

Daten bereitstellen

Coding da Vinci Süd 2019

Die Grundlage für den Erfolg des Kultur-Hackathons liegt in der Bereitstellung von geeigneten Daten. Reichhaltige und umfangreiche Datensets sind am attraktivsten. Schränken Sie deshalb bitte die Datensätze nicht schon vorab unnötig ein. Oftmals ergeben sich im Rahmen des Hackathons ungeahnte Anwendungsmöglichkeiten. Coder*innen und Entwickler*innen erkennen in Ihren Datensets möglicherweise Anknüpfungspunkte zu anderen Daten, die Sie bisher nicht gesehen haben. Metadaten und Mediendateien sind gleichermaßen spannend. Zwar beinhalten Metadaten die gewichtigeren Informationen, doch aus Sicht der Nutzerinnen und Nutzer sind Medien, Sound oder Bild, ein wichtiger Motivationsfaktor. Im besten Fall können Sie Bilder/Ton/Videomaterial o.ä. **und** umfassende Metadaten bereitstellen.

Daten der teilnehmenden Kulturinstitutionen: <http://codingdavinci.de/daten/>

Entstandene Projekte: <http://codingdavinci.de/projekte/>

1. Rechteklärung / Wahl der Lizenz

Bei der Wahl der Daten sollten Sie aus Ihrem Bestand diejenigen Datensets identifizieren, an denen Sie die vollen Nutzungsrechte haben bzw. wo eine eindeutige Rechteklärung möglich ist. Wenn Sie Argumentationshilfe benötigen oder Fragen zu Lizenzen haben, setzen Sie sich gern mit uns in Verbindung. Die benötigten Datensets der teilnehmenden Kulturinstitutionen müssen unter einer offenen Lizenz verfügbar sein, um beim Hackathon verwendet werden zu können. Achten Sie bitte auch darauf, dass die Rechte, die in den Daten stehen, identisch sind mit den Rechten, die Sie dem Datensatz als Ganzes gegeben haben.

Im Sinne von Coding da Vinci gelten diese Creative Commons-Lizenzen als offen¹:

- **CC 0** Public Domain Dedication - “no rights reserved”
- **CC BY** Namensnennung (Der Autor/die Institution muss genannt werden.)
- **CC BY-SA** Namensnennung, Weitergabe unter gleichen Bedingungen (Die Lizenz zwingt den Nutzer/in sein Folgewerk unter die gleiche Lizenz zu stellen.)
- Auch die Kennzeichnung **Public Domain** für gemeinfreie Inhalte ist akzeptiert.

2. Formatwahl

Daten, insbesondere Metadaten zu digitalen Objekten, müssen maschinenlesbar sein, damit sie in einem Hackathon eingesetzt werden können. Als leicht maschinenlesbare Formate haben sich XML (Extensible Markup Language), JSON (JavaScript Object Notation) und CSV (Comma Separated Values) etabliert. Dabei sind strukturierte Formate wie XML und JSON zu bevorzugen, wobei XML zusätzlich die Möglichkeit bietet, die syntaktische Korrektheit durch Validierung gegen ein Schema sicherzustellen. Ideal wäre es, wenn Sie ein in Ihrer Sparte etabliertes Metadaten-Austauschformat wie z.B. LIDO für Museen, EAD für Archive, etc. verwenden.

In der Regel erfolgt die Überführung in diese Formate problemlos, beispielsweise die Exportfunktion Ihrer Datenbank. Achten Sie darauf, dass die in den Daten hinterlegten Links auch wirklich zur digitalen

¹ Zu den Problemen mit NC-Lizenzen (nicht kommerzielle Nutzung) und zur Wahl der richtigen Lizenz allgemein empfehlen wir: http://irights.info/userfiles/CC-NC_Leitfaden_web.pdf

Ressource führen. Fügen Sie bitte Ihren Daten eine [Dokumentation](#) bei. Bilddateien sollten in möglichst hoher Auflösung, typischerweise mehr als 1.200 Pixel weit oder hoch, und einzeln in üblichen Formaten wie TIFF, PNG, JPEG oder GIF vorliegen, nicht als PDF oder Word-Dokument.

Achten Sie auf die Vereinheitlichung der Daten, so dass maschinell Erfassbares automatisch verarbeitet werden kann — verwenden Sie möglichst normierte Ansetzungen für Jahreszahlen statt Prosa wie „zur Zeit der Völkerwanderung“. Achten Sie bitte auf einheitliche Schreibweisen von Namen und Kategorien. Sinnvoll sind immer standardisierte Ansetzungen und einheitlich strukturierte Daten. Idealerweise verwenden Sie hierfür Normdaten-URIs, z.B. aus der [Gemeinsamen Normdatei](#), und kontrollierte Vokabulare wie z.B. den [Getty Arts & Architecture Thesaurus](#).

Natürlich können Sie sich mit Fragen direkt an uns wenden. Gemeinsam finden wir eine Lösung. Weiterführende Informationen finden Sie auch auf den Webseiten von [digiS](#)² und im Datenpartnerportal [DDBpro](#)³ der Deutschen Digitalen Bibliothek.

3. Bereitstellung

Alle Daten werden wir im Vorfeld auf unserer Webseite codingdavinci.de/events/sued/praesentieren. So bekommen die Teilnehmenden einen ersten Eindruck und können bereits Ideen für sich formulieren. Dazu benötigen wir Ihre Mithilfe, indem Sie Ihre Daten beschreiben. Ein aus unserer Sicht gutes Beispiel finden Sie in der Beschreibung des [Museums für Kunst und Gewerbe Hamburg \(MKG\)](#).

Für den Upload von Mediendateien mit Metadaten auf eine bereits existierende Plattform empfehlen wir [Wikimedia Commons](#). Für Mediendateien mit nur sehr wenig Metadateninformation bietet sich auch [flickr Commons](#) als Plattform an. Wir unterstützen Sie gern bei der Bereitstellung ihrer Daten, setzen Sie sich bitte mit uns dafür frühzeitig in Verbindung.

Metadaten und zur Vorschau bestimmte Mediendateien, die an die Deutsche Digitale Bibliothek (DDB) geliefert werden, können natürlich über die API der DDB abgerufen und auf codingdavinci.de verlinkt werden.

Kontakt: cdvsued@zd-b.de

Lucy Patterson, Projektmanagement Coding da Vinci: Lucy.Codingdavinci@zib.de

Barbara Fischer, Gesamtkoordination Coding da Vinci: Barbara.Fischer@wikimedia.de

Telefon: +49 30 219 158 26 0 (Wikimedia Deutschland Zentrale)

² <https://www.digis-berlin.de/wissenswertes/>

³ <https://pro.deutsche-digitale-bibliothek.de/>