

# NÁVRH OBECNÝCH ZÁSAD HODNOCENÍ PROGRAMŮ ÚČELOVÉ PODPORY VÝZKUMU, VÝVOJE A INOVACÍ A POTŘEBNÝCH SYSTÉMOVÝCH ZMĚN

**MARTIN SRHOLEC**

Hodnocení má poskytovat užitečné podklady pro rozhodování. Současný způsob hodnocení programů účelové podpory výzkumu, vývoje a inovací tuto úlohu neplní. Předložený dokument na základě nejlepší mezinárodní praxe navrhuje obecné zásady, podle kterých by se hodnocení mělo provádět. Po vymezení základních pojmů, východisek a konceptů hodnocení jsou diskutovány metodické otázky, je řešeno zajištění dat a rozebráno fungování systému hodnocení. Nastíněny jsou rovněž systémové změny potřebné pro dosažení mezinárodních standardů v českých podmínkách. Uvedené návrhy by měly přispět do diskuse o nadcházejících změnách rámce hodnocení. Otázky jejich praktické implementace budou řešeny v navazujících dokumentech.

**Návrh obecných zásad hodnocení programů účelové podpory  
výzkumu, vývoje a inovací a potřebných  
systémových změn**

Tento dokument vznikl v rámci Individuálního projektu národního (IPN) Metodika („Efektivní systém hodnocení a financování výzkumu, vývoje a inovací“), který byl realizován Ministerstvem školství, mládeže a tělovýchovy (MŠMT) v rámci Operačního programu Vzdělávání pro konkurenceschopnost (OP VK) a financován z Evropského sociálního fondu (ESF) a státního rozpočtu České republiky.

## Předmluva

Po přečtení tohoto dokumentu by měl čtenář získat přehled o obecných zásadách hodnocení veřejných programů účelové podpory výzkumu, vývoje a inovací a o systémových změnách potřebných pro jejich zavedení v českých podmínkách. Cílem není poskytnout podrobný návod, jak postupovat při hodnocení jednotlivých programů. Nejde o akční plán, co by se mělo udělat příští měsíc či do konce tohoto roku. Nabízíme obecnější rámec, který má sloužit jako vodítko, jak by se hodnocení mělo provádět. Otázky praktické implementace, tj. co přesně bude zavedeno, v jaké podobě, časové souslednosti, kdo za to bude odpovědný a co se zavede za rok, tři, pět či deset let, zůstávají na rozhodnutí odpovědných orgánů.

Předložený dokument je třeba chápat jako vymezení cílového stavu, od kterého je sice český systém hodně vzdálen, ale ke kterému je žádoucí začít viditelně směřovat, pokud se má hodnocení programů přiblížit nejlepší mezinárodní praxi. Někteří čtenáři zasvěcení do současného *modus operandi* české státní správy – v oblasti hodnocení nástrojů veřejné podpory zvláště – možná budou mít při čtení tohoto textu sklon vypočítávat důvody, proč to a to nelze uskutečnit. Takovým čtenářům doporučujeme raději uvažovat, jak by se tyto návrhy mohly začít implementovat. Nejedná se o žádné nevyzkoušené experimenty, nýbrž o v zahraničí dobře ověřená řešení.

Za příležitost tento dokument zpracovat a průběžné připomínkování děkuji prof. Ing. Jitce Moravcové, CSc. a doc. Ing. Danielu Münichovi, Ph.D. Za podrobné připomínky k první pracovní verzi zaslouží poděkování RNDr. Vladislav Čadil, Ph.D., prof. RNDr. Jan Hajič, Dr., Ing. Vladimír Kváča, Ph.D., Ing. Oto Potluka, Ph.D. a Ing. Vladimír Sodomka. Práce na tomto dokumentu proběhla od listopadu 2014 do září 2015. První verze prošla interní oponenturou v dubnu a květnu 2015. Revidovaný draft byl předmětem diskuse se zainteresovanými subjekty v červnu až srpnu 2015 a na semináři s experty poskytovatelů v září 2015. Konečná verze byla poskytnuta veřejně ke stažení v říjnu 2015.

## Shrnutí

Cílem tohoto dokumentu je vymezit zásady, jak by se mělo provádět hodnocení programů účelové podpory výzkumu, vývoje a inovací v českých podmínkách v návaznosti na nejlepší mezinárodní praxi. Zásady jsou navrženy v obecné rovině tak, aby mohly být zavedeny jako strategický nástroj pro řízení na úrovni systému a následně rozpracovány na konkrétní podmínky jednotlivých programů.

Předložené zásady poskytují vodítka pro provádění různých typů hodnocení, včetně ex-ante hodnocení návrhu programu, průběžného hodnocení chodu programu a ex-post hodnocení jeho dopadů, i pro různé druhy programů od grantů pro akademický výzkum přes podporu pro spolupráci ve výzkumu až po dotace zaměřené na inovace. Po vymezení základních pojmů, východisek a konceptů hodnocení jsou rozebrány metodické problémy, diskutováno zajištění odpovídající datové základny a nastíněno fungování systému hodnocení. Desatero, kterým by se zásady hodnocení měly řídit, lze shrnout do těchto bodů:

1. **Změna jako cíl.** Smyslem programu je dosažení změny, ke které by za jinak stejných podmínek nedošlo.
2. **Dopady nade vše.** Zajímají nás dopady, výsledky a výstupy v tomto pořadí důležitosti. Bez širších dopadů programů postrádá opodstatnění. Výstupy jsou pouze prostředkem k výsledkům a potažmo dopadům.
3. **Stříh na míru.** Každý program má vlastní metodiku hodnocení. Není přínosné se pokoušet vymezit jeden vzorec použitelný pro všechny programy.
4. **Kvalita důležitější kvantity.** Nemá smysl hodnotit za každou cenu. Jedno kvalitně provedené hodnocení má větší váhu než deset mizerných.
5. **Hodnocení není popis.** Hodnocení analyzuje přínosy podpory, podává normativní závěry, nekonstatuje pouze fakta. Není to monitorování, účetnictví, audit, statistika a ani výzkum.
6. **Důkazy, nikoliv názory.** Závěry hodnocení jsou založeny na prokázaných skutečnostech. Pro nepodložené spekulace není v hodnocení místo.
7. **Jaká data, takové hodnocení.** Hodnocení stojí a padá s kvalitou dat. Na datovou základnu pro hodnocení je nutné myslet již na prvopočátku programu.
8. **Sebehodnocení pro učení.** Poskytovatel hodnotí vlastní program, pokud je hlavní funkce hodnocení učící, pokud motivace pro jeho provedení vychází zevnitř.
9. **Externí hodnocení pro kontrolu.** Pokud je motivace pro hodnocení externí, pokud je ve hře rozdělování peněz mezi poskytovateli, hodnocení provádí externí subjekt.
10. **Naučit se učit se.** Bez zpětné vazby na rozhodování tvůrců politik a poskytovatelů je hodnocení zbytečné.

Na základě těchto zásad jsou navrženy systémové změny, které je potřeba provést, aby se úroveň hodnocení přiblížila k nejlepší mezinárodní praxi v oblastech nastavení rolí aktérů v systému, rozvoje lidských zdrojů, přístupu ke kvalitním mikrodátům pro hodnocení a financování ze státního rozpočtu.

# Obsah

|  |           |
|--|-----------|
| <b>1. Úvod</b>                             | <b>7</b>  |
| <b>2. Současná situace</b>                 | <b>9</b>  |
| <b>3. Nejlepší mezinárodní praxe</b>       | <b>10</b> |
| <b>4. Východiska hodnocení</b>             | <b>11</b> |
| 4.1 Formulace cílů                         | 11        |
| 4.2 Intervenční logika                     | 12        |
| 4.3 Metrika výkonnosti                     | 15        |
| <b>5. Koncept hodnocení</b>                | <b>16</b> |
| 5.1 Funkce hodnocení                       | 16        |
| 5.2 Použití hodnocení                      | 16        |
| 5.3 Typy hodnocení                         | 17        |
| 5.4 Předmět hodnocení                      | 19        |
| 5.5 Hodnoticí zpráva                       | 20        |
| <b>6. Metodika hodnocení</b>               | <b>22</b> |
| 6.1 Hodnoticí otázky                       | 23        |
| 6.2 Design hodnocení                       | 24        |
| 6.3 Kontrafaktuální přístup                | 25        |
| 6.4 Metody hodnocení                       | 26        |
| <b>7. Datová základna</b>                  | <b>31</b> |
| 7.1 Zdroje dat                             | 31        |
| 7.2 Individuální data                      | 32        |
| 7.3 Přístup k datům                        | 33        |
| <b>8. Systém hodnocení</b>                 | <b>34</b> |
| 8.1 Role aktérů                            | 34        |
| 8.2 Zpětná vazba                           | 37        |
| 8.3 Organizace průběhu                     | 39        |
| 8.4 Financování                            | 40        |
| 8.5 Režimy provádění                       | 41        |
| 8.6 Ukotvení v zákoně                      | 41        |
| <b>9. Nezávislost, nestrannost a etika</b> | <b>43</b> |
| 9.1 Nástrahy hodnocení                     | 43        |
| 9.2 Způsobilost hodnotitele                | 44        |
| 9.3 Nároky na zadavatele                   | 44        |
| 9.4 Hodnověrnost hodnocení                 | 45        |
| <b>10. Návrh systémových změn</b>          | <b>46</b> |
| 10.1 Nastavení systému                     | 46        |

|                      |    |
|----------------------|----|
| 10.2 Lidské zdroje   | 49 |
| 10.3 Datové zdroje   | 50 |
| 10.4 Finanční zdroje | 53 |
| Seznam zkratk        | 54 |
| Seznam literatury    | 55 |

## Seznam obrázků a tabulek

|  |    |
|--|----|
| Tab. 1 Komu a jak by měl tento dokument posloužit?   | 8  |
| Tab. 2 Zahraniční příručky, směrnice a manuály použité k vytvoření tohoto dokumentu                      | 10 |
| Obr. 1 Příklad schematického vyjádření modelu intervenční logiky programu na podporu výzkumu v podnicích | 14 |
| Obr. 2 Co se hodnotí a kdy?  | 18 |
| Tab. 3 Hodnocení programů podle typu   | 19 |
| Obr. 3 Kritéria hodnocení podle I-O-O-I modelu   | 24 |
| Tab. 4 Proveditelnost kvantitativního vyhodnocení kontrafaktuální situace                                | 29 |
| Tab. 5 Navrhované rozdělení rolí v českých podmínkách  | 48 |

# 1. Úvod

---

Veřejné výdajové programy jsou pod silicím tlakem na ospravedlňování rozpočtu, zvyšování výkonnosti a prokazování přínosů pro společnost. Hodnocení pomáhá nejen zlepšovat programy samotné, ale poskytuje i svědectví o jejich přínosech, které ilustrují smysluplnost vládních výdajů. K tomu, aby hodnocení účelně plnilo své funkce, je však nutné, aby se drželo určitých zásad, což vyžaduje nejenom politickou podporu, ale i porozumění napříč zainteresovanými subjekty. Cílem tohoto dokumentu je na základě nejlepší mezinárodní praxe navrhnout obecné zásady pro hodnocení programů a skupin grantových projektů (dále jen programy) účelové podpory výzkumu, vývoje a inovací (VVI) v českých podmínkách a načrtnout z nich plynoucí systémové změny.

Hodnocení (anebo evaluace) bylo zavedeno jako nástroj, jehož účelem je poskytovat poznatky pro posuzování vhodnosti veřejných výdajových programů. Hodnocení je podstatou jejich zdravé existence a rozvoje. Jedná se o široký pojem chápání, jehož význam se může výrazně lišit mezi lidmi a organizacemi. Například je často mylně zaměňováno s monitorováním, auditem anebo empirickým výzkumem. Může to být způsobeno tím, že hodnocení zastřešuje rozličnou škálu aktivit, procesů a metod, které si mimo jiné vypůjčují od výše uvedených. Hodnocení může být v podstatě kdokoli a na něm potažmo cokoliv, avšak ne jakkoliv.

Konkrétně v oblasti VVI lze uvažovat o hodnocení projektů, programů, nástrojů, politik, jednotlivců, institucí anebo systému jako celku. I když řada zde prezentovaných zásad je obecných, a tudíž platí bez ohledu na to, co je přesně hodnoceno, pro dosažení nejlepší funkčnosti vyžaduje každá z těchto úrovní hodnocení vlastní postupy. Z tohoto důvodu je nutné mít na paměti, že tento dokument se týká pouze hodnocení programů účelové podpory. Některé níže uvedené zásady nemusí být například vhodné pro hodnocení na úrovni projektů anebo institucí, avšak pro hodnocení programů účelové podpory mohou být zásadní.

Programy účelové podpory VVI zahrnují všechny přímé nástroje financování prostřednictvím veřejných soutěží. Hlavní zaměření tohoto dokumentu je na programy primárně financované z českého státního rozpočtu, a tudíž pod dohledem vlády České republiky, nikoliv na programy financované ze zahraničí, například z fondů EU, ačkoliv prezentované zásady lze snadno zobecnit i pro tuto oblast. Mezi další nástroje veřejné podpory VVI, které mohou mít podobné cíle, ale tento dokument se jich rovněž přímo netýká, patří zejména nepřímá podpora v podobě daňových odpočtů, garance za úvěry, veřejné zakázky anebo majetkové vstupy.

Hlavním znakem, který odlišuje účelovou podporu VVI od vládních výdajových programů v jiných sektorech, je její opodstatnění na základě tržních a systémových nedokonalostí, které jsou neoddělitelně spjata s procesem vytváření, přenášení a využívání znalostí. Prostor pro vládní zásady v této oblasti tkví především v tom, že znalosti vykazují charakteristiky veřejných statků a že tyto procesy probíhají v rámci velmi komplexního politického, tržního a inovačního systému, jehož fungování je neustále zevnitř i z vnějšku ovlivňováno zásahy rozličných aktérů, a to nejenom z veřejné sféry, kteří se nemusí vždy chovat v zájmu systému jako celku. Míra této komplexnosti ve světle změn, jimiž český inovační systém prochází, představuje velkou výzvu i pro hodnocení.

Avšak tím spíše je třeba zjišťovat, zda programy účelové podpory VVI mají smysl, jestli jejich poskytování stojí za peníze daňových poplatníků a jaké mají v konečném důsledku přínosy pro společnost. Dokládat užitečnost těchto programů vyžaduje nejen rostoucí tlak ze strany vlády a potažmo veřejnosti, ale i stále se zpřísňující požadavky ze strany EU. Neméně důležitou přidanou hodnotou hodnocení je učení se z předchozích nezdarů a úspěchů, a tudíž jejich promítnutí

do rozhodování o budoucích intervencích. Kvalitní hodnocení, které se provádí systematicky, spravedlivě a dlouhodobě, je toho nezbytným předpokladem.

Čím více poznatků z oblasti hodnocení VVI programů je veřejně diskutovaných, přijímaných a obecně uznávaných, tím lepších výsledků je možno dosahovat. Znalost principů hodnocení je nepostradatelná zejména pro všechny, kdo se podílejí na řízení VVI systému. Předkládané návrhy by tudíž měly v první řadě sloužit Radě pro výzkum, vývoj a inovace (RVVI), sekci pro VVI při Úřadu vlády ČR a poskytovatelům při zlepšování rámce hodnocení, ale mohou být rovněž užitečné pro případné hodnotitele jako zrcadlo toho, jak by se hodnocení mělo provádět.

Tab. 1 Komu a jak by měl tento dokument posloužit?

|                                       |  |
|---------------------------------------|--|
| <b>Tvůrci politik a poskytovatelé</b> | <ul style="list-style-type: none"><li>• Při tvorbě nových výdajových programů.</li><li>• Při formulaci zadávací dokumentace hodnocení.</li><li>• Při výběru hodnotitelů.</li><li>• Při plánování práce ve vztahu k hodnocení.</li><li>• Při zlepšování řízení programů.</li><li>• Při implementaci závěrů hodnocení.</li><li>• Při spolupráci s hodnotiteli.</li></ul> |
| <b>Hodnotitelé</b>                    | <ul style="list-style-type: none"><li>• Při zpracování návrhů na projekty hodnocení.</li><li>• Při řešení otázek ohledně metod a dat.</li><li>• Při komunikaci se zainteresovanými subjekty.</li></ul>   |



## 2. Současná situace

---

Stávající hodnocení programů účelové podpory VVI se omezuje na monitorování vstupů a výstupů, jejich mechanické počítání a prezentaci popisných přehledů. Nedokáže postihnout přínosy programů, a tudíž ani není schopno ověřit, do jaké míry dochází k naplňování jejich cílů. Souhrnné vyhodnocení výsledků ukončených programů, které vychází pod hlavičkou RVVI, je toho nejlepším příkladem (RVVI 2013b). Jedná se o komentovanou statistiku, která má hodnocením v pravém slova smyslu málo společného. Skutečné hodnocení začíná teprve tam, kde nynější končí.

Na tento nedostatek upozorňuje řada strategických vládních dokumentů včetně Národní politiky výzkumu, vývoje a inovací České republiky na léta 2009–2015 (RVVI 2009), její aktualizace (RVVI 2013a) anebo Národní inovační strategie České republiky (MPO 2011). Důrazně na to poukázal i Arnold (2011), tj. Mezinárodní audit systému VVI v České republice, podle kterého není stávající systém hodnocení programů užitečným nástrojem pro jejich řízení, protože:

- *Systém hodnocení programů VVI je spíše jen monitorováním dosažených výstupů a vstupů potřebných k jejich dosažení.*
- *Zprávy z hodnocení jednotlivých ukončených programů VVI nelze považovat za „standardní“ hodnotící zprávy, protože obsahují jen popis ukončených programů a jejich výsledků.*
- *Za velice diskutabilní je považována spolehlivost uváděných analýz, přičemž používaná data většinou pocházejí pouze z IS VaVal.*
- *Ačkoliv zprávy o hodnocení uvádějí jasné závěry, většina z nich není dostatečně a spolehlivě doložena.*
- *Doporučení pro tvorbu navazujících či jiných programů VVI většinou nejsou ve zprávách obsažena.*
- *Na rozdíl od vyspělých zemí EU není v ČR hodnocení programů považováno za důležitou součást procesu přípravy a realizace politik v oblasti VVI.*

Čadil (2014) dokládá, že od té doby se toho bohužel příliš nezměnilo. Byla sice vytvořena revidovaná Metodika 2013 (RVVI 2013c), která však průlom v hodnocení programů nepřinesla. Nejvýznamnějším posunem je práce pracovní skupiny při RVVI k systémovému opatření č. 18 aktualizované národní politiky VVI (RVVI 2013a), která v průběhu roku 2014 připravila dokument „Základní principy přípravy a hodnocení programů a skupin grantových projektů výzkumu, vývoje a inovací“ (RVVI 2015). Následně se tento materiál stal přílohou Usnesení vlády č. 351 ze dne 13. května 2015, které ukládá vypracovat do roku 2016 obecně závazný dokument upravující přípravu a hodnocení programů. S obsahem těchto principů nelze než souhlasit, nicméně se jedná o nejmenší společný jmenovatel standardů, na kterých mohla být dosažena široká shoda, a jejichž plnění je možné v rámci stávajícího systému hodnocení. Úkolem tohoto textu je tyto minimální standardy dále rozvinout a nastínit potřebné systémové změny.

Hodnocení národních programů účelové podpory VVI dnes již výrazně zaostává oproti jiným oblastem, v nichž úroveň hodnocení v posledních letech zřetelně pokročila kupředu, jako například v oblasti fondů EU (MMR 2014), kde jsou postupně zaváděny principy z nejlepší mezinárodní praxe, a z jejichž zkušeností je tudíž nasnadě se učit. Ke zlepšení české kultury hodnocení rovněž přispívá činnost České evaluační společnosti (ČES), jejíž materiály využívají i zadavatelé hodnocení ve výše zmíněné oblasti a z nichž rovněž čerpáme (ČES 2011 a 2013).

### 3. Nejlepší mezinárodní praxe

K zemím s dlouholetou tradicí vysoké úrovně hodnocení účelových programů podpory pro VVI, od kterých je radno se učit, se řadí například severské země Dánsko, Finsko, Norsko a Švédsko, ale i sousední Německo s Rakouskem a tradičně Nizozemí, Velká Británie a Spojené státy americké. Pozoruhodný je zejména pokrok v Rakousku, které bylo před dvaceti lety v podobné situaci a které je nám blízké nejenom z hlediska historie, kultury a velikosti, ale i strukturou ekonomiky. Užitečná jsou i doporučení mezinárodních institucí a asociací jako Evropská komise, OECD a European Network of Innovation Agencies (TAFTIE). Konkrétně byly využity zejména průřezové texty, tj. obecné příručky, směrnice a manuály pro hodnocení, jejichž přehled je uveden v tabulce 2. Použitelná literatura je rozsáhlá, takže se nejedná o vyčerpávající seznam, ale pouze o úzký výčet nejužitečnějších zdrojů v anglickém jazyce.

Tab. 2 Zahraniční příručky, směrnice a manuály použité k vytvoření tohoto dokumentu

| ODKAZ                       | ZEMĚ           | ORGANIZACE   | NÁZEV   |
|-----------------------------|----------------|--|---|
| BOEKHOLT, P. et al. (2014)  | Nizozemí       | Technopolis Group  | Evaluation Reference Model (for TAFTIE's Taskforce Benchmarking Impact, Effectiveness and Efficiency of Innovation Instruments) |
| CPB (2012)                  | Nizozemí       | Netherlands Bureau for Economics Policy Analysis                         | Dare to measure: Evaluation designs for industrial policy in the Netherlands  |
| Evropská komise (2013a)     | EU             | Evropská komise  | Evaluation in the Field of State Aid: Methodology Guidance Paper  |
| Evropská komise (2013b)     | EU             | Evropská komise  | EVALSED: The resource for the evaluation of Socio-Economic Development  |
| ESF (2012)                  | EU             | European Science Foundation  | Evaluation in Research and Research Funding Organisations: European Practices   |
| FTEVAL (2012)               | Rakousko       | Austrian Platform for Research and Technology Policy Evaluation (FTEVAL) | Research and Technology Policy Evaluation, Evaluation Standards in Research and Technology Policy                               |
| GUTHRIE, S. et al. (2013)   | USA            | RAND Corporation   | Measuring research, a guide to research evaluation frameworks and tools   |
| MloIR (2014)                | Velká Británie | Manchester Institute of Innovation Research (MloIR)                      | Compendium of Evidence on Innovation Policy   |
| OECD (1991)                 | OECD           | OECD   | DAC Principles for Evaluation of Development Assistance   |
| Research Councils UK (2011) | Velká Británie | Research Councils UK   | Evaluation: Practical Guidelines. a guide for evaluating public engagement activities   |
| VINNOVA (2008)              | Švédsko        | VINNOVA  | VINNOVA's Focus on Impact – a Joint Approach for Logic Assessment, Monitoring, Evaluation and Impact Analysis                   |

## 4. Východiska hodnocení

---

Program účelové podpory je nástroj implementace národní politiky VVI. Je to soubor vzájemně se podporujících opatření, která jsou společně zaměřená na dosažení jednoho anebo více konkrétních a předem vymezených cílů. Program má obvykle časově omezené trvání, svůj vlastní rozpočet a jediného poskytovatele. Přímá podpora je příjemcům poskytována formou grantů a dotací, které jsou rozdělovány prostřednictvím veřejné soutěže. Program se skládá z řady projektů. Podobné programy tvoří nástroj podpory.

