

M. Herczeg, W. Prinz, H. Oberquelle (Hrsg.): *Mensch & Computer 2002: Vom interaktiven Werkzeug zu kooperativen Arbeits- und Lernwelten*. Stuttgart: B. G. Teubner, 2002, S. 373-382.

## Matrix Browser – Visualisierung und Exploration vernetzter Informationsräume

Jürgen Ziegler, Christoph Kunz, Veit Botsch

Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO), Stuttgart

### Zusammenfassung

In diesem Beitrag beschreiben wir einen neuen Ansatz für die Visualisierung und Exploration großer Informationsnetze. Netzdarstellungen können beispielsweise zur Visualisierung hypermedialer Informationsressourcen oder für Metadaten-Strukturen wie z.B. Ontologien herangezogen werden. Der Ansatz beruht auf der Verwendung einer interaktiven Adjazenzmatrix für die Darstellung der Beziehungen zwischen den Konzepten, welche die Knoten des Netzes bilden, wobei die darzustellenden Konzepte dynamisch aus einer zugrunde liegenden Netzrepräsentation extrahiert werden. Diese Konzepte werden entweder als Listen oder Hierarchien an den beiden Achsen der Matrix dargestellt. Unterschiedliche interaktive Markierungen in den Matrixzellen stellen die Relationen zwischen den Konzepten dar. Erste Benutzbarkeitsanalysen zeigen Leistungsvorteile und eine Verkürzung visueller Suchpfade gegenüber den herkömmlichen Graphendarstellungen von Netzen.

## 1 Einleitung

Vernetzte Informationsstrukturen werden für die Exploration und Navigation in komplexen Informationsräumen wie Internetseiten oder Wissensdatenbanken immer wichtiger. Konzeptnetze bzw. formalisierte Ontologien werden auch zunehmend als Metadaten eingesetzt, um Informationen zu indexieren und eine inhaltsbezogene Suche zu unterstützen. Solche komplexen ontologischen Informationen können zum Beispiel in Formalismen wie Topic Maps (Biezunski et al 1999) oder DAML+OIL (Hendler 2001) zum Ausdruck gebracht werden. Hierbei stellt allerdings die Visualisierung und Navigation solcher Netzstrukturen ein erhebliches Problem bei der Gestaltung geeigneter Benutzungsschnittstellen dar. Gestaltungsfragen beziehen sich etwa auf die Minimie-

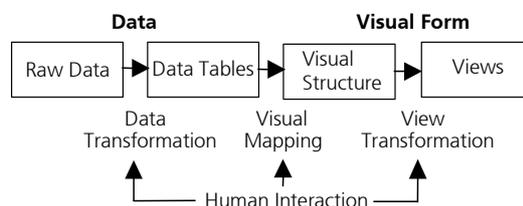


Abbildung 1: Referenzmodell der Informationsvisualisierung

rung der visuellen Suchpfade, die Übersichtlichkeit der Gesamtstruktur oder die Effizienz interaktiver Explorationsmöglichkeiten.

Ein bekanntes Referenzmodell der Informationsvisualisierung (Card et al. 1999), das in Abbildung 1 dargestellt ist, zeigt schematisch den Ablauf, wie Rohdaten bis zu der eigentlichen Visualisierung aufbereitet werden. Die Rohdaten werden aufbereitet und in eine geeignete visuelle Struktur überführt. Aus dieser visuellen Struktur können unterschiedliche Sichten auf die Daten gebildet werden. Die gebräuchlichste visuelle Struktur für Netze ist die Repräsentation als Graph mit Knoten und Kanten. Diese Repräsentation wird zwar häufig verwendet, besitzt jedoch folgende Nachteile:

Ist das Netzwerk groß und besitzt einen hohen Grad an Vernetzung, so wird es schwierig, den Überblick zu behalten und bestimmte Ausschnitte zu finden. Auch wird die visuelle Suche durch meist beliebige und unstrukturierte Anordnungen der Knoten behindert. Aus Platzgründen können oft auch assoziierte Knoten nur weit voneinander entfernt platziert werden. Während es bei dieser Darstellungsform meist noch einfach ist, die direkte Nachbarschaft von Knoten zu erfassen, wird es schwierig und fehleranfällig, größere Beziehungsmuster zu erkennen oder visuell längeren Pfaden zu folgen.

