

Metadata jak dál

Karel Charvát, Jan Vrobel

Plan4all

Metadata představují základní vrstvu datových infrastruktur: umožňují data vyhledat, pochopit, posoudit kvalitu a použít je v dalších procesech. Současná praxe je však založena na souběžném používání většího počtu standardů a profilů. Vedle rodiny ISO 19115/19139 se uplatňují DCAT a jeho geografický profil GeoDCAT-AP, SpatioTemporal Asset Catalog (STAC) a OGC API – Records. Tyto přístupy vznikaly pro rozdílné potřeby a technologická prostředí. Jejich interoperabilita proto není automatická a převody mohou vést ke ztrátě významu, nejednoznačností nebo duplikaci údajů. Výsledkem jsou vyšší náklady na implementaci katalogů, datových prostorů a nových informačních systémů i omezená možnost kombinovat data z více zdrojů.

Referát se bude opírat o současnou práci Open Geospatial Consortium (OGC), které hledá praktické cesty, jak jednotlivé standardizační ekosystémy propojit, vymezit jejich role a podporovat strojově zpracovatelná metadata. Shrne hlavní témata diskusí z OGC Metadata Summit 2026, konaného v Bolzanu pod mottem „From Metadata Strategy to Practical Data Infrastructure“. Pozornost bude věnována přechodu od formálního plnění profilů k jednodušší, škálovatelné a udržitelné infrastruktuře podporující principy FAIR, datové prostory, digitální suverenitu a důvěryhodné využití dat napříč sektory.

Naléhavost tématu zvyšuje rozvoj generativní umělé inteligence, velkých jazykových modelů a agentních systémů. Automatizovaný agent může z neúplných nebo heterogenních metadat odvodit chybný význam, zvolit nevhodný zdroj nebo nesprávně sestavit analytický postup. Metadata pro prostředí AI proto musí být jednoznačná a strojově použitelná. Musí zachycovat původ dat, kvalitu a aktuálnost, prostorové a časové pokrytí, licenci, přístupová omezení, použité metody a vazby na služby, procesy i řízené slovníky. Současně je třeba určit, do jaké míry lze metadata pomocí AI vytvářet a kontrolovat a kde musí zůstat odborné ověření člověkem.

Praktická část představí zkušenosti s návrhem a využitím metadat v platformě Jackdaw, kde metadata podporují vyhledávání zdrojů, interpretaci uživatelských dotazů a sestavování prostorových analytických postupů, a dále zkušenosti z projektů FAIR2Adapt a FOCAL. Na těchto příkladech budou ukázány problémy s mapováním profilů, popisem datových služeb, kvality a provenience i s přípravou podkladů pro jazykové modely.

Závěrem bude navržena možná cesta pro Českou republiku a Slovensko: koordinovat kompatibilní aplikační profily, systematicky mapovat ISO, DCAT/GeoDCAT-AP a STAC, zavádět perzistentní identifikátory a sdílené slovníky, zpřístupňovat metadata prostřednictvím moderních API a využívat společné validační nástroje. Důležitým krokem by měla být referenční implementace ověřená na konkrétních národních a přeshraničních případech. Cílem není vytvořit další izolovaný standard, ale postupně vybudovat interoperabilní, vícejazyčné a důvěryhodné metadatové prostředí využitelné lidmi, informačními systémy i AI agenty.