

**Fachbereich Wirtschaftswissenschaft**

**PatMining – Wege zur Erschließung textueller Patentinformationen  
für das Technologie-Monitoring**

**Dissertation  
zur Erlangung der Doktorwürde  
durch den  
Promotionsausschuss Dr. rer. pol.  
der Universität Bremen**

**vorgelegt von  
Dipl.-Wi.-Ing. Jan M. Gerken**

**Bremen, 20.06.2012**

- 1. Gutachter: Prof. Dr. Hans-Dietrich Haasis**
- 2. Gutachter: Prof. Dr. Ralf Isenmann**

**Datum des Kolloquiums: 10.12.2012**

## Vorwort

*„Wenn ich weiter geblickt habe, so deshalb, weil ich auf den Schultern von Riesen stehe“*

(Isaac Newton, 1643-1727)

Die vorliegende Arbeit ist während meiner Tätigkeit am Institut für Projektmanagement und Innovation der Universität Bremen entstanden. Zentrale Ergebnisse wurden im Rahmen eines Forschungsprojektes mit der Forschungsvereinigung Antriebstechnik erzielt. Mein Dank gilt daher der Forschungsvereinigung Antriebstechnik für die Finanzierung meiner Stelle und den Mitgliedern der Forschungsvereinigung Antriebstechnik für die konstruktive Zusammenarbeit.

Zahlreiche Menschen haben zum Erfolg dieser Arbeit beigetragen, indem ich an ihren Erfahrungen und ihrem Wissen teilhaben durfte, sie mir den Freiraum zur Erstellung dieser Arbeit gaben und sie mich in vielerlei Hinsicht während der Erstellung dieser Arbeit unterstützt haben. Nur so war das Gelingen dieser Arbeit möglich. Mein besonderer Dank gilt folgenden Personen:

Prof. Dr. Martin G. Möhrle, der mir die Möglichkeit zur Erstellung dieser Arbeit gegeben, die Freude am wissenschaftlichen Arbeiten vermittelt und mir den Weg gewiesen hat, zudem in dessen Co-Autorenschaft die in dieser Arbeit berücksichtigten Publikationen entstanden sind.

Dr. Lothar Walter, der mich ebenso intensiv von Beginn an betreut und mir wichtige Anregungen und Denkanstöße gegeben hat. Darüber hinaus war er auch eine große moralische Stütze und fand immer wieder motivierende Worte.

Prof. Dr. Hans-Dietrich Haasis und PD Dr. Ralf Isenmann für die Übernahme der Gutachten.

Dipl.-Ing. (FH) Jens Potthast für die programmiertechnische Umsetzung und die Unterstützung verschiedenster Experimente. Dr. Anja Dreßler und Dr. Isumo Bergmann, die durch ihre Dissertationen die semantische Patentanalyse maßgeblich entwickelt und mir so den Weg zu dieser Arbeit geebnet haben. Sämtlichen derzeitigen und ehemaligen Kollegen des Instituts für Projektmanagement und Innovation für die konstruktiven und interessanten Diskussionen im „Patentclub“ sowie für das angenehme und inspirierende Arbeitsklima.

Meiner Lebensgefährtin Mareike für ihre motivierende Unterstützung, den Freiraum und das von ihr in diesem Zusammenhang entgegengebracht Verständnis, welches zur Erstellung einer solchen Arbeit notwendig ist.

Meinen Eltern Uta und Eckhard Gerken, die mir diesen Weg in vielerlei Hinsicht erst ermöglicht, mich darin bestärkt, unterstützt und meine Ausbildung uneingeschränkt gefördert haben, und auch meiner Schwester Prof. Dr. Birte Nicolai, die mir diesen Weg vorgelebt hat.

