

# Crowdsourcing zur Georeferenzierung alter topographischer Karten – Ansatz, Erfahrungen und Qualitätsanalyse

Ralf Bill und Kai Walter

## Zusammenfassung

Umfassende topographisch-geographische Aufnahmen durch Vermessung und maßstäbliche Darstellung eines Landesgebiets in (amtlichen) Kartenwerken verschiedener Maßstäbe entstanden ab Ende des 18. Jahrhunderts als Vorstufe der heutigen topographischen Kartenwerke. Deren Aufbereitung und Georeferenzierung hat für raum-zeitlich arbeitende Disziplinen eine große Bedeutung, um z. B. den Landschaftswandel über lange Zeiträume zu studieren. Im DFG-Projekt »Virtuelles Kartenforum 2.0« werden rund 6.000 gescannte Messtischblätter des Deutschen Reiches im Maßstab 1:25.000 georeferenziert und über WebMapService (WMS) bereitgestellt. Für den Schritt der Georeferenzierung wird in diesem Beitrag ein Crowdsourcing-Ansatz vorgestellt und die erreichten Ergebnisse werden hinsichtlich der Datenqualität untersucht.

## Summary

*Comprehensive topographic and geographic recordings of countries by surveying and by cartographic representation in (official) map series at different scales emerged in the late 18th century as precursors of the nowadays used topographic maps. For spatio-temporal working disciplines the scanning and georeferencing of older topographic maps is of great importance to study landscape changes over long periods of time. In the DFG project »Virtuelles Kartenforum 2.0« more than 6,000 scanned topographic maps at a scale of 1:25,000 in the borders of the former Deutsches Reich are georeferenced and disseminated as WebMapServices (WMS). For the step of georeferencing the huge amount of maps, a crowdsourcing approach is presented in this paper. The results achieved are studied in terms of data quality.*

**Schlüsselwörter:** Crowdsourcing, Georeferenzierung, topographische Karten, Qualitätsanalysen

## 1 Motivation

### 1.1 Altkarten

Alte Karten üben seit eh und je eine große Faszination auf den Betrachter aus. Sie zeigen, wie eine Landschaft einmal aussah, sind in frühen Zeiten oftmals künstlerisch geprägt und sehr ästhetisch anzusehen. Zumeist sind sie aber in Bibliotheken und Museen nur aus der Ferne zu betrachten oder in Archiven konserviert. Für eine wissenschaftliche oder praktische Nutzung, um z.B. Landschaftsveränderungen über die Zeit auswerten zu können, müssen die Altkarten digital aufbereitet werden, wozu einerseits das Scannen zur Analog-Digital-Wandlung und andererseits eine Georeferenzierung des dann als Rasterbild vorliegenden Kartenbildes gehört. Gemeinsam mit neueren Daten kann die Karte danach in einem Geo-Informationssystem (GIS) verarbeitet werden. Der Zugriff auf Altkarten in digitaler georeferenzierter Form stellt eine wichtige Basis für kollaboratives Arbeiten in der heutigen datengetriebenen Wissenschaft (e-Science) dar (Bill 2014).

Der Begriff »Altkarten« umfasst im vorliegenden Beitrag topographische oder thematische Karten seit Beginn der ersten topographischen Landesaufnahmen mittels Vermessung und maßstäblicher Darstellung gesamter Landesgebiete in (amtlichen) Kartenwerken. Diese entstanden ab Ende des 18. Jahrhunderts. Die auf dem Gebiet Deutschlands in den Nationalstaaten durchgeführten Aufnahmen wurden mit der Gründung des Deutschen Reiches im Laufe der Zeit in ein mittel- bis kleinmaßstäbliches amtliches topographisches Landeskartenwerk überführt (Maßstäbe 1:5.000 bis 1:1.000.000), welches auch noch im digitalen Zeitalter der amtlichen Geoinformationen – dann als Digitale Topographische Karte – Bestand hat und fortgeführt wird.

### 1.2 Kartenforum

Neben Bibliotheken und Archiven verfügen auch nationale Vermessungsbehörden über umfangreiche Sammlungen alter Karten. Maßstäbe im internationalen Vergleich hinsichtlich der digitalen Verfügbarmachung setzen insbesondere die privatbetriebenen Portale »David Rumsey Map Collection« (DR 2014) und »Old Maps Online« (OMO 2014), die thematische Suchen und räumlich-zeitliche Recherchen nach Karten ermöglichen. Möglichkeiten zur Online-Recherche nach Altkarten haben sich im letzten Jahrzehnt fest in Deutschland etabliert. Hierbei sind insbesondere die sammlungsübergreifenden Datenbanken IKAR (IKAR 2014) und GEO-LEO (GEO-LEO 2014) zu nennen.

Eine der größten und wichtigsten Sammlungen von alten Karten in Deutschland befindet sich mit ca. 177.000 Einzelblättern im Besitz der Sächsischen Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB).

Um den Zugriff auf diese umfangreiche Kartensammlung zu verbessern und gleichzeitig die wertvollen Bestände zu schonen, entwickelte die SLUB seit 2005 das Kartenforum ([www.deutschefotothek.de/cms/kartenforum.xml](http://www.deutschefotothek.de/cms/kartenforum.xml)). Laut SLUB (2014) ist »... das Kartenforum ... ein von der Deutschen Fotothek betreutes und von der Deutschen Forschungsgemeinschaft (DFG) gefördertes Informationsportal von Bibliotheken, Museen und Archiven. Gegenwärtig werden rund 23.500 der wichtigsten, hochauflösend digitalisierten kartographischen Quellen – insbesondere zur Geschichte und Landeskunde Sachsens – aus den Sammlungen der beteiligten Partner angeboten. Im Rahmen des DFG-Projekts »Exemplarische Digitalisierung und Erschließung historischer Karten und Ansichten (2009–2011)« wurden u. a. rund 6.000 Messtischblätter im Maßstab 1:25.000 sowie die Karten des Deutschen Reiches im Maßstab 1:100.000 im Sinne eines flächendeckenden kartographischen Basisangebots für die wissenschaftliche Nutzung aus allen Disziplinen bereitgestellt.«

