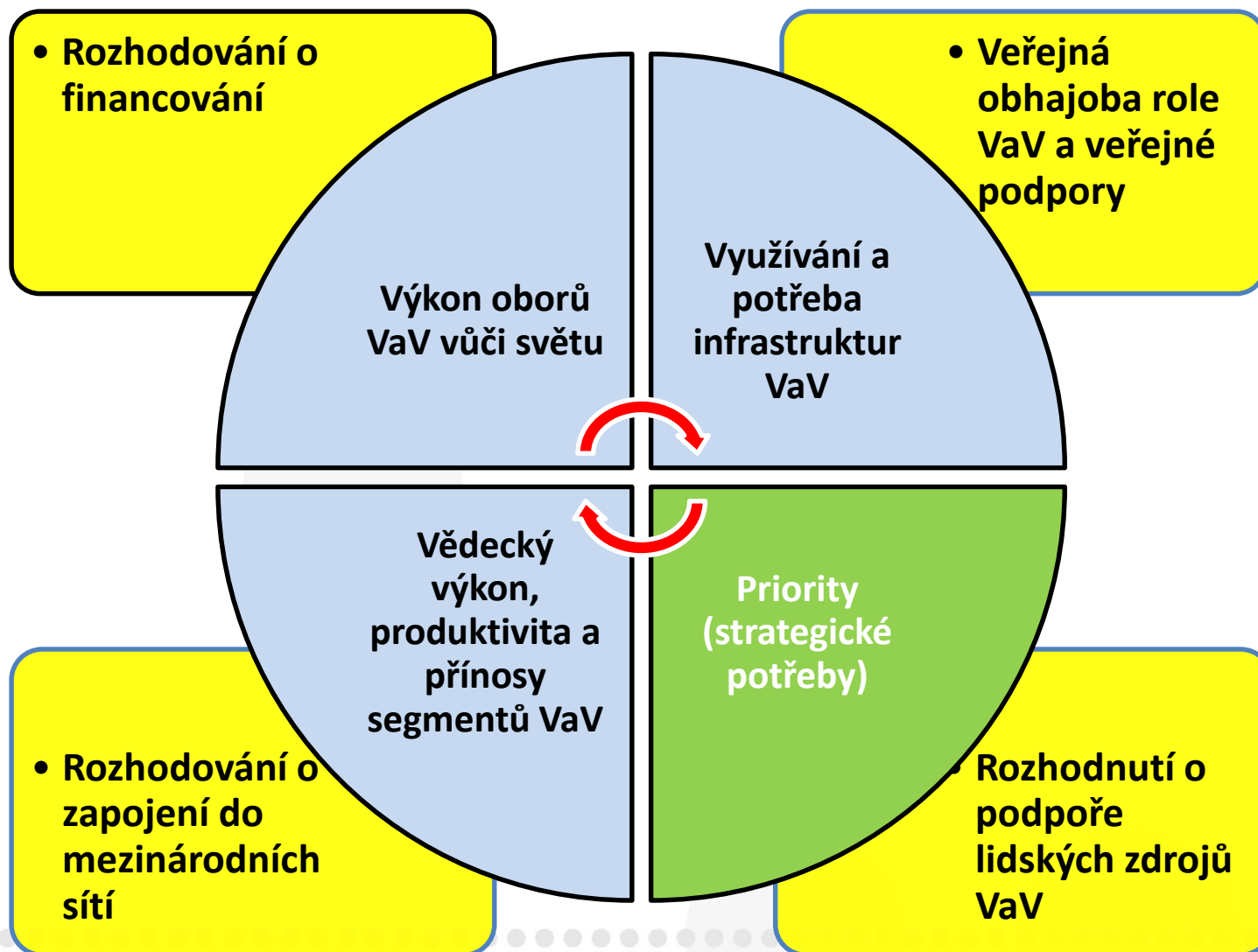


Odstíny modré odlišují míru toho, jak o daném aspektu informuje stávajícím systém institucionálního hodnocení.

Rozdílné úrovně řízení systému VaV



Informace v řízení na vyšších úrovních systému VaV



Institucionální hodnocení (rozlišujme!)



- **Úroveň hodnocení**

- vysoká škola, součást, katedra, tým, jednotlivci
- ústav / pracoviště, tým, oddělení, jednotlivci

- **Dimenze hodnocení**

- úroveň vstupů (lidé, peníze, zařízení)
- řízení / fungování (*internal governance*)
- vize rozvoje
- výstupy

The role and function of the EM



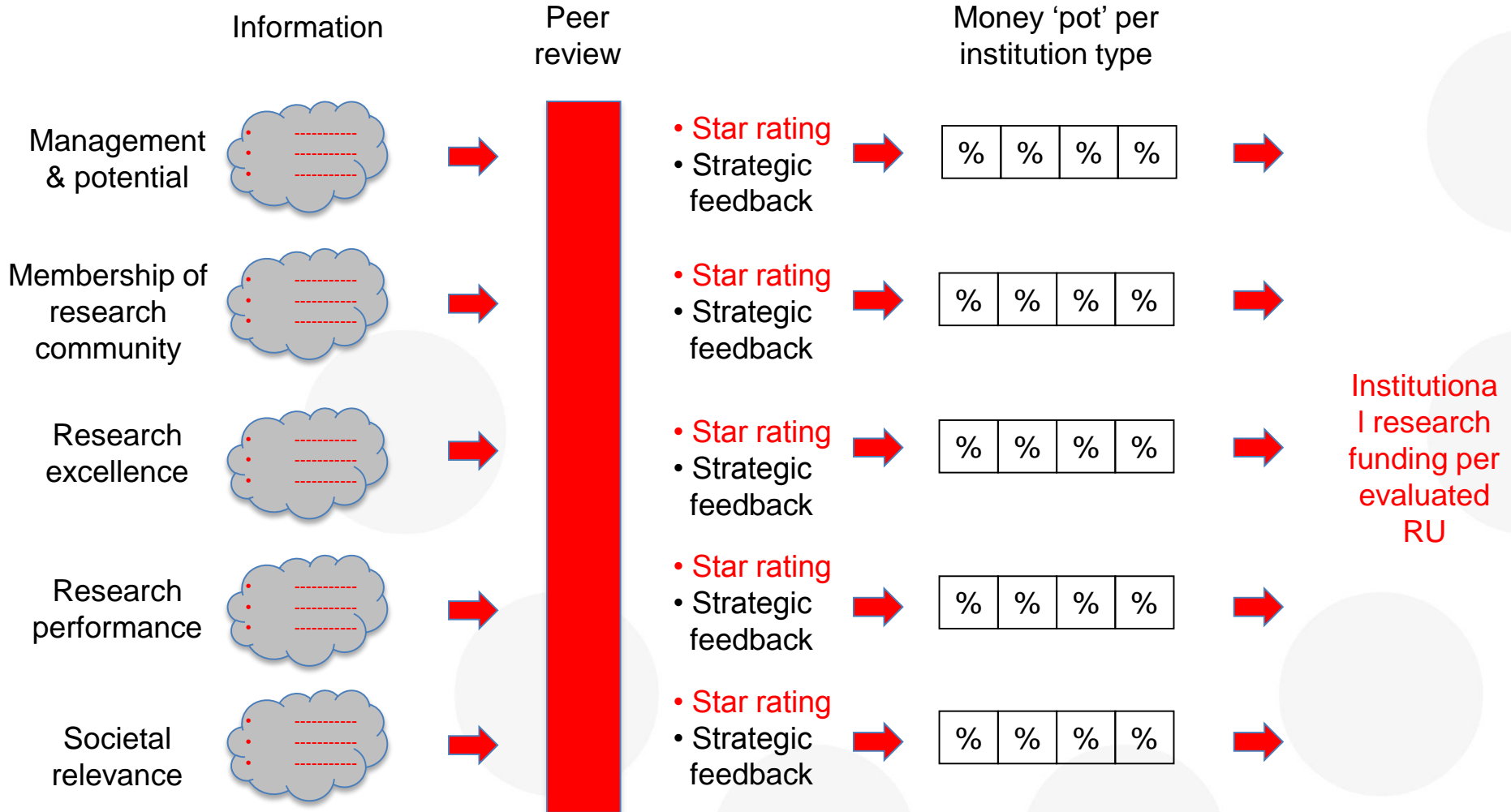
- **The key function of an evaluation system is to support public R&D governance in the attainment of its strategic objectives, i.e.**
 - To strengthen **R&D capacity**
 - To foster **excellence** in research
 - To foster its alignment with the **societal needs**
 - To support the **growth and competitiveness** of the Czech Republic
 - **By:**
 - Assessing past research performance
 - Supporting future performance improvement
-
- Provide **incentives** for positive change in the R&D system reflecting the strategy
 - Provide **strategic information** for policy making and R&D management

