

Metodika hodnocení datové otevřenosti zastupitelských sborů

Kamil Gregor



Tato metodika byla vytvořena v rámci projektu Advokacie otevřených dat na základě *Declaration on Parliamentary Openness* podpořeného nadací Open Society Fund.

KohoVolit.eu 2014

CC-BY 4.00

Úvod

Tento dokument popisuje metodiku umožňující komplexní sběr a analýzu informací zachycujících datovou otevřenost zastupitelského sboru a její následnou kvantifikaci pomocí několika indikátorů. V závěru dokumentu je metodika aplikována na Parlamentu České republiky. Tato metodika byla vytvořena primárně pro hodnocení datové otevřenosti krajských zastupitelstev v České republice, nicméně tento dokument i k němu příslušící elektronický výzkumný dotazník byly záměrně navrženy tak, aby je bylo možné co nejnázve aplikovat i na jiné zastupitelské sbory, zejména na národní parlamenty států a na lokální zastupitelské sbory na různých stupních místní samosprávy.

Tato metodika vychází z dokumentu nazvaného Deklarace otevřenosti parlamentů (*Declaration on Parliamentary Openness*)¹. Jedná se mezinárodní deklaraci definující obecné standardy sběru a publikování informací, které by měly zastupitelské sbory splňovat, aby je bylo možné považovat za transparentní. Deklarace byla vytvořena v rámci mezinárodní komunity parlamentních monitorovacích organizací, tedy organizací, které se dlouhodobě zaměřují na sledování činnosti v národních parlamentech a dalších zastupitelských sborech v různých zemích. Hlavním garantem Deklarace je americká organizace National Democratic Institute². Deklarace byla vytvořena v roce 2012 a k červnu 2014 se k ní přihlásilo více než 100 parlamentních monitorovacích organizací a několik desítek parlamentních komor z řady zemí světa.

Tato metodika odkazuje pouze na část celé Deklarace, která se bezprostředně týká problematiky takzvaných otevřených dat (*open data*), konkrétně otevřených parlamentních dat (*open parliamentary data*). Otevřená data lze definovat jako informace publikované takovým způsobem, že jejich opětovnému použití nebrání žádné nepřiměřené technické ani právní překážky. Prakticky může být podmínka otevřenosti konkrétního datového souboru naplněna v různé míře, záleží zejména na technickém způsobu publikace a použité licenci.

¹ <http://www.openingparliament.org/declaration>

² <https://www.ndi.org/>

Proto je možné míru otevřenosti vyjádřit číselným indikátorem. Čím je tento indikátor vyšší, tím více se konkrétní datový soubor blíží ideálnímu stavu.

Indikátory datové otevřenosti definované v této metodice vychází zejména z indikátorů *Open Data Index*³, které vyvíjí britská organizace Open Knowledge Foundation⁴. Tato metodika se zabývá pouze parlamentními daty, tedy daty procedurálního charakteru popisujícími činnost zastupitelského sboru a jeho členy. Nepokrývá tedy samotný obsah činnosti konkrétního sboru, v tomto případě zastupitelstva kraje, ani další složky krajské samosprávy (například transparentnost veřejných zakázek a tak podobně).

Protože je tato metodika vyvinuta primárně pro hodnocení datové otevřenosti krajských zastupitelstev a protože všechny kraje v České republice sdílí stejnou úpravu autorských práv vztahujících se na jimi publikovaná data, nevěnuje se tato metodika případným právním překážkám otevřenosti dat. Je zaměřena výhradně na technické překážky. Dále se metodika omezuje pouze na proaktivně publikovaná data, a to na data publikovaná v elektronické podobě na oficiálních webových stránkách jednotlivých krajských zastupitelstev. Nezabývá se tedy problematikou přístupu k informacím na vyžádání ani problematikou publikování informací v papírové podobě.

³ <https://index.okfn.org/>

⁴ <https://okfn.org/>

Dosavadní metodiky

Tato metodika byla inspirována několika již existujícími metodikami hodnotícími otevřenost parlamentů. Všechny tyto metodiky se na rozdíl o této věnovaly primárně jiným tématům, než jsou otevřená parlamentní data. Pokud obsahovaly otázky týkající se parlamentních dat, omezovaly se pouze na zachycení toho, zda příslušný národní parlament nebo jiný zastupitelský sbor příslušná data zveřejňuje či nikoli, nezabíhaly tedy do další podrobností o způsobu jejich publikování.

Zatím zřejmě nejrozvinutější metodikou hodnocení otevřenosti zastupitelských sborů je metodika vytvořená asociací latinskoamerických parlamentních monitorovacích organizací Latin American Network for Legislative Transparency⁵. Tato metodika byla aplikována na pět národních parlamentů latinskoamerických zemí (Argentina, Chile, Kolumbie, Mexiko a Peru) v roce 2011. Další kolo sběru dat je naplánováno na rok 2014. Metodika se zaměřuje především na oblast transparentnosti, boje proti korupci a občanské participace na legislativním procesu a obsahuje například otázky týkající se majetkových příznání zákonodárců, rozpočtu parlamentu nebo zákonodárné iniciativy občanů.

