



## Vorbemerkung

Dieses Dokument richtet sich an Personen, die sowohl ein gewisses Verständnis von XML und seinem Umfeld als auch von TEI und der Handschriftenkatalogisierung haben. Ausgangspunkt sind die Richtlinien zu TEI P5, Version 0.3.1 vom Januar 2006. (Die Version 0.4.1 vom 4. Juli 2006 scheint für die Handschriftenbeschreibung keine größeren Änderungen zu bringen.)

Es wird lediglich das *msDescription*-Element mit seinen Kindern behandelt, nicht aber das TEI-Gesamtdokument.

Ziel ist, eine möglichst klare Anleitung für die Auszeichnung von Katalogisaten zu bieten, damit sie anschließend mit den bereitgestellten Skripten nach HTML oder PDF gewandelt werden können.

Dabei werden – von den datumswertigen Attributen abgesehen – stets die Möglichkeiten von TEI P5 eingeschränkt, nicht erweitert. Diese Zusatzbedingungen werden am einfachsten mit einem Schematron-Schema überprüft. Entsprechend werden bei fehlenden empfohlenen Angaben Warnungen ausgegeben, während bei fehlenden verpflichtenden Vorgaben die Mitteilung erfolgt, daß die Beschreibung nicht konform ist.

## Inhalt

- [Vorbemerkung](#)
- 1. [Allgemeine Beschreibung einer Handschrift](#)
  - 1.1. [Aufbau eines Katalogisats](#)
  - 1.2. [msDescription](#)
  - 1.3. [msIdentifier](#)
  - 1.4. [head](#)
  - 1.5. [msContents](#)
    - 1.5.1. [msItem](#)
  - 1.6. [physDesc](#)
    - 1.6.1. [objectDesc](#)
      - 1.6.1.1. [supportDesc](#)
    - 1.6.1.2. [layoutDesc](#)
      - 1.6.2. [handDesc](#)
  - 1.6.3. [musicNotation](#)
  - 1.6.4. [decoDesc](#)
  - 1.6.5. [additions](#)
  - 1.6.6. [bindingDesc](#)
  - 1.6.7. [accMat](#)
    - 1.7. [history](#)
  - 1.8. [additional](#)
  - 1.9. [msPart](#)
- 2. [Diskussion einiger Elemente und Attribute, die an verschiedenen Stellen auftreten können](#)
  - 2.1. [altIdentifier](#)
  - 2.2. [author](#)
  - 2.3. [q und andere Zitatelemente \(explicit, finalRubric, incipit, rubric\)](#)
  - 2.4. [listBibl](#)

- 2.5. [n](#)
- 2.6. [note](#)
- 2.7. [ref](#)
- 2.8. [rs \(reference string\)](#)
- 2.9. [type](#)
- 3. [Allgemeine Konventionen](#)
  - 3.1. [Zeichensetzung](#)
- 3.2. [Optik vs. Inhalt](#)
- 3.3. [Leerräume](#)
- 3.4. [Reihenfolge](#)
- 4. [Informationskategorien und wie sie Elementen zugeordnet werden](#)
  - 4.1. [Suchrelevante Begriffe und Registereinträge](#)
  - 4.2. [Signaturen anderer Handschriften](#)

# 1. Allgemeine Beschreibung einer Handschrift

## 1.1. Aufbau eines Katalogisats

Prinzipiell besteht ein Katalogisat aus dem *msDescription*-Element, das in der angegebenen Reihenfolge folgende Kinder hat:

- *msIdentifier*, das genau einmal vorkommen muß: enthält alle Angaben zur Identifikation der gesamten Handschrift;
- *head*, das genau einmal vorkommen soll: enthält allgemeine Titelangaben;
- *msContents*, das genau einmal vorkommen kann: enthält Angaben zu den vorkommenden Texten;
- *physDesc*, das genau einmal vorkommen kann: enthält die äußere Beschreibung;
- *history*, das genau einmal vorkommen kann: enthält die Diskussion von Entstehung, Besitzgeschichte und letzter Erwerbung der Handschrift;
- *additional*, das genau einmal vorkommen kann: enthält Literaturangaben zur gesamten Handschrift, zu Digitalisaten und Verwaltungsangaben einschließlich Entstehung des Katalogisats;
- *msPart*, das beliebig oft vorkommen kann: sie enthalten die Beschreibung von Faszikeln und Fragmenten.

