

Die Digitalisierung der Katasterarchive des BEV – ein Einblick in das Projekt „VhwhIST“

Georg Topf

Bundesamt für Eich- und Vermessungswesen, Wien · georg.topf@bev.gv.at

Zusammenfassung: Das Bundesamt für Eich- und Vermessungswesen (BEV) verfügt über 4,2 Millionen Katasterpläne mit einem Gesamtvolumen von 50 Millionen Seiten. Diese Katasterpläne wurden über die letzten 132 Jahre in verschiedensten Formaten auf unterschiedlichen Trägermaterialien erstellt und in den lokalen Katasterarchiven der Vermessungsämter archiviert. Das BEV hat sich nun das Ziel gesetzt, alle Planunterlagen zu digitalisieren und im Archivformat PDF A/1b zu speichern und für den Online-Zugriff zur Verfügung zu stellen. Dieser Praxisbericht gibt einen Einblick in das Projekt, zeigt außerdem die Schritte des semiautomatischen Workflows und endet mit dem derzeitigen Stand der Umsetzung.

Schlüsselwörter: Kataster, Digitalisierung, Archiv, QR-Code

Abstract: 4.2 million cadastral maps of the last 132 years with a volume of more than 50 million pages in different layouts on a variety of materials are managed by the local cadastral authorities of the federal office of metrology and surveying (BEV). So far only available local, the BEV set the objective to digitize these enormous amounts of maps and provide them in the standardized long term archive format PDF A/1-b. This report delivers insight into the project, shows the steps of the semiautomatic workflow and ends with the current state of implementation.

Keywords: Digitalisation, cadastre, archive, qr-code

1 Motivation

Alle verfügbaren Katasterpläne der letzten 132 Jahre sind entsprechend der räumlichen Lage des Inhalts einem Vermessungssprengel zugeordnet. Das BEV hat 50 lokale Katasterarchive, die sich auf 41 Vermessungsämter und neun Dienststellen verteilen. Interne und externe Tätigkeiten, bei denen diese Planunterlagen benötigt werden, waren bisher zeitlich und örtlich an die jeweilige Archivstelle gebunden.

Das BEV startete im Jahr 2012 das Projekt „Digitalisierung der Katasterarchive“, um diese zeitlichen und räumlichen Barrieren zu überwinden. Dabei wurden folgende Ziele für das Projekt festgelegt:

1. Digitale Abgabe der Katasterpläne an alle Kunden (Umsetzung von Kundenanforderungen).
2. Abbau lokal gebundener Aufgaben und Tätigkeitstransfer zu allen Organisationseinheiten des BEV (Wandern der Arbeit anstatt der Mitarbeiter).
3. Sicherung der analogen Planurkunden (wertvolles, gefährdetes Archivgut).

2 Herausforderung

Die Digitalisierung (Datenerfassung mittels Scanner) analoger, genormter Unterlagen wie diverse Schriftstücke, Tabellen und Bilder stellt in der heutigen Zeit grundsätzlich keine Herausforderung mehr dar. Wurden Rahmenbedingungen wie Auflösung, Farbtiefe und geltende Standards für Langzeitarchivierung bedacht, ist der reine Analog-digital-Transfer mit geeigneten Scannern relativ simpel.

Was ist aber alles zu bedenken, wenn ein Katasterarchiv mit Unterlagen der letzten 132 Jahre digitalisiert werden sollte? Wenn 4,2 Millionen Pläne in unterschiedlichen, teilweise nicht genormten Formaten klassifiziert und aufbereitet werden müssen? Wenn unterschiedliche Trägermaterialien in den verschiedensten Ausprägungen von zarten Bleistiftskizzen bis hin zu hunderten Feldskizzen einander abwechseln? Welche Erfahrungswerte gibt es bei der Digitalisierung grob heterogener Daten und gibt es Werkzeuge, die diesen Anforderungen gerecht werden?

Diese und viele andere Fragen galt es vorab zu klären, bevor mit dem eigentlichen Scanprozess der Katasterarchive gestartet werden konnte. Daher wurde dieses Projekt in drei Phasen gegliedert:

- Vorprojektphase: Evaluierung der Methoden und Verfahrenstechniken zur digitalen Erfassung der Planurkunden, der Infrastruktur sowie der Hard- und Softwaretools.
- Pilotprojekt: Aufbau der Logistik, Anwenden der Methoden auf drei Pilotämter, Adaptierung des Workflows und Abschätzung der Ressourcen für Zeit- und Personalbedarf.
- Umsetzungsprojekt: Definition des Gesamtprojektes, Einbindung aller Archivstellen des BEV und Einrichtung der Datenabgabe über das BEV-Geodatenportal.

