

02 - Der Weg digitaler Unterlagen ins Magazin

Ein Erfahrungsbericht zur Übernahme von Geodaten

Das Staatsarchiv Freiburg übernahm 2013 digitale Unterlagen des Landesamtes für Geologie, Rohstoffe und Bergbau (LGRB). Dort entstehen unter anderem Karten und Geodaten, die teilweise Grundlage für Veröffentlichungen sind. Die Daten sind einzelnen Projekten zugeordnet, die das LGRB ausführt. Aus der Menge von 240 verschiedenen Projekten wurden für eine erste Ablieferung fünf Projekte ausgewählt, die im Landesarchiv dauerhaft verwahrt werden sollen. Da das LGRB für ganz Baden-Württemberg und nicht nur für den Regierungsbezirk Freiburg zuständig ist, wurde hierbei vor allem darauf geachtet, Projekte auszuwählen, die größere Teile Baden-Württembergs betreffen und nicht zu stark regional eingegrenzt sind. Bis zu diesem Punkt handelt es sich um eine ganz normale archivistische Bewertung.

Eine Ablieferung und drei Informationspakete

Abgeliefert wurde eine DVD mit insgesamt 10.884 Dateien, die in 228 Ordnern gespeichert waren. Hier zeigen sich erste Probleme für Neueinsteiger in der digitalen Archivierung: Wie geht man mit diesen Dateimengen um? Und was ist jetzt überhaupt zu tun, bevor man die Dateien zur dauerhaften Speicherung in einem digitalen Magazin ablegt?

Wer sich mit digitaler Langzeitarchivierung beschäftigt, lernt bald die gängigen theoretischen Konzepte kennen, allem voran das Referenzmodell OAIS (Open Archival Information System). In diesem Modell werden die Organisation und der Ablauf der digitalen Langzeitarchivierung beschrieben; besonders wichtig ist dabei die Beschreibung von digitalen Daten als verschiedene Informationspakete. Was Behörden an die Archive abliefern, wird als ‚Submission Information Package‘ (SIP) bezeichnet, die Form, in welcher die Daten dauerhaft gespeichert werden als ‚Archival Information Package‘ (AIP). Zuletzt wird noch das ‚Dissemination Information Package‘ (DIP) beschrieben, welches einem Nutzer zur Verfügung gestellt wird, auf das hier aber nicht näher eingegangen wird. Dem OAIS-Modell zufolge muss aus dem SIP erst das AIP gebildet werden, diese Informationspakete sind nicht unbedingt miteinander identisch. Hierzu gehört auch die Migration einzelner Dateien in ein dauerhaft lesbares Dateiformat und die Anreicherung der Ablieferung mit Metadaten, was hier allerdings auch nicht näher behandelt werden soll.

Anders und vereinfacht ausgedrückt: das, was von einer Behörde an digitalen Daten übernommen wird, muss nicht zwangsläufig auch genau so archiviert werden.

Es muss nicht jede einzelne Datei am Ende übernommen werden und die Unterlagen müssen nicht in genau dem Zustand und in der Ordnung dauerhaft archiviert werden, wie sie abgeliefert wurden - ein gewaltiger Unterschied zur klassischen, analogen Archivierung und für Einsteiger eine wichtige Information.

Ein Paket wird geschnürt

Wie läuft nun die Bildung eines AIPs in der Praxis ab?

Die Frage ist, welche Dateien für die Überlieferung tatsächlich relevant sind und bewahrt werden sollen und in welcher Ordnung die Dateien dauerhaft gespeichert werden sollen.

Im Fall der Ablieferung des LGRB war es relativ einfach, relevante von unnötigen Dateien zu unterscheiden. Die Ablieferung war nach den fünf übernommenen Projekten gegliedert, hier wurde jedes Projekt einzeln betrachtet: Worum geht es in diesem Projekt? Was war hier die Aufgabe des LGRB? Welches Produkt ist am Ende entstanden? In den meisten Fällen entstanden Geodaten in Form von Shapefiles, also Dateien, die einen bestimmten Umriss auf einer Karte darstellen und mit weiteren Dateien verknüpft sind, die hierzu eine Beschreibung liefern. In einigen Fällen kamen zu den Geodaten noch Karten im PDF-Format und einmal eine

umfangreiche Tabelle im Excelformat. Damit sind also die unbedingt relevanten und erhaltenswerten Informationen identifiziert.