Návrh programu je základním podkladem pro rozhodování tvůrce politik a v konečném důsledku vlády o tom, jestli bude program financován. Z pohledu hodnocení by měl návrh programu přinejmenším obsahovat:

- hlavní a dílčí cíle,
- model intervenční logiky,
- metriku výkonnosti,
- rozpočet se samostatnou položkou na hodnocení,
- plán hodnocení.

Plán hodnocení by měl již v návrhu projektu specifikovat požadavky na data a základní představu o metodickém mixu, který bude v budoucnu použit. Před spuštěním programu je třeba vědět, jaká data pro použití těchto analytických metod jsou k dispozici z existujících zdrojů, jaká data bude třeba sledovat v průběhu implementace programu skrze monitorování a nejlépe i jaká data bude nutné získat prostřednictvím sběru primárních dat, včetně údajů za srovnávací skupinu. Je potřebné o tom mít představu před schválením programu, aby se na data pro hodnocení mohly vyčlenit odpovídající prostředky v jeho plánovaném rozpočtu.

Plán dalšího hodnocení (průběžného, ex-post a dopadů) bude sám o sobě předmětem ex-ante hodnocení návrhu programu. Navrhovatel tím bude nucen o průběhu hodnocení přemýšlet již od okamžiku vypracování konceptu programu, a tudíž že nebude moci tyto otázky odsunout na později. Při ex-ante hodnocení by měl být tento plán brán jako samostatná položka k vyhodnocení s vlastní hodnoticí stupnicí a především s nutností překonat určitou minimální úroveň k tomu, aby mohl být program jako celek doporučen k financování. Špatně navržený plán hodnocení by bez ohledu na cokoliv dalšího měl být důvodem pro vrácení návrhu programu k přepracování.

### 4.1 FORMULACE CÍLŮ

Základním předpokladem pro zhodnocení programu je jasné vymezení jeho cílů. Již v návrhu programu musí být uvedeno, v čem bude spočívat úspěch daného programu a jakým způsobem to bude možné sledovat. Konkrétní cíle v podobě výstupů, výsledků a především dopadů se následně odvodí z této obecné představy o úspěšném programu. Cíle je třeba formulovat obezřetně, protože na nich závisí shoda na tom, čeho má být v principu dosaženo, na jejich obsahu je do značné míry založeno rozhodnutí o financování programu a v neposlední řadě se od nich bude odvíjet hodnocení. Pokud jsou cíle vágní, nekonkrétní a neověřitelné, nemá se na jedné straně hodnotitel o co opřít a na druhé straně se nemá poskytovatel na co odvolat, pokud by mu hodnocení nepřipadalo spravedlivé.

Smyslem vládního výdajového programu je dosažení změny, například v podobě otevření nových možností rozvoje, zlepšení neuspokojivého stavu či napravení trhnic a systémových selhání, ke kterému by jinak nedošlo. Cíle programu tudíž musí být vymezeny dynamicky ve srovnání s výchozím stavem, nikoliv ve statickém smyslu pouze jako stav, kterého má být dosaženo. Z toho plyne, že tento nevyhovující stav a jeho žádoucí změna k lepšímu musí být co možná

nejkonkrétnějším způsobem popsány, a to pokud možno pomocí kvantitativních indikátorů, popřípadě s odkazem na důvěryhodné zdroje. Pokud to lze, je třeba stanovit vstupní hodnoty indikátorů, které poslouží pro určení změny.

Při formulaci cílů je třeba klást si následující otázky:

- Jaká je mise programu? Jaká je motivace pro jeho zavedení? Čeho chceme podporou dosáhnout?
- Jaké jsou obecné cíle programu? Jak lze konkrétně vymezit změnu, jejímž dosažením lze tento veřejný výdaj opodstatnit?
- Jaké dílčí cíle můžeme vyvodit s odkazem na toto obecné zadání? Jak je možné je operacionalizovat? Je smysluplné tyto dílčí cíle kvantifikovat? Pokud nikoliv, je možné nastavit procedury, které umožní přezkoumat jejich splnění kvalitativním způsobem?
- Skrze jaké mechanismy mají být dosaženy zamýšlené přínosy? Na základě jakých předpokladů, například ohledně chování příjemců, vývoje prostředí či systému jako celku, je fungování těchto mechanismů založeno? Kdy, kde, jak a na koho budou působit?

Na základě zodpovězení těchto otázek, což by mělo být jednou z klíčových částí návrhu programu, se vytvoří obecné a dílčí cíle, z nichž následně přímo vychází kritéria pro hodnocení programu. Obecné a dílčí cíle by na sebe měly navazovat a společně tvořit vnitřně logicky konzistentní celek. Pokud je cílů větší množství, je žádoucí uvést, které z nich jsou prioritní. Nicméně návrh programu by se měl v každém případě vyvarovat zahlcení příliš velkým množstvím komplexních cílů.

Splnění cílů by mělo být ověřitelné. Žádoucí je pokud možno kvantifikovat cíle v podobě indikátorů. Nicméně je třeba zdůraznit, že počítání „bodů“ za bezprostřední výstupy nesmí přehlušit otázku, jaká je hlavní podstata daného vládního zásahu, protože se často jedná pouze o nepřímé proxy indikátory hlavních efektů, kterých má být dosaženo. Nelze přeskakovat vysvětlení hlubších důvodů, proč je program zaváděn, jaké změny má být konkrétně dosaženo a proč právě tímto způsobem, jakkoliv to navrhovatelé může připadat zřejmé a měřitelné cíle k tomu mohou svádět.

Snažit se za každou cenu kvantifikovat cíle v podobě měřitelných indikátorů rovněž není vždy ku prospěchu věci. Některé aspekty úspěchu programu nemusí být ze své podstaty měřitelné pomocí tvrdých dat. Lepší než používat špatné proxy indikátory je nechat zhodnocení splnění obtížně měřitelných cílů na kvalitativních metodách. Je na tvůrci politik, aby při projednávání návrhu programu učinil politické rozhodnutí, zda je takový postup oprávněný. Pokud jsou přínosy kvalitativní, není to ještě důvod zamítnout program, který by takové latentní problémy řešil.

## 4.2 INTERVENČNÍ LOGIKA

Model intervenční logiky vysvětluje principy působení programu. Jeho účelem je vysvětlit, jak by mělo být s navrženými vstupy dosaženo zamýšlených cílů. Uvádí, na dosažení konkrétně jaké změny se program zaměřuje, upřesňuje mechanismy, prostřednictvím kterých k tomu má dojít, a zřetelně vymezuje, na jakých předpokladech je působení těchto mechanismů založeno.

Rozlišovat je třeba mezi následujícími prvky:

- Vstupy – finanční, lidské a další zdroje použité v programu (výdaje na výzkum, počet výzkumníků či nákup nového přístroje).
- Výstupy – přímé výstupy aktivit podporovaných s použitím vstupů programu (publikace, patenty, prototypy apod.).

- Výsledky – bezprostřední přínosy podpory pro příjemce (renomé špičkového pracoviště, rozběhnutí nového směru výzkumu či uvedení inovovaného produktu na trh).
- Dopady – širší přínosy programu, a to nejen pro příjemce, ale i přínosy pro společnost jako celek (nižší úmrtnost na rakovinu, tvorba nových pracovních míst či vyšší produktivita práce).

Model intervenční logiky v návaznosti na deklarované cíle programu ozřejmuje jeho vstupy, výstupy, výsledky a dopady a především vysvětluje očekávané příčinné souvislosti mezi nimi, tj. jak je zamýšleno, že budou ze vstupů tvořeny výstupy, jak budou výstupy přeměněny ve výsledky a jakým způsobem mají mít výsledky za následek dopady. Za vysvětlení těchto vztahů se nepovažují obecné fráze, nýbrž konkrétní tvrzení podložená výsledky renomovaných studií, včetně existujících hodnocení podobně zaměřených programů u nás či v zahraničí, které je prokazují.

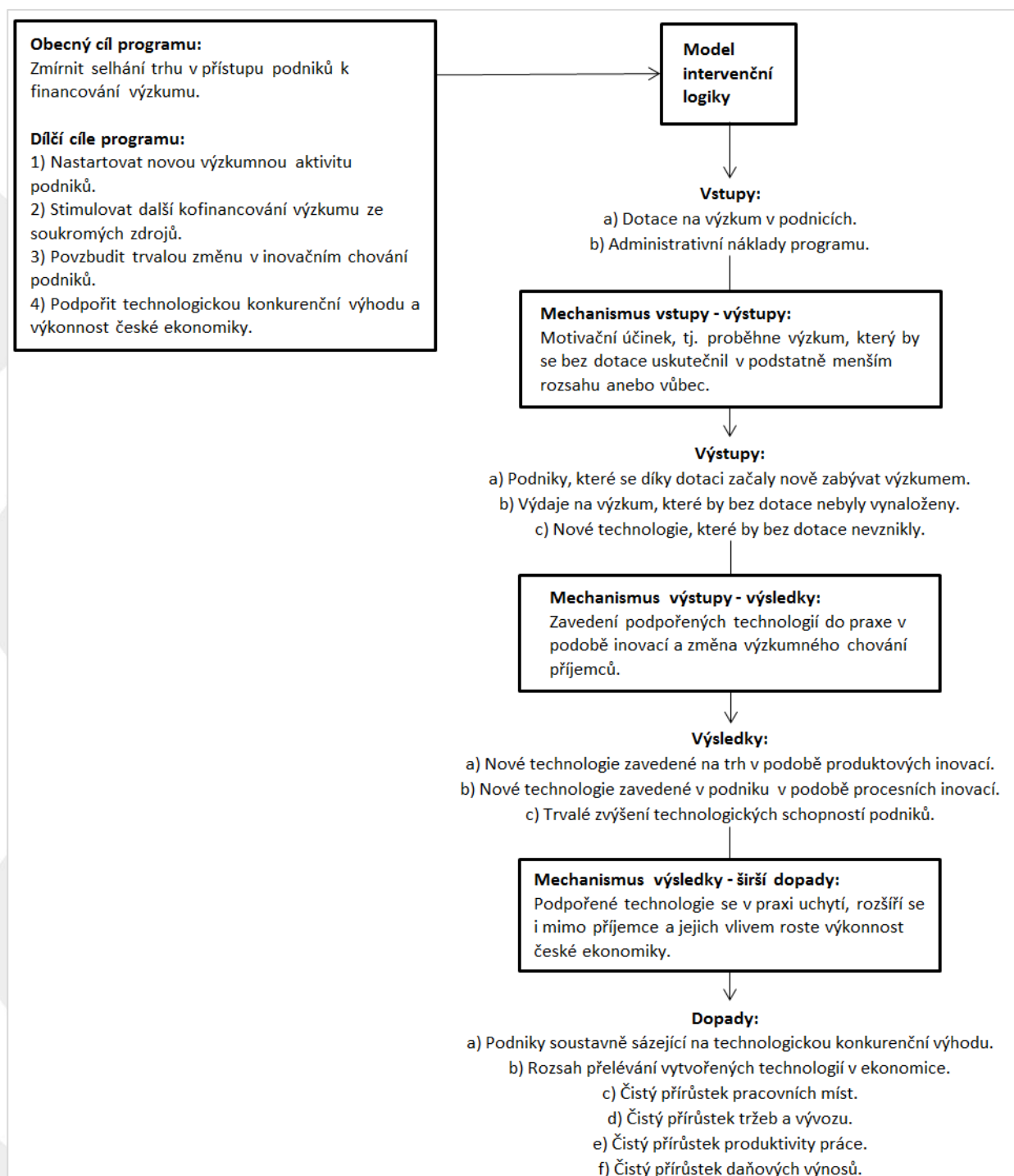
Zároveň platí, že důležitost těchto kategorií, a tudíž závažnost argumentů o příčinných souvislostech mezi nimi, pro hodnocení programu roste shora dolů. Konečné dopady, tj. široké přínosy pro společnost nad rámec prospěchu jednotlivých příjemců podpory, jsou důležitější než výsledky zatímco dosažené výsledky mají větší váhu než počty vytvořených výstupů. Každý program z podstaty financuje nějaké vstupy, jež pravděpodobně povedou k vytvoření určitých výstupů, ale to samo o sobě ještě neznačí úspěch. Až výsledky a především dopady mohou dosáhnout změn, na jejichž základě je zavedení programu odůvodněno.

Zdůraznit je třeba rozdíl mezi tím, co jsme schopni v daném okamžiku již přímo spočítat na jedné straně, a mezi tím, na co máme při hodnocení klást důraz, co máme mít v první řadě na zřeteli na straně druhé. Požadavek, aby konečné dopady podpory byly pro hodnocení nejdůležitější, nelze smést ze stolu tvrzením, že jsou příliš daleko. Je pravda, že skutečný rozsah širších dopadů můžeme obvykle určit až několik let po skončení programu. Nicméně to ještě neznamená, že na ně nemáme myslet už při ex-ante, průběžném a ex-post hodnocení. Širší dopady i mimo samotné příjemce jsou hlavní důvod, proč programy vůbec zavádíme. Bez nich nelze programy ospravedlnit, a tudíž je třeba mít širší dopady neustále na paměti.

Obrázek 1 nabízí příklad grafického vyjádření základní kostry intervenční logiky programu na podporu výzkumu v podnicích. Uvedeny by měly být odkazy na studie prokazující, že podnikový výzkum je podfinancován ze soukromých zdrojů. Doloženo by mělo být tvrzení, že výzkum je pro podniky zásadním zdrojem konkurenční výhody, že inovace založené na vlastním výzkumu mají vysokou návratnost a že nové technologie se časem přelévají i do jiných místních podniků, z čehož má prospěch ekonomika jako celek. Přidány by měly být odkazy na hodnocení ukazující, že dopady takové podpory jsou dlouhodobé. Při navrhování programu je nutné uvedené souvislosti podrobně vysvětlit. Za model intervenční logiky nelze považovat obecná tvrzení o zvyšování konkurenceschopnosti.

Bez dobře zpracované intervenční logiky by program neměl být financován, a to mimo jiné z toho důvodu, že bez explicitně vyjádřené logiky jeho fungování je v podstatě nemožné provést smysluplné vyhodnocení. Pokud by totiž z nějakého důvodu v návrhu programu jeho intervenční logika vysvětlena nebyla, musel by si hodnotitel před začátkem samotného hodnocení nejprve tento logický model zpětně zrekonstruovat. Práci hodnotitele však není spekulovat, proč je vlastně program zaváděn, případně vymýšlet argumenty, proč by měl fungovat, protože to musí především vědět a při schvalování programu prezentovat poskytovatel. Ledabyle zpracovaná intervenční logika v návrhu programu představuje pro hodnocení velký handicap ještě předtím, než vlastní hodnotící práce začnou.

Obr. 1 Příklad schematickeho vyjádření modelu intervenční logiky programu na podporu výzkumu v podnicích



Zdroj: Vlastní zpracování na základě BOEKHOLT, P. et al. (2014, str. 8).



## 4.3 METRIKA VÝKONNOSTI

Metrika výkonnosti v návaznosti na model intervenční logiky upřesňuje vstupy, výstupy, výsledky a dopady. Čím přesněji jsou určeny úkoly programu v tomto směru, tím lépe je možné hodnotit jeho úspěšnost. Programy s nejasnými úkoly často trpí i nejasným hodnocením, protože není zřejmé, co představuje jejich (ne)úspěch. Je v zájmu všech zúčastněných, včetně poskytovatelů, aby tato hranice byla zřetelná, aby v tom návrh programu byl konkrétní. Z toho důvodu by indikátory používané pro sestavení metriky výkonnosti měly splňovat následující kritéria:

- Konkrétní (specific) – indikátor je důkladně definován a váže se na konkrétní cíl programu.
- Měřitelný (measurable) – indikátor umožňuje sledování, je souměřitelný napříč příjemci.
- Dostupný (attainable) – indikátor je založen na existujících či získatelných datech.
- Relevantní (relevant) – indikátor je podstatný pro hodnocení pokroku v plnění cílů.
- Včasný (timely) – indikátor vychází z dat, která budou k dispozici včas.
- Ekonomický (economical) – indikátor je použitelný z hlediska rozpočtu.

Základní metrika výkonnosti, a tudíž vymezení klíčových indikátorů, by měla být prezentována již v návrhu programu, aby bylo možné podle toho sbírat potřebná data během monitorování v jeho průběhu. Nicméně na místě je varování, protože vědomí, že na základě této metriky bude prováděno hodnocení, s sebou nese riziko strategického chování zainteresovaných subjektů, včetně manažerů programu. Jejich přílišná orientace na plnění podrobné metriky, typicky v podobě počítání počtu bezprostředních výstupů, může vést k oslabení zřetele na plnění podstaty programu v podobě dopadů.

Zrádná je zejména skutečnost, že data za vstupy a výstupy jsou dostupná podstatně dříve, a to již v průběhu programu, než data za výsledky a následně dopady, která často přichází až s několikaletým zpožděním po skončení programu. Svádí to k mechanickému počítání výstupů při řízení programu. Avšak zdaleka ne všechny výstupy přinášejí výsledky vedoucí ke kýženým dopadům a ne všechny výstupy by se bez podpory z programu neuskutečnily. Účelová funkce manažera programu se vlivem metriky nesmí zúžit pouze na „nahánění bodů“ za výstupy.

Pro úspěšné naplnění cílů programu je zásadní, aby se poskytovatel držel nejenom plnění metriky výstupů, ale měl neustále na paměti i nutnost dosažení výsledků a především dopadů. Z tohoto důvodu je důležité, aby na jedné straně poskytovatelé byli srozuměni s tím, že úspěšnost jejich programů bude v prvé řadě hodnocena s ohledem na dosažené výsledky a dopady, nikoliv účetní přehledy výstupů, aby chápali alespoň základní podstatu metod hodnocení, včetně principů kontrafaktuálního přístupu, a aby na druhé straně bylo takové hodnocení důsledně prováděno.

## 5. Koncept hodnocení

---

Hodnocení je nástrojem pro vytváření cílených, včasných a spolehlivých podkladů pro rozhodování. Neocenitelné je zejména v oblasti programů, jejichž dopady neprochází trhem, a pro které tudíž neexistuje objektivní zpětná vazba, z níž by se daly odvodit jednoznačné indikátory úspěšnosti. Podstatou hodnocení je co možná nejspolehlivější zodpovězení hodnotících otázek za pomoci v praxi dobře prověřených metod a odpovídající datové základny. Hodnocení není monitorování, účetnictví, audit, statistika ani empirický výzkum, ačkoliv z nich čerpá data, metody a poznatky. Hodnocení neprobíhá až po ukončení programu. Hodnocení naopak začíná ještě před jeho spuštěním, postupuje skrze jeho implementaci a končí jeho zpětným posouzením.

### 5.1 FUNKCE HODNOCENÍ

Hodnocení programů může plnit řadu nejrůznějších funkcí, avšak obvykle má splnit jednu či více z následujících:

- **Funkce legitimizační** – poskytovat opodstatnění pro použití veřejných zdrojů za daným účelem.
- **Funkce informační** – přinášet informace o tom, jak se nakládá s veřejnými prostředky, co z toho má společnost, ukazovat dopady programu.
- **Funkce učící** – poučit se z předchozích úspěchů a nezdarů, tj. postupné zlepšování programů.
- **Funkce usměrňovací** – pomáhat při stanovování strategií, cílů a plánů programů do budoucna.
- **Funkce mediální** – kultivovat diskusi a vyhrazovat mantinely během střetu konkurenčních zájmů jednotlivých aktérů.

Hodnocení se liší v důrazu na jednotlivé funkce. Aby však později nedošlo ke zbytečným nedorozuměním, očekávání ohledně toho, jaké funkce mají být primárně splněny, je třeba zřetelně deklarovat již v zadávací dokumentaci hodnocení a následně hned na začátku hodnotícího procesu vyjasnit mezi zadavatelem a hodnotitelem.

### 5.2 POUŽITÍ HODNOCENÍ

Hodnocení může proběhnout v kterémkoliv okamžiku životního cyklu programu, a tudíž se může dívat zpětně do minulosti, tj. zhodnotit, co bylo uděláno, ale také může být orientováno do budoucnosti, tj. poradit, co by mělo být děláno jinak. Z toho plynou dva základní možné způsoby užití:

- **Sumativní hodnocení** – je zahleděno do minulosti, zpětně hodnotí program v určitém okamžiku, zaměřuje se na zodpovězení otázek ohledně již pozorovaných skutečností, jako je hospodárnost programu a rozsah jeho výstupů, výsledků a dopadů, zajímá ho, zda došlo ke splnění cílů, a tudíž jestli program byl úspěšný.
- **Formativní hodnocení** – dívá se do budoucnosti s cílem přispět ke zlepšení programu, klade si otázky proč, jak a za jakých podmínek něco funguje či nikoliv a snaží se z toho vyvodit poučení pro chod programu, zajímá ho, zda je program nasměrován ke splnění cílů a co je třeba případně změnit, aby byl program úspěšný.

Sumativní a formativní hodnocení si kladou rozdílné typy otázek, a proto vyžadují jiné typy podkladových údajů i odlišné metody. Hodnotitel potřebuje jasně vědět, jak má být jeho hodnocení následně použito, a tudíž jaké hodnocení se od něj očekává. Je velmi důležité předepsat zamýšlené použití hodnocení, případně vymežit potřebnou kombinaci jeho sumativních a formativních prvků již v zadávací dokumentaci před započítím samotného hodnotícího procesu.



## 5.3 TYPY HODNOCENÍ

**Hodnocení ex-ante** probíhá před zahájením programu, je formativní a orientováno do budoucnosti. Cílem je vyhodnotit návrh programu předložený poskytovatelem z hlediska schopnosti dosáhnout změny, na kterou je zaměřen. Jeho výsledkem jsou doporučení pro nastavení programu a jeho procesů. Je podkladem pro rozhodování tvůrce politik ohledně financování tohoto programu.