Das Kartenforum der SLUB ermöglicht mit Hilfe einer interaktiven Suchmaske den Zugriff auf einen schnell wachsenden Bestand digitalisierter Karten und Ansichten. Dazu gehört der gescannte Bestand topographischer Karten, u. a. flächendeckend für das Deutsche Reich ab 1868 die Karten 1:100.000 (674 Blatt) und die Messtischblätter 1:25.000 (über 6.000 Blatt), die recherchierbar aufbereitet und mit einer Vielzahl anderer Bilddokumente verknüpft sind. Ein solcher Datenfundus ist für Wissenschaftsdisziplinen, die räumlich und zeitlich arbeiten, hochinteressant (z.B. Historiker, Geographen, Planer, Ökologen, Statistiker, Demographen), wenn diese Altkarten georeferenziert und per interoperabler Dienste verfügbar gemacht werden. Das Kartenforum der SLUB bietet Forschern neben Werkzeugen zur Recherche und Visualisierung von Karten die Möglichkeit, die originalen Karten auf Basis eines digitalen Bestellsystems für eigene Arbeiten sowohl in digitaler als auch analoger Form zu ordern. Somit zeigt es auch ein Geschäftsmodell für Altkarten in der digitalen Welt auf, das allerdings in Teilen dem Open-Access-Prinzip gegenübersteht. Es passt aber hervorragend in die Ideen für »Geschäftsmodelle für Open-Data-Strategien des amtlichen Geoinformationswesens« (Ladstätter 2015).

Mit dem Aufkommen von GIS in der raumwissenschaftlichen Forschung stellen sich weitere Anforderungen an digitale Sammlungen alter Karten. So ermöglichen georeferenzierte Karten in Kombination mit GIS-Werkzeugen die Aufstellung neuer Forschungsfragen und -methoden, wie sie beispielsweise in den Spatial Humanities – also den mit modernen IT- und Geoinformatikmethoden arbeitenden Geisteswissenschaften – zur Geltung kommen (Bodenhamer, Corrigan und Harris 2010). Der Mehrwert solcher georeferenzierter Karten lässt sich bei Bill (2012) nachlesen bzw. live im Virtuellen Kulturlandschaftslaboratorium (VKLandLab 2014) erfahren.

### 1.3 DFG-Projekt »Virtuelles Kartenforum 2.0«

Im DFG-Projekt »Virtuelles Kartenforum 2.0 – Eine Service-orientierte virtuelle Forschungsumgebung in der Deutschen Fotothek« wurde ein solcher integrativer und nutzerfreundlicher Zugang zu der dann georeferenziert vorliegenden Kartensammlung entwickelt. Die gescannt vorliegenden topographischen Karten in der Gebietsausdehnung des ehemaligen Deutschen Reiches werden georeferenziert für Raum-Zeit-Forscher über einen standardisierten Dienst, dem vom Open Geospatial Consortium (OGC) spezifizierten WebMapService (WMS), bereitgestellt. Eine betriebsfähige virtuelle Arbeitsplattform für Raum-Zeitforscher steht flächendeckend für das Gebiet des ehemaligen Deutschen Reiches bereit, die neben Werkzeugen zur Georeferenzierung von Altkarten auch ein deutschlandweites Ortsnamensverzeichnis umfasst, welches unterschiedliche Ressourcen wie GeoNames, OpenStreetMap Nominatim, das Genealogische Ortsverzeichnis (GOV) und das Geographische Namensverzeichnis des BKG einbezieht.

Das im Kartenforum bereitstehende flächendeckende kartographische Basisangebot umfasst eine Vielzahl hochauflösend gescannter Originale der Kartensammlung der SLUB (ca. 120 Megapixel, 24 Bit color, TIF unkomprimiert, Pixelauflösung von ca. 400 dpi bzw. ca. 160 Pixel/cm bei einer Bodenauflösung von 1,6 m im Maßstab 1:25.000).

Aufgrund der bisher fehlenden Georeferenzierung dieser Digitalisate im Kartenforum war allerdings eine raum-zeitliche Suche, wie sie heute von vielen Wissenschaftsdisziplinen gewünscht wird, nicht möglich. Auch der Zugriff mit Hilfe interoperabler Webdienste oder die Recherche auf Basis standardisierter räumlicher Meta-informationen (ISO 19115) wurde nicht unterstützt. In der Summe beschränkten diese Faktoren den Zugriff auf die Daten und verhinderten auch die ad-hoc-Einbindung der Daten in lokale Arbeitsumgebungen von Wissenschaftlern.

Aus den beschriebenen Vorgaben und Defiziten ergeben sich somit eine Reihe von Anforderungen, vorangestellt die aufwändige Überführung der Digitalisate in einen räumlichen Kontext durch Georeferenzierung. Danach sollen die Produkte dieser Verarbeitung für die Nutzung bezogen auf räumliche und zeitliche Fragestellungen aus der Wissenschaft und Praxis über die Verwendung einschlägiger interoperabler Schnittstellenstandards der räumlichen Datenverarbeitung, wie z.B. des OGC, bereitgestellt werden. Der Zugang für Nutzer soll dabei sowohl über eine zentrale Portal-/Forschungsumgebung wie auch über direkte Einbeziehung in die eigene Arbeitsumgebung dienstebasiert oder durch herunterladbare Inhalte zur Verfügung stehen.

## 2 Datenlage und Georeferenzierungsmethode

### 2.1 Topographische Karten 1:25.000

Die heute insbesondere von den Landesvermessungsbehörden bereitgestellten aktuellen Topographischen Karten 1:25.000 sind in ihrer zweihundertjährigen Geschichte auf die ehemaligen Messtischblätter zurückzuführen. Als Rahmenkartenwerke decken sie ein Landesgebiet flächendeckend und nach einheitlichen Erfassungs- und Darstellungsvorschriften ab, weshalb sie für Auswertungen der Veränderungen von Landschaften von hoher Bedeutung sind. Zwar haben sich Inhalt und Darstellung über die Jahrhunderte durchaus verändert, dennoch sind die wesentlichen Landschaftselemente/Objektklassen heute wie auch damals enthalten.

