

AudioVault

Erschliessung von Audioaufnahmen

Lösungsmöglichkeiten

Bei Aufnahmen von Ratsdebatten klaffen die Granularitäten der Primär- und Metadaten sichtbar auseinander. In der Regel sind detaillierte deskriptive Metadaten verfügbar (Granularität Redebeitrag), aber sie ermöglichen keinen präzisen Zugriff auf die Primärdaten, da die Aufzeichnungen in der Regel eine ganze Ratssitzung umfasst. Für dieses Problem sind grundsätzlich verschiedene Lösungen denkbar.

| | | |
|----|---------------|---|
| 1. | Einleitung | 1 |
| 2. | Vier Konzepte | 1 |
| 3. | Umsetzung | 3 |
| 4. | Empfehlung | 4 |

1. Einleitung

Wir haben im Archiv in der Regel zwei sehr unterschiedliche Gruppen von Audioaufnahmen. Das eine sind Audioaufnahmen, die als solche geschaffen wurden oder Aufzeichnungen entsprechender Veranstaltungen sind, also z.B. Musikaufnahme für die Tonträgerproduktion, Hörspiele, Konzerte etc. In diesen Fällen, wo der Audioaufnahme ein künstlerisches Ereignis zugrunde liegt, ist die Verzeichnung bzw. Beschreibung meist synchron mit der Aufzeichnung. So enthält ein Tonträger, z.B. eine Schallplatte, mehrere Musikstücke, und auf dem Umschlag sind diese Stücke verzeichnet und beschrieben. Zu dem hier vorliegenden Beziehung Ereignis – Aufzeichnung – Kopie gibt es aus dem Umfeld von Dublin Core detaillierte Modellierungsvorschläge.

Anders sieht es aus mit Audioaufnahmen von Ereignissen, die nicht primär zum Zweck ihrer Aufzeichnung stattgefunden haben. Ein klassisches Beispiel aus unserem Umfeld: die Kantonsratsdebatte. Hier stammen die Aufnahme und die beschreibenden Metadaten in der Regel aus unterschiedlichen Quellen und sind von unterschiedlicher Granularität. Die Tonaufnahme ist normalerweise ein fortlaufender Mitschnitt nach Sitzung und Tagen geordnet, die Erschliessung im Protokoll ist nach Traktanden und Redebeiträgen gestaltet, d.h. die gewünschte Granularität der Metadatenererschliessung ist feiner als die Granularität der primären Tondaten.

Grundsätzlich sind vier Konzepte denkbar, wie Audioaufnahmen in der richtigen Granularität erschlossen werden können. Das Ziel dabei ist es, dass jeder Verzeichnungs-

einheit ein identifizierbares und zugängliches Primärdokument entspricht, d. h. genauso wie wir es von Musikaufnahmen auch gewohnt sind.

2. Vier Konzepte

a) *Audiostream in die Granularität der Verzeichnung schneiden*

Der erste Ansatz besteht darin, den Audio-stream in kleinere Einheiten zu schneiden, welche den Verzeichnungseinheiten entsprechen (beispielsweise Geschäfte oder Redebeiträge, oder Radioaufnahmen in einzelne Sendebiträge).

Vorteile: Jeder granulären Verzeichnungseinheit entspricht ein eigenständiges Dokument, das sich in einer einzigen Datei manifestiert. Die Handhabung der Primärdaten entspricht so derjenigen der anderen Datenkategorien. Da die Dateien eine geringere Grösse haben, kann davon ausgegangen werden, dass sie einfacher handhabbar sind.

Nachteile: Durch das Schneiden des Audio-streams wird eine Veränderung und Interpretation der Originaldaten vorgenommen mit allen damit verbundenen Risiken für die Authentizität, Integrität und Verstehbarkeit der Daten.

Tools: Audacity, SoX, FFmpeg und VLC media player o. ä.

