

## Aufgabe 09: Erstellen einer Manuskript-Beschreibung

Christof Schöch

DARIAH-DE

Creative Commons Attribution 3.0 (CC BY-SA).

Das vorliegende Dokument beruht weitgehend auf den folgenden Materialien: James Cummings, Renée Baalen, Ylva Berglund-Prytz: "An Introduction to XML and the Text Encoding Initiative", exercise 5, Digital.Humanities@Oxford Summer School 2012, University of Oxford, 2012, <http://digital.humanities.ox.ac.uk/dhoxss/2012/>, CC-BY-SA.

### Überblick

#### 1. Inhalte

In dieser Übung werden wir für ein Dokument aus der Wilfred Owen-Sammlung eine Manuskript-Beschreibung erstellen. Das Dokument enthält bereits einen recht komplexen `teiHeader`, den Sie weiter bearbeiten werden.

#### 2. Lernziele

Nach dieser Übung sollten Sie folgendes gelernt haben:

- \* Eine vorhandene, einfache Manuskript-Beschreibung strukturieren
- \* Die grundlegenden Kategorien der Manuskript-Beschreibung in TEI kennen
- \* Mehr Erfahrung damit, einen komplexen Header zu bearbeiten

#### 3. Materialien

- \* `Manuskript-Start.xml`
- \* `Manuskript-MsDesc.xml`

### Aufgabenstellung

#### 1. Vorbereitung

Öffnen Sie die Datei `Manuskript-Start.xml` mit Oxygen. In dem Bereich von Oxygen auf der linken Seite (mit dem Titel Outline), können Sie sich über die Struktur des Dokuments orientieren:

- \* Öffnen Sie die kleinen blauen Symbole mit einem Klick auf den `Rechts-Pfeil`, um weitere Hierarchie-Stufen zu sehen.
- \* Hinweis: Wird dieser Bereich nicht angezeigt? Sie können ihn über das Menü `Windows > Show View > Outline` aktivieren; dort können Sie auch weitere Ansichten aktivieren.
- \* Finden Sie das Ende der `sourceDesc` und gehen Sie im Hauptfenster mit dem Cursor dorthin.

\* Öffnen Sie außerdem die Datei `Manuskript-msDesc.xml` ebenfalls mit Oxygen.

## 2. Eine einfache Manuskript-Beschreibung hinzufügen

Wir werden diesem Header nun eine Manuskript-Beschreibung hinzufügen. Das „Original“ ist übrigens hier zu finden, wo es selbstverständlich aus einer XML-Datei für die Anzeige generiert wurde: <http://www.oucs.ox.ac.uk/ww1lit/collections/document/5243> Eine minimal strukturierte Variante dieser Beschreibung beinhaltet die Datei `Manuskript-msDesc.xml`. Sie besteht aus einem `msIdentifier`, der als einziges Element einer `msDesc` in TEI obligatorisch ist, sowie aus mehreren Prosa-Absätzen, wie man sie vielleicht in einem traditionellen, print-basierten Findmittel vorfinden würde. Eine solche Manuskript-Beschreibung ist für differenzierte Suchanfragen bspw. in Katalogen nicht geeignet. Unsere Übung ist insofern recht „realistisch“, als genau solche Umwandlungs-Prozeduren bei der Digitalisierung bzw. Retrokonversion von Findmitteln oft notwendig ist.

\* Haben Sie den Cursor schon direkt nach dem schließenden Tag `</listOrg>` platziert? Hier fügen wir nun die komplette `msDesc` aus der Datei `Manuskript-msDesc.xml` ein.

\* Ist Ihr Dokument weiterhin valide? Sehr gut!

\* Speichern Sie das Dokument unter einem neuen Namen ab.

## 3. Das Element `msIdentifier` ergänzen

Jetzt werden wir das Element `msIdentifier` ergänzen. Einige der Schritte dabei werden weniger genau beschrieben sein, als das bisher (teils) der Fall war, da Sie ja jetzt schon einige Erfahrung mit TEI und Oxygen haben.

\* Sehen Sie, dass der erste Absatz der `msDesc` vor allem Informationen enthält, die sich auf den Aufbewahrungsort des Manuskripts beziehen? Diese Informationen sind Teil der Identifikation eines Manuskripts.

\* Nehmen Sie diese Informationen also nach und nach aus dem Absatz heraus und fügen Sie sie in strukturierter Form dem `msIdentifier` hinzu.

\* Nutzen Sie die Hilfen von Oxygen: wenn Sie an einer beliebigen Stelle „<“ tippen, schlägt Ihnen Oxygen die dort zulässigen Elemente vor.

\* Fügen Sie nach und nach die relevanten Informationen ein, nach dem öffnenden `msIdentifier` Tag, aber noch vor dem `msName` Element.

\* Folgende Elemente sollten Sie verwenden; die Reihenfolge ist vorgeschrieben, wie Sie merken werden:

\*

\* `country`

\* `region`

\* `settlement` (= z.B. Stadt, Dorf, etc.)

