

Schlußbericht: Ontologien in GETESS

Zuwendungsempfänger:

Universität Karlsruhe

Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB)

Förderkennzeichen:

01 IN 901 D 2

Vorhabensbezeichnung:

GERman Text Exploitation and Search System

Laufzeit des Vorhabens:

01. April 1999 – 30. Juni 2001

Autoren:

Alexander Mädche, Steffen Staab, Rudi Studer

1 Wissenschaftlich-technische und andere wesentliche Ereignisse

GETESS (German Text Exploitation and Search System) ist der Name eines Projektes mehrerer Partner aus Forschung, Lehre und Industrie, das unterschiedlichste Fachgebiete der Informatik zusammenführt wie Datenbanken, Künstliche Intelligenz und Linguistik. Ziel der Projektes war die Entwicklung eines prototypischen Informationssystems, das Text- (HTML-) Dokumente mit neuen Technologien einsammelt, abspeichert und mittels natürlichsprachlichen Dialogsystems für die Informationsrecherche verfügbar macht. Partner des Projektes neben der Universität Karlsruhe, Institut AIFB, Forschungsgruppe Wissensmanagement waren die GECKO mbH Rostock, der Lehrstuhl Datenbanken und Informationssysteme der Universität Rostock und das DFKI in Saarbrücken.

Die Kernaufgaben des AIFBs im Projekt GETESS bestanden aus folgenden Schwerpunkten:

- Die Entwicklung einer Ontologie zum Bereich Tourismus und einer Ontologie zur Domäne “Finanz”
- Die Entwicklung geeigneter Konzepte zur Integration von Ontologien und
 1. Textanalyse/Parsing (d.h. mit DFKI Saarbrücken)
 2. Datenbanken (d.h. mit Universität Rostock)
 3. User Interface (d.h. mit GECKO mbh, Rostock)
- Die Entwicklung von geeigneten Ontologie-Werkzeugen für

1. die Erstellung und Wartung der Ontologien
2. die Anbindung der Ontologien an die Module der Partner

In diesem Kontext wurden umfangreiche Forschungen und Entwicklungen durchgeführt und angestoßen, die zu folgenden weittragenden Resultaten geführt haben:

Im Bereich **Ontologie-Engineering** wurden existierende Methodologien untersucht, überarbeitet und auf die spezielle Situation im GETESS-Kontext angepasst. Integriert wurden Erkenntnisse, die mittels Fragebögen gewonnen wurden. Über das ursprüngliche Projektziel hinaus, wurden Techniken des maschinellen Lernens benutzt zur besseren Konstruktion der Ontologie.

Im Bereich **Zusammenhang von Linguistik und Ontologie** wurde eine erweiterte Ontologierepräsentation entwickelt, die die direkte Anbindung von Semantik an Lemmata erlaubt. Für die Integration wurde ein Werkzeug "OntoRel" entwickelt, welches anhand von Ontologiestrukturen die semantische Validität von Parsingergebnissen verifiziert.

Im Bereich **Inferenzmechanismen** wurde die existierende Ontobroker-Inferenzmaschine weiterentwickelt, so dass aus der existierenden Alpha-Version eine stabil lauffähige Beta-Version entstand. Mit dem Konzept der Semantic Patterns wurde eine Methodik entwickelt, die nicht nur das Ontology-Engineering erleichtert, sondern auch die Übertragbarkeit von Schlussfolgerungen auf andere Inferenzmaschinen erleichtert.

Des Weiteren wurde konkrete **Werkzeuge** entwickelt, die es erlauben, das Ontologie-Engineering einfacher und nachvollziehbarer zu gestalten. *OntoEdit* erlaubt die direkte Modellierung von Ontologiestrukturen. *OntoKick* erlaubt die Traceability von Modellierungsentscheidungen, die anhand von Kompetenzfragen getroffen werden. *Text2Onto* erweitert *OntoEdit* so um Techniken des maschinellen Lernens, dass das Ontologie-Engineering erleichtert wird.

Es sei darauf hingewiesen, dass der Leser in den Veröffentlichungen eine umfangreiche Dokumentation der Projekthalte findet (vgl. Abschnitt 2.1)

1.1 Zusammenfassung des Projektfortschritts im ersten Halbjahr 2001

Die Arbeiten zum 5. Meilenstein (Juni 2001) des Verbundprojekts "GETESS" an der Universität Karlsruhe behandelten vor allem die Verfeinerung einer Ontologie für die gewählte Domäne "Finanzen". In einer größeren Evaluations- und Vergleichsstudie wurde die Ontologiemodellierung evaluiert und neue Evaluationsmethoden erforscht.

In letzten Halbjahr wurde in der Arbeit vor allem Wert gelegt auf die spezifische Domäne "Finanz" und die Integration mit den anderen Partner zur Vervollständigung eines lauffähigen Demonstrators.

Speziell zum Meilenstein Juni 2001 endeten die Arbeitspakete:

- AP 1: Analyse
- AP 4: Linguistische/ontologische Aspekte
- AP 8: Dialogführung
- AP 12: weitere Domäne
- AP 13: Multilingualität

Die dort und in den anderen Arbeitspaketen entstandenen Ergebnisse werden nun einzeln beschrieben.

1.2 AP 1: Untersuchung der Tourismus-Domäne

Die Tourismusdomäne wurde detailliert untersucht, um zu einer gültigen Ontologie zu gelangen. Hierzu wurde in verschiedenen Phasen vorgegangen.

In der **Phase I** wurde zunächst explorativ mit verschiedenen Methoden gearbeitet. Fragebögen an potentielle Kunden wurden ausgewertet, mit Tourismusexperten (Reisebüro, Hotelbranche, Tourismusdienste der Stadt Baden-Baden) wurden Interviews durchgeführt. Die Domänentexte wurden manuell untersucht. Auf diese Weise gelangten wir zu einer umfangreichen Kernontologie. Außerdem wurden in dieser Phase Anforderungen an zu entwickelnde Werkzeuge transparent, die so nicht von existierenden Werkzeugen abgedeckt wurden.

In der **Phase II** wurden die bereits erwähnten Werkzeuge OntoEdit, Text2Onto und OntoKick entwickelt und eingesetzt. Dabei ergaben sich — natürlicherweise — aufgrund der Unreife der Werkzeuge Probleme, die teilweise zu leichten Verzögerungen im Projektfortschritt führten, die aber im wesentlichen wieder eingeholt werden konnten.