- Vyhodnotí návrh programu před jeho zahájením s cílem co nejvíce zvýšit jeho relevanci.
- Přezkoumá předpokládanou funkčnost intervenční logiky, jestli bude možné s navrženými aktivitami dosáhnout zamýšlených cílů.
- Zhodnotí systém měření, sledování a monitorování.
- Zkoumá přiměřenost finanční alokace, zda plánované finanční zdroje odpovídají zamýšleným cílům, resp. přínosům.
- Ohodnotí nastavení systému a procesů, zda umožní úspěšnou implementaci program.
- Posoudí plán dalšího hodnocení.
- Formuluje formativní doporučení pro design programu a jeho financování.

**Průběžné hodnocení** probíhá během programu, je převážně formativní a zaměřené do budoucna, avšak s prvky sumativního posouzení dosavadního průběhu. Cílem je zefektivnění implementace programu. Jeho závěry dávají doporučení pro další průběh programu, případně změny v organizaci, managementu a zaměření. Je podkladem pro rozhodování poskytovatele, zda program nechat běžet či zastavit.

- Používá se během programu, obvykle zhruba v jeho polovině, pro vyhodnocení jeho dosavadního průběhu.
- Hodnotí průběžnou efektivnost v naplňování cílů.
- Zabývá se procesy implementace, zejména určení úzkých míst.
- Klade důraz na poučení z rozběhu programu pro zlepšení jeho fungování.
- Posoudí následky případných změn vnějšího prostředí a politických priorit.
- Formuluje formativní doporučení pro zlepšení implementace programu.

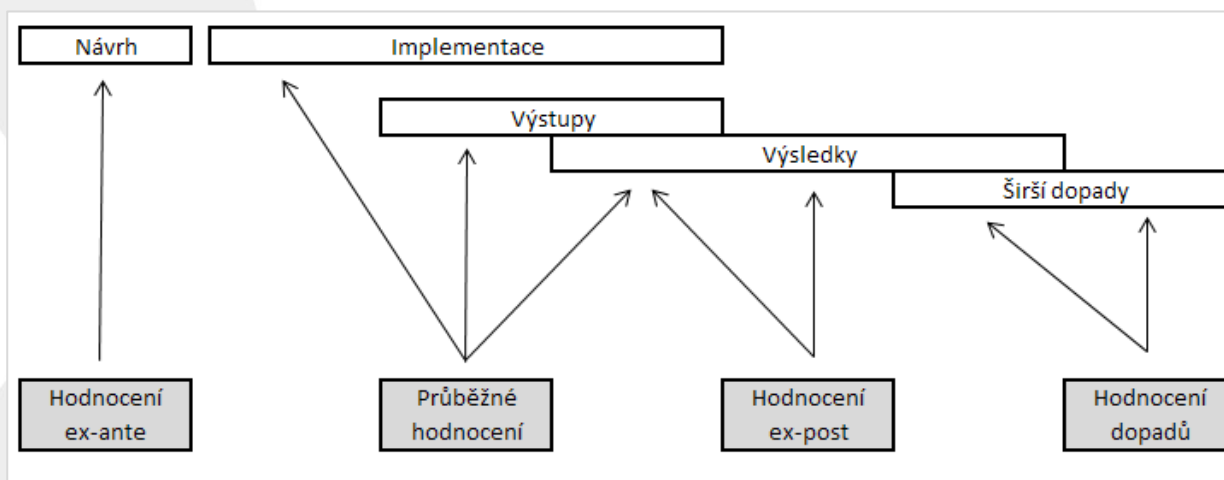
**Hodnocení ex-post** (závěrečné hodnocení) probíhá hned po skončení programu, je sumativní a orientované do minulosti. Cílem je předběžně zhodnotit úspěšnost programu na základě doposud pozorovaných skutečností. Zabývá se implementací programu a jeho účinky na vstupy, výstupy a bezprostřední výsledky. Jeho závěry poskytují doporučení pro navazující programy. Je podkladem pro rozhodování poskytovatele, zda pokračovat v podobně zaměřených programech, a pro tvůrce politik, zda návrhy na další podobné programy podporovat.

- Používá se krátce po skončení programu, zhruba do jednoho roku, pro zpětné vyhodnocení jeho průběhu.
- Posoudí, které cíle programu se již podařilo naplnit a jaký je výhled pro splnění zbývajících cílů.
- Určí bezprostřední efekty, k nimž v důsledku programu došlo, a to zejména přímé účinky na příjemce.
- Formuluje sumativní závěry z hlediska programu samotného a zároveň doporučení pro design a implementaci navazujících programů.

**Hodnocení dopadů** probíhá s delším odstupem po ukončení programu, kdy jsou již dostatečně pozorovány jeho širší přínosy. Cílem je zhodnotit konečné splnění cílů programu, a tudíž jeho celkovou úspěšnost. Zaměřuje se na trvalost účinků programu na vstupy a výstupy a konečné výsledky, ale zejména na jeho dopady. Jeho závěry slouží jako doporučení pro strategické směřování příštích programů. Je podkladem pro rozhodování poskytovatele o změnách v portfoliu jím používaných nástrojů podpory a pro rozhodování tvůrce politik ohledně prioritních oblastí a rozdělení zdrojů mezi poskytovatele.

- Souhrnně zhodnotí celý program a jeho konečné přínosy.
- Posoudí, do jaké míry se podařilo naplnit hlavní i dílčí cíle programu.
- Bere v úvahu i širší efekty, ke kterým v důsledku programu došlo, ať už přímé nebo nepřímé a zamýšlené nebo nezamýšlené.
- Formuluje sumativní závěry o daném programu, ale především poskytuje doporučení pro strategická rozhodnutí ohledně budoucích programů.

Obr. 2 Co se hodnotí a kdy?



Tab. 3 Hodnocení programů podle typu

|                  | Hodnocení ex-ante   | Průběžné hodnocení   | Hodnocení ex-post  | Hodnocení dopadů  |
|------------------|---|--|--|---|
| <b>Otázky</b>    | Je program vhodným nástrojem pro dosažení požadované změny? Jak prioritní je dané řešení? Je intervenční logika programu promyšlená? Je možné splnit cíle programu? | Jak může být zlepšeno fungování programu? Pracuje dobře jeho management? Jaké změny je třeba udělat? Jak by se měly lišit příští výzvy? Měl by program pokračovat? | Zapůsobil program podle předpokladů?<br><br>Které cíle se již podařilo naplnit, jaký je výhled?<br>Jaké měl program bezprostřední následky?<br>Co je třeba příště udělat lépe?<br><br>Měl by mít program následovníka? | Podařilo se vyřešit problém, na který se program zaměřoval?<br><br>Splnil program svoje cíle? Jaké měl program zamýšlené a nezamýšlené následky? Jak by se měly změnit příští programy? |
| <b>Uživatelé</b> | Tvůrce politik, poskytovatel  | Poskytovatel   | Tvůrce politik, poskytovatel   | Tvůrce politik, poskytovatel, veřejnost   |
| <b>Zaměření</b>  | Koncepční a formativní  | Formativní (s důrazem na učení)  | Sumativní (formativní pro podobné programy)  | Sumativní (formativní pro příští programy)  |
| <b>Předmět</b>   | Smysluplnost, významnost a design   | Průběh implementace  | Dosavadní splnění cílů, efektivnost a užitečnost   | Konečné splnění cílů a celkové přínosy  |
| <b>Obsah</b>     | Analýza konceptu, logická konzistentnost, splnitelnost cílů, zhodnocení rizik, přiměřenost nákladů, posouzení plánu dalšího hodnocení                               | Posouzení organizace, managementu a průběhu implementace, analýza dosavadních výstupů a výsledků   | Zpětné posouzení implementace a přímých efektů na vstupy, výstupy a bezprostřední výsledky   | Zhodnocení konečných efektů na vstupy, výstupy, konečné výsledky a dopady   |
| <b>Následky</b>  | Financovat či nikoliv, popřípadě změny v nastavení  | Pokračovat či zastavit, popřípadě změny v organizaci, managementu a zaměření   | Navázat podobným programem či nikoliv  | Změny v nastavení programů a ve strategii VVI politiky  |

Zdroj: Vlastní zpracování na základě FTEVAL (2012, str. 28).

## 5.4 PŘEDMĚT HODNOCENÍ

Na veřejném výdajovém programu může být hodnoceno v podstatě cokoliv. Nicméně v praktické rovině se hodnocení zpravidla zabývá následujícím:

- 1. Hodnocení konceptu** se zabývá misí, předpoklady, hypotézami a rámcem, na nichž je program založen a na základě kterých je financování programu opodstatněno. Smyslem je posoudit, zda je vynaložení veřejných prostředků tímto způsobem ospravedlnitelné pro dosažení požadované změny. Jedná se zpravidla o zásadní část hodnocení ex-ante.
- 2. Hodnocení designu** se zaměřuje na hospodárnost a organizační strukturu programu. Hodnotí se, jestli operační cíle programů odpovídají změnám, kterých má být dosaženo, jestli zvolené nástroje podpory jsou vhodné ke splnění těchto cílů a do jaké míry je jejich splnění pravděpodobné. Hodnotí se také vhodnost

organizační struktury a načasování, fázování a způsobu řešení případných problémů. Jedná se o součást všech typů hodnocení.

3. **Hodnocení procesů** se zaměřuje na provedení programu, což zahrnuje jeho implementaci, fungování kontrolních mechanismů a také hospodárnost vynakládaných prostředků. Jedná se o hlavní část průběžného hodnocení a součást ex-post hodnocení. Je zejména důležité u velkých programů s komplikovanou strukturou.
4. **Hodnocení přínosů** spočívá v určení, do jaké míry se podařilo splnit cíle programu, a ve zjištění jeho přímých i nepřímých následků. Za tímto účelem je přínosné rozlišovat mezi účinkem na vstupy, chováním příjemců, bezprostředními výstupy, následnými výsledky a konečnými dopady. Přínosy často mají, avšak nezbytně nemusí mít, ekonomickou podstatu. Jedná se o typickou součást hodnocení ex-post a hlavní předmět hodnocení dopadů.

## 5.5 HODNOTICÍ ZPRÁVA

Hodnoticí zpráva je z formálního hlediska hlavním výstupem hodnocení. Jedná se z hlediska obsahu o ucelený a samostatný dokument, který má stručnou, přehlednou a vyčerpávající formou vylíčit, co, jak a proč bylo hodnoceno, co bylo v procesu hodnocení zjištěno a zejména jaká z toho plynou doporučení. Dobrou praxí je v hlavním dokumentu shrnout a zdůraznit zásadní sdělení pro širší veřejnost a v přílohách do hloubky rozebrat podrobnosti, jež zajímají odbornou veřejnost. Po jejím přečtení by neměly zůstat pochybnosti ohledně provedení empirických analýz. Hodnoticí zpráva by měla přinejmenším obsahovat:

- požadavky na náležitosti uvedené v zadávací dokumentaci hodnocení,
- vyčerpávající popis kontextu programu (opodstatnění, cíle, průběh atd.),
- představení účelu a intervenční logiky programu,
- teoretický rámec hodnocení a z toho plynoucí předpoklady, na kterých je založeno,
- vysvětlení strategie a metodiky hodnocení,
- rozbor silných a slabých stránek zvolených analytických metod,
- popis dat, jejich zdrojů, případně sběru dat,
- analytické výsledky a podrobnosti o použitých postupech,
- konkrétní závěry a doporučení.

Hodnocení je věrohodné jen do té míry, do jaké míry jsou spolehlivá data, na nichž je založeno. Uvedena by měla být omezení použitých dat, a tudíž rozebrán rozdíl mezi daty ve skutečnosti dostupnými ve srovnání s nejlepšími možnými a konstatováním, proč se taková data nepodařilo získat. Pokud hodnocení založeno na primárních datech, je nutné zodpovědět otázky ohledně stratifikace, provedení sběru, odezvy a reprezentativnosti. Ke zprávě je třeba přiložit podkladové dokumenty, jako jsou pokyny pro řízené rozhovory, dotazník a technický projekt šetření či slovníček s definicemi používaných pojmů. Hodnoticí zpráva musí rozptýlit obavy, že při získávání dat došlo ke zkreslení, které ovlivňuje výsledky.

Podobně je třeba prezentovat podrobnosti o použitých metodách. Účelem hodnoticí zprávy není rozebírat podstatu metod samotných, ale je třeba vysvětlit nuance zvoleného řešení do té míry, aby mohla být daná analýza reprodukována třetí osobou. Například i základní kvantitativní metody jsou založeny na předpokladech, jejichž validitu je třeba testovat a výsledky zveřejnit. Pokud je použit statistický či ekonometrický software, je nutné citovat konkrétní verzi, uvést přesné

parametry použitých příkazů a zvolené varianty způsobu odhadu, a to zejména v případě odchylky od obvyklých postupů. Podrobnosti tohoto typu nemusí zatěžovat hlavní dokument, ale neměly by chybět v přílohách.

Hodnoticí zpráva by měla obsahovat konkrétní závěry a doporučení, a to buď na základě minulosti v případě sumativního hodnocení, anebo zaměřené do budoucnosti v případě formativního hodnocení. Závěry by se měly vztahovat k původně vymezeným hodnoticím otázkám, měly by být navzájem logicky konzistentní a musí být založeny na konkrétních výsledcích empirických analýz. Pro nepodložené spekulace není v hodnoticí zprávě místo. Závěry by měly být srovnány s již dostupnými výsledky, ať už ve světové literatuře, anebo s výsledky předchozích hodnocení podobných programů, a měly by poskytovat doporučení, jak zlepšit hodnocení do budoucna. Závěry by měly být prodiskutovány s dotčenými subjekty. Širší veřejnost by měla být upozorněna na úskalí předkládaných závěrů a mělo by být uvedeno, co se bohužel nepodařilo zjistit, anebo zůstává nejasné.

## 6. Metodika hodnocení

---

Metodika hodnocení je postup zvláště vytvořený za účelem vyhodnocení jednoho konkrétního programu. Každý program má svoji vlastní metodiku hodnocení, jež se v závislosti na jeho nastavení, cílech, průběhu a kontextu nebo z důvodu dostupnosti dat větší či menší měrou liší od jiných programů. Není tudíž možné vymezit jedinou nejlepší metodiku, nějakou jednu formulku, kterou by bylo vhodné použít pro všechny programy. Nicméně je možné vymezit obecné náležitosti, které by mělo hodnocení z hlediska metodiky splňovat. K tomu si je nejprve třeba uvědomit:

1. **Rozdíl mezi monitorováním a hodnocením** – sběr dat o průběhu programu a celkově vytváření datové základny je sice zásadní podpůrnou činností, nicméně samotné hodnocení začíná teprve tam, kde monitorování končí, hodnocení programů nelze zužovat na prosté počítání bodů za výstupy uplatněné pod jejich hlavičkou; monitorování není hodnocení, nelze je zaměňovat a první vydávat za druhé.
2. **Hodnocení má být průkazné** – hodnocení má být tzv. „evidence-based“, což znamená, že zodpovězení hodnotících otázek není založeno na domněnkách, spekulacích a nepodložených tvrzeních, ale na základě výsledků získaných pomocí metod k tomu určených, které jsou schopny na dané otázky uspokojivě odpovědět; soudy ohledně splnění cílů programu, které nejsou spolehlivě prokázány, nelze nazývat hodnocením.
3. **Rozmanitost použitelných metod** – hodnocení může zahrnovat jednu metodu anebo daleko lépe kombinaci metod, tzv. triangulaci, které mohou být kvantitativní i kvalitativní; různé metody mohou zodpovědět stejně položené otázky, nejdříve je třeba vymezit hodnotící otázky a až pak vybírat metody, nikoliv naopak žádoucí je použít metod, které se dokáží přiblížit prokázání příčinné souvislosti.

BOEKHOLT, P. et al. (2014) rozpracovali procesní přístup k hodnocení (*Evaluation Reference Model*), podle kterého lze rozlišit šest základních kroků:

1. Rozbor modelu intervenční logiky.
2. Vymezení cílů hodnocení a hodnotících otázek.
3. Příprava hodnocení.
4. Určení vhodných metod a provedení analýzy.
5. Formulace závěrů a odevzdání hodnotící zprávy.
6. Použití závěrů hodnocení při rozhodování.

Zdůrazněn je šestý krok, totiž aby kvalita hodnocení nebyla vnímána pouze z pohledu použitých metod, ale rovněž z hlediska užitečnosti předkládaných závěrů pro uživatele hodnocení při rozhodování. Šestý krok vede zpětně ke kroku prvnímu, a tudíž by poskytovatelé měli sestavovat model intervenční logiky v návrzích nových programů na základě závěrů předchozích hodnocení.

## Co je to metodika hodnocení programu?

Metodika hodnocení je postup zvláště vytvořený za účelem hodnocení jednoho konkrétního programu. Může zahrnovat jednu metodu anebo častěji kombinaci různých – kvantitativních i kvalitativních – metod. Každý program má svoji vlastní metodiku hodnocení, jež se v závislosti na jeho nastavení, cílech, průběhu, kontextu a dostupných datech liší od jiných programů. Není přínosné snažit se vymezit jednu metodiku použitelnou pro všechny programy.

## 6.1 HODNOTICÍ OTÁZKY

Hodnoticí otázky by měly být odvozeny od cílů programu, měly by mířit na podstatu fungování programu a ověřovat tvrzení o jeho očekávaném působení. Neměly by být pouze popisné, a tudíž sklouzávat po povrchu problému, nýbrž jít do hloubky fungování intervenční logiky programu. Je třeba se mimo jiné ptát na příčinné souvislosti mezi vstupy, výstupy, výsledky a dopady, zkoumat empirickou platnost předpokladů, na kterých jsou mechanismy působení programu založeny, a přezkoumat hypotézy o působení alternativních mechanismů.

Při formulaci hodnoticích otázek může pomoci tzv. I-O-O-I model (z anglického *Inputs-Outputs-Outcomes-Impacts*) podle doporučení BOEKHOLT, P. et al. (2014, str. 10), který nabízí následující základní hodnoticí kritéria programů:

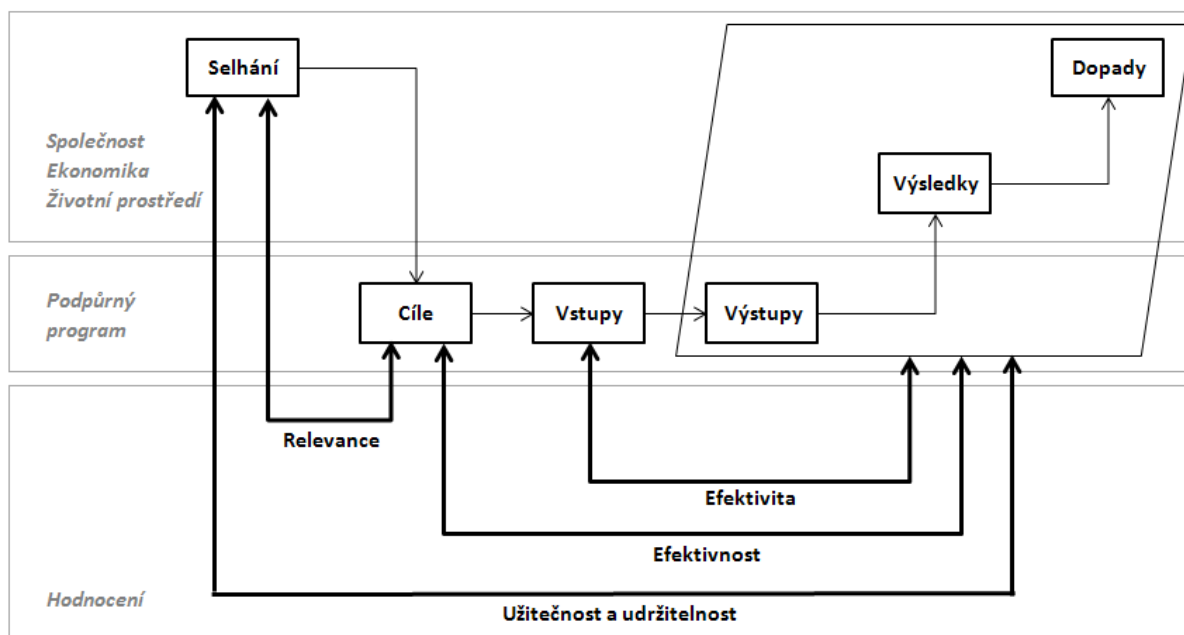
- **Relevance** (relevance) – Ptáme se, zda se cíle programu zaměřují na dostatečně závažné problémy, které je prioritní řešit s odkazem na strategické vládní dokumenty (Děláme správnou věc?).
- **Efektivnost** (effectiveness) – Zjišťujeme, zda přínosy podpořených aktivit jsou v souladu s cíli programu, jak velký byl motivační (pákový) efekt podpory a do jaké míry lze přínosy přiřknout podpoře (Co se stalo v důsledku podpory?).
- **Efektivita** (efficiency) – Zkoumáme, zda jsou veřejné zdroje vynaloženy hospodárně, tj. zda dopady, výsledky a výstupy odpovídají vstupům (Bylo možné dosáhnout stejného účinku s méně vstupy?).
- **Užitečnost** (utility) – Šetříme, zda program dosáhl (anebo pomáhá dosáhnout) změny, kvůli které byl spuštěn, včetně nezamýšlených následků, a komu to je ku prospěchu (Je ještě třeba udělat něco dalšího?).
- **Udržitelnost** (sustainability) – Chceme vědět, do jaké míry bude působení programu pokračovat i po jeho skončení (Bude mít podpora trvalé přínosy?).

Nejedná se o vyčerpávající přehled kritérií, která je možné použít, protože z cílů konkrétního programu mohou vyplynout i jiné otázky. Nicméně pro všechny hodnoticí otázky platí, že pokud jsou řádně formulovány, mělo by být možné z nich odvodit konkrétní hypotézy o fungování programu, jež je možné empiricky testovat s použitím odpovídajících analytických metod. Hodnoticí otázka, na kterou lze odpovědět zběžně bez hlubší analýzy, patrně nebyla dostatečně promyšlena.

Hodnoticí otázky mimo jiné závisí na funkci, typu a použití hodnocení, takže jejich seznam se bude lišit případ od případu. Hodnocení dopadů a ex-post se zaměřují zejména na efektivnost, udržitelnost a užitečnost, průběžná hodnocení kladou důraz na efektivitu a užitečnost, zatímco ex-ante hodnocení spíše na relevanci. Nicméně plnohodnotné hodnocení by mělo program vždy posoudit komplexně z více hledisek. Nedoporučuje se, aby se hodnocení zabývalo pouze vybranými kritérii a ostatní zcela opomíjelo, protože potom hrozí, že výsledkem bude neúplný obrázek o úspěšnosti programu.



