

# Zweitveröffentlichung



Eckstein, Raiko; Henrich, Andreas; Weber, Nadine

## Das LFRP-Framework : ein Ansatz zur interaktiven Suche für unternehmensspezifische Informationen

Datum der Zweitveröffentlichung: 07.03.2025

Akzeptiertes Manuskript (Postprint), Zeitschriftenartikel

Persistenter Identifikator: urn:nbn:de:bvb:473-irb-1069445

### Erstveröffentlichung

Eckstein, Raiko; Henrich, Andreas; Weber, Nadine (2009): Das LFRP-Framework : ein Ansatz zur interaktiven Suche für unternehmensspezifische Informationen, in: Datenbank-Spektrum : Zeitschrift für Datenbanktechnologie und Information Retrieval ; Organ der Fachgruppe Datenbanken der Gesellschaft für Informatik e.V., Berlin ; Heidelberg: Springer, Jg. 9, Nr. 30, S. 20–29.

### Verlagshinweis

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections.

### Rechtehinweis

Dieses Werk ist durch das Urheberrecht und/oder die Angabe einer Lizenz geschützt. Es steht Ihnen frei, dieses Werk auf jede Art und Weise zu nutzen, die durch die für Sie geltende Gesetzgebung zum Urheberrecht und/oder durch die Lizenz erlaubt ist. Für andere Verwendungszwecke müssen Sie die Erlaubnis der Rechteinhaberinnen und Rechteinhaber einholen.

Für dieses Dokument gilt das deutsche Urheberrecht.

# Das LFRP-Framework – ein Ansatz zur interaktiven Suche für unternehmensspezifische Informationen

Der Bedarf an unterschiedlichen »Informationsträgern«, nachfolgend als Artefakte bezeichnet, sowie eine unklare Vorstellung über die benötigte Information beim Nutzer führen zu komplexen Suchsituationen, wie sie heutzutage in vielen Unternehmen vorzufinden sind. So werden z.B. bei der Entwicklung technischer Produkte unterschiedlichste Artefakte wie Dokumente, Produkte oder Materialien benötigt, wobei Informationsbedürfnisse oft nur vage und ungenau beschrieben werden können. Zusätzlich erschweren die Heterogenität der vorhandenen Ablagesysteme und das Fehlen einer übergreifenden Suchfunktionalität das Auffinden relevanter Information. Aus diesem Grund wird mit dem LFRP-Framework ein Forschungsansatz vorgestellt, der durch die Bereitstellung eines interaktiven Retrieval-Modells zur Beherrschung derartiger komplexer Suchsituationen beiträgt. LFRP steht hierbei für die vier Basiselemente, die zu einem übergreifenden Framework integriert werden: Durch die Verwendung mehrerer Artefaktenebenen (Layer) wird der Vielfalt nachgefragter Artefakte begegnet. Die Suche nach den Artefakten selbst erfolgt gemäß dem Paradigma der facettierten Suche. Allerdings erlaubt eine reine Filterung der Ergebnismenge keine Aussage hinsichtlich der Relevanz der Ergebnisse. Folglich erweitern wir das bekannte Konzept um Möglichkeiten zur Erstellung eines Rankings auf Basis sowohl von Facettenwerten als auch Query-by-Example-(QbE-)Ansätzen. Zusätzlich schlagen wir eine visuelle Form der Anfrageformulierung mittels paralleler Koordinaten vor, die Einblicke in die Charakteristika und Abhängigkeiten der gefundenen Ergebnisse bietet.

## 1 Motivation und Problemstellung

Die Ressource Wissen ist heutzutage in allen Branchen zu einem wichtigen Wettbewerbsfaktor geworden, deren Nutzung wesentlich dazu beiträgt, dass Unternehmen auf dem Markt bestehen und erfolgreich agieren können. Allerdings ist der Arbeitsalltag der Mitarbeiter aufgrund des zunehmenden Einsatzes der Computertechnologie durch eine stetig wachsende Menge an digitalen Informationen geprägt, die in unterschiedlichsten Systemen gespeichert und verwaltet werden. Diese Heterogenität der Systemlandschaft erschwert die Suche nach relevanter Information. Ergebnisse einer Studie (vgl. [Braschler et al. 2009]) zeigen, dass durchschnittlich 18% der Arbeitszeit bei Entscheidungsträgern allein mit der Tätigkeit des Suchens verbracht werden, wobei nicht garantiert ist, dass am Ende auch tatsächlich die benötigte Information gefunden wird. Unklare und vage Vorstellungen der Nutzer über ihr Informationsbedürfnis sowie das Fehlen einer übergreifenden Suchfunktion, die sämtliche Informationsquellen mit einbezieht, tragen dazu bei, dass bereits existierendes Wissen ungenutzt bleibt und Tätigkeiten redundant ausgeführt werden [Braschler et al. 2009].

Diese Situation trifft auch auf Unternehmen aus dem Maschinenbau zu. Der Produktentwicklungsprozess (PEP) dieser Unternehmen wurde und wird im Forschungsverbund FORFLOW<sup>1</sup> näher betrachtet. FORFLOW bildet den Hintergrund des in diesem Beitrag beschriebenen LFRP-Frameworks.

Um die Wettbewerbsfähigkeit der Unternehmen zu sichern, sind innovative Produkte in kurzer Zeit bis zur Marktreife bzw. Produktion zu entwickeln. Sowohl durch die zunehmende Einbindung elektrischer, elektronischer und computergesteuerter Komponenten (z.B. ABS-, ESP-, Navigationssysteme in Fahrzeugen) als auch durch zunehmend dynamische Kundenanforderungen (Stichwort: kundenindividuelle Produktion) werden diese jedoch immer komplexer [Schichtel 2002]. Dies spiegelt sich auch im PEP wider, in dem viele verschiedene Artefakte erzeugt und benötigt werden. Neben Produktdaten und Dokumenten sind auch projektspezifische Informationen, Materialdaten und Informationen über Lieferanten oder Experten von Nutzen für einen Entwickler. Allerdings nimmt die Informationssuche speziell in diesem Bereich laut [Grabowski & Geiger 1997] ca. 20 bis 30% der gesamten Entwicklungszeit in Anspruch, was vor allem auf die Vielzahl der eingesetzten Verwaltungssysteme zurückzuführen ist. Neben Product-Data-Management-(PDM-)Systemen werden Enterprise-Resource-Planning-(ERP-)Systeme, Document-Management-Systeme (DMS), Datenbanken (DB) und selbst einfache Dateisysteme zur Archivierung der Artefakte verwendet. Obwohl jedes dieser Systeme Suchfunktionen anbietet, haben Befragungen der im FORFLOW-Projekt beteiligten Industriepartner gezeigt, dass Entwickler häufig nicht die benötigte Information finden. Dies liegt zum einen daran, dass Entwickler oft selbst nicht genau beschreiben können, wonach sie eigentlich suchen. Zum anderen ist häufig aufgrund einer fehlenden übergreifenden Suchfunktionalität nicht bekannt, dass eine Information existiert oder in welcher Quelle sie enthalten ist.

Obwohl in der Literatur bereits verschiedene Möglichkeiten zur Verbesserung der Informationsversorgung von Produktentwicklern erforscht wurden (vgl. Kap. 2), handelt es sich hierbei größtenteils um Ansätze, die nur ein spezielles Informationsbedürfnis adressieren. Diese Einzellösungen sind zwar in manchen Situationen durchaus hilfreich; jedoch sind sie nicht für die Befriedigung komplexer Informationsbedürfnisse geeignet. Befindet sich ein Entwickler z.B. in einer frühen Phase des PEP, in der er ein Lösungskonzept für definierte Produkthanforderungen erstellen muss, so interessieren ihn in dieser Situation eher Lösungskonzepte

1. Bayerischer Forschungsverbund für Prozess- und Workflowunterstützung zur Planung und Steuerung der Abläufe in der Produktentwicklung ([www.forflow.org](http://www.forflow.org)).

aus früheren Projekten, die mit ähnlichen Anforderungen verbunden waren. In späteren Phasen wie z.B. der eigentlichen Produktkonstruktion hingegen, suchen Entwickler vorrangig nach verwendbaren CAD-Modellen oder nach Informationen über Materialeigenschaften.

