


## OTEVŘENÁ DATA

Ing. JIŘÍ ŠUNKA

	<p><b>TÉMA 1</b></p> <p><b>OD ZRODU K SOUČASNOSTI</b></p> <ul style="list-style-type: none"> <li>- Historie otevřených dat</li> <li>- Charakteristika</li> <li>- Otevřená data v ČSSZ</li> </ul>
---	--

Data se dnes dostávají do centra pozornosti, ať už se jedná o otevřená data nebo data velkých objemů (tzv. big data). V případě otevřených dat jde o základní předpoklad pro vytváření hodnoty závislé na schopnosti sdílení těchto dat třetími stranami, nikoliv o jejich objem, jako je tomu u dat velkých objemů. Publikace otevřených dat se stává významnou aktivitou stále většího počtu států a jejich vlád. Celosvětově roste množství oblastí nabízejících sady otevřených dat.

### HISTORIE OTEVŘENÝCH DAT

Pojem Open Data (otevřená data) se poprvé objevil v roce 1995 v dokumentu Americké vědecké agentury, která se zabývala popisem geofyzikálních a environmentálních údajů. Zdůrazňovala se v něm mimo jiné podpora úplné a otevřené výměny vědeckých informací mezi různými zeměmi jako hlavní předpoklad pro analýzu a pochopení globálních jevů.

Již dlouho předtím, než byla definována otevřená data, vědci vnímali výhodu otevřenosti a sdílení dat v oblasti vědeckých objevů a informací. Již v roce 1942 Robert King Merton (významný americký sociolog) vysvětlil, že je důležité, aby výsledky výzkumu byly volně přístupné všem. Každý výzkumný pracovník musí přispět do „společného hrnce“ a vzdát se práv duševního vlastnictví v zájmu široké distribuce a rozvoje nových poznatků. V prosinci 2007 uspořádali internetoví propagátoři v Kalifornii setkání, jehož cílem bylo před nastávající prezidentskou kampaní definovat pojem otevřené veřejné údaje. Základní myšlenkou bylo, že veškeré veřejné údaje jsou společným majetkem, hrazeným ze společných prostředků, proto musejí být veřejné a svobodně dostupné. V roce 2007 to znělo jako sen. Ale výsledek časem překonal veškerá očekávání. Přibližně o rok později prezident Barack Obama po nástupu do Bílého domu podepsal tři prezidentská memoranda. Dvě z nich se týkala otevřeného vládnutí, pro které je otevírání dat jedním z hlavních pilířů. Otevřenost je koncept společný pro počítačové programy s otevřeným zdrojovým kódem (tzv. open source), otevřenou vládou i otevřená data. Je to filozofie činnosti, praxe a způsobu dosahování stanovených cílů.

Dnes jsou veřejná data ve většině zemí definována zákonem, který se dotýká veřejných a soukromých subjektů zapojených do veřejné služby. Dostupnost mnoha údajů je





stále ve vývoji: většina vlastníků dat uděluje prioritu takovým datovým souborům, které lze nejsnadněji otevřít (technicky, právně i politicky). Údaje, které jsou vnímány jako cenné, nebo ty s vyšším potenciálem sociálního a společenského vlivu zůstávají do značné míry mimo rámec otevřených dat. Nehledě na to si myšlenka otevřených dat jasně otevřela cestu do společnosti. Nicméně jejich ambice se postupně mění: cílem už není snažit se změnit svět prostřednictvím dat, ale skromněji zkusit zmodernizovat veřejnou moc.

## CO JSOU OTEVŘENÁ DATA

Jednoduše řečeno, jsou to úplná, snadno dostupná, strojově čitelná data, používající standardy s volně dostupnou specifikací, zpřístupněna za jasně definovaných podmínek s minimem omezení a dostupná uživatelům při vynaložení minima nákladů.