Obr. 3 Kritéria hodnocení podle I-O-O-I modelu



Zdroj: Vlastní zpracování na základě BOEKHOLT, P. et al. (2014, str. 11).

Je třeba zdůraznit, že I-O-O-I model neznámá, že hodnocení je možné provádět prostým poměřováním vstupů a výstupů programu. Jednoduché poměrové ukazatele, jako byl kupříkladu tzv. index státního rozpočtu v oficiální metodice hodnocení programů používané v letech 2007–2012 (RVVI 2012), jsou zavádějící. Poskytování podpor pro VVI nelze hodnotit jako výrobní proces, do kterého jdou na jedné straně vstupy a na druhé straně vychází výrobky, které se jednoduše spočítají, ohodnotí a poměří k nákladům. Představa, že se do VVI nasypaní peníze a ve zhruba ve stejném období se objeví dobře viditelné přínosy, a to ve stejných organizacích, do kterých plynuly peníze, je neúměrně zjednodušující.

## 6.2 DESIGN HODNOCENÍ

Na základě hodnotících otázek sestavíme design hodnocení. Pro vymezení designu hodnocení je klíčové, zda zodpovězení hodnotících otázek vyžaduje pátrání po příčinných souvislostech, tj. určení příčiny a následku, mezi poskytnutím podpory a plněním kritéria hodnocení. Obecně lze rozlišit tři základní možnosti pro design hodnocení:

- **Neexperimentální design** – k zodpovězení hodnotících otázek není třeba určovat příčinné souvislosti, tj. postačuje deskriptivní přístup, což je typické pro ex-ante a obvykle i průběžné hodnocení.
- **Kvaziexperimentální design** – zodpovězení hodnotících otázek vyžaduje pátrání po příčinných souvislostech, což je rostoucí měrou součástí ex-post hodnocení a podstatou hodnocení dopadů, avšak podpora nebyla mezi žadatele záměrně rozdělena náhodným způsobem, tj. nahodilost musíme simulovat pomocí kvaziexperimentálních metod.
- **Experimentální design** – hodnotící otázky se ptají po příčinných souvislostech, přičemž podpora byla žadatelům přidělena náhodně, což sice v případě veřejných výdajových programů není zcela vyloučeno, avšak v praxi zřídka proveditelné.



## 6.3 KONTRAFAKTUÁLNÍ PŘÍSTUP

Zásadním, ale zároveň nejnáročnějším je v hodnocení programů použití kontrafaktuálního přístupu. Kontrafaktuální situace popisuje stav, jenž by nastal, pokud by podpora nebyla poskytnuta a naopak. Kontrafaktuální přístup se snaží určit tento hypotetický stav, a tudíž odpovědět na otázku, kolik z toho, co se stalo, by bylo i bez podpory. Spočívá v testování hypotéz o (ne)existenci účinku podpory. Jde o to rozklíčovat, které efekty je možné přiřknout podpůrnému programu, zatímco se nám podaří vyhnout tomu, abychom programu přiřadili efekty, které pouze působí souběžně s ním, a tudíž by nastaly, i kdyby podpora nebyla čerpána. Z výsledků je potom možné odvodit čistý účinek veřejného výdajového programu.

Náročné je to z toho důvodu, že kontrafaktuální přístup znamená použití vědecké metody v kontextu hodnocení, což s sebou přináší vysoké nároky nejen na hodnotitele samotné, nýbrž i na zadavatele, poskytovatele a obecně všechny zainteresované. Pokud totiž tito nechápou principy vědecké metody, nemohou docenit její nadřazenost všem ostatním ani pořádně porozumět jejím výsledkům. Přitom získání průpravy ve vědecké metodě vyžaduje usilovné studium, které obvykle spadá až do nejvyšších pater vysokoškolského vzdělávání, což od širší veřejnosti nelze očekávat. Nicméně řadu klíčových hodnoticích otázek nelze uspokojivě zodpovědět jinak než s pomocí kontrafaktuálního přístupu. Pokud někdo hodlá hodnocení skutečně porozumět, nezbyvá nic jiného než principy kontrafaktuálního přístupu rozlousknout.

Analytických metod, které lze použít pro kontrafaktuální přístup, je značné množství. Není v silách tohoto textu vysvětlovat jejich podrobnosti. Nicméně všechny spojuje princip, že porovnáváme podporovaná pozorování se srovnávací skupinou co možná nejjpodobnějších pozorování, která podporu nečerpaly. Pokud by jediný rozdíl mezi podpořenou a srovnávací skupinou bylo čerpání podpory, potom by srovnávací skupina přímo ukazovala, co by se stalo s podpořenými pozorováními, kdyby podporu nedostaly. Srovnání tímto způsobem je možné pouze pomocí řízeného experimentu, který je hojně využíván v přírodních vědách, což však bohužel až na výjimky není možné pro hodnocení programů. Nezbyvá tudíž než simulovat kontrafaktuální situace pomocí kvaziexperimentů, které k prokázání příčinných souvislostí využívají určitých nahodilostí při udělování podpory.

Při kontrafaktuálním hodnocení podpůrných programů nás konkrétně zajímá efekt podpory na vybrané charakteristiky příjemců, jež jsou odvozené od plnění cílů programu a vztahují se k mechanismům jeho působení. Například může jít o určení motivačního (anebo pákového) efektu podpory na vstupy, tj. kolik výdajů na výzkum by podporovaný subjekt vynaložil i bez poskytnutí podpory, jinými slovy kolik nových výdajů na výzkum, které by se jinak neuskutečnily, podpora stimulovala. Může jít o stanovení účinku podpory na změnu chování, tj. například s jakou pravděpodobností by spolupracovala firma s univerzitou na inovacích i bez podpory. Podobně můžeme určovat efekt podpory na výstupy, výsledky a dopady, tj. co z toho by nevzniklo, kdyby podpora poskytnuta nebyla.

Kontrafaktuální hodnocení programů je podobné výzkumu v sociálních vědách. Používány jsou stejné analytické metody, avšak s tím zásadním rozdílem, že definičním znakem výzkumu je prvek novosti, který se po hodnocení nepožaduje. Při hodnocení jde v první řadě o spolehlivé zodpovězení hodnoticích otázek, přičemž je podružné, jestli jsou při tom vytvářeny nové znalosti. Hodnocení s použitím metod sociálních věd pro rozhodování ve veřejném sektoru lze připodobnit poskytování služeb měření a kalibrace v oblasti přírodních věd pro průmysl. Z výzkumného hlediska může jít o zcela nezajímavé rutinní cvičení, což však hodnocení nijak neubírá na užitečnosti, ba naopak může to být vnímáno jako přínosné, protože se jedná o prověřené metody. Společné mají hodnocení i výzkum požadavky na objektivitu, důkazy a reprodukovatelnost. Není náhodou, že v zemích s vysokou úrovní sociálních věd se provádí i kvalitní hodnocení.

Z principů kontrafaktuálního přístupu mimo jiné plyne rozdíl mezi hodnocením úspěšnosti projektů v rámci programu a úspěšnosti programu jako celku. Pokud podpořené projekty splní, co slíbily, nezbyvá než je považovat za úspěšné, protože jednotlivé příjemce nelze trestat za nedomyšlené nastavení programu. Při hodnocení programu jako takového záleží na tom, do jaké míry se vytvořilo, co by jinak nevzniklo. Program je více než prostý součet projektů, a tudíž jeho hodnocení nemůže spočívat ve shrnutí závěrečného hodnocení projektů. I když jsou nakrásně všechny projekty v rámci programu shledány jako úspěšné a v jednotlivých projektech byly vytvořeny velké počty čehokoliv, ještě to nutně neznamená, že se program jako celek neminul účinkem, pokud by tyto přínosy vznikly i bez podpory.

Kontrafaktuální vyhodnocení je nasnadě u podpor směřujících do soukromého sektoru, kdy se nabízí možnost financování z neveřejných zdrojů. Alternativy financování jsou však téměř vždy možné i pro více „fundamentální“ výzkum ve veřejném sektoru, který by se sice bez veřejné podpory z velké části neuskutečnil, ale to ještě neznamená, že je nejlepší jej financovat právě skrze daný program. Jako alternativa se nabízí například financování z jiného instrumentu, z národních versus evropských fondů anebo z institucionálních zdrojů. Pro program, který poskytuje podporu pro návrat špičkových vědců ze zahraničí, se jako srovnávací skupiny nabízí kromě výzkumníků, kteří o podporu žádali, ale nedostali ji (vrátili se stejně?), i výzkumníci, kteří o ni nežádali, a i přesto se vrátili (proč i bez podpory?), a čeští výzkumníci působící v zahraničí, kteří nežádali a zůstávají tam (proč vůbec nepožádali?). Kontrafaktuální situaci k vyhodnocení je možné vymezit i v řadě případů pro dotace směřující mimo soukromý sektor.

Kontrafaktuální myšlení musí být vodítkem při prvotní formulaci cílů a intervenční logiky, které jsou předmětem ex-ante hodnocení, přes manažerské rozhodování při řízení implementace programu, kterým se může zabývat průběžné i ex-post hodnocení, až po otázky konečných přínosů, jimž se věnuje hodnocení dopadů. Není to pouze otázka použití jednotlivé metody v určitém typu hodnocení, což lze případně vynechat, a potom takový přístup opustit, nýbrž způsob myšlení o podstatě poskytování podpor. Nemá sice smysl přímo požadovat kontrafaktuální analýzu, pokud to není nutné k zodpovězení hodnotících otázek, avšak i když se zrovna v hodnocení nezjišťuje motivační efekt podpory, stále je třeba se držet kontrafaktuálního myšlení. Již cíle programu, od kterých se v hodnocení odvíjí vše ostatní, musí být vymezeny ve smyslu dosažení určité změny. Pokud kontrafaktuální přístup není použit, jsou poznatky prezentované v hodnotící zprávě slabým stínem toho, co bylo možné zjistit.

## 6.4 METODY HODNOCENÍ

Metod používaných pro hodnocení programů existuje široká škála od čistě kvantitativních po ryze kvalitativní přes přístupy, které spojují prvky obou. Kvantitativní metody zahrnují například ekonometrickou analýzu, bibliometrickou analýzu či síťovou analýzu. Mezi kvalitativní metody patří případové studie, expertní hodnocení a peer review. Nejvhodnější metoda anebo lépe mix metod pro vyhodnocení konkrétního programu závisí kromě hodnotících otázek na účelu hodnocení, na tom kolik času je na hodnocení, na rozpočtu pro hodnocení a bohužel také často na dostupnosti dat.

Příklady metod hodnocení:

- ekonometrická analýza (econometric analysis),
- analýza sociálních sítí (social network analysis),
- srovnávací analýza (benchmarking),
- kontextová analýza (context analysis),
- analýza zainteresovaných subjektů (stakeholder analysis),
- mapování výsledků (outcome mapping),

- případové studie (case studies),
- expertní hodnocení (expert assessment),
- mezinárodní peer review (international peer review),
- bibliometrická analýza (bibliometric analysis).

Kvantitativní metody použité na tvrdá reprezentativní mikrodata mají nesporné přednosti, a to mimo jiné ohledně objektivnosti, reprodukovatelnosti a zobecňování výsledků. Zdaleka nejsou omezeny pouze na tradiční kvantitativní analýzy číselných proměnných, jako jsou peněžní indikátory, údaje o zaměstnanosti či počty patentů. Kvantitativní metody, jako například tobit, logit a probit ekonometrické modely, jsou velmi dobře použitelné i pro analýzu nespojitých, ordinálních, kardinálních či binárních proměnných, které zachycují měřitelné charakteristiky chování, výběrů a rozhodování jednotlivých subjektů, což dokládá stále širší použití těchto metod i v oblastech sociologie, politologie či psychologie. Záběr použitelnosti kvantitativních metod je tudíž podstatně širší, než se může na první pohled zdát.

Zároveň je však třeba varovat před přílišným spoléháním na kvantitativní analýzy. Předmětem hodnocení je dosažení kvalitativní změny, ať už v podobě vědeckého objevu, technologického průlomu či inovace produktu na trhu, která je ze své kvalitativní povahy obtížně souměřitelná mezi jednotlivými pozorováními. Každý výsledek výzkumu je svým způsobem unikátní, anebo by alespoň měl být, pokud má striktně dostát definici, a lze tudíž pouze vymezit jeho určité proxy měřitelné charakteristiky, pro které lze sbírat tvrdá data. Některé hodnotící otázky ani nelze s pomocí kvantitativních metod zodpovědět.

Kvantitativní metody jsou vhodné pro:

- analýzu velkých souborů pozorování,
- vyvození sumativních závěrů založených na tvrdých datech,
- určení hlavních rozdílů, vztahů a trendů,
- testování hypotéz ohledně nenáhodnosti (tzv. statistické významnosti) pozorovaných korelací a v lepším případě příčinných souvislostí,
- získání odpovědí na hodnotící otázku v podobě čísla.

Kvalitativní metody jsou vhodné pro:

- analýzu menších souborů,
- získání formativních závěrů,
- zachycení měkkých aspektů chování,
- pochopení komplexních problémů těžko zredukovatelných do matematického modelu,
- rozklíčování příčin a následků silně závisících na obtížně měřitelném kontextu,
- zodpovězení hodnotící otázky v podobě stanoviska.

Dobrou praxí je kombinování kvantitativních a kvalitativních metod, které umožňuje vzájemné vykrytí jejich silných a slabých stránek. Pokud jsou výsledky z jedné metody přímo ověřovány jinou metodou, například když jsou zjištění ekonometrických analýz na tvrdých mikrodatech potvrzována pomocí případových studií postavených na řízených rozhovorech, hovoří se o tzv. triangulaci (*triangulation*) analytických metod. Přínosné je zejména kombinovat metody,

kteře mají široký záběr, tj. kvantitativní analýzy na velkém počtu pozorování, a metody zaměřené na podrobnosti, tj. kvalitativní analýzy vybraných pozorování. Pokud oba přístupy vedou ke stejnému výsledku, jsou závěry hodnocení podstatně lépe podloženy, než pokud jsou odkázány pouze na jednu metodu.

Kontrafaktuální hodnocení je obvykle kvantitativní, ale může být použito i na semi-kvantitativní indikátory, jako jsou binární proměnné, anebo může být i kvalitativní. Použití logiky kontrafaktuální analýzy totiž samo o sobě nijak nepředjímá, na jakém typu metod má být hodnocení založeno. Například se zapomíná na to, že lze zvažovat situaci, co by se stalo, kdyby podpora udělena nebyla, rovněž v případových studiích na základě poznatků získaných během rozhovoru s příjemcem podpory ve srovnání s obrázkem plynoucím z rozhovoru s žadatelem, který podporu nedostal. Nicméně je třeba varovat, že v každém případě ke spolehlivému prokázání příčinných souvislostí musí hodnotitel zvládat pokročilé metody. Není nic horšího než jednoduchá popisná analýza, která se pouští do dalekosáhlých závěrů ohledně příčin a následků.

Zadavatelé by se neměli bát požadovat použití nejmodernějších analytických metod a v přiměřené míře experimentovat s novými přístupy k hodnocení. Namísto je v zadávací dokumentaci požadovat alespoň mírný pokrok oproti předchozímu podobně zaměřenému hodnocení, aby byli hodnotitelé motivováni se učit. Český systém ve kvalitě hodnocení výrazně zaostává za světovou špičkou, takže při zadávání hodnocení je nutné mít na paměti, že existuje velký prostor pro dotahování směrem k úrovni běžné v zahraničí. Zadavatelé by se neměli uspokojit se stávajícím stavem hodnocení a přijímat výmluvy ze strany hodnotitelů na přílišnou obtížnost provedení anebo neexistenci potřebných mikrodat, ale naopak by měli neustále vytvářet tlak na zlepšení.

Tab. 4 Proveditelnost kvantitativního vyhodnocení kontrafaktuální situace

|  | VÍCE PROVEDITELNÉ   | MÉNĚ PROVEDITELNÉ   |
|--|---|---|
| <b>Povaha, rozsah a délka očekávaného účinku</b> | <p>Lze vymezit přímočarou vazbu mezi poskytnutím podpory a účinkem, který nás zajímá.</p> <p>Očekává se velký, dobře rozpoznatelný účinek podpory v porovnání s ostatními souběžně působícími vlivy.</p> <p>Účinek se soustředí do kratšího časového úseku, ale rychle nevyrchá.</p>  | <p>Mezi podporou a účinkem je komplexní a zprostředkovaná vazba, kterou ovlivňuje spousta dalších souběžně působících vlivů.</p> <p>Očekává se spíše menší, komplikovaný a obtížněji rozeznatelný účinek.</p> <p>Účinek postupně pomalu roste v dlouhém období.</p>   |
| <b>Měřitelnost a dostupnost dat</b>              | <p>Data jsou dostupná na individuální úrovni, tj. mikrodata.</p> <p>Dostatečný počet pozorování pro daný typ statistického testování.</p> <p>Podpora má u příjemců za následek ostře vymezenou a podobnou změnu.</p> <p>Data jsou dostupná před i po poskytnutí podpory.</p> <p>Data jsou dostupná v potřebném časovém rozmezí.</p> | <p>Dostupná jsou pouze nahrubo agregovaná data.</p> <p>Malý počet pozorování, který neumožňuje spolehlivý odhad modelu.</p> <p>Podpora přináší změny, které je obtížné ohraničit a porovnat mezi příjemci.</p> <p>S hledáním dat se začalo pozdě, časové řady nejsou dostupné.</p> <p>Není zřejmé, kdy podpora přesně působila.</p> |
| <b>Srovnávací skupiny</b>                        | <p>Rozhodnutí o udělení podpory je na základě objektivních kritérií, například známe pořadí žadatelů nad a pod čarou na základě hodnotícího skóre.</p> <p>Lze najít prvek nahodilosti, který ovlivnil přidělení podpory, tj. identifikovat model.</p>   | <p>Rozhodnutí o udělení podpory je subjektivní, neznáme přesné důvody rozhodnutí anebo k tomu nelze dohledat data.</p> <p>Nahodilost, která by umožnila vymezení srovnávací skupiny, není známa.</p>  |

Zdroj: Vlastní zpracování na základě BOEKHOLT, P. et al. (2014, str. 14).

Zdůraznit je třeba také skutečnost, že zvláště v oblasti VVI je přínosný mezioborový přístup k hodnocení. Změny, na jejichž dosažení programy cílí, jsou často složité s přesahy do ekonomické, technické i sociální problematiky, a které lze stěží hodnotit pouze z jednoho pohledu. Například ekonomický pohled je velmi přínosný, ale často by bylo ke škodě věci, pokud by byl jediný anebo silně dominantní, protože všechno neprochází trhem a nelze převádět na peníze. Mezioborový přístup k hodnocení je potřebný zejména u programů, které mají velmi komplexní a vícevrstevnaté cíle.

Konečně je třeba připomenout rozpočtová omezení, podle kterých je nutné nastavit metodický mix hodnocení. Případové studie založené na osobních rozhovorech jsou velmi časově náročné, a tudíž drahé ve srovnání s internetovým průzkumem. Ekonometrické a bibliometrické analýzy mohou být také drahé, pokud je nutné nově nakupovat velké objemy mikrodat a před první analýzou je zdlouhavě čistit, harmonizovat a propojovat. Je třeba zmínit, že vytváření sdílených databází mikrodat, z nichž by hodnotitelé mohli čerpat, znamená nejenom úsporu na nákladech, jak bude podrobněji diskutováno níže, ale v řadě případů rovněž otevírá možnosti pro provedení analýz, které by jinak nebyly myslitelné. Bez systematického řešení přístupu k mikrodatům zůstanou bez ohledu na pokroky v metodách kvalitně provedená hodnocení velkou vzácností.

### **Kvantitativní či kvalitativní metody?**

Kvantitativní metody mají řadu nesporných výhod, ale kvalitativní metody mají v hodnocení rovněž místo. Na hodnotitele je často vyvíjen tlak, aby jejich závěry byly co možná nejvíce kvantifikovatelné. Nicméně kvantitativní výsledky zpravidla samy o sobě neposkytují dostatečné sdělení pro strategické rozhodování, protože k tomu je nejdříve musíme správně kvalitativně interpretovat a zasadit do ekonomického, společenského a politického kontextu. Kvantitativní výsledky jsou navíc zpravidla ex-post, zatímco pro rozhodování je třeba i kvalitativní ex-ante vzhled do problému. Nejlepší je umět oba přístupy správně vyvážit a o hodnocení uvažovat jako o mixu různých metod.

## 7. Datová základna

---

Úroveň hodnocení stojí a padá s daty, na nichž je založeno. Dostupnost relevantních dat je nejčastějším problémem, který hodnocení omezuje a jehož vlivem zůstává za svými možnostmi. Kvalita dat je důležitější než jejich kvantita. Příliš nepomůže mít k dispozici rozsáhlou databázi, pokud v ní chybí klíčové proměnné pro vyhodnocení daného programu. Stejně tak je nutné mít na paměti, že používaná data musí být kvalitní z hlediska jejich reprezentativnosti ve vztahu k cílové populaci, protože bez toho nelze zobecňovat nad rámec souboru pozorování, který máme k dispozici.

Podle rozsahu a kvality dat je zároveň nutné volit metody analýzy. Ani největší expert na metody nic nezmuže, pokud nemá v ruce data, jež potřebuje. Rozhodování o datech, případně o nutnosti rozšířit stávající datovou základnu, a výběr analytických metod je třeba synchronizovat. Dostupná data určují vhodné metody a zpětně ze znalosti použitelných metod odvozujeme požadavky na data, přičemž mezi nimi hledáme průnik. Nejvhodnější řešení je takové, které dokáže vytěžit nejvíce z obou. Proto je třeba otázkám dat věnovat velkou pozornost.

### 7.1 ZDROJE DAT

Předběžná představa o zdrojích dat dostupných pro hodnocení by měla být předložena již v návrhu programu. Podrobněji by otázka dat měla být vyjasněna ještě před spuštěním projektu hodnocení. Není radno nechat tuto otázku otevřenou s tím, že se zdroje dat podaří vyřešit až za pochodu, protože může lehce dojít k nepříjemným překvapením, a protože s rozpisem nákladů na pořízení dat je nutno pracovat již při sestavování rozpočtu. Na začátku hodnocení je třeba udělat vyčerpávající přehled dat relevantních k zodpovězení hodnotících otázek, která již jsou k dispozici, která sice nejsou k dispozici, ale existují někde jinde, a která ještě neexistují, a bylo by tudíž přínosné je sesbírat.

