Archivschule Marburg
Hochschule für Archivwissenschaft

Checkliste für die Durchführung von Crowdsourcing-Projekten

Eine Handreichung des DFG-Produktivpiloten

Digitalisierung von archivalischen Quellen

2013–2015

Projektpartner:

Landesarchiv Nordrhein-Westfalen

LWL - Archivamt für Westfalen

LANDESARCHIV BADEN-WÜRTTEMBERG

SÄCHSISCHES STAATSARCHIV | Freistaat SACHSEN

STADTARCHIV MANNHEIM
INSTITUT FÜR STADTGESCHICHTE

Ansprechpartner für diese Handreichung:

Dr. Clemens Rehm, Landesarchiv Baden-Württemberg

clemens.rehm@la-bw.de / Tel: +49 (0)711 212 4288

Bei der Planung und Durchführung von Crowdsourcing-Projekten sollten – abhängig von den jeweiligen Rahmenbedingungen des durchführenden Archivs – folgende Punkte bedacht werden:

1. Idealerweise sollten die Crowdsourcing-Aktivitäten Teil einer umfassenden Web 2.0- oder **Crowdsourcing-Strategie** sein oder zu einer solchen hinführen.

2. Die **Projektbetreuung** muss geregelt sein; personelle Ressourcen müssen zur Verfügung stehen. Crowdsourcing bindet in jedem Fall personelle und ggf. auch finanzielle Ressourcen, die in ihrem Umfang nicht zu unterschätzen sind. Archivische Crowdsourcing-Projekte sind insbesondere dann erfolgreich, wenn gezielt und ausreichend **Ressourcen** für ihre Konzeption und Durchführung zur Verfügung gestellt werden. Personelle oder finanzielle Ressourcen werden z.B. gebraucht für die Konzeption, Koordinierung und Betreuung des Projekts, die technische Umsetzung (z.B. Programmierung und Pflege einer Plattform oder Datentransfer), die Erstellung von Editionsrichtlinien oder anderen Anweisungen für die Teilnehmer, sowie Öffentlichkeitsarbeit. Crowdsourcing gleichsam „nebenbei“ zu betreiben, ist nicht erfolgversprechend!

3. Die **Projektziele** sollten klar definiert sein. Denkbare Ziele können beispielsweise sein
    - Verbesserung der Erschließung des (Teil)Bestandes
    - Verbesserung der Zugänglichkeit des (Teil)Bestandes
    - Verbesserung der Wahrnehmung des (Teil)Bestandes und/oder des Archivs
    - Ansprechen neuer Nutzergruppen

4. Es sollten konkrete **Bestände oder Teilbestände** ausgewählt werden. Alternativ können auch geeignete Einzelstücke aus mehreren Beständen in einem Projekt zusammengefasst werden, wenn sich dies z.B. thematisch anbietet.

5. **Rechtsfragen** im Zusammenhang mit den ausgewählten (Teil)Beständen (z.B. Urheberrecht, Datenschutz) und den zu generierenden Inhalten (z.B. Forenhaftung) müssen vor Projektbeginn geklärt sein.

