

Zur Differenzierung der XML-Objekte bei Nietzsche

2000-12-06 (I)

XMLDIFF

In diesem Papier werden die wichtigsten Objekttypen und Regeln in verbaler Form beschrieben. Bezüglich der formalen Strukturierung wird auf die Dokumentation XML1 verwiesen.

Gliederung:

- 1 Differenzierung der Objekte
- 2 Globale Strukturelemente
- 3 Markierungen nicht-struktureller Art (mit Beginn und Ende)
- 4 Markierungen nicht-Struktureller Art (lokal)

1. Es werden folgende Objekte unterschieden:

| | | |
|----------|----------------|---|
| <faq> | FAQ-Objekt | <i>Elektronische Darstellung von Fragen und zugehörigen Antworten mit den Teilen:</i> |
| | <frage> | Frage |
| | <antwort> | Antwort |
| <quelle> | Quellen-Objekt | <i>Beschreibung einer physikalischen Quelle (etwa eines Buches, eines Aufsatzes). Zusammenhängende Quellen können durch ein nicht-physikalisches Quellenobjekt aufeinander bezogen werden (dies gilt etwa für den Zarathustra, der in vier physikalisch differenzierten Teilen erschienen ist). (Physikalische) Aufsätze als Quellen-Objekt verweisen dabei ggf. mit Heft- und Seitenangaben auf das nicht-physikalische <i>Quellen-Objekt</i> der 'Zeitschrift' (es sind also bei einem <i>Aufsatz</i> aus einer Zeitschrift, einem Sammelband, einer Festschrift mindestens zwei Quellenobjekte zu erfassen). AutorInnen, Titel- wie Untertitelangaben, Jahrgangangaben, Publikationsorte, Verlagshinweise ... werden <i>nur</i> im Quellen-Objekt verzeichnet. Dies gilt auch für <i>Abstracts</i> und <i>Annotationen</i>. Vor- und Nachworte, Inhaltsverzeichnisse, Register sind dagegen (neben den reinen Texten) als E-Werk-Objekte zu erfassen, dort entsprechend zu <i>typisieren</i> und mit vor- oder nachfolgenden E-Werk-Objekten entsprechend zu verketteten. Gibt es Quellen-Objekte mit mehreren <i>Auflagen</i> bzw. <i>Ausgaben</i>, kann eine Referenzverweis <Grundlage> auf ein ausführlich beschriebenes Quellenobjekt gegeben werden, um Wiederholungen von Angaben zu vermeiden. In diesen Fällen werden nur die spezifischen Daten des Quellenobjekts (u.a. eine eigenständige Ident-Kennung)</i> |

| | | |
|----------|-----------------|--|
| | | angegeben, die übrigen gelten als vom Grundlagen-Objekt 'vererbt'. |
| <ewerk> | E-Werk-Objekt | <p><i>Elektronisch präsentiertes Werk</i>, das <i>nicht</i> vom Typ 'Brief' ist. Es muss zu jedem elektronisch präsentierten Werk eine Beschreibung des Quellen-Objekts existieren. Es wird in beiden Richtungen verwiesen (durch Angabe der Ident-Kennung).</p> <p>Quellen-Objekte können auf <i>kein, eines</i> oder <i>mehrere</i> E-Werk-Objekte bezogen werden. beziehen sich mehrere E-Werk-Objekte auf <i>ein</i> Quellen-Objekt, so ist ihre Anordnung durch eine entsprechende <i>Vorgänger-Nachfolger-Verkettung</i> der E-Werk-Objekte untereinander beschrieben. In diesem Fall verweist das Quellen-Objekt auf das erste Element dieser Kette von E-Werk-Objekten, die einzelnen E-Werk-Objekte auf die Positionen / Orte im Quellen-Objekt (z.B. bei Kapiteln als E-Werk-Objekt auf die Beginn-Seite des Kapitels in dem verwendeten Quellen-Objekt).</p> |
| <brief> | Brief-Objekt | <p><i>Elektronisch präsentiertes 'Brief'-Objekt</i>. Auch hier wird - soweit <u>vorhanden</u> - auf eine entsprechende Quelle (Briefsammlung) verwiesen (die dementsprechend als Quelle erfasst sein muss). Allerdings erfolgt dieser Verweis nur aus Richtung des einzelnen Briefes auf den Ort der physikalischen Sammlung.</p> <p>Briefe haben - wie üblich - einen Absender, einen Adressaten, ein Datum, einen Absende- und Zielort. Dementsprechend müssen zumindest diese Namen als Personen-Objekte und Orts-Objekte (ggf. nur als 'solche' beschrieben sein, um über ihre Ident-Kennung verwiesen werden zu können. Zu den Brief-Objekten rechnen auch - entsprechend typisiert - Postkarten, Telegramme, jegliche Informationen oder intendierte Informationen (also auch nicht abgeschickte Brief-Fragmente, Entwürfe), die einen oder mehrere spezifische(n) Adressaten haben.</p> |
| <person> | Personen-Objekt | <p>Hier werden <i>alle im Informationssystem vorkommenden Personen</i> beschrieben. Dazu gehören folgende durch ein entsprechendes Merkmal unterschiedene 'Typen':</p> |
| | 'bezugsperson' | Bezugsperson(en) des Informationssystems (hier: Alfred Gulden; Friedrich Nietzsche) |
| | 'hist' | historische Person (darunter Verwandte, Freunde, Bekannte der Bezugsperson, aber auch Künstler, Wissenschaftler mit entsprechendem Bezug ...). Falls eine Person in einer Doppelrolle auftritt (z.B. Lou Salomé), wird sie nur an einer Stelle (meist 'hist') beschrieben. |
| | 'expert' | Expertin / Experte (etwa ein fachlicher Kommentator, Autor einer wiss. Arbeit ...) |
| | 'designer' | GestalterIn, Mitarbeiterin bei der Entwicklung des Informationssystems |
| | 'psonst' | z.Zt. nur als Merkposten |
| <ort> | Orts-Objekt | Beschreibung eines (meist) realen Ortes, einer (meist) realen Lokalität (z.B. eines Berges ...). |