In dieser Phase wurde auch besonders deutlich, dass noch erheblicher Forschungsbedarf für Ontologieentwicklungswerkzeuge bestand — und auch immer noch besteht. Durch die Forschungsarbeiten an GETESS konnte aber auch gerade solcher Bedarf erkannt und angegangen werden. In der **Phase III** wurden die Erfahrungen in einer Methodologie festgehalten (vgl., z.B., [54]).

1.3 AP 3.1: Angemessener Repräsentationsformalismus

Im Arbeitspaket 3.1 ging es darum einen angemessenen Repräsentationsformalismus zu definieren, der den Projektanforderungen gerecht wird. Frühzeitig im Projekt wurde F-Logic als zentrale Ontologierepräsentationssprache gewählt, die sich im weiteren Projektverlauf auch als weitestgehend geeignet erwiesen hat.

Allerdings mussten wir nach einiger Zeit feststellen, dass F-Logic, ebenso wie andere Repräsentationssprachen, nicht sehr geeignet ist, um die semantische Ebene von Worten direkt zu beschreiben. Es wurde daher notwendig, die Ontologiedefinition um eine geeignete lexikalische Ebene zu erweitern, die auch für *Mehrsprachigkeit* und *Benutzerinteraktion* ausgelegt wurde (vgl. [25]).

1.4 AP 3.2: Inferenzmechanismen

Für die Durchführung von Inferenzen wurde die Ontobroker Inferenzmaschine benutzt. Diese basiert auf deduktiven Datenbankmethoden, lag zu Beginn des Projektes in einer Alpha-Version vor und bietet als Schnittstelle Java-Aufrufmöglichkeiten mit F-Logic als Anfragesprache. Im Laufe des Projekts wurde die Inferenzmaschine in eine stabile Beta-Version überführt, um den Anforderungen von Seiten der Projektpartner gerecht zu werden.

Darüberhinaus wurde allerdings im Laufe des Projektes deutlich, dass es nicht genügte die Inferenzmaschine den anderen Partnern als Dienst zur Verfügung zu stellen. Vielmehr trat der Bedarf auf, gezielt Schlussfolgerungsmuster auch auf andere Weise (z.B. in der verwendeten

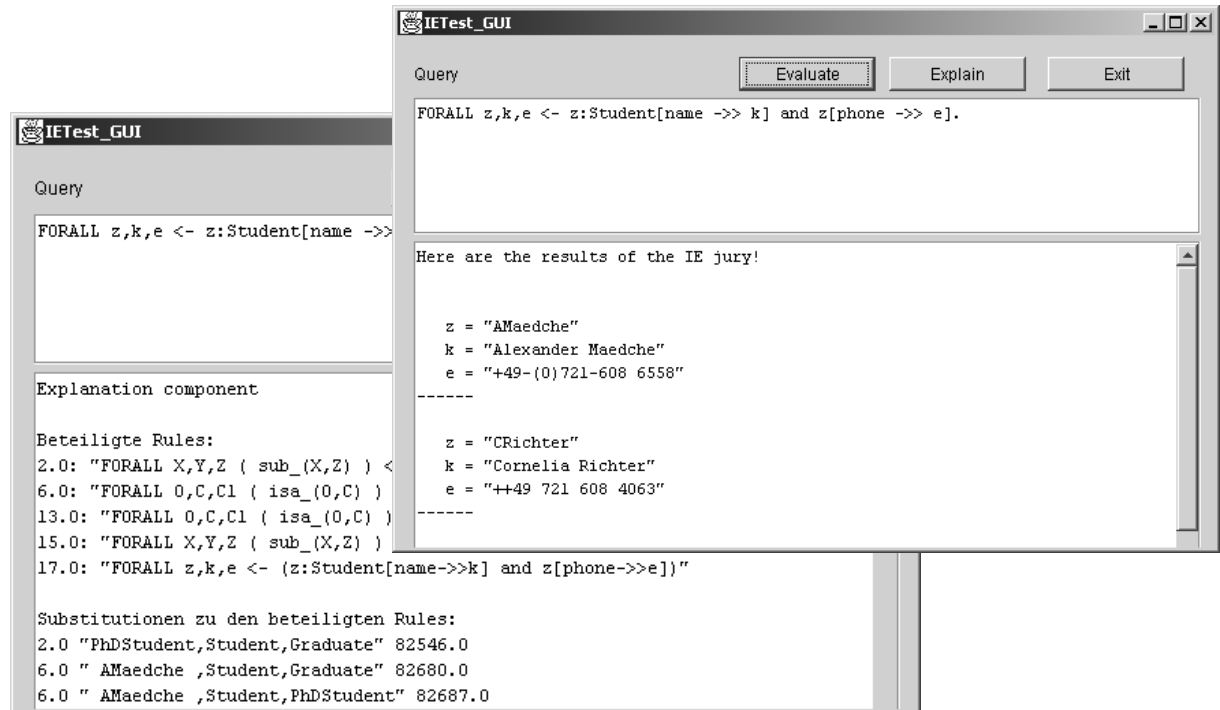


Abbildung 1: F-Logic Query Interface

Objektrelationalen Datenbank der Universität Rostock) zu verwenden. Hierfür fehlte allerdings eine geeignete Theorie zur Überführung von F-Logic-Spezifikationen in objektrelationale Muster.

Aus diesem Grund wurde in der zweiten Hälfte des Projekts "Semantic Patterns" entwickelt (vgl. [41]). Diese erlauben die — weitestgehend — zielsprachenunabhängige Beschreibung von Inferenzmustern und wurden in einer einfachen Form auch in OntoEdit integriert. Aufgrund der zu entwickelnden theoretischen Fundierung war es allerdings nicht möglich, auch noch im Rahmen des GETESS-Projektes volle Schlussfolgerungsmöglichkeiten auf den Daten in einer hochskalierbaren Weise zu realisieren. Deswegen war der Einsatz von Inferenzen in der Datenbank nur sehr eingeschränkt möglich (z.B. für die Taxonomie von Konzepten aber nicht für die geographische *enthalten-sein*Relation).

1.5 AP 3.3 Konstruktion des Domänen-Modells

Das Arbeitspaket 3.3 umfasste die Erstellung der Tourismus-Ontologie. Hierfür wurden zunächst Top-Level-Ontologien untersucht. Aufgrund ihrer Abdeckung wurde der Top Level von HPKB & Sensus herangezogen, aus dem eine geeignete Untermenge selektiert wurde. Mit den bereits erwähnten Methoden und Werkzeugen entstand basierend auf diesem Top Level eine Tourismus-Ontologie mit 1087 Konzepten, 199 Relationen, sowie 2690 lexikalischen Beschreibungen für die Linguistik und 2174 Beschreibungen für die Benutzungsschnittstelle. Für Inferenzen wurden 29

Axiome modelliert.

