

Dauer der Arbeit: 4 Monate  
Abschluss: Oktober 2014  
Betreuer: Dr.-Ing. Volker Walter

## Entwicklung eines Webinterfaces zur Crowd-basierten Erfassung von Geodaten aus Luftbildern

DOMINIK LAUPHEIMER, LAUPHEIM, GERMANY

**Keywords:** Crowdsourcing, Geodaten, Webinterface

**Zusammenfassung:** Ziel dieser Bachelor-Thesis ist die Entwicklung eines webbasierten Programmes zur Crowd-basierten Erfassung von raumbezogenen Daten aus Luftbildern. Konkret bedeutet dies, dass eine Webseite entwickelt wird, die die Digitalisierung von Luftbildern durch Laien ermöglicht. Im Gegensatz zu den existierenden Crowdsourcing-Ansätzen der Geodäsie und Geoinformatik (siehe Kapitel 1 Einführung) handelt es sich hierbei um bezahltes Crowdsourcing: Die Benutzer werden monetär entlohnt. Per Mausclick können Straßen, Gebäude und Wälder erfasst werden. Des Weiteren wird eine Kontrollwebseite entwickelt, die der Überprüfung und Kategorisierung der Digitalisierungen dient. Die Kategorisierung erfolgt nach dem Schulnoten-System. Die Digitalisierungsseite wird der Crowd über Microworkers (<https://microworkers.com>), eine 600.000 Mitglieder starke Crowdsourcing-Plattform, zugänglich gemacht. Prinzipiell ist auch eine autonome Nutzung des Webprogrammes denkbar, jedoch ermöglicht Microworkers viele Einstellmöglichkeiten (Herkunftsland, Geschwindigkeit des Projektes, Belohnung/Aufgabenstellung, ...) und wickelt die Geldtransfers ab. Es wird untersucht, inwieweit diese Einstellungen Einfluss auf die Qualität der Ergebnisse nehmen.

**Summary:** *Development of a web interface for crowd-based acquisition of geodata from aerial images* - The aim of this bachelor thesis is the development of a web-based programme for the crowd-based acquisition of spatial data from aerial images. The developed website allows the digitalization of aerial images by non-experts. Compared to already existing approaches of crowdsourcing within Geoinformatics this project realizes paid crowdsourcing. This means, that users of this website will be paid. By mouse click one can measure streets, buildings and forests. For control purposes a second website will be developed. On the control site the submitted digitalisations can be checked and categorized. The categorization is based on the grading system used in German schools. The digitalisation's site will be made available more widely by the crowdsourcing platform Microworkers (<https://microworkers.com>), which has approximately 600.000 members. In principle, the developed websites can be operated autonomously. However, Microworkers enables many setting possibilities (targeting countries, speed of project, payment/work, ...) and bears the money transfer. It will be investigated, to which extent these settings have an influence on the quality of the results.

---

### 1 Einführung

Der von Jeff Howe geprägte Neologismus Crowdsourcing (Crowd + Outsourcing) beschreibt eine stetig populärer werdende Methode zur Problemlösung (HOWE 2006). Die Idee, die sich hinter Crowdsourcing verbirgt, ist bereits Jahrtausende alt, doch erst seit dem Zeitalter des interaktiven Webs 2.0 entfaltet Crowdsourcing sein vollständiges Potential (GASSMANN 2013). Noch nie in der Geschichte der Menschheit ist es so einfach gewesen, eine Problemstellung auf schnellem Wege vielen Lösungswilligen zu unterbreiten. Anders als beim traditionellen Outsourcing wird beim Crowdsourcing der Problemlöser nicht vom Problemsteller nach bestimmten Kriterien ausgewählt. Stattdessen

wird die Problemstellung einer anonymen Masse (Crowd) übermittelt, die sich der Problemstellung annimmt. Bei den Lösenden handelt es sich in den seltensten Fällen um Experten. Aufgrund des Interesses des Löser an der Thematik können dennoch qualitativ hochwertige Lösungen erzielt werden. Besonders für Aufgabenstellungen, die sich durch Nicht-Innovativität, Nicht-Automatisierbarkeit sowie schnelle Bearbeitung charakterisieren lassen, findet Crowdsourcing Anwendung. Die bekannteste Umsetzung des Crowdsourcing-Prinzips ist die Online-Enzyklopädie Wikipedia. Auch im Bereich der Geodäsie und Geoinformatik existieren bereits Crowdsourcing-Anwendungen. Die bekanntesten sind OpenStreetMap und Wheelmap.