Bei den etwa ab 1820 entstandenen Messtischblättern im Maßstab 1:25.000 handelt es sich um ein erstes flächendeckendes topographisches Kartenwerk, welches dann mit der Gründung des Deutschen Reiches zwischen den National- und Bundesstaaten harmonisiert und vereinheitlicht wurde und für das Gebiet in den Grenzen des ehemaligen Deutschen Reiches vorliegt. In der Weimarer Republik wurde dieses Kartenwerk vom Reichsamt für Landesaufnahme unter dem Begriff Topographische Karten 1:25.000 (TK25) geführt. Der Maßstab 1:25.000 blieb in der DDR wie auch in der alten Bundesrepublik bestehen, sodass über diesen Maßstab eine raum-zeitliche Analyse von Landschaftsveränderungen über nunmehr 150 bis 200 Jahre möglich ist.

Die Kartenblätter sind als Gradabteilungskarten nach geographischen Koordinaten beschnitten. Der Blattschnitt jeder Karte umfasst geographisch je 10 Längenminuten mal 6 Breitenminuten, also etwa 11 × 11 km. Die gedruckten Topographischen Karten im Maßstab 1:25.000 tragen in Deutschland zur eindeutigen Zuordnung seit 1937 eine meist vierstellige Blattnummer und eine Benennung, meist nach dem jeweils größten oder (historisch) bedeutendsten Ort. Die Nummerierung folgt dabei einem tabellarischen System: Die ersten beiden Ziffern geben die Zeile an (von Nord nach Süd durchnummeriert), die letzten beiden (oder drei) Ziffern die Spalte (von West nach Ost durchnummeriert). Das Nummernsystem wurde ausgerichtet auf die Grenzen des Deutschen Reiches bis 1914 (bzw. Mitte 1939), das sich weiter nach Osten und Norden erstreckte als die heutige Bundesrepublik Deutschland in den Grenzen seit 1990. Über die eindeutige Identifizierung des einzelnen Messtischblattes mittels Blattnummer sind die vier Eckpunktkoordinaten in geographischen Koordinaten auf dem damals verwendeten Bessel-Ellipsoid bekannt.

Im topographischen Kartenbestand der SLUB zum Maßstab 1:25.000 liegen diese Kartenblätter vor, für bestimmte Gebiete teilweise bis zu 10 Zeitschnitte über den Zeitraum von 1868 bis 1965 (Abb. 1). Gut 40 % der Altkarten entstammen der Zeit des kaiserlichen Deutschlands, gut ein Drittel von 1933 bis 1945, hier insbesondere

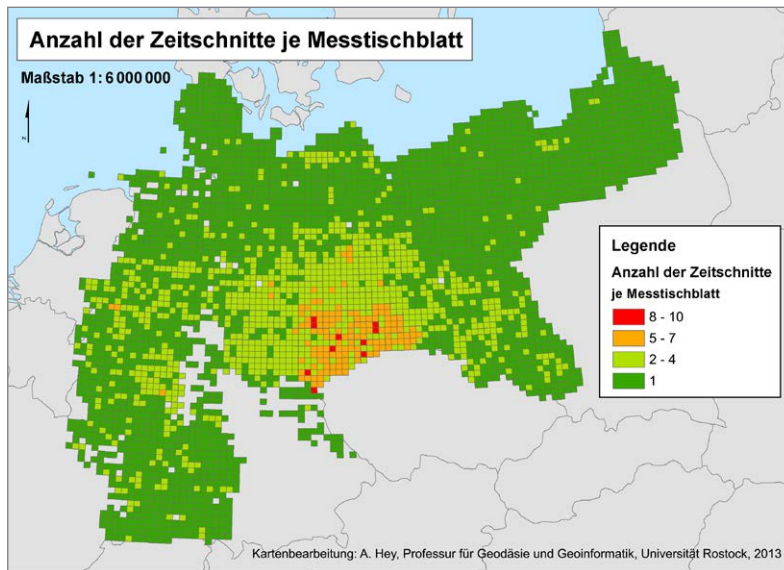


Abb. 1: Gebietsabdeckung und Anzahl der Zeitschnitte je Messtischblatt resp. TK25

die östlichen Gebiete. Die ältesten drei Karten stammen von 1868, etwa 800 Karten liegen für die Zeit der DDR bzw. der alten Bundesrepublik vor.

## 2.2 Georeferenzierung

»Georeferenzierung kann als räumliches Metakzept betrachtet werden, womit räumliche Referenzinformation einem Datensatz mitgegeben wird. Hierzu gehören die Wahl des geodätischen Bezugssystems und die Festlegung der Passpunkte, die zur Überführung verwendet werden sollen. Den eigentlichen Überführungsschritt leistet dann die Geokodierung. [...] Geokodierung behandelt den tatsächlichen Transformationsschritt, der notwendig ist, um Daten verschiedenartiger Georeferenzierung in ein gewünschtes Referenzsystem umzurechnen. Bei Rasterdaten schließt dies z.B. das Resampling der Bildelemente mit ein. [...]« (Bill und Zehner 2001).

Im Ablauf bezogen auf das »Virtuelle Kartenforum 2.0« bedeutet dies die Festlegung des Koordinatenreferenzsystems (hier EPSG-Code 4314: Deutsches Hauptdreiecksnetz, geographische Koordinaten mit Datum Potsdam und dem Ellipsoid von Bessel). Überführungen in andere Systeme wie das ETRS89/UTM sollen Transformationsdienste übernehmen. Konkret werden zur Georeferenzierung die vier Gitterpunkte am Kartenrand gemessen und als Passpunkte verwendet. Untersucht werden vier ebene überbestimmte Transformationen mit vier bis sechs Parametern. Das Rasterbild wird anschließend einem Resampling mit der Methode »nearest neighbour« unterzogen. Das einzelne georeferenzierte Kartenblatt kann dann via WMS-Schnittstelle in eigene GIS-Umgebungen eingebunden werden.

