

# Referat 08: Dokument-zentriertes Kodieren mit TEI

**DARIAH-DE Tutorial *Textedition mit TEI***

Redaktion: Christof Schöch (Univ. Würzburg)

Version 1.0, 02/2014

Grundlage der Folien: Elena Pierazzo, King's College London

Lizenz: [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) (BY-SA)



# Gliederung

1. Vorüberlegungen

2. Dokument-zentriertes Kodieren: "zones"

# 1. Vorüberlegungen

# TEI und Manuskripte

## Etablierte Komponenten von TEI

- Modul 11, "transcr", für relativ "saubere" Manuskripte
- Modul 10, "mdDesc", für die Beschreibung von Manuskripten
- Modul 12, "apparatus", für Emulation von print-basierten kritischen Apparaten
- <figure> in Modul 14, "Tables...", um Transkriptionen mit digitalen Faksimiles zu verknüpfen

Das klingt gut, aber...

# TEI und Manuskripte (2)

...aber diese Module wurden entwickelt:

- für (idealisierte) mittelalterliche Manuskripte
- (zumindest teilweise) im Geiste der Print-Ära
- (zumindest teilweise) im Geiste des "einen/besten" Textes
- Text als "ordered hierarchy of content objects" (OHCO, Allen Renear)

Grenzen der vorhandenen Mechanismen:

- Text als Spur eines dynamischen Schreibprozesses (*critique génétique*)
- Text in seiner räumlichen (statt logischen) Struktur
- Enge Verbindung von Text und Bild
- Varianten in ihrem Eigenwert

# Neu! genetic/documentary editing

## Cf. Elenas Vortrag: Arbeitsgruppe und proposal

- hohe Priorität für "genetic editions"
- Revision der "Manuscript Description"

## Drei Komponenten

- Dokument-Kodierung: topographische Sicht auf das Dokument
- Erweiterte Transkription: verfeinerte Mechanismen
- Genetische Kodierung: prozessuale Sicht auf Schreibvorgang

## Neues Kapitel 11

- Seit November 2011 teil von TEI

# Das neue Modul 11: drei Teile

## Dokument-Kodierung

- räumlich-topographische Sicht auf das Faksimile
- `<surface>`, `<zone>` und `<line>`
- `<layout>`, `<space>`

## Erweiterte Transkription

- Erweiterte Element für mikrostrukturelle, genetische Phänomene
- `<damage>`, `<supplied>` (mit `<gap>` und `<unclear>`)
- `<restore>`, `<retrace>`, `<redo>`, `<undo>`, `@instant`
- `<transpose>`, `<metamark>`

## Genetische Kodierung

- Schreibprozess mit „revision campaigns“
- `<listChange>` und `<change>`

# 1. Dokument-Kodierung

## <surface>

- Seite, Stein, anderes "text-tragendes Objekt"
- kann beinhalten: <zone> und <line>
- hat räumliche Koordinaten

## <zone>

- beliebiger, definierter Bereich auf einer Seite etc.
- Zonen können überlappen (via Koordinaten; cf. overlap problem)
- kann beinhalten: <line> oder Text
- hat räumliche Koordinaten; Rechtecke oder Polygone!

## <line>

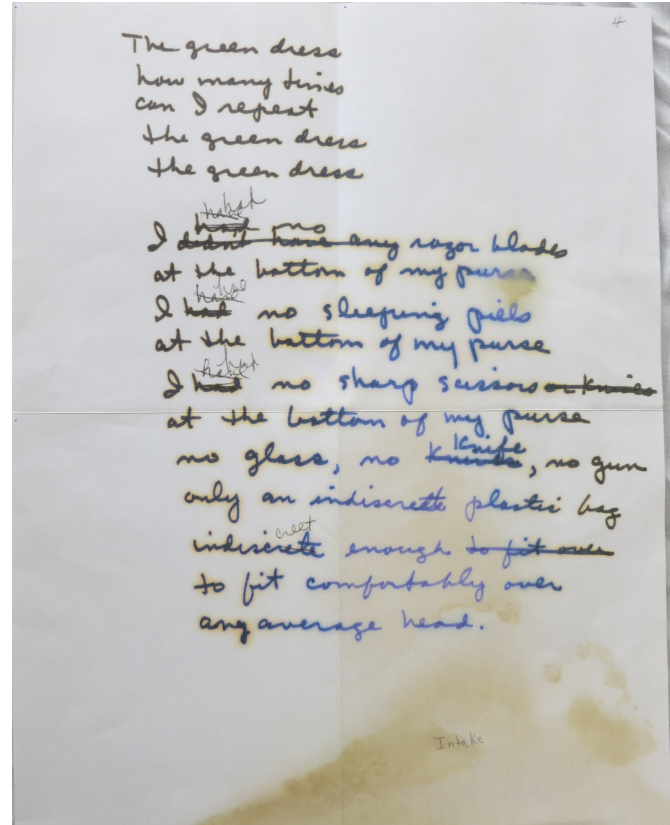
- eine Zeile Text (keine metrischer Vers o.ä.)
- enth. in: <surface> oder <zone>; kann enth.: Text oder <zone> (!)
- hat keine Koordinaten



# Beispiel: <surface> und <zone>

Welche Oberfläche, welche Zonen/Bereiche könnte man hier unterscheiden?