Empfehlung: Als Kompromiss, um von den Vorteilen des Schneidens zu profitieren, ohne die Risiken in Kauf nehmen zu müssen, empfiehlt die KOST die Erstellung von Gebrauchskopien. Dabei wird einerseits die originale Audiodatei unverändert archiviert; andererseits werden davon den Verzeichnungseinheiten entsprechende kleinere Dateien geschnitten. Da diese in einem solchen Konzept den Status von Gebrauchskopien haben, ist es möglich, dafür eine Datenkomprimierung zu wählen, weil das unkomprimierte Original weiterhin zugänglich bleibt.

b) Timecode-Referenz auf den Audiostream in den Metadaten

Der zweite Ansatz behält die Primärdateien in ihrer originalen Form bei. Im AIS wird zu jeder relevanten Verzeichnungseinheit nicht nur die Referenz auf die Primärdatei, sondern zudem eine Timecode-Referenzⁱ auf den richtigen Ort in der Primärdatei verzeichnet. Der Zugang zur Primärdatei ist also zweistufig insofern, als dass die gesamte Audiodatei ausgeliefert wird und dann zur richtigen Position in der Tonspur vorwärtsgespult wird. Es wird keine spezielle Software benötigt, ausser man möchte das cueing oder Vorspulen zur richtigen Stelle automatisierenⁱⁱ.

Vorteile: Die Originaldatei wird unverändert erhalten; es müssen überhaupt keine Eingriffe vorgenommen werden. Beim Zugriff auf die Primärdaten ist der grössere Zusammenhang jederzeit gegeben.

Nachteile: Der Zugriff auf die Primärdaten erfordert die Handhabung tendenziell grosser Dateien, und ohne spezielle Software muss der Benutzer meist mit einem Schiebeschalter selber zur entsprechenden Stelle vorspulen. Gängige Abspielsoftware unterstützt das Springen zu einem Timecode nur unzureichend.

Tools: VLC media player o. Ä.

Empfehlung: Diese Lösung verschiebt den Erschliessungsaufwand hin zur Benutzung. Vor dem Ingest kann allenfalls durch Stichproben geprüft werden, ob die richtigen Audiodateien abgeliefert wurden. Fehler bei der Zuordnung von Metadaten zu Audiodaten werden nicht erkannt. Dieses Vorgehen ist nur bei fehlenden Erschliessungsressourcen zu empfehlen.

c) Externes Cue Sheet zur Audiodatei

Eine weitere Möglichkeit besteht darin, die Timecode Metadaten in einem sogenannten Cue Sheet festzuhalten. Ein Cue Sheet ist eine Ascii-Datei in einer wohldefinierten Syntaxⁱⁱⁱ. Sie erlaubt es dem Abspielprogramm, aus einer grossen Audiodatei einzelne Tracks oder Musikstücke darzustellen. Auf die Tracks kann wie auf einzelne Audio-

dateien zugegriffen werden. Die Tracks können mit Metadaten genauer bezeichnet werden (Title, Creator, Comment).

Vorteile: Die Originaldatei wird unverändert erhalten; es müssen überhaupt keine Eingriffe vorgenommen werden. Beim Zugriff auf die Primärdaten ist der grössere Zusammenhang jederzeit gegeben. Im Abspielprogramm erscheint die Audiodatei als mehrere Tracks oder einzeln abspielbare Audiodateien die auch einzeln entsprechend bezeichnet werden können.

Nachteile: Der Zugriff auf die Primärdaten erfordert die Handhabung tendenziell grosser Dateien zusammen mit dem Cue Sheet. Die Abspielsoftware muss die Cue Sheet Syntax kennen und richtig interpretieren. Beim Ingest muss ein Cue Sheet erzeugt werden.

Tools: Audacity und VLC media player o. Ä.

Empfehlung: Diese Lösung kann aus archiverischer Sicht als optimal bezeichnet werden. Originaldaten werden nicht verändert; dennoch kann auf einzelne Teile des Tondokumentes direkt zugegriffen werden. Vor dem Ingest beim Anlegen des Cue Sheet, was bei Punkt eins dem Schneiden entspricht, wird die Synchronizität von Metadaten und Audiodaten verifiziert.

d) Cue-Points im Audio-Container

Eine weitere Möglichkeit ist, via Metadaten im Audiocontainer auf einzelne Teile eines Audiostroms zu verweisen. Im WAVE-File können dazu im *cue Chunk* sogenannte *cue points* gesetzt werden^{iv}, das sind Marker auf den Audiostrom. Zu diesen Markern kann mit entsprechender Software direkt gesprungen werden, oder die Marker können mit einem entsprechenden Label im Player angezeigt werden. Der Benutzer muss nicht zu einem bestimmten Timecode vorwärts spulen, sondern kann im Player direkt zum entsprechenden korrekt gesetzten Marker springen. Beim Setzen von *cue points* wird der Audiostrom nicht bearbeitet; alle notwendigen Ergänzungen werden in einen Metadaten-Chunk geschrieben.