1.6 AP 4: Zusammenhang zwischen linguistischen und ontologischen Aspekten

Als grundlegend für den Zusammenhang zwischen linguistischen und ontologischen Aspekten erwies sich die Erarbeitung eines linguistisch-kognitiven Grundverständnisses, das mit den Systemaspekten von GETESS vereinbar war. Basierend auf der Unterscheidung von Zeichen, Begriff und Bezeichnetem wurde die Ontologiedefinition angepasst, die Ontologiestruktur wurde um ein Lexikon erweitert (vgl. [26]).

Darauf basierend wurde das Werkzeug *OntoRel* entwickelt. Die Grundidee von *OntoRel* basiert darauf, dass das Parsing syntaktische Strukturen vorschlägt, welche durch die Ontologie verifiziert oder abgelehnt werden (vgl., z.B., [40]). Eine erste Version von *OntoRel* konnte frühzeitig bereitgestellt werden, im Laufe der Zeit konnten seine Fähigkeiten durch Interaktion mit den Projektpartnern deutlich erweitert werden.

1.7 AP 8.1 Werkzeugunterstützung zur Modellierung von Ontologien

In einer ersten Projektphase wurden gängige Werkzeuge für die Ontologiemodellierung wie Protege untersucht. Allerdings fehlten diesen Werkzeugen Fähigkeiten wie z.B. Multilingualität. Deswegen wurde, basierend auf den aus den Untersuchungen und Experimenten, gewonnenen Erfahrungen, ein eigenes Werkzeug *OntoEdit* entwickelt.

Über den ursprünglichen Projektplan hinaus wurde *OntoEdit* im Laufe des Projekts durch die Komponenten *OntoKick* und *Text2Onto* erweitert. *OntoKick* unterstützt eine Ontologieentwicklungsmethodik, die Kompetenzfragen — wie sie aus dem Software-Engineering bekannt sind — und Fragebögen berücksichtigt. Neue Konzepte werden anhand der Antworten auf Kompetenzfragen und Fragebogenresultaten eingefügt und ihre Herkunft wird anhand dieser Antworten protokolliert. Damit ist eine verbesserte Nachvollziehbarkeit von Ontologiestrukturen möglich.

Text2Onto realisiert verschiedene Lernmethoden, die das Explorieren der Texte (z.B. HTML-Seiten aus der Tourismusdomäne) erleichtern (vgl., [23, 16]). *Text2Onto* erwies sich als wegweisend im Bereich des Ontology-Engineerings. Parallel zur Entwicklung von *Text2Onto*, wurde eine Workshop-Serie begründet. Die nun vor ihrem dritten erfolgreichen Event steht (vgl. Abschnitt 2.3).

1.8 AP 8.2: Unterstützung der Ontologie für die Anfrageunterstützung

Das Arbeitspaket 8.2 schlug sich vor allem auf zwei Arten in der Ontologie und den Ontologiewerkzeugen nieder.

1. Neben den lexikalischen Einträgen für das Domänenlexikon der Parsingkomponente wurden zweisprachig Einträge für das Dialoginterface bereitgestellt. Während erstere vor allem auf die Abstraktion zu Gunsten der Linguistik fokussierten, lag der Schwerpunkt bei zweiteren in der ansichtorientierten Darstellung von sprechenden Konzept- und Relationsbeschreibungen. *OntoEdit* wurde mit Fähigkeiten versehen diese Einträge zu verwalten.

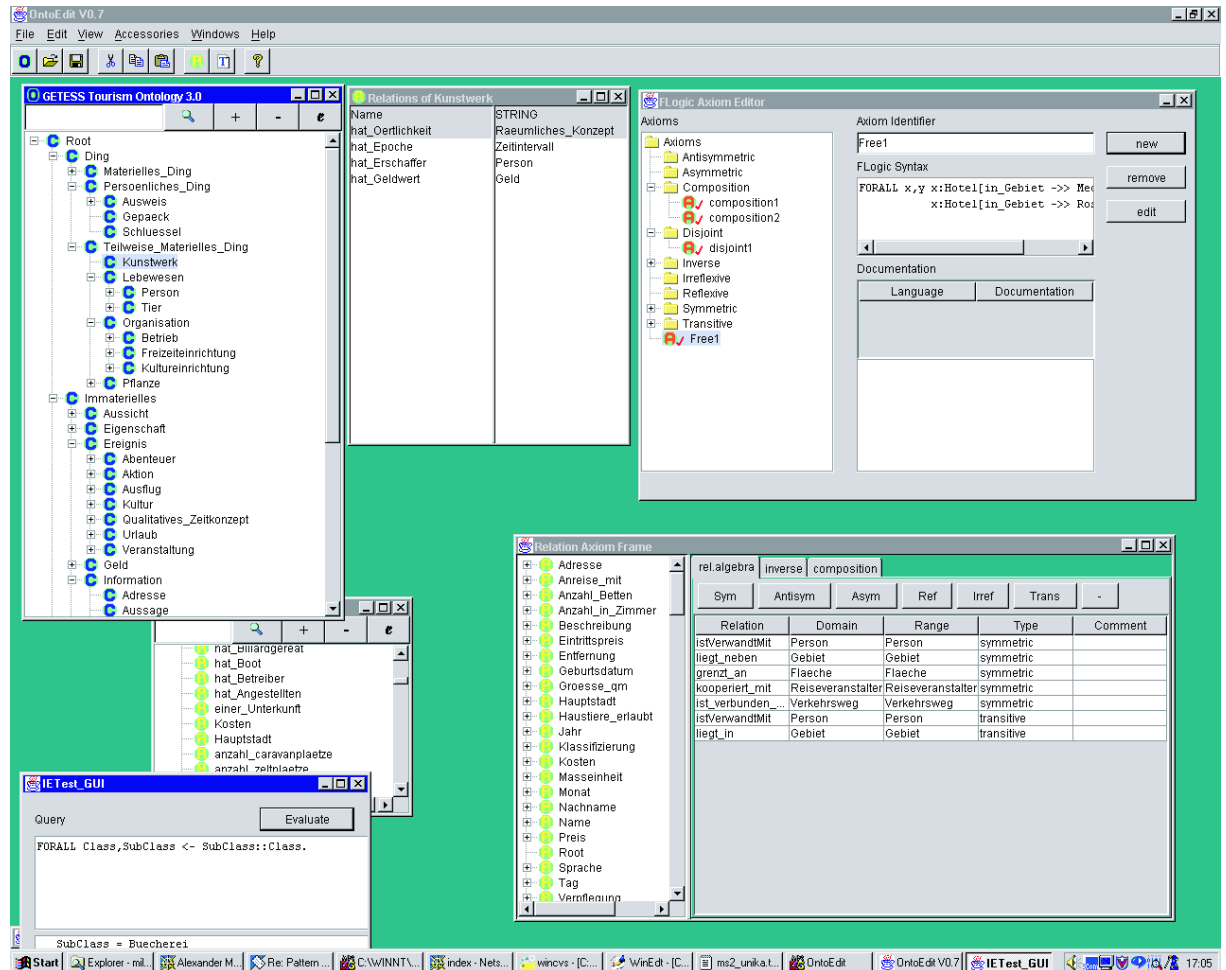


Abbildung 2: OntoEdit

2. OntoRel wurde so gestaltet, dass die Anfragebedürfnisse (z.B. nach Vererbung von Relationen) behandelt werden konnten.

