

Efektivní způsob sdílení statistických informací

Už jste někdy připravovali a předávali data Mezinárodnímu měnovému fondu, Evropské centrální bance, OECD nebo Eurostatu? Jsou národní účty, průmyslová produkce či údaje o zaměstnancích náhodně vybrané statistické oblasti? Předávání statistických informací na mezinárodní úrovni není jednoduché.

Náročná práce vyplněním tabulek často nekončí. Je třeba doplnit tzv. referenční metadata – obsáhlý text s informacemi o zjišťování, dostupnosti a kvalitě dat. Různorodá je i struktura předávaných tabulek. Liší se jejich formáty, použité SW aplikace i číselníky. Doprovodné textové informace nejsou jednoduše srovnatelné, je třeba je vytvářet ad hoc a nutně dochází ke zvýšené chybovosti. Sporná je tak i interpretace předávaných informací.

Toho si všimly dokonce i výše uvedené nadnárodní organizace. Někdy na přelomu tisíciletí proto vznikla iniciativa, která si klade za cíl sjednotit formu i obsah předávaných informací, harmonizovat obsah a standardizovat nástroje na předávání dat a doprovodných metadata. Tak se zrodil projekt **Statistical Data and Metadata eXchange** (dále SDMX), který odstartoval 14. června 2002, kdy se hlavy statistických útvarů různých mezinárodních organizací – budoucích sponzorů projektu – sešly, aby dohodly další spolupráci. BIS, ECB, Eurostat, IMF, OECD, OSN a Světová banka se spojily s cílem efektivnějšího postupu při sdílení statistických informací. V březnu 2007 bylo podepsáno memorandum o porozumění a dohodnuto rotující předsednictví (každé dva roky). Statistická komise OSN v únoru 2008 vyhlásila, že SDMX považuje za přednostní standard

pro výměnu a sdílení dat a metadata v celé statistické komunitě. Zdůraznila potřeby dalšího zapojení národních a mezinárodních agentur a význam vybudování kapacit a šíření povědomí o SDMX pomocí seminářů, workshopů, manuálů, školení a technické pomoci.

SDMX sestává ze statistických a technických norem, návodů, příruček a softwarových aplikací, které jsou propojeny infrastrukturou IT nástrojů a služeb. Jedná se zejména o statistické standardy a pokyny, modely dat a metadata, architekturu pro výměnu dat a metadata a SW aplikace na podporu a užívání standardů. SDMX není v současnosti jen formát pro předávání údajů, ale může posloužit i jako nosný systém od sběru až po publikování dat pro celé statistické zpracování.

Současný Evropský statistický systém (dále ESS) je založen na heterogenní směsi různorodých právních předpisů. Heterogenita právních předpisů často vede k dodatečným nákladům pro zúčastněné statistické organizace. Tato situace by se měla zlepšit s vytvořením nové generace průřezových statistických předpisů, které by stimulovaly integraci statistických procesů v rámci ESS (viz také společné vize ESS přijaté ESSC – Výborem pro evropský statistický systém – v květnu 2010). Na to navázalo i jednání ESSC v květnu

2011 svojí komplexní legislativní politikou pro ESS.

První dvě oblasti, pro které má vzniknout průřezová legislativa, jsou metadata a datové a metadataové standardy výměny dat – s využitím standardů a nástrojů SDMX. Návrh nařízení (regulation) Rady a Evropského parlamentu byl připraven na podzim 2011 a dán do připomínkového řízení v členských zemích EU. Došlé připomínky byly (podle zástupců Eurostatu) k legislativnímu návrhu velmi kritické a jsou v současnosti analyzovány. Nový, upravený návrh nařízení bude předložen nejdříve k připomínkám členským zemím (únor–březen 2012). Eurostat plánuje projednaný návrh nařízení předložit k projednání a odsouhlasení na podzimním jednání ESSC. Poté bude návrh nařízení představen současně Radě i Evropskému parlamentu k projednání a schválení. Předpokládá se, že nařízení by mohlo vstoupit v platnost v druhé polovině roku 2013.

V návaznosti na výše uvedené aktivity se ve dnech 26. a 27. ledna 2012 uskutečnilo pracovní jednání zástupců Eurostatu, Augusta Götzfrieda, vedoucího Unit B5 (Management of statistical data and metadata exchange), a Marca Pellegrina z Unit B3 (IT and standards for data and metadata exchange) direktorátu B. Tématem jednání byla informace o současném

stavu SDMX, legislativních prací na návrhu nařízení o harmonizaci statistických metadata, o implementovaných statistických a technických standardech a také diskuse o způsobu aplikace těchto standardů při předávání statistických dat a metadata ve formátech SDMX z ČSÚ do Eurostatu. Jednání se zúčastnil okruh zaměstnanců ČSÚ zaměřených na tvorbu statistických metadata, problematiku kvality dat a poskytování dat i metadata Eurostatu a jiným mezinárodním organizacím.