Evaluation (assessment) for a PRFS is necessarily more narrow



'General' evaluation	Evaluation in a PRFS
Informs policy making on the failures in the system and recommends possible policy interventions	Is part of a policy intervention : it acts upon previously identified failures and steers research behaviour to tackle these by providing incentives
Has no effects directly linked to the evaluation	Is intended to create effects
Has no consequences of gaming or unintended effects	Has inevitable consequences of gaming and may lead to unintended effects
Gives information on the positioning of the evaluated objects in the national/international context	Sets the evaluated objects in competition to each other

Evaluation in the context of a funding mechanism



Ratings tend to use 4-5 points

Score	Meaning	Research quality	Relevance to society	Viability
1	World leading/ excellent	The research unit has been shown to be one of the few most influential research groups in the world in its particular field.	The research unit makes an outstanding contribution to society.	The research unit is excellently equipped for the future.
2	Very good	The research unit conducts very good, internationally recognised research.	The research unit makes a very good contribution to society.	The research unit is very well equipped for the future.
3	Good	The research unit conducts good research.	The research unit makes a good contribution to society.	The research unit makes responsible strategic decisions and is therefore well equipped for the future.
4	Unsatisfactory	The research unit does not achieve satisfactory results in its field.	The research unit does not make a satisfactory contribution to society.	The research unit is not adequately equipped for the future.

Example: Netherlands

Absolutní kritéria hodnocení REF'14



4* *Quality that is world-leading in term of originality, significance, and rigour*

3* *Quality that is internationally excellent in terms of originality, significance, and rigour but which falls short of the highest standards of excellence*

2* *Quality that is recognized internationally in terms of originality, significance, and rigour*

1* *Quality that is recognized nationally in terms of originality, significance, and rigour*

U/C *Quality that falls bellow the standard of nationally recognized work. Or work which does not meet the published definition of research for the purpose of this assessment.*

Some effects of PRFS

- **Positive**

- Improved research and management attention to productivity and quality
- Performance focus and individual and collective levels
- Restructuring towards high-quality performers
- Greater transparency about performance

- **Negative**

- Focus on producing indicators rather than science, encouraging 'gaming'
- Narrowing of the career path, with growing influence of research managers
- Squeezing out the heterodox and interdisciplinarity
- Matthew Effect encourages lock-ins
- Research indicators often used in relation to teaching

Who does what



	Australia (2014)	Austria (2014)	Belgium (FL) – BOF (2014)	Finland (2014)	Italy (2014)	Netherlands (2014)	New Zealand (2014)	Norway / evaluations (2014)	Norway / PRFS (2014)	Sweden (2014)	UK – (2014)
Purpose											
Performance assessment	X	X	X	X	X	X	X	X	X	X	X
Inform funding	X		X	X	X		X		X	X	X
Main function											
Formative	X	X		X		X	X	X		X	
Summative			X		X				X		X
Formative function											
National R&D governance	X		X		X			X	X		X
Institutional R&D management	X	X		X		X	X	X		X	
Summative function											
R&D quality	X	X	X	X	X	X	X	X	X	X	X
R&D capacity building	X		X	X	X		X	X	X		X
Research excellence	X										X
Societal relevance	X				X	X					X
Notes											
		Performance contracts	Performance agreements	Performance contracts		Performance contracts		Separate PRFS – simple, 10% of funding		New PRFS in 2015	

National and institutional evaluations tend to be distinct



- **National evaluations typically**

- Are commissioned by national research funders or their principals
- Compare performance with other countries at the field level
- Collect strategic information to support research policy interventions
- Are summative with respect to the institutional level
- Lean heavily on bibliometric methods – with selective use of peer review

- **Institutional evaluations**

- Are commissioned by institutions or their principals
- If they benchmark (often they don't), compare with other institutions
- Try to explain performance at a level of detail that is formative, in order to support institutional research management
- Tend to use informed peer review (though note that modern research management routinely monitors performance by bibliometric means)

It's unusual to compare different types of research organisations (RO)



- **Types of organisation**
 - Universities / Higher Education Institutions
 - Scientific research institutes
 - Research and Technology Organisations (RTOs)
 - Government laboratories
 - NB that **hybrid** forms are increasingly appearing
- **Different types of RO are rarely benchmarked against each other**
 - Differences in mission or function
 - Different principals and accountability
 - Different mix among types of research and between research and non-research activities
- **Legal form is irrelevant – evaluation is based on function**
 - See, for example, the Norwegian institute PRFS, which separates institutes from universities but ignores legal form

Typical institutional research evaluation approaches by RO type



- **Universities**
 - Focus on quality, interaction with teaching
 - Preference for peer review
- **Scientific research institutes**
 - Ditto, but no emphasis on teaching unless PhD education is provided
- **RTOs**
 - Economic performance, social impacts, customer feedback, research quality
 - Predominance of social scientific methods; peers if quality is a major issue
- **Government labs**
 - Mission performance, usefulness to policy formation processes, social impacts, research quality
 - Predominance of social scientific methods; peers if quality is a major issue

Why use peers?