Vorteile: Der Audiostrom der Originaldatei wird unverändert erhalten; es werden nur Ergänzungen im Metadatenteil des Containers

ⁱ In der Regel als mm:ss (Minuten-Sekunden) oder als mm:ss:ff (minute-second-frame) mit 75 frames pro Sekunde Audiodata, spezifiziert.

ⁱⁱ Die gesamte Datei wird ausgeliefert und das Abspielen beginnt beim entsprechenden Timecode. Ein Schneiden *on access* ist nicht realistisch.

ⁱⁱⁱ Ein *Cue Sheet* ist im Prinzip eine Art Gegenteil einer *Playlist*, welche einzelne Musikdateien in einen grösseren Abspielzusammenhang zusammenfasst. Das *Cue Sheet* zerlegt eine Musikdatei in mehrere Musikstücke. Das Konzept stammt

noch aus der Zeit der Umwandlung von Audio-CDs in Audiodateien. Eine Audiodateien enthält nach dem Rippen in der Regel mehrere Tracks (Musikstücke)

http://wiki.hydrogenaud.io/index.php?title=Cue_sheet „The official cue sheet specification is widely accepted to be Appendix A of the CDRWIN User's Guide“

<http://web.archive.org/web/20070221154246/http://www.goldenhawk.com/download/cdrwin.pdf>

^{iv} *Cue points* können auch zu einer *Playlist* zusammengefasst werden, darüber kann das Abspielen des Audiostroms in beliebiger Reihenfolge gesteuert werden. Diese Möglichkeit bietet aber für unser Problem keine Lösung.

vorgenommen werden. Beim Zugriff auf die Primärdaten ist der grössere Zusammenhang jederzeit gegeben. Die Marker sind in der Audiodatei inhärent und ohne Verweis aus den externen Metadaten im AIS lesbar.

Nachteile: Die Originaldatei wird verändert. Der Zugriff auf die Primärdaten erfordert die Handhabung tendenziell grosser Dateien, und es wird spezielle Software benötigt um die entsprechenden Marker anzuzeigen.

Tools: Traktor Pro, Serato DJ v

Empfehlung: Diese Lösung ist aus archivischer Sicht nicht zu empfehlen. Einerseits wird die Originaldatei verändert, und der Umstand, dass der Audioteil unverändert bleibt, ist nicht einfach zu verifizieren. Andererseits ist sowohl für das Setzen der Cue Points als auch zum Abspielen Spezialsoftware notwendig, und die Spezifikation des Inhalts eines „cue“ vi Chunks ist tendenziell produkteabhängig

3. Umsetzung

a) Zeitpunkt für Schneiden oder Timecode-Referenzsetzen

Der ideale Zeitpunkt für das Schneiden oder Timecode-Referenzen Setzen ist vor dem *Ingest* – es handelt sich dabei um eine eigentliche *Pre-Ingest*-Erschliessungsaufgabe. Die Aufgabe des Schneidens bzw. des Referenzsetzens kann in der Regel nicht auf die aktenbildende Stelle abgewälzt werden. Im Idealfall werden jedoch *Timecode*-Referenzen bereits beim Erzeugen der Datei gesetzt oder festgehalten. Das Archiv soll mit der aktenbildenden Stelle zusammen abklären, ob dies möglich ist und in welcher Form dies geschehen soll.