Demzufolge stellen wir einen Forschungsansatz vor, der eine übergreifende, interaktive Suche nach unternehmensspezifischen – speziell entwicklungspezifischen – Informationen ermöglicht. Dabei werden unterschiedliche Artefakte aus verschiedenen Quellsystemen einbezogen und Nutzer in zweierlei Weise unterstützt. Zum einen wird eine zielorientierte Suche gemäß dem Konzept der *Known-Item Search* [Reitz 2004] benötigt, die in Situationen, in denen der Produktentwickler eine definierte Vorstellung über sein Suchobjekt besitzt, dieses zielgerichtet bereitstellt. In Situationen hingegen, in denen der Entwickler nicht genau weiß, wonach er sucht, ist eine explorative Vorgehensweise erforderlich, bei der der Anfragende durch *Browsen* des Datenbestandes auf Informationen stößt, die für sein Informationsbedürfnis relevant sind [Marchionini 2006].

Erste Ideen zu diesem Retrieval-Modell wurden in [Eckstein & Henrich 2008] skizziert und in [Eckstein & Henrich 2009] mit dem Fokus auf der Anfragekomponente und der Umsetzung der grafischen Benutzeroberfläche vertieft. Der vorliegende Beitrag präsentiert einen Gesamtüberblick über das entwickelte Framework, indem es sowohl die Indexierungs- als auch die Anfrageseite berücksichtigt. Nach einem Überblick über bereits existierende Retrieval-Ansätze stellen wir in Kapitel 3 unser *LFRP-Framework* für komplexe Suchsituationen anhand einer Beispielanfrage im Prototyp unserer grafischen Benutzeroberfläche (GUI) vor. Ausgehend von der Grundidee einer facettierten Suche erläutert Kapitel 4 im Anschluss die für ein derartiges Suchframework notwendige Indexierung, die Artefaktbeschreibungen auf Basis von Artefakttyphierarchien aus diversen Informationsquellen generiert (Kap. 4) und dabei Beziehungen zwischen Artefakten berücksichtigt (Abschnitt 4.2). Im Anschluss beschäftigt sich Kapitel 5 detaillierter mit der Suchseite des Frameworks. Dabei werden die visuelle Anfrageformulierung mittels paralleler Koordinaten ( $\parallel$ -coords), die Integration von Ranking- und QbE-Kriterien (Abschnitt 5.2 und 5.3) sowie die Nutzung von Artefaktbeziehungen zur Verbesserung der Suchfunktionalität mithilfe eines Ebenenkonzeptes (Abschnitt 5.5) erklärt.

## 2 State of the Art

In den letzten beiden Jahrzehnten beschäftigten sich bereits viele Forscher mit der Herausforderung, die Informationsversorgung von Produktentwicklern im Speziellen, aber auch innerhalb von Unternehmen im Allgemeinen zu verbessern. Dabei sind verschiedene Lösungsansätze entstanden, die im Folgenden skizziert werden.

Da das Auffinden ähnlicher, bereits existierender Produkte während der Entwicklung wesentlich dazu beiträgt, redundante Tätigkeiten zu vermeiden und so sowohl Entwicklungszeiten als auch -kosten zu reduzieren, bildete sich ein wichtiger Forschungsschwerpunkt im Bereich der Ähnlichkeitssuche. Aufgrund der zunehmenden Verfügbarkeit von 3D-CAD-Modellen konzentrierte man sich hierbei hauptsächlich auf die Entwicklung von Retrieval-Ansätzen, die einen Vergleich von Produkten auf Basis ihrer

dreidimensionalen Geometriebeschreibungen ermöglichen. Somit sind in der Literatur unterschiedliche Möglichkeiten zur Repräsentation dieser Geometriemodelle zu finden, die u.a. auf Tiefenpuffer-Informationen [Vranic 2004], Abstandsverteilungen beliebiger Objektoberflächenpunkte [Osada et al. 2002] oder Skelettgraphen [Sundar et al. 2003] beruhen. Des Weiteren wurden derartige Anstrengungen auch für technische Zeichnungen unternommen, zu denen wir in [Weber & Henrich 2007] einen Überblick gegeben haben. Obwohl oder vielleicht gerade weil diese Ansätze nur einzelne, spezielle Informationsbedürfnisse fokussieren, konnten sich einige Konzepte – vor allem im 3D-Bereich – auch kommerziell durchsetzen. Beispiele sind die geometrische Suchmaschine für 3D-Daten *Geolus Search*<sup>2</sup> der Siemens AG oder *CADFind*<sup>3</sup> der Applied Search Technology Ltd.

Des Weiteren hat man versucht, derartige Einzellösungen in umfassendere Ansätze zu integrieren. Ein Beispiel hierfür ist das *Design Navigator System* von [Karnik et al. 2005], das zur Unterstützung der Produktentwicklung im militärischen Bereich prototypisch entwickelt wurde. Neben Möglichkeiten zur Suche nach Produktinformationen (Anfragemöglichkeiten hinsichtlich der Geometriebeschreibung, Produktfunktion und Produktstruktur) wurde hierbei insbesondere die Designhistorie einer Produktentwicklung bei der Suche mit berücksichtigt, um so vor allem neue und unerfahrene Entwickler bei der Arbeit zu unterstützen.

Obwohl alle diese Ansätze interessante Möglichkeiten der Unterstützung bieten, fand bisher kein Konzept breite Anwendung in der Praxis. Stattdessen ist die unternehmensinterne Softwarelandschaft durch heterogene Systeme geprägt, die jeweils nur bestimmte Kernaspekte fokussieren. Während bei PDM/PLM<sup>4</sup> und ERP-Systemen das Produkt mit all seinen Informationen im Vordergrund steht, wird der Fokus bei DMS auf Dokumente und bei CRM<sup>5</sup>- und SRM<sup>6</sup>-Systemen auf Kunden- und Lieferantendaten gelegt. Dabei verfügt jedes einzelne System standardmäßig über Suchfunktionalitäten, die sich allerdings eher am Konzept der in Kapitel 1 genannten *Known-Item Search* orientieren.

Um dieser Heterogenität von Insellösungen und damit der Begrenztheit von verfügbaren Informationen zu begegnen, wurden in den letzten Jahren immer mehr Enterprise-Search-Lösungen entwickelt. Durch die Nutzbarmachung des gesamten in einem Unternehmen verfügbaren Wissens ist es Ziel dieser Systeme, eine unternehmensweite Informationsbereitstellung zu gewährleisten. Derartige Systeme berücksichtigen zwar neben allgemeingültigen Artefakten wie E-Mails oder Office-Dokumenten auch spezielle Artefakttypen (bspw. CAD-Modelle, technische Zeichnungen), allerdings werden diese vorrangig textuell analysiert. Dieser Ansatz reicht jedoch gerade bei speziellen Dokumententypen nicht aus, da wichtige Informationen eben nicht nur in der Textform, sondern vor allem in grafischen Beschreibungen enthalten sind.