Podle této definice se otevřeným datům přiřazují tzv. stupně otevřenosti, které říkají, do jaké míry splňují výše uvedené parametry. Zveřejňování otevřených dat si klade za cíl zejména umožnit jejich opakované použití a zajistit k nim svobodný a neomezený přístup. Z toho vyplývá možnost zpracování těchto dat k libovolným účelům. A v tom se skrývá (a je postupně odkrýváno) jejich obrovský potenciál. Vytvářejí se tak podmínky k tvorbě nezávislých programů, které tato data zpracovávají, k vypracování studií a hledání souvislostí, které by za jiných okolností znamenaly obrovské úsilí. Současně je umožněn i vznik nových pracovních příležitostí a podporován trh práce. Rovněž otevřenost a dostupnost tohoto typu dat je velkým přínosem pro zvyšování transparentnosti státní správy.

Z hlediska poskytovatele otevřených dat je jeho primárním úkolem data zveřejnit, publikovat, nikoliv se zabývat aplikacemi pro jejich využití. To je úlohou okolního prostředí, schopného s velkou pravděpodobností přijít na možnosti jejich využití, které poskytovatele dat ani nenapadnou. Lze jen souhlasit s konstatováním Rufuse Pollocka, zakladatele nadace Open Knowledge, že „nejzajímavější využití vašich dat za vás vymyslí někdo jiný“.

## OTEVŘENÁ DATA V EU A ČR

O otevřených datech (původně *opakovatelně využitelných datech*) se začalo diskutovat někdy kolem roku 2000 na evropské úrovni v souvislosti s postupným zveřejňováním různých databází členských států EU. V roce 2003 byla vydána směrnice č. 2003/98/ES „o opakovaném využívání informací veřejného sektoru“ konstatující, že mají existovat data veřejné správy, která by byla sdílena napříč veřejnou správou a dala se automatizovaně využívat. Stanovuje rovněž několik definic, zásad a pravidel týkajících se jejich zveřejňování a poskytování a také možnost vytvoření katalogu takových informací v jednotlivých státech. Dále umožňuje na národní úrovni přijmout legislativní opatření na podporu zveřejňování sdílených dat. Od roku 2009 se takovýmto datům říká „opendata“ – tedy otevřená data. Technologický pokrok je nezastavitelný a v jeho souvislosti se začíná v posledních letech pracovat i s pojmem „linkedata“, tedy tzv. propojitelná data jako vyšší stupeň otevřených dat, která je možno vzájemně propojovat a nalézat tak v nich nové souvislosti.

V počátcích otevřených dat v České republice poněkud převládal trochu jednostranný pohled na „opendata“ jen jako na protikorupční opatření a nebyl zcela doce-

něn původní smysl otevřených dat jako základu pro poskytování otevřených služeb. Situace se však postupně stabilizovala a v současné době se Česká republika řadí mezi státy, které výrazně zrychlily svou cestu k otevřeným datům, mají zásadně vyvinutou politiku v této oblasti nebo portál (zpráva Evropské komise „Open Data Maturity in Europe 2017“). Právní úprava otevřených dat v ČR je obsažena v zákoně č. 106/1999 Sb., o svobodném přístupu k informacím, a doplněna o nařízení vlády č. 425/2016 Sb., o seznamu informací zveřejňovaných jako otevřená data. Zároveň jsou otevřená data zahrnuta do několika strategických dokumentů a akčních plánů (jejich seznam a další relevantní dokumenty naleznete na <https://opendata.gov.cz/legislativa:start>).

## OTEVŘENÁ DATA V ČSSZ

Česká správa sociálního zabezpečení se v roce 2014 podílela jako partner na projektu aplikovaného výzkumu *Publikace dat statistických ročenek ve standardu otevřených dat (TD020121)*, který řešila FIS VŠE Praha ve spolupráci se společností KOMIX, s. r. o., a jenž byl podpořen Technologickou agenturou České republiky. V rámci tohoto projektu byl navržen, a na datech statistických ročenek z oblasti důchodového pojištění za roky 2008–2012 pilotně ověřen, způsob publikace těchto dat ve formátu RDF s využitím principů propojených dat. Využití těchto principů umožňuje vzájemně propojovat data v prostředí webu a dávat je tak do souvislostí. Takový přístup k publikování otevřených dat v prostředí internetu je metodicky a technologicky inovativní a zatím jej využívají jenom technicky nejvyspělejší organizace.