Eine grobe Abschätzung vor Projektbeginn ergab bei etwa 10 Minuten Arbeitszeit pro Georeferenzierung für die händische Messung der vier Blattecken und Rech-

nung etwa 75 Arbeitstage Aufwand. Da in Zukunft noch weitere Altkarten georeferenziert werden sollen, rechtfertigt dieser Aufwand die Auseinandersetzung mit alternativen Verfahren zur Georeferenzierung. Die Messtischblätter sollten ursprünglich mit Hilfe von Bildverarbeitungsverfahren weitestgehend automatisiert georeferenziert werden. Nach einer ersten umfangreichen Sichtung der Kartengrundlage ergab sich jedoch eine starke Heterogenität insbesondere in der Rahmgestaltung der Kartenblätter, die trotz der Verwendung von automatisierter Bilderkennung zu erheblichen manuellen Nacharbeiten führen könnte. Als aussichtsreiche Alternative wurde daher, neben dem Automatisierungsversuch, ein Crowdsourcing-Ansatz nach internationalem Vorbild verfolgt, bei welchem eine Infrastruktur aufgebaut wird, die es interessierten Nutzern er-

möglichen soll, die Altkarten selbst zu georeferenzieren. Internationale Projekte wie z.B. das der British Library ([www.bl.uk/maps](http://www.bl.uk/maps)) zeigen dabei eindrucksvoll, wie mit Hilfe von interessierten und engagierten Freiwilligen große Mengen von Altkarten verlässlich und in kürzester Zeit georeferenziert werden können.

Dieser Beitrag beschränkt sich auf die Crowdsourcing-Variante. Der automatische Bildverarbeitungsansatz wird in Koldrack und Bill (2015) beschrieben und dort mit der Crowdsourcing-Variante verglichen.

## 2.3 Georeferenzierung mittels Crowdsourcing

Der Begriff »Crowdsourcing« – zusammengesetzt aus den Begriffen Crowd und Outsourcing – beschreibt laut Gabler Wirtschaftslexikon (Springer Gabler Verlag 2014), eine interaktive Form der Wertschöpfung unter Nutzung moderner Informations- und Kommunikationstechniken (IuK-Techniken (Web 2.0)). Einzelne Aufgaben, die bisher intern bearbeitet wurden, werden an eine Vielzahl von Interessenten ausgelagert und häufig in Form eines Wettbewerbes ausgeführt. Die Aufgabe kann sich dabei sowohl auf eine Innovation beziehen oder aber auch bereits bestehende operative Aktivitäten oder Produkte. Bekanntester Vertreter für die Anwendung des Crowdsourcing dürfte das Online-Lexikon Wikipedia sein.

Crowdsourcing-Projekte sollen gemäß Gabler Wirtschaftslexikon (Springer Gabler Verlag 2014) folgende grundlegende Anforderungen erfüllen: a) Klare Aufgaben- und Zieldefinition, b) Auswahl der richtigen Crowd (Zielgruppe/Community), c) Respekt vor den Bearbeitern und ihren Ergebnissen, d) eindeutige Klärung der Rechtslage.

Für die Georeferenzierung mittels Crowdsourcing sind diese Anforderungen gegeben. Für die Georeferenzierung durch mehr oder weniger Fachfremde wurde im Projekt »Virtuelles Kartenforum 2.0« ein möglichst in-



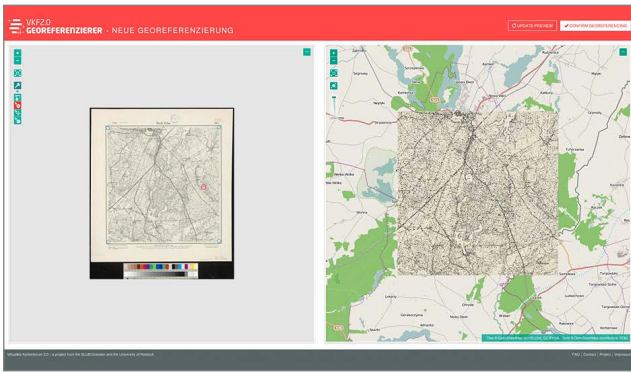


Abb. 2: Georeferenzierungs-Client: gescannte TK mit Editierwerkzeugen (links), georeferenzierte TK überlagert über OpenStreetMap-Daten mit Werkzeugen zur visuellen Überprüfung (rechts)

tuitiv bedienbarer Web-Client zur Georeferenzierung erstellt (<http://kartenforum.slub-dresden.de/vkviewer>). Der Nutzer wählt nach Interessenslage über Ortsname oder Blattnummer und Zeitangabe ein zu georeferenzierendes Kartenblatt aus. Auf Basis der per Zoomify-Technologie – einer Methode zum interaktiven performanten webbasierten Betrachten von Bilddaten – kachelbasiert (Abb. 2) ausgelieferten Scans der Kartenblätter ermöglicht der Georeferenzierungs-Webclient die Erfassung durch Messung der vier Eckpunktkoordinaten der Kartenbildrahmen. Die bekannte Methodik des Aufbaus des Blattschnittmusters der Messtischblätter, mit jeweils einer Ausdehnung von 10 Gradminuten geographischer Länge und 6 Gradminuten geographischer Breite, gestattet die Realweltkoordinaten der vier Eckpunkte im Voraus zu berechnen, in einer Datenbank zu hinterlegen und diese mit den vier gemessenen Eckpunkten der jeweiligen Kartenblätter in Referenz zu setzen. Die vier erfassten Koordinatenpaare werden nebst weiteren Metadaten in einer Datenbank abgelegt. Die Berechnung der Georeferenzierung nach der Messung der vier Eckpunkte wird mittels einer Funktion der Softwarebibliothek Geospatial Data Abstraction Library (GDAL) als polynomiale Transformation ersten Grades durchgeführt. Die Ergebnisse dieses Vorgangs können sowohl dynamisch zur Validierung an den Nutzer des Georeferenzierungs-Webclients zurückgeliefert, als auch im Anschluss persistent in der Datenhaltung (Backend) abgelegt werden. Dem Nutzer wird sofort ein visuelles Ergebnis seiner Arbeit auf dem räumlich passenden aktuellen OpenStreetMap-Datenbestand gezeigt, in dem er sich selbst von der Qualität seiner Arbeit überzeugen kann. Ausgerichtet auf die Verwendung in einem Crowdsourcing-Umfeld werden die Leistungen einzelner Nutzer systematisch gesammelt, in Form einer Nutzerhistorie visualisiert und in Form von Akkreditierungspunkten angerechnet, um so eine Art Wettbewerb zwischen den Nutzern zu stimulieren. Nach einer ersten Testphase fand eine Evaluierung der Georeferenzierungswerkzeuge statt, die zu einer Vereinfachung der Arbeitsabläufe, Verbesserung der Nutzeroberfläche und Performanzoptimierung der Georeferenzierungswerkzeuge führte.