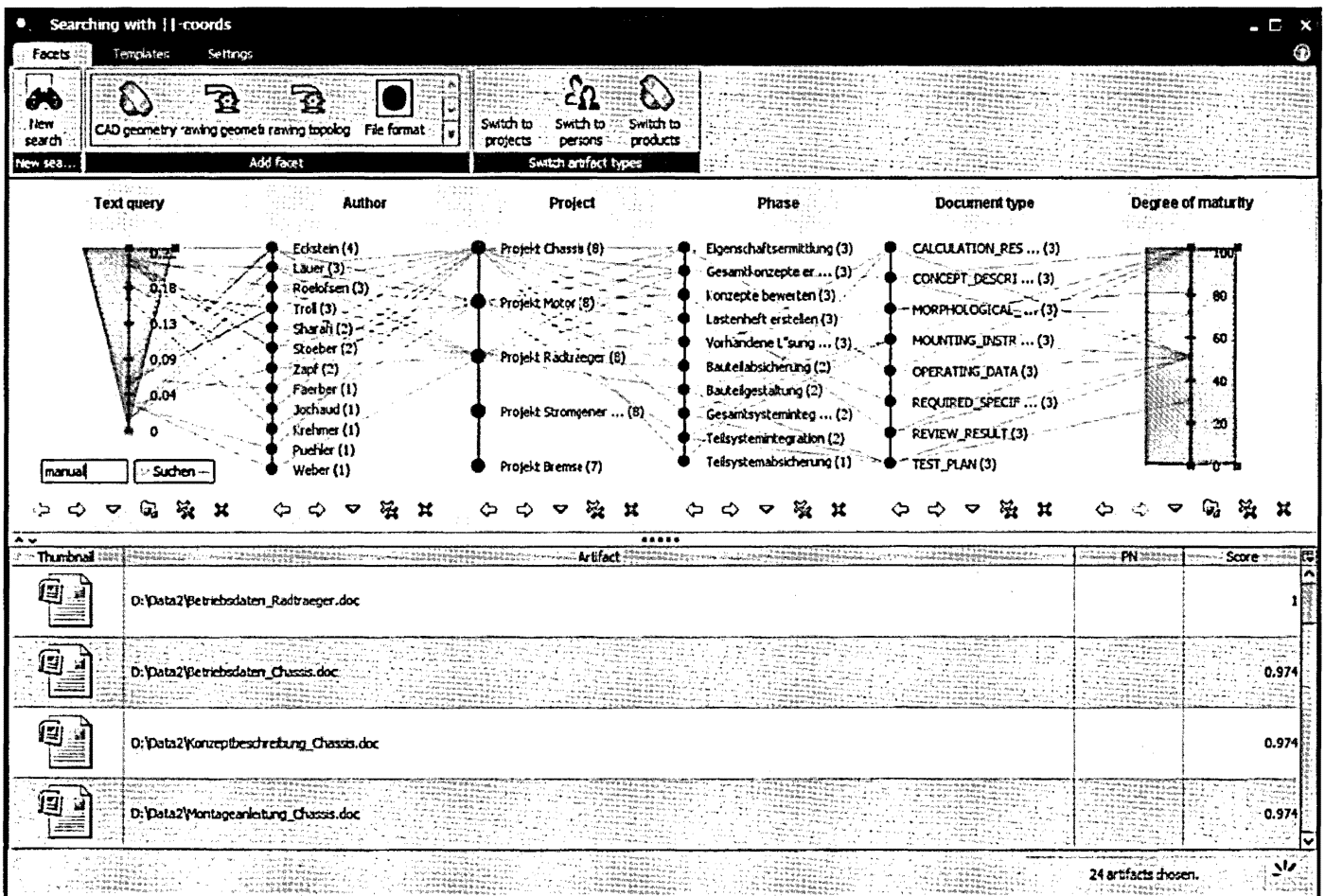
2. [www.plm.automation.siemens.com/de/\\_de/products/\open/geolus/index.shtml](http://www.plm.automation.siemens.com/de/_de/products/\open/geolus/index.shtml)
3. [www.sketchandsearch.com](http://www.sketchandsearch.com)
4. Product Lifecycle Management
5. Customer Relationship Management
6. Supplier Relationship Management

### 3 LFRP-Framework für komplexe Suchsituationen

Für die Realisierung sowohl eines zielgerichteten als auch eines explorativen Suchvorgehens stützen wir unser LFRP-Framework auf das Konzept der *facettierten Suche*, das den Nutzer durch Selektionskriterien, sogenannte Facetten, bei seiner Anfrageformulierung unterstützt [Yee et al. 2003]. Eine Facette beschreibt einen Aspekt eines Artefakts (z.B. Erstellungsdatum bei Dokumenten) und kann während der Indexierung ermittelt werden. Die facettrierte Suche stellt dem Nutzer Anfragevorschauen zur Verfügung. Diese Vorschauen bestehen aus der Anzahl der Artefakte, die im Ergebnis verbleiben würden, wenn der Nutzer die aktuelle Anfrage mit der betrachteten Facettenausprägung weiter einschränkt. So wird wirksam verhindert, dass der Nutzer durch zu eng gestellte Filterungen eine leere Ergebnismenge zurückgeliefert bekommt, da Facettenausprägungen, für die es bei der aktuellen Selektion keine Artefakte gibt, ausgeblendet werden. Die Registerkarte *Facets* im oberen Bereich unseres Prototyps (vgl. Abb. 1) gibt eine Übersicht über die Facetten, indem unter *Add facet* eine dynamisch aktualisierte Liste von wählbaren Facetten angeboten wird. Aus dieser Liste kann der Nutzer die für seine Anfrage relevanten Facetten auswählen und so seine Anfrage schrittweise verfeinern.

Die Darstellung der gewählten Facetten erfolgt dabei im mittleren Bereich der GUI. Im Gegensatz zu herkömmlichen facettierten Suchansätzen verwenden wir hierfür einen visuellen Ansatz mittels  $\|$ -coords. Gemäß [Inselberg 1985] wird jede Facette als eine vertikale Achse dargestellt, die die Facettenausprägungen als Achsenpunkte beinhaltet. Folglich wird eine beliebige Anzahl von Facetten durch parallele Achsen im Zusammenhang abgebildet, was die Veranschaulichung einer komplexen, aus mehreren Unteranfragen bestehenden Anfrage in einer übersichtlichen Form ermöglicht. Abbildung 1 zeigt eine Beispielanfrage nach Dokumenten unter Verwendung der Facetten *Text query*, *Author*, *Project*, *Phase*, *Document type* und *Degree of maturity*. Dabei unterscheiden wir zum einen Attributfacetten, die die Auswahl einer oder mehrerer Facettenausprägungen zulassen und somit als reine Filterkriterien dienen (vgl. Facette *Project*: Hier interessieren nur Dokumente, die in mindestens einem der drei im oberen Bereich der Achse gelisteten Projekte erstellt oder benötigt wurden). Diese Filterung bewirkt allerdings nur eine Einschränkung der Ergebnismenge, sodass für den Nutzer keine Rangfolge der Ergebnisse hinsichtlich ihrer Relevanz zur Anfrage ersichtlich ist. Aus diesem Grund wurden Präferenzfunktionen in das Retrieval-Modell integriert, mit deren Hilfe der Nutzer ein Ranking der Ergebnisse anhand seiner Präferenzen für bestimmte Facettenwerte oder -wertebereiche ermitteln kann. Abbildung 1 zeigt eine solche

Abb. 1: Screenshot der grafischen Benutzeroberfläche mit Darstellung einer Beispielanfrage



Präferenzfunktion für die Facette *Degree of maturity*, bei der jeder Facettenwert im Intervall von 0 bis 100 mit dem gleichen Präferenzwert belegt wird. Obwohl der Nutzer hier prinzipiell jede beliebige Präferenzfunktion definieren kann, stellt das System für gängige Funktionen Vorlagen zur Verfügung (vgl. Abschnitt 5.2).

Zusätzlich ermöglicht unser Framework die Verwendung von Ähnlichkeitskriterien für die Anfrageformulierung. Demzufolge können auch Ansätze zum Vergleich nicht nur der textuellen Ähnlichkeit, sondern auch z.B. der 3D-Geometrieähnlichkeit in einem derartig übergreifenden Framework berücksichtigt und angewendet werden. Wie Abbildung 1 bei der Facette *Text query* zeigt, werden diese Ähnlichkeitsfacetten ebenfalls als Achsen dargestellt. Allerdings veranschaulichen ihre Achsenpunkte die ermittelten Retrieval-Statuswerte, weshalb sie defaultmäßig mit einer Präferenzfunktion überlagert sind. Eine detaillierte Erläuterung hierzu wird in Abschnitt 5.3 gegeben.

Die Ergebnismenge, die durch jede Aktion des Nutzers unmittelbar verändert wird, wird auf zwei Arten visualisiert. Zum einen wird jedes Artefakt in Form eines Linienzuges zwischen den parallelen Koordinaten dargestellt. Dies gibt sowohl Einblicke in die Charakteristika der Ergebnisse als auch in die Abhängigkeiten zwischen den gewählten Facetten. Zum anderen wird die Ergebnismenge im unteren Bereich der GUI zusammen mit einer etwas detaillierteren Beschreibung aufgelistet. Diese Beschreibung besteht aus einem Identifizierungsschlüssel (z.B. dem Dokumentpfad), evtl. einem Retrieval-Statuswert (sofern verfügbar), der die Relevanz eines Ergebnisses angibt, und einem Vorschaubild, falls vorhanden.

