

Varianz in Templates und Annotationen

Templates sind Textschablonen, vergleichbar Formularen. In ein festes Textgerüst sind Leerstellen eingebaut, die manuell oder z.B. aus Datenbanken automatisch ausgefüllt werden. Sie sind wichtig, weil sie Arbeitsteilung und -strukturierung ermöglichen. Sie sind interessant, weil sie ineinander verschachtelt werden und so komplexe Strukturen erzeugen können. Varianz kommt hier zumindest auf zwei Weisen ins Spiel: Zum einen wird durch Templates die Varianz von Texten gehandhabt, nicht die eines Texts, sondern die eines Texttyps. Zum anderen kann umgekehrt ein Text durch variante Templates erzeugt werden, auch können diese unterschiedliche Abstraktionsebenen repräsentieren.

Wie wird mit dieser Varianz auf Text- und Template-Ebene informatisch umgegangen? Dafür wurden eigene Templatesprachen entwickelt, die sozusagen Variablen innerhalb von Markupssprachen bereitstellen. An den dafür benutzten speziellen Tags für die Leerstellen/Variablen stören sich die üblichen Markupeditoren. Das wird durch attributorientierte Template-Sprachen vermieden. Template-Sprachen unterstützen die für Templates wichtigsten Strukturen, teilweise z.B. auch listenartige Strukturen.

Gleichwohl liefern sie sozusagen nur die Elemente der Template-Programmierung. Für den Umgang damit bieten sich Grammatiken an. Diese haben ja gerade die Aufgabe, die unendliche Vielfalt einer Sprache in einer endlichen Reihe von Regeln wiederzugeben. Damit hätten wir dann vielleicht auch eine Brücke zu den inhaltlichen Themen des Projekts geschlagen.

Templates sind eindeutig Werkzeuge der Textproduktion. Meine Grundidee ist nun die folgende: könnten wir nicht, zumindest in manchen Fällen wie z.B. Wörterbuchartikeln, auch für die Textanalyse auf eben diese Strukturen, Sprachen und

Formalisten zurückgreifen? Sofern wirklich Strukturen und nicht nur einzelne Begriffe annotiert werden sollen, sind wir ja darauf angewiesen, dass bei der Produktion eine Art Template, sei es als Textschablone, als Regelwerk oder nur im Kopf des Autors, vorhanden war. Auch hier wäre das Ziel übrigens die Ermöglichung einer besseren Arbeitsteilung zwischen verschiedenen Aufgaben bei der Erschließung von Texten.

Was wir dafür brauchen, ist eine in Analogie zu den Templatesprachen gestaltete Annotationssprache und eine Möglichkeit, die Verschachtelungen nach Art einer Grammatik zu handhaben. Im Detail betrachtet, erweist sich allerdings, dass die Standardgrammatikformalisten nicht ausreichen, sondern modifiziert beziehungsweise erweitert werden müssen.

Die erfreuliche Tatsache ist nun, dass sich die Elemente für all das bereits im Campe-Teil des Projekts finden, einschließlich der Möglichkeit der Handhabung iterativer Strukturen, wie sie auch in Templatesprachen nicht selbstverständlich sind, und einschließlich eigens entwickelter Grammatikerweiterungen (alles nicht von mir!) Es bedürfte nur noch einiger Schritte, um daraus eine vollständige Annotationssprache und -grammatik zu machen. Dazu anzuregen war ein weiteres Motiv für diesen Beitrag