

DASMOD Workshop Februar/März 2007

ADIB

Adaptive Informationsbereitstellung

Heiko Maus, DFKI FB Wissensmanagement

DFKI
und

AG Bildverstehen und Mustererkennung

Prof. Dr. Thomas Breuel

AG Wissensbasierte Systeme

Prof. Dr. Andreas Dengel

<http://www.dasmod.de/twiki/bin/view/DASMOD/ADIB>



Grunddaten

- ▶ **Projekt:** ADIB – Adaptive Informationsbereitstellung
- ▶ **Projektlaufzeit:** 01.12.2005 – 31.12.2007

- ▶ **Beteiligte Personen**

Prof. Dr. Andreas Dengel	(TU Kaiserslautern, DFKI FB WM)
Prof. Dr. Thomas Breuel	(TU Kaiserslautern, DFKI FB IUPR)
Heiko Maus	(DFKI FB WM)
Joost van Beusekom	(TU Kaiserslautern)
Kinga Schumacher	(DFKI FB WM)
Dr. Armin Stahl	(DFKI FB IUPR)
Dr. Harald Holz	(DFKI FB WM)

- ▶ **Kooperationspartner:**

EU IP Projekt NEPOMUK
BMBF Projekte Mymory, IPeT;
Prof. Hiroshi Tsuji, Osaka Prefecture University, Japan
Mark Siebert, SBS München



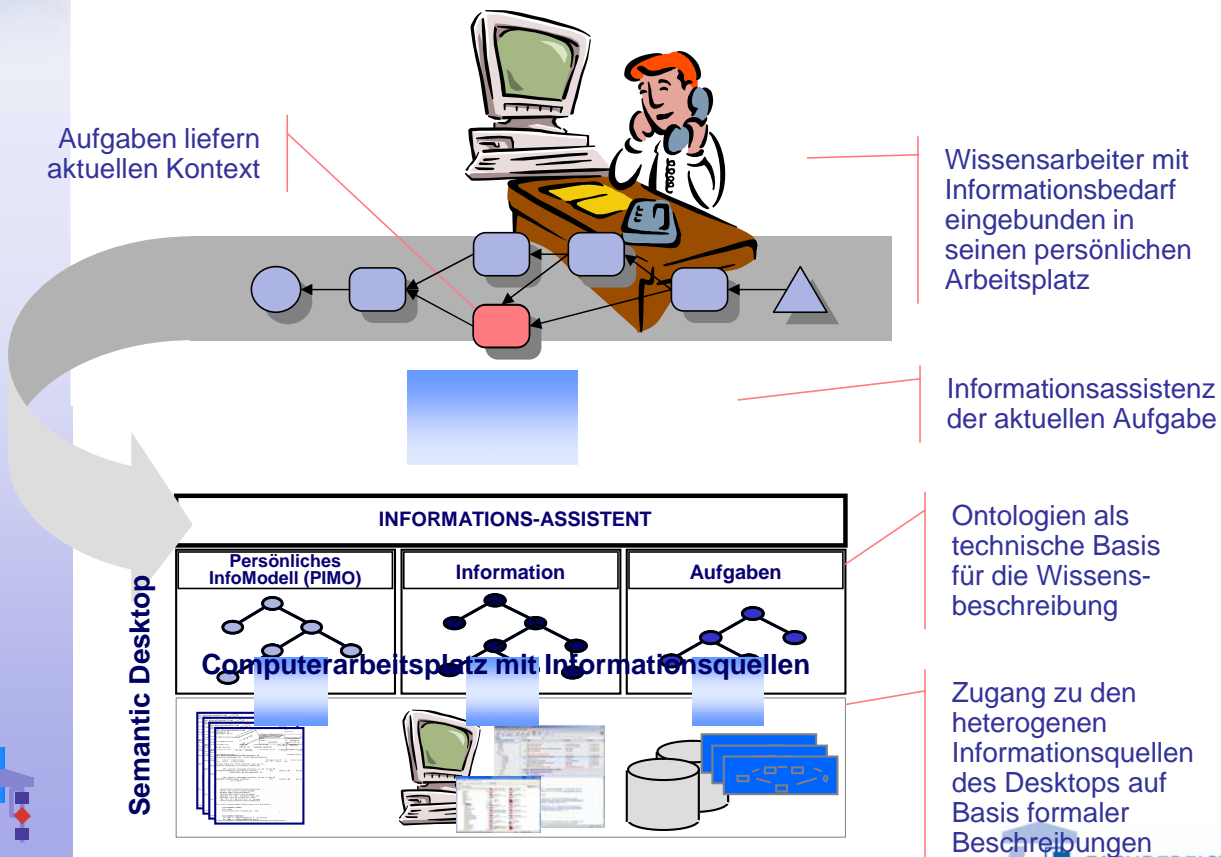
ADIB realisiert eine an den Nutzer angepasste Assistenz für wissensintensive Aufgaben im Semantic Desktop