## Inhaltsverzeichnis

Vorwort .....	i
Inhaltsverzeichnis .....	iii
Abbildungsverzeichnis .....	v
Tabellenverzeichnis .....	v
Liste der berücksichtigten Publikationen .....	vi
1 Einleitung und Relevanz des PatMining .....	1
2 Technologie-Monitoring als Rahmen für das PatMining .....	5
2.1 Theoretische Motivationen für das Technologie-Monitoring .....	5
2.2 Technologie-Monitoring als grundlegende Form technologieorientierter Analysen .....	6
3 Grundlagen und Rahmenmodell des PatMining .....	9
3.1 Patente als Informationsquelle .....	10
3.2 Textuelle Patentinformationen .....	11
3.3 Text-Mining .....	13
3.4 Rahmenmodell des PatMining .....	15
4 Konkretisierung der Forschungsfragen und Zuordnung der Beiträge .....	16
4.1 Umfang und Frühzeitigkeit von Patenten .....	16
4.2 Methodische Umsetzung des PatMining für das Technologie-Monitoring .....	18
4.3 Zuordnung der berücksichtigten Publikationen zu den Forschungsfragen F1.1, F1.2, F2.1 bis F2.4 .....	20
4.4 Darstellung der berücksichtigten Publikationen .....	22

4.4.1	F1.1 und F1.2: Gerken, Möhrle und Walter 2012 (under review): One Year Ahead! Investigating the Time Lag between Patent Publication and Market Launch: Insights from a Longitudinal Study in the Automotive Industry .....	22
4.4.2	F2.1 und F2.2: Möhrle und Gerken 2012: Measuring Textual Patent Similarity on the Basis of Combined Concepts: Design Decisions and their Consequences .....	27
4.4.3	F2.3: Gerken, Möhrle und Walter, 2010a,b: Patents as an Information Source for Product Forecasting: Insights from a Longitudinal Study in the Automotive Industry & Semantische Patentlandkarten zur Analyse technologischen Wandels: Eine Längsschnittstudie aus der Allradtechnik .....	33
4.4.4	F2.4: Gerken und Möhrle 2012: A New Instrument for Technology Monitoring: Novelty in Patents Measured by Semantic Patent Analysis .....	37
5	Zusammenfassung und Ausblick .....	41
	Literaturverzeichnis .....	47
	Anhang: Erklärung über Eigenanteil .....	54

## **Abbildungsverzeichnis**

Abbildung 1: Drei Ebenen des Technologie-Monitorings (Quelle: Eigene Darstellung).....	9
Abbildung 2: Generelle Unterscheidung von textuellen und nicht-textuellen Patentinformationen (Quelle: Möhrle et al. 2012 mit leichten Änderungen).....	13
Abbildung 3: Rahmenmodell des PatMining für das Technologie-Monitoring (Quelle: Eigene Darstellung basierend auf der Text-Mining Architektur von Feldman, Sanger 2007).....	16

## **Tabellenverzeichnis**

Tabelle 1: Zuordnung der in dieser Arbeit berücksichtigten Publikationen in das Rahmenmodell des PatMining unter Berücksichtigung der drei Ebenen des Technologie-Monitorings .....	20
Tabelle 2: Darstellung meiner Mitwirkung bei der Erstellung der hier berücksichtigten Publikationen .....	54

## Liste der berücksichtigten Publikationen

Diese kumulative Dissertationsschrift beruht auf den nachfolgend aufgeführten Publikationen. Von den insgesamt fünf Publikationen sind vier nach einem doppelblinden Begutachtungsverfahren bereits publiziert oder zur Publikation angenommen. Eine weitere Publikation befindet sich derzeit im Begutachtungsverfahren. Die folgenden Publikationen sind bereits veröffentlicht:

- Gerken, J. M.; Möhrle, M. G. und Walter, L., (2010a): Patents as an Information Source for Product Forecasting: Insights from a Longitudinal Study in the Automotive Industry. *R&D Management Conference 2010 Proceedings*. Manchester, UK.
- Gerken, J. M.; Möhrle, M. G. und Walter, L. (2010b): Semantische Patentlandkarten zur Analyse technologischen Wandels: Eine Längsschnittstudie aus der Allradtechnik, in: Gausemeier, J. (Hrsg.): *6. Symposium für Vorausschau und Technologieplanung*. Paderborn: Heinz-Nixdorf-Institut, S. 325-349.
- Gerken, J. M. und Möhrle, M. G. (2012): A New Instrument for Technology Monitoring: Novelty in Patents Measured by Semantic Patent Analysis. *Scientometrics*, 91(3), 645-670.
- Möhrle, M. G. und Gerken, J. M., (2012): Measuring Textual Patent Similarity on the Basis of Combined Concepts: Design Decisions and their Consequences. *Scientometrics*, doi: 10.1007/s11192-012-0682-0.

