

Museum Analytics: Ein Online-Tool zur vergleichenden Analyse musealer Datenbestände

Schneider, Stefanie

stefanie.schneider@itg.uni-muenchen.de

Ludwig-Maximilians-Universität München, Deutschland

Kohle, Hubertus

hubertus.kohle@lmu.de

Ludwig-Maximilians-Universität München, Deutschland

Burg, Severin

severin.jo.burg@gmail.com

Università di Bologna, Italien

Küchenhoff, Helmut

kuechenhoff@stat.uni-muenchen.de

Ludwig-Maximilians-Universität München, Deutschland

Museen produzieren in den letzten Jahren verstärkt digitale Inventare ihrer Bestände, die sie zuweilen auch im Internet veröffentlichen: Manche stellen Teilbestände zur Verfügung, andere den ganzen Besitz; wie etwa das Metropolitan Museum in New York, das über 375.000 Objekte datenbankmäßig erschlossen präsentiert.¹ Üblicherweise werden einzelne Objekte aus diesen Inventaren gefiltert; entweder solche, die bekannt sind, oder andere, auf die man über erschließende Metadaten stößt. „Museum Analytics“, kurz „MAX“, ein an der Ludwig-Maximilians-Universität München am Institut für Kunstgeschichte und Institut für Statistik entwickeltes Online-Tool, sieht etwas anderes vor: Es ermöglicht es, im Unterschied zu anderen Systemen (Stack 2018), die vorhandenen Metadaten als „Massendaten“ statistisch zu erschließen, zu analysieren und zu visualisieren – ohne programmieren zu können. Damit eignet es sich besonders für die Lehre in geisteswissenschaftlichen Kontexten.²

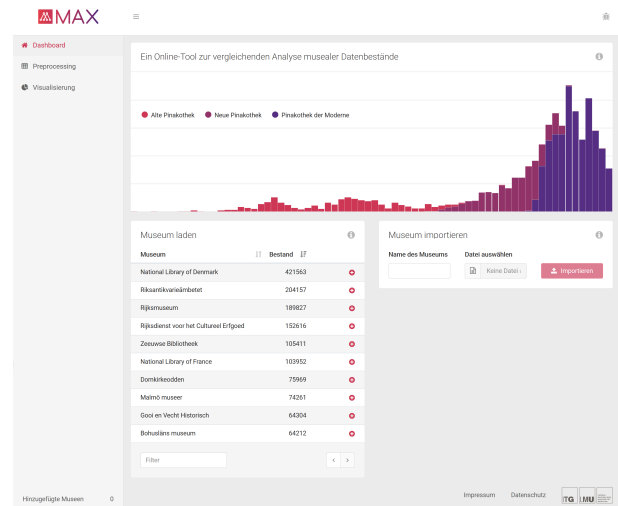


Abbildung 1: Screenshot des Moduls

Dashboard, das das Laden und Importieren von Museen ermöglicht.

„MAX“ implementiert folgende Komponenten:

1. Ein *Graphical User Interface*, das responsiv ist und ohne exzessive Einarbeitung schnelle Fortschritte und eine flüssige Analyse ermöglicht. Es ist mit der Open-Source-Programmiersprache *R* und dem auf *R* aufsetzenden Webapplikationspaket *Shiny* realisiert, die eine leichte Modularisierung im *Backend* garantieren (R Core Team, 2017; Chang et al., 2017). Hilfestellungen werden durch eine allgemeinverständliche Dokumentation und den jeweiligen Funktionen zugeordnete *Screencasts* gegeben, die einzelne Verfahrensschritte näher erläutern.
2. Eine Schnittstelle zu musealen *Application Programming Interfaces (APIs)*, um bestehende Daten möglichst einfach in das Tool importieren und weiterverarbeiten zu können. Momentan bereitgestellt werden fast 4.000.000 Objekte aus 200 Institutionen. Auch eigene, etwa per *Web Scraping* extrahierte und als CSV- oder RDS-Datei vorliegende Bestände können eingespeist und bearbeitet werden. Daher löst sich die Beschränkung auf museale Inventare: Wer sich für die Entwicklung von Bevölkerungszahlen interessiert, kann „MAX“ ebenfalls nutzen.
3. Komplementiert werden diese Komponenten mit der Javascript-Bibliothek *Highcharts* durch dynamische und interaktive Grafiken, die bei *Mouseover* weitere Kennzahlen anzeigen oder mittels *Zoom* interessierende Bereiche vergrößern oder selektieren können.³ Sie bereichern die statistische Analyse, indem sie komplexe Zusammenhänge attraktiv abbilden und die Nutzerinnen und Nutzer unmittelbar in die Analyse einbinden.