Vor diesem Hintergrund wurde mit den hier berichteten Arbeiten das Ziel verfolgt, die Übersichtlichkeit komplexer Netzdarstellungen zu verbessern, die visuelle Suche nach bestimmten Konzepten und Assoziationen zu unterstützen und intuitive Interaktionsmöglichkeiten zur Exploration auf unterschiedlichen Detaillierungsstufen nach dem Prinzip des „progressive disclosure“ bereitzustellen. Die dargestellten Arbeiten wurden im Rahmen des vom BMBF geförderten Leitprojektes INVITE im Programm Mensch-Technik-Interaktion durchgeführt.

## 2 Verwandte Arbeiten

Eine neuartige Form der visuellen Darstellung komplexer Informationsstrukturen wurde durch den Hyperbolic Browser (Lamping & Rao 1994) eingeführt, der gleichzeitig eine kontinuierliche, direkt-manipulative Interaktion bietet. Der Ansatz ist jedoch aufgrund seines Verzerrungsmechanismus, bei dem nur Beziehungen zwischen räumlich benachbarten Knoten im Display-Fokus vollständig zu erkennen sind, nur für die Visualisierung hierarchischer Strukturen geeignet. Er wurde jedoch in einigen Fällen auch für die Darstellung von Topic Maps und Ontologien eingesetzt. Ein Beispiel hierfür ist OntoBroker (Fensel et al. 1998). Hierfür ist er allerdings nur bedingt einsetzbar, da bei nichthierarchischer Vernetzungsstruktur manche Knoten mehrfach in der Darstellung auftauchen können.

Eine nichthyperbolische Darstellung, die dieses Problem vermeidet, wird innerhalb des OntoViz Plugins (Abbildung 2) für das Ontologie-Autorenwerkzeug Protegee-2000 (Noy et al. 2001) verwendet. Es basiert auf GraphViz (Gansner & North 1999), einer von AT&T entwickelten Umgebung zur Graphenvisualisierung.

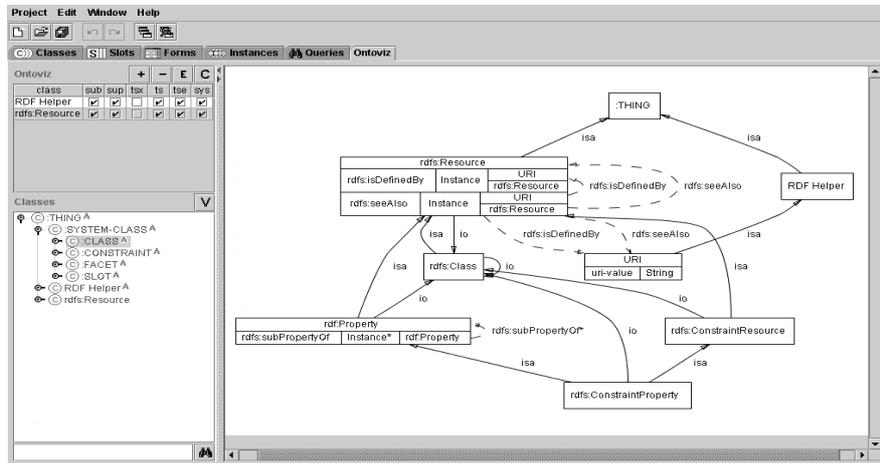


Abbildung 2: Ontologievisualisierung mit OntoViz

Ein dreidimensionales Visualisierungsverfahren (Le Grand & Soto 2000) für XML Topic Maps, das am Laboratoire d'Informatique de Paris 6 entwickelt wurde, zeigt Abbildung 3. Es beruht auf Cone Trees (Robertson et al. 1991), die um interaktive Möglichkeiten erweitert wurden. Die Anzahl der dargestellten Knoten kann durch Filterungs- und Klassifizierungsalgorithmen reduziert werden. Mittels zweier Sichten kann die visualisierte Topic Map erkundet werden. Zum einen können mehrstellige Beziehungen zwischen Knoten als Klassen dargestellt werden, zum anderen können binäre Relationen als tatsächliche Kanten repräsentiert werden.