## 1.2. *msDescription*

Es sind nur die globalen Attribute erlaubt. Empfohlen wird die Verwendung von *xml:lang*; auch *xml:id* bietet sich an. Bei letzterem kann in aller Regel nicht die Signatur verwendet werden, da sie meist keine XML-konforme ID darstellt. Darüberhinaus sollte der Wert so gewählt werden, daß seine Eindeutigkeit auch beim Zusammenstellen *beliebiger* Katalogisate gewährleistet ist. Daher kann dieses Attribut nur dann sicher verwendet werden, wenn sowohl durch bibliotheksinterne Vorgaben als auch – etwa durch Vorschaltung eines Institutionskürzels – diese Eindeutigkeit übergreifend ermöglicht wird. Es ist aber auch darauf zu achten, daß bei der Umwandlung der Signaturen in XML-konforme IDs die Eindeutigkeit gewährleistet bleibt. Bei dem Sprachattribut, das insbesondere für die Einbindung der korrekten Trennmuster Hilfestellung bieten kann, müssen die aktuellen Werte von RFC 3066 genommen werden, die die Unterscheidung zwischen alter und neuer Rechtschreibung erlauben. (Für die Übersetzung registerrelevanter Auszeichnungen reicht die entsprechende Kurzform beim Element. Im allgemeinen kann bei einer so spezialisierten Materie davon ausgegangen werden, daß die wesentlichen Informationen auch in einer fremden Sprache gefunden werden können.)

### 1.3. *msIdentifier*

Um eine gute Austauschbarkeit der Katalogisate ohne Zusatzaufwand zu ermöglichen, wird die alleinige Verwendung von *msName* ausgeschlossen. Daher sieht der Aufbau folgendermaßen aus:

- *country*, kann einmal vorkommen
- *region*, kann einmal vorkommen
- *settlement*, muß einmal vorkommen: enthält die Angabe zur Gemeinde
- *institution*, kann einmal vorkommen: enthält Angaben zu der Institution, zu der die besitzende Institution gehört
- *repository*, muß einmal vorkommen: enthält den Namen der besitzenden Institution
- *collection*, kann einmal vorkommen: enthält die Angabe, zu welcher Sammlung die Handschrift gehört
- *idno*, muß einmal vorkommen: enthält die vollständige, aktuell gültige Signatur
- *msName*, kann beliebig oft – auch gemischt mit *altIdentifier* – vorkommen: enthält Namen, unter denen die Handschrift bekannt ist (aber keine früheren Signaturen)
- *altIdentifier*, kann beliebig oft – auch gemischt mit *altIdentifier* – auftreten: vgl. den Abschnitt [altIdentifier](#)

### 1.4. *head*

Dieses optionale Element sollte genau einmal vorkommen und zumindest den Katalogisattitel der Handschrift aufnehmen. Normalerweise sollen Autorenangaben nur in *msContents* gemacht werden, allerdings wird ein eventuell vorhandener *author* ausgewertet. Wird kein *title* angegeben so wird der Katalogisattitel aus allen *author*- und *title*-Elementen von *msContents* erzeugt.

Wird die Schlagzeile eines Katalogisats angegeben, so ist sie in einem *note* mit dem *type*, das den Wert *caption* erhält, erfaßt. Es ist darauf zu achten, daß alle Angaben der Schlagzeile auch anderweitig erfaßt und ausgezeichnet sind. Wird die Schlagzeile nicht explizit angegeben, so wird sie aus allen *material*-Nachkommen von *support*, *measure@type='leavesCount'*-Nachfahren von *extent*, *measure@type='leavesSize'*-Nachfahren von *layout* (falls vorhanden wird auch *dimension* ausgewertet), *origPlace* aus *head* oder *origin* und *origDate* entsprechend aufgebaut; dies führt aber besonders bei Faszikelhandschriften zu unschönen Ergebnissen.