DJ
Zlvvhqvedvlhuwh
V|vwhph

SZENARIO



© 2007 DFKI S. 3



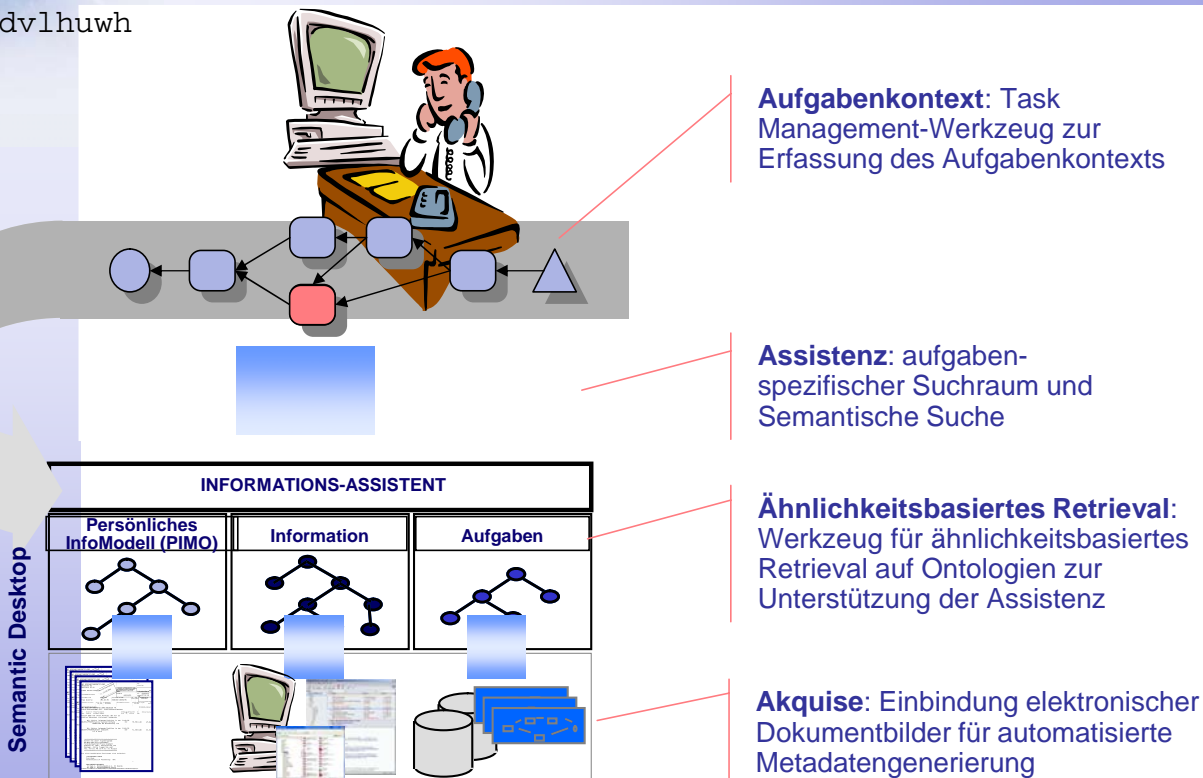
Wesentliche Arbeitspakete im Gesamtzusammenhang