## 2 Realisierung des Webprogrammes

In diesem Projekt wurden folgende Technologien eingesetzt:

- HTML → Struktur digitaler Inhalt
- CSS → Design (Farbe, Schriftart, Alignment, etc.) digitaler Inhalte
- JavaScript → Funktionalität; clientseitige Veränderung digitaler Inhalte
- PHP → Funktionalität; serverseitige Veränderung digitaler Inhalte

Mit HTML und CSS können statische Webseiten erzeugt werden. Wird auf einer statischen Webseite JavaScript und/oder PHP integriert, so entsteht eine interaktive und dynamische Webseite. Für eine clientseitige Veränderung der Webseite spielen *Event Listener*, die die Handlungen des Users "abhören", eine große Rolle. Wird eine bestimmte Aktion vom User getätigt, so greift der Event Listener und führt die zum Event gehörenden Befehle aus.

Die Client-Server-Kommunikation erfolgt über das Hypertext Transfer Protocol (http). AJAX (Asynchronous JavaScript and XML) ermöglicht, dass http-Requests an den Server gesendet werden können, während die Webseite angezeigt wird. Mittels dieser Technologie können einzelne Seitenelemente verändert werden, ohne dass ein komplettes Neuladen der Seite notwendig wäre. Die Asynchronität besteht darin, dass das Senden und Empfangen von Daten nicht zeitgleich stattfindet. Das Warten auf die Daten blockiert somit den Ablauf des Prozesses nicht. Beispielsweise bleibt die Webseite (mit Ausnahme des vorgesehen Platzes für das Luftbild) unverändert, während der Client per AJAX ein Luftbild vom Server anfordert. Damit wird der Datenverkehr zwischen Server und Client minimal gehalten.

## 3 Webseiten-Aufbau

Im Rahmen dieser Bachelor-Thesis wurden zwei verschiedene Hauptseiten entwickelt (siehe Abbildung 1):

### 3.1 Aufbau der Digitalisierungsseite

Der Benutzer kann auf der Digitalisierungsseite (<http://lauphedo.bplaced.net/Digi/index.htm>, siehe Abbildung 2) zwischen zwei Programm-Modi



Abb. 1: Funktionen der Webseiten

auswählen: Digitalisieren und Löschen. In beiden Programm-Modi gibt es je drei Optionen. Mit den Buttons *Streets*, *Buildings* und *Forests* kann zwischen den drei verschiedenen Objekt-Typen gewählt werden. Folglich kann nach der Auswahl per Button-Klick eine der drei Objektarten digitalisiert oder gelöscht werden - je nachdem, welcher Programm-Modus aktiv ist. Der aktive Modus und Objekttyp wird durch eine farbliche Kennzeichnung des entsprechenden Buttons sichtbar gemacht. Straßen werden als Linien-, Gebäude als Punkt- und Wälder als Flächenobjekte erfasst.

Zentrales Element ist der mit einer Luftbild-Kachel gefüllte Canvas. Der schwarze Rahmen markiert die innere Grenze des Überlappungsbereiches. Sinn und Zweck dieses Überlappungsbereiches ist die Erzeugung von Verknüpfungspunkten zwischen den einzelnen Luftbild-Kacheln. Auf diese Weise können abgeschlossene Digitalisierungen miteinander verknüpft werden. Der Überlappungsbereich ist besonders für Objekte, die nicht vollständig innerhalb einer Luftbildkachel liegen, von Bedeutung. Mindestens ein Digitalisierungspunkt von derartig abgeschnittenen Objekten muss innerhalb des Überlappungsbereiches liegen.