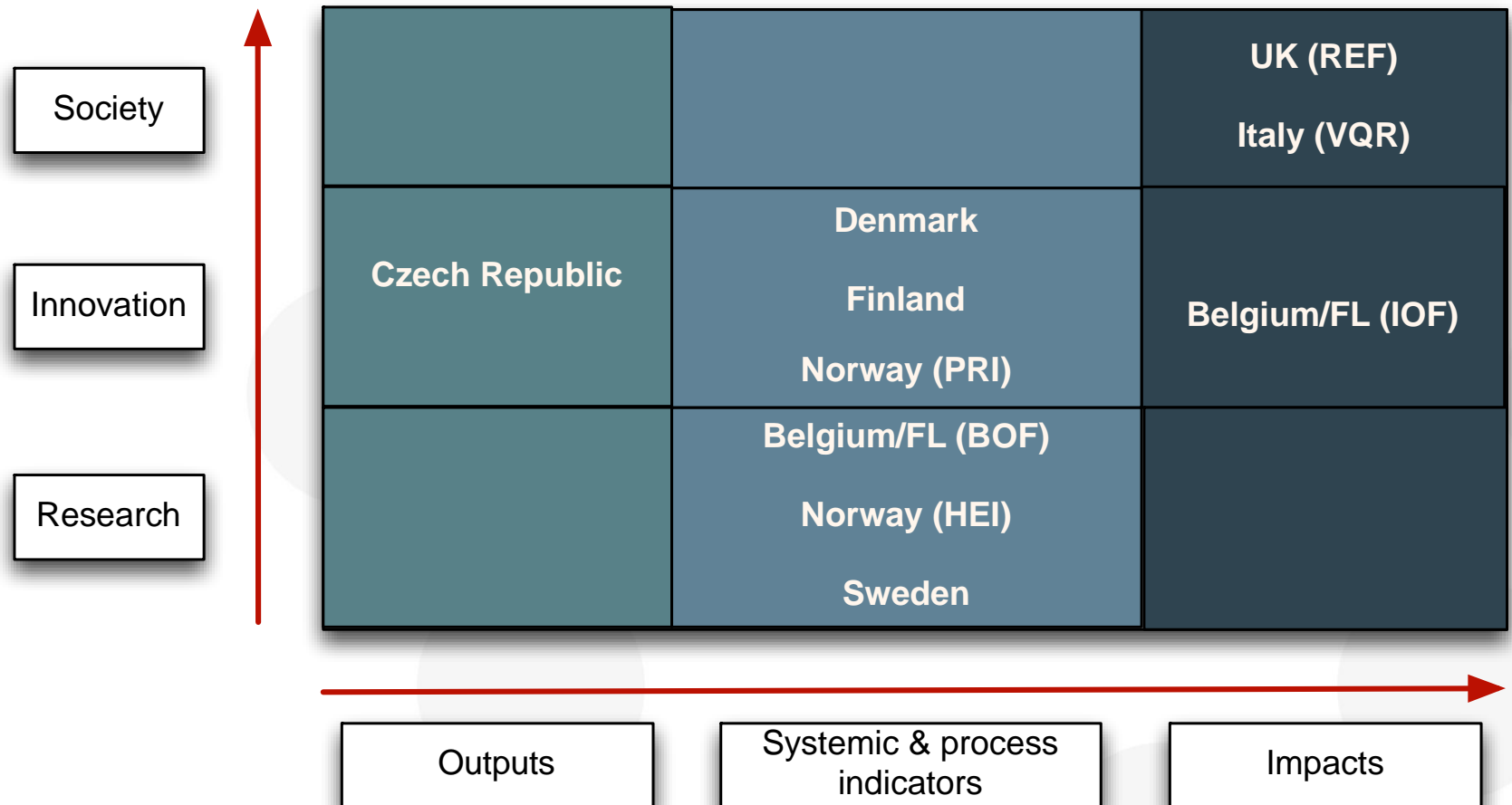
- **Legitimacy in the scientific community**
- **Ability to address context, for example at institutional level**
- **Understanding inter-field differences**
- **Ability to be formative**
- **Drawbacks include**
 - Those evaluated generally have to do work as part of the process
 - Need to ensure independence (research funders devote significant effort to this)
 - Expensive, especially if site visits are involved
 - Shortage of peers
 - Occasional misbehaviour at the individual level – typically constrained by the use of panels