DJ
Zlvvhqvedvlhuwh
V|vwhph

GLIEDERUNG



© 2007 DFKI S. 4

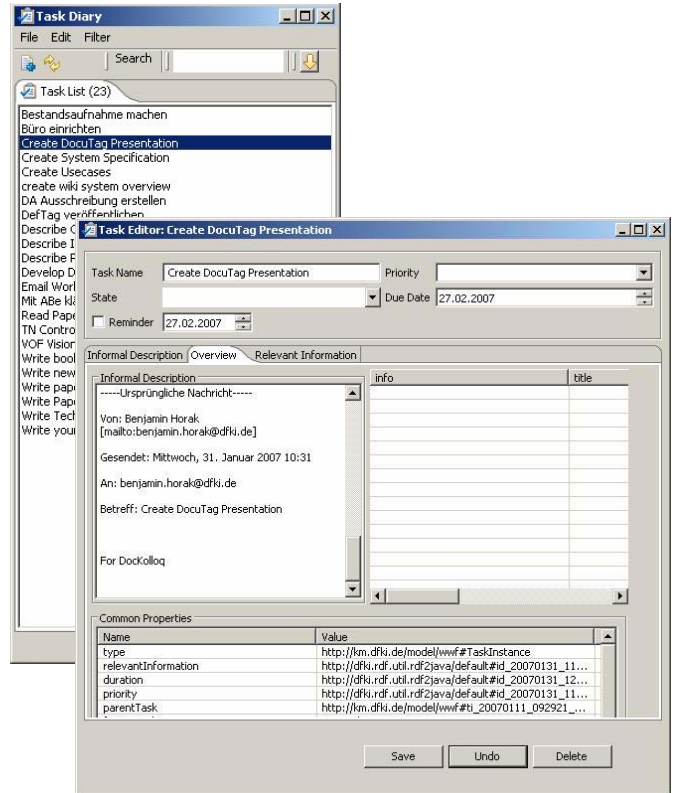


Der Aufgabenkontext wird mit dem TaskPad erfasst

DJ
Zl|vvhqvedvlhuwh
V|vwhph

AUFGABENKONTEXT

- ▶ TaskPad ist ein Werkzeug für das persönliche Aufgaben-Management
- ▶ Einbindung und Synchronisation verschiedener Aufgaben-quellen
 - ▶ bspw. Workflow-System, Textquellen (,TODO' in Quelltexten)
- ▶ Lieferung des Aufgabenkontexts
 - ▶ Name der Aufgabe, textuelle Beschreibung, Status, Informationsobjekte, Attribute, ...
 - ▶ Kontext der Nutzung (Suchen, geöffnete Dokumente, gebrowste Seiten, etc.)



Assistenz findet hier statt



© 2007 DFKI S. 5

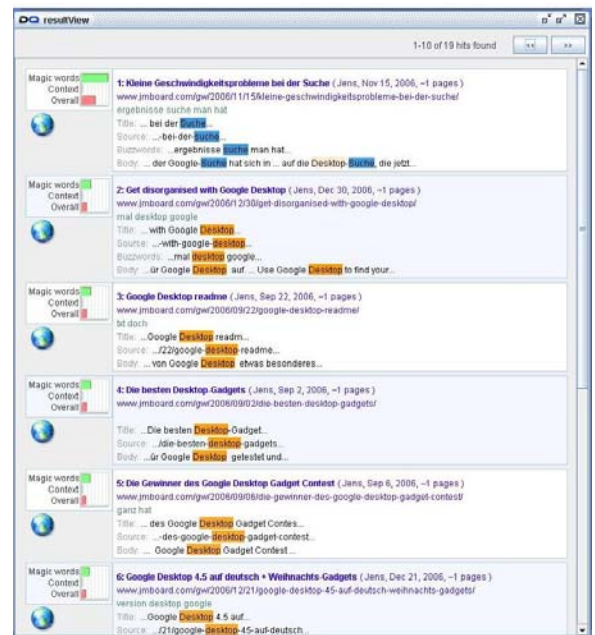


Aufgabenkontext unterstützt die Suche

DJ
Zl|vvhqvedvlhuwh
V|vwhph

ASSISTENZ

- ▶ Einbindung von DynaQ (Dynamic Queries)
 - ▶ Suche als Prozess mit unterschiedlicher Gewichtung der Suchterme
 - ▶ Berücksichtigung von Kontextdokumenten
- ▶ Ermöglicht proaktive kontextabhängige Informationsbereitstellung und Suchen im Kontext der Aufgabe
 - ▶ Ergebnisdokumente können an die Aufgabe angehängt werden
- ▶ Ziel:
 - ▶ Suchschritte werden zum Kontext der Aufgabe hinzugefügt
 - ▶ Nutzung als Erfahrungswissen
 - ▶ Fälle für ähnlichkeitsbasiertes Retrieval