3 Projektphasen

3.1 Vorprojektphase

Die Vorprojektphase (2012/2013) war von der Evaluierung der Methoden und Verfahrenstechniken zur digitalen Erfassung der Planurkunden geprägt. Dabei konnten auch die Erfahrungen des Südtiroler Katasteramts in Bozen genutzt werden. Dieses hat selbst historische Unterlagen digitalisiert und wertvolle Tipps von der Vorbereitung bis hin zur digitalen Nachbereitung der Daten bereitgestellt. Weitere Besuche von Organisationen im In- und Ausland lieferten Hinweise auf Problemfälle. Diese wurden bei der Planung berücksichtigt.

In der Evaluierungsphase wurden vier unterschiedliche Testreihen durchgeführt, wo anhand der Testbeispiele unterschiedliche Hardware- und Softwaretools externer Firmen getestet und bewertet wurden. Ein wichtiges Ergebnis aus den Testreihen war, dass die heterogenen Datenbestände der Katasterarchive nur mit verschiedenen Scangerättypen digitalisiert werden können. Alleine schon diese Erkenntnis war für die weitere Entwicklung des Scanworkflows bedeutend. Die Information der Zusammengehörigkeit einzelner Planteile muss somit bereits bei der Aufbereitung des Scanmaterials definiert, als Metainformation erfasst und dem Folgeprozess bereitgestellt werden.

Für die Erfassung der Metadaten wurde eine Applikation entwickelt, in der die erforderlichen Attribute eingegeben, verwaltet und die Ergebnisse als QR-Codes ausgedruckt werden können. Pro Plan können ein oder mehrere QR-Code-Blätter entstehen, je nachdem, ob ein Plan auf einem oder wegen eines Überformates auf mehreren Geräten gescannt werden muss. Als Überformate wurden jene Planunterlagen definiert, welche das Format 29,7 cm × 150 cm überschreiten oder ein Trägermaterial eine gewisse Stärke überschreitet.

Mit dem QR-Code werden Informationen wie

- Dateiname,
- die Anzahl der Seiten pro Scannertyp,
- die Anzahl der Überformate und
- die Position der Überformate

an die Scansoftware übergeben. Durch Einlesen der QR-Codes beim Scannen werden die Metadaten ausgelesen, die Anzahl der gescannten Seiten mit der Seitenanzahl der Metainformation überprüft und die restlichen Metadaten an die Serversoftware weitergeleitet. Aufgrund dieser Daten setzt die speziell adaptierte Serversoftware die getrennt gescannten Plan-teile zusammen und generiert daraus den digitalen Plan im Format PDF/A1b. In Abbildung 1 wird dieser Scanworkflow schematisch dargestellt.

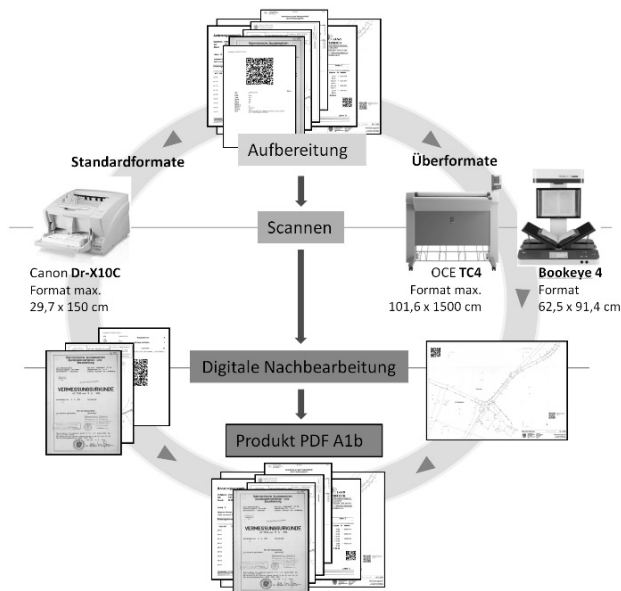


Abb. 1: Schematische Darstellung des Scanworkflows

3.2 Pilotprojekt

Das Pilotprojekt wurde Ende 2013 gestartet. Ziel war, bis Ende 2014 die Unterlagen dreier ausgewählter Archivstellen zu digitalisieren, den Scanworkflow zu testen und messbare Ergebnisse für das Umsetzungsprojekt zu sammeln. Um das Projekt in seinem Umfang greifbar zu machen, wurden folgende Themenbereiche als eigene Projektsäulen eingerichtet:

- Logistik,
- Klassifizierung,
- Scannen und
- Langzeitarchiv.