- \* `institution`
- \* `repository`
- \* `collection`
- \* `<idno type="folio">`
- \* `altIdentifier`

\* Hinweis: Falls Sie nicht mehr sicher sind, welche Elemente für welche Informationen gedacht sind, können Sie selbst in den TEI Guidelines nachschauen; der hier folgende Link führt zum Inhaltsverzeichnis der Guidelines: <http://kurzlink.de/tei-guidelines>. Finden Sie dort das Kapitel 10.4 The Manuscript Identifier.

\* Wenn Sie fertig sind sollten Sie sicherstellen, dass Ihr Dokument weiterhin valide ist (`Strg+Shift+V`)

\* Außerdem können Sie nun den ersten Absatz `p` der `msDesc` löschen.

\* Und vergessen Sie nicht, das Dokument abzuspeichern.

#### 4. Informationen zu Inhalten mit `msContents`

Der zweite Absatz beinhaltet Informationen, die für die Beschreibung des intellektuellen Inhaltes des Manuskripts, also den Bereich `msContents` relevant ist. Dieser Bereich fasst den Inhalt des Manuskripts zusammen; man kann dies mit dem Element `msItem` auch in einer Form abbilden, die bspw. die Informationen aus einem Inhaltsverzeichnis wiedergibt.

\* Benennen Sie den zweiten Absatz `p` in `msContent` um. Das Dokument ist nun nicht mehr valide. (Logisch, oder?)

\* Zunächst fügen wir eine Zusammenfassung des Inhalts ein. Fügen Sie hierfür ein Element `summary` ein und kopieren Sie den passenden Teil des zweiten Absatzes dort ein (von „Letter to Leslie Gunston / The Wrestlers“ bis „...Collected Letters“).

\* Das Brief-Manuskript enthält im Grunde zwei inhaltliche Teile; den Brieftext selbst und den Teil eines Entwurfs für das Gedicht The Wrestlers. Daher unterscheiden wir zwei "manuscript items". Fügen Sie also zunächst einen neuen `msItem` ein.

\* Innerhalb des ersten `msItem` geben Sie nun die relevanten Informationen in separaten Elementen an, das heißt einen Titel, den Autor, und die Sprache.

\* Geben Sie dem ersten Item den Titel Letter to Leslie Gunston und benutzen Sie hierfür das Element `title`.

\* Löschen Sie „Authored by“ und kodieren Sie Wilfred Owen (1893-1918) mit dem Element `author`.

\* Kodieren Sie English mit dem Element `textLang`, und geben Sie als Attribut „mainLang“ mit dem Wert „en“ an („en“ ist der ISO-Code für „English“)

\* Fügen Sie dem Element `author` ebenfalls in Attribut, „ref“ hinzu; wählen Sie aus der Liste der „xml:ids“, die Oxygen Ihnen vorschlägt, das passende aus (Tipp: Initialen). Dieses Attribut verweist auf den Eintrag mit dieser `xml:id` in

der Liste der Personen, weiter oben im Header (`listPerson` in `sourceDesc`).

- \* Fügen Sie nun einen weiteren `msItem` mit dem gleichen Aufbau ein, und kodieren Sie entsprechend den "Partial draft of 'The Wrestlers'" mit Titel, Autor und Sprache.

Das `msContents` Element sollte nun folgendermaßen aussehen: `msContents <summary> Letter To Leslie Gunston / The Wrestlers. Talks about forthcoming address to the 'Field Club'. Includes a partial draft of 'The Wrestlers'. This is letter no. 535 in Ed. 'Wilfred Owen Collected Letters'. </summary> <msItem> <title> Letter to Leslie Gunston </title> <author ref="#WO"> Wilfred Owen (1893-1918) </author> <textLang mainLang="en"> English </textLang> </msItem> <msItem> <title>Partial draft of 'The Wrestlers'</title> <author ref="#WO">Wilfred Owen (1893-1918)</author> <textLang mainLang="en">English</textLang> </msItem></msContents>`

## 5. Eine `physDesc` einfügen

Der nächste Absatz beinhaltet vor allem Informationen über die materiellen Aspekte des Manuskripts. Diese können wir in eine strukturierte `physDesc` verwandeln.

- \* Fügen Sie zunächst direkt nach dem schließenden Tag von `</msContents>` das Element `physDesc` ein
- \* Direkt innerhalb, d.h. direkt nach dem öffnenden `physDesc` Tag, fügen Sie das Element `objectDesc` ein
- \* innerhalb dessen wiederum ein Element `supportDesc`
- \* Innerhalb dessen `supportDesc` wiederum das Element `support`
- \* Als Inhalt des `support` Elements ist der Satz „A single folio of paper in the collection as ff504 recto and verso“ geeignet; kopieren Sie ihn dort ein.