Unter den Lösch-Buttons befinden sich dynamische Auswahlmensüs, die alle bis dahin digitalisierten Objekte des entsprechenden Objekt-Typs auflisten. Per Mausklick auf ein Objekt der Liste wird das Objekt im Bild farblich markiert. Mit einem Doppelklick wird das ausgewählte Objekt gelöscht. Das darunter gelegene Textfeld bietet dem Benutzer die Möglichkeit, ein Feedback oder sonstige Anmerkungen zu geben. Mit dem Button *Submit job* wird die Bearbeitung der Aufgabenstellung abgeschlossen und die digitalisierten Daten werden an den Server übermittelt. Falls sich der User dazu entscheiden sollte, den Job nicht zu bearbeiten, wird ihm diese Möglichkeit mit dem Button *Reject job* explizit ermöglicht. Es kann auch lediglich das Fenster bzw. der Browser-Tab geschlossen werden. In der Navigationsleiste kann der User Informationen über die Seite (*About this website*) oder bei eventuellen Unklarheiten Hilfestellung (*Help*) erhalten.

### 3.2 Aufbau der Kontrollseite

Das dynamische Auswahlménú auf der linken Seite der Kontrollseite (siehe Abbildung 3) listet alle digitalisierten Luftbild-Kacheln auf. Per Mausklick kann ein Luftbild ausgewählt werden, das dann zweifach angezeigt wird: mit und ohne Digitalisierungen. Je nach Qualität der Erfassung wird die Digitalisierung nach dem Schulnoten-System (sehr gut, gut, befriedigend, ausreichend, mangelhaft) mit den Buttons auf der rechten Seite bewertet. Als mangelhaft bewertete Luftbilder werden auf der Digitalisierungsseite erneut zur Digitalisierung angeboten.

Auf der Kontrollseite werden der Kommentar des Benutzers und ein Proof-Code, der nach Betätigen des Buttons *Submit job* generiert und dem User zugänglich gemacht wird, angezeigt. Der Proof-Code dient der eindeutigen Zuordnung von eingereicherter Digitalisierung und Benutzer auf Microworkers, um die Benutzer entlohnen zu können.

## 4 Ergebnisse

Die Auswirkungen des Parameters „Bezahlung/Aufgabenstellung“ auf die Qualität der Ergebnisse wurden untersucht (siehe Abbildung 4). Die wenigen Tests liefern noch keine 100 prozentig aussagekräftigen Ergebnisse. Dennoch lassen sich durch die Analyse der Kampagnen mit unterschiedlichen Löhnen Tendenzen feststellen.

Grundsätzlich kann durch einen erhöhten Lohn das Qualitätsniveau angehoben werden. Bei den Kampagnen 3 und 4 werden mehr Digitalisierungen als sehr gut kategorisiert als bei den ersten beiden Kampagnen. Der Relativanteil der akzeptierten Digitalisierungen steigt mit zunehmendem Lohn.

Ein erhöhter Lohn bedeutet nicht notwendigerweise, dass wenige Digitalisierungen als mangelhaft kategorisiert werden. In Kampagne 4 mussten im Vergleich zu Kampagne 3 wesentlich mehr Digitalisierungen abgelehnt werden. Die Aussicht auf gute Bezahlung lockt viele Benutzer an – ungeachtet deren Eignung oder Engagements.

Die vier Kampagnen zeigen, dass ein erhöhter Lohn die Kampagnen-Dauer verkürzt. Auch bei höherem Lohn werden zwar viele Digitalisierungen abgewiesen, jedoch werden die als mangelhaft kategorisierten Digitalisierungen innerhalb kürzester Zeit von der Crowd erneut digitalisiert.

Der erhöhte Lohn führt dazu, dass weniger Bilder ohne jegliche Digitalisierung eingereicht werden. In Kampagne 3 und 4 sind die als mangelhaft kategorisierten Digitalisierungen qualitativ geringwertig. Nur wenige eingereichte Aufgabenstellungen sind komplett unbearbeitet.