© 2007 DFKI S. 6



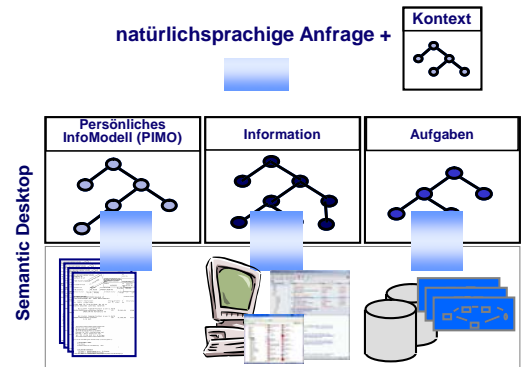
Aufgabenspezifische Semantische Suche

DJ
Zlvvhqvedvlhuwh
V|vwhph

ASSISTENZ



- ▶ Hybride Suche auf Ontologien und Dokumenten unter Berücksichtigung des Aufgabenkontextes des Nutzers
- ▶ Ergebnisse sind
 - ▶ Instanzen bzw. Teilgraphen aus den Ontologien
 - ▶ Dokumente
- ▶ Erster Schritt: Suche auf Ontologien
 - ▶ Führe mit den Anfragetermen eine Suche in der PIMO durch
 - ▶ Bestimme passende Relationen, Klassen und Instanzen
 - ▶ Finde mögliche Aussagen dazu (z.B. $(i_i r_j ?)$)
 - ▶ Gewichte die Ergebnisse entsprechend dem Aufgabenkontext
- ▶ Nächster Schritt
 - ▶ Kontext der Aufgabe als Ausgangspunkt der Suche

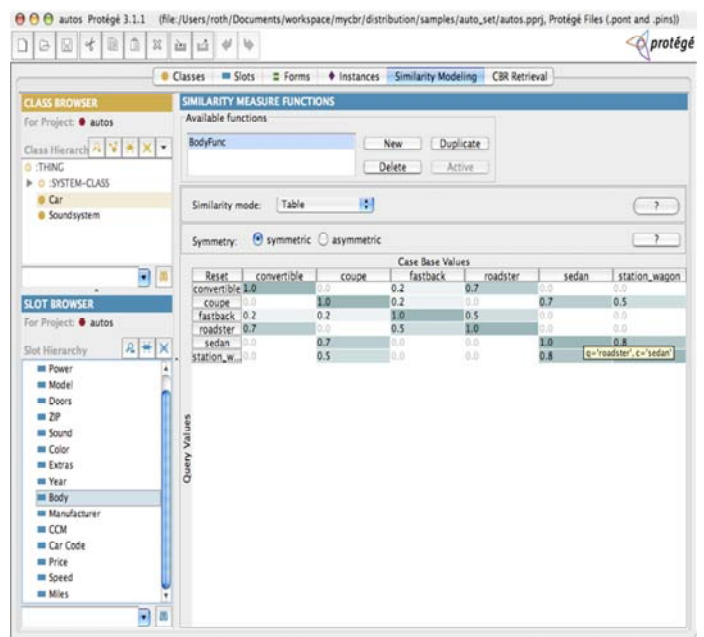


myCBR – Werkzeug zum ähnlichkeitsbasierten Retrieval auf Basis von Ontologien

DJ
Zlvvhqvedvlhuwh
V|vwhph

ÄHNLICHKEITSBASIERTES RETRIEVAL

- ▶ myCBR ist ein Plugin für den Ontologie-Editor Protégé
- ▶ GUIs zur Modellierung von wissensintensiven Ähnlichkeitsmaßen auf Ontologien
- ▶ Erlaubt ähnlichkeitsbasiertes Retrieval
- ▶ Basis zur Bestimmung u.a. von Aufgabenähnlichkeiten zur Informationsassistentz
- ▶ Open Source; Download unter <http://myCBR-project.net>



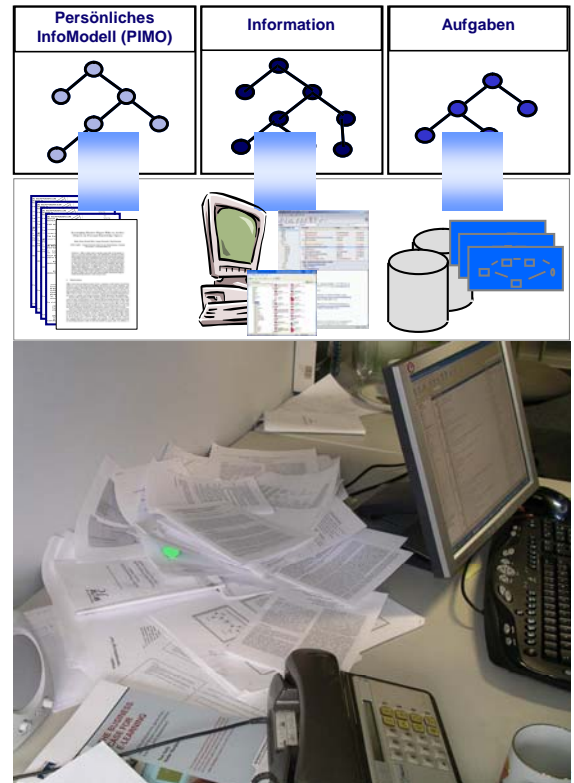
Einbindung elektronischer Dokumentbilder in den Semantic Desktop

