arbido

2019/3 Zugang

Fornaro Peter, Leitungsteam DHLab / DaSCH

Tags: Benutzung Zugang,

Zugang zum digitalen Archiv

Die wachsende Menge an digitalem Archivgut und die damit verbundenen Kosten zur Speicherung wecken zunehmend den Wunsch nach Funktion und Anwendbarkeit der digitalen Daten. Open Access und Interoperabilität sind in diesem Zusammenhang von Wichtigkeit. Ein Blick auf die Qualität der digitalen Daten ist um so wichtiger.

An einem normalen Tag werden ca. 80 Millionen Fotografien auf die Plattform Instagram geladen, auf youtu.be sind digitale Video mit einer gesamthaften Spieldauer von 58'000 Jahre gespeichert, einzelne digitale Fotografien können heute problemlos grösser als 1.5 GByte gross sein und eine Stunde Film in Kinoqualität benötigt einen Speicherplatz von mehr als 6 TByte. Die Kosten für die Speicherung dieser Datenmengen sind auch heute noch immens und um so mehr stellt sich Frage nach der Nutzbarmachung von gespeicherten Inhalten, nach der Auffindbarkeit und Hilfsmitteln zur Sichtbarmachung. Eine Funktionalisierung von digitalen Daten erleichtert zudem das Rechtfertigen der mit der Speicherung verbundenen Kosten. Das Öffnen von Archiven ist ein wichtiger Schritt in der Wissensvermittlung und es erlaubt die Nutzung digitaler Quellen für die Forschung. Open Data bedeutet, dass Daten uneingeschränkt genutzt und wiederverwendet werden dürfen. Über welche Kanäle die Daten bereitgestellt werden ist hierbei nicht genauer spezifiziert. Präziser beschreibt Linked Open Data (LOD) den Mechanismus. Hiermit sind Daten gemeint, die durch das World Wide Web (www), über eine eindeutige Identifikation, einem Unique Resource Identifier (URI), mit dem HyperText Transfer Protocol (http) abgerufen und transportiert werden können. Mittels solchen Schnittstellen können Repositorien geöffnet werden, um so die gespeicherten Daten zur Nutzung bereitzustellen.

Schnittstellen sind im Prinzip nichts anderes als standardisierte Protokolle, die das Abfragen und Übertragen von digitalen Inhalten erlauben. Solche Schnittstellen werden heute meist als Application Programming Interface (API) bezeichnet. Im Bereich von digitalen Bildern definiert das International Image Interoperability Framework (IIIF) ein solches Interface, welches bereits heute weit verbreitet und schnell wachsend ist. Das IIIF Konsortium zählte bereits zu Beginn 40 namhafte Institutionen zu seinen Mitgliedern, darunter Namen wie die Stanford oder Harvard Universität. Das Konsortium will mit den Standarisierungsanstrengungen nicht nur digitale Repositorien öffnen, sondern den Zugriff auf die Bild-Daten so vereinheitlichen, dass dieselben Werkzeuge für Daten aus mehreren, kompatiblen Quellen eingesetzt werden können. Ein bekanntes Beispiel hierzu ist die Virtuelle Handschriftenbibliothek der Schweiz e-codices (www.e-codices.ch). Digitalisate von mittelalterlichen Handschriften, gespeichert in international verteilten Datenbanken, können in einer einheitlichen Web-Applikation (z. B. https://projectmirador.org/demo/) geladen, betrachtet und bearbeitet werden. Genau genommen definiert IIIF sogar vier standardisierte Schnittstellen mit unterschiedlicher Funktion:

- Eine Authentication API, zur Zugangskontrolle
- Die Image API, die eine Bilddatei in einem gewünschten Format "on demand" liefert
- Die Presentation API, zur Definition von komplexen mehrseitigen Dokumenten, welches in einem sogenannten Manifest, einer Meta-Datei, beschrieben wird.
- Die Search API, mit der nach Bilddateien standardisiert gesucht werden kann.