Werden Entstehungszeit oder -ort nicht weiter diskutiert, so sind sie mit *origDate* bzw. *origPlace* hier anzugeben.

Da *head* ein allgemein verwendbares Element ist – etwa auch als Überschrift von Listen oder Abschnitten – sollte entsprechend der Funktion für deutsche Handschriftenbeschreibungen in der Regel der Titel und die Schlagzeile angegeben werden und darf außer Autor, Entstehungszeit oder -ort kein anderes Element verwendet werden.

### 1.5. *msContents*

Neben den globalen Attributen steht noch *class* zur Verfügung, dessen Werte jedoch an anderer Stelle (im *teiHeader* in *textClass*) spezifiziert sein müssen.

Die Struktur ist:

- *summary*, kann genau einmal vorkommen: enthält eine zusammenfassende Wertung des Inhalts;
- *textLang*, kann genau einmal vorkommen: enthält eine Diskussion der in der Handschrift verwendeten Sprachen, die in den Attributen *mainLang* und *otherLangs* nach RFC 3066 (oder einem Nachfolger) recherchierbar kodiert werden können.

- *msItem*, muß mindestens einmal vorkommen: enthält die Diskussion eines Textes.

Das stärker strukturierte *msItemStruct* wird hier nicht weiter berücksichtigt.

### 1.5.1. msItem

Folgende Elemente sind erlaubt (die Reihenfolge nach *locus* alphabetisch):

- *locus*, das immer an erster Stelle stehen muß und genau einmal vorkommen soll: Seitenreferenz;
- *author*
- *editor*
- *bibl*
- *colophon*
- *decoNote*
- *explicit*
- *filiation*
- *finalRubric*
- *incipit*
- *listBibl*
- *msItem*
- *note*
- *q*
- *rubric*
- *respStmt*
- *textLang*
- *title*

Ein *msItem* kann die obigen Elemente – abgesehen von *locus*, das höchstens einmal vorkommen darf und an erster Stelle stehen muß – in beliebiger Zahl und Reihenfolge enthalten. Für die Übersichtlichkeit sollte aber stets ein *title*-Element nach *locus* stehen. Auch ist darauf zu achten, daß jeder Text ein eigenes *msItem* bekommt, andererseits aber eine zu tiefe Verschachtelung vermieden wird.

*decoNote* sollte nur in Ausnahmefällen verwendet werden und die Information ist stets auch in *decoDesc* festzuhalten.

## 1.6. physDesc

Die äußere Beschreibung gliedert sich folgendermaßen, wobei jedes Element genau einmal vorkommen kann (die nach der Einbandbeschreibung mögliche Diskussion von Siegeln wird weggelassen):

- *objectDesc*: beschreibt die physischen Aspekte der Handschrift;
- *handDesc*: beschreibt die Schrift und die Schreiber;
- *musicNotation*: gibt Hinweise zu verwendeten Musiknotationen;
- *decoDesc*: beschreibt den Buchschmuck;
- *additions*: beschreibt spätere Hinzufügungen;
- *bindingDesc*: beschreibt den Einband;
- *accMat*: beschreibt beiliegende, aber nicht direkt zur Handschrift gehörige Materialien, wie eingelegte Blätter.

### 1.6.1. objectDesc

Die physischen Aspekte gliedern sich in zwei Gruppen, wobei jede einmal vorkommen darf (das nach den Richtlinien erlaubte Beschreiben in Absätzen wird – wie auch meist sonst – nicht unterstützt). Die

Reihenfolge ist festgelegt:

- *supportDesc*: enthält die verschiedenen Beschreibungen des Materials (Art, Umfang, Follierung, Kollationierung, Zustand);
- *layoutDesc*: enthält alle Angaben zur Gestaltung der Seiten.

#### 1.6.1.1. supportDesc

Die Beschreibung des Materials gliedert sich in die folgenden Punkte, wobei die Reihenfolge festgelegt ist und jede Einheit höchstens einmal vorkommen kann:

- *support*: enthält die Beschreibung der Materialart;
- *extent*: enthält die Beschreibung der Blattzahl;
- *foliation*: enthält die Angaben zur Follierung einschließlich deren Entstehung;
- *collation*: enthält die Kollationierung;
- *condition*: enthält Angaben zum Zustand der Handschrift bzw. von Teilen einschließlich von Leerseiten.