Nachfolgende Publikation befindet sich derzeit im Begutachtungsverfahren:

- Gerken, J. M.; Möhrle, M. G. und Walter, L. (2012): One Year Ahead! Investigating the Time Lag between Patent Publication and Market Launch: Insights from a Longitudinal Study in the Automotive Industry. *R&D Management*, under review.

Es folgt die Darstellung des Forschungszusammenhangs der hier berücksichtigten Publikationen. Die einzelnen Publikationen sind in dieser Fassung nicht enthalten. Sie sind jedoch der begutachteten Fassung beigelegt und liegen dem zentralen Prüfungsamt der Universität Bremen vor.

## 1 Einleitung und Relevanz des PatMining

Seit Schumpeters (1927, 1939, 1996) Arbeiten besteht weitgehend Konsens darüber, dass Technologien maßgeblich den Unternehmenserfolg beeinflussen. Gleichzeitig können aber neu aufkommende, substituierende Technologien auch zum Niedergang etablierter Unternehmen führen (z. B. Foster 1986). Historische Beispiele hierfür sind das Unternehmen Keuffel & Esser, welches als führender Hersteller von Rechenschiebern das Aufkommen von Taschenrechnern verpasste, und das Unternehmen Smith Corona, welches als führender Hersteller von elektronischen Schreibmaschinen die zunehmende Bedeutung der Computertechnologie unterschätzte (siehe hierzu Schilling 1998). Auch dies hatte Schumpeter (1996) bereits beobachtet und als Prozess der kreativen Zerstörung thematisiert:

*„the process of Creative Destruction is the essential fact about capitalism ... it is not [price] competition which counts but the competition from ... the new technology... which strikes not at the margins of profits of existing firms but at their foundations and their very lives“* (Schumpeter 1996).

Da Technologien die Wettbewerbs- und Überlebensfähigkeit eines Unternehmens maßgeblich beeinflussen können, stehen Unternehmen vor der Herausforderung, Technologien und deren Veränderungen kontinuierlich zu überwachen. Dafür stellen Patente eine geeignete Informationsquelle dar, da sie im Vergleich zu anderen Informationsquellen verschiedene Vorteile aufweisen. Hervorgehoben werden in diesem Zusammenhang häufig der Umfang, die Aktualität, die Objektivität, die Klassifikation, der Marktbezug, die Verfügbarkeit, die Vollständigkeit, die Standardisierung und die Analysierbarkeit (siehe z. B. Griliches 1990, Grupp 1997, Debackere et al. 2002, Ernst 2003, Sternitzke 2008). Daher ist es nicht verwunderlich, dass Patente

häufig als wertvolle Informationsquelle im Technologie-Monitoring betrachtet werden (z. B. Yoon, Park 2004, Tseng et al. 2007).

Jedoch stellt heute nicht die Verfügbarkeit von Informationen, sondern deren Menge und Komplexität sowie die Selektion von Informationen die zentrale Herausforderung im Umgang mit Informationen dar (Willke 2007). Dies trifft auch auf Patentinformationen zu. Die in Patenten dokumentierten technologischen Informationen nehmen permanent zu. Die Schwierigkeit der Auswertung und Analyse der Patentinformationen liegt (i) in der jährlich großen Anzahl von neuen und (ii) dem Umfang einzelner Patentanmeldungen begründet.

- (i) In den letzten Jahren ist ein deutliches Wachstum an Patentanmeldungen weltweit zu verzeichnen. Während im Jahr 2001 beim US-amerikanischen Patentamt insgesamt ca. 324.000 Patentanmeldungen gezählt wurden, waren es im Jahr 2010 bereits über 479.000 Patentanmeldungen (U.S. Patent and Trademark Office (Hrsg.) 2012). Ein ähnliches Bild ist beim Europäischen Patentamt zu erkennen: Waren es 2001 noch ca. 101.000 Patentanmeldungen (European Patent Office (Hrsg.) 2005), wurden 2010 bereits ca. 143.000 Patentanmeldungen eingereicht (European Patent Office (Hrsg.) 2012).
- (ii) Die durchschnittliche Seitenzahl je Patentanmeldungen beim Europäischen Patentamt stieg zwischen 1988 und 2004 von ca. 14 Seiten auf nahezu 30. Die Anzahl der Ansprüche stieg in dem Zeitraum von ca. 11 Ansprüche auf über 21 (van Zeebroeck et al. 2009).