Logickým pokračováním bylo pak rozhodnutí vedení ČSSZ o zveřejňování otevřených dat na vlastních webových stránkách. V souvislosti s tímto úkolem byly připraveny odpovídající organizační a technické podmínky. V rámci přípravy bylo přijato zásadní rozhodnutí, že data budou publikována v maximální otevřenosti, že bude vytvořeno vlastní řešení pro publikaci a že kromě samotných dat bude ČSSZ zveřejňovat i jejich vizualizace ve formě interaktivních grafů, což usnadní běžnému netechnickému uživateli jejich interpretaci.

Výsledkem všech těchto aktivit bylo, že ČSSZ zveřejnila 2. listopadu 2015 na adrese <https://data.cssz.cz> portál pro publikaci a vizualizaci otevřených dat a připojila se tak k mezinárodní iniciativě World Wide Web Foundation.

Poskytování informací zveřejněním na portálu otevřených dat je náročným procesem, který má své jednotlivé etapy a před spuštěním vlastního portálu musely být zabezpečeny následující úkony:

- Vypracování návrhu vnitřní organizační směrnice upravující oblast otevřených dat.
- Zajištění podmínek technického charakteru pro realizaci procesu zveřejnění otevřených dat v prostředí ČSSZ, a to jak pro vytváření datových sad, tak pro jejich publikaci a vizualizaci.
- Zajištění bezpečnosti technického řešení i otevřených dat.
- Vypracování publikačního plánu pro zveřejňování otevřených dat na základě vyhodnocení přínosů a rizik.
- Vypracování licenčních podmínek (jinak též podmínek užití) otevřených dat.
- Vytvoření prostředí internetových stránek s otevřenými daty ČSSZ.





- Implementace lokálního katalogu otevřených dat nebo jeho ekvivalentu na webových stránkách ČSSZ.
- Zajištění evidence v Národním katalogu otevřených dat (dále jen NKOD), který provozuje Ministerstvo vnitra.

Tento proces zabral přibližně 6 měsíců.

Před publikací otevřených dat byla provedena detailní analýza dat a informací, které ČSSZ spravuje, včetně informací dostupných v současné době prostřednictvím jejích webových stránek. Na základě této analýzy byly určeny potenciální datové sady, u kterých byly vyhodnoceny přínosy, možná rizika a odhadovaná náročnost jejich publikace. Tím bylo zajištěno, že byla zvolena pouze data, jejichž publikace je v souladu s platnou legislativou, a že byly vhodným způsobem nastaveny priority jejich publikace s ohledem na očekávané přínosy a náročnost samotného zveřejnění dat.

V současné době je publikováno celkem 63 datových sad z oblastí:

- Důchodové pojištění
- Lékařská posudková služba
- Nemocenské pojištění
- Osoby samostatně výdělečně činné
- Personální
- Sociální pojištění
- Správa dat sociálního pojištění
- Správní řízení

ČSSZ důsledně chrání osobní údaje svých klientů a publikovaná data představují statistické údaje, tj. agregovaná data. Proto se zde nevyskytují žádné informace o konkrétních osobách a organizacích.

Velkým přínosem je, že ČSSZ nezůstala pouze u publikace samotných dat, ale připravila také interaktivní vizualizace vybraných datových sad v podobě grafů (sloupcových, spojnicových, koláčových) nebo map na úrovni krajů a okresů. U většiny vizualizací lze interaktivně měnit parametry pro výběr dat pro zvolený graf (například rok, druh důchodu, pohlaví ...) a získané výstupy uložit ve formě tabulky ve formátu CSV pro další zpracování nebo jako obrázek ve formátu PNG.