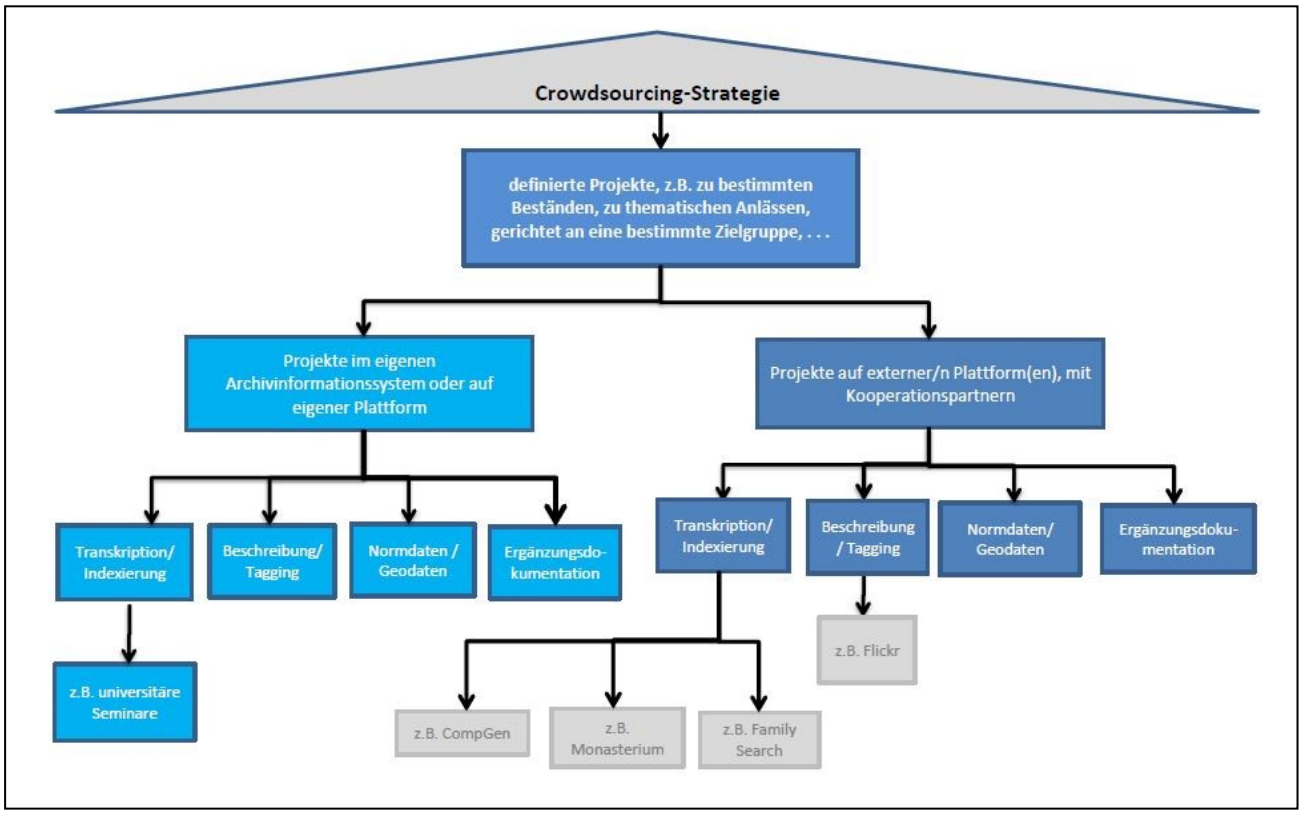
6. Der zu generierende **Inhalt** sollte definiert sein, z.B. als
    - Indices
    - Transkriptionen
    - Normdaten
    - Freitextkommentare
    - Schlagworte/Tags
    - Geodaten
    - Unterlagen zum Aufbau einer Ergänzungsdokumentation

Die verschiedenen Arten des zu generierenden Inhalts stellen unterschiedliche Anforderungen an die Eingabeoberfläche.

7. Die **Zielgruppe** – also die Crowd – sollte definiert sein, beispielsweise als

   - geschlossene Gruppe (Heimatverein, Studierende, Firmenangehörige, Wikipedianer, etc.)
   - offene Gruppe (z.B. Adressierung aller Interessierten über eine offene Plattform, z.B. Flickr)

8. Es sollte eine bewusste Entscheidung für eine **Plattform** getroffen werden. Je nach Art des Projekts ist zu unterscheiden zwischen Projekten, die innerhalb des eigenen Archivinformationssystems oder auf einer vom Archiv selbst betriebenen Plattform durchgeführt werden und Projekten, die zusammen mit einem Kooperationspartner auf dessen externer und ggf. kommerzieller Plattform angesiedelt sind (vgl. Grafik).



9. Wird eine externe Plattform gewählt, so muss der **Transfer der generierten Daten** ins eigene Archivinformationssystem technisch sichergestellt werden. Denn der Mehrwert nutzergenerierter Daten kommt erst dann richtig zum Tragen, wenn diese in die elektronischen Findmittel integriert sind und somit bei einer Suchanfrage einbezogen werden können.

10. Es sollten **Qualitätsstandards** definiert werden; dies gilt insbesondere dann, wenn die nutzergenerierten Daten ins eigene Archivinformationssystem übernommen werden sollen, ggf. sind redaktionelle Bearbeitungsschritte einzuplanen (vgl. Nr. 11-13).

11. Der Umfang der **qualitätssichernden Maßnahmen** muss projektbezogen definiert und die Umsetzung geregelt werden. Es gibt verschiedene Vorgehensweisen, die von einer engmaschigen und aufwändi-