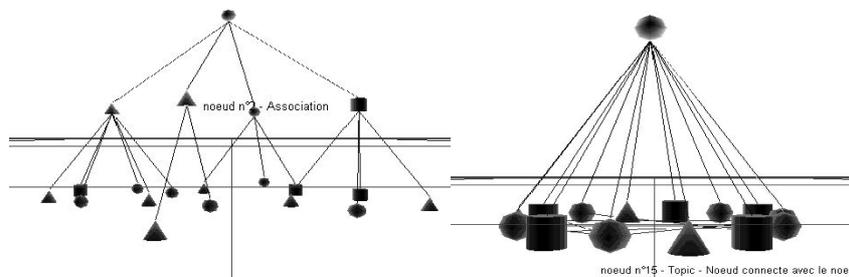


Abbildung 3: Topic Map Visualisierung mit Relationen als Klassen oder Kanten

Alternativ wurden statische Netzvisualisierungen einschließlich der Repräsentation als Matrix schon von Bertin (Bertin 1981) vorgeschlagen. Interaktive Möglichkeiten wurden jedoch in dieser Arbeit nicht betrachtet. Keiner der vorgestellten Ansätze unterstützt in ausreichender Form die Darstellung beliebig vernetzter Informationsstrukturen im Hinblick auf unterschiedliche Nutzeraufgaben wie die Suche spezifischer Knoten und Kanten, das Entdecken der Beziehungen zwischen beliebigen Knoten, oder die Erfassung aller Beziehungen eines bestimmten Knotens. Wei-

terhin fehlen wesentliche interaktive Eigenschaften wie ein systematischer Drill-Down bzw. die Verdichtung von Teilstrukturen. Aus diesen Gründen sind die existierenden Techniken nicht ausreichend geeignet, um komplexe Netzstrukturen zu visualisieren und zu explorieren.

### 3 Gestaltungsprinzipien des Matrix Browsers

Die zentrale Idee für die Konzeption des Matrix Browsers ist es, den zugrundeliegenden Graphen auf eine hoch interaktive Adjazenzmatrix ab zu bilden (Abbildung 4). Die aus der Graphentheorie bekannten Adjazenzmatrizen sind eine alternative Form der Graphendarstellung. Die Knoten des Graphen werden an der horizontalen und vertikalen Achse der Matrix angeordnet, wobei die Zellen der Matrix die Kanten repräsentieren. Eine Zelle kennzeichnet also eine Relation, wenn sich die Linien, die rechtwinklig zu den Achsen von zwei Knoten aus gehen, sich innerhalb einer Zelle schneiden und diese beiden Knoten durch eine Kante verbunden sind. Durch Verwendung von Pfeilen und graphischen Symbolen können sowohl die Richtung der Relation im Falle von gerichteten Graphen als auch verschiedene Typen von Relationen visualisiert werden. Unterstützend können dabei auch zusätzliche Techniken wie die Beschreibung von unterschiedlichen Beziehungen durch Tool Tips eingesetzt werden.