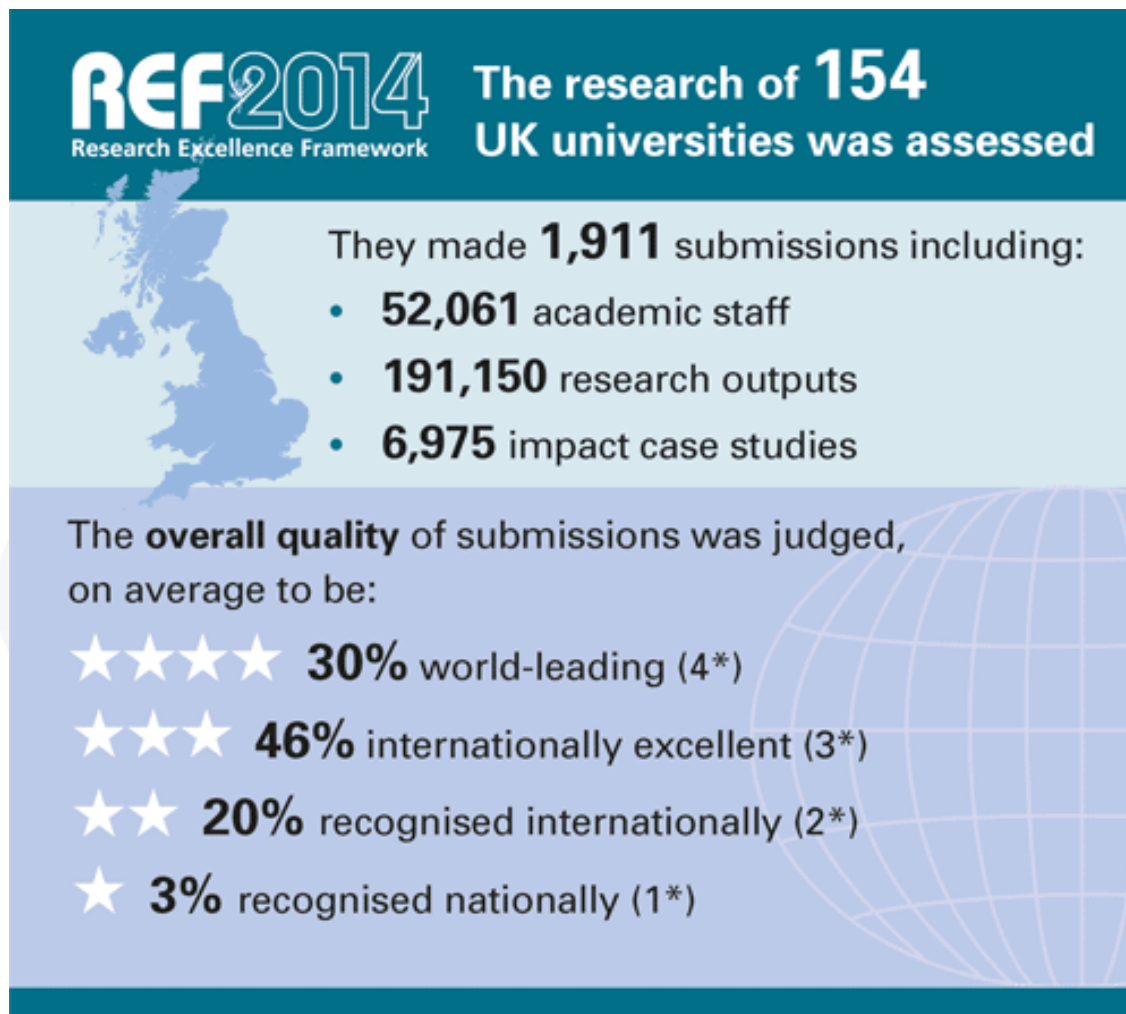
Unit of analysis

- Depends on the purpose of the evaluation
- Tendency to avoid the individual level – that produces erratic results and undermines institutional autonomy
- Some systems require tactical decisions about including individual researchers, with intra-institutional and personal consequences
- Decisions to include organisations are normally policy-based. You can't 'volunteer' (and usually you can't avoid assessment, either)

Recent Czech focus on outputs is unusual



Britské hodnocení REF 2014



Britské hodnocení REF 2014

METODIKA

The screenshot shows a web browser window displaying the REF2014 Research Excellence Framework website. The browser's address bar shows the URL: [results.ref.ac.uk/\(S\(utq0q3n35r1i1wfr3edo0r\)\)/Results/SelectUoa](https://results.ref.ac.uk/(S(utq0q3n35r1i1wfr3edo0r))/Results/SelectUoa). The website has a teal header with the REF2014 logo and a search bar. Below the header is a navigation menu with links: Publications, Results & submissions, Expert panels, Equality & diversity, About the REF, FAQs, and Contact. The main content area is titled 'View results by UOA' and includes a breadcrumb trail: Home » Results & submissions » Select UOA. On the left, there is a sidebar with links: Results & submissions, Introduction, and Analysis. The main content area lists 23 units of assessment (UOA) as a bulleted list:

- [1 - Clinical Medicine](#)
- [2 - Public Health, Health Services and Primary Care](#)
- [3 - Allied Health Professions, Dentistry, Nursing and Pharmacy](#)
- [4 - Psychology, Psychiatry and Neuroscience](#)
- [5 - Biological Sciences](#)
- [6 - Agriculture, Veterinary and Food Science](#)
- [7 - Earth Systems and Environmental Sciences](#)
- [8 - Chemistry](#)
- [9 - Physics](#)
- [10 - Mathematical Sciences](#)
- [11 - Computer Science and Informatics](#)
- [12 - Aeronautical, Mechanical, Chemical and Manufacturing Engineering](#)
- [13 - Electrical and Electronic Engineering, Metallurgy and Materials](#)
- [14 - Civil and Construction Engineering](#)
- [15 - General Engineering](#)
- [16 - Architecture, Built Environment and Planning](#)
- [17 - Geography, Environmental Studies and Archaeology](#)
- [18 - Economics and Econometrics](#)
- [19 - Business and Management Studies](#)
- [20 - Law](#)
- [21 - Politics and International Studies](#)
- [22 - Social Work and Social Policy](#)
- [23 - Sociology](#)

The Windows taskbar at the bottom shows the Start button and several open applications, including Total Commander, Calendar, and Microsoft Word. The system clock indicates the time is 6:26 PM on 6/26/2014.

Britské hodnocení REF 2014

METODIKA

Results &

submissions

18 - Economics and Econometrics

Introduction

 [UOA summary sheet](#)

Analysis

Overall profiles

Outputs

Impact

Environment

Select institution link
to view submission



Expand
all
profiles

% of the submission
meeting the standard for:

4* 3* 2* 1* U/C

4* 3* 2* 1* U/C

FTE

Category

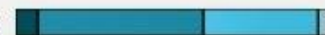
A staff
submitted

[Birkbeck College](#)



Overall

7 52 36 4 1



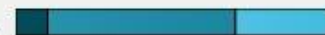
25.15

[University of
Birmingham](#)



Overall

10 59 30 1 0



24.20

[University of Bristol](#)



Overall

44 44 12 0 0



18.60

[Brunel University
London](#)



Overall

1 30 57 12 0



26.20

[University of Cambridge](#)



Overall

47 49 3 1 0



27.00

Outputs

54.5 39.4 5.1 1.0 0.0



Impact

50.0 50.0 0.0 0.0 0.0



Environment

12.5 87.5 0.0 0.0 0.0



Druhý pilíř současné Metodiky hodnocení 2013-2015



- Rozšíření počtu panelů.
- Cca 1 promile výsledků je třeba navýšit.
- Počet předkládaných výsledků za instituci by se musel vztáhnout k počtu jejích vědeckých pracovníků v rámci oboru.
- Na VŠ hodnotit na úrovni součástí: tedy fakult a ústavů.
- Výstupy hodnotit na absolutní škále kvality světovým standardů oborů.
- V panelech zvýšit zastoupení solidní vědců ze zahraničí.
- Honoráře panelistů zvýšit.
- Logistická podpora práce panelů profesionálního charakteru.
- Hodnocení dělat jen jednou za několik let.

Why use bibliometrics?

- **Comparatively low cost – and the cost is declining**
- **Need not involve work by those evaluated**
- **Can produce fine-grained ratings and rankings – though arguably this is spurious precision**
- **However**
 - Inter-field comparisons are problematic
 - Use of indicators is generally unsophisticated and sometimes problematic (for example, Journal Impact Factors)
 - Lack of context indicators and conventions for handling them
 - Hence, they provide a much more partial picture than peer review
 - Unhelpful for formative evaluation
- **Emerging conclusion: use ‘informed’ peer review**

Bibliometrie informuje o



- **Počtu a kvalitě publikací (v časopisech a sbornících; WoS/Scopus)**
- **Rozsahu a kvalitě akademického ohlasu (citace)**
- **Rozsahu a kvalitě spolupráce a mobility**
- **Mezioborovosti**



Bibliometrická opora inst. hodnocení (pracovní verze pro hodnocení AV ČR)

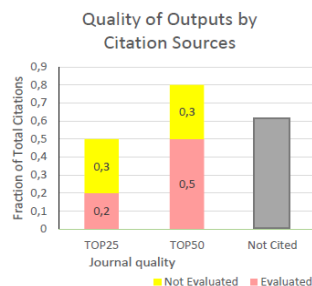
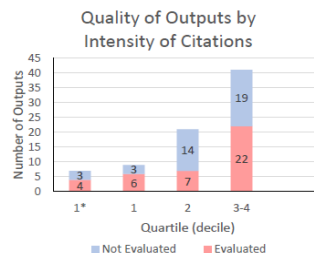
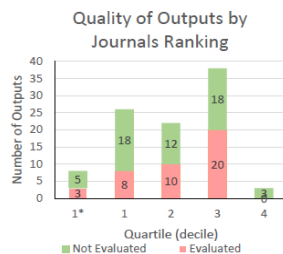
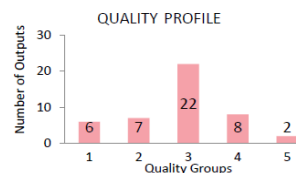


