

## DOPORUČENÍ PRO TVŮRCE PLATFOREM

Doporučení se týkají otázek spojených s návrhem online platforem pro výzkum a nejsou spojena s konkrétním technickým řešením. Online platformami pro výzkum rozumíme digitální knihovny, databáze, nebo geografické informační systémy.

Nejedná se o vyčerpávající popis ani obecně platná pravidla; spíše oblasti, nad kterými je vhodné se zamyslet.

1) Stanovit účel a žádoucí funkce .....	2
2) Popsat obsah .....	2
3) Dokumentovat postup i výsledek .....	3
4) Zajistit autorská práva k vystavovanému obsahu .....	3
5) Určit, pod jakou licencí jsou data poskytována a jak je citovat .....	3
6) Myslet na uživatele.....	4
7) Sdílet informace a (meta)data .....	4

### Volně zpracováno podle

- FAULDER, Erin, Jim DELROSO, Jenn COLT, et al. Cornell University Library Repository Principles and Strategies Handbook. *Cornell University Dashboard* [online]. Cornell University, 2018 [cit. 2019-07-25]. Dostupné z: <https://confluence.cornell.edu/x/18Z0F>
- MILLER, Steven J. *Metadata for digital collections: a how-to-do-it manual*. New York: Neal-Schuman publishers, [2011]. ISBN 978-1-55570-746-0.
- NATIONAL INFORMATION STANDARDS ORGANIZATION. *A Framework of Guidance for Building Good Digital Collections*. US, 2007. ISBN 978-1-880124-74-1. Dostupné také z: <http://framework.niso.org>
- What is a Relational Database?. In: *Nodegoat* [online]. 2017-03-28 [cit. 2019-07-25]. Dostupné z: <https://nodegoat.net/blog.p/82.m/20/what-is-a-relational-database>

## 1) Stanovit účel a žádoucí funkce

Základní otázky můžeme shrnout jako: Proč by měla platforma existovat? Co má být jejím obsahem? Kdo s ní bude pracovat a jakým způsobem?

Důvod existence platformy je pravděpodobně daný výzkumným projektem a absencí jiného vyhovujícího zdroje. S platformou obvykle pracuje její správce (výzkumný tým) při přidávání a úpravě obsahu. Tento výzkumný tým ji dále využívá i po naplnění obsahem. Pokud je platforma veřejně přístupná, rozšiřuje se skupina potenciálních uživatelů a jejich potřeb. Potřeby uživatelů a žádoucí funkce je možné popsat více způsoby, od prostého textového popisu po různé typy modelů. Existuje také množství metod pro zjištění těchto potřeb (více bod 6). Pro funkčnost jsou zásadní také rozhodnutí týkající se popisu obsahu (více bod 2).

Další požadavky se mohou týkat dat o využívání platformy. Například je vhodné určit, zda chceme sledovat počty stažení souborů nebo zobrazení jednotlivých záznamů a zda tato data chceme zobrazit veřejně přímo v systému.

## 2) Popsat obsah

Aby bylo možné s objekty ve sbírce (naskenovanými materiály, daty, fotografiemi, texty, ...) dále pracovat, je potřeba je vhodně popsat.

V případě fiktivní sbírky dopisů se jejich popis může týkat samotných dopisů (jména osob, které spolu korespondovaly, popis jejich obsahu, data odeslání, formát papíru, barvu inkoustu, ...), nebo odesílatelů a příjemců (země původu, oblasti působení, ...). Nakolik komplexní popis bude záleží na potřebách uživatelů, v první řadě samotného výzkumného týmu, který sbírku tvoří. Určujícím faktorem je také množství času, který je pro tvorbu k dispozici.

Při formulování pravidel popisu je potřeba rozhodnout jaké údaje zahrnout, odkud je zjišťovat a jakým způsobem je zapisovat.

### **Příklady potenciálně problematických oblastí**

- Objekty ve sbírce mohou mít autory nebo jiné s nimi asociované osoby, které by měly být jednoznačně identifikovatelné. Problematické mohou být pseudonymy, různé varianty zápisu jména (J. R. R. Tolkien vs John Ronald Reuel Tolkien) nebo více osob stejného jména.
- Podobné problémy mohou nastat při uvádění geografických oblastí. Například pokud fiktivní sbírka obsahuje dopisy odeslané ze Zlína v období, kdy se jmenoval Gottwaldov, je potřeba zvážit, jaký název použít.
- Popis obsahu může být velmi subjektivní a často jej pro jednu sbírku provádí více osob. Je proto vhodné zvážit sepsání povolených termínů a instrukcí pro jejich použití, včetně toho, kolik jich je nutné zadat. Stejný objekt lze popsat z různých hledisek a volba termínů by proto měla být přizpůsobena účelu sbírky.
- Problémy mohou způsobovat různé způsoby zápisu data (4. únor 2020 vs 2020-02-04 vs February 4, 2020), rozsahy (1800–1900 vs 19. století), textově označená období

(pozdní středověk, jura) a odhadovaná data (cirka 1800 vs 1800?). Ať už bude zvolený formát jakýkoliv, je důležité být při jeho používání konzistentní.

- Existují volně dostupné seznamy osob, míst i témat, které mohou zahrnovat také synonyma a vztahy mezi obsaženými termíny. Například pro místní názvy lze využít Getty Thesaurus of Geographic Names (TGN); záznam pro Zlín uvádí i alternativní název Gottwaldov. Využití těchto slovníků může při vhodné implementaci zajistit jednoznačnou identifikaci za hranicemi konkrétní platformy.