- Surface: die gesamte Seite
- Zone 1: Nummer oben rechts
- Zone 2: Textblock oben
- Zone 3: Textblock unten
- Zone 4: Titel unten rechts
- evtl. Zone 5: Verfärbung (Polygon)



# Beispiel: <surface> und <zone>

```
<sourceDoc>
  <surfaceGrp>
    <surface ulx="0" uly="0" lrx="600" lry="900"
      xml:id="surface-1">
      <graphic url="Applebroog.jpg" height="900px"
        width="600px" />
      <zone xml:id="zone-1" ulx="100" uly="100" lrx="350"
        lry="250">
        <line>How many times</line>
        <!-- [weitere Zeilen] --!>
      </zone>
      <zone xml:id="zone-2" ulx="100" uly="260" lrx="500"
        lry="650">
        <line>I had no...</line>
        <!-- [weitere Zeilen] --!>
      </zone>
    </surface>
  </surfaceGrp>
</sourceDoc>
```

## 2. Erweiterte Transkription

Siehe das Input-Referat “Input 04b – Transkription von Primärquellen und Manuskripten (erweitert)”

- Erläuterungen und Beispiel dort

# 3. Genetische Kodierung

## Schreibprozess / „revision campaigns“

- `<listChange>` - beinhaltet eine Liste von Bearbeitungsphasen
- `<change>` - für die Beschreibung der einzelnen Bearbeitungsphasen

### `<listChange>`

- im `<teiHeader>`, innerhalb der `<profileDesc>` und `<creation>`
- aufeinander folgende oder ungeordnete Bearbeitungsphasen

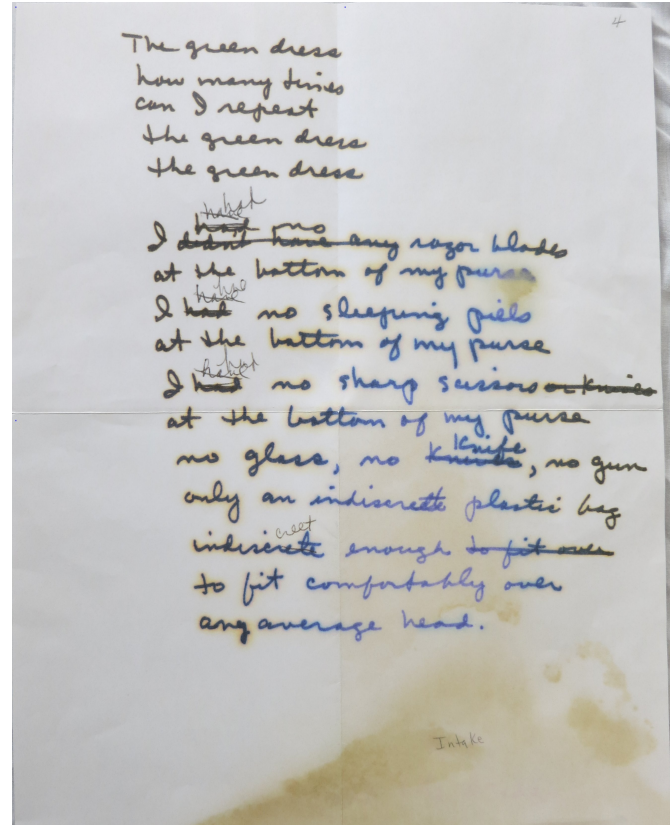
### `<change>`

- innerhalb eines `<listChange>`, mit `@xml:id` zur Identifikation
- `<lines>` oder `<zones>` können einem `<change>` zugeordnet werden
- einzelne Modifikationen (`<del>`, `<add>`, etc.) ebenfalls
- Zuordnung jeweils mit `@change` und `xml:id`

# Beispiel: <change>

Welche Phasen könnte man hier unterscheiden?

- Zuerst: Erste Niederschrift
- Dann, recht bald: Korrekturen mit Filzstift
- Dann, evtl. etwas später: Korrekturen mit Bleistift



# Beispiel: <listChange>

```
<profileDesc>
  <creation>
    <listChange ordered="true" n="writing">
      <change xml:id="stage-1">Text wird von Hand mit
        Filz einmal geschrieben.</change>
      <change xml:id="stage-2">Text wird von Hand mit
        Filz korrigiert: mehrere Ersetzungen.</change>
      <change xml:id="stage-3"> Weitere Korrekturen, mit
        Bleistift.</change>
    </listChange>
  </creation>
</profileDesc>
```