RESULTS OF THE I. STAGE OF EVALUATION AND BIBLIOMETRIC PARAMETERS

INSTITUTE: Institute of Mathematics AS CR, v.v.i.	TEAM: Team no. 1	HEAD: Jaroslav Novák
TOTAL NUMBER OF OUTPUTS: 24	EVALUATED OUTPUTS: 14	EVALUATED RESULTS: 3

QUALITY GROUPS OF OUTPUTS AND RESULTS

QUALITY	1	2	3	4	5
OUTPUTS	6	7	22	8	5
RESULTS	1	2	0	0	1



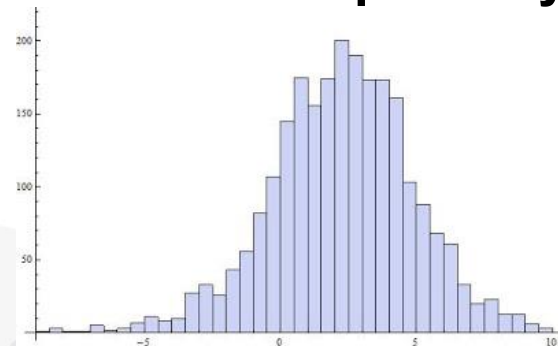
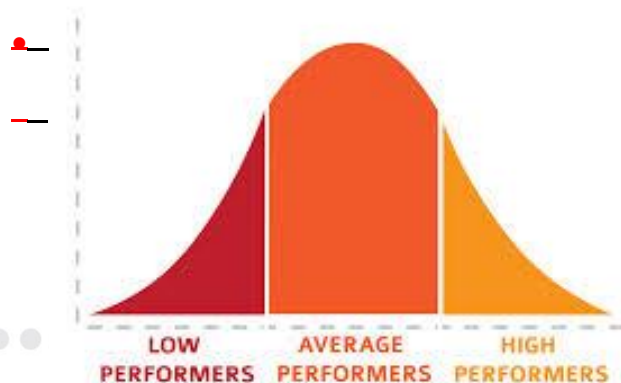
Field Structure of Outputs	Outputs (Evaluated)	Outputs (not Evaluated)	Average number of co-authors
Electrochemistry	4	6	7
Biochem. Mol.Biol	8	15	10
Mol.Oncology	4	7	6
Chem.Analytical	15	21	5
Chem. Physical	4	3	6

Quality Profile: number of outputs vs quality groups (5 groups) Quality of Outputs by Journals Ranking: number of outputs in quartiles (1-4) or top decile (1*) by AIS of journals. Quality of Outputs by Intensity of Citations: number of outputs in quartiles (1, 2, 3-4) or top decile (1*) of the list of outputs ordered by the number of citations. Quality of Outputs by Citation Sources: fraction of citations of all outputs in top quartile (TOP25) or top half (TOP50) of list of journals ordered by AIS; fraction of not cited outputs is added. Field Structure of Outputs: number of outputs of the team in different categories.

Zásady bibliometrického hodnocení



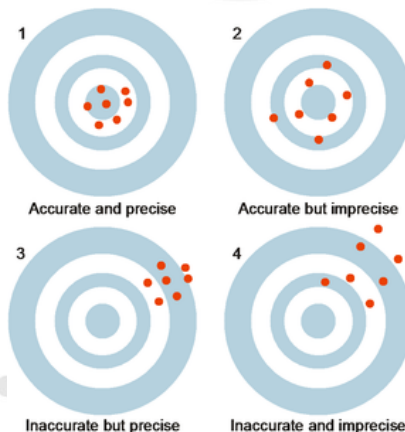
- **Vypovídací schopnost a věrohodnost podmiňuje**
 - kvalita vstupních dat (ověření a ověřitelnost)
 - adekvátní zohlednění oborových specifík
 - adekvátní benchmarking (srovnávat srovnatelné)
- **Ukazovat spíše výkon (tj. počty výstupů)**
 - ~~než produktivitu (výstupy/vstupy)~~
- **Sledovat spíše profil (distribuci) indikátorů a ne průměry**



Bibliometrie vs. peer-review výstupů



- **Kvalita a význam výsledků výzkumu je zatížena neurčitostí**
- stochastický (chybovost) element nelze vyloučit,
- ale lze ho redukovat počtem pozorování,
- ale pozorování jsou **drahá** → nutný kompromis
- **-> využití zákona velkých čísel a kombinace metod**
- **impaktní faktor čas** ~ renomé & náročnost peer-review ~ **odhad kvality/významu**
- **citovanost** ~ ohlas akademické komunity ~ **odhad kvality / významu**
- **dodatečné peer-review** ~ zlepšení odhadu (snížení míry chybovosti)



		Bibliometry (IF,...)			
		1	2	3	4
# Citations	1	A	A	U	U
	2	B	B	B	U
	3	U	C	C	C
	4	U	U	D	D

Bibliometrická opora inst. hodnocení

(pracovní verze pro hodnocení AV ČR)



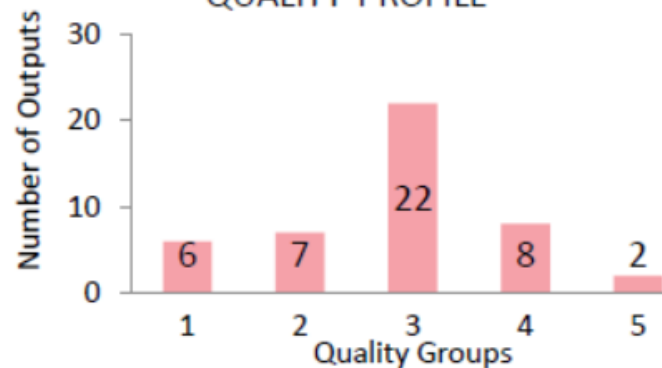
RESULTS OF THE I. STAGE OF EVALUATION AND BIBLIOMETRIC PARAMETERS

INSTITUTE: Institute of Mathematics AS CR, v.v.i.	TEAM: Team no. 1	HEAD: Jaroslav Novák
TOTAL NUMBER OF OUTPUTS: 24	EVALUATED OUTPUTS: 14	EVALUATED RESULTS: 3