Der Zeitpunkt ist auch deshalb ideal, weil beim Visionieren vii und Anhören der Audiodateien und dem Vergleichen mit den mitgelieferten Metadaten schnell Unstimmigkeiten oder fehlerhafte Ablieferungen erkannt werden.

b) Vorgehen in einem Fallbeispiel

Hier wollen wir kurz das Vorgehen an einer Audiodatei erläutern, die in diesem Projekt vom Staatsarchiv Uri zur Verfügung gestellt wurde. Es handelt sich dabei um eine Audioproduktion der Dezembersession 2007 des

v *Traktor Pro*, *Serato DJ* sind Softwarelösungen aus dem Bereich Musik sampeln, produzieren oder arrangieren. Beim Sampeln werden die die *Cue Points* in einer Playlist zu beliebig hintereinander oder wiederholt abgespielten Teilen arrangiert

vi Der Leerschlag hinter „cue“ ist Absicht, ein *RIFF Chunk Tag* besteht aus vier Zeichen.

Urner Landrates. Die Metadaten liegen in einem Excelsheet vor, für uns hier interessant: Sprecher, Laufzeit und Geschäft:

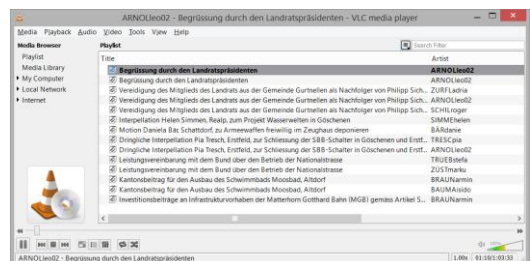
Metadaten in 2007-24.xlsx

Aus dieser Metadatenliste lässt sich auf einfache Art ein *Cue Sheet* zur entsprechenden Audiodatei, hier 2007-24.mp3 viii, erstellen. In einem grösseren Übernahmезusammenhang muss dieser Schritt natürlich automatisiert werden.

Im *Cue Sheet* verzeichnet sehen wir jeden der 14 Redebeiträge mit den Metadaten *GENRE*, *TITLE* und *PERFORMER*. *INDEX* schlussendlich verweist auf die Position in der Audiodatei. Wir könnten diese ganzen Informationen als eingebettete Metadaten in die Audiodatei schreiben; dabei würden wir die Datei aber verändern, wir hätten also eine eigentlich unnötige Migration beim *Pre-Ingest*.

Cue Sheet 2007-24.cue

Mit einem Mediaplayer, hier VLC Media Player, können wir jetzt 2007-24.cue öffnen und direkt zu den einzelnen Redebeiträgen springen, wie wenn es sich um einzelne Audiodateien in einer Playlist handeln würde. Damit lässt sich der Inhalt der Audiodatei beim *Pre-Ingest* schnell verifizieren und im Benutzungsfall haben wir den gewünschten Komfort.

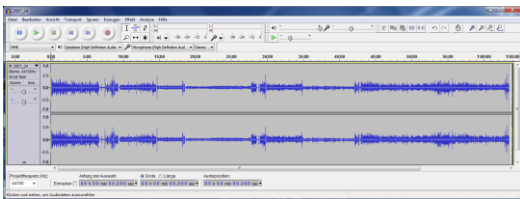


vii Wir können hier von Visionieren sprechen, weil uns Programme wie *Audacity* die Audiodatei auch grafisch aufbereitet darstellen können, siehe Beispiel weiter unten.

viii Der Dateiname leitet sich aus der ursprünglichen CD-Signatur ab. Bei der Migration von Audio-CD zu Audiodateien wurde MP3 als Archivformat gewählt, für unser Beispiel ist das Dateiformat ohne Bedeutung.

VLC media player zeigt den Inhalt der Audiodatei 2007-24.mp3

Betrachten wir die Tonspur optisch, z.B. mit Audacity, können wir die einzelnen Segmente ebenfalls erkennen.