Všechna výše uvedená rozhodnutí o způsobu popisu by měla být popsána a dostupná všem, kdo budou s obsahem pracovat. Tato rozhodnutí budou ovlivněná mimo jiné oborovými zvyklostmi a měla by být činěna s ohledem na budoucí využitelnost a udržitelnost platformy a jejího obsahu.

Lišit se bude také způsob, kterým budou popisné informace uloženy, a to v závislosti na zvoleném technickém řešení.

### 3) Dokumentovat postup i výsledek

Na tvorbu dokumentace často nezbyvá čas, přesto je nezbytná pro plné využití potenciálu pracně budované platformy. Technická dokumentace je zásadní pro případné strojové zpracování obsažených dat nebo jejich přesun do jiného systému (například pokud stávající řešení zastará a přestane správně fungovat). Pro inspiraci lze využít nabízenou Šablonu dokumentace (PDF) se základními okruhy, kterým je vhodné se věnovat.

Pro lepší využitelnost platformy je důležitá také dokumentace rozhodnutí, která vedla k jejímu vzniku a formovala její podobu. Základní informace o obsahu a cílech projektu by měly být dostupné rovněž na stránkách platformy, má-li veřejnou adresu.

### 4) Zajistit autorská práva k vystavovanému obsahu

Pokud je součástí platformy digitalizovaný obsah, je nutné vědět, zda a jak s ním lze nakládat. Je také nezbytné uvést vlastníky autorských práv. Speciální pozornost je pak potřeba věnovat osobním a citlivým datům, jsou-li ve sbírce obsažena.

Pokud uběhlo více než 70 let od úmrtí autora digitalizovaného obsahu, pak lze s dílem automaticky nakládat jako s volným dílem. Pokud však uběhlo méně, pak je třeba dodržet zákonné důsledné vyhledávání podle § 27b odst. 3 Autorského zákona. V případě, že vlastníka autorských práv nedohledáme ve zmíněných zdrojích, pak postupujeme podle § 37a a zažádáme o Licenci pro určitá užití osiřelého díla.

### 5) Určit, pod jakou licencí jsou data poskytována a jak je citovat

Nejen u volně dostupných sbírek se nabízí otázka dalšího využití zahrnutých dat. Pokud jste sami tvůrci jejího obsahu nebo jeho části, je vhodné uvést na stránkách, pod jakou licencí je tento obsah poskytován. Je možné využít například licence Creative Commons. Jako nejméně restriktivní je doporučována varianta CC BY.

Licencovat lze také samotnou databázi jako celek. Ta může být chráněna autorským právem, pokud jde o tzv. originální databázi, nebo právem pořizovatele databáze. Více informací a názorný diagram lze najít [na stránkách Otevřená data](#).

Pro další práci s obsahem je důležité jednoznačně určit, kdo je jeho tvůrcem a kdy byl obsah vytvořen a publikován. Dále je zvykem uvést, jak by měla být platforma správně citována, ať už jako celek, nebo jednotlivá obsažená data. Citování dat jsme věnovali [přehledový článek](#) (z června roku 2020).

## 6) Myslet na uživatele

Uživateli platform pro výzkum jsou obvykle samotní tvůrci. Pokud má platforma potenciál dalšího využití, je vhodné zapojit do jejího vzniku více osob a zjišťovat, jaké jsou jejich potřeby. Důležitou součástí tvorby webového obsahu a služeb je také [přístupnost](#) pro uživatele s různými druhy omezení.

Díky prototypům (průběžným, nehotovým verzím) je možné platformu otestovat a zjistit, zda je pro uživatele srozumitelná. Pro uživatelské testování je vhodné si připravit zadání úkolů, které mají testeři provést: například zobrazit všechny dopisy, které si vyměnila osoba A s kýmkoliv ze Zlína. Základní informace o realizaci uživatelského testování nabízí například web [100 metod](#), spolu s dalšími metodami pro zjišťování uživatelských potřeb.

Po spuštění lze zkoumat, jak uživatelé platformu používají. Využít lze data generovaná samotným systémem (např. s využitím nástrojů Google Analytics nebo HotJar), nebo metody jako pozorování či stínování uživatelů.

## 7) Sdílet informace a (meta)data

Aby se o platformě dozvěděli potenciální uživatelé, vyplatí se zjistit, zda neexistuje rejstřík platform podobného zaměření, kam by bylo možné ji zaregistrovat. Platformy vzniklé na Filozofické fakultě MU shromažďujeme [v katalogu](#) na stránkách projektu LINDAT/CLARIAH-CZ. Dále existují společné vyhledávací služby, které sbírají metadata\* z více platform. V takovém případě je obvykle potřeba, aby metadata splňovala stanovené požadavky.

\* Americká národní organizace pro informační standardy (NISO) metadata definuje jako strukturované informace, které popisují, vysvětlují, lokalizují či jinak zjednodušují vyhledání, využití nebo správu informačních zdrojů.

---

Budeme vděční za vaše zkušenosti, komentáře a doplnění: [lindat-clariah@phil.muni.cz](mailto:lindat-clariah@phil.muni.cz). Pokud nějakou platformu připravujete, rádi se s vámi setkáme a budeme se hlouběji věnovat vašim potřebám.