1.9 AP 11: Evaluation

Für die Evaluierung der Ontologien wurden verschiedenen Methoden verwendet und teilweise neu entwickelt:

1. Mittels Fragebögen wurde die Vollständigkeit der bereits modellierten Konzepte überprüft. Hierbei konnte die Abdeckung der wichtigsten Kategorien bestätigt werden. Darüberhinaus ergaben sich aus den Fragebögen wertvolle Hinweise für die Vervollständigung durch Spezialbegriffe.

Abbildung 3: Text2Onto

- Es wurde eine Kreuzevaluierung von Ontologien anhand von paralleler Modellierung durchgeführt. Verschiedenen Personen entwickelten parallel an Tourismusontologien. Deren Resultate wurden in einem zweiten Schritt miteinander verglichen (vgl. [22]). Dadurch konnte festgestellt werden, dass die Sichten dieser Ontologien in wesentlichen Punkten übereinstimmte, allerdings im Detail und im Top Level auch wesentlich voneinander abwich.

1.10 AP 12: Anpassung/Erweiterung der Ontologie für Finanzdomäne

Im Rahmen von Arbeitspaket 12 wurde der bereits erwähnte Top Level wiederverwendet. Es wurden für die Domäne Finanzen 1257 Konzepte und 270 Relationen erhoben. Zu diesen Konzepten wurden 2964 lexikalische Einträge modelliert und 1257 Einträge für Darstellungen in der Dialogkomponente. Auf Axiome wurde in diesem Fall verzichtet, da umfangreiche Test — anders als in der Tourismusdomäne — in der Finanzdomäne nicht mehr möglich waren.

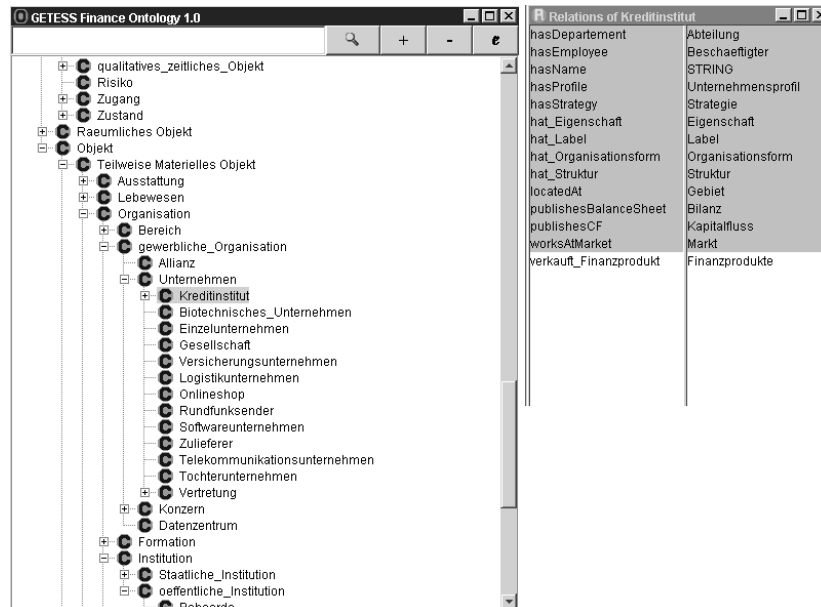


Abbildung 4: Ausschnitt aus der Taxonomie der Finanzontologie (links), sowie der Sicht auf die nicht-taxonomische Relationen (rechts)

1.11 AP 13: Berücksichtigung multilingualer Aspekte

In der Tourismusdomäne wurden deutsche und englische Bezeichnungen modelliert. Grundlage hierfür bildete die Ontologiedefinition wie in Abschnitt 1.3 beschrieben. Nötig waren wiederum die Erweiterung von OntoEdit um multilinguale Aspekte.

1.12 Zusammenfassung der Projektergebnisse

Im Projekt GETESS entstanden zahlreiche theoretische und praktische Erkenntnisse in den Bereichen:

- Ontologieentwicklung
- Interaktion von Ontologie mit Linguistik (z.B., Bootstrapping beider Komponenten, [17])
- Interaktion von Ontologie mit Dialogkomponente (z.B. Lexikalische Einträge für Ontologiestruktur, [25])
- Interaktion von Ontologie mit Datenbank (z.B. Semantic Patterns, [48])

Viele der angegangenen Problemstellungen konnten wie im Projektplan vorgezeichnet realisiert werden. Einige Probleme im Bereich Ontologien und seiner Randgebiete entstanden daraus, dass Wissen und Werkzeuge teilweise erst im Projekt entwickelt werden mussten. Dazu gehörten

z.B. eine Ontologieentwicklungsmethodik oder ein Werkzeug wie OntoEdit. Einige Probleme blieben offen wie z.B. die Skalierung der Inferenzmaschine auf Datenbankfähigkeiten oder eine Gesamtevaluierung des Systems in Hinblick auf die Leistungsfähigkeit des Systems. An anderen Stellen konnte dafür ein deutlicher wissenschaftlicher und praktischer Mehrwert erzielt werden. Zu nennen sei hier nur beispielhaft das Ontology Learning oder OntoKick.