#### 4 Generierung spezifischer Artefaktbeschreibungen

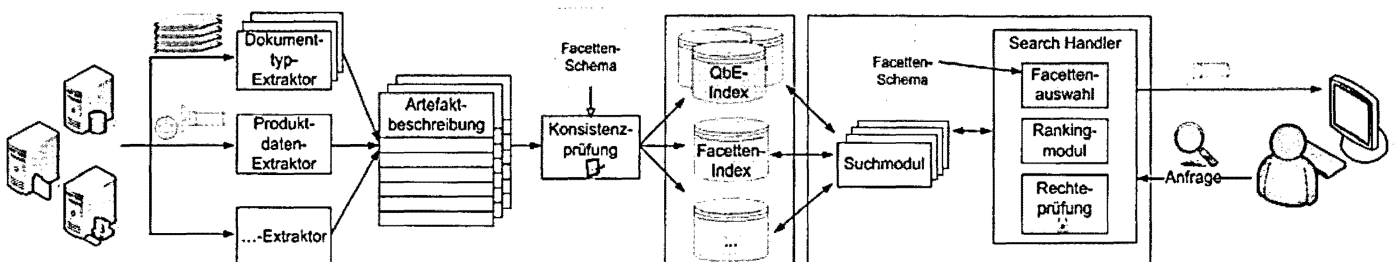
Zur Realisierung des in Kapitel 3 vorgestellten LFRP-Frameworks wird eine Indexierungskomponente benötigt, die eine Beschreibung der Suchobjekte auf Basis von Facetten generiert. Dabei ist zu berücksichtigen, dass in komplexen Suchsituationen diverse Artefakte von Bedeutung sein können. So werden gerade in der Produktentwicklung nicht nur Dokumente, sondern vor allem auch Produktdaten und Materialinformationen nachgefragt. Für jeden dieser Artefakttypen sind andere charakteristische Facetten für die Suche relevant. Während beispielsweise für Dokumente der Dokumenttyp, das Dateiformat oder das Erstellungsdatum wichtige Filterkriterien darstellen, sind Produkte durch eine Sachnummer, ihre Produktgruppe, ihr Gewicht und andere beschreibende Eigenschaften charakterisiert. Folglich sind bei der Indexierung typspezifische Artefaktbeschreibungen zu erstellen. Abbildung 2 zeigt in der linken Hälfte den prinzipiellen Ablauf der Indexierung.

In einem ersten Schritt sind alle Informationen aus den im Unternehmen existierenden Verwaltungssystemen und Anwendungen zu sammeln. Das umfasst zum einen die Indexierung sämtlicher Dokumente aus DMS, DB und Dateiverzeichnissen. Aber auch Produktdaten aus PDM- oder ERP-Systemen, Informationen aus Materialdatenbanken und Projektmanagementsystemen sowie Personendaten aus CRM-/SRM-Systemen und internen Organisationsstrukturen sind bei der Indexierung mit aufzunehmen. Die Indexierung selbst sollte dabei sowohl in regelmäßigen Zeitabständen als auch beim Einstellen neuer oder modifizierter Artefakte erfolgen.

Für jedes der erfassten Artefakte ist in einem zweiten Schritt eine geeignete Artefaktbeschreibung zu generieren. Diese kann einerseits Repräsentationen wie z.B. Featurevektoren oder Histogramme enthalten, die einen Ähnlichkeitsvergleich des Artefakts mit anderen Artefakten des gleichen Typs auf Basis z.B. der Geometrie oder Topologie erlauben (im Weiteren als QbE-Repräsentationen bezeichnet). Zum anderen besteht sie aus Attributfacetten, die für das jeweilige Artefakt charakteristisch sind. Dabei ist zu beachten, dass ein Artefakt durchaus auch Informationen über andere Artefakte beinhalten kann. Insbesondere Dokumente enthalten oft zusätzliche Informationen wie z.B. den oder die Autoren des Dokuments. Somit können prinzipiell sowohl dokument- als auch personenspezifische Facetteninformationen extrahiert werden. Darüber hinaus dienen Dokumente vor allem in der Produktentwicklung dazu, Produkte näher zu beschreiben (z.B. durch CAD-Modelle, technische Zeichnungen, Stücklisten). Dies ermöglicht ein Auslesen von Produktfacetten oder – wenn vorhanden – auch von materialspezifischen Informationen. Da es neben diesen sogenannten Produktmodellen noch andere Dokumentarten wie Projektdokumente und allgemeine Dokumente (z.B. Richtlinien, Normen) gibt, sind abhängig vom Dokumenttyp Beschreibungen für jeweils alle in einem Dokument adressierten Artefakte zu erstellen.

Demzufolge werden für jeden Artefakttyp eine oder mehrere spezifische Extraktorkomponenten benötigt, die geeignete Artefaktbeschreibungen erstellen. Dabei stellt sich die Frage, wann eine Artefaktbeschreibung geeignet ist. Hierfür empfehlen wir im Rahmen der Indexierung die Durchführung einer Konsistenzprüfung, die in einem dritten Schritt die erstellten Artefaktbeschreibungen validiert und verifiziert. Dies erfolgt auf Basis eines Facettenschemas, das aus zwei Definitionsabschnitten besteht. Zum einen sind alle zu berücksichtigenden Facetten zusammen mit ihrem Facettentyp und ihrer Kardinalität<sup>7</sup> (einwertig vs. mehrwertig) festzulegen (siehe Kap. 5). Zum anderen enthält dieses Schema eine Artefakttyphierarchie, die beschreibt, welche Facetten

Abb. 2: Architektur des LFRP-Frameworks



für einen bestimmten Artefakttyp verfügbar sind. Auf Basis dieser beiden Informationen ist es schließlich möglich, zu prüfen, ob eine Artefaktbeschreibung den Vorgaben des Schemas entspricht. Eine genauere Erläuterung hierzu geben wir im folgenden Abschnitt 4.1.

Zusätzlich muss bei dieser Prüfung berücksichtigt werden, dass eine Information aus mehreren Quellen extrahiert werden kann. Beispielsweise sind die aus Produktmodellen extrahierten Produktinformationen (wie z.B. der Produktname) häufig in PDM-Systemen vorhanden. Obwohl die Produktinformation im Dokument normalerweise mit der im PDM-System übereinstimmen sollte, ist dies nicht immer der Fall, weshalb ein Abgleich der Informationen stattfinden muss. Gleiches gilt z.B. auch, wenn Informationen geändert werden und somit zu aktualisieren sind. Für die Beseitigung möglicher Unstimmigkeiten empfehlen wir die Definition von Regeln, die z.B. auf Basis der Extraktionsquelle oder des Extraktionszeitpunktes festlegen, welche Information aus welcher Quelle bei der Indexierung eines Artefakts zu verwenden ist.

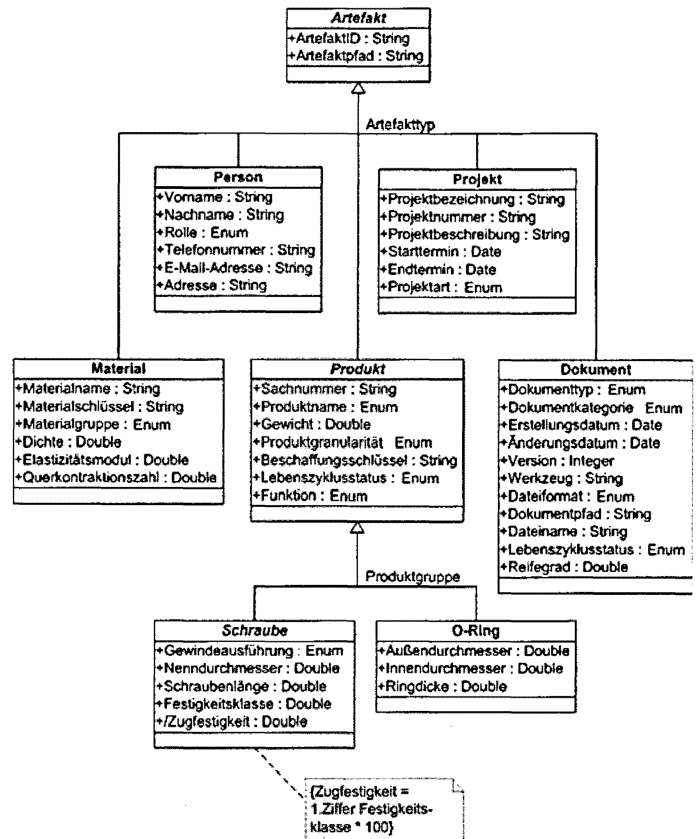
Nach einer erfolgreichen Konsistenzprüfung können die Artefakte schließlich im Index gespeichert werden. Hierzu empfehlen wir die Unterteilung des Index in Subindizes, die die unterschiedlichen Aspekte der Artefaktbeschreibung enthalten. Demzufolge werden die Artefaktfacetten in einem Facettenindex und die QbE-Repräsentationen in entsprechenden QbE-Indizes abgelegt.