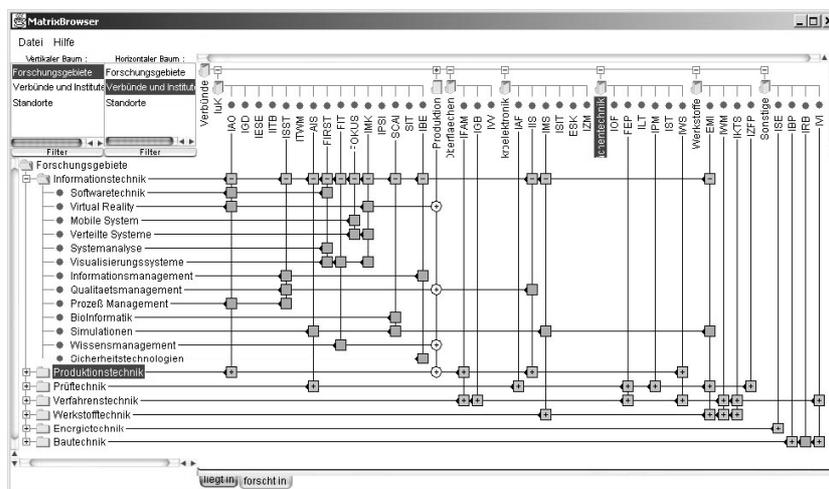


Abbildung 4 : Prototyp des Matrix Browser

Die zweite Grundidee ist, diejenigen Teile des Netzes, die eine hierarchische Struktur aufweisen, durch bekannte Baummetaphern dar zu stellen. Diese werden auf die beiden Achsen der Matrix platziert. Viele netzartige Strukturen, wie beispielsweise die Vernetzungsstruktur einer Webseite und ihrer Teilseiten, beinhalten zahlreiche Knoten, die durch transitive Relationen verbunden sind. Transitive Relationen können z.B. „Teil-von-“ oder Subklassen-Beziehungen sein. Diese können mittels vertrauten Windows Explorer ähnlichen Werkzeugen, die hierarchische Informationen visualisieren, effektiv erfahrbar gemacht werden. Werden hierarchische Teilstrukturen bereits durch interaktive Bäume an den Matrixachsen dargestellt, wird die Markierung in den Zellen lediglich für weitere, zwischen den beiden Substrukturen bestehende Relationen benötigt. Mit

dem Matrix Browser können die Knoten und Teilhierarchien, die an den Achsen dargestellt werden, auf flexible Art angewählt und gefiltert werden.

## 4 Interaktives Verhalten

Durch Expandieren und Kollabieren der korrespondierenden Bäume können die Hierarchien auf den Achsen direkt exploriert werden. Diese vertraute und effektive Art des Explorierens ist auf beiden Achsen möglich. Die in den Zellen der Matrix gezeigten Relationen zwischen den beiden, gerade auf den Achsen angezeigten Knotenmengen, sollten immer mit dem interaktiven Status der Achsen konsistent sein. Um diese Konsistenz der Visualisierung zu gewährleisten, müssen eine Reihe von Problemen gelöst werden. Das erste Problem betrifft die Sichtbarkeit der Relationen von Knoten, die gerade in dem entsprechenden Baum nicht sichtbar sind, weil sich ihr übergeordneter Ast in einem kollabierten Zustand befindet. Das zweite Problem tritt bei subsumierten Beziehungen auf, bei denen die Relationen zwischen zwei Knoten nach oben oder unten innerhalb der Hierarchie, die einen der Knoten beinhaltet, weitergegeben werden können. Innerhalb eines Beispiels, das die Organisationsstruktur einer Firma mit den entsprechenden Standorten in Beziehung setzt, liegt der Standort einer Zweigstelle in einer Stadt. Dieses ist die explizite Verbindung im Netz. Der Standort liegt aber auch in dem Land, wie auch innerhalb des Kontinents, in dem diese Stadt sich befindet. Ob solche Inferenzen gemacht werden können oder nicht, hängt von den zu Grunde liegenden Regeln innerhalb der Ontologie und einem passenden Inferenzmechanismus ab. Die Visualisierung sollte jedoch auf diese Fälle ausgelegt sein.

Um mit den verschiedenen auftretenden Darstellungssituationen umgehen zu können, verfügt der Matrix Browser über unterschiedliche Arten von visualisierten Relationen, welche auch über interaktives Verhalten verfügen. Eine Relation in einer Zelle der Matrix kann dabei einen der folgenden Fälle repräsentieren :

