



Audiovisuelle Kulturgüter erhalten

Kolloquium: Datenkomprimierung bei Bild, Audio und Video

Colloque: Compression des données images, audio et vidéo

11. November 2009, 13:15-16:45 Uhr

Bern, Hauptgebäude der Universität, Kuppelsaal

(siehe <http://www.bau.unibe.ch/plaene/hgexwiunis.htm>)

Programm

1. Begrüssung

Georg Büchler, KOST

2. Einführung: Eine Begriffsklärung

- Komprimierung – Datenreduktion
- verlustfrei – verlustbehaftet
- *visually lossless*
- Digitalisate – originär digitale Daten

Martin Kaiser, KOST

3. Best practices im audiovisuellen Bereich

Kurt Deggeller, Memoriav

4. Fallstudien im Bildbereich

- a. Bild-Datenkomprimierung aus Sicht der Nutzer: Experimente und Befunde aus der Wissenschaftspraxis

*Ralf Stockmann, Göttinger Digitalisierungszentrum,
Staats- und Universitätsbibliothek Göttingen*

- Kaffeepause -

- b. La numérisation, une étape sensible à intégrer dans la préservation des collections photographiques

*Christophe Brandt,
Institut Suisse pour la conservation de la photographie*

- c. Qualitätskriterien digitaler Bilder im Langzeitarchiv

Peter Fornaro, Imaging&Media Lab, Universität Basel

5. Schlussdiskussion und Zusammenfassung der Erkenntnisse

Wir bitten um Anmeldung bis am 6. November an kolloquium@kost-ceco.ch.

Zur Ausgangslage

In ihrem Katalog archivtauglicher Dateiformate (<http://www.kost-ceco.ch/wiki/whelp/KaD/>) schloss sich die KOST dem archivarischen Konsens zur Datenkomprimierung an: Komprimierung wird grundsätzlich kritisch gesehen; verlustbehaftete Komprimierung grundsätzlich abgelehnt.

In den seitherigen Diskussionen zu Formatfragen, Digitalisierung und Komprimierung sind jedoch besonders im Bereich der Bildformate Veränderungen in dieser Position aufgetaucht. Diese konzentrieren sich auf zwei Fragestellungen:

- Bei Retrodigitalisierungen von Bildern, Codices, Urkunden etc. ist der Begriff "Verlust" besser zu klären. Ein erster, unwiederbringlicher Datenverlust findet nämlich beim Digitalisierungsvorgang statt. Darauf folgende Datenkomprimierungen beziehen sich immer auf den digitalen Master, nicht auf das analoge Original. Datenkomprimierung kann also nicht losgelöst von den Parametern der Digitalisierung und von entsprechenden *best practices* gesehen werden.
- In verschiedenen Digitalisierungsvorhaben der letzten Jahre wurde vom bisherigen archivarischen Grundsatzentscheid für TIFF abgewichen. Erwähnt seien die Überlegungen der KB in den Niederlanden sowie das Projekt MONOLITH des Staatsarchivs Zürich. Es zeigt sich, dass bei sehr grossen Datenmengen der Speicherplatzbedarf unkomprimierter TIFF-Dateien zusehends als Problem wahrgenommen wird.

Diese Fragestellungen nimmt die KOST zum Anlass, das Thema der Datenkomprimierung in einem Kolloquium genauer zu untersuchen. Die Bereiche Audio und Video sollen dabei summarisch vorgestellt werden. Anschliessend soll die Situation bei Bilddaten (besonders Digitalisaten) an Hand von Fallstudien beschrieben, diskutiert und analysiert werden. Schwarz-Weiss-Antworten werden keine erwartet. Ziel ist jedoch eine Identifikation der Entscheidkriterien.

Die KOST führt das Kolloquium in Zusammenarbeit mit Memoriav durch.

Situation actuelle

Dans son catalogue sur les formats de données d'archivage (<http://www.kost-ceco.ch/wiki/whelp/Cfa/>), le CECO a rallié le consensus archivistique relatif à la compression des données : d'une manière générale, il évalue la compression de manière critique, et la rejette lorsqu'elle est synonyme de perte de données.

Au cours des dernières discussions relatives aux formats, à la numérisation et à la compression, cette position a toutefois quelque peu évolué, notamment dans le domaine des formats d'images. Deux problématiques se dégagent :

- Il s'agit de définir plus précisément le terme de « perte » dans le cadre de la rétroumatisation d'images, de manuscrits et d'actes officiels, etc. Une première perte de données irrévocable a en effet lieu lors du processus de numérisation. La compression de données qui s'ensuit se fonde toujours sur le format numérique et non sur l'original analogique. On ne peut donc évaluer la compression des données sans prendre en compte les paramètres de la numérisation et les bonnes pratiques correspondantes.
- La solution archivistique en faveur du format TIFF qui prévalait n'a pas été systématiquement appliquée dans les projets de numérisation des dernières années. A titre d'exemple, on peut citer les réflexions de la Bibliothèque royale aux Pays-Bas ou le projet MONOLITH des Archives d'Etat de Zurich. L'expérience a montré que, pour les très grandes quantités de données, la place de stockage nécessaire aux données TIFF non comprimées constituait manifestement un problème.

Ces problématiques ont motivé le CECO à se pencher sur le thème de la compression des données en lui consacrant le présent colloque. Il est prévu de présenter un aperçu des domaines audio et vidéo puis de décrire, de discuter et d'analyser la situation dans le domaine des images (et notamment des données numériques primaires) en se fondant sur des études de cas. Il ne s'agit pas d'apporter une réponse claire mais de déterminer des critères de décision.

Ce colloque est organisé par le CECO en collaboration avec Memoriav.