#### 4.1 Integration von Artefakttyphierarchien

Abbildung 3 zeigt ein einfaches Beispiel einer Artefakttyphierarchie in UML-Notation, bei der Artefakte zunächst in die Typen *Material*, *Person*, *Produkt*, *Dokument* und *Projekt* unterteilt werden. Jeder einzelne Artefakttyp besitzt dabei spezifische Facetten. So werden z.B. Materialien durch ihren Materialschlüssel, einen Materialnamen, die zugehörige Materialgruppe, ihre Materialdichte und andere Facetten beschrieben. Eine Person hingegen besitzt einen Vor- und Nachnamen, eine Kontaktadresse und eine Rolle (z.B. Mitarbeiter, Lieferant oder Kunde). Derartige in den Blattknoten der Hierarchie befindliche Facetten gelten somit nur für Artefakte des jeweiligen Typs. Facetten des Wurzelknotens hingegen werden an alle Unterklassen vererbt. Demzufolge wird jedes Artefakt durch drei Parameter spezifiziert: eine ArtefaktID zur eindeutigen Identifizierung, einen Artefaktpfad, der angibt, wo das Artefakt verwaltet wird (z.B. Pfadangabe ins PDM für Produkte oder ins DMS für Dokumente), und einen (gemäß UML) »virtuellen« Artefakttyp, der die Spezialisierung der Artefakte auf oberster Ebene definiert.

Darüber hinaus zeigt Abbildung 3, dass auch Artefakte desselben Typs unterschiedliche Facetten besitzen können. Während beispielsweise Produkte wie Schrauben anhand ihrer Gewindeausführung, ihres Nenndurchmessers oder ihrer Schraubenlänge klassifiziert werden, interessieren bei O-Ringen eher Werte für Außen- und Innendurchmesser. Somit hat jede einzelne Produkt-

Abb. 3: Beispiel einer einfachen Artefakttyphierarchie mit den »spezialisierenden« Facetten Artefakttyp und Produktgruppe



gruppe selbst wieder eigene charakteristische Facetten, was in einer eigenen Produkthierarchie resultiert.

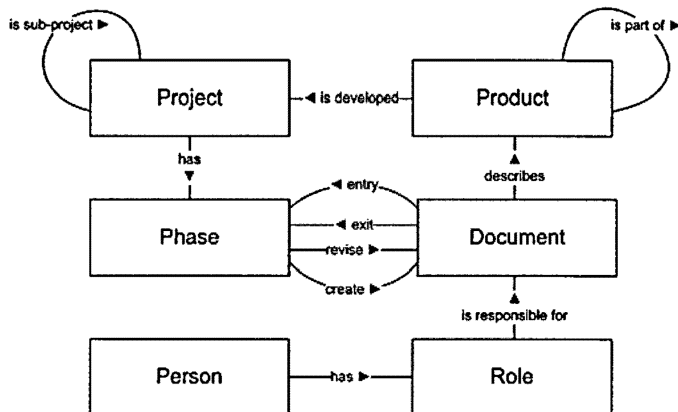
Das bedeutet, dass wir grundsätzlich zwischen Facetten unterscheiden müssen, die eine Spezialisierung von Artefakten definieren, und Facetten, die dies nicht tun. Dabei stellen die »spezialisierenden« Facetten (Artefakttyp, Produktgruppe, ...) sozusagen Muss-Facetten dar, die in einer Artefaktbeschreibung enthalten sein müssen. Nur mit ihrer Hilfe kann die Indexierungskomponente bestimmen, welcher Subtyp in der Hierarchie betrachtet werden muss. Fehlen diese Informationen, sollten sie entweder aus anderen Quellen bezogen oder vom Nutzer erfragt werden, um einen eventuellen Informationsverlust zu vermeiden. Zusätzlich ist eine zweite Differenzierung von Facetten zu beachten. Zum einen gibt es Facetten bzw. Facettenwerte, die direkt aus der Quelle extrahiert werden können, wie z.B. das Erstellungsdatum eines Dokumentes. Zum anderen sind häufig aber auch abgeleitete Facetten, deren Wert von anderen Facetten und ihren Ausprägungen abhängt, als Selektionskriterien notwendig. So lässt sich beispielsweise die Zugfestigkeit einer Schraube aus ihrer Festigkeitsklasse ermitteln (vgl. Einschränkung in Abb. 3). Indem Berechnungsvorschriften bei der jeweiligen Facettendefinition im Facettenschema mit angegeben werden, können derartige Abhängigkeiten berücksichtigt werden.

7. Kardinalität meint die Vorkommenshäufigkeit einer Facette für ein Artefakt. Ein Produkt kann z.B. durch mehrere Funktionen, aber nur durch eine Sachnummer beschrieben werden.

## 4.2 Beziehungen zwischen Artefakten

Wie in Kapitel 4 bereits angesprochen, kann ein Artefakt nicht nur Informationen über sich selbst, sondern auch über andere Artefakte beinhalten. Abbildung 4 zeigt beispielhaft verschiedene Artefakttypen und deren Beziehungen auf.

Abb. 4: Beziehungen zwischen Artefaktebenen (vereinfachtes Schema)



In Domänen wie der Produktentwicklung können derartige Beziehungsnetzwerke sehr komplex sein. Hier werden Projekte initiiert, um neue Produkte zu entwickeln oder um bestehende Produkte anzupassen. Bei einem Produkt handelt es sich entweder um ein Einzelteil oder um eine Baugruppe, die selbst wieder aus mehreren Produkten besteht. Die Durchführung der Projekte erfolgt durch Personen, die in den einzelnen Prozessphasen unterschiedlichste Dokumente erstellen und ändern. Unter anderem gehören hierzu Produktmodelle, die das zu entwickelnde Produkt näher beschreiben. Des Weiteren wird zur späteren Herstellung des Produktes ein bestimmtes Material festgelegt, dessen Eigenschaften den Anforderungen des Produktes genügen und das eventuell von diversen Lieferanten geliefert wird.

All diese Beziehungen tragen wertvolle Informationen, die bei der Befriedigung komplexer Informationsbedürfnisse zu berücksichtigen sind. So sollten beispielsweise Anfragen wie »Finde alle zu einem Produkt verfügbaren Dokumente« oder »Finde alle Projekte, in denen für das entwickelte Produkt Material Z von Lieferant XY verwendet wurde« bearbeitet werden können. Hierzu ist es erforderlich, die Artefaktbeschreibung um Beziehungen zu erweitern, die in einem separaten Index verwaltet werden. Die Definition möglicher Beziehungen zwischen Artefakten erfolgt dabei ebenfalls innerhalb der Artefakttyphierarchie, sodass sie u.a. bei der Erstellung mehrerer Artefaktbeschreibungen aus einer Quelle gesetzt werden können. Die Suchseite kann anschließend diese Beziehungen im Rahmen eines Ebenenkonzepts ausnutzen, um so die Suchfunktionalitäten für den Nutzer zu erweitern und damit die Retrieval-Qualität zu verbessern (vgl. Abschnitt 5.5).

## 5 Anfrageformulierung mit II-coords

Die rechte Hälfte von Abbildung 2 stellt den grundsätzlichen Aufbau des LFRP-Anfrageframeworks dar und illustriert den Ablauf einer Suchanfrage. Hierbei ist insbesondere hervorzuheben, dass

die Anfragestellung interaktiv erfolgt, d.h., der Nutzer muss nicht die gesamte Anfrage formulieren, bevor er diese an das System stellt. Der Nutzer hat die Möglichkeit, Unteranfragen in Form von weiteren Facettenselektionen zur aktuellen Anfrage hinzuzufügen bzw. vorhandene Unteranfragen zu modifizieren. Stellt sich heraus, dass eine durch eine Unteranfrage vorgenommene Filterung nicht zielführend ist, kann diese Filterung modifiziert oder entfernt werden. Die Darstellung wird nach jeder Änderung der Suchanfrage angepasst, d.h., die aktuell gültigen Facettenausprägungen und die Linienzüge werden neu berechnet sowie die Suchergebnisliste angepasst, um nur die aktuell gewählten Artefakte zu präsentieren.

Bei jeder Änderung der Anfrage wird diese an das Suchframework übergeben. Das zentrale Modul ist der *Search Handler*, der die Steuerung und das Delegieren der einzelnen Aufgaben übernimmt. Die einzelnen Unteranfragen werden an die jeweils zuständigen *Suchmodule* weitergeleitet. Bei reinen Attributfacetten ist dies die Facettierungskomponente, die die Anfragevorschaueinblendungen für die einzelnen Facetten bestimmt. Bei Ähnlichkeitsfacetten kommen spezialisierte Module zum Einsatz, die die Ähnlichkeitsbestimmung übernehmen. Dies kann beispielsweise ein Modul für die 3D-Geometrieähnlichkeit oder eine Volltextsuche für textuelle Inhalte sein. Diese Module liefern jeweils ein Ranking zurück, das für die Achsendarstellung in den parallelen Koordinaten notwendig ist. Zusätzlich bestimmt die Facettierungskomponente noch die Linienzüge zur Repräsentation der einzelnen Artefakte in den II-coords. Das *Rankingmodul* erstellt basierend auf den Einzelrankings der Unteranfragen sowie deren Präferenzfunktionen ein finales Ranking. Das Modul *Facettenauswahl* bestimmt auf Basis der aktuellen Anfrage und der Artefakttyphierarchien (inkl. der Facettenbeschreibungen) die Liste der aktuell wählbaren Facetten. Das Modul *Rechteprüfung* stellt sicher, dass der Nutzer nur Artefakte im Ranking zurückgeliefert bekommt, für die er mindestens Leserechte besitzt.