## 2.4 Crowd-bezogene Auswertungen

Mit dem bereitgestellten Client haben insgesamt 124 Nutzer 5.681 Messtischblätter georeferenziert. Die meisten Blätter (mit 1.648 Blatt mehr als 25 % der Kartensammlung) hat ein einzelner Nutzer generiert. Nur 10 Nutzer haben mehr als 100 Blatt georeferenziert. Die meisten Nutzer haben je 10 Blätter georeferenziert. Darunter sind eine große Anzahl Studierender im Masterstudiengang Umweltingenieurwissenschaften an der Universität Rostock, die dieses Projekt zu verschiedenen Zeitpunkten im Rahmen kleiner Übungsaufgaben begleitet haben.

Im Schnitt dauert eine Messung und Georeferenzierung über die vier Karteneckpunkte weniger als 40 Sekunden. Der Nutzer sieht kurz nach dem Klicken des vierten Eckpunktes in einem zweiten Fenster die georeferenzierte Karte einem OSM-Datensatz, der gegenüber den amtlichen Geodaten auch für die ehemaligen Gebiete des Deutschen Reiches vorliegt, überlagert und kann somit die Qualität prüfen, bevor er die Abspeicherung der Ergebnisse veranlasst.

## 3 Qualitätsanalysen der Georeferenzierung

### 3.1 Gewählte Transformationsansätze

Zur Bestimmung der Transformationsparameter zwischen den Bildkoordinaten und den Weltkoordinaten wurden – neben der in GDAL gewählten Variante der 6-Parametertransformation – mehrere überbestimmte Transformationsverfahren untersucht. Schlussendlich geht es um die Bestimmung von zwei Translationen, ein oder zwei Rotationen und einem oder zwei Maßstäben, also je nach gewählter Methode um vier bis sechs unbekannte Parameter, die zu schätzen sind. Vier überbestimmte 2D-Transformationsansätze wurden mittels der Methode der kleinsten Quadrate an allen georeferenzierten Kartenblättern getestet und verglichen (ebene Ähnlichkeitstransformation mit vier Parametern und ebene Affintransformation mit fünf und sechs Parametern (Luhmann 2003)). Hierzu wurde die in Java programmierte Open Source Software »CoordTrans – The Open Source Similarity Transformation Program« von Michael Lösler genutzt (Dokumentation siehe <http://derletztekick.com>). Diese überbestimmten Transformationen erlauben neben der Koordinatenüberführung noch die Berechnung von Qualitätsmaßen, die einerseits die Genauigkeit der Koordinatentransformation angeben und andererseits die Deformationen der Altkarten beschreiben. Die erreichte Genauigkeit lässt sich am besten durch Angabe der mittleren Koordinatengenauigkeit transformierter Punkte angeben, ein in der Geodäsie und Kartographie übliches Genauigkeitsmaß, welches auch leicht interpretierbar ist, da es einen Bezug zum Kartenmaßstab herstellen lässt.

Die Messtischblätter und topographischen Karten haben den Maßstab 1:25.000, d. h. eine Zeichengenauigkeit

von 0,1 bis 0,2 mm unter heutigen Qualitätsmaßstäben würde zu einem Lagefehler von etwa 2,5 bis 5 m führen. Die Scanqualität der Messtischblätter liegt mit 1,6 m ebenfalls in dieser Größenordnung. Die Qualität der Kartenherstellung vor fast 150 Jahren war jedoch geringer, zudem sind noch Lagerungseffekte, Deformationen des Originals sowie Qualitätsverluste beim Scannen zu berücksichtigen. Daher ist für diese Altkarten bestenfalls eine Koordinatengenauigkeit von etwa 5 bis 10 m zu erwarten, die im Einzelfall auch deutlich schlechter sein kann.

Die vier untersuchten Transformationsvarianten sind:

- M1: 2 Translationen, 1 Rotation, 1 Maßstab (Helmert-Transformation mit 4 Parametern);
- M2: 2 Translationen, 2 Rotationen, 1 Maßstab;
- M3: 2 Translationen, 2 Rotationen, 2 Maßstäbe;
- M4: 2 Translationen, 1 Rotation, 2 Maßstäbe.

Zwei Maßstäbe bilden unterschiedliche Deformationen der Kartenblätter in Längs- und Querrichtung nach, zwei Rotationen können eine Scherung des Kartenblattes, d. h. die Nichtrechtwinkligkeit des Koordinatensystems des alten Kartenblatts berücksichtigen. Die Variante M3 entspricht der polynomialen Transformation ersten Grades bzw. der oben genannten Affintransformation mit sechs Parametern, die im Georeferenzierungs-Client genutzt wird.

### 3.2 Statistische Auswertungen der Stichprobe

Untersucht wird die für die vier Gittereckpunkte aus der Ausgleichung ermittelte Koordinatengenauigkeit  $\sigma_x$  und  $\sigma_y$  sowie eine daraus ermittelte Punktgenauigkeit  $\sigma_p$ . Bei den untersuchten 5.681 Blättern ergeben sich durchschnittliche Koordinatengenauigkeiten zwischen 3,5 m und 5.811 m (s. Tab. 1). Die wenigen großen Abweichungen sind auf einen einzelnen Nutzer zurückzuführen, der grobe Fehler produzierte.