- Eine *explizite Relation* stellt eine direkt spezifizierte Verbindung zwischen zwei Knoten des zugrunde liegenden Graphen dar. Diese Relation kann semantisch typisiert sein.
- Eine *versteckte Relation* ist ein Indikator für eine weiter unten in der Hierarchie vorhandene explizite Relation, die aber gerade nicht sichtbar ist, da der übergeordnete Ast kollabiert ist. Diese trägt aber keine weitere Bedeutung.
- Eine *implizite Relation* zeigt eine Beziehung, welche durch einen Inferenzmechanismus von dem zu Grunde liegenden Graphen abgeleitet wurde. Diese kann entweder durch Vererbung oder Generalisierung entstehen. Bei Vererbung wird die Beziehung nach unten innerhalb der Hierarchie durchgereicht, während bei Generalisierung die Weitergabe nach oben erfolgt. Speziell im zweiten Fall bedeutet diese Art von Beziehung, dass durch Expansion der entsprechenden Äste die explizite Relation, welche der Grund für die Inferenz war, weiter unten in der Hierarchie gefunden werden kann.

Diese verschiedenen Relationstypen werden visuell durch die Darstellung mit unterschiedlichen Symbolen in den Zellen unterschieden. Sowohl versteckte als auch implizite Relationen im Fall der Generalisierung können wie die Äste der Bäume durch Mausklicks aufgeklappt werden. Um diese Art der Interaktion zu verdeutlichen, sind beide Arten durch ein „+“-Symbol ausgezeichnet. Durch eine enge Kopplung von Zellen und Hierarchien auf den Achsen wird erreicht, dass durch das Expandieren einer Relation auch der zugehörige Ast expandiert wird. Durch Selektion und Filterung des vollständigen Graphen und durch interaktives Herunterbrechen der Grobkonzepte

auf Feinkonzepte und ihren Relationen auf unterster Ebene, kann der Benutzer auf diese Weise denjenigen Teil, der gerade angezeigt wird, flexibel untersuchen.

## 5 Extraktion von Teilhierarchien

Um die interaktiven Möglichkeiten des Matrix Browsers optimal zu nutzen, muss das zu Grund liegende Netz so aufbereitet werden können, dass Teilstrukturen hierarchischer Art, beziehungsweise transitive Teilhierarchien erkannt und extrahiert werden. Hierdurch wird es möglich, semantisch zusammenhängende Teilstrukturen dynamisch zu filtern und zu visualisieren. Betrachtet man im gesamten Graphen jeweils nur einen Relationstyp, ergeben sich Teilgraphen, die wiederum durch andere Beziehungstypen verbunden sind. Dabei können sich die Teilgraphen überlappen, d.h. sie besitzen gemeinsame Knoten. Zwar können diese Ausschnitte des Netzes mittels geeigneter Algorithmen in Baumstrukturen überführt werden, für den Benutzer sind jedoch insbesondere semantisch zusammenhängende Teilstrukturen von Interesse, die durch einen einzigen und gerichteten Beziehungstyp verbunden sind. Die meisten formalen Repräsentationen für Themennetze wie XML Topic Maps, RDFS und DAML+OIL besitzen mindestens einen explizit ausgezeichneten Relationstyp, die Subklassenbeziehung, welche ohne aufwendigere Inferenzmechanismen dazu verwendet werden, Teilhierarchien in der Netzstruktur zu erkennen. Weiterführende Beziehungstypen können jedoch nur durch deduktive Methoden zur Schlussfolgerung erkannt werden.

Zu diesem Zweck wird hier der Ansatz verfolgt, die zu Grunde liegenden Netzstrukturen in einer Wissensrepräsentation ab zu bilden, aus der die interessierenden Teilstrukturen durch Inferenzmechanismen extrahiert werden können. Dadurch können auch komplexere Abfragen auf die Wissensbasis ausgeführt werden, um implizite Verbindungen herzustellen. Die formalen Repräsentationen von Ontologien werden hierzu in F-Logik (Kiefer et al. 1995) überführt. Die Relationen müssen nun noch durch Axiome ausgezeichnet werden, um ihre Semantik zu definieren und Inferenzen zu ermöglichen. Axiome sind nur in Fällen wie DAML+OIL in der formalen Repräsentation hinterlegt, der Topic Map Standard verfügt nicht über eine solche Modellierungsmöglichkeit. Daher müssen sie auf geeignete Weise annotiert werden. Abbildung 5 zeigt den Informationsfluss von der Ontologie bis zum Matrix Browser.