DJ
Zlvvhqvedvlhuwh
V|vwhph

AKQUISE



- ▶ Ziel ist die Erweiterung der Informationsakquise auf elektronische Dokumentbilder
- ▶ Extraktion & Explizierung dort vorhandener Informationen:
 - ▶ Layoutelemente
 - ▶ Literaturreferenzen
 - ▶ Dokumentähnlichkeiten



© 2007 DFKI S. 9

FACHBEREICH INFORMATIK

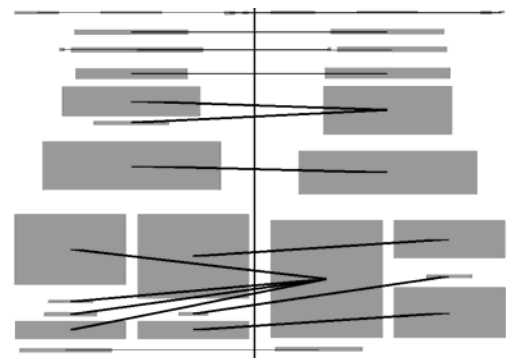
Strukturelle Ähnlichkeit von Dokumenten

DJ
Zlvvhqvedvlhuwh
V|vwhph

AKQUISE



- ▶ Bestimmung der Ähnlichkeit zwischen Dokumenten anhand von Layout-Informationen
 - ▶ Blockbasierte Extraktion geometrischer Merkmale
 - ▶ Bipartite Graphpaarung der Blöcke
- ▶ Ziel ist die Informationsassistentz mittels Dokumenten, die ähnlich zu bereits bei der Aufgabe vorhandenen sind
- ▶ Basis für die Identifikation von Layoutobjekten



© 2007 DFKI S. 10

FACHBEREICH INFORMATIK

Kennzeichnung von Dokumentbildteilen zur Metadatengenerierung

DJ
Zlvvhqvedvlhuwh
V|vwhph

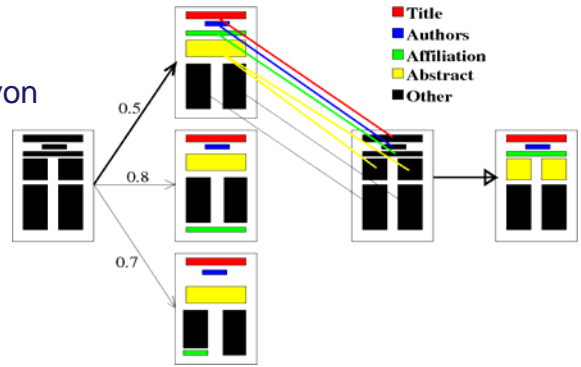
AKQUISE



© 2007 DFKI S. 11

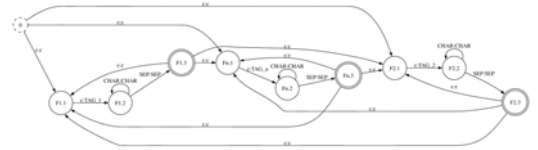
Ähnlichkeitsbasierte Kennzeichnung von Titel, Autor, Abstract und Affilierung

- ▶ Basiert auf ähnlichkeitsbasiertem Retrieval bereits ausgezeichnete Dokumente



Extraktion strukturierter Literaturreferenzen

- ▶ Einsatz von probabilistischen Finite State Transducern zum Lernen und Parsen von Referenzen
- ▶ Ausgabe ist BibTeX



Anreicherung der Metadaten im Semantic Desktop

- ▶ Verbesserung der Ergebnisse der semantischen Suche

Title
Differential effects of 5-HT_{2A} repeated acquisition and performance sequences in monkeys

Author
P.J. Winsauer and J.M. Moerschbaecher

Affiliation
Department of Pharmacology and Experimental Therapeutics
Ochsner, LA, USA

FACHBEREICH INFORMATIK

Semantische Annotation papierbasierter Information



Annotation von Papierdokumenten mit Konzepten aus der persönlichen Wissensbasis

Kombination einer Dokumentkamera (sceye), OCR (Single Click Entry, Océ) und dem Semantic Desktop