Druhým příkladem je metodiky vyvinutá srbskou parlamentní monitorovací organizací Centre for Research, Accountability and Transparency.⁶ Ta sleduje dostupnost několika vybraných druhů dat, jako jsou zápisy z plenárních zasedání, výsledky hlasování nebo texty navrhované legislativy v několika parlamentech jihovýchodní Evropy (Turecko, Srbsko, Chorvatsko, Slovinsko, Bosna a Hercegovina, Albánie, Černá Hora, Makedoni, Rumunsko, Bulharsko). Zajímavá je skutečnost, že tato metodika byla aplikována hned dvakrát v poměrně krátké době, takže umožňuje zachytit změny v míře dostupnosti informací mezi oběma časovými body.

⁵ <http://www.transparencialegislativa.org/indice/indice-2011/>

⁶ <http://www.otvoreniparlament.rs/statistika-i-zanimljivosti/report-on-parliamentary-openness-in-the-region-of-western-balkans-and-turkey/>

Další metodiku vytvořila americká organizace National Democratic Institute pro zachycení dostupnosti různých druhů parlamentních informací v zemích východní Evropy (Albánie, Bosna a Hercegovina, Česká republika, Makedonie, Černá Hora, Kosovo, Polsko, Srbsko, Slovensko).⁷ Tato metodika je poutavá skutečností, že výzkumný dotazník nebyl vyplňován experty z parlamentních monitorovacích organizací, jak je tomu zvykem u jiných metodik, ale samotnými zaměstnanci parlamentu. Umožňuje tedy srovnat jejich odpovědi s názory expertů z nevládních organizací.



⁷ <http://blog.openingparliament.org/post/80065479246/survey-on-openness-in-central-and-eastern-europe>

Struktura výzkumného dotazníku

Základem této metodiky je výzkumný dotazník, který je vyplňován expertem na daný zastupitelský sbor. Může jím být buď pracovník tohoto sboru, jiný zaměstnanec veřejného sektoru, člen parlamentní monitorovací organizace nebo jiné organizace občanské společnosti, akademický pracovník nebo jiný jednatel. Odpovědi na otázky v dotazníku jsou posléze analyzovány a na jejich základě je stanovena hodnota jednotlivých indikátorů datové otevřenosti.

Tato kapitola popisuje a komentuje jednotlivé otázky v dotazníku.

Obecná struktura okruhů otázek

Výzkumný dotazník obsahuje několik okruhů otázek. Okruhem otázek se myslí skupina otázek, které jsou tematicky příbuzné, zpravidla protože se týkají jednoho druhu parlamentních dat.

Dotazník je rozdělen na tři segmenty. První segment obsahuje jeden okruh obecných uvozovacích otázek, konkrétně otázku na jméno a příjmení experta a název organizace, kterou zastupuje. Kontaktní údaje jsou pokryty pouze otázkou na e-mailovou adresu experta, nicméně v případě potřeby lze doplnit i další otázky (například otázku na telefonní číslo). Následuje otázky zjišťující, jaký sektor expert zastupuje. To je užitečné pro případné databázování odpovědí. Další otázka zjišťuje, kterého krajského zastupitelstva se odpovědi týkají, a dotazuje se na odkaz na oficiální webovou stránku kraje, aby bylo možné rychle dohledat zdroj odpovědí. První segment je ukončen textovým polem pro případné otázky.

Druhý segment se týká plenárních schůzí zastupitelstva a postupně obsahuje okruhy otázek o jednotlivých druzích parlamentních dat. Třetí a poslední segment obsahuje tytéž otázky, ale tentokrát se týká schůzí jednotlivých výborů. Druhý a třetí segment je vždy zakončen

samostatně stojícím textovým polem umožňujícím expertovi poskytnou dodatečné komentáře k danému segmentu.

Nutno podotknout, že segment týkající se výborů záměrně umožňuje poskytnout u každé otázky pouze jedinou odpověď týkající se všech výborů daného krajského zastupitelstva. Dotazník je v případě potřeby možné snadno rozšířit o další identické segmenty týkající se jednotlivých výborů. Ve většině případů to ale není nutné, protože datová otevřenost se mezi výbory jednoho zastupitelského sboru zpravidla nijak moc neliší. Pokud existují rozdíly, expert by měl vyplnit odpověď popisující převládající stav u většiny výborů a případné výjimky popsat v polích pro komentáře.

Dotazník neobsahuje segment týkající se komisí nebo jiných orgánů krajského zastupitelstva, protože ty mají na rozdíl od výborů zpravidla pouze poradní roli. Příslušný segment je nicméně možné snadno doplnit zkopírováním druhého nebo třetího segmentu.