Ein solcher Satz kann mit strukturierenden oder klassifizierenden Elementen angereichert werden

- \* Kodieren Sie Papier als `material`

Zusatzinformationen können auch auf der Ebene der Elemente hinzugefügt werden:

- \* Bspw. können Sie dem Element `supportDesc` das Attribut „material“ mit dem Wert „paper“ geben
- \* Oder Sie können dem Element `objectDesc` das Attribut „form“ mit dem Wert „folio“ geben (Übrigens: in beschreibender Prosa kürzt man gerne „Attribut xyz“ als `@xyz` ab (nicht in der Kodierung!))

Nach dem schließenden `</supportDesc>` Tag können wir nun noch eine `layoutDesc` hinzufügen.

- \* Oberstes Element: `layoutDesc`
- \* Darin: `layout`
- \* Der Satz: „Written full width as a single column, with approximately 20 lines per page“ gehört hier hinein.
- \* Strukturierte Informationen für das Element `layout`:

- \* @columns mit Wert „1“
- \* @writtenLines mit Wert „20“

Damit ist die `objectDesc` abgeschlossen. Nach dem schließenden `objectDesc` können wir nun aber noch eine `handDesc` einfügen.

- \* `handDesc` einfügen, mit `@hands`, Wert „1“ (es gibt nur eine Schrift/Hand in dem Manuskript)
- \* Innerhalb von `handDesc: handNote` einfügen
- \* Der restliche Text, „Written in Wilfred Owen’s hand in pen.“ kann hier als Inhalt eingefügt werden.
- \* Außerdem können Sie wieder einige Informationen explizit kodieren
  - \* „Wilfred Owen“ als `persName`
  - \* `persName` mit `@ref` und Wert „#WO“, um wiederum auf die einheitliche Personenbeschreibung weiter oben im Header zu verweisen

Hinweis: Da dies nun eine recht komplexe `physDesc` geworden ist, haben wir am Ende dieses Dokuments sozusagen die „Lösung“ für diesen Teil der Aufgabe eingefügt; erst nachschauen, wenn Sie fertig sind!

## 6. Die Geschichte des Manuskripts mit `history` angeben

Das Element `history` dient dafür, den Ursprung, die Provenienz und die Akquise des Manuskripts festzuhalten, mit den Elementen `origin`, `provenance`, und `acquisition`. In diesem Fall liegt immerhin etwas Information über die Entstehung des Manuskripts vor, die wir nun kodieren sollten.

- \* Benennen Sie den vorletzten Absatz `p` um in `history`.
- \* Fügen Sie ein Element `origin` ein
- \* Fügen Sie dort den Satz „This letter was written by Wilfred Owen in July 1917 at Craiglockhart War Hospital.“ ein.

Einige Informationen können auch hier wieder explizit kodiert werden:

- \* „July 1917“ ist das Entstehungsdatum, markieren Sie es mit dem speziellen Element `origDate`.
- \* Damit diese Angabe auch wirklich maschinen-lesbar ist, sollte Sie standardisiert kodiert werden: geben Sie dem Element `origDate` ein `@when` mit dem Wert „1917-07“
- \* Den Ort „Craiglockhart War Hospital“ können Sie als `origPlace` kodieren
- \* Hier ist der explizite Verweis auf den Eintrag für diesen Ort möglich: mit `@ref` und einem Wert von „#craiglockhart“

## 7. Weitere Informationen mit dem Element `additional` festhalten

Am Ende der `msDesc`, also nach dem schließenden `</history>` Tag, können wir zusätzliche Informationen festhalten, bspw. administrative Informationen (mit `adminInfo`), Forschungsliteratur über das Manuskript (mit `listBibl`), oder weitere

(digitale oder analoge) Repräsentationen des Manuskripts (mit `surrogates`). In unserem Falle haben wir Informationen über eine Abbildung an einem anderen Ort.

- \* Fügen Sie also das Element `additional` ein
- \* Innerhalb dessen: `surrogates`
- \* Fügen Sie den letzten Satz hier ein
- \* Der Link muss als Element `<ptr/>` mit `@target` und der URL als Wert eingetragen werden
- \* Ist das Dokument nun wieder valide? Dann haben Sie auf jeden Fall einiges richtig gemacht!

## 8. Speichern

Speichern Sie Ihre Arbeit, aber überprüfen Sie nun noch einmal:

- \* Ist das Dokument valide?
- \* Haben sie es ordentlich formatiert? (Automatisch mit `Strg+Shift+F`)

Geschafft! Das ist das Ende der Übung.

### Anhang: Die kodierte `physDesc`

```
<physDesc> <objectDesc form="folio"> <supportDesc material="paper">
<support>A single folio of <material>paper</material> in the collection as ff504
recto and verso</support> </supportDesc> <layoutDesc> <layout columns="1"
writtenLines="20">Written full width as a single column, with approximately 20
lines per page</layout> </layoutDesc> </objectDesc> <handDesc hands="1">
<handNote>Written in <persName ref="#WO">Wilfred Owen's</persName> hand
in pen.</handNote> </handDesc> </physDesc>
```