2 Veröffentlichungen, Vorträge und Veranstaltungen

2.1 Veröffentlichungen

- [1] J. Angele, H.-P. Schnurr, S. Staab, and R. Studer. The times they are a-changin' — the corporate history analyzer. In D. Mahling and U. Reimer, editors, *Proceedings of the Third International Conference on Practical Aspects of Knowledge Management. Basel, Switzerland, October 30-31, 2000*, 2000. <http://www.research.swisslife.ch/pakm2000/>.
- [2] T. Barnekow, S. Staab, J. Ziegler, and R. Studer. An architecture for recovering business events bottom-up. In *HCI '99 - Proceedings of the 8th International Conference on Human-Computer Interaction, Munich, Germany, August, 22-27, 1999*, pages 614–618, London, 1999. Lawrence Erlbaum.
- [3] Nils Bräunlich. Visualisierung von Konzeptgraphen mittels Abbildungen von einer hyperbolischen in die Euklidische Ebene. Master's thesis, Institut AIFB, Universität Karlsruhe, 2000.
- [4] Stefan Decker, Jan Jannink, Prasenjit Mitra, S. Staab, R. Studer, and Gio Wiederhold. An information food chain for advanced applications on the www. In *ECDL 2000 — Proceedings of the Fourth European Conference on Research and Advanced Technology for Digital Libraries. Lisbon, Portugal, September 18-20, 2000*, LNCS. Springer, 2000.
- [5] M. Erdmann, A. Maedche, H.-P. Schnurr, and S. Staab. From manual to semi-automatic semantic annotation: About ontology-based text annotation tools. In P. Buitelaar and K. Hasida, editors, *Proceedings of the COLING 2000 Workshop on Semantic Annotation and Intelligent Content*, 2000.
- [6] M. Erdmann, A. Maedche, H.-P. Schnurr, and S. Staab. Semantic annotation: About ontology-based text annotation tools. volume 6, 2001.
- [7] D. Fensel, J. Angele, S. Decker, M. Erdmann, H.-P. Schnurr, S. Staab, and R. Studer. On2broker: Semantic-based access to information sources at the www. In *Proceedings of the World Conference on the WWW and Internet (WebNet 99), Honolulu, Hawaii, USA, October 25-30, 1999*, 1999.
- [8] S. Handschuh, S. Staab, and A. Maedche. Cream — creating relational metadata with a component-based, ontology-driven annotation framework. In *SWWS 2001 — Semantic Web Working Symposium. Stanford, CA, July 30-August 1, 2001*, 2001.

- [9] I. Horrocks, D. Fensel, J. Broekstra, S. Decker, M. Erdmann, C. Goble, F. Van Harmelen, M. Klein, S. Staab, and R. Studer. The ontology interchange language oil: The grease between ontologies. Technical report, Dep. of Computer Science, Univ. of Manchester, UK/ Vrije Universiteit Amsterdam, NL/ Administrator, Nederland B.V./ AIFB, Univ. of Karlsruhe, DE, 2000. <http://www.cs.vu.nl/~dieter/oil/>.
- [10] A. Hotho, A. Maedche, and S. Staab. Ontology-based text clustering. In *Proceedings of the IJCAI-01 Workshop on "Text Learning: Beyond Supervision"*. Seattle, WA, USA, August 6, 2001, 2001.
- [11] A. Hotho, A. Maedche, and S. Staab. Text clustering based on good aggregations. In *Proc. of ICDM — IEEE International Conference on Data Mining 2001, Vancouver, December 2001*, 2001.
- [12] A. Hotho, A. Maedche, S. Staab, and R. Studer. Seal-ii — the soft spot between richly structured and unstructured knowledge. *Journal of Universal Computer Science*, 7(7):566–590, 2001.
- [13] J. U. Kietz, A. Maedche, and R. Volz. Extracting a domain-specific ontology learning from a corporate intranet. In *Second "Learning Language In Logic" LLL Workshop, co-located with the International Conference in Grammere Inference (ICGI'2000) and Conference on Natural Language Learning (CoNLL'2000)*, Lisbon, Portugal, September 13-14 2000.
- [14] M. Klettke, M. Bietz, I. Bruder, A. Heuer, D. Priebe, G. Neumann, M. Becker, J. Bedersdorfer, H. Uszkoreit, A. Maedche, S. Staab, and R. Studer. Getess — ontologien, objek-relationale datenbanken und textanalyse als bausteine einer semantischen suchmaschine. *Datenbank-Spektrum*, 1(1):14–24, 2001.
- [15] A. Mädche, H.-P. Schnurr, S. Staab, and R. Studer. Representation language-neutral modeling of ontologies. In *Modellierung-2000. Koblenz, Deutschland, April 5-7, 2000*. Fölbach-Verlag, 2000.
- [16] A. Maedche. *Ontology Learning for the Semantic Web*. Kluwer Academic Publishers, 2002.
- [17] A. Maedche, G. Neumann, and S. Staab. Bootstrapping an ontology-based information extraction system for the web. In P.S. Szczepaniak, J. Segovia, J. Kacprzyk, and L.A. Zadeh, editors, *Intelligent Exploration of the Web*, Studies in Fuzziness and Soft Computing, Heidelberg, 2002. Physica-Verlag.
- [18] A. Maedche and S. Staab. Learning ontologies for the semantic web. In *Proc. of Semantic Web 2001 — Workshop at WWW-10. May 1, Hongkong, China*, pages 51–60.
- [19] A. Maedche and S. Staab. Discovering conceptual relations from text. In *Proceedings of ECAI-2000. Berlin, August 21-25, 2000*. IOS Press, 2000.
- [20] A. Maedche and S. Staab. Mining Ontologies from Text. In *Proceedings of EKAW-2000, Springer Lecture Notes in Artificial Intelligence (LNAI-1937), Juan-Les-Pins, France, 2000*. Springer, 2000.