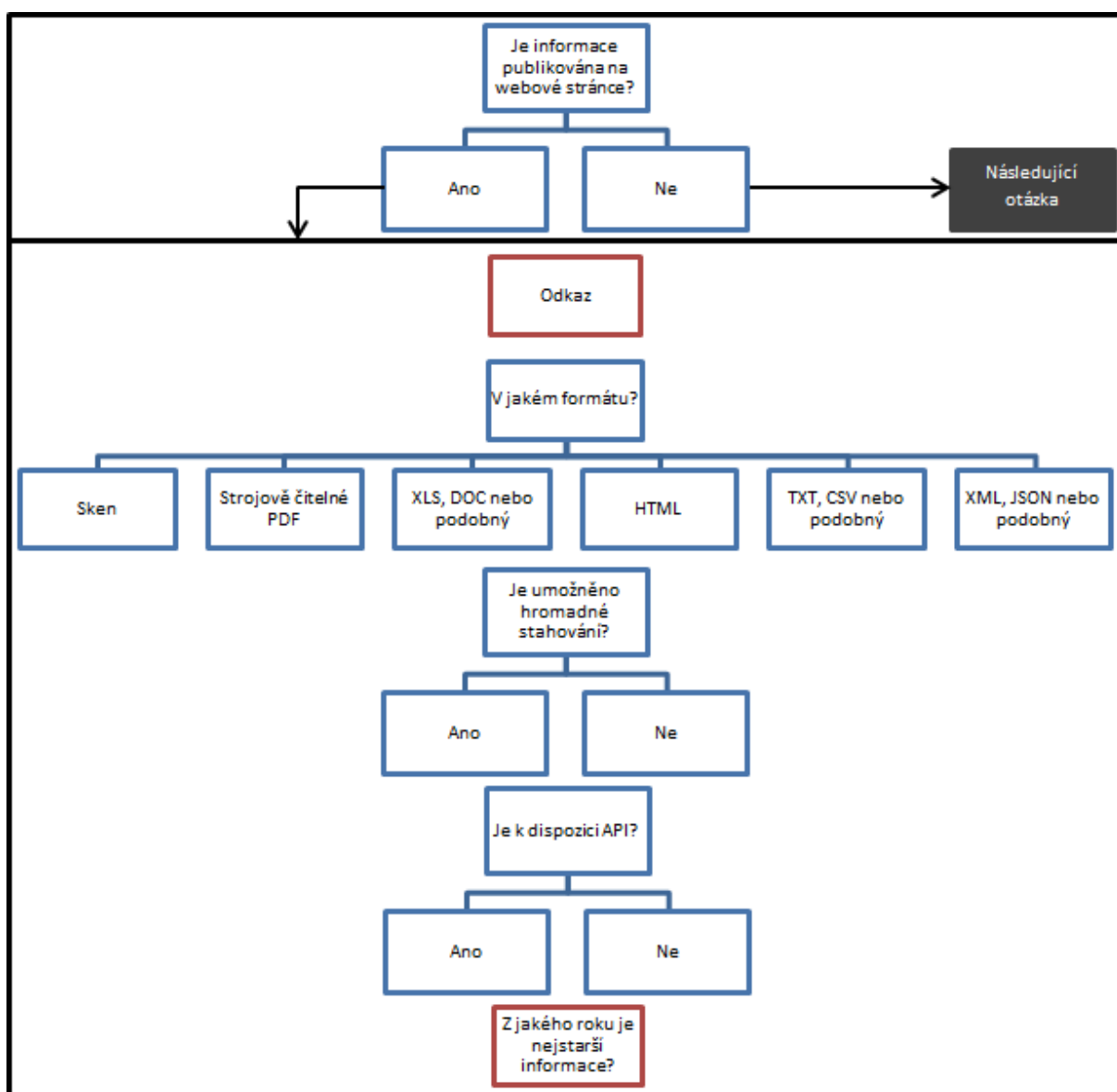
Každý okruh otázek v rámci druhého a třetího segmentu dotazníku se týká jednoho druhu parlamentních dat. Konkrétně dotazník pokrývá tyto druhy dat:

- Zápisy ze schůzí
- Zvukové a obrazové záznamy ze schůzí
- Výsledky hlasování
- Podklady ke schůzi
- Usnesení ze schůzí
- Informace o členech orgánu

„Členy orgánu“ se ve druhém segmentu týkajícím se plenárních chůzí myslí všichni členové krajského zastupitelstva. Ve třetím segmentu se jimi myslí členové výborů. Dotazník nepokrývá celou řadu jiných druhů parlamentních dat, zejména dat týkajících se návrhů zákonů a jejich průběhu legislativním procesem. Tyto druhy dat jsou opomenuty záměrně, a to protože je tato metodika vytvořena primárně pro hodnocení datové otevřenosti krajských zastupitelstev, která v České republice přijímají pouze podzákoné normy a nemají definovaný legislativní proces odpovídající legislativnímu procesu v národním parlamentu.