## 5.1 Beschreibung des Anfragemodells

Aufgrund der unterschiedlichen Visualisierungsmöglichkeiten und der verschiedenen Ausprägungen der Anfragearten ist es zunächst notwendig, auf die Charakteristika von Facetten genauer einzugehen. Grundsätzlich werden Attributfacetten und Ähnlichkeitsfacetten unterschieden. Attributfacetten stellen einen Aspekt eines Artefakts dar und können während der Indexierung bestimmt werden. Ähnlichkeitsfacetten hingegen werden zum Anfragezeitpunkt dynamisch ermittelt. Der Nutzer muss in diesem Fall eine Ähnlichkeitsanfrage starten (z.B. eine Schlagwortanfrage oder eine QbE-Anfrage mit einem Beispieldokument). Dieses Vorgehen wird in Abschnitt 5.3 detailliert.

Zusätzlich zu dieser Unterscheidung wird jeder Facette noch ein Facettentyp zugeordnet, der das Skalenniveau des Attributs beschreibt. Der LFRP-Ansatz unterscheidet *nominale*, *ordinale* und *metrische* Merkmale. Die Ausprägungen von nominalen und ordinalen Facetten sind unterscheidbar, wobei für letztere zusätzlich noch eine Rangfolge definiert ist. Wenn zusätzlich zur Rangfolge die Abstände zwischen den Facettenausprägungen bestimmbar sind, spricht man von metrischen Facetten.

Sowohl für die Indexierung als auch die Anfragestellung ist die Kardinalität von Facetten zu berücksichtigen. Facetten können *einwertig* oder *mehrwertig* sein, was sich auf die Kombinationsmöglichkeiten von Facettenwerten in der Anfrage auswirkt.

Eine Suchanfrage besteht aus einzelnen Unteranfragen für jede Facette, die durch eine Achse in den  $\|\text{-coords}$  dargestellt wird. Die verschiedenen Unteranfragen  $Q_i$  werden per Konjunktion (logisches UND) verknüpft. Die Verknüpfung der einzelnen Achsen mittels Disjunktion (logisches ODER) wird absichtlich nicht unterstützt, da diese erhöhte Komplexität der Anfrage das Verständnis für den Nutzer erschweren würde.

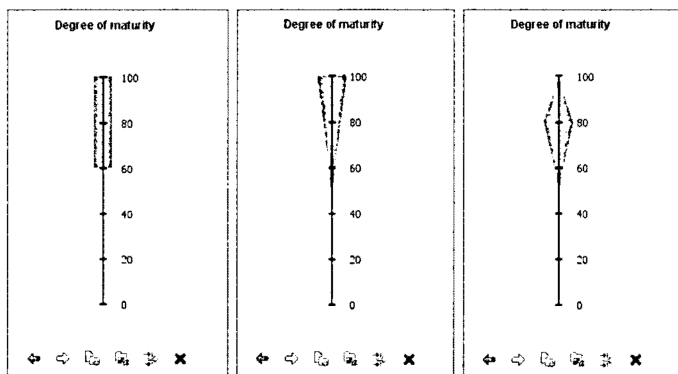
Die Anfragemöglichkeiten für die einzelnen Unteranfragen  $Q_i$  sind abhängig vom Facettentyp und von der Kardinalität einer Facette.

Bei einwertigen Facetten wird bei der Selektion von mehreren Facettenausprägungen automatisch die ODER-Verknüpfung verwendet, da die UND-Verknüpfung zu einer leeren Ergebnismenge führen würde. Eine einwertige Beispielfacette ist der *Dokumententyp*, der für jedes Dokument eindeutig ist. Dokumente, die gleichzeitig mehreren Dokumententypen entsprechen, existieren nicht.

Bei mehrwertigen Facetten sind sowohl die Konjunktion als auch die Disjunktion anwendbar. Betrachtet man die mehrwertige Dokumentfacette *Autor*; äußert sich die unterschiedliche Behandlung von Mehrfachselektionen. Bei Anwendung der Konjunktion werden nur die Dokumente zurückgeliefert, die von allen selektierten Autoren erstellt wurden. Die Disjunktion dagegen schwächt dieses Kriterium ab und liefert Dokumente, die von mindestens einem der selektierten Autoren erstellt wurden.

Die Selektion einzelner Werte bei metrischen Facetten ist möglich, allerdings wird die Filterung auf Intervallabschnitten der häufigere Anwendungsfall sein. Das Anfragemodell des LFRP-Anfrageframeworks unterstützt den Nutzer durch die Möglichkeit, mehrere Intervalle auf einer Facette anzugeben. Ein sinnvoller Einsatzzweck für diese Funktionalität ist beispielsweise der Vergleich von bestimmten Produkten mit verschiedenen Preisgruppen. Ein Beispiel für eine Intervallselektion findet sich in Abbildung 5 in der linken Achse. In diesem Beispiel werden nur die Dokumente zurückgeliefert, deren Reifegrad in einem Bereich zwischen 60% und 100% liegt.

Abb. 5: Beispiele für Präferenzfunktionen



## 5.2 Möglichkeiten zur Erstellung und Beeinflussung eines Rankings

Die klassische facettenierte Suche ist eine Repräsentation des booleschen Retrieval-Modells, also mengenbasiert. Das heißt, das Ergebnis besteht aus einer Menge von Artefakten, die keinerlei Rangfolge beinhalten. Da die Ergebnismengen im Unternehmensumfeld sehr groß werden können, ist es notwendig, dem Nutzer die Möglichkeit zur Verfügung zu stellen, Prioritäten für Facetten anzugeben, nach denen ein Ranking erfolgen soll. Der Nutzer kann über sogenannte *Präferenzfunktionen* festlegen, welche Facettenausprägungen bzw. Wertebereiche bei metrischen Facetten im Ranking höher gewichtet werden sollen. Mithilfe der nachfolgenden Formel kann eine Punktzahl  $score(a_j)$  für jedes Artefakt  $a_j$  berechnet werden, nach der die Sortierung im Ranking vorgenommen wird.

$$score(a_j) = \sum_{i=1}^n \alpha_i \cdot f_i(x_{i,j}) \quad (1)$$

mit  $\alpha_1 + \alpha_2 + \dots + \alpha_n = 1$

Hierbei beschreibt  $\alpha_i$  die Gewichtung der Facette  $i$ , die der Nutzer über die dritte Schaltfläche von rechts in der Symbolleiste unter jeder Rankingfacettenachse einstellen kann. Abbildung 1 zeigt dies für die beiden äußersten Achsen. Der Nutzer kann damit eine niedrigere oder höhere Gewichtung für einzelne Rankingkriterien festlegen. Standardmäßig beträgt die Gewichtung für jede Facette  $1/n$ , wobei  $n$  der Anzahl der gewählten Facetten entspricht. Die Funktion  $f_i(x_{i,j}) \in [0,1]$  beschreibt die Nutzerpräferenzen für den Wert  $x_{i,j}$  der Facette  $i$  für das Artefakt  $a_j$  und kann vom Nutzer grafisch definiert werden.

Der Nutzer kann die Gestaltung der Funktion  $f_i$  frei nach seinen Präferenzen vornehmen, d.h., das System gibt keine Restriktionen vor. Zum besseren und einfacheren Verständnis stellt das System jedoch ein paar einfache »Funktionsvorlagen« zur Verfügung, die häufiger benötigt werden, wie beispielsweise die Dreiecks- und Vierecksfunktionen aus Abbildung 5. Diese Vorlagen können nach dem Hinzufügen durch direkte Manipulation weiter verfeinert und verändert werden. Jede Änderung der Funktion äußert sich in einer Neuberechnung des Rankings und der Facettenwerte.

## 5.3 Einbindung von QbE-Bedingungen

Zusätzlich zur Filterung von Artefakten benötigt der Produktentwickler die Möglichkeit, mithilfe von Beispielartefakten nach ähnlichen Artefakten zu suchen. Das LFRP-Framework ermöglicht diese QbE-Suche durch die Bereitstellung von Ähnlichkeitsfacetten.