gen Prüfung der nutzergenerierten Inhalte bis zum bewussten gänzlichen Verzicht auf eine redaktionelle Kontrolle und Bearbeitung reichen können. Generell gilt, dass nutzergenerierte Inhalte als solche gekennzeichnet sein und für die Nutzer eindeutig von den archivarischen Erschließungsinformationen unterscheidbar sein sollten. So könnte bei einem Projekt mit dem Ziel der Bildbeschreibung über Freitextkommentarfelder beispielsweise im Sinne der Ressourcenschonung gänzlich auf eine Überprüfung der Nutzerkommentare auf sachliche Richtigkeit verzichtet werden, solange diese eindeutig als nutzergeneriert gekennzeichnet bleiben und nicht in die archivische Titelaufnahme integriert werden. Auch Art und Umfang weiterer qualitätssichernder Maßnahmen sind immer in Relation zu den Zielen des jeweiligen Projekts und zu dem für dieses Projekt als vertretbar erachteten Ressourcenaufwand zu sehen. Grundsätzlich gibt es verschiedene Möglichkeiten, die Qualität der nutzergenerierten Inhalte positiv zu beeinflussen. Hier müssen wenige Beispiele genügen: Bei Transkriptionsprojekten empfiehlt es sich jedoch, **Editions- oder Transkriptionsrichtlinien** zu erstellen. Darüber hinaus könnte eine Qualitätssicherung durch das sogenannte **„Double-Keying“** – also die Erfassung von Texten durch zwei voneinander unabhängige Teilnehmer – erfolgen. Es kann sinnvoll und arbeitsersparend sein, Antworten auf die häufigsten Fragen der Teilnehmer in **FAQs**[1] zur Verfügung zu stellen. Auch ein **Forum**, in dem sich die Teilnehmer bei Problemen an die Projetbetreuer wenden aber sich auch untereinander austauschen können, kann bei größeren Projekten zur Qualitätssteigerung beitragen.

12. Es ist zu entscheiden, ob sich die Projektteilnehmer registrieren müssen, es Ihnen als eine Option angeboten wird, oder ob auf eine **Registrierung** bewusst verzichtet wird. Je nach Art des Projekts kann eine Registrierung Anreize zur wiederholten Mitarbeit bieten (vgl. auch Nr. 14). Andererseits kann sie aber auch die Hemmschwelle für eine Teilnahme erhöhen, da spontane Beiträge nicht möglich sind. Eine Registrierung kann auch eine qualitätssichernde Maßnahme sein (vgl. Nr. 11).

13. Entsprechend den angesetzten Qualitätskriterien und in Abhängigkeit vom Ablauf des ggf. zu leistenden Datentransfers und der zur Verfügung stehenden Ressourcen ist ein **Redaktionsworkflow** zu definieren. Dabei ist zwischen der Veränderung (z.B. Modifikation des Titels, Berichtigungen, Anlage von Enthält-Vermerken etc.) und der Ergänzung schon vorhandener archivischer Erschließungsinformationen (z.B. Hinzufügen von Normdaten) zu unterscheiden.

14. Es sollte darüber nachgedacht werden, **Anreize** für eine (längerfristige) Teilnahme zu schaffen. Es kann z.B. sinnvoll sein, die Teilnehmer durch die **Vergabe von Punkten** zu einer wiederholten Teilnahme zu animieren.

15. Der Erfolg eines Crowdsourcing-Projekts hängt auch von der **Öffentlichkeitsarbeit** ab. Es sollte daher ein Werbekonzept für das Projekt geben. Möglicherweise bietet sich als Aufhänger ein aktuelles Thema oder ein Jubiläum etc. an. Je nach Bestand lassen sich eventuell auch geeignete Multiplikatoren identifizieren (z.B. genealogische Vereine).

[1] Frequently Asked Questions.

Diese Checkliste stellt überblicksartig die Kriterien dar, die erfüllt sein müssen, um Crowdsourcing-Projekte erfolgreich durchzuführen:

1	**Strategie** (Das Projekt sollte Teil einer umfassenden Web 2.0- oder Crowdsourcing-Strategie sein oder dort hinführen)
2	**Projektbetreuung** geregelt und **Ressourcen** in ausreichendem Maße vorhanden
3	**Projektziele** definiert
4	**Beständeauswahl** getroffen
5	**Rechtsfragen** geklärt (z.B. zu Urheberrecht, Datenschutz, Lizenzfragen)
6	**Zu generierender Inhalt** definiert
7	**Zielgruppe (Crowd)** definiert
8	**Plattform ausgewählt** • (archiv)interne Plattform (z.B. beschränkter Zugriff der Erfasser auf das Erschließungs- oder Findmittelsystem) • externe Plattform (kommerziell; Plattform eines Kooperationspartners, etc.)
9	**Datentransfer** sichergestellt (z.B. von einer externen Plattform ins eigene Archivinformations-system)
10	**Qualitätsstandards** definiert
11	Umfang der **Qualitätssicherung** festgelegt
12	**Authentifizierung/Registrierung** der Projektteilnehmer geregelt
13	**Redaktionsworkflow** definiert
14	Entscheidung, ob es **Anreize** für die freiwilligen Helfer geben soll, getroffen
15	Konzept für **Öffentlichkeitsarbeit** erarbeitet