Diese Komponenten sind in drei Modulen integriert: *Dashboard*, *Preprocessing* und *Visualisierung*. *Dashboard*

dient der Auswahl und dem Import der jeweils interessierenden Museen, die in *Preprocessing* diversen statistischen Operationen unterzogen werden können, um bspw. heterogene Datierungsangaben zu vereinheitlichen. *Visualisierung* präsentiert die Ergebnisse in grafischer Form, wobei die verschiedenen Diagrammart, z. B. Histogramm oder *Bubble Chart*, automatisch je nach Datentyp der zu visualisierenden Spalte eines oder mehrerer Museen angewandt werden. Ein besonderes *Feature* ist mit der *Historie* verbunden: Einzelne Arbeitsschritte – oder ganze Arbeitsschrittketten – können mit ihr reversibel gemacht oder auf andere zu bearbeitende Museen übertragen werden. Es wurden sowohl Standardmethoden der Statistik implementiert als auch für die geisteswissenschaftliche Disziplin spezifische Schnittstellen, um bspw. die Entstehungsländer der Kunstwerke einer Sammlung in einer Karte visualisieren zu können.

Je nach Erschließungstiefe der Bestände kann untersucht werden, welche Sammlungskonjunkturen es in bestimmten Museen gegeben hat, welche Gattungen zu welcher Zeit besonders beliebt waren oder ob es Zusammenhänge zwischen der Sammlungstätigkeit und gesellschaftspolitischen Bedingungen gab. Bei hinreichend repräsentativer Datenlage lassen sich auch Fragen adressieren, die die Kunstgeschichte als solche betreffen: Welche künstlerischen Techniken waren im historischen Wandel vorherrschend? Welche Themen zu welcher Zeit beliebt? Ist die Säkularisierung an der Entwicklung der Ikonographie der Kunstwerke abzulesen? Die Werke bzw. deren Digitalisate können auch direkt adressiert und statistisch analysiert werden: Ist die Verwendung von bestimmten Farben bzw. Farbzusammenstellungen historisch beschreibbar? Lässt sich Heinrich Wölfflins Bildanalyseinstrumentarium, in dem etwa das Renaissance-Kunstwerk als linear und klar, das barocke als malerisch und unklar beschrieben wurde, automatisieren und auf seine historische Berechtigung befragen?

Die Funktionalitäten des Tools wurden in einem Seminar mit geisteswissenschaftlichen Studierenden im Sommersemester 2018 an der Ludwig-Maximilians-Universität überprüft und evaluiert, worauf mit der Integration der genannten *Screencasts* reagiert wurde. „MAX“ ist ein Teil des „Digital Humanities Virtual Laboratory“ („DHVLab“), einer modularen Lehr- und Forschungsinfrastruktur zur Ausbildung von Studierenden der Kunst-, Geschichts- und Sprachwissenschaften in Anwendungen und Methoden der *Digital Humanities*.⁴

Rahmen des Programms Lehre@LMU zur Stärkung der Forschungsorientierung in der Lehre gefördert.

3. <https://www.highcharts.com/> (23.09.2018).

4. <https://dhvlab.gwi.uni-muenchen.de/> (24.09.2018); weitere Einblicke in Klinke (2018).

Bibliographie

Chang, Winston / Cheng, Joe / Allaire, JJ / Xie, Yihui / McPherson, Jonathan (2017): *shiny: Web Application Framework for R*, <https://cran.r-project.org/package=shiny>.

Klinke, Harald (Hrsg., 2018): *#DigiCampus. Digitale Forschung und Lehre in den Geisteswissenschaften*. München: Universitätsbibliothek der Ludwig-Maximilians-Universität.

R Core Team (2017): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, <https://www.r-project.org/>.

Ridge, Mia (2012): *Mia Ridge Explores the Shape of Cooper-Hewitt Collections*, <https://labs.cooperhewitt.org/2012/exploring-shape-collections-draft/>.

Stack, John (2018): *Exploring Museum Collections Online. Some Background Reading*, <https://lab.sciencemuseum.org.uk/exploring-museum-collections-online-some-background-reading-da5a332fa2f8>.

Fußnoten

1. <https://www.metmuseum.org/art/collection> (23.09.2018).

2. <https://www.max.gwi.uni-muenchen.de/>, das Online-Tool selbst unter <https://dhvlab.gwi.uni-muenchen.de/max/> (beide 23.09.2018). Es wird im