Die verschiedenen Kampagnen haben aufgezeigt, dass die Benutzer vorwiegend aus dem indischen Subkontinent und Osteuropa stammen (siehe Abbildung 10).

### 4.1 Eingereichte Erfassungen

Alle erfassten Kacheln wurden nach folgendem Bewertungsschema in fünf Klassen eingeteilt:

#### *Sehr gut*

Perfekte Erfassungen ohne erkennbare Fehler werden als sehr gut kategorisiert (Beispiel siehe Abbildung 5).

#### *Gut*

Wenn fast alle Objekte richtig erfasst sind, wird die Digitalisierung als gut kategorisiert. Bei einer guten Digitalisierung gibt es wenige Objekte, die fehlen oder zu viel sind. Bei der Digitalisierung bestehen aufgrund der Qualität des Luftbildes Unsicherheiten (Beispiel siehe Abbildung 6).

#### *Befriedigend*

Bei einer als befriedigend kategorisierten Digitalisierung fehlen offensichtlich ein paar Objekte. Im Großen und Ganzen ist die Erfassung jedoch in Ordnung (Beispiel siehe Abbildung 7).

#### *Ausreichend*

Sind bereits auf den ersten Blick Fehler in der Digitalisierung zu erkennen, wird die Digitalisierung als ausreichend eingestuft - vorausgesetzt, es handelt sich um eine erkennbar sinnvolle Erfassung (Beispiel siehe Abbildung 8).

#### *Mangelhaft*

Wurden keine Digitalisierungen vorgenommen, Daten unsinnig oder mehr als 50% falsch erfasst, so wird die Digitalisierung als mangelhaft kategorisiert. Die eingereichte Digitalisierung wird abgelehnt und auf der Digitalisierungsseite erneut angeboten (Beispiel siehe Abbildung 9).

# Digitalization of aerial images

Help  
About this  
website



Luftbild 'mosaic\_rgb\_1m\_(1.6).jpg'

Digitize forests by clicking several points of the edge of wood on the aerial image. Double-clicking will finish and fill a forest polygon automatically.

Digitize objects

Delete objects

Streets	Buildings	Forests
Street 1	Building 1	Forest 1
Street 2	Building 2	
Street 3	Building 3	
Street 4	Building 4	
Street 5	Building 5	
Street 6	Building 6	
Street 7	Building 7	
	Building 8	

Your feedback

Here's some space for your feedback...

**Abb. 2:** Digitalisierungsseite (inklusive Beispielerfassung)

## Control Site

**ifp**  Universität Stuttgart

Choose which aerial image you'd like to check...

Please select an image

[1mosaic\\_rgb\\_1m\\_\(1.6\).jpg](#)