Je třeba ovšem podotknout, že dostupnost dat o legislativním procesu lze v případě potřeby reflektovat do určité míry i pomocí této metodiky, a to zejména pomocí okruhu otázek týkajících se podkladů ke schůzím a usnesení ze schůzí. Podklady ze schůzí mohou totiž zahrnovat i texty návrhů zákonů a pozměňovacích návrhů a usnesení ze schůzí mohou zahrnovat i texty schválené legislativy.

Okruhy otázek týkající se jednotlivých druhů parlamentních dat sdílí jednotnou strukturu, kterou zobrazuje obrázek dole:



Okruh je uvozen otázkou na dostupnost dat. Odpovědi mohou být omezeny na „Ano“ či „Ne“, nebo mohou být rozděleny do více možností. Ve druhém případě se možné odpovědi

mohou, nebo nemusí vzájemně vylučovat. Podrobnosti o každém okruhu otázek popisující následující podkapitoly.

Pokud určité krajské zastupitelstvo daný druh dat nepublikuje na své oficiální webové stránce, vede odpověď „Ne“ k dalšímu okruhu otázek. Odpověď „Ano“ vede k sadě otázek zachycující datovou otevřenost. Nejprve expert poskytuje odkaz na příklad daného druhu dat, aby byla možná snadná kontrola odpovědí. Poté je dotázán na formát, v němž je daný druh dat k dispozici. Nutno podotknout, že jednotlivé možnosti se vzájemně nevylučují.

První možnost („Sken“) se týká strojově nečitelných dat v jakémkoli formátu, nejčastěji skenů dokumentů. Třetí možnost („XLS, DOX nebo podobný“) se týká dat publikovaných v oddělených souborech. Dotazník přitom pro jednoduchost nerozlišuje proprietární a neproprietární formáty. Možnost „HTML“ popisuje situaci, kdy jsou daná data publikována přímo v textu webové stránky. Možnost „XML, JSON nebo podobný“ je vhodná pro formáty vysoce strukturovaných dat nebo o formáty obohacené o metadata. Pokud jsou data v jiném formátu, než obsahuje výčet, měl by expert vždy uvést přesný formát v komentáři.

Další dvě otázky se týkající existence možnosti hromadného stahování daného druhu dat za delší časové období a existence API. V případě kladné odpovědi je vhodné doplnit popis těchto možností do komentáře, zejména poskytnout odkaz na tyto služby. Poslední otázka zachycuje bohatství historických dat. Každý okruh je ukončen textovým polem pro případné komentáře.

Okruh otázek k zápisům ze schůzí

První okruh otázek se týká zápisů ze schůzí, a to buď plenárních schůzí zastupitelstva kraje, nebo schůzí jednotlivých výborů. Tento okruh otázek umožňuje expertovi označit, jaký druh zápisů je pořizován:

- Zápis ve formě shrnutí projednávaných bodů – Obsahuje pouze výčet jednotlivých bodů, není patrné, kdo na schůzi vystoupil a co říkal

- Zápis ve formě shrnutí rozpravy o jednotlivých bodech – Shrnuje více či méně vyčerpávajícím způsobem rozpravu, je uvedeno, kdo vystoupil a popřípadě stručně shrnut obsah jeho rozpravy
- Zápis ve formě doslovných přepisů rozpravy – Je přesným přepisem rozpravy

Pokud se praxe zápisů liší, měl by expert označit převládající praxi zápisů a případné rozdíly popsat pomocí komentáře.

Okruh otázek ke zvukovým a obrazovým záznamům ze schůzí

Druhý okruh otázek se týká toho, zda jsou na oficiální webové stránce krajského zastupitelstva publikovány záznamy schůzí zastupitelstva nebo výborů. Nutno zdůraznit, že tento okruh otázek popisuje pouze situaci, kdy je možné na webové stránce dohledat záznamy již uplynulých schůzí. Jestliže zastupitelstvo zprostředkovává zvukový či obrazový přenos pouze v době schůze a následně nearchivuje záznam, což je běžná praxe v řadě zastupitelských sborů, má se za to, že zastupitelstvo záznamy nepublikuje vůbec.

Tento okruh otázek umožňuje expertovi označit, zda krajské zastupitelstvo publikuje pouze zvukové, pouze obrazové, nebo zvukové i obrazové záznamy.

Okruh otázek k výsledkům hlasování na schůzích

Třetí okruh otázek se týká publikování výsledků hlasování. Okruh otázek opět umožňuje rozlišit mezi několika druhy publikovaných informací:

- Součty hlasů v celém orgánu – Jsou publikovány pouze počty zastupitelů, kteří hlasovali pro, proti, zdrželi se, a popřípadě nebyli přítomní. Nelze rozlišit, jak hlasovali jednotliví zastupitelé, ani jak hlasovaly jednotlivé strany.

