

ARCHIVACE DAT V SOCIÁLNÍCH VĚDÁCH Z POHLEDU ČESKÉHO SOCIÁLNĚVĚDNÍHO DATOVÉHO ARCHIVU

Martin Vávra

Sociologický ústav AV ČR

Konference o šedé literatuře a repozitářích 2018

24. 10. 2018

Archivace sociálněvědních dat v ČR jako výsledek zahraniční inspirace a mezinárodní spolupráce

- V prezentaci se chci podívat na současnou podobu archivace sociálněvědních dat v ČR s důrazem na to, jak její podobu určovaly vlivy z jiných zemí i současný trend mezinárodní spolupráce

Východiska a principy otevřeného přístupu k vědeckým datům

Zásady „open access“ k vědeckým informacím se vztahují i k datům. Tyto zásady byly formulovány na mezinárodní úrovni - například OECD (v roce 2007 Principles and Guidelines for Access to Research Data from Public Funding)

- *„Otevřenost znamená rovný přístup pro mezinárodní vědeckou komunitu za nejnižší možné náklady, pokud možno nejvýše rovné mezním nákladům na poskytnutí dat. Otevřený přístup k vědeckým datům by měl být jednoduchý, uživatelsky přátelský (s preferencí online přístupu) a bez zbytečných průtahů.*

Od doporučení a nevymahatelných principů se posouváme k pravidlům, zakotveným v pravidlech poskytování veřejných peněz na výzkum – viz EU Horizon 2020, kde v principu platí, že k datům, která v rámci něj vznikla, by měl být zajištěn otevřený přístup

Význam otevřeného přístupu k vědeckým datům

- umožňuje stavět na předchozích výsledcích vědy a tím podporuje kumulativnost a kvalitu poznání
- usnadňuje kontrolu vědeckých postupů (skrze replikace analýz)
- podporuje spolupráci (i na mezinárodní úrovni) a snižuje pravděpodobnost duplikace výzkumů na stejné téma a tím podporuje efektivitu výdajů na vědu
- usnadňuje zahrnutí širšího spektra aktérů do oblasti vědy (viz současný trend „citizen science“)

Český sociálněvědní datový archiv

- Založen roku 1998 pod názvem Sociologický datový archiv
 - Přímá inspirace zahraničními archivy, například německým Zentral Archivem při institutu GESIS
- Postupný růst a ustavení jakožto samostatné oddělení SOU AV ČR
- V roce 2011 změna názvu na Český sociálněvědní datový archiv – širší zaměření akviziční politiky
- V současné době více než 800 datových souborů, především ze sociologických dotazníkových výzkumů
- ČSDA není pouze archivem, ale také metodologickým centrem
- Základem pro další směřování je začlenění do organizace CESSDA (Consortium of European Social Science Data Archives)

Přístup k sociálněvědním datům na evropské úrovni – organizace CESSDA

- CESSDA byla původně neformálním sdružením archivů
- V současnosti je CESSDA součástí evropské výzkumné infrastruktury (European Research Infrastructure Consortium - ERIC)
- Do CESSDA nyní patří 18 evropských archivů
- CESSDA usiluje o vytvoření jednoho místa („one stop shop“), které bude uživatelům sloužit jako přístupový bod ke všem datům, jež jsou uloženy v jednotlivých členských archivech
- CESSDA se hlásí k iniciativě European Open Science Cloud (EOSC) a k principům FAIR Data (findability, accessibility, interoperability a reusability)

Standardy archivace

Vytváření a prosazování nadnárodních standardů je nutným krokem v procesu integrování datových archivů

- Metadatový standard DDI – jednotný způsob popisu dat, jeho východiskem byl v počátcích Dublin core
- Tezaurus pro sociální vědy ELSST – strukturovaný slovník klíčových slov
- Systém certifikace CoreTrustSeal – certifikace důvěryhodných digitálních archivů

Výzvy pro archivaci dat v současnosti

Přestože v sociálních vědách v ČR (a obecně ve všech zemích, jejichž datové archivy jsou členy CESSDA) je situace archivace a zpřístupňování dat vcelku dobrá, stále platí, že je:

- potřeba zlepšovat kulturu sdílení dat
- nutné zvýšit povědomí výzkumníků o potřebě vytvářet “data management plány” a obecně lépe pracovat s daty
- potřebné postoupit v oblasti archivace a zpřístupňování nových typů dat – například dat ze sociálních sítí na internetu

Závěry

- Základní infrastruktura už existuje (i když je samozřejmě nadále potřeba ji rozvíjet, zejména na celoevropské úrovni)
- Potřeba stanovit (což lze samozřejmě pouze na úrovni vědní politiky státu, ale pokud možno na základě debaty zainteresovaných aktérů) jakou podobu otevřeného přístupu k vědeckým datům chceme a jak ji budeme podporovat a také jak budeme vědce a jejich instituce motivovat aby na tomto spolupracovali.

- Děkuji za pozornost

- Pozdější připomínky a dotazy možno směřovat na martin.vavra@soc.cas.cz