



Anleitung für die Bereitstellung von Daten für Hackathons

Version 2, 11. Dezember 2015

Auswahl der Datensets und der Bild-/Ton-/Video-Daten

Die Grundlage für den Erfolg eines Kultur-Hackathons liegt in der Bereitstellung von geeigneten Daten. Dazu sollten Sie in Ihrem Bestand diejenigen Datensätze identifizieren, für die Sie die Nutzungsrechte haben bzw. wo die Rechtklärung möglich ist.

Bitte schränken Sie die Datensätze nicht schon vorab unnötig ein, sondern prüfen Sie intern, wie viel Zeit Sie für den Hackathon aufwenden können und möchten. Suchen Sie den Dialog mit potentiellen NutzerInnen, um zu eruieren, welche Datensätze aus Nutzer-Sicht besonders interessant erscheinen.

Für den Hackathon sind grundsätzlich folgende Arten von Daten von Interesse:

- Metadaten von Überlieferungsobjekten
- Digitale Bilder, Videos, Tonaufnahmen / digitale Reproduktionen von Überlieferungsobjekten
- Vokabulare / Schlagwortverzeichnisse
- Normdaten
- Daten aus dem Bereich der Kultur- und Kreativwirtschaft: z.B. Daten zu kulturellen Ereignissen (Theateraufführungen u.ä.), Nutzungsstatistiken, welche über die Nutzung von Kulturgütern Aufschluss geben, usw.

Bitte beachten Sie, dass aus Sicht der Nutzerinnen und Nutzer die eigentlichen Inhalte (Texte, Bilder, Audio-, Videodateien) oftmals ebenso wichtig sind wie die Metadaten. Im Idealfall stellen Sie daher Metadaten zusammen mit dem dazugehörigen Bild-, Ton- oder Videomaterial bereit (insbesondere wenn keine urheberrechtlichen Einschränkungen bestehen).

Rechteklärung / Wahl der Lizenz

Klären Sie die Rechte an den ausgewählten Daten und wählen Sie eine passende Lizenz für die Veröffentlichung. Wenn Sie Argumentationshilfen benötigen oder Fragen zu Lizenzen haben, können Sie sich gerne mit uns in Verbindung setzen. Zu Fragestellungen rund um die "Public Domain" hat das Schweizerische Institut für Geistiges Eigentum ein "FAQ" mit den häufigsten Fragen und Antworten veröffentlicht.

Ideal:

- Creative Commons - Public Domain Mark
Die Kennzeichnung als "Public Domain" empfiehlt sich für alle Werke, deren Urheberrechtsschutz abgelaufen ist.
- Creative Commons - Public Domain Dedication - "no rights reserved"
Mit dieser Lizenz wird soweit gesetzlich zulässig auf alle urheberrechtlichen Ansprüche verzichtet. diese Urheberrechtslizenz wird beispielsweise bei der Veröffentlichung von Metadaten empfohlen, um den Nutzern zu signalisieren, dass auf die jeweiligen Metadaten kein Urheberrechtsanspruch geltend gemacht wird.
- Creative Commons - Namensnennung
Bei Verwendung dieser Lizenz muss der Urheber/die Institution genannt werden. Geben Sie bitte an, in welcher Form auf den Urheber / die Institution verwiesen werden soll. Beachten Sie dabei, dass die Werke urheberrechtlich geschützt sein müssen und die Institution über ausreichende Nutzungsrechte verfügen muss, um die Werke freizugeben.

Auch akzeptiert:

- Creative Commons - Namensnennung, Weitergabe unter gleichen Bedingungen
Diese Lizenz zwingt die Nutzerinnen und Nutzer, ihre darauf aufbauenden Werke unter die gleiche Lizenz zu stellen. Beachten Sie dabei, dass die Werke urheberrechtlich geschützt sein müssen und die Institution über ausreichende Nutzungsrechte verfügen muss, um die Werke freizugeben.

Bitte keine Einschränkung der kommerziellen Nutzung

Sehr problematisch sind in der Regel Lizenzen, welche die kommerzielle Nachnutzung ausschliessen und somit oft legitime Nutzungsszenarien wie die meiste journalistische Berichterstattung oder die Verwendung in der Wikipedia unmöglich machen.

Formatwahl und Bereitstellung

Daten, insbesondere Metadaten zu Überlieferungsobjekten, müssen maschinenlesbar sein, damit sie auf einem Hackathon eingesetzt werden können. Als leicht lesbare Formate haben sich JSON (JavaScript Object Notation) und CSV (Comma Separated Values) etabliert. In der Regel erfolgt die Überführung in diese Formate problemlos, beispielsweise über Excel oder eine Exportfunktion der Datenbank. Wenn Sie intern eine Anlaufstelle für technische Fragen haben, dann werden Sie damit sicherlich keine Probleme haben. Natürlich können Sie uns bei Fragen zu konkreten Fällen gern kontaktieren und wir schauen, wie Ihre Dokumente sauber in eine maschinenlesbare Datei überführt werden können.