- Součty hlasů jednotlivých stran – Jsou publikovány počty zastupitelů, kteří hlasovali pro, proti, zdrželi se, a popřípadě nebyli přítomní, a to pro jednotlivé strany zvlášť. Nelze rozlišit, jak hlasovali jednotliví zastupitelé.
- Výsledky hlasování po jménech jednotlivých zastupitelů.

V řadě zastupitelských sborů je běžné, že se různá hlasování zaznamenávají různým způsobem. Pokud tomu tak je, měl by expert označit převládající možnost a pokud možno upřesnit okolnosti těchto rozdílů v textovém poli pro komentáře.

Okruh otázek k podkladům ke schůzím

Čtvrtý okruh otázek se týká široce pojatého pojmu „podklady ke schůzím“. V dotazníku jsou výslovně rozlišeny informace o místě a čase konání schůze, o programu schůze (to platí zejména pro situaci, kdy je zveřejněn seznam jednotlivých bodů programu) a další související dokumenty. Tato kategorie může zahrnovat například návrhy projednávaných norem, zprávy, existující právní předpisy, smlouvy a další dokumenty finanční povahy a tak dále. Je vhodné, aby expert při označení této možnosti upřesnil povahu obvykle příkládaných dokumentů v textovém poli pro komentáře.

Okruh otázek k usnesením ze schůzí

Pátý okruh se týká textů závazných rozhodnutí, která jsou na schůzích přijata. Tato usnesení mohou mít povahu právních norem nebo jiných rozhodnutí. Mohou být publikována buď odděleně, nebo v rámci zápisu ze schůze. Pokud jsou různé druhy usnesení publikovány různým způsobem, měl by expert označit převládající způsob a případné rozdíly upřesnit v textovém poli pro komentáře.

Okruh otázek k informacím o členech

Šestý a poslední okruh otázek se týká informací o členech zastupitelstva a jeho orgánů. Z velkého množství možných druhů dat jsou ve výzkumném dotazníku zahrnuty pouze následující druhy:

- Jméno a příjmení zastupitele
- Stranická příslušnost zastupitele
- Členství zastupitele v orgánech kraje – To se týká zejména členství ve výborech a komisích
- Členství zastupitele v orgánech mimo kraj – To se může týkat členství v dozorčích a správních radách firem s majetkovou účastí veřejné správy, spojení se soukromými subjekty a tak dále.
- Další životopisné informace – Tato možnost je nejčastěji relevantní v případě, kdy webová stránka zastupitelského sboru obsahuje životopisy jednotlivých členů, a tedy i například kontaktní údaje, informace o vzdělání nebo o pracovních zkušenostech.



Technické řešení dotazníku

Výzkumný dotazník byl záměrně vytvořen tak, aby ho bylo možné replikovat pomocí nástroje Google Form. Tento nástroj umožňuje zdarma vytvářet a používat elektronické dotazníky. Odpovědi jsou automaticky zapisovány společně s údajem o času zaslání odpovědi do tabulky Google Spreadsheet. Odpovědi je následně možné prohlížet, stanovat či publikovat na Internetu. Nástroj Google Form umožňuje omezit počet odpovědí jednoho uživatele, prohlížet a editovat jednou odeslané odpovědi a provádět jednoduché analytické operace (zejména přehledy četnosti jednotlivých odpovědí).

Vzorová verze dotazníku je k dispozici v českém jazyce zde: <http://bit.ly/1n3DvCz> a v anglickém jazyce zde: <http://bit.ly/1y8olou>.



Indikátory datové otevřenosti

Uvažujeme-li o určitém druhu dat, který nějaký zastupitelský sbor zveřejňuje, uvědomíme si, že způsob publikování popisuje celá řada vzájemně nezávislých vlastností. Především jde o následující koncepty:

- **Kvalita** – Tento koncept postihuje, jak bohatá a vyčerpávající data jsou publikována, neboli k jak velké ztrátě informací dochází od okamžiku vzniku informace do okamžiku jejího publikování.
- **Použitelnost** – Tento koncept postihuje, jak snadno je daný druh dat znovu použitelný. Prakticky se týká především formátu publikovaných souborů, možnosti data hromadně stáhnout a tak dále
- **Úplnost** – Tento koncept postihuje, jak jsou publikovaná data úplná. To se týká nejen otázky, zda například krajské zastupitelstvo publikuje zápisy všech schůzí nebo jen některých, ale i otázky, kolik let činnosti zastupitelstva publikovaná data pokrývají.
- **Včasnost** – Tento koncept postihuje, s jakým časovým předstihem krajské zastupitelstvo publikuje informaci o budoucí schůzi před samotnou dobou trvání schůze. Je zřejmé, že tento koncept se netýká druhů dat, které vznikají přímo na schůzích (zápisy, záznamy, výsledky hlasování, usnesení), ani druhů dat zcela nezávislých na schůzích (informace o členech).
- **Prodleva** – Tento koncept postihuje, s jakým časovým zpožděním jsou dat publikována na oficiální webové stránce krajského zastupitelstva, měřeno od okamžiku jejich vzniku. Tento koncept tedy obsahuje i informaci o tom, jak jsou data publikovaná v dané chvíli na webové stránce krajského zastupitelstva aktuální.