QUALITY GROUPS OF OUTPUTS AND RESULTS

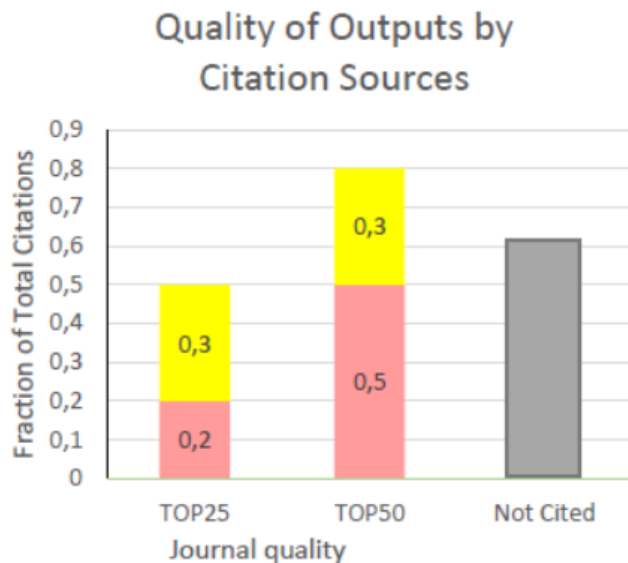
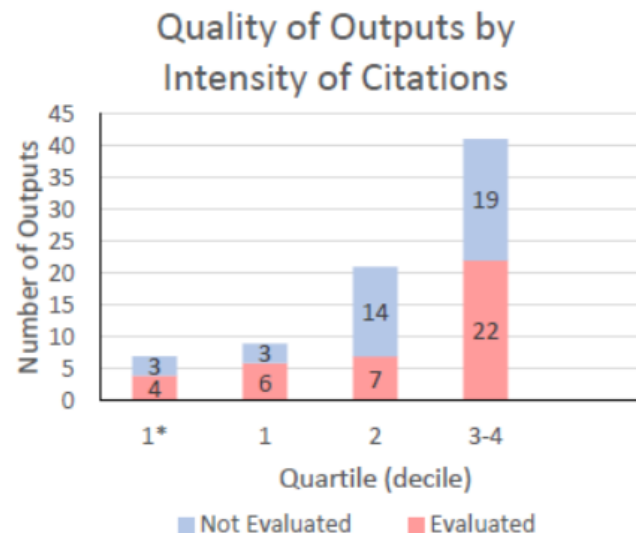
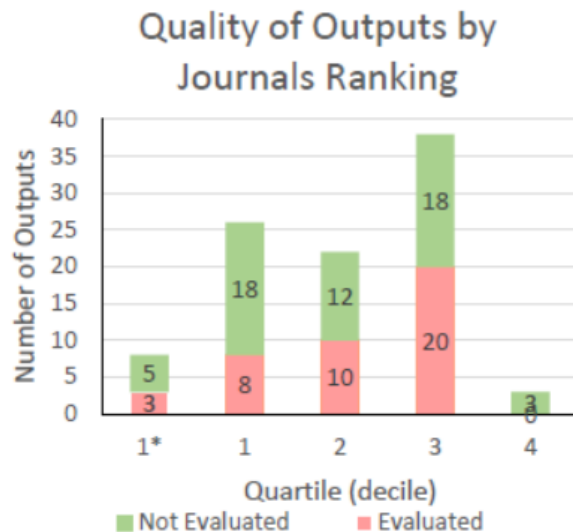
QUALITY	1	2	3	4	5
OUTPUTS	6	7	22	8	5
RESULTS	1	2	0	0	1

QUALITY PROFILE



Bibliometrická opora inst. hodnocení (pracovní verze pro hodnocení AV ČR)

METODIKA



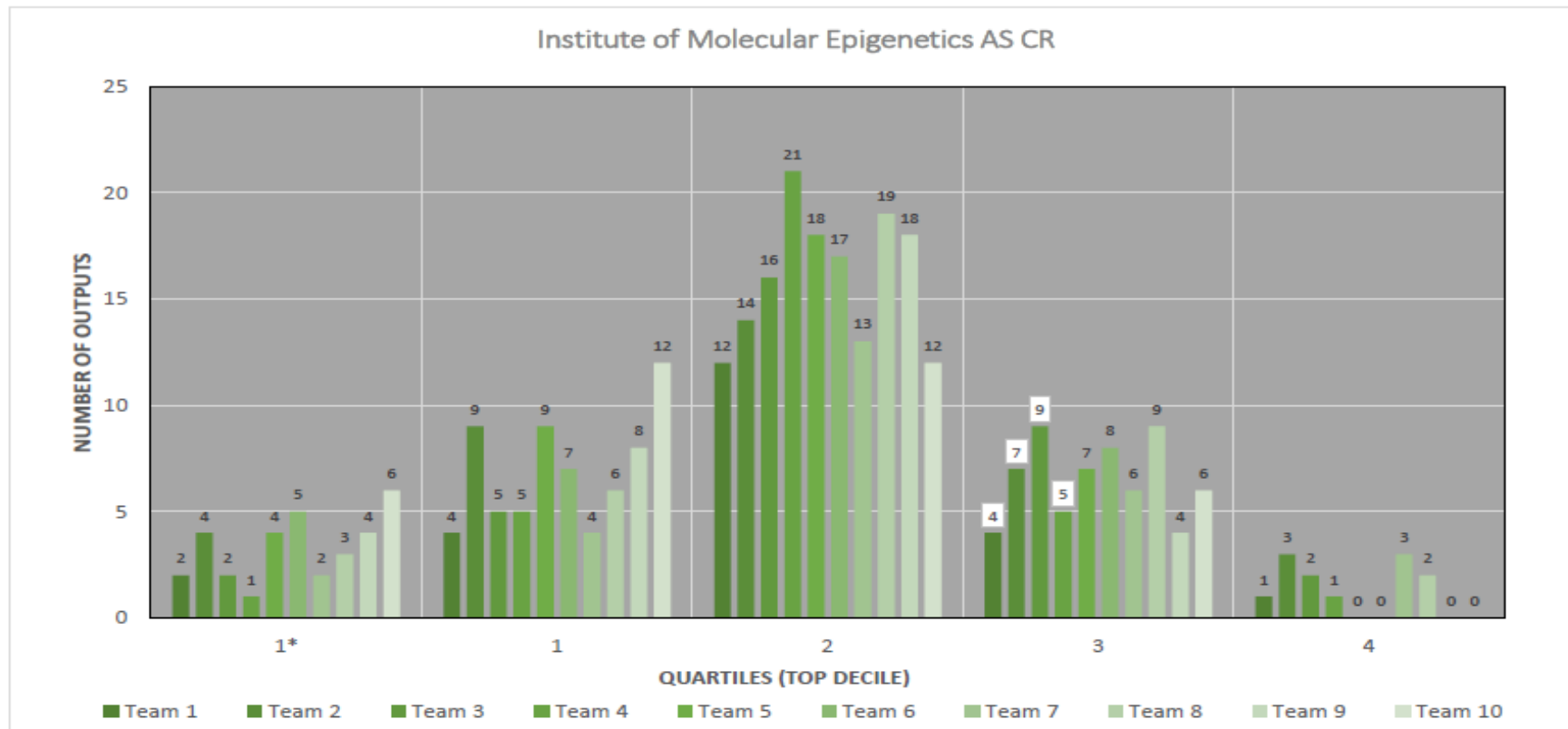
Field Structure of Outputs	Outputs (Evaluated)	Outputs (not Evaluated)	Average number of co-authors
Electrochemistry	4	6	7
Biochem. Mol.Biol	8	15	10
Mol.Oncology	4	7	6
Chem.Analytical	15	21	5
Chem. Physical	4	3	6