Mittels Globaltest (Niemeier 2008, gewählte Irrtumswahrscheinlichkeit  $\alpha = 0,1\%$  und Güte des Tests  $\beta = 80\%$ ) werden die gewählte a priori Varianz und die sich aus der Ausgleichung ergebende a posteriori Varianz zueinander getestet. Die Nullhypothese wird bei den Varianten M1 und M2 in über einem Drittel der gemessenen Blätter ver-

worfen, d. h. die einfachen Modelle mit 2 Translationen, 1 oder 2 Rotationen und einem Maßstab reichen nicht zur Koordinatenüberführung aus. Hingegen erbringt die Einführung eines zweiten Maßstabs (M4, 3,3 % verworfen) oder zusätzlich einer zweiten Rotationskomponente (M3, 3,3 % verworfen) eine deutliche Verbesserung zwischen Modellannahmen und Messung. Erwartungsgemäß fängt die Variante M3 mit sechs zu bestimmenden Parametern die Verschiebungen, Scherungen und Maßstabsunterschiede am besten auf. Die beste Koordinatengenauigkeit liegt in allen Varianten bei unter 4 m, welches bei den optimistischsten Schätzungen der aus der Lagerung und dem Scannen abgeleiteten Qualitätsaussagen liegt. Bezogen auf die Variante M3 haben weniger als 1 % der Georeferenzierungen eine Genauigkeit schlechter als 100 m. Damit kann die Georeferenzierung mittels Crowdsourcing als erfolgreich interpretiert werden.

Für jede der vier Varianten wurden zwei Histogramm- auswertungen bezüglich der erreichten Punktgenauigkeit durchgeführt (hier dargestellt am Beispiel der Variante M3):

1. Gesamtstichprobe Punktgenauigkeit (alle 5.681 Kartenblätter) (Abb. 3),
2. Punktgenauigkeit der Stichprobe mit den 5.558 Kartenblättern, die den Globaltest bestanden haben, also nach Elimination der als Ausreißer detektierten Georeferenzierungen (Abb. 4).

### 3.3 Nutzerbezogene Auswertungen

Die meisten Karten (fast 3.200) wurden nach Freischaltung des Portals Ende Juli 2014 georeferenziert. Eine größere Anzahl von Messtischblättern wurde daneben im Rahmen studentischer Projektarbeiten im Masterstudien- gang Umweltingenieurwissenschaften im Sommersemester 2013 (Monate 06–07) und 2014 (Monate 04–06) sowie Anfang des Wintersemesters 2013 (Monate 10–11) erfasst. Der Nutzer mit der größten Anzahl georeferenzierter Blätter zeigt auch ein sehr gutes Qualitätsverhalten. Im Mittel liegt die Qualität der georeferenzierten Blätter dieses Nutzers bei 15,7 m in der Variante M3. Deutlich davon unterscheidet sich die Qualität des Nutzers mit der schlechtesten Georeferenzierung,

der grobe Fehler im Bereich von mehreren 1.000 m erzeugte. Wie sich in der Einzelblattanalyse zur Fehlerursacheneingrenzung zeigte, hat dieser Nutzer den Blattschnitt sehr kreativ interpretiert und mehrere Zentimeter weggelassen (vgl. Abb. 6).

Tab. 1: Statistische Auswertung aller Georeferenzierungen mittels Crowdsourcing

	M1	M2	M3	M4
Stichprobenumfang	5.681	5.681	5.681	5.681
Mittlere Koordinatengenauigkeit	30,65	19,52	11,27	9,89
Mittlere Punktgenauigkeit	43,34	27,60	15,93	13,99
Standardabweichung Koordinatengenauigkeit	289,36	40,25	6,52	6,23
Standardabweichung Punktgenauigkeit	409,21	56,98	9,22	8,82
Maximale Koordinatengenauigkeit	5810,76	2277,71	238,75	186,72
Minimale Koordinatengenauigkeit	3,72	3,53	3,44	3,58
Globaltest: $H_0$ verworfen	34,7 %	34,4 %	2,1 %	3,3 %