-
- [21] A. Maedche and S. Staab. Semi-automatic engineering of ontologies from text. In *Proceedings of the Twelfth International Conference on Software Engineering and Knowledge Engineering (SEKE'2000)*. Chicago, July 6-8, 2000, 2000.
- [22] A. Maedche and S. Staab. Comparing ontologies – similarity measures and a evaluation study. Technical report, 2001.
- [23] A. Maedche and S. Staab. Ontology learning for the semantic web. *IEEE Intelligent Systems*, 16(2):72–79, 2001.
- [24] A. Maedche and S. Staab. Ontology learning from text. In M. Bouzeghoub, Z. Kedad, and E. Metais, editors, *Natural Language Processing and Information Systems. 5th International Conference on Applications of Natural Language to Information Systems, NLDB 2000, Versailles, France, June 28-30, 2000*, LNCS 1959, page 364. Springer, 2001.
- [25] A. Maedche, S. Staab, N. Stojanovic, R. Studer, and Y. Sure. SEMantic portAL — the seal approach. In D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster, editors, *Creating the Semantic Web*, Cambridge, MA, 2002. MIT Press.
- [26] A. Maedche, S. Staab, and R. Studer. Ontologien. *Wirtschaftsinformatik*, 43(4):393–396.
- [27] S. Melzig, J. Pretzel, G. Neumann, M. Becker, H. Uszkoreit, A. Maedche, S. Staab, R. Studer, Y. Sure, A. Düsterhöft, A. Heuer, M. Klettke, and D. Priebe. BMBF-Projekt GETESS. GERman Text Exploitation and Search System. Arbeitsbereich zum Meileinstein 3. Technical report, Juni 2000.
- [28] D. Oberle. Semantic community web portals - personalization. Master's thesis, Institute AIFB, University of Karlsruhe, 2000.
- [29] U. Reimer and D. Mahling, editors. *PAKM-2000 — Proceedings of the Third Conference on Practical Aspects of Knowledge Management*, volume 34 of *CEUR Workshop Proceedings*, 2000. Available from <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/>.
- [30] H.-P. Schnurr, S. Staab, and R. Studer. Ontology-based process support. In *Workshop on Exploring Synergies of Knowledge Management and Case-Based Reasoning (AAAI-99)*, Menlo Park, CA. AAAI. Technical Report.
- [31] H.-P. Schnurr, S. Staab, R. Studer, G. Stumme, and Y. Sure, editors. *WM2001 — Professionelles Wissensmanagement — Erfahrungen und Visionen. Proceedings der Konferenz WM-2001. Baden Baden, Maerz 14-16, 2001*.
- [32] H.P. Schnurr and S. Staab. A proactive inferencing agent for desk support. In Staab and O'Leary [51].
- [33] S. Staab. *Grading Knowledge — Extracting Degree Information from Texts*. LNCS 1744. Springer Verlag, Heidelberg, 1999.
- [34] S. Staab. From binary temporal relations to non-binary ones and back. *Artificial Intelligence*, 128(1-2), 2001.

- [35] S. Staab. Semantic Web — Das Web der nächsten Generation. *Karlsruher Transfer*, 15:18–23, 2001.
- [36] S. Staab, J. Angele, S. Decker, M. Erdmann, A. Hotho, A. Maedche, H.-P. Schnurr, R. Studer, and Y. Sure. Semantic community web portals. *Proc. of WWW9 / Computer Networks*, 33(1-6):473–491, 2000.
- [37] S. Staab, J. Angele, S. Decker, M. Erdmann, A. Hotho, A. Maedche, H.-P. Schnurr, R. Studer, and Y. Sure. AI for the Web — Ontology-based community Web portals. In *Proc. of AAAI 2000/IAAI 2000 - Proceedings of the 17th National Conference on Artificial Intelligence and 12th Innovative Applications of Artificial Intelligence Conference, Austin/TX, USA, July 30-August 3, 2000*, Menlo Park/CA, Cambridge/MA. AAAI Press/MIT Press.
- [38] S. Staab and T. Barnekow. Towards learning notification triggers. In *Proceedings of the international Workshop on Intelligent Workflow and Process Management: The New Frontier for AI in Business (IJCAI-99)*. Stockholm, Sweden, 1999, 1999.
- [39] S. Staab, C. Braun, I. Bruder, A. Düsterhöft, A. Heuer, M. Klettke, G. Neumann, B. Prager, J. Pretzel, H.-P. Schnurr, R. Studer, H. Uszkoreit, and B. Wrenger. A system for facilitating and enhancing web search. In *IWANN '99 - Proceedings of International Working Conference on Artificial and Natural Neural Networks. Volume 2. Alicante, ES, June 2-4, 1999*, LNCS 1607, pages 706–714, Berlin, Heidelberg, 1999. Springer.
- [40] S. Staab, C. Braun, A. Düsterhöft, A. Heuer, M. Klettke, S. Melzig, G. Neumann, B. Prager, J. Pretzel, H.-P. Schnurr, R. Studer, H. Uszkoreit, and B. Wrenger. GETESS — searching the web exploiting german texts. In *CIA '99 — Proceedings of the 3rd Workshop on Cooperative Information Agents. Upsala, Sweden, July 31-August 2, 1999*, LNCS 1652, pages 113–124, Berlin, 1999. Springer.
- [41] S. Staab, M. Erdmann, and A. Maedche. Engineering ontologies using semantic patterns. In A. Preece and D. O'Leary, editors, *Proceedings of the IJCAI-01 Workshop on E-Business & the Intelligent Web. Seattle, WA, USA, August 5, 2001*, 2001.
- [42] S. Staab, M. Erdmann, and A. Maedche. Ontologies in RDF(S). *ETAI Journal*, 6, 2001.
- [43] S. Staab, M. Erdmann, A. Maedche, and S. Decker. An extensible approach for modeling ontologies in RDF(S). In V. Christophides and D. Plexousakis, editors, *Proceedings of the ECDL-2000 Workshop — Semantic Web: Models, Architectures and Management*, 2000.
- [44] S. Staab, M. Erdmann, A. Maedche, and S. Decker. An extensible approach for modeling ontologies in rdf(s). In R. Grütter, editor, *Knowledge Media in Healthcare: Opportunities and Challenges*, page 231, Hershey, USA, 2001. Idea Group Publishing.
- [45] S. Staab and U. Hahn. Scalable temporal reasoning. In *IJCAI-99 — Proceedings of the 16th International Joint Conference on Artificial Intelligence, Stockholm, Schweden, August, 1-6, 1999*, pages 1247–1252, San Francisco/CA, 1999. Morgan Kaufmann.