Schließt die Lücke zwischen papierbasierter Information und dem Semantic Desktop



© 2007 DFKI

INFORMATIK

DJ
Zlvvhqvedvlhuwh
V|vwhph



- ▶ **Distance Measures for Layout-Based Document Image Retrieval**
Joost van Beusekom, Daniel Keysers, Faisal Shafait, Thomas M. Breuel. In: *International Conference on Document Image Analysis for Libraries (DIAL 2006)*. IEEE, 2006
- ▶ **Personalized document retrieval in multi-party environments with the Semantic Desktop**
Mark Siebert, Pierre Smits, Leo Sauermann, Andreas Dengel. In: *Proceedings EDM, IEEE Int. Workshop on the Electronic Document Management in an Enterprise Computing Environment, Hong Kong, China*. IEEE Computer Society, 2006
- ▶ **Combining Case-Based and Similarity-Based Product Recommendation**
Armin Stahl. In: *Proceedings of the 8th European Conference on Case-Based Reasoning (ECCBR 2006)*. Springer, 2006
- ▶ **Ricoh/Sun-Contest Europe: Embedded Java Applications on MFPs**
 - ▶ Student gewann Deutschlandauswahl mit Applikation
 - ▶ Semantische Annotation gescannter Dokumente direkt an MFPs (Multifunktionalgeräten)
 - ▶ Finale Ende April in Prag

DJ
Zlvvhqvedvlhuwh
V|vwhph



- ▶ **Einbindung der Benutzerobservation aus Mymory**
 - ▶ Liefert detailliertere Kontextinformationen über die Aktionen des Benutzers bei der Aufgabenbearbeitung
- ▶ **Ähnlichkeitsbasiertes Retrieval von angereicherten Aufgaben mittels myCBR**
- ▶ **Integration der Extraktionswerkzeuge in den Semantic Desktop**
- ▶ **Integration weiterer Verfahren zur Informationsassistentz in das TaskPad**

Ähnlichkeitsbasiertes Retrieval mit myCBR

DJ
Zl|vvhqvedvlhuwh
V|vwhph

ÄHNLICHKEITSBASIERTES
RETRIEVAL

Reset	convertible	coupe	fastback	roadster	sedan	station_wagon
convertible	1.0	0.0	0.2	0.7	0.0	0.0
coupe	0.0	1.0	0.2	0.0	0.7	0.5
fastback	0.2	0.2	1.0	0.5	0.0	0.0
roadster	0.7	0.0	0.5	1.0	0.0	0.0
sedan	0.0	0.7	0.0	0.0	1.0	0.8
station_w...	0.0	0.5	0.0	0.0	0.8	1.0

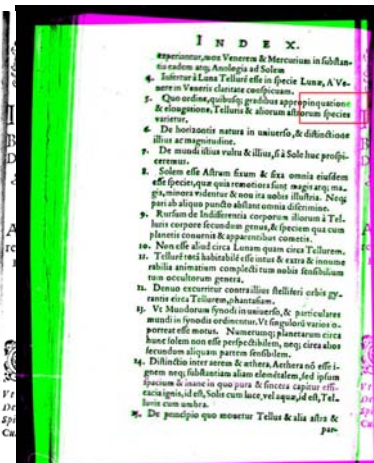
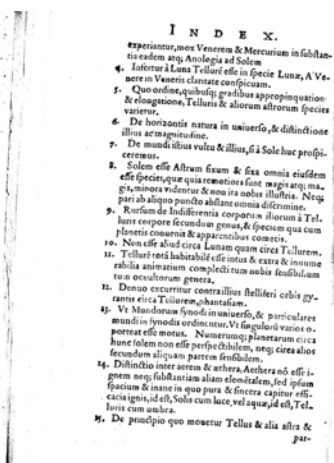
The screenshot shows the myCBR software interface. At the top, there is a similarity matrix for car body styles. Below it, the 'DETAILS AND QUERY' section shows a query for a 'Car' with various attributes like 'Gas', 'Power', 'Model', etc. The 'QUERY RESULTS' table on the right lists retrieved cases with their similarity scores. The interface includes buttons for 'Reset Query', 'Query from Case', and 'Start Retrieval'.

IUPR

Bildbasierte Erkennung textueller Unterschiede in Dokumenten

DJ
Zl|vvhqvedvlhuwh
V|vwhph

- ▶ Methode:
- ▶ Bestimmung der optimalen Transformationsparameter



IUPR

opinqatione
orum species