

Digitalisat

- Seiten TIFF
- Seiten JPG
- Figures

raw text

gültiges XML

XML gültig nach Schema

XML vorbereitet für scholarly process

XML mit GIS-Daten

Daten für ein Dokument

- evtl. Seiten-XMLs
- Index-Daten

(später nicht mehr benutzte Objekte)

(wird jede Seite versioniert?)

Digitalisierung

- scannen
- identifizieren
- TIFF, JPG hochladen

Transkription

- Special Instructions
- schicken mit Begleitbrief
- Fragen beantworten
- überprüfen
- wiki pflegen

XML-Erstellung

pre-processing

- closing tags: <pb/>
- Vorkorrektur
- escape sequences
- ** unknown characters
- Normalisierung (best practices)

Grobstruktur

- * Metadaten
- * <div> numerieren
- * <pb> links (Konkordanz Seitennummer - Dateiname)
- ergänze <lb/>
- * <s>
- floats (interne links)
- Normalisierung (best practices) --> MPDL authoring?

Feinstruktur

- * <reg>
- * Bindestriche
- <foreign>
- * <num>, <var>
- * MathML
- * figures, tables nachbearbeiten
- Wort- und Satzgrenzen in Chinesisch
- * errata?
- * IDs
- Normalisierung (best practices)

GIS-Daten

- GIS-Metadaten
- Zeit, Person, Ort, Ereignis

Josefs System

- Indexierung
- ...

Falks System

Lucene

Donatus

"Donatus für Chinesisch"

minimale Satzanalyse (z.B für <reg>, hyphens)

Orts-Datenbanken

Personen-Datenbanken lat., chin., etc.

Metadaten: authority file

Rolle von Arboreal?

- Korrekturen am XML ermöglichen (z.B. zwei <reg> zusammen)

"fiktive Services"

DESspecs

Schema

Service zum Korrigieren von Transkriptionsfehlern: spontan benutzbar, evtl. erst als Kommentar

maintenance services: Unicode 5.2, Schema 1.1 (auch: editorische Entsch.)

unterschiedlich starke Automatisierung

Services auch standalone

scholarly process: Korrekturen insbesondere in den Schritten mit * (und Interaktion bei unknown characters!)

diary-System für Veränderungen im Text

"Playlists" für workflows? (filter chains, extended)

Ersatz für wiki, genauso einfach

Aktualisierung der Website (bzw. Website als Frontend von eSciDoc?)

systemweites und

- Speichern des Weges zusätzlich zum Ergebnis
- Sprache zum Notieren von Abhängigkeiten?
- intelligentes merging (Services-aware)

Texte bekanntmachen (Verzeichnisse, wikipedia)

Benachrichtigungssystem

- Annotationen
- technische Korrekturen
 - Seiten einfügen/löschen
 - Seiten vertauschen
- nachträgliche editorische Entscheidungen
 - einzelner Text
 - global
- "out of sync" mit anderem Objekt
- object needs maintenance