Základní druhy zdrojů dat jsou následující:

- **Data z monitorování** – posbíraná poskytovatelem v průběhu implementace programu – klíčový zdroj pro hodnocení průběžné, ex-post i dopadů, jež zahrnuje nejenom data povinně zadávaná do IS VaVal, ale i přímo od poskytovatele, včetně dat z procesu hodnocení žádostí o projekty, jako jsou údaje vyplněné do žádostí, seznamy financovaných i nepodpořených projektů, jejich pořadí nad „čarou“ financování a pod ní, rozdělení žádostí mezi zpravodaje apod.
- **Sekundární databáze** – získatelná z existujících databází – záleží na kvalitě datové infrastruktury v zemi, zahrnuje data z předchozích hodnocení, volně přístupných databází, komerčních databází a data sbíraná veřejnými institucemi, jako je statistický úřad; výhodou je dostupnost dat i za kontrolní skupinu, tj. nežadatele o podporu, a zpravidla za delší časové řady, avšak problém může nastat při zajištění přístupu.
- **Primární šetření** – v písemné (dotazníkové) podobě anebo pomocí rozhovorů – používá se pro získání dat jinak nedostupných, jež je nutné nově sesbírat v průběhu hodnocení, většinou jsou z důvodu nákladů na sběr omezena na příjemce podpory, takže chybí údaje za kontrolní skupinu, problematická je míra odezvy respondentů, a tudíž reprezentativnost sesbíraných dat.

Při sestavování datové základny se postupuje shora dolů. Nejdříve je třeba plně vytěžit monitorovací data od poskytovatele, následně se pokusit získat data ze sekundárních databází a až potom uvažovat o vykrytí zbývajících mezer pomocí primárních dat. Kromě vysokých nákladů na sběr primárních dat a možná zbytečného zatěžování



případných respondentů, pokud dané údaje existují již někde jinde, je hlavním důvodem pro takový postup, že míra návratnosti vlastního šetření bývá velmi nízká. Monitorovací data bývají téměř úplná, sekundární data mívají solidní pokrytí, i když záleží na zdroji a nějaké ztráty pozorování jsou nevyhnutelné, avšak pomocí primárního šetření se většinou podaří získat data za menšinu, ne-li pouze zlomek cílového souboru. Primární data mohou být cenná, ale to je vyváženo rizikem, že mohou být zatížena výběrovým zkreslením, a tudíž obtížně zobecnitelná.

Nicméně pokud není vyhnutí, což je zvláště případ programů zaměřených na řešení úzce vymezených problémů, pro které se relevantní data běžně nesbírají, je pro důvěryhodnost hodnocení kriticky důležité, aby sběr primárních dat proběhl profesionálně a byl odpovídajícím způsobem zdokumentován v hodnotící zprávě. Sběr primárních dat pomocí řízených rozhovorů, expertních panelů anebo fokusních skupin se zástupci zainteresovaných subjektů pro zodpovězení otázek ohledně relevance, užitečnosti a udržitelnosti probíhá ve většině hodnocení, protože jinak ani nemusí být možné se ke kýženým postřehům dostat. Řízené rozhovory mohou být hloubkou získaných poznatků velmi přínosné i pro hodnocení efektivnosti a efektivity, a to zejména u menších programů, v jejichž případě mohou pokrýt významnou část příjemců. Nezbytnou součástí takového hodnocení by však vždy měla být analýza reprezentativnosti sesbíraných dat.

## 7.2 INDIVIDUÁLNÍ DATA

Pokročilejší analytické metody zpravidla vyžadují individuální údaje, tj. mikrodata, ať už na úrovni jednotlivých organizací, jako firem, výzkumných ústavů a univerzit, anebo ještě podrobněji na úrovni pracovišť, výzkumných týmů či projektů. Za většinou agregovaných statistik, které souvisí s programem, lze nalézt mikrodata, na jejichž základě byly spočítány, takže individuální údaje v určité podobě zpravidla existují. Problém tkví v tom, že při agregaci mikrodat na úroveň určitých skupin, jako mohou být ekonomická odvětví, sektory provádění, velikostní skupiny, regiony či dokonce program jako celek, nevyhnutelně dochází ke ztrátě potenciálně cenných informací o jednotlivých zpravodajských jednotkách. Jinými slovy se agregací snižuje variabilita, která může být zásadní pro vyhodnocení účinku podpory, jenž nás nejvíce zajímá.

Konečným příjemcem podpory nikdy není odvětví, sektor anebo region, ale konkrétní mikrosubjekt, který se rozhoduje o tom, jestli o ni bude žádat, do jehož rozpočtu peníze z podpory plynou, pod svou hlavičkou provádí podpořené činnosti a ve většině případů má z podpory i přímý prospěch. Spojení mezi vstupy, výstupy, výsledky a dopady probíhá na mikroúrovni a podcenění různorodosti mezi jednotlivými subjekty v tomto procesu může vést k zavádějícím závěrům. Pokud analyzujeme agregovaná data, nemůžeme si být nikdy jisti, že zjištěné výsledky nejsou pouze statistickým klamem (tzv. *ecological fallacy*) vzniklým právě z důvodu seskupení mikrodat. Koneckonců účinky na agregátní úrovni, které nás rovněž mohou zajímat, lze do modelů definovaných na individuální úrovni zabudovat, takže je v každém případě lepší používat přímo mikrodata.

Příklady rozsáhlých zdrojů sekundárních mikrodat:

- Informační systém výzkumu, experimentálního vývoje a inovací (IS VaVal) – veřejně přístupná data.
- Centrální registr dotací (IS CEDR III) – veřejně přístupná data.
- Obchodní a živnostenský rejstřík – veřejně přístupná data.
- Registr ekonomických subjektů ČSÚ – veřejně přístupná data.
- Šetření ČSÚ (průzkum o inovacích, roční šetření o VaV, šetření o licencích apod.) – chráněná mikrodata dostupná pouze v anonymizované podobě.
- Bisnode (Artemis, Albertina, Magnusweb), Bureau van Dijk (Amadeus) – komerční databáze.



- Bibliometrické údaje (Web of Science, Scopus, Google Scholar apod.) – komerční databáze.
- Úřad průmyslového vlastnictví, PATSTAT apod. – poloveřejné patentové databáze.
- Informační systémy veřejných vysokých škol, Akademie věd ČR, ministerstev apod. – ad hoc přístupná data veřejných institucí.

Je třeba mít na paměti, že kromě těchto velkých sekundárních databází s celostátní působností existují i menší lokální zdroje mikrodat na úrovni jednotlivých poskytovatelů i mimo monitorovací systém hodnoceného programu, včetně historických údajů z již provedených hodnocení předchozích programů, a rovněž mikrodata sesbíraná v rámci nespočtu výzkumných projektů v oblasti sociálních věd financovaných z veřejných zdrojů, která se mohou ukázat jako cenná pro hodnocení konkrétních programů.

### 7.3 PŘÍSTUP K DATŮM

Nad rámec mikrodat z monitorování programu by měl hodnotitel identifikovat i všechny ostatní relevantní zdroje. Pokud mikrodata nejsou volně dostupná anebo k dispozici skrze navrhovatele či poskytovatele, měl by se pokusit si k nim zajistit přístup. Nemusí se to vždy z důvodů právních, procesních či finančních podařit, ale je žádoucí, aby hodnotitel vyjádřil poptávku po těchto mikrodtech, a tudíž aby jejich držitelé byli srozuměni s tím, že existuje potenciál jejich využívání pro účely hodnocení. Případné tvrzení, že hlubší a pokročilejší analýzu není možné provést, protože k tomu nejsou mikrodata, by zadavatel neměl akceptovat, pakliže hodnotitel neprokáže, že na získání přístupu k těmto datům vynaložil přiměřené úsilí.

Stejná data by neměla být sbírána opakovaně. Zejména je třeba se vyvarovat zbytečného sběru dat, která jsou již dostupná v sekundárních zdrojích. Úspory z rozsahu plynoucí z propracované datové infrastruktury, jež pro účely hodnocení umožňuje sdílený přístup k již existujícím sekundárním datům a jejich propojování s monitorovacími a primárními daty, jsou enormní. Jedná se v podstatě o nutnou podmínku k tomu, aby mohlo být hodnocení pravidelně prováděno na vysoké kvalitativní úrovni, protože náklady na opakované získávání, čištění a propojování stejných mikrodat pro každé jednotlivé hodnocení by byly příliš vysoké. Pro zajištění přístupu k již existujícím datům je nutné odstranit přežitě právní a technické překážky, přičemž v tomto směru se lze poučit ze zemí s vyspělou kulturou hodnocení, kde systémy pro sdílený přístup k mikrodátům, včetně chráněných údajů, rutinně fungují.

Dobrou praxí je vytvoření centrální databáze mikrodat, obvykle pod kuratelou národního statistického úřadu, ve které jsou data ze sekundárních zdrojů propojována s monitorovacími a primárními daty, zajištěna je jejich vzájemná kompatibilita, jsou vytvářeny dlouhé časové řady a do níž mají při dodržování přísných pravidel ochrany osobních údajů umožněn přístup nejenom výzkumní pracovníci, nýbrž i hodnotitelé podpůrných programů. Při dnešní úrovni informačních technologií je technické zabezpečení pouze okrajový problém k řešení, avšak významné překážky je nutné překonat v oblasti právní, a to zejména ohledně spolupráce mezi organizacemi.

#### **Kdy myslet na data pro hodnocení?**

Přístup k datům je nutné řešit ještě před spuštěním programu a následně vyjasnit v zadávací dokumentaci hodnocení. Je pozdě začít se zabývat otázkami dat až v průběhu hodnocení, protože to obvykle vede ke špatně provedené práci s výmluvou, že „pro lepší hodnocení nebyla data“. Z toho také plyne, že otázku potřebných dat nelze zcela delegovat až na hodnotitele, protože nutnost jejich zajištění musí mít na paměti hned od prvopočátku tvůrce politik i poskytovatel programu.

## 8. Systém hodnocení

---

K hodnocení programů účelové podpory VVI je třeba přistupovat systémově. Měl by to být zásadní zdroj podkladů pro rozhodování, nedílná součást řízení VVI politiky, a to zejména z toho důvodu, že finanční a ostatní signály odvozené od úspěchu na trhu často chybí. Na jednotlivá hodnocení by se nemělo nahlížet jako na jednorázové činnosti pro konkrétní účely, které spolu vzájemně nesouvisí, nýbrž jako na součásti provázaných činností zapadajících do celku, jehož účelem je vytvářet relevantní podklady pro VVI politiky. Jednotlivá hodnocení, jakkoliv mohou být kvalitní, kýženou změnu fungování celku nepřinesou, pokud nebudou integrována do soukolí řízení politik VVI.

Systém hodnocení uspořádává jednotlivé procesy, předepisuje jejich souslednost, poskytuje vodítka pro zodpovězení otázek, kdo hodnotí, jak a kdy, jak se hodnocení financuje a jaké má hodnocení důsledky. Páteř systému hodnocení by měla být vymezena v zákonech, nařízeních a směrnících, jejichž implementace je podrobněji rozvedena v příručkách, metodikách či jiných pokynech. Nastavení systému hodnocení by mělo přinejmenším pokrýt následující témata:

- **rozdělení rolí** mezi aktéry systému,
- **souslednost** jednotlivých **procesů** hodnocení,
- **funkce** hodnocení v systému,
- **kdo provádí** hodnocení (externí anebo interní hodnocení),
- **kdo zadává** hodnocení a jakým způsobem,
- **jak se** hodnocení **financuje** (zdroje vyčleněné pro hodnocení),
- **důsledky** hodnocení (jak se jeho výsledky projeví v praxi),
- způsob **zveřejnění výsledků** hodnocení,
- **ukotvení** hodnocení (v legislativě apod.).

### 8.1 ROLE AKTÉRŮ

Systém hodnocení zahrnuje aktéry s následujícími rolemi:

1. **Tvůrce politik** – má na starosti strategické směřování programů VVI. Společenskou poptávku po dosažení určité změny převádí do priorit pro vládní zásahy, čímž vymezuje životní prostor pro programy, tj. nastavuje „terč“, do kterého se strefují návrhy programů od poskytovatelů. Činí politická rozhodnutí, zajišťuje koordinaci a doporučuje návrhy programů k financování. Jeho hlavním nástrojem jsou strategické vládní dokumenty, nařízení a rozhodnutí.
2. **Poskytovatel** – provádí taktické řízení programů VVI. Přichází s řešením v rámci priorit vymezených tvůrcem politik. Předkládá návrhy programů a plní roli jejich správce. Rozhoduje o podpoře pro jednotlivé projekty v programech a tuto podporu poskytuje. Jeho hlavním nástrojem jsou návrh programu, zadávací dokumentace programu a obecně vše, co souvisí s jeho implementací.
3. **Zadavatel** – zadává hodnocení programů VVI buď interně pro svůj vlastní program, anebo externě pro programy jiných aktérů. Stanovuje základní parametry hodnocení, mezi které patří cíle, funkce, metody

rozpočet a časový rozvrh hodnocení. Jeho hlavním nástrojem je zadávací dokumentace hodnocení, případně smlouva s externím hodnotitelem.

4. **Hodnotitel** – provádí hodnocení programů VVI ať už interně v rámci poskytovatele, anebo externě. Podle zadávací dokumentace hodnocení navrhuje jeho projekt a zajišťuje hodnotící práce. Na základě odborného posouzení poskytuje zpětnou vazbu k programům. Přináší podložená doporučení pro rozhodování. Jeho hlavním nástrojem je hodnotící zpráva.
5. **Veřejnost** – sleduje, kontroluje a vyjadřuje se k dění v systému. Tvoří ji všechny ostatní zainteresované subjekty v oblasti programů VVI, mezi které patří i cílová skupina programu, příjemci podpory, uživatelé přínosů programu, zájmové skupiny, nezávislí odborníci i široká veřejnost. Jejich hlavním nástrojem jsou sdělení pro ostatní aktéry, mezi které může patřit i pro bono hodnocení programů.

Systém hodnocení potřebuje ústředního tvůrce politik, hierarchicky vrcholného aktéra, který formuluje národní politiku VVI, sestavuje národní rozpočet VVI, synchronizuje programy VVI napříč poskytovateli, rozhoduje případné spory mezi aktéry a koordinuje systém jako celek. Resortní politiky VVI by měly být ústředním aktérem koordinovány, aby vytvářely synergie a řídily se národními prioritami. V konečném důsledku je takovým aktérem vláda, nicméně je přínosné, aby byl těmito pravomocemi pověřen specializovaný subjekt, který bude mít k jejich aktivnímu vykonávání odpovídající zdroje, administrativní podporu a odborné zázemí. Ústřední tvůrce politik musí mít silné výkonné kompetence a legitimitu pro politická rozhodnutí, takže by se mělo jednat o součást vlády.

Hodnotitel musí být nestranná autorita, což je podstatně snadněji dosažitelné při externím než interním hodnocení. Uvnitř poskytovatele je třeba striktně rozlišovat organizační jednotku, která zodpovídá za hodnocení, tj. hodnotící jednotku, od organizační jednotky, která program spravuje, tj. implementační jednotku. Hodnocení, pokud je zadáno interně, musí provádět jiná jednotka než ta, jež implementuje program. Podobně zadavatel hodnocení může být interní, tj. vedoucí orgán poskytovatele, anebo externí, tj. nadřazený aktér v systému, přičemž druhá možnost je opět příslibem větší nestrannosti.

Pro hladké fungování systému je přínosné zřetelně rozdělit role mezi aktéry. Konkrétní aktéři, tj. jednotlivé instituce či některé jejich části, v praxi často zastávají více než jednu roli. Nicméně v obecné rovině lze doporučit, aby se aktéři soustředili na svou hlavní úlohu, kterou se naučí plnit nejlépe, a nesnažili se kumulovat příliš mnoho dalších pravomocí. Systém by měl být nastaven tak, aby mezi jednotlivými aktéry zbytečně nedocházelo k překrývání rolí, jež může vytvářet třetí plochy. Souběh některých rolí, jako tvůrce politik a zadavatele hodnocení, je přirozený, a tudíž žádoucí. Naopak souběh jiných rolí vede k přímému střetu zájmů, který by měl být záměrně omezen.

Ústřední tvůrce politik by neměl zastávat roli významného poskytovatele, protože v systému není hierarchicky výše postavený aktér s odpovídajícími odbornými kapacitami, jenž by mohl mít odpovědnost za hodnocení jeho programů. Neměl by tímto způsobem také konkurovat poskytovatelům, protože by byl v pozici tvořit národní politiku šitou na míru svým programům, a tím získávat rostoucí část národního rozpočtu pro sebe. Jeho ústřední role v systému vyžaduje, aby nebyl ostatními aktéry vnímán jako jeden z nich, nýbrž jako aktér hierarchicky nadřazený, jehož mise je z tohoto titulu hájit zájmy systému jako celku. Role jednoho z poskytovatelů by plnění této úlohy narušovala.

K silnému střetu zájmů dochází mezi rolemi poskytovatele, zadavatele a hodnotitele. Nikoho by neměl překvapit požadavek, aby návrhy jednotlivých projektů v rámci programu byly hodnoceny na základě nezávislých posudků. Na úrovni projektů je takový proces hodnocení široce přijímán jako dobrá praxe. Stěží si lze představit, že by si žadatel o projekt určoval, jak hodnocení proběhne, podle jakých kritérií anebo kdo konkrétně bude jeho žádost posuzovat,

nemluvě o možnosti, že by si ji dokonce posoudil sám. Závěry hodnocení ohledně toho, zda takovou žádost financovat, by byly nasnadě. Sebehodnocení je ze stejných důvodů problematické i při hodnocení programů.

Nejzřetelnější je tento střet zájmů pro ex-ante hodnocení. Nedává smysl, aby poskytovatel sám vybíral hodnotitele anebo přímo prováděl hodnocení vlastního návrhu programu, pokud ze závěrů hodnocení plynou rozpočtové požadavky. Ex-ante hodnocení tudíž musí interně provádět anebo externě zadávat nadřízený aktér, což je ústřední tvůrce politik. Sebehodnocení je podobně problematické i u hodnocení ex-post a nejvíce u konečného verdiktu ohledně dopadů, které rovněž mohou mít závažné rozpočtové důsledky. Obecně lze říci, že riziko zkreslení je v případě sebehodnocení veliké, jestliže je převládající motivace k hodnocení externí. Poskytovatel nebude sám sebe hodnotit kriticky, anebo to přinejmenším nebude vyjádřeno ve veřejné části hodnotící zprávy, pokud to může vést ke krácení jeho rozpočtu.

Sebehodnocení má naopak nesporné výhody, pokud je hlavní motivací pro jeho provedení interní snaha poskytovatele zlepšit svoje fungování. Pokud hodnocení probíhá pod střechou stejné organizace, může těžit z lepší znalosti zkoumaného programu, může být lépe zasvěcené do kontextu a může vykazovat lepší chápání možností změn. Sebehodnocení má rovněž lepší přístup k interním informacím, může poskytnout kýžené odpovědi rychleji a v podobě, která je pro uživatele lépe stravitelná, protože komunikace uvnitř úřadu je jednodušší. Zesiluje se tím zpětná vazba, a tím i celková užitečnost hodnocení. Bohužel interní učící funkce není jediná, kterou musí většina hodnocení zastávat. Pro splnění ostatních funkcí, jako legitimizační, informační, usměrňovací a mediační, má zásadní přednosti externí hodnocení v podobě nezávislosti, nadhledu a odstupu od zájmů poskytovatele.

Jediný typ hodnocení, který má převážně funkci učící, který je úzce propojen s řízením a pro který tudíž převažují výhody interního provádění, je hodnocení průběžné. Poskytovatel by měl mít možnost průběžné hodnocení provádět operativně, zaměřit se na témata, jež sám považuje za klíčová pro zlepšení chodu programu, způsobem, který zaručuje nejlepší využití výsledků pro jeho interní potřeby. Průběžné hodnocení je rovněž nejvíce založeno na interních datech z monitorování. Pokud by za průběžné hodnocení byl zodpovědný externí aktér, vyvstal by problém střetu kompetencí, protože by v podstatě zasahoval do řízení programu, za něj je ovšem zodpovědný poskytovatel.

Sebehodnocení může fungovat i v případě hodnocení ex-post, avšak pouze v zemích s vyspělou kulturou hodnocení, potažmo vysokou profesionální úrovní státní správy a politickou kulturou jako takovou. Nutkání ovlivňovat hodnocení, které je organizováno uvnitř poskytovatele, ale které má v systému plnit externí kontrolní funkce, je enormní. Pro země, které nejsou v úrovni hodnocení na špici, je vhodnější hodnocení ex-post zadávat externě anebo použít kombinovaný přístup, který co nejlépe využívá omezenou odbornou kapacitu, ať už se v systému nachází kdekoliv, a který v sobě spojuje funkce učící i kontrolní, například formou smíšených týmů interních a externích hodnotitelů, včetně aktérů ze zahraničí.

## 8.2 ZPĚTNÁ VAZBA

Hodnocení se nesmí provádět samoúčelně, za účelem hodnocení samotného, ale musí sloužit jako podklad pro rozhodování. Smyslem hodnocení je poskytnout poznatky, které povedou k navazující akci. Zda k tomu dojde, záleží na tom, jestli se s hodnocením seznámí lidé v rozhodovacích pozicích. Pro lepší užitečnost výsledků hodnocení je přínosné, aby aktér zodpovědný za hodnocení měl na paměti následující:

- **Vyžadovat kvalitní hodnocení** – trvat na použití nejlepších dostupných metod, aby výsledky byly obhájitelné i před mezinárodní odbornou veřejností, požadovat formulaci konkrétních doporučení a zajistit jejich relevanci pro nadcházející rozhodovací procesy.
- **Podporovat kritický přístup** – oceňovat, nikoliv zahlazovat popsání problémů, podporovat kritickou diskusi na základě věcných argumentů, požadovat průkazná stanoviska.
- **Zabudovat do rozhodování** – provádět hodnocení systematicky a pravidelně, načasovat hodnocení v souladu s rozhodováním a provázat výsledky hodnocení s plánováním a rozdělováním rozpočtu.
- **Sledovat implementaci závěrů** – organizovat na hodnocení navazující diseminační aktivity (konference, semináře, workshopy atd.), závěry převádět do akčního plánu a kontrolovat jeho realizaci.
- **Stimulovat zájem o výsledky** – zajistit si pozornost vysoce postavených rozhodovatelů, vtáhnout zainteresované subjekty do procesu hodnocení, aby byly aktivně zapojeny, brát v úvahu jejich potřeby, a to již při formulaci zadání, a dát jim vědět, co mají čekat.
- **Zajistit celoživotní vzdělávání** – seznámit s metodami hodnocení nikoliv pouze hodnotitele, nýbrž i tvůrce politik a poskytovatele, aby měli alespoň základní pochopení, co jim výsledky sdělují, jak je správně interpretovat.
- **Srozumitelnost, atraktivnost a čitelnost** – zajistit, aby byla hodnocení zpráva zajímavá i pro širší laickou veřejnost, nepodceňovat formátování, design a rozvržení textu, ale zároveň požadovat i zveřejnění technických podrobností ohledně dat, modelů, výpočtů atd. v přílohách, které zajímají specializované odborníky.
- **Široký okruh čtenářů** – hodnocení zprávy zveřejnit, nejlépe je poskytnout volně ke stažení on-line, pokud tomu nebrání vážné důvody, jako utajení.