-
- [46] S. Staab, S. Handschuh, and A. Maedche. Metadata and the semantic web — and cream. In *Proceedings of the DELOS-2001 workshop. September 8-9, 2001, Darmstadt*. ECRIM, 2001.
- [47] S. Staab and A. Maedche. Knowledge portals — ontologies at work. *AI Magazine*, 22(2):63.
- [48] S. Staab and A. Maedche. Axioms are objects, too - ontology engineering beyond the modeling of concepts and relations. In V.R. Benjamins, A. Gomez-Perez, and N. Guarino, editors, *Proceedings of the ECAI 2000 Workshop on Ontologies and Problem-Solving Methods. Berlin, August 21-22, 2000*, 2000.
- [49] S. Staab, A. Maedche, and S. Handschuh. An annotation framework for the semantic web. In S. Ishizaki, editor, *Proc. of The First International Workshop on Multimedia Annotation, January 30 - 31, 2001. Tokyo, Japan*, 2001.
- [50] S. Staab, A. Maedche, C. Nedellec, and P. Wiemer-Hastings, editors. *ECAI'2000 Workshop on Ontology Learning (OL'2000). Berlin, Germany, August 25, 2000*, volume 31 of *CEUR Proceedings*, 2000.
- [51] S. Staab and D. O'Leary, editors. *Bringing Knowledge to Business Processes. Papers from 2000 AAAI Spring Symposium*, Technical Report SS-00-03, Menlo Park, CA, 2000. AAAI Press.
- [52] S. Staab and H.-P. Schnurr. Knowledge and business processes: Approaching an integration. In *OM '99 — Proceedings of the international Workshop on Knowledge Management and Organizational Memory (IJCAI-99). Stockholm, Sweden, 1999*, pages 140–159, 1999.
- [53] S. Staab and H.-P. Schnurr. Smart task support through proactive access to organizational memory. *Knowledge-based Systems*, 13(5):251–260, September 2000.
- [54] S. Staab, H.-P. Schnurr, R. Studer, and Y. Sure. Knowledge processes and ontologies. *IEEE Intelligent Systems*, 16(1):26.
- [55] S. Staab and R. Studer. Intelligente (symbolische) Methoden für das Wissensmanagement. In Radermacher, editor, *Management von nicht-explizitem Wissen: Noch mehr von der Natur lernen. Abschlussbericht — Teil 3 - Die Sicht verschiedener akademischer Fächer zum Thema des nicht-expliziten Wissens*, pages 165–180. FAW Ulm, 2001.
- [56] S. Staab, R. Studer, and Y. Sure. Knowledge processes and meta processes in ontology-based knowledge management. In C. Holsapple, editor, *Handbook on Knowledge Management*, International Handbooks on Information Systems, Heidelberg, 2002. Springer Verlag.
- [57] L. Stojanovic, S. Staab, and R. Studer. eLearning based on the semantic web. In *WebNet 2001 — World Conference on the WWW and the Internet. Orlando, Florida, USA, Oct. 23-27, 2001*, 2001.
- [58] N. Stojanovic, A. Maedche, and R. Studer. Seal — a framework for developing SEMantic portALs. In *K-CAP 2001 — Proceedings of the First International ACM Conference on Knowledge Capture. October 21-23, 2001, Victoria, B.C., Canada*. ACM Press, 2001.

- [59] R. Studer, S. Decker, D. Fensel, and S. Staab. Knowledge engineering — Survey and future. In *Knowledge Engineering and Agent Technologies*. IOS Press, 2000.
- [60] R. Studer, S. Decker, D. Fensel, and S. Staab. Situation and prospective of knowledge engineering. In J. Cuenca, Y. Demazeau, A. Garcia, and J. Treur, editors, *Knowledge Engineering and Agent Technology*, IOS Series on Frontiers in Artificial Intelligence and Applications 52. IOS Press, 2000.
- [61] Y. Sure, A. Maedche, and S. Staab. Leveraging corporate skill knowledge - From ProPer to OntoProper. In D. Mahling and U. Reimer, editors, *Proceedings of the Third International Conference on Practical Aspects of Knowledge Management. Basel, Switzerland, October 30-31, 2000*, 2000. <http://www.research.swisslife.ch/pakm2000/>.
- [62] R. Volz. Ontologieakquisition aus Intranetdokumenten. Master's thesis, Institute AIFB, University of Karlsruhe, 2000.

2.2 Vorträge

- *Eingeladener Vortrag von S. Staab: Ontology-based Knowledge Management. Seminar of the Knowledge Management Research Group, British Telecom, Ipswich, 20.1.2000*
- *Vortrag von S. Staab: A Proactive Inferencing Agent for Desk Support (Vgl. [32]). AAAI Spring Symposium on Bringing Knowledge to Business Processes, Stanford, CA, 20.3.2000*
- *Vortrag von Prof. Dr. R. Studer: Semantic for the Web. IuK Meeting "Information, Knowledge and Knowledge Management", Darmstadt, 27.03.2000*
- *Vortrag von S. Staab: Representation Language-Neutral Modeling of Ontologies (Vgl. [15]). Modellierung-2000, St. Goar, 7.4.2000*
- *Eingeladener Vortrag von S. Staab: Intelligente Techniken für das Wissensmanagement. 5. AIK-Symposium "Wissensmanagement", Verein für Angewandte Informatik Karlsruhe e.V. (AIK), Karlsruhe, 5.5.2000*
- *Vortrag von S. Staab: Semantic Community Web Portals (Vgl. [36]). WWW9 - 9th International World Wide Web Conference, W3C, Amsterdam, The Netherlands, 17.5.2000.*
- *Vortrag von Y. Sure: Internet- und Intranetbasiertes Wissensmanagement. Unischau 2000, Universität Karlsruhe, Karlsruhe, 1.7.2000*
- *Vortrag von S. Staab: Ontology-based Knowledge Management. Dagstuhl Seminar on Knowledge Management, Dagstuhl, 11.7.2000*
- *Vortrag von S. Staab: AI for the Web - Ontology-based Community Web Portals (Vgl. [37]). AAAI 2000 / IAAI 2000, American Association for Artificial Intelligence, Austin, TX, 1.8.2000*

- *Vortrag von S. Staab: Axioms are objects, too - Ontology Engineering Beyond the Modeling of Concepts and Relations. (Vgl. [48]). ECAI 2000 Workshop on Ontologies and Problem-Solving Methods, Berlin, 21.8.2000 .*
- *Vortrag von A. Maedche: Discovering Conceptual Relations from Text. ECAI 2000, Berlin, 22.8.2000 .*
- *Tutorium von S. Staab: ECAI-2000 Tutorial "AI Techniques for Knowledge Management". European Conference on Artificial Intelligence, ECCAI, Berlin, 21.8.2000*
- *Vortrag von S. Staab: An extensible approach for Modeling Ontologies in RDF(S) (Vgl. [43]). ECDL-2000 Workshop "Semantic Web: Models, Architectures and Management", Lisbon, Portugal, 21.9.2000*
- *Tutorium von A. Maedche: Entwicklung und Anwendung von Ontologien. Fachgruppentreffen Maschinelles Lernen FGML-2000, Bonn, September 2000*
- *Vortrag von Y. Sure: Per Anhalter durchs Internet. WIKOM Sommerfest, WIKOM Verlag, Wegscheid, 22.9.2000*
- *Vortrag von S. Staab: The Times They Are A-Changin' - The Corporate History Analyzer (Vgl. [1]). Third International Conference on Practical Aspects of Knowledge Management, Basel, CH, 30.10.2000*
- *Vortrag von Y. Sure: Leveraging Corporate Skill Knowledge - From ProPer to OntoPro-Per (Vgl. [61]). PAKM 2000 - Third International Conference on Practical Aspects of Knowledge Management, Reimer, U., Basel, Switzerland, 31.10.2000*
- *Vortrag von A. Maedche: Mining Ontologies from Text. EKAW 2000, Juans-Les-Pins, Frankreich, 12.10.2000*
- *Vortrag von Y. Sure: On-To-Knowledge: Content-driven Knowledge Management through Evolving Ontologies. KM Europe 2000, Brussels, Belgium, 20.11.2000*
- *Vortrag von Prof. Dr. R. Studer: Beyond RDF(S): The Ontology Perspective for the Semantic Web, Semantic Web Technologies EU-Workshop Luxembourg, 22-23.11.2000*
- *Tutorium von S. Staab: Wissensdatenbanken für den Help Desk. Wissensmanagement in der IT, Management Circle, Oberursel, 13.12.2000*
- *Vortrag von Prof. Dr. R. Studer: The Semantic Web – Ringvorlesung E-Commerce, E-Business, RWTH Aachen, 18.12.2000*
- *Vortrag von A. Maedche: An Annotation Framework for the Semantic Web. First International Workshop on Multimedia Annotation, Tokio, Japan, 29.01.2001*
- *Eingeladener Vortrag Knowledge Portals. Auf dem zweiten HERMES Workshop über Knowledge-based systems for the delivery and adaptation of educational courseware, Patras, Greece, 17 Februar 2001*