In der Ablieferung waren aber noch wesentlich mehr Dateien in verschiedenen Formaten enthalten: Zu den „eigentlichen Kerndateien“ kamen noch weitere Dokumente, die für das Verständnis der Geodaten notwendig sind, zum Beispiel „readme“-Textdateien, in denen das betreffende Projekt und die zugehörigen Dateien beschrieben werden sowie Projektberichte des LGRB. Auch diese Unterlagen wurden Teil des AIP.

In manchen Ordnern waren Kopien von bereits an anderer Stelle vorhandenen Dateien enthalten, zum Beispiel die Karten im JPG-Format, sowie zahlreiche kleine Hilfsdateien, die aber für die Darstellung der Karten und Shapefiles nicht notwendig sind. Diese wurden entfernt und nicht mit in das digitale Magazin übernommen.

Die Ordnerstruktur der Ablieferung wurde grundlegend überarbeitet. Die einzelnen Dateien lagen in einer Vielzahl von Ordnern mit teilweise bis zu sechs Unterordnern sowie mit nicht unmittelbar verständlichen Ordnernamen. Nach der Bearbeitung des Zugangs gibt es für jedes übernommene Projekt im Wesentlichen zwei verschiedene Ordner: Einen mit den Geodaten oder Karten im PDF-Format und ein weiterer mit der zugehörigen Dokumentation. Bei den Ordnernamen wurden unverständliche Kürzel aufgelöst, so dass der Inhalt nun unmittelbar zu erkennen ist. Am Ende blieben 2.209 Dateien in 92 Ordnern, wobei die Ordner Ebenen auf höchstens vier reduziert wurden.

Fazit

Die Umsetzung eines kleinen Ausschnitts des OAIS-Modells – der Übergang vom SIP zum AIP – bedeutet in der Praxis eine Menge Arbeit, nämlich die sorgfältige Überlegung, welche Teile einer Ablieferung wirklich relevant sind und erhalten werden sollen und können.

Gleichzeitig liegt hier ein großer Unterschied zur klassischen analogen Archivierung. Dort kann man davon ausgehen, dass eine Ablieferung einer Behörde nur die bereits bewerteten und als archivwürdig eingestuften Unterlagen enthält und dass diese einigermaßen gut geordnet sind (zum Beispiel als durchnummerierte Aktenbände mit einer Ablieferungsliste). Nur in den seltensten Fällen, am ehesten bei Nachlässen, erfolgt eine Neuordnung bis hin zum Einzelblatt und das auch erst bei einer späteren Verzeichnung.

Bei der digitalen Langzeitarchivierung ist es hingegen notwendig, sich bald nach der Übernahme intensiv mit einer Ablieferung zu beschäftigen; sie kann nicht „erst mal ins Magazin eingelagert werden“, wie das bei einer analogen Ablieferung meist ohne Verluste möglich ist.

Eine so intensive Beschäftigung ist sicher auch nicht bei jeder digitalen Übernahme notwendig. Erhält man zum Beispiel eine einfache Serie von PDF-Dokumenten, so können diese relativ schnell und einfach in ein digitales Magazin übernommen werden.

In anderen Fällen wird man möglicherweise davon ausgehen, dass eine Überlieferung derart wichtig ist, dass man jede auch noch so kleine und scheinbar unwichtige Hilfsdatei erhalten möchte. Doch auch in diesem Fall ist es je nach Speichersystem oft notwendig, Ordner anzulegen oder umzubenennen, um mit einer gewissen Menge an Dateien überhaupt umgehen und sie sinnvoll speichern zu können. In jedem Fall müssen digitale Daten nicht nur den Weg von der Behörde ins Archiv meistern, sondern auch innerhalb des Archivs den teilweise mühsamen Weg von der Ablieferung bis zu dauerhaften Aufbewahrung im digitalen Magazin.

Annette Riek, Landesarchiv Baden-Württemberg - Abt. Staatsarchiv Freiburg

Weitere Kurzartikel aus der Reihe "Thema in 500 Wörtern" finden sie auf www.langzeitarchivierung.de - der Webseite von **nestor – Kompetenznetzwerk Langzeitarchivierung**.