Zásadním předpokladem pro účelné využívání výsledků hodnocení je jeho správné načasování. Hodnocení zprávy by měly být připraveny v okamžiku, kdy výsledky v nich obsažené mají největší šanci být užitečné v rozhodování. Například to znamená, že výsledky hodnocení musí být k dispozici s dostatečným předstihem, než má padnout rozhodnutí. Zveřejnění sebelépe provedeného hodnocení o jeden den později by již vyznělo do prázdna. Aby byly výsledky ve správný čas, je nutné hodnocení plánovat v návaznosti na budoucí termíny rozhodování, a to střízlivě v dostatečném předstihu, nikoliv je zadávat na poslední chvíli. Následující doporučení je přínosné vzít v úvahu při plánování časového rozvrhu hodnocení:

- Výsledky hodnocení jednotlivých projektů je možné využít jako vstup do (zejména průběžného a ex-post) hodnocení programu jako celku.
- V případě dlouhotrvajících programů může průběžné hodnocení hrát z části i úlohu ex-ante hodnocení plánů pro podobu jeho další fáze (nově připravovaných výzev) a zároveň i úlohu ex-post hodnocení dosavadních přínosů.

- Načasování termínů pro politická rozhodnutí by mělo brát v úvahu časové nároky a omezení při hodnocení, tj. například co je schůdné vyhodnotit do daného termínu z hlediska dostupnosti aktuálních dat, a tudíž by plánované kroky v hodnocení a rozhodování měly být synchronizovány.
- Pokud je hlavní pozornost zaměřena na přínosy programu, je prospěšné rozložit hodnocení do několika období po jeho skončení, protože různé druhy účinků mají odlišná časová zpoždění, například působení na vstupy či chování příjemců probíhá souběžně s poskytováním podpory, zatímco konečné přínosy mohou být pozorovány až s několikaletým zpožděním.

Dlouhá časová zpoždění je třeba očekávat zejména ohledně hodnocení dopadů. Pro většinu programů jsou dopady pozorovatelné v průběhu několika let po skončení podpořených projektů, nicméně jsou oblasti výzkumu, u kterých jsou časová zpoždění daleko delší. Zároveň širší přínosy i mimo příjemce podpory, jež například vyžadují šíření vytvořených znalostí ve společnosti, trvají podstatně déle než přínosy přímo pro příjemce. Nemá smysl organizovat hodnocení dopadů rok po skončení programu, když jeho hlavní přínosy budou pozorovatelné až za pět let. Zpětná vazba na rozhodování bude opožděna, nicméně počkat si na poučení z nich se vyplatí, protože právě dosažení širokých, hlubokých a trvajících dopadů je to hlavní, oč by programům mělo jít. Zpoždění by rozhodně neměl být důvod dlouhodobé dopady nehodnotit.

Z hlediska implementace je důležitou otázkou, do jaké míry by se měl tvůrce politik řídit výsledky, doporučeními a závěry hodnocení, tj. zda by mělo být hodnocení určující pro následný rozhodovací proces. Na jednu stranu by byl přehnaný požadavek, aby se rozhodnutí muselo řídit výsledkem hodnocení, protože pak by tvůrci politik nezůstávala žádná pravomoc. Závěry hodnotitele by plně nahrazovaly politické rozhodnutí, což nelze doporučit. Na druhou stranu by tvůrce politik neměl hodnocení ignorovat, protože to by bylo proti samotnému smyslu, proč se hodnocení provádí, a zdroje vynaložené na hodnocení by tím byly promrhány.

V konečném důsledku záleží na kultuře rozhodování, která se v přístupu k hodnocení nastaví. Tvůrce politik by měl mít možnost ve velmi pečlivě odůvodněných případech rozhodnout jinak, než doporučuje hodnocení, avšak mělo by to být vnímáno jako nestandardní postup pro výjimečné situace. Rozhodnutí proti závěrům hodnocení by mělo být omezeno zejména na situace, kdy má tvůrce politik pochybnosti ohledně provedení samotného hodnocení, pokud je shledáno z odborného hlediska jako obzvláště nezvládnuté, či dokonce záměrně zkreslené. Nekvalitně provedené hodnocení by mělo mít pro hodnotitele vážné následky.

Mezi aktéry s různými rolami by měly být nastavené kanály pro komunikaci, vzájemnou spolupráci a synchronizaci úsilí. Samotná hodnocení jsou jednou z forem komunikace, protože jejich závěry je třeba diskutovat napříč aktéry, tj. sdílet nahoru i dolů po hierarchické linii, aby byla kultivována společná paměť a probíhal proces učení o tom, co funguje a co nikoliv. Pokud takový cyklus nefunguje, hrozí v lepším případě riziko selhání vlády v podobě neúčelně vynaložených veřejných prostředků na hodnocení samotné a v horším případě riziko zavádění škodlivých programů, nežádoucí pokračování programů, jejichž opodstatnění již vypršelo, skončení programů, které je třeba dále rozvíjet, a celkově rozhodování na základě horších znalostí, než které jsou, anebo by mohly být k dispozici.



## 8.3 ORGANIZACE PRŮBĚHU

Kvalita hodnocení, ať už interního anebo externího, závisí na schopnostech zadavatele celý proces dobře připravit a řídit. Zadavatel by měl na hodnocení v jeho průběhu aktivně dohlížet. Z pohledu zadavatele lze rozlišit čtyři fáze:

1. Příprava zadávací dokumentace – specifikace parametrů hodnocení, co se bude hodnotit a co je cílem.
2. Výběr vhodného hodnotitele – organizace výběru subjektu, který provede hodnocení samotné, tj. jestli bude interní, nebo externí a jakou formou bude případně vybrán externí hodnotitel.
3. Dohled nad prací hodnotitele – dodržování zvoleného postupu hodnocení, jeho kvality, plnění termínů atd.
4. Zpětná vazba na rozhodování – napojení závěrů hodnocení zpátky do rozhodovacího procesu.

Zadávací dokumentace vysvětluje, co je potřeba z hodnocení zjistit, uvádí obecné hodnoticí otázky a vymezuje mantinely z hlediska rozpočtu, termínů či metod, v nichž se hodnocení musí pohybovat. Příprava zadávací dokumentace vyžaduje dobrou znalost řešené problematiky, a tudíž investice do lidských zdrojů na straně zadavatele, pokud má obsahovat informace, které hodnotitelé potřebují vědět k tomu, aby připravili kvalitní projekt hodnocení. Zadávací dokumentace by měla přinejmenším splnit následující:

- Představit cíle, funkce a rozsah hodnocení, včetně uvolnění zdrojů k tomu určených.
- Určit hlavní kritéria hodnocení a hodnoticí otázky.
- Navést hodnocení určitým směrem, aby závěry mohly jít do hloubky.
- Zajistit posouzení, nikoliv pouze popis situace.
- Rozebrat kontext, ve kterém se hodnocení provádí.

Zadavatel musí být schopen vybrat externího hodnotitele na základě posouzení, zda splňuje parametry specifikované v zadávací dokumentaci. Zásadní je nastavit požadované parametry hodnocení tak, aby bylo zajištěno, že schopnosti vítězného hodnotitele budou na dostatečné úrovni, aby bylo hodnocení proveditelné v odpovídající kvalitě. Jako klíčový se v této souvislosti jeví rozsah dosavadních zkušeností z hodnocení programů, odborný profil členů hodnotitelského týmu, případně prokázání znalosti praktického použití pokročilých analytických metod, pokud je jejich uplatnění požadováno v hodnoticím procesu. K tomu je třeba hlubší pochopení problémů s daty, metodických otázek anebo technického provedení, pro jejichž rozklíčování může zadavatel potřebovat externí konzultaci.

Od začátku projektu až po odevzdání konečné verze hodnoticí zprávy by měl zadavatel aktivně provádět dohled nad průběhem hodnoticích prací, a to zejména sledovat, zda práce probíhají v termínu, podle rozpočtu a v odpovídající kvalitě. K tomu je potřeba, aby pracovníci zadavatele rozuměli praktickému provádění hodnocení. Jejich schopnost vyjednávat s hodnotitelem je důležitým předpokladem pro kvalitní hodnocení. Doporučuje se rozfázovat dodávání hodnoticí zprávy takovým způsobem, aby byl dostatečný prostor pro připomínkování a následné revize pracovních verzí. Přinejmenším by mělo proběhnout jedno kolo komentářů k první průběžné verzi a druhé kolo komentářů k draftu konečné zprávy.

Dobrou praxí je ověřit spolehlivost závěrů hodnocení skrze oponentní řízení, do kterého se zapojí externí experti na použitá data, metody a proces hodnocení obecně, anebo přinejmenším prostřednictvím diskuse v rámci řídicího výboru hodnocení. Je nutné mít se na pozoru zejména v případech, kdy jsou výsledky empirických metod překvapivé, či

dokonce v rozporu s renomovanými analýzami, výzkumy a světovou literaturou v dané oblasti, včetně předtím provedených hodnocení podobných programů.

Zadavatel hodnocení by se měl postarat o to, aby byly závěry hodnocení patřičně využity v rozhodovacím procesu. Hodnotící zpráva se musí dostat do rukou relevantních uživatelů, a tudíž musí být prezentována atraktivním způsobem. S tím souvisí organizace diseminačních aktivit, jejíž plán by měl být součástí návrhu projektu hodnocení. Nicméně diseminační aktivity by měl zaštitit, ve spolupráci s hodnotiteli organizovat a konečné slovo v tomto směru mít zadavatel, který mimo jiné drží duševní práva k hodnotící zprávě.

Hodnocení by mělo probíhat participativním způsobem. Hodnotitel by měl aktivně pracovat s lidmi a organizacemi, kterých se dotknou jeho závěry, a pro potřeby hodnocení využívat znalosti, názory a zkušenosti zainteresovaných subjektů. Během externího hodnocení je zejména důležitá dobrá komunikace s poskytovatelem, protože to je na jedné straně subjekt, který toho o praktickém řízení programů ví nejvíce, a na druhé straně se ho závěry hodnocení nejvíce dotknou. Poskytovatel by měl být srozuměn s účelem externího hodnocení, co se bude hodnotit, podle jakých kritérií, s případnými požadavky na data a pokud možno porozumět metodám hodnocení, aby v ideálním případě seznal, jaký může mít hodnocení užitek pro něho samotného.

## 8.4 FINANCOVÁNÍ

Na fungování systému hodnocení je rozumné vyčlenit zdroje, které odpovídají řádově 0,5 až 1,0 % z výdajů na programy účelové podpory. Může se to zdát hodně, ale pokud to zajistí, aby se nedělaly zbytečné chyby v utrácení toho zbývajícího celku, tak to jsou efektivně vynaložené peníze. Hodnocení, pokud je děláno kvalitně se zpětnou vazbou na rozhodování, je investicí do budoucnosti, do procesu učení, jak dělat lepší veřejné výdajové programy, nikoliv pouze provozní výdaj státního rozpočtu, který je třeba minimalizovat. Jedná se o investici, jež se může několikanásobně vrátit v podobě vyšších přínosů podpory, anebo ve zrušení zbytečných programů. Hodnota dobře provedeného hodnocení je sice obtížně vyčíslitelná, ale obecně lze tvrdit, že se vyplatí.

Kvalitní systém hodnocení nemůže být levný, protože je k tomu mimo jiné třeba vysoce kvalifikovaná, a proto drahá pracovní síla. Čím vyšší jsou požadavky na kvalitu, tím je obvykle celý systém rovněž dražší, ale bohužel nikoliv nutně naopak. Samotná výše rozpočtu na hodnocení tudíž sama o sobě nic neřeší, pokud tyto peníze protékají špatně fungujícím systémem, který nedokáže kvalitu vyžadovat, rozeznat a produkovat. Z toho plyne, že změny systému hodnocení a rozpočtu na jeho financování by měly proběhnout současně. Na jedné straně by bylo pošetilé se pokoušet o zavedení lepšího systému bez odpovídajících zdrojů na jeho provoz, ale na druhé straně navýšit rozpočet na hodnocení programů bez řezu do systému jejich provádění by nebylo efektivní.

Důležitou proměnnou při stanovování rozpočtů pro konkrétní hodnocení je jeho očekávaná délka. Ex-ante hodnocení jsou zpravidla kratší, bez větších nároků na data, a tudíž levnější. Důkladně provedené hodnocení dopadů však může zabrat hodně dlouho. Do hry vstupuje řada proměnných, včetně otázky dostupnosti dat, takže existuje velké rozpětí a rozpočet je nutné posuzovat jednotlivě. Dobře provedené monitorování programu, při kterém měl poskytovatel na paměti i sbírání dat, která budou potřeba k hodnocení, může jeho délku výrazně zkrátit. Naopak hodnocení se může prodražit, pokud musí napravovat ledabylost v předchozích fázích programového cyklu, například rekonstruovat intervenční logiku, protože chybí v návrhu programu, anebo zpětně sbírat data, jež chybí v monitorování, a tudíž zahrnovat i činnosti, které do hodnocení jako takového nepatří.

K tomu, aby se hodnocení mohlo začít provádět na solidní úrovni, je tudíž nutné zreformovat systém hodnocení jako celek. Dílčí řešení nebude fungovat, protože hodnocení pak bude narážet na úzká místa tam, kde to bylo ponecháno postaru. Není možné očekávat kvalitní hodnocení, pokud nemůže stavět na dobře zpracovaném návrhu programu,



odpovídajících datech z monitorování, přístupu k potřebným mikrodátům z externích zdrojů anebo spolupráci se zasvěcenými odborníky na straně navrhovatele a poskytovatele. Pokud by mělo hodnocení vyžadovat zdroje na vykrývání vnitřního dluhu systému jako celku, bylo by prohibitivně drahé. Hodnocení nemá napravovat nedostatky ostatních činností při zadávání, správě a implementaci programu, na něž musí navazovat.

## 8.5 REŽIMY PROVÁDĚNÍ

Z hlediska hospodárnosti nakládání s veřejnými prostředky musí být zdroje vynaložené na hodnocení přiměřené šíři, velikosti a důležitosti programu. Hodnocení programu podle určitých zásad by mělo být povinné. Nicméně tvůrce politik by měl mít možnost poskytnout úlevy z těchto požadavků. Hlavním důvodem by měl být nedostatečný rozpočet pro provedení plnohodnotného hodnocení. Zvláště výrazné úlevy lze předpokládat u menších a kratší dobu trvajících programů. Přínosné by bylo odstupňovat tři režimy hodnocení:

1. **Plnohodnotný režim** (pro velké programy) – celý cyklus hodnocení, mělo by to být povinné podle určitých měřítek jako délka v letech anebo výše rozpočtu, například pro programy trvající více než pět let s rozpočtem více než 1 mld. Kč.
2. **Zjednodušený režim** (pro střední programy, co do délky a velikosti) – úlevy z vybraných požadavků, například možnost vynechání průběžného hodnocení pro krátké programy, odpuštění hodnocení dopadů anebo obecně nižší požadavky na data a metody, avšak ex-post hodnocení by mělo zůstat povinné.
3. **Minimalizovaný režim** (pro malé a krátké programy) – možnost výrazných úlev v hodnocení průběžném, ex-post i dopadů anebo jejich úplné odpuštění, například pro programy trvající do tří let s rozpočtem méně než 100 mil. Kč.

Za standardní by se měl považovat plnohodnotný režim. Případné požadavky na úlevy by měly být předem odůvodněny v plánu hodnocení průběžného, ex-post a dopadů, který je součástí návrhu programu, a tudíž by měly být předmětem ex-ante hodnocení, na základě jehož o úlevách následně rozhodne tvůrce politik. Z toho rovněž plyne, že ex-ante hodnocení by mělo zůstat povinné pro všechny programy bez výjimky.

Programy mohou trvat až deset (či více) let, v jejichž průběhu se mohou změnit společenské potřeby, okolnosti poskytování podpory a rovněž vyvinout programy samotné, takže je třeba, aby v režimu hodnocení mohlo docházet ke změnám oproti plánu v návrhu projektu. Doporučení změn v hodnocení ex-post jsou jedním z možných závěrů průběžného hodnocení a podobně doporučení změn v hodnocení dopadů mohou vycházet z hodnocení ex-post. K provedení změn by mělo sloužit změnové řízení, které zpravidla iniciuje poskytovatel, ale tuto možnost, například pro zpřísnění režimu hodnocení rozšířeného programu, by měl mít i tvůrce politik.

## 8.6 UKOTVENÍ V ZÁKONĚ

Požadavek hodnotit programy ex-ante, průběžně, ex-post a jejich dopady na úrovni mezinárodních standardů by měl být ukotven jako zákonná povinnost při poskytování podpory. Neměla by existovat možnost se hodnocení zcela vyhnout, ačkoliv jak je také diskutováno výše, musí být zároveň možnost provádět hodnocení ve zjednodušeném režimu, pokud jsou k tomu objektivní důvody. Zákonný požadavek na hodnocení zaručí jeho pravidelnost a kontinuitu. Nebudou vznikat mezery v hodnocení programů, přičemž to bude právně vynutitelné. Zákonná povinnost zaručuje dlouhodobý horizont, což povzbudí investice do systému hodnocení, umožní vytvoření potřebné infrastruktury a celkově podpoří změnu kultury hodnocení k lepšímu, protože to zvýší hodnotu dobrého hodnotitelského jména.

Avšak rizikem ukotvení v zákoně je, že se na hodnocení bude ve stejném duchu jako doposud nahlížet jako na nechtěnou povinnost, která bude plněna pouze formálně. Hrozí odcizení zainteresovaných subjektů od podstaty hodnocení a jeho provádění bez vědomí hlubšího smyslu a úlohy v systému řízení. Forma může vítězit nad obsahem a zpětná vazba na rozhodování může zůstat zanedbatelná. Ačkoliv dobrovolnost a iniciativa zdola obecně funguje lépe než plnění povinností, ještě větší riziko než formální přístup spočívá v tom, že v případě dobrovolnosti by se hodnocení neprovádělo vůbec. Z tohoto důvodu se zákonná povinnost jeví jako záruka, že se hodnocení bude pravidelně provádět.

Nicméně je třeba zdůraznit, že vymezení zákonné povinnosti musí být dostatečně flexibilní, aby nepůsobilo rigidně jako „svěrací kazajka“, do které by se pak muselo vejít každé hodnocení. Musí být nechán dostatečný prostor pro použití rozličných přístupů a metod, protože jak je podrobněji rozvedeno níže, jednotlivá hodnocení se v podrobnostech provedení výrazně liší. Nelze zákonem stanovit jediný přístup, který by pasoval na všechny programy a musel by být použit stejným způsobem. Zákon nesmí vnucovat příliš velké detaily, na které pak bude nutné šroubovat každé hodnocení, i když to v konkrétní situaci nebude přínosné. Tvůrci politik musí být při zadávání ponechána dostatečná volnost, aby mohl vybrat nejlepší podobu hodnocení.

Zákon by měl zejména určit, jaké typy hodnocení bude povinnost provádět, režimy hodnocení, kolik z rozpočtu programu by mělo být použito na hodnocení, a vymežit role jednotlivých aktérů v systému hodnocení, tj. kdo ho bude zadávat, provádět a za jakých podmínek bude vyloučeno sebehodnocení. Ostatní by mělo být předmětem podzákonných norem, které by se postupem času měly vyvíjet s tím, jak se systém bude zlepšovat, usazovat a vyvíjet.

## 9. Nezávislost, nestrannost a etika

---

Pro spravedlivé hodnocení musí všechny zainteresované strany dodržovat určitá pravidla chování. Ne všechna pravidla, kterými se hodnocení řídí, však mohou být explicitně sepsána. Stejně důležitá jsou i nepsaná pravidla, která je třeba dodržovat z podstaty procesu hodnocení. Zainteresované strany musí mít na paměti, že nemají pouze formální vztah vůči sobě navzájem, nýbrž i neformální závazky vůči společnosti jako celku, protože je hodnoceno nakládání s penězi daňových poplatníků, a tudíž přijetím práce na hodnocení veřejného výdajového programu na svá bedra nepřímo přijímají i spoluzodpovědnost za nakládání s těmito penězi. Níže uvedené požadavky platí bez rozdílu pro externí i interní hodnotitele, respektive zadavatele.

### 9.1 NÁSTRAHY HODNOCENÍ

Evropská komise (2013b, str. 67) upozorňuje, že se hodnotitelé mohou ocitnout pod tlakem silných a zároveň protichůdných zájmů, jimž se nezávislost hodnocení a nebezpečí toho či onoho závěru protíví. Hodnotitelé musí být při vykonávání své profese připraveni čelit neoprávněným pokusům vměšovat se do způsobu, jakým je hodnocení prováděno, mezi které patří například:

- **Nevyvážená zadávací dokumentace** – snaha předepsat hodnocení takovým způsobem, aby výsledky byly v souladu s parciálními zájmy.
- **Nátlak od poskytovatele** – přehnaný vliv na proces hodnocení ze strany těch, kteří nesou za program odpovědnost.
- **Předvýběr hodnotitele** – pokusy prosadit do hodnotícího týmu (anebo konsorcia hodnotitelů) subjekty, které přijdou se „správnými“ závěry.
- **Přepisování závěrů** – tlak na hodnotitele, aby přetvořili výsledky, a tudíž změnili závěry na méně (či více) kritické.
- **Cenzura dat** – poskytnutí zkreslených, neúplných anebo vybraných údajů z procesu monitorování programu.

Je zřejmé, že ničivé vlivy tohoto typu na hodnocení působí. Nejsou k tomu imunní ani země s nejvyspělejší kulturou hodnocení, protože nějaký problém se občas vyskytne všude. Z toho důvodu je zásadní nastavit systém hodnocení takovým způsobem, aby v něm výše uvedené vzorce chování měly co nejmenší prostor, jako jsou kupříkladu zásady, že pro výběr externího hodnotitele probíhá transparentní výběrové řízení, že interní hodnotitel je oddělen od implementační jednotky, že do systému jsou zabudovány kontrolní a čisticí mechanismy, které případné přešlapy odhalí a uvalí odpovídající sankce.