In der Logistik wird die Bereitstellung des Materials (Archivboxen, Transportkisten, Schneidemaschinen, Etiketten), Organisation der Transporte zwischen Archiv, Klassifizierungsstelle und Scanstraße, Abstimmung der Mengen zwischen Vorbereitung und Digitalisierung sowie die Kommunikation mit den Klassifizierungsstellen wahrgenommen.

Bei den Klassifizierungsstellen werden die Unterlagen in verschiedene Kategorien klassifiziert und für den Scanprozess aufbereitet. Sämtliche Heft- und Büroklammern, Schnellhefter, Schnüre usw. werden aus dem Archivgut entfernt, geklebte Unterlagen aufgetrennt sowie Beschädigungen, Risse und Knicke beseitigt. Besonders wichtig bei dieser Tätigkeit sind die korrekte Erfassung der Metadaten, die Erzeugung der QR-Codes sowie die Identifikation der Überformate.

Beim Scannen werden die aufbereiteten Archivpläne an die Scangeräte verteilt, die Unterlagen aufgefaltet, gescannt und gleichzeitig das Ergebnis am Bildschirm kontrolliert. Die Anzahl der Seiten aus dem QR-Code wird mit dem Ergebnis des Scannens verglichen und bei Abweichung überprüft. Die speziell adaptierte Serversoftware führt die Planunterlagen, die auf mehreren Gerätetypen gescannt wurden, zusammen. Alle Pläne werden im Format PDF/A1b bereitgestellt und jede Nacht auf ein zentrales Laufwerk transferiert.

Das Langzeitarchiv wurde eingerichtet, um einerseits die dauerhafte Sicherung des digitalen Archivguts zu gewährleisten und andererseits, die Datenabgabe über das BEV-Portal zu unterstützen. Im Langzeitarchiv werden zusätzlich Metainformationen gespeichert, um unabhängige Recherchen nach Plänen, zum Beispiel über Grundstücksnummern, zu ermöglichen.

3.2.1 Ergebnisse des Pilotprojektes

Die Digitalisierung der Unterlagen der drei Pilotämter mit einer Anzahl von 259 Katastralgemeinden wurde mit Jahresende 2014 abgeschlossen. Es wurden 103.688 Archivpläne mit einer Anzahl von 1,42 Millionen Seiten und 28.769 Überformaten aufbereitet, gescannt und digital bereitgestellt.

Folgende Scanner wurden im Pilotprojekt eingesetzt:

- 2x Dokumentenscanner Canon DR-X10c
- 1x Flachbettscanner Canon Océ TC 4
- 1x Bookeye 4 Professional DinA1+

Das Ergebnis hat gezeigt, dass alle bisherigen Unterlagen mit diesen Geräten gescannt werden konnten. Dabei werden ca. 98 % der Unterlagen auf den Dokumentenscannern, die restlichen 2 % auf den beiden Großformatscannern digitalisiert. Mit den beiden Dokumentenscannern konnte bei einem durchschnittlichen 7h-Betrieb eine tägliche Scanleistung von 620 Plänen erreicht werden. Die Auslastung der Überformatscanner betrug ca. 2 h pro Tag, auch wenn manche Pläne, wie in Abbildung 2 ersichtlich ist, mehr Zeit in Anspruch nahmen.

Für die Aufbereitung und Klassifizierung der Pläne im Pilotprojekt wurde eine durchschnittliche Leistung von 800 Plänen pro Tag erreicht. Es wurde eine detaillierte Richtlinie erstellt, welche die eindeutige Zuordnung der Planbeilagen zu den vordefinierten Planabschnittstypen

(Planteilen) beinhaltet. Die Verfahrensabläufe wurden anhand der Richtlinie definiert und eingerichtet.



Abb. 2:
Digitalisierung von Überformaten am Bookeye 4

3.3 Umsetzungsprojekt

Mit dem Jahr 2015 wurde das Umsetzungsprojekt gestartet. Als Rahmenbedingungen wurden folgende Ziele festgelegt:

1. Vollständige Digitalisierung der Unterlagen von vier Archivstellen pro Jahr.
2. Digitalisierung aller Katasterarchive bis zum Ende des Jahres 2024.