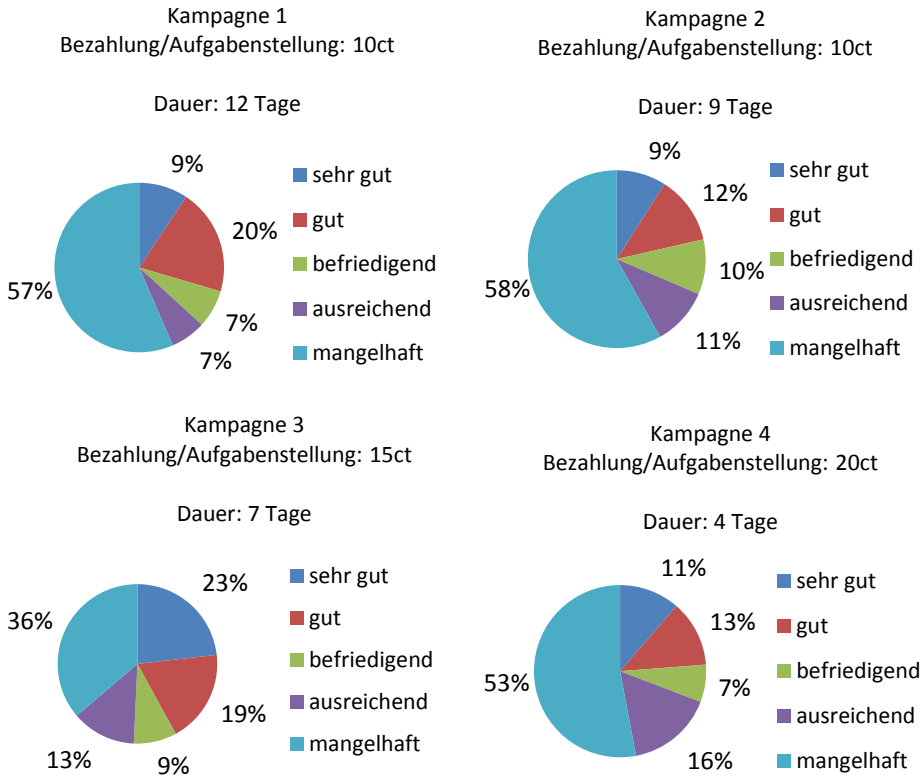
1mosaic\_rgb\_1m\_(1.6).jpg

Commentar des User:  
kein Kommentar abgegeben

Proof-Code:  
189p6e44



**Abb. 3:** Kontrollseite (inklusive Beispielerfassung)



**Abb. 4:** Kategorisierte Digitalisierungen (Als mangelhaft kategorisierte Digitalisierungen werden erneut auf der Digitalisierungsseite angeboten und können demnach mehrmals als mangelhaft kategorisiert werden)



**Abb. 5:** Als sehr gut kategorisiertes Luftbild



**Abb. 6:** Als gut kategorisiertes Luftbild



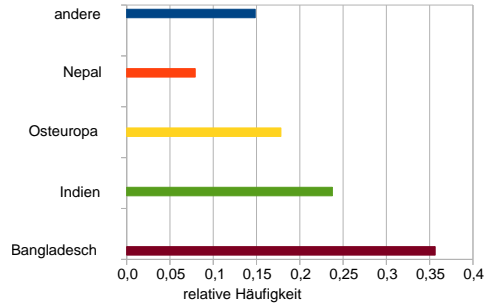
**Abb. 7:** Als befriedigend kategorisiertes Luftbild



**Abb. 8:** Als ausreichend kategorisiertes Luftbild



**Abb. 9:** Als mangelhaft kategorisiertes Luftbild



**Abb. 10:** Herkunftsländer der Benutzer

## 5 Ausblick

In der Bachelor-Thesis wurde aufgezeigt, dass sich die Erfassung von Geodaten aus Luftbildern mit Crowdsourcing realisieren lässt. Teilweise wurden dabei sehr gute Ergebnisse erzielt. Allerdings gibt es eine Vielzahl an Erfassungen, die schlecht oder gar unsinnig sind. Die erzielten Ergebnisse sind hochgradig heterogen. Zukünftige Tests auf Microworkers sollen im Sinne der Zuverlässigkeitssteigerung die bislang gewonnenen Erkenntnisse festigen. Außerdem sollen Tests mit verschiedenen Rahmenbedingungen durchgeführt und deren Einfluss auf die Qualität überprüft werden. Im nächsten Schritt sollen die Ergebnisse homogenisiert werden. Vorangeschaltete Tests, bei denen geeignete Arbeiter ermittelt werden, sind an dieser Stelle denkbar. Darüber hinaus soll der Auswertungsprozess automatisiert werden. Dies kann ebenso mit Crowdsourcing realisiert werden: die Crowd überprüft die Arbeit der Crowd.

## Literatur

HOWE, J., 2006: The Rise of Crowdsourcing. WIRED Magazin, verfügbar online auf: <http://archive.wired.com/wired/achive/14.06/crowds.htm> (aufgerufen am 10. Oktober 2014).

GASSMANN, O., 2013: Crowdsourcing - Innovationsmanagement mit Schwarmintelligenz. 2. Auflage, Carl Hanser Verlag GmbH & Co. KG, Regensburg, 212S.