Profesní standardy hodnocení obvykle přinejmenším zahrnují příslib, a to na straně hodnotitelů i zadavatelů hodnocení, dodržovat následující:

Používat **nejvhodnější**, spolehlivé a ověřené **metody**.

- Udržovat si **nezávislý přístup**, který je nestranný a věrně odráží výsledky.
- **Chovat se eticky** vůči různým stranám, jako jsou zadavatelé, zdroje dat či příjemci podpory.
- **Brát vážně postřehy** zainteresovaných subjektů, včetně poskytovatelů.
- Aktualizovat a **rozvíjet** svou **odbornost**.

Podrobněji jsou tyto požadavky rozvedeny například v etickém kodexu evaluátora a formálních standardech evaluací od České evaluační společnosti (ČES 2011 a 2013), German Evaluation Society (BEYWL, W. 2003) anebo American Evaluation Association (AEA 2004), které jsou v komunitě hodnotitelů obecně přijímány.

## 9.2 ZPŮSOBILOST HODNOTITELE

Hodnotitel se musí chovat zodpovědně. Hodnotitel by neměl slibovat něco, co není schopen splnit. Hodnotitelský tým má mít dostatečné odborné znalosti pro úspěšné provedení daného způsobu hodnocení. Pokud hodnotiteli chybí některá z kvalifikací nezbytných pro provedení hodnocení, musí dát dané omezení na vědomí zadavateli. Zásadní je, aby zadavatel i hodnotitel byli vzájemně srozuměni se záležitostmi používaných analytických metod, s riziky ohledně dat a jinými omezeními, na které mohou narazit. Je žádoucí, aby mezi nimi v průběhu hodnocení o případných problémech probíhala otevřená a transparentní komunikace.

Hodnotitel má ze své podstaty zvláštní vztah k subjektu, který hodnocení zadává a financuje, nemluvě o vztahu k poskytovateli. Na jedné straně by se měl snažit v co možná největší míře vyjít vstříc legitimním požadavkům zainteresovaných stran. Na druhou stranu se hodnotitel pohybuje na tenkém ledě, protože se může dostat do situace, že i dobře míněný požadavek může být v rozporu se zájmy třetích stran či s veřejným zájmem. Nezáviděníhodná je zejména situace, kdy vyjde najevo, že splnění požadavků zadavatele by bylo v rozporu s etickými principy hodnocení. Profesionální povinností hodnotitele je bránit svoji nezávislost. Nejtransparentnějším řešením případného konfliktu je jeho zveřejnění.

Hodnotitel musí dodržovat základní etické principy, na kterých je provádění jeho profese postaveno. Zejména se nesmí dostat v průběhu hodnocení do střetu zájmů. Hodnotitel tudíž musí zadavateli oznámit veškeré svoje vlastní zájmy ve vztahu k zainteresovaným subjektům, včetně vazeb politických, obchodních, finančních či kariérních. Hodnotitel rovněž ručí za to, že nejsou porušena legitimní práva třetích stran, včetně práv na ochranu osobních údajů.

## 9.3 NÁROKY NA ZADAVATELE

Zadavatel musí mít k dispozici dostatečnou úroveň odbornosti v požadovaném způsobu hodnocení, která bude zárukou toho, že zadávací dokumentace hodnocení bude zpracována kvalitně, že bude schopen o konkrétním provedení s hodnotitelem zasvěceně jednat, bude správně interpretovat výsledky a v neposlední řadě bude umět je dále využít v procesu strategického rozhodování. Pokud zadavatel potřebnou odbornou kvalifikaci postrádá, měl by své pracovníky nechat odpovídajícím způsobem vyškolit a následně jim vytvořit podmínky, aby si mohli udržovat přehled o nejlepší mezinárodní praxi.

Zadavatel musí respektovat nezávislost hodnotitele. Měl by se vystríhat přímých zásahů do procesu hodnocení a zejména se nesmí pokoušet ovlivňovat jeho výsledky, vyvíjet v tomto směru na hodnotitele nátlak, nebo mu dokonce podsouvat určité závěry. Komunikace mezi zadavatelem a hodnotitelem by měla probíhat transparentně, obě strany by se měly striktně držet podmínek uvedených v zadávací dokumentaci, potažmo ve smlouvě o hodnocení, a případné požadavky na odklon od původního záměru musí být podloženy věcnými argumenty a implementovány pouze na základě vzájemné dohody.

## 9.4 HODNOVĚRNOST HODNOCENÍ

Hodnocení musí být provedeno nestranně. Závěry hodnocení musí být podložené, spravedlivé a transparentní. Hodnoticí zpráva dokumentuje způsob jeho provedení v rozsahu, který umožňuje reprodukovatelnost. Odborně způsobilý čtenář by měl být schopen pochopit, jakým způsobem bylo přesně hodnocení provedeno. Silné, ale i slabé stránky, včetně hlavních předpokladů, na kterých bylo hodnocení postaveno, by měly být vysvětleny. Kontroverzní rozhodnutí ohledně designu, dat a metod hodnocení by měla být omezena na minimum. Účelem hodnocení není oslnit experimenty s nejnovějšími metodami, ale zajistit, aby výsledky byly spolehlivé, vyvážené a srozumitelné, a tudíž použitelné pro rozhodování na straně uživatele hodnocení. Pokud hodnotitel narazí na překážky, které nebylo možné odstranit, mělo by to být uvedeno, aby se na jejich překonání mohlo pracovat alespoň do budoucna.

Zahraniční hodnotitel není sám o sobě zárukou hodnověrnosti. U zahraničního hodnotitele je sice obecně nižší riziko přímých osobních vazeb či skrytých forem střetu zájmů, avšak na mezinárodním trhu se službami tohoto typu působí pouze omezený počet dostatečně velkých dodavatelů, takže vytváření osobních vazeb s českými subjekty je pouze otázkou času. Jedná se navíc převážně o ziskově orientované soukromé konzultační firmy, kterým záleží na velikosti svého tržního podílu a které naprosto přirozeně v první řadě maximalizují výnosy pro své vlastníky. Pokud by se navíc domácí hodnotitel zachoval tendenčně, mohl by se ve svém vlastním prostředí nevratně zdiskreditovat s fatálními důsledky pro budoucnost. Do hodnověrnosti tak vstupuje větší množství proměnných, přičemž některé z nich, jako morální profil, odborný kredit či motivace k hodnocení, mohou být důležitější než země původu.

### **Dodržování pravidel se musí vynucovat!**

Nezávislost, nestrannost a etické principy hodnocení je třeba brát velmi vážně, protože jinak se systém hodnocení může přeměnit na nástroj pro výrobu argumentů k prosazování parciálních zájmů. Na dodržování profesních standardů v konečném důsledku celý systém stojí a padá. Z tohoto důvodu je nutná nejen znalost těchto pravidel napříč aktéry v systému, ale i jejich kolektivní dodržování a především striktní vynucování. Klíčový význam v tomto směru hrají zadavatelé, ale svoji roli by měla sehrát i samočisticí schopnost hodnotitelské komunity, případně profesní organizace, která je sdružuje.

## 10. Návrh systémových změn

---

### 10.1 NASTAVENÍ SYSTÉMU

Současné nastavení systému hodnocení programů účelové podpory VVI umožňuje výrazné překrývání a kumulaci rolí jednotlivých aktérů, tj. rolí tvůrce politik, poskytovatele, hodnotitele a zadavatele. Řada aktérů plní hned několik, anebo dokonce všechny čtyři role najednou. Poskytovatel programu má být zpravidla i zadavatelem jeho hodnocení a často i samotným hodnotitelem, čímž dochází k vážnému střetu zájmů a vede to k tomu, že hodnocení nemůže důvěryhodně splnit všechny požadované funkce.

Kumulace rolí mezi aktéry by se měla omezit. Zejména role poskytovatele, zadavatele a hodnotitele je nutné v základním nastavení systému oddělit. Mohou existovat výjimky, pokud si to situace vyžaduje, například v případě průběžného hodnocení. Nicméně standardním požadavkem by mělo být rozdělení těchto rolí mezi různé aktéry. Oddělit by se měla i role ústředního tvůrce politik a poskytovatele, aby se prvně jmenovaný nedostával do střetu zájmů při rozdělování národního rozpočtu na programy.

Ústředním tvůrcem politik by měl být ustanoven jeden zastřešující aktér, který bude mít pro vykonávání této role vytvořeny podmínky. Ústřední roli by měla plnit sekce pro VVI při Úřadu vlády ČR, respektive z ní nově vznikající Úřad pro vědu, výzkum a inovace (ÚVVI), anebo výhledově Ministerstvo pro vědu, výzkum a inovace (MVVI), s RVVI jako poradním orgánem pro strategické otázky. Zde by se mělo soustředit kritické množství administrativních a odborných kapacit, měl by být k zajištění této role k dispozici odpovídající rozpočet a přiděleny potřebné kompetence.

Neznamená to, že by měl ÚVVI rozhodnutí diktovat. Tvorba národních politik musí probíhat za aktivního zapojení ostatních aktérů v systému, včetně například poskytovatelů, hodnotitelů i příjemců podpory. Úlohou ústředního aktéra by mělo být v lepším případě tvorbu politiky spíše moderovat, rozhodovat spory mezi ostatními aktéry a prostřednictvím doporučení pro schválení navrhovaných programů pro vládu dávat poskytovatelům mandát pro její vykonávání. Nicméně některý aktér musí dbát na chod systému hodnocení jako celku a vymezovat priority napříč systémem, protože v rámci omezeného rozpočtu nemohou být uspokojeny požadavky všech zainteresovaných stran.

ÚVVI bude moci rovněž lépe využívat úspory z rozsahu při centrálním řízení dat pro hodnocení. Centralizace již dnes existuje v podobě správy IS VaVal, na který by se však měly napojovat i další databáze. Při centralizovaných datech bude podstatně jednodušší vytvářet srovnávací skupiny nepodporovaných firem pro hodnocení jednotlivých programů, než kdyby si každý poskytovatel udržoval vlastní databázi, nakupoval externí data zvláště, a státní správa tudíž za stejná data platila opakovaně. Z pozice centralizované správy bude rovněž jednodušší jednání s ostatními vlastníky dat pro zajištění přístupu k jejich datovým zdrojům pro účely hodnocení.

Pod hlavičkou ÚVVI, a potažmo RVVI, by měl být zřízen samostatný poradní orgán pro hodnocení programů účelové podpory VVI. RVVI může tuto roli sama o sobě zastávat jen s velkými obtížemi, protože v ní musí být zastoupeni odborníci na široké spektrum témat, z nichž většina tudíž problematice hodnocení programů nemůže podrobně rozumět. Současná Komise pro hodnocení výsledků výzkumných organizací a ukončených programů (KHV), ustavená pod RVVI, je zase primárně orientována na hodnocení výzkumných organizací, což je sice příbuzné, ale z hlediska problémů k řešení odlišné téma. Nemá smysl, aby v KHV zasedali odborníci na hodnocení programů, pokud se velká část její agendy bude věnovat něčemu jinému. Není žádoucí, aby v KHV o hodnocení programů rozhodovali odborníci na jiné téma, kteří programům nerozumí. Pokud se programy začnou hodnotit kvalitně, bude mít samostatná Komise pro hodnocení programů (KHP) agendy k vyřizování více než dost.



Tabulka 5 ukazuje navrhované rozdělení rolí mezi aktéry. ÚVVI zodpovídá za ex-ante hodnocení návrhů programů, které provádí interně, anebo zadává externímu hodnotiteli. Sebehodnocení v principu nedoporučujeme, a to s jedinou výjimkou průběžného hodnocení, jehož hlavní funkce je interní učení. Poskytovatelé mají tudíž na starosti pouze průběžné hodnocení, které standardně provádí jejich interní hodnotící jednotka. Nicméně i v tomto případě by poskytovateli měla být ponechána možnost zadat průběžné hodnocení nebo jeho část externě, pokud se tak rozhodne učinit, protože k jeho provedení nemá dostatečnou interní kapacitu. ÚVVI nese odpovědnost i za hodnocení ex-post a dopadů, které zadává externě, anebo organizuje smíšeným způsobem. ÚVVI rovněž celkově dohlíží na kvalitu hodnocení v systému, včetně průběžného hodnocení.

Povědomí o tom, že hodnocení může být použito k externí kontrole, má zpětnou vazbu na chování poskytovatele. Funkce učící je tudíž obtížně slučitelná s ostatními funkcemi hodnocení, tj. legitimizační, informační, usměrňovací a mediační, které jsou kontrolní. Nicméně to neznamená, že učící funkce na úrovni poskytovatele i mimo něj není relevantní i pro jiné typy hodnocení než průběžné, a to zejména v systému, jenž v úrovni hodnocení výrazně zaostává, a tudíž se musí učit. Zásadní otázkou je, jak dosáhnout toho, aby nebyla příliš oslabena funkce učící, zatímco budou věrohodně zajištěny i ostatní funkce, které musí hodnocení plnit.

Z tohoto důvodu navrhuje, aby byla zavedena možnost, že hodnocení ex-post sice bude probíhat pod kuratelou ÚVVI, nicméně samotné hodnotící činnosti bude provádět smíšený tým (anebo konsorcium) skládající se ze zástupců různých aktérů, tj. hodnotící jednotky ÚVVI, interní hodnotící jednotky poskytovatele a externích hodnotitelů. Zapojení externích hodnotitelů může probíhat například formou veřejné zakázky na dílčí část hodnocení, anebo prostřednictvím dohod o pracovní činnosti (či provedení práce) vybraných nezávislých odborníků. Smíšený tým pro ex-post hodnocení má oproti sebehodnocení poskytovatele na straně jedné a čistě externímu hodnocení na straně druhé přinejmenším tyto výhody:

- Při hodnoticích pracích probíhá proces interaktivního učení mezi aktéry.
- Jednotliví aktéři mohou snadněji sdílet zdroje, včetně odborných kapacit.
- Pomůže to překlenout nedostatek interních hodnoticích kapacit ve státní správě.
- Bude zajištěn jednodušší přístup k interním datům poskytovatele.
- Důvěryhodnost závěrů hodnocení bude vyšší než při sebehodnocení.
- Závěry hodnocení budou považovat za vlastní všichni zainteresovaní.

Zpětná vazba hodnocení na poskytovatele bude v konečném důsledku probíhat přes sestavování národního VVI rozpočtu, které bude v kompetenci ÚVVI, tj. pokud budou poskytovatelé předkládat návrhy nových programů, které nebudou v souladu s výše uvedenými požadavky, včetně kvalitního plánu na následné hodnocení programu a využití poznatků z dříve provedených hodnocení, nebudou tyto návrhy doporučovány k financování. Poskytovatelé tímto budou tlačeni k tomu, aby rozvíjeli interní odborné kapacity, které budou procesu hodnocení rozumět, a zároveň aby patřičně zdůvodňovali navrhované programy. Případné odmítnutí poskytovatele připojit se do smíšeného hodnocení ex-post je nutno rigorózně odůvodnit. Samotná nechuť k hodnocení či hodnocení pomocí metod, které zatím nejsou příliš rozšířeny, není opodstatněná, a měla by tudíž mít pro poskytovatele rozpočtové následky.



Tab. 5 Navrhované rozdělení rolí v českých podmínkách

|  | Hodnocení ex-ante  | Průběžné hodnocení   | Hodnocení ex-post   | Hodnocení dopadů   |
|--|--|--|---|--------------------|
| <b>Kdo nese za hodnocení odpovědnost?</b>        | ÚVVI   | Poskytovatel   | ÚVVI  | ÚVVI               |
| <b>Kdo hodnocení provádí?</b>                    | ÚVVI, anebo externí hodnotitel (může být zadáno externě) | Poskytovatel, anebo externí hodnotitel (může být zadáno externě) | Smišený tým z ÚVVI, poskytovatele a externího hodnotitele, anebo pouze externí hodnotitel | Externí hodnotitel |
| <b>Kdo zadává externí hodnocení?</b>             | ÚVVI   | Poskytovatel   | ÚVVI  | ÚVVI               |
| <b>Kdo zajišťuje uplatnění závěrů hodnocení?</b> | ÚVVI   | Poskytovatel   | ÚVVI  | ÚVVI               |
| <b>Kdo dohlíží na kvalitu hodnocení?</b>         | ÚVVI   | ÚVVI   | ÚVVI  | ÚVVI               |

ÚVVI bude rozhodovat o režimu hodnocení, tj. bude mít možnost programu poskytnout úlevy z požadavků na hodnocení. Hodnocení ex-ante by mělo být povinným podkladovým dokumentem pro rozhodování ÚVVI, zda doporučit schválení návrhu programu vládě. Úlevy se tudíž budou týkat pouze hodnocení průběžného, ex-post a dopadů. ÚVVI by měl mít relativně velkou volnost při rozhodování, do jaké míry bude nutné takto hodnotit menší programy. Je třeba počítat s tím, že nový systém hodnocení bude nabíhat postupně, takže alespoň ze začátku bude obtížné zajistit kvalitní hodnocení většího počtu programů. Navíc bude nejprve nutné provést pilotní ověření metodik hodnocení pro různé druhy programů. Obecně doporučujeme v prvních několika letech provádět spíše menší počet hodnocení ex-post a dopadů, aby nebyl rodící se systém přetížen, tj. řádově jedno až tři hodnocení ex-post a dopadů ročně.

Najdou se výjimky, na které navrhované postupy nebude možné beze zbytku uplatnit. Hranice mezi hodnocením průběžným, ex-post a dopadů se stírá v případě dlouhotrvajících nebo v praxi trvale běžících programů, což je například podpora standardních projektů GA ČR, která nemá určen termín ukončení. Pokud bychom se striktně drželi definic, hodnocení ex-post a dopadů, které má být provedeno po ukončení programu, by nemuselo proběhnout nikdy. Z tohoto důvodu by měl mít ÚVVI pravomoc v takových případech vyhlásit hodnocení ex-post nebo dopadů i před formálním ukončením programu, anebo možnost vyhlásit kombinované hodnocení s prvky různých typů. Dalším zvláštním případem jsou programy bezpečnostního výzkumu, ve kterých se pracuje s utajovanými skutečnostmi, pro přístup k nimž je nutná bezpečnostní prověrka. Případy programů, které nezapadají do obecných pravidel, bude nutné řešit jednotlivě.

Shrnutí navrhovaných opatření ohledně rolí aktérů v systému hodnocení programů účelové podpory pro VVI:

- ÚVVI se stane ústředním aktérem v systému hodnocení.
- ÚVVI vymezí obecné zásady hodnocení programů, které budou závazné pro ostatní aktéry.
- Poskytovatelé budou mít pravomoc předkládat návrhy programů k ex-ante hodnocení.
- ÚVVI bude mít povinnost zajišťovat ex-ante hodnocení návrhů programů od poskytovatelů a bude rozhodovat o jejich doporučení k financování před předložením do vlády.
- Poskytovatelé budou mít odpovědnost za průběžné hodnocení svých programů.
- ÚVVI bude mít odpovědnost za hodnocení ex-post a dopadů vybraných programů.

- ÚVVI bude určovat režim hodnocení programů a organizovat jeho případné změnové řízení.
- ÚVVI bude dohlížet na kvalitu veškerého hodnocení programů a případně stanovovat odpovídající sankce.
- ÚVVI bude zajišťovat centrální správu mikrodat, tj. shromažďovat veškerá mikrodata z monitorování a primární data, na kterých jsou založena hodnocení průběžná, ex-post a dopadů.

## 10.2 LIDSKÉ ZDROJE

Kvalifikované lidské zdroje pro hodnocení jsou v českých podmínkách velmi omezené. Slabá je zejména průprava v metodách hodnocení, což souvisí s nízkou úrovní sociálních věd. Je třeba si uvědomit, že najmout na trhu práce, popřípadě relativně rychle vyškolit, je možné lidi, kteří budou schopni provádět jednoduché popisné analýzy, ale pokud chceme kvalitní hodnocení na mezinárodní úrovni, které bude produkovat hodnotné znalosti pro řízení systému VVI, mluvíme o kvalifikaci na kvalitativně vyšší úrovni. Kromě přípravy odborníků na samotné hodnocení je nutné mít na zřeteli také nutnost zlepšit povědomí o hodnocení na straně poskytovatelů, konkrétně manažerů programů, a tvůrců politik tak, aby závěry hodnocení byly využity v praxi. Sebelepší hodnocení je zbytečné, pokud je ignorováno lidmi v rozhodovacích pozicích.

K provádění kvalitního hodnocení jsou potřeba lidé, kteří mají načtenou mezinárodní literaturu, znají pokročilé analytické metody a mají s hodnocením praktické zkušenosti. I když domácích expertů s mezinárodními zkušenostmi přibývá, lidí s takovým profilem zatím vskutku není mnoho, a to zejména ve státní správě. Bohužel vybudování odpovídajících odborných kapacit bude trvat řadu let. Z tohoto důvodu je třeba zajištění lidských zdrojů pro hodnocení začít řešit ve třech rovinách:

1. Zavést na některé z vysokých škol stadium VVI problematiky, které bude – podobně jako například [European programme on Society](#), Science and Technology v několika jiných evropských zemích – pravidelně produkovat nové odborníky.
2. Spustit program celoživotního vzdělávání, který bude školit lidi, kteří již v systému hodnocení působí (včetně financování jejich účasti na krátkodobých kurzech pořádaných zahraničními univerzitami, jako je například [Evaluation of Science and Innovation Policies](#) na MIOIR v Manchesteru a nebo R&D Evaluation Course na University of Twente).
3. Ze začátku provést pilotní hodnocení programů s důrazem na získávání praktických dovedností a procesy učení se „za pochodu“ pod vedením externích odborníků. Následně organizovat ex-post hodnocení ve smíšených týmech s cílem podporovat interaktivní učení mezi různými aktéry v systému.

Jako krátkodobé řešení doporučujeme najímání zahraničních hodnotitelů v kombinaci s domácími kapacitami. Dlouhodobě je však třeba podporovat vytvoření domácí komunity hodnotitelů, na niž bude možné fungování systému postavit. Zahraniční hodnotitelé mají řadu výhod například v podobě obecně vyšší nezávislosti na místních podmínkách. Nicméně domácí hodnotitelé mohou mít oproti zahraničním rovněž některé přednosti:

- Rozumí domácímu kontextu, znají historii programů, institucí i politik, nemusí se snažit pokaždé všemu porozumět od nuly.
- Znají specifika domácích mikrodat, s řadou z nich již mají analytické zkušenosti, vyvarují se zbytečných přehmatů plynoucích z rozdílů v datech mezi zeměmi.