Oftmals ist es für den Nutzer zielführend, nach Schlagwörtern zu suchen. Bei diesen Volltextsuchen kommt häufig das Vektorraummodell zum Einsatz [Salton et al. 1975]. Bei der Auswahl einer Ähnlichkeitsfacette wird initial eine Achse ohne Facettenbeschriftung als Platzhalter in die  $\|\text{-coords}$  eingefügt. Der Nutzer muss im zweiten Schritt die Beispielanfrage füllen, d.h. bei einer Textanfrage Schlagwörter übergeben. Dies geschieht über einen Eingabebereich unter der Achse (vgl. Achse *Text query* in Abb. 1).

Im Bereich der Produktentwicklung tritt häufig die Notwendigkeit auf, dass bereits vorhandene Komponenten wiederverwendet werden sollen, um so eine erneute kostspielige Entwicklung zu vermeiden. Um dies zu unterstützen, bieten wir zusätzlich zu den textuellen Ähnlichkeitsfacetten QbE-Facetten an, die es dem Nutzer ermöglichen, einen Ähnlichkeitsvergleich von Produkten auf Basis der im Beispieldokument enthaltenen Geometrie oder auch der Topologie durchzuführen. Bei der Indexierung der Artefakte werden diese Informationen aus den zugehörigen Produktmodellen gewonnen (CAD → 3D, techn. Zeichnung → 2D). Hinsichtlich der Artefakttypen werden diese Informationen der Produktebene zugeordnet, um die produktorientierte Denkweise des Entwicklers zu unterstützen. Zur Repräsentation dieser Facetten wird auf bestehende Ansätze der Forschung zurückgegriffen (vgl. Kap. 2).

Für den Nutzer arbeitet die Verwendung der Ähnlichkeitsfacetten analog zu den Attributfacetten. Dies ist von Vorteil, da der Nutzer die grundsätzliche Funktionsweise der Anfragesteuerung auf QbE-Anfragen übertragen kann und nur die Metapher der Facettenachse verstehen muss.

#### 5.4 Dynamische Bereitstellung von Facetten

Im Enterprise-Search-Umfeld findet sich eine Vielzahl von (semi-)strukturierten Artefakten, die mithilfe einer übergreifenden Suchfunktionalität zugänglich gemacht werden sollten. Da die Anzahl der beschreibenden Facetten aller Artefakte sehr groß werden kann, ist es notwendig, den Nutzer bei einer Informationssuche so zu unterstützen, dass ihm nur die jeweils aktuell sinnvollen und gültigen Facetten für eine weitere Verfeinerung der Suchanfrage angeboten werden. Der Inhalt der Liste der verfügbaren Facetten wird durch Auswertung der aktuellen Suchergebnisse und der Artefakttyphierarchien nach jeder Änderung der Suchanfrage aktualisiert.

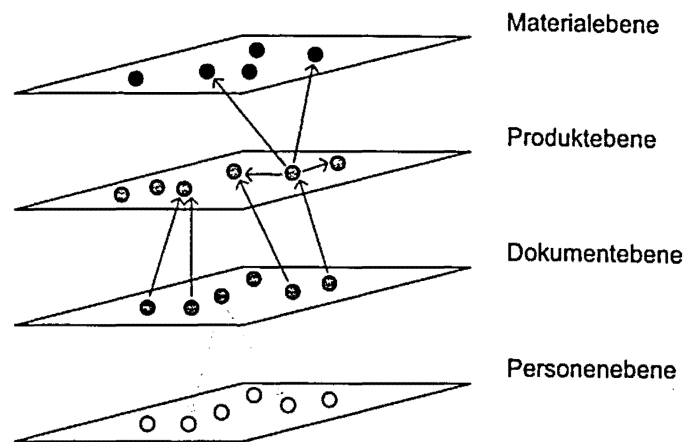
Bei einer Anfrage auf der Produktebene, die der Nutzer über die Facette Produktgruppe »Schrauben« einschränkt, werden zusätzlich zu den allgemeinen produktspezifischen Facetten noch Facetten, wie beispielsweise die Gewindeausführung, die Schraubenslänge und der Nenndurchmesser, angeboten (vgl. Abb. 3).

Die Bestimmung der aktuell gültigen Facetten wird vom Modul *Facettenauswahl* (siehe Abb. 2) übernommen. Dies umfasst auch die Bestimmung der aktuell gültigen Verbindungen zu anderen Artefaktebenen.

#### 5.5 Wechsel des Artefakttyps

Das LFRP-Suchframework unterstützt die Suche auf verschiedenen miteinander verbundenen Artefaktebenen. Die dazu notwendigen Beziehungen wurden in Abbildung 4 visualisiert. Eine Ebene wird durch einen einzelnen Artefakttyp definiert und enthält alle indixierten Artefakte dieses Typs. Die Beziehungen zwischen den verschiedenen Artefakttypen werden für den Ebenenübergang verwendet. Abbildung 6 zeigt beispielhaft Artefakte auf vier Ebenen. Auf der Produktebene sind Beziehungen zwischen Produkten sichtbar (bspw. *is-part-of*-Beziehungen zwischen einzelnen Bauteilen und dem Gesamtprodukt) sowie Beziehungen zwischen Ebenen (Produkt *is-made-of*-Material).

Abb. 6: Beziehungen zwischen und innerhalb von Artefaktebenen



Suchanfragen im LFRP-Suchframework können prinzipiell in jeder unterstützten Artefaktebene beginnen. Der Nutzer kann die Verlinkungen zwischen den Artefakten in zwei verschiedenen Anwendungsfällen nutzen.

Der erste Anwendungsfall besteht in der Suche nach Artefakten einer Ebene, indem der Nutzer seine Anfrage gemäß Abschnitt 5.1 formuliert. Um die Suchergebnismenge weiter zu verringern, kann es hilfreich sein, Filterungen mithilfe von Facetten, die von direkt verbundenen Ebenen stammen, vorzunehmen. Bei einer Suche auf der *Dokumentebene* kann es sinnvoll sein, die Ergebnisdokumente über die Produktfacette *Produktname* weiter einzuschränken. Gemäß Abbildung 4 können Dokumente Produkte beschreiben, d.h., es existiert eine Verbindung zwischen diesen beiden Artefakttypen, die es dem LFRP-Framework ermöglicht, Facetten der direkt verbundenen Ebenen auf Basis der aktuellen Ergebnismenge zu bestimmen und diese zur Filterung anzubieten. Das Anfrageframework generiert automatisch die dazu notwendige geschachtelte Anfrage. Dabei werden ausgehend von den aktuellen Artefakten des Suchergebnisses, in diesem Fall Dokumente, die damit verbundenen Produkte bestimmt. Basierend auf dieser Menge wird dann die Facette *Produktname* berechnet. Dem Nutzer wird damit die Möglichkeit gegeben, die Artefaktverlinkung zu nutzen, um weitere Filterkriterien zu definieren. Das Suchergebnis enthält nach wie vor nur Dokumente, die allerdings durch Facetten, die von direkt verbundenen Ebenen stammen, gefiltert werden können.

Der zweite Anwendungsfall beschreibt den Übergang zwischen Artefaktebenen. Analog zum ersten Anwendungsfall wählt der Nutzer Facetten zur Filterung des aktuellen Suchergebnisses aus und nimmt Selektionen darauf vor. Bei bestimmten Informationsbedürfnissen kann es sinnvoll sein, die Artefaktebene basierend auf den aktuellen Suchergebnissen zu wechseln. Ein Beispiel ist die Suche auf der Dokumentebene mithilfe eines Beispieldokuments. Der Nutzer kann die erhaltene Suchergebnismenge mit Dokumentfacetten weiter einschränken. Ist er mit den Ergebnisdokumenten zufrieden, kann der Wunsch aufkommen, alle durch die Dokumente beschriebenen Produkte zu sehen. Das Anfrageframework bestimmt jetzt basierend auf den Dokumenten in der Ergebnismenge die damit verbundenen Produkte. Das heißt,

ausgehend von der aktuellen Ergebnismenge der Ursprungsebene navigiert das LFRP-Framework anhand der spezifizierten Beziehung zur Zielebene und erstellt die neue Ergebnismenge auf Basis der Artefakte in dieser Ebene. Das neue Suchergebnis kann weiter über Facettenselektionen eingeschränkt werden. Zusätzlich hat der Nutzer die Möglichkeit, zur Ursprungsebene zurückzugehen, um die initialen Filterkriterien zu verringern oder zu erweitern. Die Liste der aktuell gültigen Verlinkungen zu anderen Ebenen ist direkt neben der Liste der aktuell angebotenen Facetten zu finden (vgl. Abschnitt »Switch artifact types« in Abb. 1).