Účastníci jednání z ČSÚ prezentovali zástupcům Eurostatu principy a výsledky statistického metainformačního systému (SMS) implementovaného v ČSÚ a záměry rozvoje SIS v rámci IOP projektu Redesign SIS. Dále proběhla detailní diskuse o našem návrhu, jak propojit SMS ČSÚ se strukturami dat a metadata specifikovaných podle standardů SDMX. Obě strany se dohodly, že tvorba souborů ve strukturách SDMX bude ve statistickém procesu ČSÚ začleněna jako součást fáze diseminace, kde se budou automatizovaně provádět transformace našich dat a metadata do struktur podle standardů SDMX. Realizaci bude předcházet důkladná analýza našich metadata a požadovaných metadata SDMX, která by měla být zakončena vytvořením převodníku mezi metadata ČSÚ a SDMX. Dále bude



zadána a implementována aplikace, která bude podle převodníku provádět transformaci našich metadat na metadata SDMX. Zdroji vstupních dat budou datový sklad (DWH) a při něm vytvořené specializované datové tržiště pro data a databáze SMS.

Formátu SDMX se v současné době v ČSÚ používá pro předávání dat STS (krátkodobé statistiky) do OECD, výsledků SLDB 2011 do Eurostatu (tzv. hyperkostky) a dat

zahraničního obchodu (<http://apl.czso.cz/pll/stazo/STAZO.STAZO>).

Co bude vše výše uvedené znamenat pro řadového statistika?

Zejména se bude muset o celém systému dozvědět více. Pro rok 2012 již oddělení vzdělávání a FKSP (odbor personalistiky a mezd) zařadilo do své nabídky kurz/školení „Základní informace o SDMX a jeho využití na ČSÚ“. V rámci školení budou vysvětleny principy SDMX a datové i metada-

tové struktury SDMX. Dále bude podrobněji vysvětlen postup přípravy výstupních souborů dat a referenčních metadat pro Eurostat ve strukturách SDMX.

Dalšími předpokládanými kroky bude analýza číselníků a statistických proměnných SMS a struktur SDMX s cílem vytvořit převodník mezi metadaty SMS a SDMX. Ve vazbě na převodník pak bude nutno zadat a implementovat proceduru pro transformaci metadat.

Aktuálními úkoly v únoru a březnu 2012 jsou důkladné posouzení nového návrhu nařízení o harmonizaci metadat a zavádění standardů SDMX při předávání dat a metadat do Eurostatu.

Miroslava Brchaňová,
Ebbo Petrikovits
oddělení klasifikací, číselníků
a SMS

Stav ropné nouze jsme zvládli na výbornou

Vloni v létě vyhlásila **Mezinárodní energetická agentura (IEA)** pro všechny členské země „stav ropné nouze“. Evropskému prostoru v souvislosti s vyhrcozenou situací v Libyi totiž hrozil výpadek v zásobování ropou a ropnými produkty. Pro ČSÚ to byl signál k zajištění sběru a zpracování dat z této oblasti v extrémně krátkém období.

Statistika energetiky má řadu odlišností od většiny ostatních statistických šetření. Například jedna z nich spočívá ve skutečnosti, že ve zdrojových částech energetických statistik nelze používat data získaná na základě různých způsobů dopočtů, odborných odhadů apod. Znamená to, že v případě ukazatele např. „změna stavu zásob“ se muselo zjistit, kolik bylo skutečně uskladněno a vyskladněno energetického produktu, například motorové nafty

či automobilového benzínu. Toto pravidlo se týká i dalších bilančních zdrojových ukazatelů, jako jsou těžba, výroba, dovoz, vývoz, převody poloproduktů a následně celkové spotřeby jednotlivých 26 energetických produktů.

Běžně se energetická data o ropě a ropných produktech zjišťují v pravidelných měsíčních termínech. Vloni v létě však IEA vyhlásila mimořádný sběr dat pro období od června do srpna 2011, a to

pouze 16 dnů před začátkem akce. Pro pracovníky útvaru statistiky energetiky ČSÚ to představovalo velmi náročný úkol – zajistit organizaci sběru dat a jejich zpracování v bleskovém termínu. Jenom díky enormně vysokému pracovnímu nasazení a obětavosti se to našim statistikům podařilo.

Obecně platí, že ze strany mezinárodních organizací jsou na energetickou statistiku o ropě a ropných produktech kladeny vysoké nároky, a to zejména z hlediska kvality dat. Za běžné situace se předběžná měsíční data mohou pohybovat v přípustném koridoru odchylek do 3 % od finálních dat. V případě uváděného mimořádného sběru dat byla data všech členských zemí podrobena rozsáhlému a detailnímu hodnocení. Základem tohoto hodnocení byla tři kritéria: kvalita dat s maximál-



ně přípustnou odchylkou do 5 % u každého energetického produktu včetně jejich součtových hodnot (v porovnání k finálním datům), kompletnost dat a splnění požadovaného termínu. Před koncem minulého roku obdržely všechny členské země od IEA vyhodnocení. Skvělou zprávou je, že ČSÚ dosáhl ve srovnání s ostatními evropskými státy vysoce nadprůměrného hodnocení.

Když je nouze, nastupují inspektoři

Stav ropné nouze je upraven mj. ve směrnici Rady 2009/119/ES o nouzových zásobách ropy a ropných produktů. Ta **zavádí možnost inspekce** ze strany orgánů Evropské komise, které mohou provádět kontroly s osobní účastí v kontrolovaném členském státě. Mají přístup ke všem souvisejícím dokumentům a mohou vstupovat do všech vybraných prostor, ve kterých jsou skladovány příslušné zásoby ropy a ropných produktů.

Jan Zámyslický
oddělení statistiky energetiky