Tato metodika pro jednoduchost opomíjí koncepty včasnosti a prodlevy. Každý ze tří zbývajících konceptů kvality, použitelnosti a úplnosti obsahuje prvek intenzity – konkrétní datový soubor může být nekvalitní, poměrně kvalitní, velmi kvalitní a tak dále. Podobně je to i s použitelností a úplností. Je tedy zřejmé, že tyto abstraktní koncepty je možné v případě

konkrétních zastupitelských sborů převést na číselné hodnoty a tím pádem mezi sebou srovnávat míru, do jaké se sledované sbory přibližují ideálnímu stavu, kdy jsou všechny druhy dat úplně kvalitní, použitelné a úplné.

Kvalita, použitelnost a úplnost jsou měřeny zvlášť pro každý druh dat. Kromě toho jsou ještě měřeny zvlášť pro plenární schůze a pro schůze výborů. Protože tato metodika sleduje šest druhů dat, vzniká při její aplikaci na jeden zastupitelský sbor 12 hodnot jednoho indikátoru. To je celkem 36 indikátorů popisujících jeden sbor, což je pro přehledné vyjádření a především pro vzájemné srovnávání více sborů neprakticky vysoké číslo.

Proto je nutné indikátory agregovat a vytvořit jediný ukazatel datové otevřenosti zastupitelského sboru. Nejprve je vypočten ukazatel datové otevřenosti každého druhu dat na plenárních schůzích, a to tak, že se vynásobí hodnoty indikátoru kvality, použitelnosti a úplnosti. Poté se ukazatelé datové otevřenosti všech druhů dat sečtou a vydělí číselným ukazatelem odpovídajícím dokonale datově otevřenému zastupitelskému sboru. Následně jsou stejné operace provedeny pro schůze výborů. Přesně lze tento postup vyjádřit následujícím vzorcem:

$$O = \frac{\sum_{i=1}^d \prod_{i=1}^k I_k}{\sum_{i=1}^d \prod_{i=1}^k M_k}, O \in \langle 0; 1 \rangle$$

Proměnná O udává ukazatel otevřenosti plenárních schůzí nebo schůzí výborů. Proměnná d udává počet druhů dat (v našem případě šest), proměnná k udává počet konceptů (v našem případě tři), proměnná I_k udává hodnotu indikátoru jednoho konceptu a proměnná M_k udává maximální možnou hodnotu indikátoru tohoto konceptu v dokonale otevřeném zastupitelském sboru.

Tím získáme dvě hodnoty datové otevřenosti pro zastupitelský sbor – jedna hodnota zachycuje otevřenost plenárních schůzí a druhá zachycuje otevřenost schůzí výborů. Alternativně je možné ukazatele datové otevřenosti plenárních schůzí a schůzí výborů sečíst a získat tak jednu hodnotu pro celý zastupitelský sbor. Přesně lze tento postup vyjádřit následujícím vzorcem:

$$O_s = \frac{O_p + O_v}{2}, O_s \in \langle 0; 1 \rangle$$

Proměnná O_s udává hodnotu datové otevřenosti daného zastupitelského sboru, proměnná O_p udává hodnotu datové otevřenosti plenárních schůzí a proměnná O_v udává hodnotu datové otevřenosti schůzí výborů. Z výpočtu mimo jiné vyplývá, že pokud nějaký zastupitelský sbor nepublikuje nějaký druh dat vůbec, je hodnota ukazatele datové otevřenosti pro daný druh dat automaticky nulová.

Výše uvedené vzorce dávají všem druhům dat stejnou váhu. Stejně tak dávají stejnou váhu plenárním schůzím a schůzím výborů. Expert analyzující výsledky výzkumného dotazníku si ovšem může přát udělit větší váhu určitému druhu dat, protože je pro činnost hodnoceného zastupitelského sboru z nějakého důvodu důležitější. Rovněž si může přát udělit větší váhu plenárním schůzím nebo naopak schůzím výborů podle toho, jak jsou tyto instituce v rámci zastupitelského sboru reálně významné.

Váhu lze přidělit pomocí jednoduché proměnné w_d přidělené každému druhu dat, která nabývá hodnoty od 0 do 1. Její velikost je arbitrární a záleží na rozhodnutí experta. Čím vyšší velikost, tím větší je relativní důležitost daného druhu dat. Hodnota 0 pochopitelně znamená, že se daný druh dat do výpočtu vůbec nezahrne. Výpočet váženého indikátoru popisuje následující vzorec:

$$O = \frac{\sum_{i=1}^d w_d \times \prod_{i=1}^k I_k}{\sum_{i=1}^d \prod_{i=1}^k M_k}, O \in \langle 0; 1 \rangle$$

Plenární schůze a schůze výborů lze navázat pomocí proměnných w_p a w_v , které udává relativní důležitost každého ze dvou druhů orgánů (pléna a výborů). Výpočet váženého indikátoru popisuje následující vzorec:

$$O_s = \frac{w_p O_p + w_v O_v}{w_p + w_v}, O_s \in \langle 0; 1 \rangle$$

Následující podkapitoly popisují způsob kvantifikace jednotlivých konceptů.