Z publikovaných 77 vizualizací uvádíme například:

- měsíční výše důchodů
- průměrná délka pobírání starobního důchodu v jednotlivých letech
- průměrný věk u nově přiznaných důchodů v České republice
- počet důchodců v krajích a okresech ČR
- počet důchodů s výplatou do zahraničí
- počet vyplacených dávek nemocenského pojištění podle okresů a krajů
- počet OSVČ podle okresů a krajů

Charakter zveřejněných dat vyplývá ze základních činností, které ČSSZ zajišťuje. Tím je do jisté míry určen i okruh typických uživatelů. Kromě běžného uživatele, který projeví individuální zájem o statistické informace z oblastí důchodového a nemocenského pojištění, jsou data určena každému, kdo se chce hlouběji zabývat zkoumáním závislostí a vztahů mezi daty poskytovanými různými orgány. Jde o datové analytiky, novináře, členy akademické obce, výzkumníky v sociální a demografické oblasti,

statistiky apod. Ti všichni mají možnost svobodného a volného přístupu k publikovaným datům a mohou je zpracovávat dle svého uvážení.

V obecné rovině použitý formát RDF usnadňuje propojování dat z různých zdrojů, a tím i analýzu možných souvislostí mezi jevy, které tato data popisují. Jako příklad lze uvést možnost propojení dat o počtu a struktuře důchodců v jednotlivých okresech a letech, které zveřejňuje ČSSZ, s daty o počtu obyvatel, věkové struktuře v okresech, které zveřejňuje Český statistický úřad. Výsledkem takového propojení je možnost výpočtu ukazatelů jako počet důchodců na počet obyvatel okresu, počet invalidních důchodců ve věkových pásmech na počet obyvatel apod., které lze použít pro demografické studie nebo pro rozhodování o aktivitách orgánů veřejné moci v sociální oblasti. Protože jsou data zveřejněna v časových řadách, lze zpracovávat prognózy a odhady dalšího vývoje jako jeden z podkladů pro řízení společnosti.

Jak už bylo řečeno v úvodu, otevřená data jsou vnímána jako jeden z nástrojů podporujících myšlenku tzv. otevřeného vládnutí. Otevřené vládnutí představuje snahu o posílení transparentnosti, demokracie a spolupráce mezi politiky, veřejnou správou, zástupci podnikatelské sféry a občany. Otevřené vládnutí je i jedním z témat strategie Digitální agenda pro Evropu. Aby se občané a další subjekty mohli efektivně podílet na politické diskusi, potřebují relevantní data a informace. Zvyšování dostupnosti otevřených dat tak pomáhá naplnit tento předpoklad pro vyšší zapojení občanů a dalších subjektů do politické diskuse.

Proto ČSSZ nezůstane pouze u dat zveřejněných od listopadu 2015, ale má zpracovaný podrobný publikační plán zahrnující nové datové sady, které budou zveřejněny v letošním roce. ■

Autor článku pracuje v ČSSZ.

## ÚZSVM SPUSTIL ELEKTRONICKOU AUKCI NEPOTŘEBNÉHO MAJETKU

Úřad pro zastupování státu ve věcech majetkových (ÚZSVM) spustil historicky první státní elektronický aukční systém prodeje nepotřebného majetku. V první fázi nabízí movitý majetek, následně bude spuštěn i prodej nemovitého majetku a zároveň jej budou moci využívat i jiné státní instituce.

Nové internetové stránky na adrese [www.nabidkamajetku.cz](http://www.nabidkamajetku.cz) obsahují tři základní sekce: nabídky majetku pro státní instituce, nabídky připravované a nabídky pro veřejnost aktuálně vyhlášené. Pokud má ÚZSVM nepotřebný majetek, nejprve jej podle zákona musí nabídnout dalším státním institucím (záložka Nabídky pro státní instituce). Když žádná neprojeví zájem, ÚZSVM začne připravovat převod/pronájem (např. zpracovává ocenění), majetek se nachází v záložce Nabídky připravované. Následně je majetek nabízen k prodeji či pronájmu veřejnosti. V sekci Nabídky pro veřejnost jsou všechny nabídky majetku určené k prodeji veřejnosti, a to jak majetek prodávaný v elektronické aukci, tak majetek nabízený v dražbách nebo výběrových řízeních s aukcí.

Elektronické aukce se může účastnit každá fyzická či právnická osoba, která je zaregistrovaná v systému.