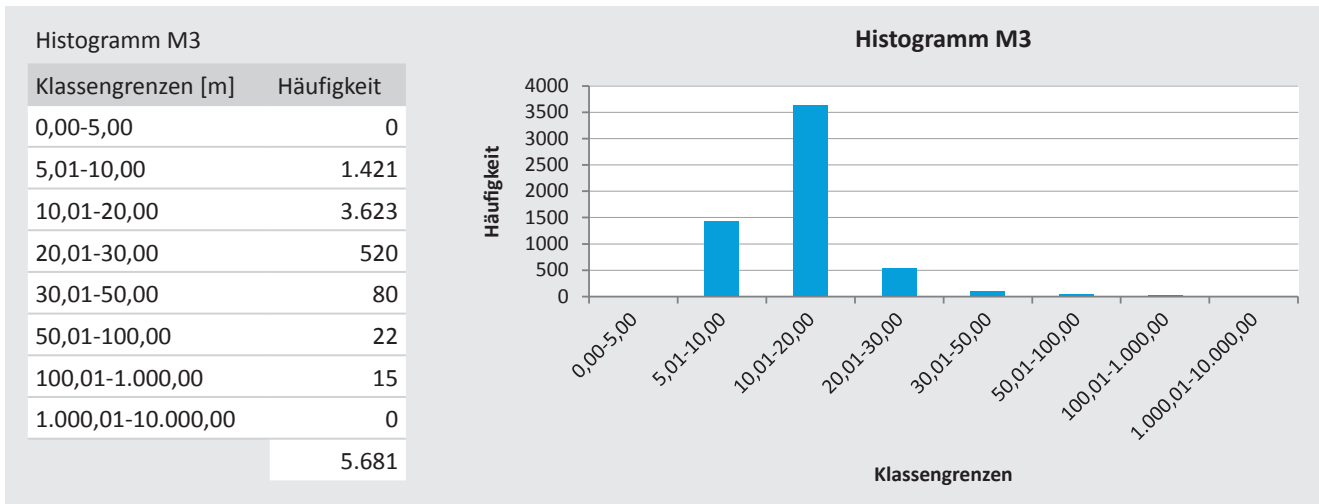


Abb. 3: Punktgenauigkeit der Gesamtstichprobe

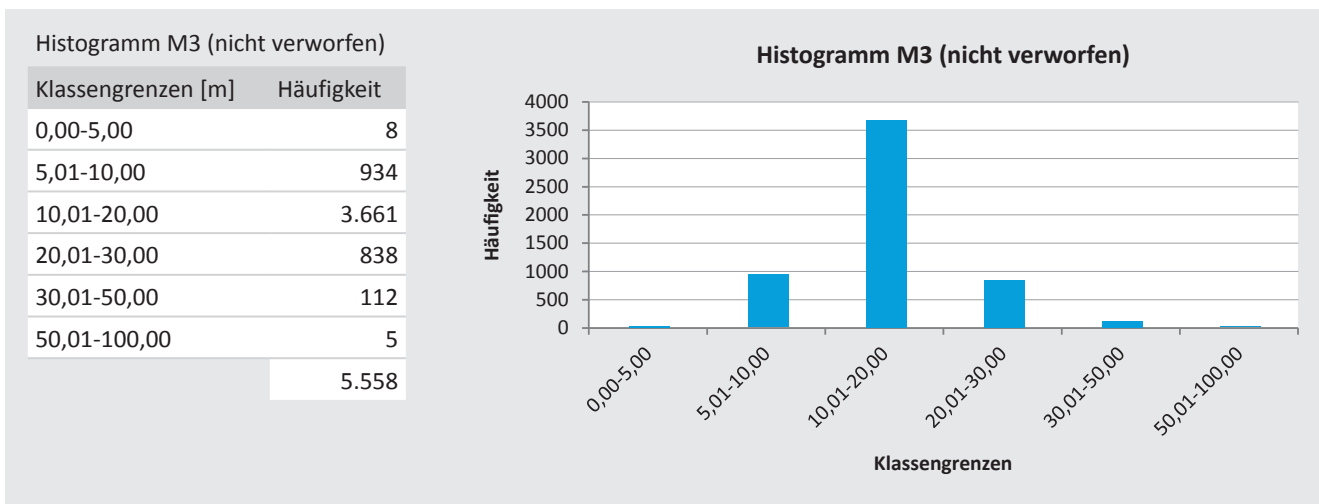


Abb. 4: Punktgenauigkeit der reduzierten Stichprobe nach Globaltest

### 3.4 Qualitätsanalyse nach Kartenalter

Die Fotothek beinhaltet Messtischblätter und Topographische Karten 1:25.000 aus den Zeiträumen von 1868 bis 1965, insgesamt darin jedoch nur aus 89 Jahren, da nicht von jedem Jahr Karten vorliegen. Hierbei dominieren die Jahrgänge ab 1930. Wird die Stichprobe in drei etwa gleich große Zeitabschnitte geteilt, so unterscheidet sich die geometrische Georeferenzierungsqualität kaum hinsichtlich des Kartenalters, welches für eine sorgfältige Lagerung und ein hochwertiges Scannen der Originale spricht. Mit 3,5 m bester Koordinatengenauigkeit wird fast die vorhergesagte beste Qualität von etwa 2,5 m erreicht (s. Tab. 2).

### 3.5 Ansätze zur Fehlerinterpretation

Erste Erkenntnisse zur Fehlerinterpretation der Georeferenzierung mittels Crowdsourcing sind:

- Nur wenige der per Crowdsourcing erzeugten Georeferenzierungen sind nicht verwendbar. Als Grenze wird in der aktuellen Lage eine mittlere Koordinatengenauigkeit von 100 m betrachtet.

Tab. 2: Statistische Analysen nach Kartenalter [Variante M3]

	komplette Stichprobe		
	1872–1899	1900–1930	1931–1965
Jahrgänge			
Stichprobenumfang	1.175	1.893	2.613
Mittlere Koordinatengenauigkeit	10,84	11,17	11,53
Mittlere Punktgenauigkeit	15,33	15,80	16,30
Standardabweichung Koordinatengenauigkeit	5,03	6,33	7,20
Maximale Koordinatengenauigkeit	64,48	87,02	238,75
Minimale Koordinatengenauigkeit	3,68	3,44	3,49
Standardabweichung Punktgenauigkeit	7,12	8,96	10,18



- Einzelne Blätter waren zur Lagerung ein- bis mehrfach gefaltet. Die Faltungsbreite beträgt etwa 2 mm, welches bei dem Maßstab von 1:25.000 etwa 50 m Versatz ausmacht.
- Visuell ist auch erkennbar, dass einzelne Blätter definitiv falsch georeferenziert wurden, so wurde z.B. statt der gewünschten inneren Kartenrandecke der äußere Kartenrand als Punkt digitalisiert (Abb. 5), welches bei einem Kartenrandgestaltungs-bereich von etwa 1 cm und beim Maßstab 1:25.000 etwa 250 m Versatz ergibt.
- Es gibt nur einen wirklichen Ausreißer, wie bereits bei der Nutzersauswertung erwähnt. Dieser Nutzer hat willkürlich die Kartenblätter eigenmächtig beschnitten, wie der Vergleich mit dem Originalmessstischblatt in der Deutschen Fotothek zeigt (Abb. 6).