Aus den hohen Anmeldezahlen und dem zunehmenden Patentumfang resultieren erhebliche Probleme bei der Erschließung des Wissens aus den Patenten. Es kommt bei verschiedenen Aspekten der Patentierung zum „Information Overflow“, der die Unsicherheit (Bergmann et al. 2007) und den Zeitaufwand bei der Analyse (Harmann

2003) auf diesem Gebiet beträchtlich erhöht, so dass der Aufwand den potentiellen Nutzungsmöglichkeiten gegenübersteht.

Als Reaktion auf diese Informationsflut wurden neue, IT-gestützte Methoden der Patentanalyse entwickelt, die die Analyse großer Datenmengen erleichtern (Harmann 2003). So ist es möglich, im Vergleich zur manuellen Analyse eine wesentlich größere Datenmenge zu durchsuchen. Die Ausweitung der Datenmengen kann durch die Vergrößerung der Suchbreite, aber auch durch eine größere Suchtiefe erfolgen. Für das Technologie-Monitoring wird dadurch insbesondere die Erkennung schwacher Signale in großen Datenmengen maßgeblich unterstützt und im Vergleich zu früheren Ansätzen vereinfacht (Bürgel et al. 2000). Allerdings wurden in der Vergangenheit hauptsächlich Methoden des Data-Mining verwendet (Harmann 2003), die auf patentstatistische Auswertungen (z. B. Comanor, Scherer 1969, Haupt et al. 2007) oder der Auswertung bibliografischer Informationen (z. B. Gupta, Pangannaya 2000, Ernst 2003, Shih et al. 2010) basieren.

Im Gegensatz zu Data-Mining-Methoden ist die Verwendung von Text-Mining-Methoden bisher weniger verbreitet. Während Data-Mining der Auswertung strukturierter Datenmaterials (z. B. bibliografischer Patentinformationen) dient, werden mit Text-Mining unstrukturierte Daten (z. B. die Beschreibung der Erfindung oder die Patentansprüche) ausgewertet. So ist eine zentrale Herausforderung des Text-Mining gegenüber dem Data-Mining, unstrukturierte Daten so aufzubereiten, dass diese ausgewertet werden können. Bisherige Ansätze zur Patentanalyse mittels Text-Mining beschränkten sich häufig auf die Auswertung von Schlüsselbegriffen (z. B. Yoon, Park 2004, Yoon, Park 2005, Kim et al. 2008). Ausnahmen stellen hierbei die Arbeiten von Dressler (2006) und Bergmann (2011) dar, die das Verfahren der

semantischen Patentanalyse im Rahmen von Mergers & Akquisition und zur Erkennung von Patentverletzungen eingesetzt haben.

Die vorliegende Arbeit knüpft an die genannten Arbeiten an und setzt sich mit der Auswertung textueller Patentinformationen - also den unstrukturierten Daten der Patente - für das Technologie-Monitoring mittels Text-Mining-Methoden auseinander. Hierfür ist eine methodische Anpassung der bestehenden Verfahren der semantischen Patentanalyse für das Technologie-Monitoring notwendig. Zuvor scheint es nützlich zu klären, inwieweit Patentinformationen den Anforderungen des Technologie-Monitorings gerecht werden und tatsächlich eine geeignete Informationsquelle darstellen. Daher werden folgende Forschungsfragen formuliert:

*F1: Inwieweit sind Patente für das Technologie-Monitoring eine geeignete Informationsquelle?*

*F2: Wie können Text-Mining-Methoden für typische Aufgaben des Technologie-Monitorings angepasst und genutzt werden?*