#### 1.6.1.2. layoutDesc

Dieses Element enthält mindestens ein *layout*-Element. Für jeden Aspekt sollte ein eigenes *layout*-Element verwendet werden. Da zur Zeit die meisten Attribute eine ganze Zahl sein müssen und andererseits sie nicht besonders rechercherelevant sind, kann auf sie verzichtet werden.

#### 1.6.2. handDesc

Für jeden behandelten Aspekt zu den Schreibern oder der Schrift sollte eine eigene *handNote* verwendet werden. Ob das Attribut *hands*, das entweder die genaue Zahl der Hände oder den Wert „many“ erhält, den Aufwand wert ist, sei dahingestellt.

#### 1.6.3. musicNotation

Enthält absatzweise Bemerkungen zu verwendeten musikalischen Aufzeichnungssystemen.

#### 1.6.4. decoDesc

Für jeden behandelten Aspekt des Buchschmucks sollte ein eigenes *decoNote*-Element verwendet werden. Für die Attribute *type* und *subtype* ist eine geeignete Kategorisierung vorzunehmen.

#### 1.6.5. additions

Hier sind absatzweise spätere Hinzufügungen festzuhalten.

#### 1.6.6. bindingDesc

Um die Einheitlichkeit zu wahren, ist auch bei der Beschreibung nur eines Einbandes das *binding*-Element zu verwenden. Werden verschiedene Einbände beschrieben, sollte für jeden ein eigenes *binding* genommen werden.

### 1.6.7. accMat

Hier sind – in Absätzen – nur die später beigelegten Materialien zu behandeln; für Fragmente ist jeweils ein *msPart* zu verwenden.

### 1.7. history

Hier wird zwischen Entstehung, Besitzgeschichte und Erwerbung unterschieden:

- *origin*, kann höchstens einmal vorkommen: enthält in Absätzen die Diskussion der Entstehung der Handschrift;
- *provenance*, kann beliebig oft vorkommen: sollte pro Absatz jeweils einen besitzgeschichtlichen Zusammenhang enthalten;
- *acquisition*, kann einmal vorkommen: enthält die Erwerbung durch die gegenwärtig besitzende Institution in Absätzen.

### 1.8. additional

Jedes der folgenden Elemente kann in der angegebenen Reihenfolge einmal vorkommen:

- *adminInfo*: Hier werden nicht nur Angaben zum Katalogisat und der Verfügbarkeit der Handschrift sowie deren konservatorische Behandlung erfaßt, sondern es können auch allgemeine Notizen festgehalten werden;
- *surrogates*: Hier finden sich die Einträge zu Sekundärformen (etwa Digitalisaten);
- *listBibl*: Hier wird die Literatur zur gesamten Handschrift gesammelt; im untergeordneten *head*-Element sollte entweder „Zur gesamten Handschrift“ stehen oder für *listBibl* in *additional* die verpflichtende Verwendung von *head* aufgehoben werden.

### 1.9. msPart

Dieses Element ist völlig analog zu *msDescription* aufgebaut mit dem einzigen Unterschied, daß statt *msIdentifier* das Element *altIdentifier* verwendet werden muß. Um unterscheiden zu können, ob es sich um einen Faszikel oder ein Fragment handelt, ist im verpflichtend vorgeschriebenen *altIdentifier* das Attribut *type* mit dem Wert „frag“ bzw. „fasc“ zu verwenden.

## 2. Diskussion einiger Elemente und Attribute, die an verschiedenen Stellen auftreten können

Hier sollen nicht die Erläuterungen der Richtlinien wiederholt, sondern nur Ergänzungen festgehalten werden.

### 2.1. altIdentifier

Enthält frühere Signaturen analog zum *msIdentifier*, wobei aber *idno* die frühere Signatur enthält sowie *msName* nicht erlaubt und alle anderen Elemente optional sind.