Audacity zeigt den Inhalt der Audiodatei 2007-24.mp3 optisch aufbereitet

Neben inhaltlichen Metadaten, die in der Regel zusätzlich zu den Audiodateien abgeliefert werden, können wir auch eine Reihe technischer Metadaten direkt aus der Audiodatei extrahieren. Sinn und Zweck dieser Metadatenextraktion^{ix} ist dabei, dass die Erschliessung nach technischen Aspekten schneller erfolgen kann. Wir extrahieren in unserem Fallbeispiel diese Metadaten mit dem Programm MedialInfo in eine MPEG-7-Datei.

```
<?xml version="1.0" encoding="UTF-8" ?>
<mpeg7 xmlns="urn:mpeg:mpeg7:schema:2004" http://standards.iso.org/IT/ISO/TC4306/Information/2004/urn:mpeg:mpeg7:schema:2004"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="urn:mpeg:mpeg7:schema:2004 http://standards.iso.org/IT/ISO/TC4306/Information/2004/urn:mpeg:mpeg7:schema:2004"
  >
  <mpeg7:Audio>
    <mpeg7:AudioFormatCR:2001:1F>
      <mpeg7:Name and lang="en">MPEG-7 Audio Layer III</mpeg7:Name and lang>
      <mpeg7:SampleRate>44100</mpeg7:SampleRate>
      <mpeg7:Channels>2</mpeg7:Channels>
      <mpeg7:MediaTime>
        <mpeg7:MediaDuration>PT13M42S231.19441000</mpeg7:MediaDuration>
      </mpeg7:MediaTime>
    </mpeg7:AudioFormatCR:2001:1F>
  </mpeg7:Audio>
  <mpeg7:MediaInformation>
    <mpeg7:MediaTitle>
      <mpeg7:MediaTitleText>MPEG-7 Audio Layer III</mpeg7:MediaTitleText>
    </mpeg7:MediaTitle>
  </mpeg7:MediaInformation>
</mpeg7:Audio>
```

MPEG-7 Sidecar Datei 2007_24.MP3.mpeg7.xml

c) Ausgestaltung von SIP und AIP

Die Audiodateien sollen in der Form und Granularität an das Archiv abgeliefert werden, in welcher sie erzeugt wurden. Im optimalen Fall kann das Archiv Vorgaben zu

Qualität, Format und Granularität der Aufzeichnung machen, im Fallbeispiel ist aber klar, dass aus technischen Gründen nur eine Aufzeichnung einer ganzen Sitzung integral in Frage kommt und nicht Aufnahmen einzelner Redebeiträge.

Die Granularität des AIP soll sich an den generellen Vorgaben des Archivs ausrichten; es ist nicht sinnvoll, an Audiodaten hier spezielle Maßstäbe anzulegen. Falls Gebrauchskopien geschnitten und diese im AIP abgelegt werden, sollte klar ersichtlich sein, dass es sich hier nicht um die originalen Ablieferungsdaten handelt. Ebenso bei Sidecar Files (Cue Sheet oder MPEG-7), wobei hier durch die Natur dieser Dateien ihr Metacharakter ersichtlich ist.

In unserem Fallbeispiel haben wir pro Sitzung eine Audiodatei. Die Landratssitzung bezeichnet hier im AIP die Dossierebene, die Erschliessung erfolgt innerhalb der Sitzung bis auf Ebene Geschäft bzw. Redebeitrag, so wie wir es von den Protokollen ebenfalls kennen. In einem Dossier „Landratssitzung“ befinden sich dann also eine Audiodatei, ein Cue Sheet und eine MPEG-7 Datei. Im Falle, dass Zugangskopien erstellt worden sind, kommen diese noch dazu.

4. Empfehlung

Fassen wir die ganzen Überlegungen noch zu einer Empfehlung zusammen:

- Das Erstellen eines *Cue Sheet* unter Beibehalten der originalen Audiodatei scheint die Lösung mit den meisten Vorteilen zu sein.
- Das Erstellen von Gebrauchskopien in der gewünschten Granularität, unter Beibehaltung der originalen Audiodatei im AIP, scheint die zweitbeste Lösung.
- Die drittbeste Lösung ist wohl das alleinige Abstellen auf den *Timecode* in den abgelieferten Metadaten, die aus der Ablieferung ins Findmittel übernommen werden.

^{ix} Technische Metadaten sind inhärenter Teil der Audiodatei und werden im Audiocontainer festgehalten, sie gehen nicht

verloren oder werden auch nicht verändert. Die Extraktion dient nur dem vereinfachten Zugriff.