## 6 Zusammenfassung und Ausblick

Das im vorliegenden Beitrag vorgestellte LFRP-Framework zeigt am Beispiel der technischen Produktentwicklung eine Suchlösung für unternehmensspezifische Informationen auf. Hierbei handelt es sich um einen interaktiven Ansatz zur Unterstützung komplexer Suchsituationen, wie sie in wissensbasierten Unternehmen diverser Branchen zu finden sind. Unser Ansatz kombiniert die etablierten Techniken der facettierten Suche, des Rankings und von parallelen Koordinaten in ein integriertes und mächtiges Werkzeug, um explorative Suchanfragen zu unterstützen. Erste Nutzerbefragungen der am Forschungsprojekt beteiligten Industriepartner zeigen, dass die grundsätzliche Ausrichtung des Ansatzes Erfolg versprechend ist.

Nichtsdestotrotz existieren weitere offene Punkte, die noch zu untersuchen sind. Wir entwickeln ein Konzept, um verschiedene Nutzergruppen zu unterscheiden. Expertennutzer sollen hierbei auf die vollständige Funktionalität des Ansatzes zurückgreifen können, um Suchvorlagen für verschiedene Situationen im PEP, die durch komplexe Informationsbedürfnisse geprägt sind, zu definieren. Diese Vorlagen werden in eine Prozessmanagementlösung integriert und können von Standardnutzern als Einstiegspunkt für interaktive Suchanfragen genutzt werden.

Ein weiterer Aspekt betrifft die Präsentation der Ergebnisliste. Hierbei sind weitere Visualisierungen vorstellbar, beispielsweise Schlagwortwolken (Tag Clouds, vgl. [Hassan-Montero & Herrero-Solana 2006]) als Repräsentation für interessante Aspekte wie Projektwolken oder Autorwolken bzw. erweiterte Vorschau-funktionalitäten für die Artefakte.

Ferner müssen auch effiziente Implementierungen des Frameworks und die sinnvolle Verteilung der Berechnungen zwischen Frontend und Backend betrachtet werden.

## Danksagung

Diese Arbeit entstand im Rahmen des von der Bayerischen Forschungsförderung geförderten Forschungsverbundes FORFLOW, der von sechs Lehrstühlen an vier bayerischen Universitäten und 21 Industriepartnern getragen wird. Für weitere Informationen verweisen wir auf die Internetseite ([www.forflow.org](http://www.forflow.org)).

## Literatur

- [Braschler et al. 2009] *Braschler, M.; Briner, N.; Fischer, H.; Häni, M.; Mandl, Th.; Schäuble, P.; Steinbach, M.; Stuker, J.*: Enterprise-Search-Systeme im internen Wissensmanagement: Ergebnisse einer Studie zu Perspektiven der Unternehmen. In: *Datenbank-Spektrum*, 9. Jg., Heft 30, 2009, S. 5-12.
- [Eckstein & Henrich 2008] *Eckstein, Raiko; Henrich, Andreas*: Reaching the Boundaries of Context-Aware IR: Accepting Help from the User. In: *Proc. of the 2nd Int. Workshop on Adaptive Information Retrieval (AIR 2008)*, BCS, London, 2008.
- [Eckstein & Henrich 2009] *Eckstein, Raiko; Henrich, Andreas*: Visual Browsing in Product Development Processes. In: *The Design Society, Proc. of the 17th Int. Conf. on Engineering Design*, 2009. (angenommen)
- [Grabowski & Geiger 1997] *Grabowski, H.; Geiger, K. (Hrsg.)*: Neue Wege zur Produktentwicklung. Dr. Josef Raabe Verlags-GmbH, Stuttgart, 1997.
- [Hassan-Montero & Herrero-Solana 2006] *Hassan-Montero, Y.; Herrero-Solana, V.*: Improving Tag-Clouds as Visual Information Retrieval Interfaces. In: *Int. Conf. on Multidisciplinary Information Sciences and Technologies*, 2006.
- [Inselberg 1985] *Inselberg, A.*: The Plane with Parallel Coordinates. *The Visual Computer*, 1(4):69-91, 1985.
- [Karnik et al. 2005] *Karnik, M. V.; Gupta, S. K.; Anand, D. K.; Valenta, F. J.; Wexler, I. A.*: Design Navigator System: A Case Study in Improving Product Development through Improved Information Management. In: *ASME Computers and Information in Engineering Conference*, Long Beach, CA, 2005.
- [Marchionini 2006] *Marchionini, Gary*: Exploratory Search: from Finding to Understanding. *Communications of the ACM*, 49(4):41-46, 2006.
- [Osada et al. 2002] *Osada, R.; Funkhouser, T.; Chazelle, B.; Dobkin, D.*: Shape Distributions. *ACM Transactions on Graphics*, 21:807-832, 2002.
- [Reitz 2004] *Reitz, Joan M.*: *Dictionary for Library and Information Science*. Libraries Unlimited, Westport, Conn., 2004.
- [Salton et al. 1975] *Salton, Gerard; Wong, Anita; Yang, Chung-shu*: A Vector Space Model for Automatic Index. *Comm. of the ACM*, 18(11):613-620, 1975.
- [Schichtel 2002] *Schichtel, Markus*: *Produktdatenmodellierung in der Praxis*. Carl Hanser Verlag, 2002.
- [Sundar et al. 2003] *Sundar, H.; Silver, D.; Gagvani, N.; Dickinson, S.*: Skeleton Based Shape Matching and Retrieval. In: *Proc. of the Int. Conf. on Shape Modeling and Applications*, IEEE Comp. Society, 2003, S. 130-139.
- [Vranic 2004] *Vranic, D. V.*: *3D Model Retrieval*. Dissertation. Universität Leipzig, 2004.
- [Weber & Henrich 2007] *Weber, N.; Henrich, A.*: Retrieval of Technical Drawings in DXF Format – Concepts and Problems. In: *Lernen – Wissen – Adaption, Workshop Proc.*, Universität Halle-Wittenberg, 2007, S. 213-220.
- [Yee et al. 2003] *Yee, K.-P.; Swearingen, K.; Li, K.; Hearst, M.*: Faceted Metadata for Image Search and Browsing. In: *Proc. of SIGCHI Conf. on Human Factors in Computing Systems*, New York, ACM Press, 2003, S. 401-408.



**Raiko Eckstein**

studierte Wirtschaftsinformatik an der Otto-Friedrich-Universität Bamberg. Seit seinem Abschluss 2006 ist er wissenschaftlicher Mitarbeiter am Lehrstuhl für Medieninformatik bei Prof. Dr. Henrich. Als Projektmitarbeiter im Forschungsverbund FORFLOW beschäftigt er sich mit den Forschungsbereichen kontextbasiertes Information Retrieval, Datenvisualisierung und Enterprise Search.



**Andreas Henrich**

studierte Wirtschaftsinformatik an der Technischen Hochschule Darmstadt; promovierte 1990 an der Universität Hagen und erhielt seine Habilitation 1997 an der Universität Siegen. Seit 1998 ist er Professor an der Otto-Friedrich-Universität Bamberg und seit 2004 hat er dort den Lehrstuhl Medieninformatik inne. Seine Forschungsinteressen liegen im Bereich Information Retrieval, Datenvisualisierung und -exploration sowie im eLearning.



**Nadine Weber**

studierte Wirtschaftsinformatik an der Otto-Friedrich-Universität Bamberg und ist seit ihrem Abschluss 2006 als wissenschaftliche Mitarbeiterin am Lehrstuhl für Medieninformatik bei Prof. Dr. Henrich tätig. Im Rahmen des Forschungsverbundes FORFLOW forscht sie im Bereich des Information Retrieval, wobei ihre Schwerpunkte auf den Aspekten Ähnlichkeitssuche, CAE-spezifische Dokumenttypen und Enterprise Search liegen.

Dipl.-Wirtsch.-Inf. Raiko Eckstein  
Prof. Dr. Andreas Henrich  
Dipl.-Wirtsch.-Inf. Nadine Weber  
Otto-Friedrich-Universität Bamberg  
Lehrstuhl für Medieninformatik  
Feldkirchenstr. 21  
96052 Bamberg  
{raiko.eckstein, andreas.henrich, nadine.weber}@uni-bamberg.de  
www.uni-bamberg.de/minf