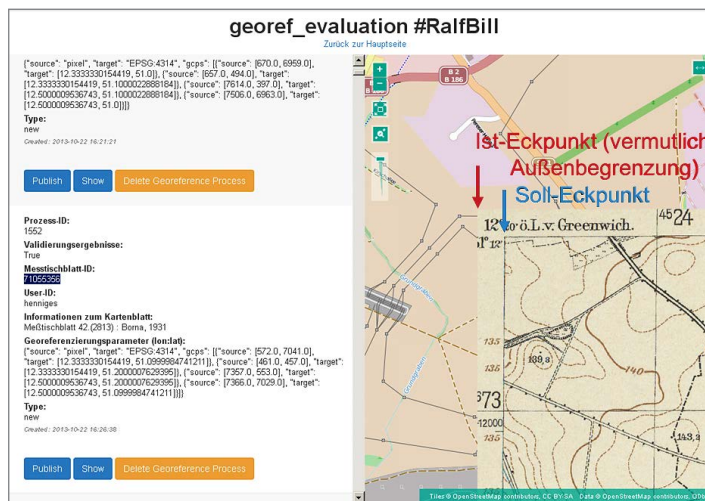


Abb. 5: Falsche Randlinie

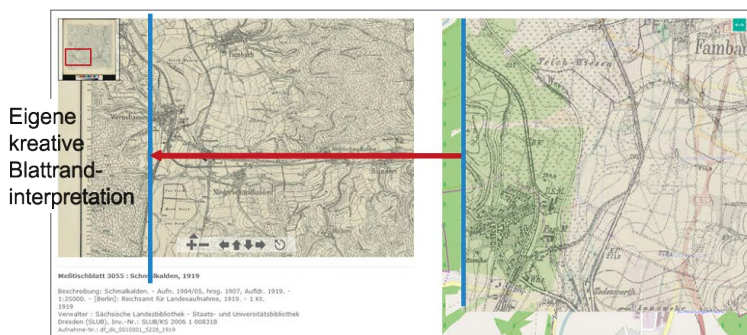


Abb. 6: Kreative Randinterpretation

## 4 Fazit

Die Georeferenzierung durch Fachfremde hat sich als verlässlicher Weg dargestellt, der zu hohen Genauigkeiten führt. Mit den im Virtuellen Kartenforum angebotenen Werkzeugen zur Administration können die Ergebnisse gut evaluiert werden und bei Bedarf kann ein Blatt auch wieder zur neuen Georeferenzierung freigeschaltet werden. In einem zukünftigen Beitrag (Koldrack und Bill, angekündigt für 2015) wird ein Vergleich zu einer auf Bildverarbeitung basierenden Lösung dargestellt.

## Dank

Die Autoren danken der DFG für die Finanzierung des Projekts »Virtuelles Kartenforum 2.0« (Förderkennzeichen Bi 467/24-1 und Bu 2228/14-1). Dank geht auch an die Kollegen des SLUB-Teams.

## Literatur

- Bill, R. (2014): Geoinformatik im Kontext der E-Science. In: GIS.SCIENCE, 2014, Nr. 4, S. 123–128.
- Bill, R. (Hrsg.) (2012): Kulturlandschaftsforschung in einer virtuellen Forschungsumgebung auf Basis von Internet-GIS-Technologien. Wichmann Verlag, Berlin, 186 Seiten.
- Bill, R., Walter, K., Mendt, J. (2014): Virtuelles Kartenforum 2.0 – Verfügbarmachung von Altkarten über eine räumliche Portalanwendung. In: Strobl, J., Blaschke, T., Griesebner, G., Zagel, B. (Hrsg.): Angewandte Geoinformatik 2014. Wichmann Verlag, Offenbach, S. 684–693.
- Bill, R., Zehner, M.L. (2001): Lexikon der Geoinformatik. Wichmann Verlag, Heidelberg, 312 Seiten.
- Bodenhamer, D.J., Corrigan, J., Harris, T.M. (Ed.) (2010): The Spatial Humanities – GIS and the Future of Humanities Scholarship. Indiana University Press, Bloomington.
- British Library (2014): <http://britishlibrary.typepad.co.uk/magnificentmaps/2014/01/done-2700-maps-georeferenced-by-volunteers.html>, letzter Zugriff 9.4.2015.

- DR – David Rumsey Map Collection (2014): [www.davidrumsey.com](http://www.davidrumsey.com), letzter Zugriff 9.4.2015.
- GEO-LEO (2014): <http://geo-leo.de/en>, letzter Zugriff 9.4.2015.
- IKAR (2014): <http://ikar.staatsbibliothek-berlin.de>, letzter Zugriff 9.4.2015.
- Koldrack, N., Bill, R. (2015): Automatisierte Georeferenzierung alter topographischer Karten – Algorithmus und Qualitätsanalyse. Erscheint in dieser Zeitschrift.
- Ladstätter, P. (2015): Geschäftsmodelle für Open-Data-Strategien des amtlichen Geoinformationswesens. zfv 140, Heft 2, S. 70–75, 2015.
- Luhmann, T. (2003): Nahbereichsphotogrammetrie. Grundlagen, Methoden und Anwendungen. Wichmann Verlag, Heidelberg, 586 Seiten.
- Niemeier, W. (2008): Ausgleichsrechnung. Statistische Auswertungsmethoden. DeGruyter Verlag, Berlin, 493 Seiten.
- OMO – Old Maps Online (2014): [www.oldmapsonline.org](http://www.oldmapsonline.org), letzter Zugriff 9.4.2015.
- SLUB – Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (2014): [www.slub-dresden.de/sammlungen/karten](http://www.slub-dresden.de/sammlungen/karten) bzw. [www.deutschefotothek.de/cms/kartenforum.xml](http://www.deutschefotothek.de/cms/kartenforum.xml), letzter Zugriff 9.4.2015.
- Springer Gabler Verlag (Hrsg.) (2014): Gabler Wirtschaftslexikon, Stichwort: Crowdsourcing, online im Internet unter <http://wirtschaftslexikon.gabler.de/Definition/crowdsourcing.html>, letzter Zugriff 9.4.2015.
- VKLandLab – Virtuelles Kulturlandschaftslaboratorium (2014): <http://kvwmap.geoinformatik.uni-rostock.de/VKLandLab/index.php?gast=45>, letzter Zugriff 9.4.2015.

## Anschrift der Autoren

Prof. Dr.-Ing. Ralf Bill | Dr.-Ing. Kai Walter  
 Universität Rostock, Professur für Geodäsie und Geoinformatik  
 18051 Rostock  
[ralf.bill@uni-rostock.de](mailto:ralf.bill@uni-rostock.de) | [kai.walter@uni-rostock.de](mailto:kai.walter@uni-rostock.de)