Zur Beantwortung der Forschungsfragen wird zunächst in **Kapitel 2** auf das Technologie-Monitoring als Anwendungskontext des PatMining im Rahmen dieser Arbeit eingegangen. Daran anschließend werden in **Kapitel 3** die Grundlagen des PatMining, nämlich Patente als Informationsquelle, textuelle Patentinformationen und Text-Mining, erläutert. In diesem Zusammenhang wird dann ein Rahmenmodell für das PatMining dargestellt, das sich an der Grundarchitektur von Text-Mining-Systemen orientiert. Anhand des Rahmenmodells und basierend auf den vorangegangenen Ausführungen zum Technologie-Monitoring sowie den Grundlagen des PatMining werden dann in **Kapitel 4** die Forschungsfragen F1 und F2 konkretisiert. Die in dieser Arbeit zu berücksichtigenden Publikationen werden den Forschungsfragen zu- und in das Rahmenmodell des PatMining eingeordnet, bevor sie im Detail vorgestellt wer-

den. Abschließend wird die Arbeit zusammengefasst und ein Ausblick für Theorie und Praxis gegeben.

## **2 Technologie-Monitoring als Rahmen für das PatMining**

In der vorliegenden Arbeit werden Aufgaben und Prozesse des Technologie-Monitorings als Rahmen für das PatMining genutzt. Zunächst wird daher die theoretische Bedeutung des Technologie-Monitorings vor dem Hintergrund bestehender betriebswirtschaftlicher Theorien erläutert. Danach wird das Technologie-Monitoring als grundlegende Form technologieorientierter Analysen dargestellt.

### **2.1 Theoretische Motivationen für das Technologie-Monitoring**

Bereits einleitend wurde drauf hingewiesen, dass die Bedeutung von Technologien und deren Veränderungen schon bei Schumpeter (1927, 1939, 1996) thematisiert wurden. Ebenso ist auch die Relevanz der Überwachung von Technologien vor dem Hintergrund verschiedener theoretischer Strömungen offensichtlich, so dass (i) Market-based View, (ii) Ressource-based View und (iii) der Ansatz der Dynamic Capabilities theoretische Motivationen für das Technologie-Monitoring liefern.

- (i) Im Market-based View wird davon ausgegangen, dass der Unternehmenserfolg maßgeblich von der Anpassung an die Umweltgegebenheiten abhängt. Die Bedrohung durch Ersatzprodukte ist eine dieser Umweltgegebenheiten (Porter 1998). Insoweit liefert der Market-based View eine unmittelbare Motivation für das Technologie-Monitoring.
- (ii) Als Reaktion auf den Market-based View wurde der Ressource-based View zunächst als Gegenpol etabliert (Wolf 2008). Unternehmenserfolg wird im Resource-based View maßgeblich durch interne Ressourcen determiniert (Barney 1991). Der Ressourcenbegriff umfasst dabei materielle und immaterielle Ver-

mögensgegenstände, so dass auch das technologische Wissen eine erfolgswirksame Ressource darstellt (Wernerfelt 1984). Die Beobachtung der Imitation und der Erosion des technologischen Wissens bilden daher aus dieser Perspektive die Motivation für das Technologie-Monitoring.

- (iii) Besondere Bedeutung erlangt das Technologie-Monitoring vor dem Hintergrund des Ansatzes der Dynamic Capabilities. Im Mittelpunkt dieses Ansatzes steht die Verwertung interner und externer unternehmensspezifischer Kompetenzen zur Sicherstellung des Unternehmenserfolgs in sich verändernden Umfeldern (Teece et al. 1997). Dynamic Capabilities sind daher *„the firm’s ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments“* (Teece et al. 1997). Dem Technologie-Monitoring kommen hierbei zwei wichtige Aufgaben zu: Die Erkennung relevanter technologischer Kompetenzen, die durch das Unternehmen internalisiert werden können, und die Überwachung des sich schnell verändernden technologischen Umfelds.