Kvalita

Jak již bylo řečeno, koncept kvality odkazuje na bohatství dat. Pokud například zastupitelstvo publikuje pouze výsledky hlasování po jednotlivých stranách, není tato informace tak kvalitní, jako kdyby dané zastupitelstvo publikovalo výsledky hlasování po jménech. V prvním případě totiž dochází při publikování dat k velké ztrátě informace (není možné zjistit, jak hlasovali jednotliví zastupitelé), což ve výsledku snižuje hodnotu dat pro uživatele.

Kvantifikace konceptu kvality se liší u jednotlivých druhů dat. Hodnota indikátoru kvality může u každého druhu dat dosahovat nejvýše 10 bodů. Konkrétní počet bodů závisí na stupni kvality daného druhu dat. Počty bodů pro jednotlivé stupně kvality ukazuje tabulka níže.

Druh dat	Stupeň kvality	Počet bodů
Zápisy ze schůzí	Shrnutí projednávaných bodů	1
	Shrnutí rozpravy o jednotlivých bodech	5
	Doslovné přepisy rozpravy	10
Zvukové a obrazové záznamy	Zvukové záznamy	10
	Obrazové záznamy	10
	Zvukové i obrazové záznamy	10
Výsledky hlasování	Součty za celý orgán	1
	Součty za jednotlivé strany	3
	Výsledky po jménech	10
Podklady ke schůzím	Místo	1
	Datum a čas	1
	Agenda	3
	Další dokumenty	5
Usnesení ze schůzí Informace o členech		10
	Jméno a příjmení	1
	Stranická příslušnost	1
	Členství v orgánech kraje	1
	Členství v orgánech mimo kraj	2
	Další životopisné informace	5

V případě zápisů, zvukových a obrazových záznamů a výsledků hlasování ze schůzí se jednotlivé stupně kvality vzájemně vylučují. V případě podkladů ke schůzím a informací o členech se jednotlivé možnosti nevylučují. Pokud zastupitelský sbor publikuje data popsaná

ve více kategoriích, počty bodů se sčítají. U usnesení ze schůzí žádné stupně kvality nejsou – zastupitelský sbor je buď publikuje, nebo ne.

Nutno podotknout, že tato tabulka je pouze doporučující a bodové hodnocení je ve výsledku vždy arbitrární. Expert analyzující výsledky výzkumného dotazníku může podle svého uvážení zvýšit nebo snížit bodovou dotaci, zejména pokud se výrazně liší dostupnost a způsob publikování dat v rámci jednotlivých druhů dat (například pokud jsou zaznamenávána pouze některá hlasování, výrazně se liší kvalita zápisů z různých schůzí a tak podobně). Přitom je třeba vždy zachovat pravidlo, že nejvyšší možný součet bodů je 10.

Použitelnost

Koncept použitelnosti odpovídá na otázku, jak snadné je data publikovaná na oficiální webové stránce zastupitelského sboru znovu použít. Jak již bylo řečeno, tato metodika se nezabývá právními překážkami používání parlamentních dat a zaměřuje se pouze na technické překážky.

Mechanismus publikování	Stupeň použitelnosti	Počet bodů
Formát	Sken	1
	Strojově čitelné PDF	2
	XLS, DOC nebo podobný	3
	HTML	4
	TXT, CSV nebo podobný	5
	XML, JSON nebo podobný	6
Možnost hromadného stahování	Součty za celý orgán	+1
API	Součty za jednotlivé strany	+1

Tabulka výše obsahuje počty bodů jednotlivých stupňů použitelnosti. Důležitý rozdíl oproti konceptu kvality spočívá v tom, že jednotlivé datové formáty se vzájemně nevylučují, ale při výpočtu indikátoru se vždy používá pouze počet bodů nejvíce bodově ohodnoceného formátu. Pokud tedy například zastupitelský sbor publikuje výsledky hlasování ve formátu HTML a zároveň CSV, použije se hodnota 5 bodů a nikoli (4 + 5) bodů. Jestliže zastupitelský sbor poskytuje možnost hromadného stažení dat, počet bodů se zvětší o jedna. Pokud zastupitelský sbor poskytuje API, počet bodů se rovněž zvětší o jedna. Je-li k dispozici jak

možnost hromadného stažení dat, tak i API, počet bodů se zvětší o dva. Pokud zastupitelský sbor možnost hromadného stažení dat a API neposkytuje, počet bodů za formát se nemění.