| | | |
|------------|-----------------|--|
| <comment> | Kommentar-Obj. | Ein nur in elektronischer Form verfügbarer Kommentar-Text. |
| <link> | Link-Objekt | Links zu Adressen im Internet mit Bezug zur Bezugsperson (hier: Nietzsche) |
| <register> | Schlagwort-Obj. | (Register-Eintrag) Nur für Schlagwörter, um zwischen ihnen Beziehungen (Relationen) herzustellen (z.B. Synonymie, Ober- / Unterbegriff, Übersetzung) oder (Gewichtungs)-Markierungen anzubringen ...). |
| <glossar> | Glossar-Objekt | Eintrag zur Begriffs-Erklärung (und Schreibweisen) |
| <proverb> | Wendung-Objekt | Meist fremdsprachliche Wendung mit Übersetzung |

2. *Hinzu kommen folgende 'globalen' Strukturelemente:*

| | | |
|------------|----------------|---|
| <datum1> | Einzeldatum | ggf. mit Angaben zur Zuverlässigkeit und Präzision (mit weiterer Strukturierung: Jahr, Monat, Tag ...) |
| <datum2> | Von-bis-Datum | Zur Angabe eines ununterbrochenen Zeitraums, ggf. mit Angaben zur Zuverlässigkeit und Präzision (mit jeweils weiterer Strukturierung: Jahr, Monat, Tag ...) |
| <bezug> | Bezugshinweise | zu der Bezugsperson. Sie können im Quellen-Objekt oder im Orts- und Personen-Objekt auftreten. |
| <design> | GestalterIn | Hinweise zum / zur verantwortlichen Gestalterin + Datumsangaben ... |
| <sw> | Schlagwort | Zur Erstellung eines Schlagwort-Registers (zunächst 'frei', schrittweise 'systematisiert') |
| <citation> | Zitat | Beginn und Ende eines <i>Zitats</i> (nicht jedoch einer Wörtlichen Rede) - mit Hinweis auf die zitierte 'Person' ... bzw. die zitierte E-Werk-Stelle |
| <speech> | Wörtliche Rede | Beginn und Ende einer Wörtlichen Rede- mit Hinweis auf die sprechende 'Person' bzw. Kunst-Person ... - und ggf. den Typ 'Normal', 'Monolog', 'innerer Monolog' und dem / der Angeredeten. |

3. *Folgende Markierungen nicht-struktureller Art mit Anfangs- und Endekennung werden vergeben, wenn dies nicht aus anderen Merkmalen hervorgeht (bzw. generiert werden kann):*

| | | |
|-----------|---------------|--|
| <text> | Text | Beginn und Ende einer Textsequenz |
| <hands> | Handschrift | Beginn und Ende einer handschriftlichen Einfügung |
| <sprache> | Sprache | Beginn und Ende einer Textsprache (mit Angabe dieser Sprache) |
| <probl> | Problem | Beginn und Ende einer problematischen Stelle (Lesbarkeit) |
| <bold> | halbfett | Beginn und Ende einer mit 'halbfett' ausgezeichneten Zeichenkette |
| <under> | unterstrichen | Beginn und Ende einer mit 'unterstrichen' ausgezeichneten Zeichenkette |
| <curs> | kursiv | Beginn und Ende einer mit 'kursiv' ausgezeichneten Zeichenkette |
| <hoch> | hochgestellt | Beginn und Ende einer mit 'hochgestellt' ausgezeichneten Zeichenkette |

| | | |
|----------|---------------|--|
| <spreiz> | gespreizt | Beginn und Ende einer mit 'gespreizt' ausgezeichneten Zeichenkette |
| <anz1> | Anf.zeichen 1 | Beginn und Ende einer einfachen 'Anführung' (bei Hervorhebung eines Wortes oder einer Wortsequenz durch den Autor mit Hilfe des Anführungszeichens) - <i>nicht</i> bei <i>Zitaten</i> oder Wörtlicher Rede zu verwenden! |
| <anz2> | Anf.zeichen 2 | Beginn und Ende einer einfachen 'Anführung' (bei Hervorhebung eines Wortes oder einer Wortsequenz durch den <i>Gestalter</i> bzw. die <i>Gestalterin</i>) |

4. Es werden folgende 'nicht-strukturellen' Markierungen zugelassen:

| | | |
|---------|-----------|--|
| <abs1/> | Absatz | ohne strukturelle Differenzierung (= Return) |
| <abs2/> | Absatz | ohne strukturelle Differenzierung (= Return und Leerzeile) |
| <gap1/> | Lücke 1 | Auslassung im Original |
| <gap2/> | Lücke 2 | Auslassung in einem Auszug / Zitat (im Original vorhanden). Muss bei einem Auszug auch am Anfang und / oder Ende stehen, wenn er sich auf einen Eintrag <ewerk> bezieht, ohne dass er mit dem Anfang oder Ende dieses <ewerk> übereinstimmt. |
| <tab/> | Tabulator | (vorgesehen) |