## 2.2. *author*

Dieses Element darf in *msItem* nur den Autor des besprochenen Textes beinhalten, in *bibl* und ähnlichen Elementen natürlich den Autor des Werkes. Wird eine nähere Kategorisierung gewünscht – etwa Person, Institution oder Pseudozuordnung – so ist das *rs*-Element als Kind mit einer entsprechenden Typisierung zu verwenden.

## 2.3. *q* und andere Zitatelemente (*explicit*, *finalRubric*, *incipit*, *rubric*)

Um eine leichte und eindeutige Zuordnung zu ermöglichen, ist eine eventuell vorhandene Blattangabe *innerhalb* des Zitats festzuhalten. Da es anzustreben ist, inhaltliche Auszeichnungen nicht durch typographische Kodierungen anzugeben, sollten stets die entsprechenden Auszeichnungen wie *sic* oder *gap* verwendet werden. Soll unbedingt genau die erfaßte Form wiedergegeben werden, ist das globale *rend*-Attribut mit dem Wert „static“ zu verwenden.

## 2.4. *listBibl*

In TEI P5 sind viele Elemente erlaubt, doch sind für die Handschriftenkatalogisierung lediglich ein oder mehrere *head*-Elemente gefolgt von beliebig vielen *bibl*-, *biblStruct*-, *biblFull*- oder *biblItem*-Elementen vorgesehen. Das *head*-Element dient zur Charakterisierung der in dieser Liste zusammengefaßten bibliographischen Angaben.

## 2.5. *n*

Dieses globale Attribut sollte nur in seltenen Ausnahmefällen numerische Werte im engeren Sinne wie Seitenzahlen enthalten (etwa bei *pb*), sondern nur zur gliedernden oder kategorisierenden Unterstützung etwa von Texten dienen.

## 2.6. *note*

Hier sollte im Hinblick auf die strikte Trennung von Inhalt und Form immer eine Typisierung mittels des *type*-Attributs verwendet werden, in *head* etwa „caption“. Das *rend*-Attribut dient nur zur Platzierungsangabe bei der Wiedergabe von erfaßten Texten, sollte aber auch im Rahmen der Retrokonversion nicht benutzt werden. Um die Verarbeitung zu erleichtern, ist auf eine Ankerlösung zugunsten der direkten Eingabe zu verzichten.

## 2.7. *ref*

Dieses Element erlaubt mit dem *target*-Attribut den Verweis an eine andere Stelle, sei es im selben Dokument oder in einem anderen. Der Verweis erfolgt stets auf eine URI (Uniform Resource Identifier, vgl. RFC 2396) bzw. eine Liste von URIs. Steht eine kanonische Referenzierung eines Werkes zur Verfügung, kann statt *target* das Attribut *cRef* verwendet werden.

Es ist wünschenswert, für häufig vorkommende Verweisarten Standards festzulegen. Zu denken ist an Kurzliteratur, Standardausgaben oder Verweise zu anderen Katalogisaten.

## 2.8. *rs* (*reference string*)

Oft genügt ein Element nicht, um den Inhalt hinreichend klar zu kennzeichnen. So kann ein Autor nicht

nur eine reale Person, sondern auch eine Institution oder ein Pseudoautor sein. Auch kann bei Abkürzungen nicht das *author*-Element benutzt werden. In diesen Fällen kann das *rs*-Element mit einem entsprechenden *type*-Attribut dienen. Es sollte generell darauf geachtet werden, daß das allgemeinere Elemente das speziellere umschließt, etwa erst die Referenzierung einer Person, dann deren Namen und dann Vor- und Nachname.

## 2.9. *type*

Dieses häufig erlaubte Attribut sollte stets englisch- oder lateinischsprachige Wortbildungen bzw. Akronyme als Wert erhalten. Sie können auch künstlich zusammengesetzt, müssen aber stets selbsterklärend sein. Da sie für gleiche Zwecke bei verschiedenen Elementen vorkommen können (man denke beispielsweise an *measure* und *dimensions*), sollten sie nicht möglichen Elementnamen entsprechen. Sinnvolle Ausnahmen sind dort gegeben, wo ein allgemeineres Element vom Typ eines spezielleren ist, etwa ein Zitat vom Typ *Initium*.