Úplnost

Koncept úplnost je oproti předchozím konceptům výrazně jednodušší. Pro jednoduchost předpokládáme, že úplnost publikování dat v rámci jednoho časového období je konstantní a jednotlivé zastupitelské sbory se liší pouze tím, kolik let jejich existence pokrývají data publikovaná na jejich oficiálních webových stránkách. Za každý jeden rok existence zastupitelského sboru, pro který jsou publikována data, je hodnota indikátoru úplnosti zvýšena o 1 bod.

Protože různé zastupitelské sbory existují různě dlouhou dobu, je počet bodů třeba ještě vydělit dobou existence daného zastupitelského sboru. Indikátor úplnosti tedy dosahuje hodnoty od 0 do 1, přičemž hodnota 1 odpovídá situaci, kdy jsou publikována data za celou dobu existence daného sboru.

Stejně jako v případě kvality je i toto pravidlo pouze doporučující. Pokud se výrazně liší bohatství dat publikovaných v různých funkčních obdobích, může expert analyzující výsledky výzkumného dotazníku ohodnotit jednotlivá období různými počty bodů.



Příklad: Parlament ČR

V této kapitole ilustrujeme metodiku hodnocení datové otevřenosti na příkladu dvou komor Parlamentu České republiky – Poslanecké sněmovny a Senátu. Při hodnocení otevřenosti byly v souladu s výše popsanými pokyny texty návrhů zákonů a pozměňovacích návrhů považovány za součást podkladů pro schůze a texty schválené legislativy za druh usnesení.

Poslanecká sněmovna

Jak je patrné z tabulky níže, otevřenost Poslanecké sněmovny je poměrně malá. To je dáno zejména tím, že neposkytuje API a hromadně je možné stáhnout pouze výsledky hlasování a informace o poslancích. Je patrné, že otevřenost plenárních schůzí je vyšší než otevřenost schůzí výborů.

Druh dat	Kvalita	Použitelnost	Úplnost	Otevřenost
Plénium				44 %
Zápisy ze schůzí	10	4	1	50 %
Záznamy schůzí	10	1	1	13 %
Výsledky hlasování	10	6	1	75 %
Poklady ke schůzím	10	4	1	50 %
Usnesení ze schůzí	10	3	1	38 %
Informace o členech	5	6	1	38 %
Výbory				25 %
Zápisy ze schůzí	5	3	1	19 %
Záznamy schůzí	0	0	0	0 %
Výsledky hlasování	1	3	1	4 %
Poklady ke schůzím	10	4	1	50 %
Usnesení ze schůzí	10	3	1	38 %
Informace o členech	5	6	1	38 %
Poslanecká sněmovna				34 %

Pokud jde o plenární schůze, kvalita dat je poměrně vysoká až na informace o členech, kde chybí doplňující životopisné údaje. S použitelností dat je to horší – řada druhů dat je publikována v samostatných textových souborech nebo pouze v textu webové stránky. Nižší otevřenost výborů je dána malou kvalitou výsledků hlasování a zápisů ze schůzí a neexistencí

zvukových nebo obrazových záznamů. Obecně horší je i použitelnost dat – daleko více druhů dat je publikováno v samostatných textových souborech, aniž by bylo zároveň publikováno v textu webové stránky. V obou případech je ale velmi dobrá úplnost dat – všude, kde data existují, pokrývají celé období existence Poslanecké sněmovny.

Senát

Tabulka dole ukazuje hodnoty indikátorů pro Senát. I otevřenost této komory je malá. Senát neposkytuje API ani hromadné stahování dat. I v případě Senátu je větší otevřenost plenárních schůzí oproti schůzím výborů. Kvalita dat a jejich úplnost je u všech druhů dat stejná jako v Poslanecké sněmovně. Použitelnost dat je ale horší.

Druh dat	Kvalita	Použitelnost	Úplnost	Otevřenost
Plénum				38 %
Zápisy ze schůzí	10	4	1	50 %
Záznamy schůzí	10	1	1	13 %
Výsledky hlasování	10	4	1	50 %
Poklady ke schůzím	10	4	1	50 %
Usnesení ze schůzí	10	3	1	38 %
Informace o členech	5	4	1	25 %
Výbory				23 %
Zápisy ze schůzí	5	3	1	19 %
Záznamy schůzí	0	0	0	0 %
Výsledky hlasování	1	3	1	4 %
Poklady ke schůzím	10	4	1	50 %
Usnesení ze schůzí	10	3	1	38 %
Informace o členech	5	4	1	25 %
Senát				30 %

Srovnání

Prezentace výsledků hodnocení prostřednictvím tabulek podobných těm, jaké jsou obsaženy v předchozích dvou podkapitolách, není příliš názorná. Mnohem lepší je jejich vizualizace pomocí sloupcových grafů, které dávají snadno představu o relativních rozdílech v otevřenosti jednotlivých druhů dat i celých zastupitelských sborů. Příklad takového grafu je obsažen níže.