- Vzhledem k přetrvávajícím rozdílům v paritě kupní síly jsou podstatně levnější než hodnotitelé najímaní z vyspělých evropských zemí.
- Praktické zkušenosti z provádění hodnocení se budou soustředit u domácích hodnotitelů, a tudíž zůstat v českém systému.
- Z domácí hodnotitelské komunity se mohou rekrutovat odborně zasvěcení zaměstnanci tvůrců politik a poskytovatelů, což může dále pomoci ke zlepšení ve fungování systému jako celku.

Rozvoji domácí komunity hodnotitelů by rovněž pomohlo založení české platformy pro hodnocení VVI podle vzoru FTEVAL v Rakousku, která začala fungovat v roce 1996 a přispěla k výraznému zlepšení tamní hodnoticí kultury. Pod hlavičkou dané platformy by kromě diseminačních aktivit ohledně zkušeností s hodnoticími metodami mohly fungovat i internetové stránky věnované informacím o hodnocení programů, tj. o proběhlých, probíhajících a nadcházejících hodnoceních, a mohly by tam být i volně ke stažení hodnoticí zprávy, tj. mohly by fungovat jako depozitář dokumentů z proběhlých hodnocení.

Shrnutí navrhovaných opatření ohledně lidských zdrojů v systému hodnocení programů účelové podpory pro VVI:

- Zavést na vysoké škole studium VVI problematiky pro vzdělávání nových odborníků.
- Rozvinout program celoživotního vzdělávání pro lidi, kteří již v systému působí.
- Provádět pilotní hodnocení s důrazem na učení se.
- Nespoléhat se pouze na zahraniční hodnotitele, zadávat hodnocení pro kombinované domácí a zahraniční týmy, a tím podpořit vytvoření domácí znalostní báze.
- Založit českou platformu pro hodnocení VVI podle vzoru FTEVAL v Rakousku.

### 10.3 DATOVÉ ZDROJE

IS VaVal je dobrý zdroj mikrodat z monitorování programů, avšak to samo o sobě pro kvalitní hodnocení nestačí. Datová základna pro hodnocení nezanedbatelné části programů (z hlediska objemu jejich rozpočtů) musí zahrnovat vyřešení přístupu ke chráněným mikrodatům z ČSÚ pro účely hodnocení. Dostupnost mikrodat z ČSÚ je však velmi omezená, protože je poskytován jen výsek dat, nikoliv například data z firemních rozvah, výsledovek a data o zaměstnanosti, a to pouze pro výzkumné účely.

Cílem v této oblasti by mělo být zavedení datové infrastruktury, ve které budou relevantní existující databáze ze sekundárních zdrojů na mikroúrovni propojovány s monitorovacími daty a uchovávány v centrální databázi zajišťující jejich vzájemnou kompatibilitu, propojitelnost a tvorbu časových řad. Hodnotitelům bude umožněn opakovaný přístup k této centrální databázi za účelem hodnocení různých programů tak, aby nebylo nutné vytvářet, propojovat, získávat a nakupovat mikrodata zvlášť pro každé hodnocení. Primární šetření rovněž propojitelné s touto databází by se mělo provádět pouze, pokud to bude nezbytně nutné pro zajištění unikátních údajů, které nikde jinde neexistují, kupříkladu pro potřeby identifikačních strategií v ekonometrických odhadech nebo pro potřeby podrobných případových studií.

Při dnešní úrovni informačních a komunikačních technologií je technické zabezpečení v podstatě triviální problém k řešení, avšak je nutné překonat organizační a právní překážky, a to zejména ohledně spolupráce mezi vlastníky dat. Zcela klíčovým předpokladem bude zapojení ČSÚ, který sbírá dlouhé časové řady zásadních dat, například mikrodata z ročního šetření VaV a z inovačních průzkumů, která z jiného zdroje není možné získat. Nejprogresivnějším řešením, které je dnes zaběhnuto již v řadě vyspělých zemí, je vzdálený přístup ke chráněným mikrodatům pod dohledem

statistického úřadu. Příkladem, ze kterého by bylo možné čerpat zkušenosti, je systém Microdata Online Access (MONA) národního statistického úřadu ve Švédsku (SCB 2015).

Zavedení systému MONA by mělo následující výhody:

1. ČSÚ by měl neustálou kontrolu nad mikrodaty a nad tím, co se s nimi děje. Mikrodata budou centrálně uložena na zabezpečeném serveru v ČSÚ, který sama o sobě nikdy neopustí. Ke zveřejnění mimo systém budou uvolňovány pouze analytické výsledky po kontrole, že neobsahují žádné chráněné informace. Mikrodata se již nebudou muset nikomu poskytovat, tj. nahrávat na nosiče, mimo tento systém.
2. Uživatelé budou moci v rámci systému mikrodata sdílet mezi sebou, tj. bude možné mikrodata centrálně propojovat v rámci šetření (vytvářet časové řady) a průřezově (mezi šetřeními), takže se to nebude muset pro každé nové vyhodnocení dělat od začátku, což mimo jiné ušetří osobní náklady na hodnocení.
3. Mikrodata bude možné udržovat neustále aktualizovaná, a to jak s ohledem na přidávání dat z nových šetření, tak s ohledem na jejich případné zpětné revize, tj. v daném okamžiku bude existovat pouze jedna verze této databáze, se kterou budou pracovat všichni uživatelé, nikoliv různé verze dat podle toho, kdy byla v minulosti přesně poskytnuta na nosiči.
4. Ušetří se náklady pro administraci opakovaného poskytování dat pro jednotlivé projekty hodnocení, a to jak na straně ČSÚ, tak na straně hodnotitelů, tj. nebude nutné pro každý projekt podepisovat novou smlouvu, sliby mlčenlivosti apod.; při naběhnutí systému bude sice více práce s vytvořením prvotního souboru uživatelů, kteří však budou následně přibývat v menším množství.
5. Pokud bude přístup do systému placený, může to být nezanedbatelný zdroj příjmů pro další rozvoj poskytování mikrodat.
6. Stejný systém přístupu k datům bude využitelný i pro vědecké a výzkumné účely.

Zavedení systému MONA by mělo následující nevýhody:

1. Náklady na vytvoření a provoz systému, tj. prvotní investice do hardwarového a softwarového řešení, administrace prvního okruhu uživatelů a následně operativní náklady na každodenní provoz.
2. Bude nutné najít nový zdroj financování nad rámec současného rozpočtu ČSÚ. Nicméně návratnost těchto výdajů z hlediska kvality fungování veřejné správy jako celku bude velmi vysoká, a to nejen z hlediska výše zmíněných přímých úspor v nákladech na hodnocení, ale zejména v podobě lepších strategických informací o účincích podpůrných programů.

Zřízení by dále mělo být centrální datové úložiště, kde budou archivována všechna primárně sbíraná mikrodata v projektech hodnocení programů financovaných z veřejných zdrojů, tj. pokud se v rámci hodnocení provede dotazníkové šetření, výsledná data a související dokumentace budou povinně uložena v tomto depozitáři. Je to potřebné provést ze dvou hlavních důvodů, a to zaprvé aby byla v případě pochybností umožněna druhotná analýza těchto dat, zadruhé aby mohla být tato data využita v budoucích hodnoceních.

Zároveň je třeba dále rozvíjet IS VaVal. Bez ohledu na to, v jakém časovém horizontu se podaří zavést systém MONA, by měl být co nejdříve zaveden režim přístupu k mikrodatům z IS VaVal pro výzkumné a hodnotící účely. Mělo by být

přínejmenším umožněno připojit k datům z IS VaVal přes IČO či rodné číslo externí data a výsledek poskytnout v anonymizované podobě. Dále by mělo být systematicky vyřešeno propojení IS VaVal s mikrodaty z bibliometrických databází Web of Science, Scopus a případně dalších oborových citačních databází. Zásadní je zavedení unikátního identifikačního kódu pro konkrétního člověka, přes který by bylo možné databáze propojovat. Jednou z možností by mohla být povinnost pro žadatele o badatelské projekty a jejich spolupracovníky založit si Open Researcher and Contributor ID (ORCID) účet a identifikovat se pomocí ORCID ID v žádosti o projekt. IS VaVal by neměl obsahovat pouze data o podpoře z národních zdrojů, ale rovněž o čerpání významných grantů ze zahraničí, například z programu Horizon 2020 nebo ESF grantů, měly by tam být přidány i údaje za smluvní výzkum, aplikovaný výzkum a komercializaci výzkumu. Důsledněji je také třeba provádět zpětnou kontrolu vyplněných údajů.

Jako standard by měla být do zadávací dokumentace programů zakotvena právně vynutitelná povinnost žadatelů o podporu poskytovat data pro potřeby hodnocení. Při podání žádosti o projekt by mělo být povinné vyplnit jako přílohu dotazník se vstupními daty pro hodnocení daného programu, přičemž uvedení nesprávných údajů by mohlo být důvodem k vyřazení ze soutěže. Následně v průběhu nebo po skončení projektu (programu) by na základě této povinnosti museli všichni žadatelé pod určitou sankcí vyplnit další dotazník, aby získané údaje mohly být použity pro vyhodnocení změny dosažené poskytnutím podpory. Zásadní je, aby se uvedená data sbírala nejenom za příjemce podpory, ale i za žadatele, kteří podporu nedostali, a kteří by tudíž mohli sloužit jako srovnávací skupina. Jinými slovy žadatelé o podporu by se podáním žádosti měli smluvně zavázat k poskytnutí dat potřebných pro vyhodnocení programu, aby návratnost dotazníků byla co nejvyšší.

Podobně by měla být zavedena podmínka, aby firmy žádající o podporu musely mít beze zbytku splněny zákonné zveřejňovací povinnosti, například zakládání rozvah a výsledovek do sbírky listin obchodního rejstříku nebo zodpovězení relevantních statistických šetření ČSÚ. Neplnění této podmínky by mělo být v zadávací dokumentaci programu uvedeno jako důvod pro vyřazení ze soutěže. Navíc by měla být zavedena podmínka plnit pod sankcí tyto zveřejňovací povinnosti i po celou dobu podpory, případně i v určitém období po jejím vypršení.

Poskytovatelé musí mít zavedeny elektronické informační systémy pro monitorování programů, ze kterých budou pro účely hodnocení snadno získatelné údaje o veřejných soutěžích, průběhu podpořených projektů, hodnoticích, dílčích a závěrečných zprávách projektů apod. Hodnotitelé by měli mít do těchto systémů umožněn přístup.

Shrnutí navrhovaných opatření ohledně datové základny v systému hodnocení programů účelové podpory VVI:

- Zavést systém pro vzdálený přístup ke chráněným mikrodatům ČSÚ, vytvořit systém pro přístup ke chráněným mikrodatům z IS VaVal a následně oba systémy propojit, tj. vytvořit centrální databázi mikrodat.
- Vytvořit depozitář primárních dat, tj. všechna primární mikrodata posbíraná v rámci hodnocení programů budou pro budoucí využití uložena na jednom místě.
- Zavést informační povinnost žadatelů o podporu, tj. příjemci i odmítnutí žadatelé se přihlášením do veřejné soutěže o podporu pod hrozbou sankce zavazují poskytovat data potřebná pro vyhodnocení programu.

## 10.4 FINANČNÍ ZDROJE

Rozpočtové zdroje na hodnocení programů nyní nejsou jasně vymezeny, jsou roztroušené mezi poskytovatele a nejsou používány hospodárně, protože kvalita hodnocení je velmi nízká. Pro zdroje určené na hodnocení by měla být vyčleněna samostatná rozpočtová položka, a to jak v rozpočtu pro jednotlivé programy, tak i ve výdajích státního rozpočtu na VVI jako celku. Na financování celého systému hodnocení programů, tedy všech typů hodnocení, doporučujeme vyčlenit zdroje, které odpovídají zhruba 0,5 až 1,0 % z výdajů na programy účelové podpory, což by bylo řádově 70 až 140 mil. Kč ročně. Rozhodnutí ohledně způsobu, jak budou zadavatelé hodnocení tuto podporu přesně nárokovat, čerpat, účtovat apod., tj. zda bude na tento účel vyčleněna institucionální podpora, anebo budou dané prostředky odčerpávány přímo z rozpočtu programů, ponecháváme na implementační fázi.

Závažné problémy je nutné nicméně v českých podmínkách očekávat při externím zadávání hodnocení prostřednictvím veřejných zakázek. Pokud má být hodnocení kvalitní, nemůže být cena dominantním kritériem pro výběr hodnotitele. Nesmí se stávat, že veřejné zakázky na hodnocení budou vyhrávat nabídky s nejnižší cenou od hodnotitelských týmů, jež však následně nebudou schopny odborně dostát požadavkům na kvalitu hodnocení. Bude to nezbytné řešit nastavením podmínek ze strany zadavatele tak, aby veřejné zakázky na hodnocení programů nevyhrávaly nekompetentní subjekty s nízkou cenou. Zadavatel bude muset být schopen vybrat nejvhodnějšího hodnotitele v závislosti na posouzení kvality jím navrženého projektu hodnocení a schopností jeho hodnotitelského týmu. K tomu bude třeba hlubší pochopení problémů s daty, metodických otázek nebo technického provedení, pro jejichž rozklíčování může zadavatel potřebovat externí konzultaci. Nutná bude kombinace kritérií ceny a kvality, přičemž zásadní váhu by měl mít kvalitativní faktor.

Shrnutí navrhovaných opatření ohledně finančních zdrojů v systému hodnocení programů účelové podpory VVI:

- Na systém hodnocení bude řádově vyčleněno 0,5 až 1,0 % z ročních výdajů na programy účelové podpory.

Bude nutné zavést takový způsob zadávání veřejných zakázek pro externí hodnocení programů, aby byla zajištěna zásadní váha faktoru kvality.



## Seznam zkratek

|                 |   |
|-----------------|---|
| <b>ČES</b>      | Česká evaluační společnost  |
| <b>ČSÚ</b>      | Český statistický úřad  |
| <b>EU</b>       | Evropská unie   |
| <b>GAČR</b>     | Grantová agentura České republiky   |
| <b>FTEVAL</b>   | Austrian Platform for Research and Technology Policy Evaluation           |
| <b>IPN</b>      | Individuální projekt národní  |
| <b>IS VaVal</b> | Informační systém výzkumu, experimentální vývoje a inovací                |
| <b>KHP</b>      | Komise pro hodnocení programů   |
| <b>KHV</b>      | Komise pro hodnocení výsledků výzkumných organizací a ukončených programů |
| <b>MIoIR</b>    | Manchester Institute of Innovation Research                               |
| <b>MONA</b>     | Microdata Online Access   |
| <b>MŠMT</b>     | Ministerstvo školství, mládeže a tělovýchovy                              |
| <b>MVVI</b>     | Ministerstvo pro vědu, výzkum a inovace                                   |
| <b>OECD</b>     | Organizace pro hospodářskou spolupráci a rozvoj                           |
| <b>ORCID</b>    | Open Researcher and Contributor ID  |
| <b>PATSTAT</b>  | Worldwide Patent Statistical Database                                     |
| <b>RVVI</b>     | Rada pro výzkum, vývoj a inovace  |
| <b>TAČR</b>     | Technologická agentura České republiky                                    |
| <b>TAFTIE</b>   | European Network of Innovation Agencies                                   |
| <b>ÚVVI</b>     | Úřad pro vědu, výzkum a inovace   |
| <b>VVI</b>      | Výzkum, vývoj a inovace   |



## Seznam literatury

- ARNOLD, E. (2011) International Audit of Research, Development & Innovation in the Czech Republic, Synthesis Report. Technopolis Group, Brighton.
- AEA (2004) Guiding Principles for Evaluators. American Evaluation Association, Washington, DC.
- BEYWL, W. (2003) Selected Comments on the Standards for Evaluation of the German Evaluation Society. German Evaluation Society (DeGEval), <http://www.degeval.de/publikationen/standards-fuer-evaluation/>.
- BOEKHOLT, P., ARNOLD, E., GIARRACCA, F. a PLOEG, M. (2014) Evaluation Reference Model (for TAFTIE's Taskforce Benchmarking Impact, Effectiveness and Efficiency of Innovation Instruments). Technopolis. Amsterdam, <http://www.technopolis-group.com/?report=taftie-s-taskforce-benchmarking-impact-effectiveness-efficiency-innovation-instruments-evaluation-reference-model>.
- CPB (2012) Dare to measure: Evaluation designs for industrial policy in the Netherlands. Netherlands Bureau for Economics Policy Analysis, Impact Evaluation Expert Working Group, <http://www.cpb.nl/en/article/dare-measure-evaluation-designs-industrial-policy-netherlands>.
- ČADIL, V. (2011) Metodický rámec pro hodnocení programů výzkumu, vývoje a inovací. Závěrečná zpráva. Technologické Centrum AV ČR. <http://www.vyzkum.cz/storage/att/18C0C6E0BEDC982432D14A62BD739099/A%203-3%20Hodnoceni%20programu.pdf>.
- ČADIL V. (2014) Metodologické nedostatky stávajícího hodnocení programů výzkumu, vývoje a inovací v ČR. ERGO, r. 9, č. 2, s. 11–16, <http://www.tc.cz/cs/publikace/periodika/seznam-periodik/ergo/ergo-roc-9-c-2-cerven-2014>.
- ČES (2011) Etický kodex evaluátora. Česká evaluační společnost, Praha, [http://czecheval.websnadno.cz/standardy\\_kodex/ces\\_eticky\\_kodex\\_.pdf](http://czecheval.websnadno.cz/standardy_kodex/ces_eticky_kodex_.pdf).
- ČES (2013) Formální standardy provádění evaluací. Česká evaluační společnost, Praha, [http://www.czecheval.cz/standardy\\_kodex/ces\\_formalni\\_standardy\\_evaluaci\\_short\\_5\\_.pdf](http://www.czecheval.cz/standardy_kodex/ces_formalni_standardy_evaluaci_short_5_.pdf).
- ESF (2012) Evaluation in Research and Research Funding Organisations: European Practices. ESF Member Organisation Forum on Evaluation of Publicly Funded Research. <http://www.esf.org/coordinating-research/mo-fora/evaluation-of-publicly-funded-research.html>.
- Evropská komise (2009) Impact Assessment Guidelines. European Commission, Brussels.
- Evropská komise (2013a) Evaluation in the Field of State Aid: Methodology Guidance Paper. European Commission, DG Competition, Brussels.
- Evropská komise (2013b) EVALSED: The resource for the evaluation of Socio-Economic Development. European Commission, Brussels, [http://ec.europa.eu/regional\\_policy/sources/docgener/evaluation/guide/guide\\_evalsed.pdf](http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf).
- Evropská komise (2013c) Public consultation on Commission Guidelines for Evaluation. European Commission, Brussels.

- FTEVAL (2012) Research and Technology Policy Evaluation, Evaluation Standards in Research and Technology Policy. Austrian Platform for Research and Technology Policy Evaluation, Vienna.
- GUTHRIE, S., WAMAE, W., DIEPEVEEN, S., WOODING, S., GRANT, J. (2013) Measuring research, A guide to research evaluation frameworks and tools. RAND Corporation, Pittsburgh.
- MloIR (2010) Inno-Appraisal, Understanding evaluation of Innovation Policy in Europe, Study for European Commission, DG Enterprise, <http://www.innovation-policy.net/compendium>.
- MloIR (2014) Compendium of Evidence on Innovation Policy. MloIR, Manchester, <http://www.innovation-policy.org.uk/compendium>.
- MMR (2014) Metodický pokyn pro evaluace v programovém období 2014–2020. Ministerstvo pro místní rozvoj, Národní orgán pro koordinaci a řízení Dohoda o partnerství, Verze: 3, červenec 2014.
- MPO (2011) Národní inovační strategie České republiky. Ministerstvo průmyslu a obchodu, Praha.
- MŠMT (2014) Hodnocení programů účelové podpory VVI z pohledu IPN projektu MŠMT „Efektivní systém hodnocení a financování výzkumu, vývoje a inovací“ (IPN Metodika). MŠMT, Praha.
- OECD (1991) DAC Principles for Evaluation of Development Assistance. OECD, Paris.
- Research Councils UK (2011) Evaluation: Practical Guidelines, A Guide for Evaluation Public Engagement Activities. Research Councils UK, London.
- RVVI (2009) Národní politika výzkumu, vývoje a inovací České republiky na léta 2009-2015. Rada pro výzkum, vývoj a inovace, Praha.
- RVVI (2012) Metodika hodnocení výsledků výzkumných organizací a hodnocení výsledků ukončených programů (platná pro léta 2010 a 2011 a rok 2012). Rada pro výzkum, vývoj a inovace, Praha.
- RVVI (2013a) Aktualizace Národní politiky výzkumu, vývoje a inovací České republiky na léta 2009 až 2015 s výhledem do roku 2020. Rada pro výzkum, vývoj a inovace, Praha.
- RVVI (2013b) Souhrnné vyhodnocení výsledků programů výzkumu, vývoje a inovací ukončených v roce 2012. Rada pro výzkum, vývoj a inovace, Praha, <http://vyzkum.cz/FrontClanek.aspx?idsekce=706005>.
- RVVI (2013c) Metodika hodnocení výsledků výzkumných organizací a hodnocení výsledků ukončených programů (platná pro léta 2013 až 2015). Rada pro výzkum, vývoj a inovace, Praha, <http://www.vyzkum.cz/FrontClanek.aspx?idsekce=685899>.
- RVVI (2015) Základní principy přípravy a hodnocení programů a skupin grantových projektů výzkumu, vývoje a inovací. Rada pro výzkum, vývoj a inovace, Praha.
- SCB (2015) MONA – Microdata Online Access. Statistics Sweden, Örebro, <http://www.scb.se/Grupp/Produkter/Tjanster/Forskare/Dokument/MONA/Produktblad-Eng.pdf>
- VINNOVA (2008) VINNOVA's Focus on Impact - A Joint Approach for Logic Assessment, Monitoring, Evaluation and Impact Analysis. VINNOVA Analysis VA 2008:01, <http://www.vinnova.se/en/Publications-and-events/Publications/Products/VINNOVAs-Focus-on-Impact/>.

**Návrh obecných zásad hodnocení programů účelové podpory výzkumu, vývoje  
a inovací a potřebných systémových změn**

---

Vydává Ministerstvo školství, mládeže a tělovýchovy, Karmelitská 7, Praha 1  
Individuální projekt národní pro oblast terciárního vzdělávání, výzkumu a vývoje:  
Efektivní systém hodnocení a financování výzkumu, vývoje a inovací (IPN Metodika)

<http://metodika.reformy-msmt.cz>

Praha 2015