- *author*: Autor
- *columns*: Zahl der Spalten
- *explicit*: Explicit
- *fasc*: Faszikel
- *frag*: Fragment
- *holdings*: Vorbesitz
- *incipit*: *Initium*
- *leavesCount*: Zahl der Blätter
- *leavesSize*: Format der Blätter
- *realPerson*: Reale Person
- *pseudo*: Fälschlich zugeordnete Person
- *vulgata*: Bibel
- *writtenLines*: Zahl der Zeilen
- *writtenArea*: Schriftspiegel

Für *decoNote* ist eine eigene Klassifikation anzustreben.

## 3. Allgemeine Konventionen

- Zeichensetzung
- Optik vs. Inhalt
- Leerräume
- Reihenfolge

### 3.1. *Zeichensetzung*

Nachdem in der aktuellen Überarbeitung der Richtlinien die verpflichtenden Absatzelemente – soweit ich zumindest sehe – nur noch an den Stellen vorkommen, wo sie tatsächlich ganze Absätze beinhalten, gilt folgende Regel: Alle Kinder von *p* behalten die Zeichensetzung, bei allen anderen wird sie bei der Ausgabe erzeugt. Im Einzelfall sind diese Vorgaben so zu ergänzen, daß das *p*-Element als Kind ausgeschlossen oder erzwungen wird. Für einige Elemente – wie den Zitatformen – kann durch Setzen des Attributs *rend* auf den Wert „static“ angezeigt werden, daß der Bearbeiter das Erscheinungsbild festgelegt hat.

### 3.2. Optik vs. Inhalt

Es sollen keine optischen Auszeichnungen – wie Ellipsen für Auslassungen oder ein geklammertes Ausrufezeichen für erkannte Fehler – verwendet werden, sondern Elemente. Insbesondere gilt:

- Ergänzungen: *add* mit dem Attribut *hand*
- Lücke: *gap* mit den Attributen *reason* (Grund), *hand* (falls einer Person zuschreibbar), *agent* (Ursache wie Radierung, Tintenfleck) und *extent* (Ausmaß)
- Unsicher: *unclear* mit den Attributen *reason*, *hand* und *agent*.

Dies ist auch bei speziellen Zitaten zu berücksichtigen, die – falls das Attribut *rend* nicht den Wert „static“ hat – automatisch um entsprechende Zeichen ergänzt werden.

### 3.3. Leerräume

Alle Inhalte müssen so erfaßt sein, daß nach *normalize-space()* (Entfernung aller führenden und abschließenden Leerräume sowie Reduzierung mehrfacher Leerräume auf einen) noch alles korrekt ist. Für die Lesbarkeit sollten neue Zeilen insbesondere mit Einrückungen nur vorkommen, wenn zwischen den jeweiligen Elementen kein zu enger Lesezusammenhang besteht.

### 3.4. Reihenfolge

Bei der Ausgabe werden dir Blöcke entsprechend der deutschen Gepflogenheiten ausgegeben. Innerhalb von *msItem* wird die Abfolge der Elemente in der Datei jedoch beibehalten.

## 4. Informationskategorien und wie sie Elementen zugeordnet werden

### 4.1. Suchrelevante Begriffe und Registerinträge

Als elektronisches Dokument kann jede Auszeichnung (auch in beliebiger Kombination) recherchiert werden. Um etwas zu finden ist also zuerst sicherzustellen, daß die Information in regularisierter Form ausgezeichnet ist. Für die Webpräsentation genügt hierfür eine Liste von recherchierbaren Elementen bzw. Attributen mit Angabe des Kontextes. Letzterer ist beispielsweise erforderlich, um den Bezug von *origDate* klarzustellen. Kommt der Begriff im Element nicht in der registerrelevanten Form vor und steht auch kein entsprechendes Attribut zur Verfügung, so ist als Kind *reg* zu verwenden. Es wird nur für die Recherche und das Register ausgewertet.

### 4.2. Signaturen anderer Handschriften

Hier sei die HAB-Lösung empfohlen.