- *Eingeladener Vortrag Knowledge Portals - Taking advantage of Human Language Technology 2001 EACL/ACL Workshop on Human Language Technology and Knowledge Management (HLT & KM), Toulouse, July 6-7 at ACL '01*

2.3 Veranstaltungen

- *Bringing Knowledge to Business Processes. 2000 AAAI Spring Symposium, Stanford, March 20-22, 2000. Steffen Staab und Dan O'Leary.*
- *First Workshop on Ontology Learning OL'2000, Berlin, Germany, August 25, 2000. Steffen Staab, Alexander Mädche, Claire Nedellec, Peter Wiemer-Hastings.*
- *Tutorial zu "AI Techniques for Knowledge Management", Berlin, Germany, August 24, 2000. Steffen Staab, Stefan Decker.*
- *Dagstuhl Workshop on Knowledge Management. Rudi Studer, Dan O'Leary, Matthias Jarke.*
- *Vorlesung "Knowledge Discovery" Wintersemester 1999/2000. Steffen Staab und Rudi Studer.*
- *Seminar "Text Mining", Wintersemester 1999/2000, Universität Karlsruhe (TH). Alexander Mädche, Steffen Staab und Rudi Studer.*
- *Seminar "Ontology Engineering", Sommersemester 2000, Universität Karlsruhe (TH). Alexander Mädche und Rudi Studer.*
- *Vorlesung "Intelligente Systeme im WWW", Sommersemester 2000, Universität Karlsruhe (TH). Steffen Staab.*
- *Vorlesung "Knowledge Discovery" Wintersemester 2000/2001. Steffen Staab und Gerd Stumme.*
- *Vorlesung "Angewandte Informatik I" Wintersemester 2000/2001. Rudi Studer und Alexander Maedche.*
- *Vorlesung "Intelligente Systeme im WWW", Sommersemester 2001, Universität Karlsruhe (TH). Steffen Staab.*
- *WM2001 Erste deutsche Konferenz für Wissensmanagement, Baden-Baden, März 14-16, 2001, Steffen Staab, Rudi Studer*
- *Semantic Web Workshop (at WWW10), Hong Kong, China, May 1, 2001, Steffen Staab, Stefan Decker, Dieter Fensel, Amit Sheth.*
- *Tutorial zu "Ontologien", Fachgruppentreffen Machine Learning, September 2000, Alexander Mädche*
- *Tutorial zu "Ontologien", European Conference on Machine Learning, Freiburg, September 3, 2001, Alexander Mädche*

- *Second Workshop on Ontology Learning OL'2001, Seattle, USA, August 4, 2001. Steffen Staab, Alexander Mädche, Claire Nedellec, Ed Hovy.*
- *Tutorial zu "AI Techniques for Knowledge Management", Seattle, USA, August 5, 2001. Steffen Staab, Stefan Decker.*

3 Erfindungen, Schutzrechtsanmeldungen, erteilte Schutzrechte

Es gab keine Erfindungen, Schutzrechtsanmeldungen oder erteilten Schutzrechte.

4 Transfer in die Praxis

Im Laufe des Projektes ergaben sich zahlreiche Kontakte mit Unternehmen, die an ontologiebasierten Techniken für zahlreiche Anwendungsfelder interessiert sind. Besonders erwähnenswert erscheinen uns allerdings folgende beiden Ereignisse:

- Im Dezember 2000 wurde von Prof. Dr. Rudi Studer am Forschungszentrum Informatik an der Universität Karlsruhe eine Abteilung "Wissensmanagement" gegründet, die seit Juni 2001 operativ ist. Die Abteilung betreibt Technologietransfer vor allem im Bereich "Ontologien und Wissensmanagement". Sie wird operativ geleitet von Herrn Dipl.-Wi.-Ing. Alexander Mädche, einem ehemaligen GETESS-Mitarbeiter. Dort erfolgt nun z.B. die Anwendung der in GETESS erzielten Ergebnisse auf Bereiche wie *Life Sciences* (in Kooperation mit Sciconis), auf *Human Resource Management* (in Zusammenarbeit mit DaimlerChrysler, Würth) und auf virtuelle Organisationen (im BMBF-Projekt Consense).
- Im Juni 1999 wurde Ontoprise GmbH als ein Spin-Off des Instituts AIFB gegründet, das Ontologiebasierte Techniken entwickelt und in die Praxis trägt. Ontoprise wurde im August 2000 operativ als die beiden Geschäftsführer Prof. Dr. Jürgen Angele und Dipl.-Wi.-Ing. Hans-Peter Schnurr, ein ehemaliger GETESS-Mitarbeiter, mit zwei Angestellten die Arbeit in der Technologiefabrik Karlsruhe aufnahmen. Ontoprise beschäftigt derzeit 14 voll angestellte Mitarbeiter — darunter einige ehemalige GETESS-Hilfswissenschaftler. Seit Juni 2001 wird Ontoprise durch die Triangle GmbH, ein Venture Capital-Unternehmen, finanziert.

Ontoprise verwendet die im GETESS-Projekt gewonnenen und publizierten Erkenntnisse. So wurde zum Beispiel basierend auf den Erfahrungen mit OntoEdit eine vollständig neue, kommerzielle Version des Werkzeugs entwickelt und Lücken bei der Integration von Inferenzmaschine und Datenbanken werden in der kommerziellen Version des Ontobrokers bald gelöst sein.