

Kurzzusammenfassung der Dissertation

„*Autonome Topic Maps*. Zur dezentralen Erstellung von implizit und explizit vernetzten Topic Maps in semantisch heterogenen Umgebungen.“ **Autor:** Lutz Maicher

Als Grundidee vieler Anwendungen des Semantic Web und des Web 2.0 kristallisiert sich die dezentrale Erstellung und Veröffentlichung sogenannter Semantic-Web-Dokumente heraus. Diese Dokumente basieren auf dem *aussagegegenstandszentrierten Modellierungsparadigma*. Entsprechend dieses Paradigmas, welches durch Topic Maps umgesetzt wird, erhält *jeder* relevante Aussagegegenstand der realen Welt *genau* einen Repräsentanten im Modell. Die Eigenschaften der Aussagegegenstände werden in den Modellen als Beziehungen zwischen deren Repräsentanten dokumentiert. Das führt zu einer *expliziten* Vernetzung innerhalb der Modelle. Zudem wird sichergestellt, dass Repräsentanten, die den gleichen Aussagegegenstand verkörpern, zusammengeführt werden. Repräsentanten des gleichen Aussagegegenstandes, die in getrennten Dokumenten vorliegen, sind somit *implizit* vernetzt. Diese implizite Vernetzung ist von Bedeutung, wenn die Modelle durch Austauschprotokolle zusammengeführt werden.

Mit *Autonomen Topic Maps* (ATM) wird die dezentrale Erstellung qualitativ hochwertiger, implizit und explizit vernetzter Modelle in semantisch heterogenen Umgebungen ermöglicht. Eine ATM ist eine Topic Map, die eine durch einen generischen Interpreter ausführbare Modellierungsmethoden formalisiert, durch deren Ausführung Beobachtungen zu einer spezifischen Domäne in implizit und explizit vernetzten Topic Maps adäquat dokumentiert werden. ATMs können leicht verteilt werden und in Interpretern für beliebige Anwendungskontexte genutzt werden.

Bei der im Allgemeinen propagierten Nutzung von Ontologien zur Erstellung von Modellen kann nicht sichergestellt werden, dass *identische* Beobachtungen über die Welt mit *identischen* Aussagen der verfügbaren Sprache ausgedrückt werden. Dieser Interpretationsspielraum wird *Semantic Gap* genannt: die Nutzung *einer* Ontologie impliziert die Möglichkeit der Erstellung von Instanzen *einer Vielzahl* unterschiedlicher Modelltypen. Konsistenz ist sichergestellt, wenn für jedes Modell Ontologie *und* Modellierungsmethode offengelegt sind.

Die *Offenlegung der Modellierungsmethoden* ist somit ein probates Mittel, um dem Ziel qualitativ hochwertiger, implizit vernetzter Modelle näher zu kommen. Der nächste Schritt ist die *Formalisierung der Modellierungsmethoden*, so dass diese durch generische Interpreter ausführbar werden. Im Rahmen der Arbeit wird mit Modelling Workflow Patterns (MWP) eine solche, auf Petri-Netzen basierende, Infrastruktur für beliebige Modellierungsmethoden geschaffen. Eine ATM ist eine als Topic Map repräsentierte MWP-Beschreibung, die Topic Maps erzeugt.

Die zweite zentrale Problemstellung ist die Sicherstellung der impliziten Vernetzung auf Typ- und Individuenebene. Die dabei entstehenden Probleme sind offensichtlich: so können mit einem gegebenen Vokabular gleiche Beobachtungen unterschiedlich beschrieben werden (*Synonymie*) bzw. unterschiedliche Beobachtungen gleich beschrieben werden (*Homonymie*).

Der Einsatz von MWP-Beschreibungen löst nicht die Problematik der Homonymie und Synonymie auf Individuenebene in terminologisch dynamischen Domänen, da die offengelegten Modellierungsmethoden nicht mit *ex ante* unbekanntem Individuen umgehen können. Im Rahmen der Arbeit wird der *Impact of Semantic Handshakes* beschrieben. Ein solcher Handshake ist die synonyme Nutzung mehrerer Terme für die Referenzierung eines Aussagegegenstandes. Es wird gezeigt, dass auf die Nutzung zentraler Ontologien (auf Individuenebene) verzichtet werden kann, wenn jedes Topic mind. zwei (populäre) Terme für die Referenzierung des Aussagegegenstands nutzt. Dieser Ansatz einer Standardisierung ohne zentrale, ordnende Instanz ist orthogonal zu dem üblichen Weg der global gültigen, kontrollierten Vokabulare.

Das Zusammenspiel von impliziter und expliziter Vernetzung illustriert die Emergenz einer globalen, stark vernetzten Faktenbasis, basierend auf der dezentralen, kleinteiligen und kollaborativen Dokumentation von Informationen entsprechend des aussagegegenstandsorientierten Modellierungsparadigmas. Durch die mit ATMs ermöglichte Dezentralität der Erstellung dieser Faktenbasis kann ihr Wachstum durch keine Kapazitätsbeschränkung irgendeiner Organisation aufgehalten werden.