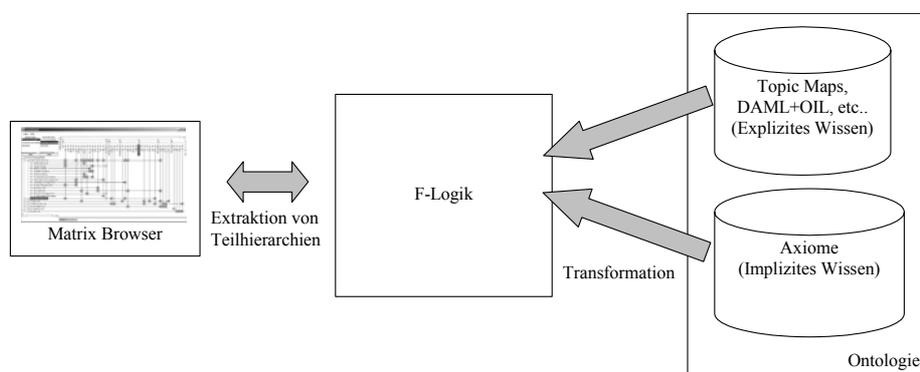


Abbildung 5: Extraktion von hierarchischen Strukturen



sind Untersuchungen in Vorbereitung, mit denen auch im interaktiven Fall die Effektivität des Matrix Browsers getestet werden soll. Insgesamt zeigen die bisherigen Ergebnisse, dass der Matrix Browser-Ansatz auf Grund seiner reguläreren Anordnungsstruktur effizienter als konventionelle Netzdarstellungen ist. Es besteht die begründete These, dass durch die interaktiven Funktionen gerade bei komplexen Netzen noch deutlichere Vorteile gegenüber der Netzdarstellung zu erzielen sind. Die Verifikation dieser Thesen bleibt jedoch noch Gegenstand weiterer Untersuchungen.

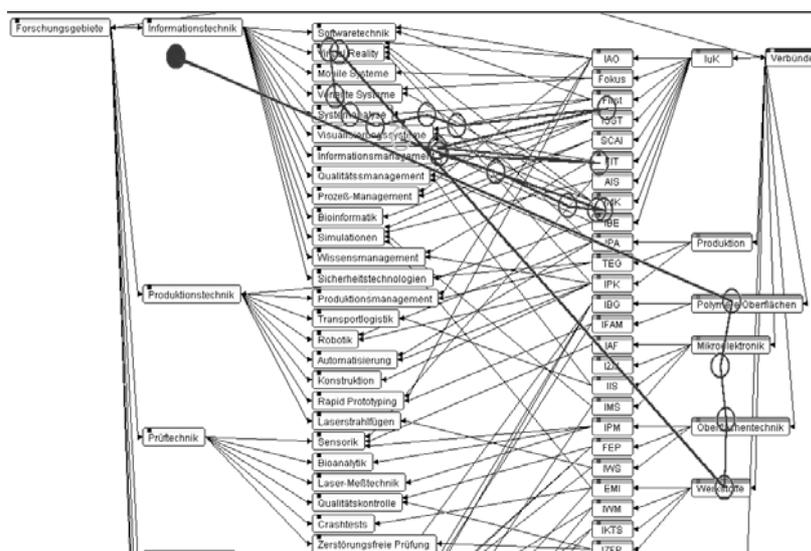


Abbildung 7: Visueller Suchpfad eines Probanden mit einer strukturierten Netzdarstellung

## 7 Zusammenfassung und Ausblick

Mit diesem Beitrag haben wir ein neuartiges Konzept und einen ersten Prototypen für die Visualisierung und Exploration komplexer, vernetzter Informationsstrukturen vorgestellt. Matrix Browser ist ein Werkzeug, um umfangreiche, stark vernetzte Informationsstrukturen effektiv zu visualisieren und durch vertraute Interaktionstechniken zu explorieren. Beim hier berichteten Stand der Arbeiten werden für die Visualisierung vordefinierte Konzeptionshierarchien und -beziehungen verwendet. Um beliebige Ontologien für die Exploration mit Matrix Browser aufzubereiten, ist gegenwärtig eine Architektur und ein Inferenzmechanismus in Entwicklung, der es erlaubt, aus einer zu Grunde liegenden wissensbasierten Netzrepräsentation Teilhierarchien und implizite Relationen zu extrahieren.