## **2.2 Technologie-Monitoring als grundlegende Form technologieorientierter Analysen**

In der Vergangenheit hat sich eine große Begriffsvielfalt in Hinblick auf Technologie-Analysen herausgebildet. Neben dem Technologie-Monitoring wurden die *„Technologiefrühaufklärung“* (Zeller 2003), die *„Technologiefrüherkennung“* (z. B. Grawatsch 2005), die *„Technologie-Vorausschau“* (z. B. Geschka et al. 2008), die *„Technology Intelligence“* (z. B. Lichtenthaler 2002), die *„Technology Foresight“* (z. B. Reger 2001), das *„Technology Assessment“* (z. B. Porter, Cunningham 2005) und das *„Technology Forecasting“* (z. B. Watts, Porter 1997, für eine umfassende Übersicht siehe Technology Futures Analysis Methods Working Group 2004) thematisiert. Das Technologie-Monitoring ist daher weiter zu konkretisieren.

Technologie-Analysen haben generell die Aufgabe der „Gewinnung entscheidungsrelevanten Wissens zur Unterstützung des strategischen (Technologie-)Managements“ (Zeller 2003). Sie liefern Erkenntnisse, mit denen die Technologiestrategie überprüft, ggf. revidiert und Ressourcen, insbesondere im Forschungs- und Entwicklungsbereich, (re-)allokiert werden (Bürgel et al. 2000, Teichert, Mittermayer 2002), so dass Chancen für das Unternehmen genutzt und Gefährdungen vom Unternehmen abgewendet werden können (Lichtenthaler 2008). Darüber hinaus können Erkenntnisse von Technologie-Analysen eine wichtige Ideenquelle für neu zu akquirierende Technologien und neue Forschungs- und Entwicklungskooperationen sein (Bürgel et al. 2000).

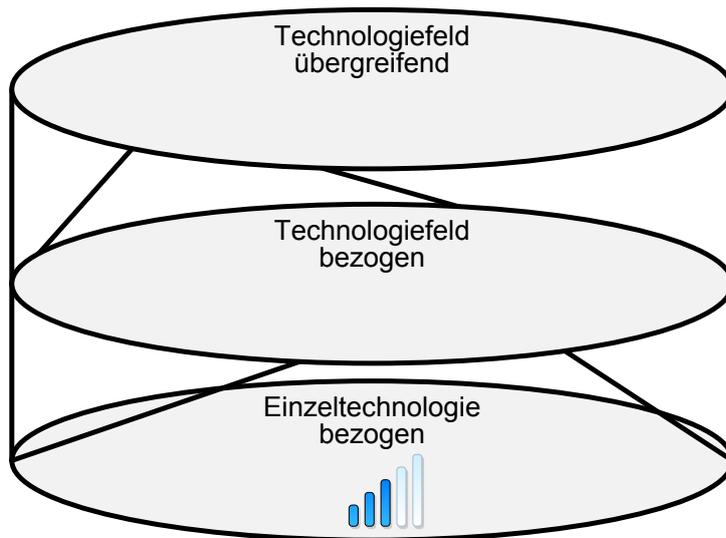
Das Technologie-Monitoring ist hierbei von besonderer Bedeutung, denn es ist die Grundlage für die übrigen Technologie-Analysen, kann aber auch isoliert genutzt werden (Porter et al. 1991, Ashton, Stacey 1995, Watts, Porter 1997, Porter, Cunningham 2005). So kommt dem Technologie-Monitoring beispielsweise bei der Trenderkennung und der Technologie-Vorausschau eine Schlüsselrolle zu (Watts, Porter 1997, Wellensiek et al. 2011). Zentrale Aufgabe des Technologie-Monitoring ist die Informationsbeschaffung und -interpretation (Porter, Cunningham 2005). Informationen werden hierbei zur Erkennung neuer, aber auch zur Überwachung bestehender Technologien beschafft und interpretiert (Porter et al. 1991), so dass der aktuelle Stand der Technik (Reger 2001) und vor diesem Hintergrund technologische Veränderungen erkannt werden können (Ashton, Stacey 1995). Das Technologie-Monitoring erfolgt dabei mit einem (nahezu) uneingeschränkten Zeithorizont, so dass Informationen unter zeitdynamischen Aspekten betrachtet werden (Wellensiek et al. 2011). Dies sollte auf unterschiedlichen Ebenen erfolgen (Porter, Cunningham 2005), die sowohl das Erkennen von Veränderungen in Technologiefeldern, aber auch von Veränderungen zwischen Technologiefeldern ermöglichen (Wellensiek et al. 2011).