Aus diesen Vorgaben folgt, dass jedes Jahr ca. 400.000 Archivpläne digitalisiert werden müssen. Um dieser Anforderung zu genügen, wurde in einem ersten Schritt die Auslastung der Scanstraßen auf 12 h erhöht. Der zweite Schritt erfolgte im September 2015 mit dem Ankauf einer dritten und vierten Scanstraße, was zu einer Verdoppelung der Kapazitäten der Scanleistung führte. Damit wurde auch erstmals ein Scandurchsatz von 2.000 Archivplänen pro Tag erreicht.

Im Bereich der Klassifizierung wurden weitere Archivstellen in den Klassifizierungsprozess eingebunden. Diese bereiten die Archivunterlagen ihres eigenen Wirkungsbereiches auf und helfen mit, die Scanstraße entsprechend auszulasten. Im November 2015 wurde damit begonnen, alle Archivstellen in den Klassifizierungs- und Scanprozess einzubinden.

3.3.1 Ergebnisse des Umsetzungsprojektes im Jahr 2015

Die Digitalisierung der vier Archivstellen mit einer Anzahl von 457 Katastralgemeinden wurde im Jahr 2015 abgeschlossen. Es wurden 217.005 Archivpläne mit einer Anzahl von 2,85 Millionen Seiten und 54.022 Überformaten digitalisiert.

Durch die Erweiterung der Kapazitäten auf der Scanstraße und den Archivstellen konnte gezeigt werden, dass mit einem Scandurchsatz von 2.000 Archivplänen pro Tag die Projektlaufzeit von 10 Jahren umgesetzt werden kann.

4 Status und Ausblick

Von den 7.850 Katastralgemeinden (KG) sind derzeit 1.356 KGs (432.501 Archivpläne) digital verfügbar, 195 KGs befinden sich in Klassifizierung und 6.299 KGs wurden noch nicht begonnen. Der Umsetzungsstand ist auch in Abbildung 2 ersichtlich (Stand: 20.04.2016). Eine wöchentlich aktualisierte Statusliste kann auf der BEV-Homepage www.bev.gv.at unter „Digitalisierung Katasterarchive“ heruntergeladen werden.

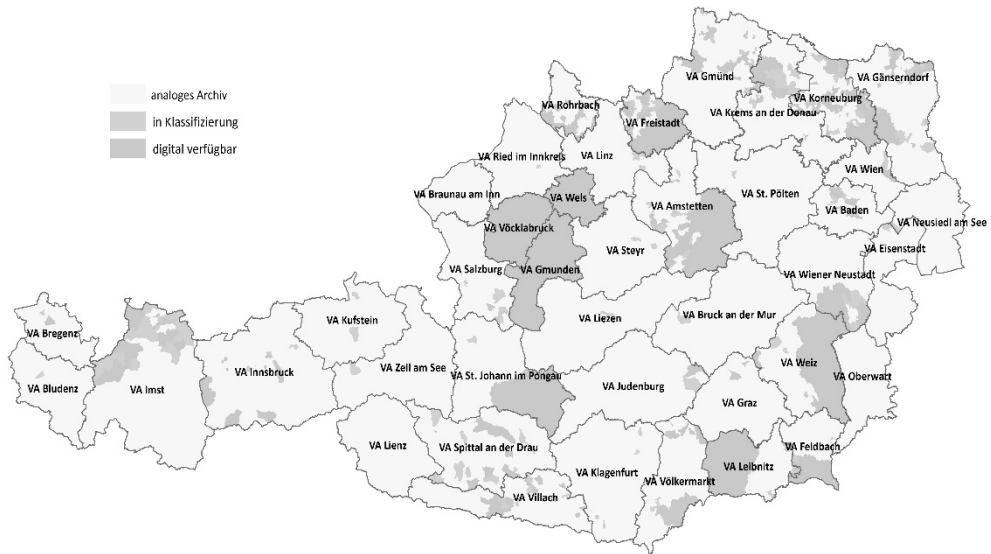


Abb. 2: Überblick der bereits digital verfügbaren Archivpläne (20.04.2016)

Derzeit werden die Pläne von den Vermessungsämtern per E-Mail oder FTP-Server an die Kunden übermittelt. Die Dateigröße der Pläne bewegt sich zwischen 0,5 Megabyte und 5,5 Gigabyte, wobei 98,8 % der Pläne kleiner als 50 Megabyte sind. Der Preis für die digitale Kopie eines Planes wurde mit 0,80 € festgelegt, unabhängig von dessen Größe und Seitenanzahl. Bis Mitte Juni 2016 wird der Abgabeprozess umgestellt und die Abgabe über den Webshop im BEV-Portal ermöglicht.