
Tipy pro sdílení dat

Ujistěte se, že data opravdu můžete sdílet

Ne všechna data je možné otevřeně sdílet, ujistěte se proto, že svá data opravdu sdílet můžete. Rozhodování o tom, zdali data můžete sdílet či nikoli ovlivňuje mnoho různých faktorů. Pokud pracujete v rozsáhlejší výzkumné skupině, musí se sdílením dat **souhlasit všichni spoluautoři**. Pokud používáte **data třetích stran**, musí jejich **autor se sdílením souhlasit**. Tento souhlas může být vyjádřený prostřednictvím přidělené licence, nebo se můžete na podmínkách sdílení domluvit přímo s autorem datasetu. Obsahují-li vaše data **osobní údaje**, je třeba je z datasetu **odstranit, anonymizovat, nebo mít k jejich sdílení souhlas**. Data spojená s **národní bezpečností** typicky není možné sdílet. V některých případech dokonce není možné sdílet ani metadatové záznamy o jejich existenci.

Zároveň s daty sdílejte dokumentaci

Zároveň s daty by vždy měla být sdílena také jejich dokumentace, aby se uživatelé v datech snadno zorientovali (např. formou Read-me souboru) a snížilo se riziko mylné interpretace dat z důvodu nedostatečného popisu, co data znamenají. Forma dokumentace se bude odvíjet od konkrétního typu dat, s jakými pracujete, může se jednat např. o tabulku s popisem jednotlivých proměnných, souhrnná demografická data účastníků výzkumu, komentáře v počítačovém skriptu a podobně. Pro dokumentaci obecně platí princip „sdílení se svým budoucím já“, tedy přiložit k datům takové informace, které byste sami k interpretaci svých dat potřebovali, kdybyste se k nim vrátili třeba za deset let.

Licencujte svá data

Pokud sdílíte výzkumná data, je dobré opatřit je vhodnou licenci, aby uživatelé věděli, jak s nimi mohou nakládat. Pro tyto účely se typicky využívají licence Creative Commons, prostřednictvím kterých může autor některá svá práva k dílu uživatelům poskytnout a jiná si naopak vyhradit. Je však možné aplikovat i vlastní licenci, která bude vyhovovat vašim požadavkům. Existuje také možnost poskytovat stejná data pod více licencemi, například zdarma pro nekomerční využití a pro komerční využití za úplatu. Pokud vaše datová sada není předmětem ochrany autorských práv, je dobré ji označit jako Public Domain Mark, aby uživatelé měli jistotu, že data mohou využít.

Využívejte perzistentní identifikátory

Aby byla data snáze nalezitelná („findable“ dle principů FAIR), měla by být opatřena perzistentním (trvalým) identifikátorem jako je například DOI nebo handle. Perzistentní identifikátor datům zpravidla nemůže autor přidělit sám, je na to třeba služba, která má oprávnění identifikátory přidělovat. U datových sad mohou tuto službu poskytovat například některé repozitáře, proto když vybíráte vhodný repozitář, zkontrolujte, zda trvalý identifikátor přiděluje.

Uveďte doporučený formát citace

Pro usnadnění citování vašich dat je dobré přiložit k datům doporučený formát citace. Tím jednak uživatelům připomenete, že pokud vaše data používají, je třeba, aby je řádně citovali, a zároveň jim citování usnadníte. Některé repozitáře uvádí doporučený formát citace na úvodní stránce s metadaty podle údajů, které zadáte, nebo je možné uvést doporučený formát citace jako součást dokumentace.