Das Konsortium um IIIF entwickelt den Standard laufend weiter. Aktuell werden vor allem audio-visuelle Inhalte in den Fokus genommen, um deren Zugriff und Darstellung zu standarisieren.
Diese Tendenz zeigt das Potential interoperablen Daten, also dem funktionalen Zusammenführen von Inhalten aus ganz unterschiedlichen Repositorien. Der Trend zeigt aber auch wichtiges Entwicklungspotential. Kann auf mehrere solche Bildquellen zugegriffen werden und werden diese nebeneinandergestellt, so wird die Vergleichbarkeit der Daten entscheidend. Vergleichbarkeit heisst in diesem Sinne, dass Bilder im selben Farbraum vorliegen, die geometrischen Attribute bekannt sind, die Auflösung gut definiert wurde und auch die Bildkomposition, also z. B. die Perspektive, identische oder zumindest vergleichbar gewählt ist. Werden fotografische Vorlagen retrodigitalisiert, so sind mehrere Standards bekannt, um die erzeugten Bilddaten in einem wohldefinierten Sinne zu erzeugen. Etablierte Standards für fotografisches Material sind z. B. die Richtlinie der Federal Agencies Digital Guidelines Initiative (FADGI) aus den USA, die europäischen Metamorfoze Richtlinien oder der Internationale Standard ISO 19264-1. Dieser Bereich der Standardisierung der Bildaufnahme wird heute z. B. in der Arbeitsgruppe "Digitalisierung von Archivgut" des VSA untersucht.

Ebenfalls wichtig im Kontext von Open Data ist die rechtliche Situation. Werden Daten offen im Netz zur Verfügung gestellt, muss sichergestellt werden, dass dies auch ohne Einschränkungen möglich ist. Oft ist diese Bestimmung allerdings langwierig und zum Teil auch kaum möglich, da die entsprechenden Eigner nicht mehr auffindbar sind oder die Erlaubnis zur Öffnung von Archivdaten nicht ohne Bedingungen gegeben wird. Aus diesem Grund ist eine detaillierte und fallspezifische Öffnung von Daten auch in Zukunft der einzig gangbare Weg. Repositorien müssen darum technisch erlauben, dass für jedes Datenobjekt spezifische Rechte vergeben werden können, mit welchen bestimmt wird wer, wann, welchen Zugang zu Daten erhält.

Entscheidend ist, dass die Bildaufnahme, die entsprechende Bildspeicherung und das zur Verfügung Stellung Hand in Hand gehen. Nur wenn bereitgestellte Daten auch inhaltlich vergleichbar sind, wird mit "Open Data" funktionale Interoperabilität gewährleistet. Es ist in diesem Sinne vorteilhaft, wenn die ganze Prozesskette des digitalen Daten-Managements integriert betrachtet wird und die Bilddatengenese gut abgestimmt ist mit der Strategie der Datennutzung und -verbreitung. Es ist allerdings so, dass die Bedürfnisse an die Bildaufnahmen zu einem grossen Grad vom Fachbereich abhängig sind. In der Archäologie werden andere Anforderungen an Bilder gestellt, als im Bereich der historischen Fotografie oder der Numismatik. Ein "community buliding" und ein Definieren von Standards ist dennoch wichtig, um Open Data über die Grenzen eines einzelnen Repositoriums hinaus sinnvoll nutzbar zu machen.

Peter Fornaro

Dr. Peter Fornaro ist Physiker und habilitierte in Digital Humanities. Er ist Privatdozent an der Universität Basel, wo er in Forschung und Lehre tätig ist.

Peter Fornaro ist im Leitungsteam des Digital Humanities Lab und des DaSCH. Er leitet zudem die AG "Digitalisierung von Archivgut" des Vereins Schweizerischer Archivarinnen und Archivare (VSA) und ist Mitglied der Eidgenössischen Kommission für Kulturgüterschutz.

Abstract

Deutsch

Mit Open Data und Interoperabilität wird die einfachere Zugänglichkeit von Daten angestrebt. Ein grosser Vorteil dieses Öffnens von Datenrepositorien ist die Möglichkeit der Nutzung einheitlicher digitaler Werkzeuge. Dies setzt allerdings auch einheitliche Daten voraus, was im Bereich von digitalen Bilddaten theoretisch möglich ist, allerdings auch das Einhalten von Bildaufnahmestandards erfordert. Im Artikel wird diese Integration dargelegt und gezeigt warum diese Verknüpfung von Genese und Nutzung von Vorteil ist.

Français

X

English

The aim of Open Data and Interoperability, is to make data easier to access. A major advantage of this opening of data repositories is the possibility of using uniform digital tools. However, this also requires uniform data, which is theoretically possible in the field of digital image data, but also requires to be compliance with image acquisition standards. The article describes this integration and shows why this combination of genesis and application is advantageous.