Zusätzlich zu den oben genannten Formaten eignen sich in bestimmten Fällen für die Daten auch folgende Formate: GeoJSON, KMZ, XML, RDF oder Shapefile. Gern gesehen wird es in der Entwickler-Community auch, wenn Daten über offene Web-Dienste bereitgestellt werden, wie z.B. RESTful API, Web Feature Service (WFS) oder Web Map Service (WMS).

Bilddateien sollten bei Originalgrösse eine Auflösung von mindestens 300 dpi haben und einzeln in üblichen freien Formaten wie PNG, JPEG, SVG vorliegen (und nicht als GIF, PDF oder Word-Dokumente). Oftmals ist es sinnvoll, auch gleich hochauflösende TIFF-Dateien zur Verfügung zu stellen (z.B. für den Upload auf Wikimedia Commons oder wenn zu erwarten ist, dass von gewissen Bildern Ausschnitte verwendet oder einzelne Bilder retouchiert werden - in all diesen Fällen ist es sinnvoll, die Bilddateien in der bestmöglichen Qualität vorliegen zu haben).

Für Audio-/Videodateien eignen sich die folgenden Formate: Ogg (in Verbindung mit dem FLAC, Speex, Opus oder Vorbis Codecs), FLAC, WAVE, sowie WebM.

Achten Sie auf eine Vereinheitlichung der Daten und darauf, dass maschinell Erfassbares automatisch verarbeitet werden kann — statt Prosa wie „zur Zeit der Völkerwanderung“ lieben Computer und Software-Entwickler Jahreszahlen. Einheitliche Schreibweisen von Namen helfen bei der statistischen Auswertung.

Die Daten werden vor der Veranstaltung auf der Webseite des Hackathons präsentiert und zum Download angeboten. So bekommen die Teilnehmenden einen ersten Eindruck und können bereits Ideen formulieren.

Die Daten und Mediendateien können Sie entweder auf der Website Ihrer Institution oder via einen externen Dienst bereitstellen. Für Daten eignen sich beispielsweise datahub.io oder das Open Government Data Portal des Bundes. Für Mediendateien eignen sich Wikimedia Commons oder Flickr. Konktatieren Sie uns, falls Sie Fragen zu einem der externen Dienste haben.

Nachhaltigkeit

Mittelfristig sollten Sie dafür sorgen, dass Ihre Daten und Mediendateien für Dritte auch nach dem Hackathon gut auffindbar bleiben. Wir empfehlen daher in der Regel eine Bereitstellung der Daten auf dem Schweizer Open Government Data Portal bzw. auf Europeana sowie eine Bereitstellung der Mediendateien (inklusive Metadaten) auf Wikimedia Commons. Bitte beachten Sie, dass die Veröffentlichung der Daten auf diesen Plattformen einiger Vorbereitungszeit bedarf und daher möglicherweise erst nach dem Hackathon erfolgen kann.

Kontaktpersonen

Ihre Ansprechpartner für Fragen im Zusammenhang mit der Bereitstellung der Daten:

Oliver Waddell
043 259 69 74
oliver@waddell.ch

Beat Estermann
031 848 34 38
beat.estermann@openglam.ch

Internetseiten

- OpenGLAM Schweiz, <http://glam.opendata.ch/>
- Webseite des Hackathon, <http://make.opendata.ch/wiki/event:2016-07>
- datahub.io, <http://datahub.io/>
- Open Government Data Portal des Bundes (ab 2. Februar 2016), <http://opendata.swiss>
- Wikimedia Commons, <https://commons.wikimedia.org/>
- Flickr, <https://www.flickr.com/>

Lizenzen

- Eidgenössische Institut für Geistiges Eigentum, FAQ zu Public Domain, <https://www.ige.ch/urheberrecht/haeufige-fragen/public-domain.html>
- Creative Commons, Public Domain Mark, <http://creativecommons.org/publicdomain/mark/1.0/deed.de>
- Creative Commons, Public Domain Dedication, <http://creativecommons.org/publicdomain/zero/1.0/deed.de>
- Creative Commons, Namensnennung, <http://creativecommons.org/licenses/by/4.0/deed.de>
- Creative Commons, Namensnennung, Weitergabe unter gleichen Bedingungen, <http://creativecommons.org/licenses/by-sa/4.0/deed.de>