Neben diesen funktionalen Erweiterungen wird gegenwärtig die grafische Darstellung des Matrix Browsers optimiert, um die Lesbarkeit der Konzeptionshierarchien, insbesondere auf der horizontalen Achse zu verbessern, die Darstellung unterschiedlicher Relationstypen effektiver zu gestalten und Möglichkeiten für den Einsatz der dritten Dimension in der Informationsvisualisierung zu untersuchen. Weiterhin erscheint der Matrix Browser auch als geeigneter Ausgangspunkt, um Sucher-

gebnisse in einer Struktur von Metadaten anzuordnen und interaktiv zu explorieren. Hierzu werden gegenwärtig Techniken, die auf einem Smart Lens-Ansatz basieren, untersucht und entwickelt.

## 8 Literaturverzeichnis

- Becker, R. A.; Eick, S. G.; Wilks, A. R. (1995): Visualizing Network Data. In: *IEEE Transactions on Visualization and Computer Graphics*, Vol. 1, Nr. 1, S. 16 - 28.
- Bertin, J. (1981): *Graphics and Graphic Information-Processing*. Berlin: Walter de Gruyter Co.
- Biezunski, M., Bryan, M., Newcomb, S. R. (1999): *ISO/IEC 13250 Topic Maps: Information Technology -- Document Description and Markup Language*. <http://www.ornl.gov/sgml/sc34/document/0058.htm>.
- Card, S.K.; Mackinley J. D.; Shneiderman, B. (1999): *Readings in Information Visualization : Using Vision to Think*. San Francisco: Morgan Kaufmann.
- Fensel, D., Decker, S., Erdmann, M., Studer, R. (1998): Ontobroker: The Very High Idea, In: *Proceedings of the 11th International Flairs Conference*, Florida.
- Gansner, E.R., North, S.C. (1999): An open graph visualization system and its applications to software engineering. In: *Software-Practice and Experience*, Vol. 0, S. 1 – 5.
- Hendler, J.(2001): *DAML: The DARPA Agent Markup Language Homepage*. <http://www.daml.org>
- Kifer, M., Lausen, G., Wu, J. (1995): Logical Foundations of Object-Oriented and Frame-Based Languages, In: *Journal of the ACM*, Vol. 42, Nr. 4, S. 741-843.
- Lamping, J.; Rao, R. (1994): Laying Out and Visualizing Large Trees Using a Hyperbolic Space. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*, S. 13 – 14.
- Le Grand, B., Soto, M., (2000): Information Management – Topic Maps Visualization. In: *XML Europe '2000*, Frankreich.
- Noy, N. F., Sintek, M., Decker, S., Crubezy, M., Ferguson, R.W., Musen. M. A. (2001): Creating Semantic Web Contents with Protege-2000. In: *IEEE Intelligent Systems*, Vol. 16, Nr. 2. S. 60 - 71.
- Robertson, G., Mackinlay, J., and Card, S. (1991): Cone trees: Animated 3D visualizations of hierarchical information. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 189-194.

## **Kontaktinformationen**

Bei Fragen oder zu weiterführenden Information wenden Sie sich bitte an :

Dr.-Ing. Jürgen Ziegler

Email : [juergen.ziegler@iao.fhg.de](mailto:juergen.ziegler@iao.fhg.de)

Dipl.-Ing. Christoph Kunz

Email : [christoph.kunz@iao.fhg.de](mailto:christoph.kunz@iao.fhg.de)

Dipl.-Phys. Veit Botsch

Email : [veit.botsch@iao.fhg.de](mailto:veit.botsch@iao.fhg.de)

Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO)

Nobelstrasse 12

D-70569 Stuttgart