Příklad bibliometrické opory

inst. hodnocení (pracovní verze pro hodnocení AV ČR)

METODIKA

QUALITY OF OUTPUTS BY JOURNALS



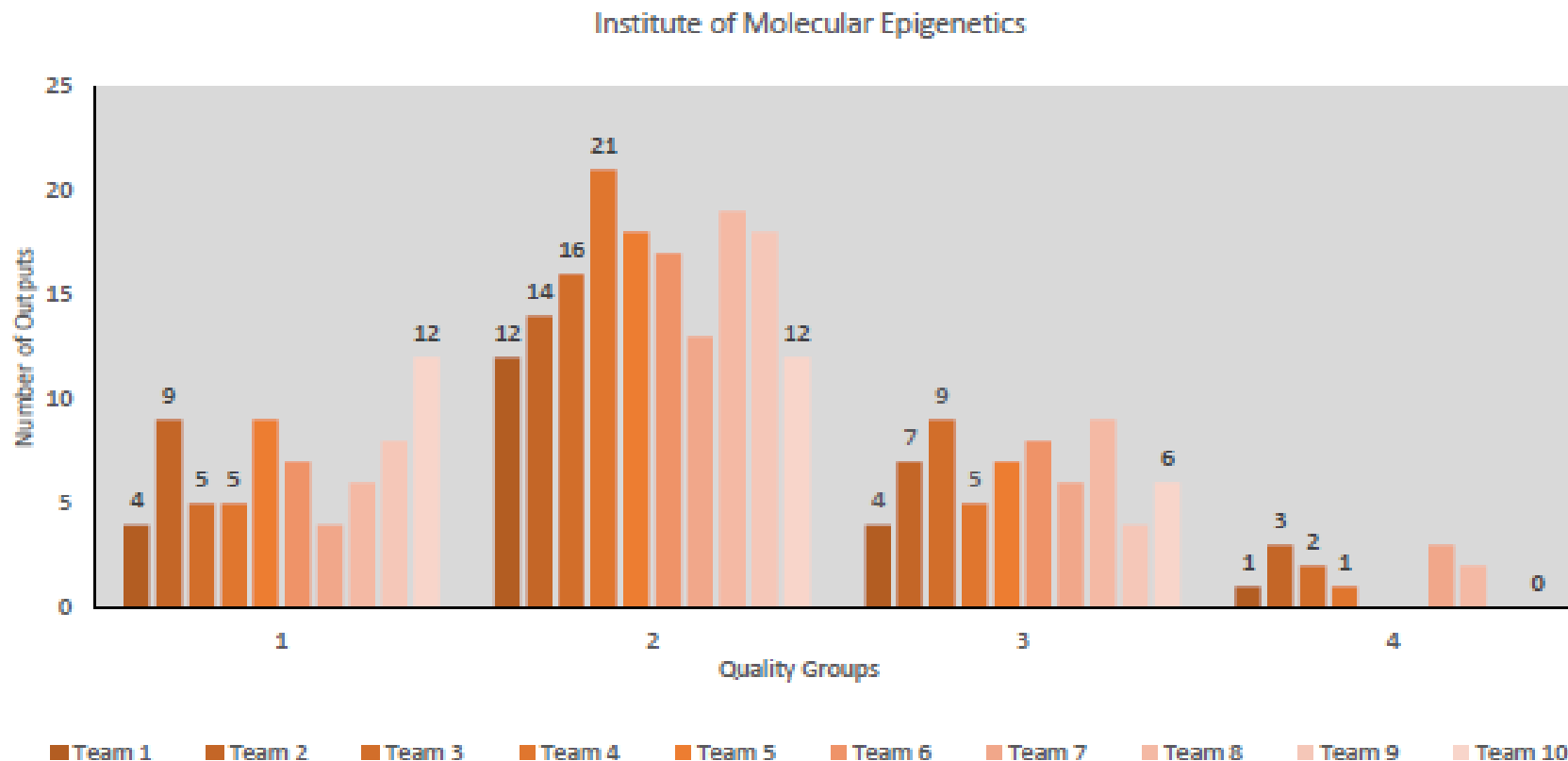
This plot is presented as an aggregate of bibliometric data for convenience of evaluators; the columns represent outputs (not productivity) and cannot be directly compared to each other.

Příklad bibliometrické opory

inst. hodnocení (pracovní verze pro hodnocení AV ČR)



QUALITY PROFILES



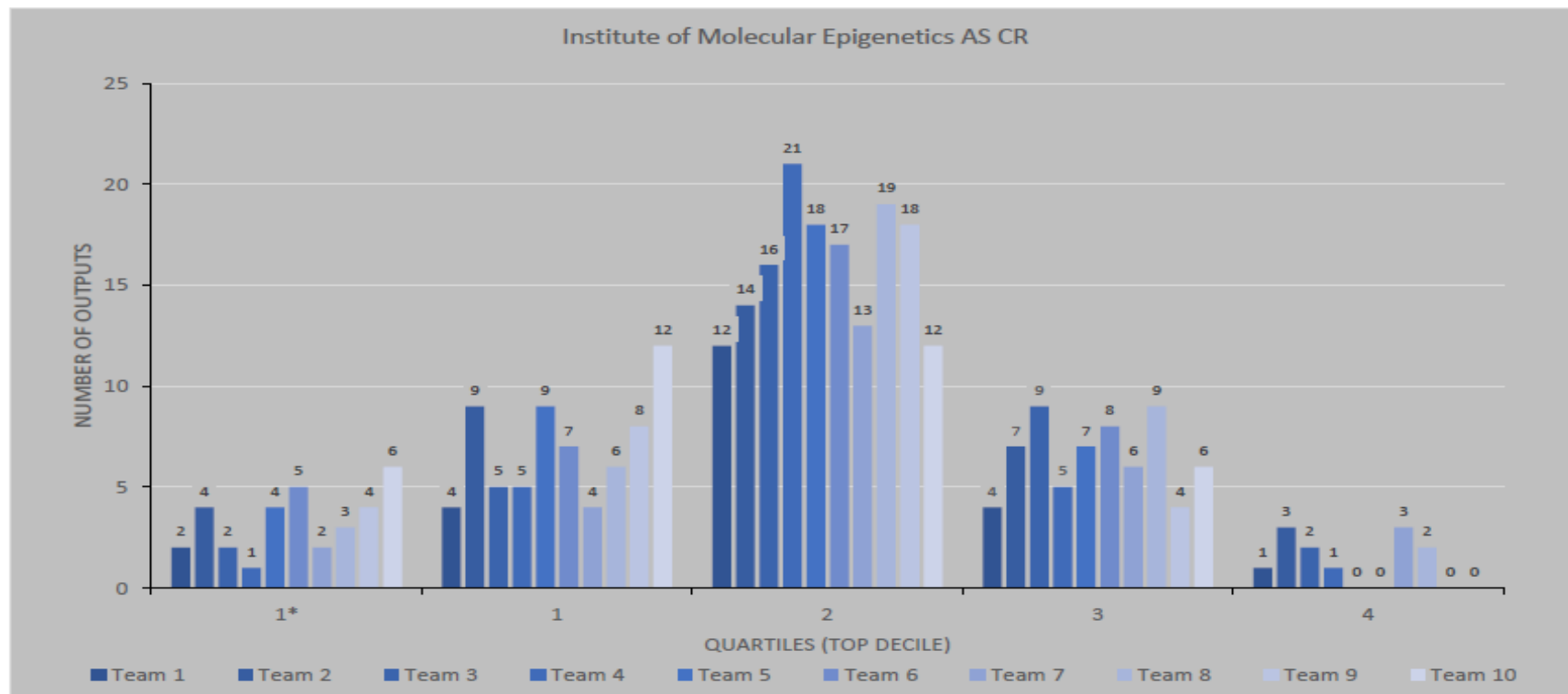
This plot is presented as an aggregate of data from the I. stage of evaluation for convenience of evaluators in the II. stage; the columns represent outputs (not productivity) and cannot be directly compared to each other.