Zu beachten ist dabei, dass (i) einzelne Technologien maßgeblich für technologische Veränderungen sind, (ii) neue Technologien häufig auf der Rekombination bestehender Technologien beruhen und dass (iii) technologische Veränderungen durch angrenzende Technologiefelder beeinflusst werden.

- (i) Radikal (Chandy, Tellis 1998, Dahlin, Behrens 2005), disruptiv (Christensen, Overdorf 2000), bahnbrechend (Ahuja, Lampert 2001) und diskontinuierlich (Anderson, Tushman 1990) sind nur einige Adjektive, die in der Vergangenheit verwendet wurden, um Technologien mit besonderer Bedeutung für technologische Veränderungen zu bezeichnen. Typisch für diese Technologien ist im Hinblick auf deren technologischen Charakteristiken eine herausragende Neuheit. Da diese Technologien technologische Veränderungen in einem Technologiefeld maßgeblich mitbestimmen, müssen diese durch das Technologie-Monitoring erkannt werden.
- (ii) Neue Technologien erwachsen aus der Rekombination bestehenden technologischen Wissens. Insbesondere radikale Innovationen beruhen stärker als nicht-radikale Innovationen auf der Rekombination von Wissen (Schoenmakers, Duysters 2010). Daher bedarf es nicht nur einer Überwachung von einzelnen Technologien, sondern auch der Überwachung der Relationen zwischen Technologien und damit auch der Strukturen in einem Technologiefeld.
- (iii) In der jüngeren Vergangenheit hat sich gezeigt, dass technologischen Veränderungen nicht nur innerhalb eines Technologiefelds stattfinden. Diese kann auch durch die Verflechtung und Kombination bestehender Technologien und Technologiefelder bedingt werden (Zeller 2003), so dass technologische Veränderung aus dem Zusammenwachsen von Technologiefeldern resultieren kann (Curran, Leker 2009). Dies wird als Konvergenz bezeichnet. Beispiele hierfür sind die sog. „Smartphones“, die aus dem Zusammenwachsen von Mobilfunk

und anderen Technologien wie MP3, Organizer und Internet entstanden sind. Hierbei entstehen keine grundlegend neuen Technologien. Vielmehr verschmelzen bestehende Technologien. Das Technologie-Monitoring muss entsprechend auch für die Erkennung Technologiefeld übergreifender Veränderungen ausgestaltet werden.

Aufgrund der Bedeutung einzelner Technologien, der Veränderungen in Technologiefeldern sowie der Technologiefeld übergreifenden Veränderungen, müssen durch das Technologie-Monitoring Einzeltechnologie bezogene, Technologiefeld bezogene und Technologiefeld übergreifende Technologieindikatoren bereitgestellt werden (Abbildung 1).



**Abbildung 1: Drei Ebenen des Technologie-Monitorings (Quelle: Eigene Darstellung)**

### **3 Grundlagen und Rahmenmodell des PatMining**

Möhrle et al. (2012) haben für Verfahren, bei denen es um die Erschließung und inhaltliche Analyse textueller Patentinformationen geht, den Begriff PatMining eingeführt. Das PatMining fügt sich in das Technologie-Monitoring ein und kann zur Analyse von Einzeltechnologien, von Technologiefeldern sowie zur Technologiefeld über-

greifenden Analyse genutzt werden. Dabei werden Patente als Informationsquelle (Kapitel 3.1) genutzt und die textuellen Patentinformationen (Kapitel 3.2) mit Text-Mining-Methoden (Kapitel 3.2) ausgewertet. Patentinformationen und Text-Mining stellen die Elemente des Rahmenmodells des PatMining dar (Kapitel 3.4).

### **3.1 Patente als Informationsquelle**

Patente sind mittlerweile eine Standardinformationsquelle für Technologie-Analysen (Harmann 2003). Die herausragende Stellung von Patenten als Informationsquelle ist auf den wechselseitigen Austausch zwischen Staat und Erfinder während des Erteilungsverfahrens zurückzuführen (Behrmann 1998).