Příklad bibliometrické opory

inst. hodnocení (pracovní verze pro hodnocení AV ČR)



QUALITY OF OUTPUTS BY INTENSITY OF CITATIONS



This plot is presented as an aggregate of bibliometric data for convenience of evaluators; the columns represent outputs (not productivity) and cannot be directly compared to each other.

Příklad bibliometrické opory

inst. hodnocení (pracovní verze pro hodnocení AV ČR)



Quality Profile: number of outputs vs quality groups (5 groups) Quality of Outputs by Journals Ranking: number of outputs in quartiles (1-4) or top decile (1*) by AIS of journals. Quality of Outputs by Intensity of Citations: number of outputs in quartiles (1, 2, 3-4) or top decile (1*) of the list of outputs ordered by the number of citations. Quality of Outputs by Citation Sources: fraction of citations of all outputs in top quartile (TOP25) or top half (TOP50) of list of journals ordered by AIS; fraction of not cited outputs is added. Field Structure of Outputs: number of outputs of the team in different categories.

Příklad bibliometrické opory



Studie 3/2012

Kde se v ČR dělá nejlepší výzkum¹

9.8. 2012 (DOPLNĚNO 20.12.2012)²

ŠTĚPÁN JURAJDA a DANIEL MÜNICH



CC-Organická chemie (58)

PRACOVISTĚ	Top	H	S	D	Celkem	Pořadí	Alt.pořadí
AVCR-Ustav organické chemie a ..	56	223	154	23	400	1	1
UPa-Fakulta chemicko-technolog..	14	36	36	23	95	4	3
UK-Prirodovedecká fakulta	12	41	18	8	67	3	4
VSCHT-Fakulta chemické technol..	9	51	36	28	115	2	2
MU-Prirodovedecká fakulta	8	27	15	9	51	5	6
AVCR-Ustav chemických procesu ..	3	20	13	6	39	6	7
UP-Prirodovedecká fakulta	1	17	14	21	52	7	5
UK-Farmaceutická fakulta v Hra..	1	11	11	5	27	8	9
VSCHT-Fakulta potravinářské a ..	1	5	7	2	14	9	10
AVCR-Mikrobiologický ústav AV ..	1	5	4	2	11	10	11
UTB-Fakulta technologická	0	4	9	20	33	11	8

CD-Makromolekulární chemie (30)

PRACOVISTĚ	Top	H	S	D	Celkem	Pořadí	Alt.pořadí
AVCR-Ustav makromolekulární ch..	66	234	114	41	389	1	1
UK-Prirodovedecká fakulta	8	24	9	1	34	3	4
UTB-Fakulta technologická	3	32	40	12	84	2	2
AVCR-Ustav fotoniky a elektron..	3	4	0	0	4	7	9
AVCR-Mikrobiologický ústav AV ..	3	3	0	0	3	9	10
VUT-Fakulta chemická	2	12	7	6	25	5	5
MU-Prirodovedecká fakulta	2	5	1	0	6	6	7
VSCHT-Fakulta chemické technol..	1	21	18	15	54	4	3
UPa-Fakulta chemicko-technolog..	0	3	6	2	11	8	6

Tabulka A5: Doplněk studie CERGE-EI IDEA *Kde se v ČR dělá nejlepší výzkum?*

Oborové výsledky podle počtu článků v horním decilu dle IF

Sloupec **Top**, podle kterého jsou doplněné tabulky řazeny, nově uvádí počet článků (v přepočtu za spoluautorství institucí a zaokrouhlený na celá čísla) v horní desetině časopisů dle IF. **Top** je tedy cílenější ukazatel excelence než původní sloupec **H**, který i nadále udává počet článku v impaktovaných časopisech WoS daného pracoviště v letech 2006-2010 a to časopisech v horní třetině v řazení podle výše IF časopisu v daném oboru. Sloupce **S** a **D** uvádějí počty článků v časopisech ve střední a dolní třetině dle IF. Sloupec **Celkem** sčítá počet všech článků sloupců **H**, **S**, **D**. Sloupec **Pořadí** připomíná pořadí pracovišť v řazení podle sloupce **H**. Sloupec **Alt.pořadí**, stejně jako v původní studii udává řazení pracovišť podle počtu všech článků (**Celkem**). Sloupec **Oborů** udává počet oborů, ve kterých se dané pracoviště nachází mezi prvními deseti v ČR při řazení podle **H**. Tabulky uvádí hodnoty pro typicky deset pracovišť s nejvyšším počtem **H** v oboru. Výsledky jsou uváděny pro obory podle dělení používaného v informačním systému RIV. Dvoupísmenné zkratky oborů odpovídají kódům RIV, viz www.vyzkum.cz. Dále v závorce za názvem oboru uvádíme počet pracovišť, která v daném oboru publikovala alespoň jeden článek v časopise s impakt faktorem. U společenskovedních oborů, tj. AD až AO, uvádíme v rámci poznámkového aparátu některé příklady sporných řazení do RIVu. Jde o ilustraci významné role expertního vhledu, který by měl bibliometrické údaje doprovázet. Nicméně expertní posouzení všech výsledků ve všech oborech jde již daleko za hranice možností této jednoduché přehledové práce.

Pozor na !



- **Malá čísla**
- **Sebecitovanost**
 - samotných autorů a spřízněnců
 - časopisů a jejich svazků (zvyšuje IF)
- **Publikace ve spřízněných časopisech a sbornících**
- **Velké megakolaborace = velké autorské týmy**
- **Vedlejší role spoluautorů**
- **Malé obory**
- **Obory nedostatečně pokryté WoS a Scopus**
-



METODIKA



www.metodika.reformy-msmt.cz



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