Während des Erteilungsverfahrens wird dem Anmelder durch den Staat ein Verbotungsrecht verliehen, das dem Patentinhaber das Recht einräumt, Dritten zu untersagen, die durch das Patent geschützte Erfindung gewerblich zu nutzen, also beispielsweise herzustellen, zu gebrauchen, anzubieten, zu lagern, zu importieren oder zu verkaufen. Dieses Recht ist territorial und zeitlich beschränkt. Der territoriale Schutz erstreckt sich auf die Länder, für die bei den jeweiligen regionalen Patentämtern die Erfindung zum Patent angemeldet wurde. Zeitlich beschränkt sich der Schutz in den meisten Ländern auf maximal 20 Jahre. Im Gegenzug zu dem gewährten Schutz wird die Erfindung während des Patentierungsverfahrens veröffentlicht, so dass diese für die Allgemeinheit einsehbar ist und so den Stand der Technik bereichert (siehe z. B. Granstrand 2000, Gassmann, Bader 2007). In den USA und in Europa wird die Erfindung spätestens 18 Monate nach Patentanmeldung offenbart, unabhängig davon, ob das Patent zu diesem Zeitpunkt bereits erteilt ist oder nicht (Debackere et al. 2002). Die Veröffentlichung von erteilten Patenten erfolgt in sogenannten Patentschriften, die von Patentanmeldungen in sogenannten Offenlegungsschriften. Um die Durchsetzung des Ausschlussrechts zu gewähren, müssen

Schutzwirkung und -bereich klar und umfassend im Patentdokument, dies sind Patent- und Offenlegungsschriften, dargestellt werden (Behrmann 1998). Die in Patentdokumenten offenbarten Erfindungen stellen daher eine umfangreiche technische Informationsquelle dar (Knight 2004) und bieten Dritten die Möglichkeit, sich Kenntnisse über den technologischen Fortschritt anzueignen (Grupp 1997).

### **3.2 Textuelle Patentinformationen**

Patentdokumente folgen einer einheitlichen Struktur, die durch die Patentämter vorgegeben wird. Die Struktur kann zwar von Patentamt zu Patentamt variieren, jedoch ist die grundsätzliche Struktur vergleichbar.

Für das Rechtsgebiet des Deutschen Patent- und Markenamts sind die Inhalte von Patentdokumenten in §32 des Patentgesetzes geregelt (Benkard 2006). Patentschriften umfassen Patentansprüche, eine Beschreibung der Erfindung und Zeichnungen, die grundlegend für die Erteilung des Patents waren, sowie die Druckschriften, die das Patentamt für die Patentfähigkeit begutachtet hat. Sofern eine Zusammenfassung des Patents noch nicht veröffentlicht wurde, wird auch diese in die Patentschrift aufgenommen. Die Offenlegungsschrift weist eine vergleichbare Struktur auf. Sie enthält die Unterlagen zur Anmeldung, die für jedermann zur Einsicht zur Verfügung stehen, sowie eine Zusammenfassung. Die Anmeldung muss wiederum nach §34 des Patentgesetzes den Namen des Anmelders, einen Antrag auf Erteilung des Patents einschließlich einer Bezeichnung der Erfindung, ein oder mehrere Patentansprüche, eine Beschreibung der Erfindung sowie ggf. Zeichnungen enthalten (Benkard 2006). Außerdem umfassen Patentdokumente in der Regel ein Titelblatt, welches die bibliografischen Informationen, wie Angaben über den Anmelder, den Erfinder, das Anmeldejahr und das Erteilungsjahr, enthält.

Diese grundsätzliche Struktur, bestehend aus bibliografischen Angaben, Bezeichnung oder Titel, Ansprüchen, Zusammenfassung, Zeichnung und einer Beschreibung der Erfindung, die beispielsweise Informationen über den Stand der Technik und das technische Gebiet sowie eine detaillierte Beschreibung der Erfindung und der Zeichnung enthält, ist auch international vergleichbar (Behrmann 1998, Knight 2004).

Patentdokumente enthalten textuelle und nicht-textuelle Patentinformationen (Abbildung 2). Der strukturell-grammatischen Textauffassung folgend (siehe hierzu im Detail Gansel 2009) zeichnen sich Texte insbesondere durch die Verknüpfung von mehreren Sätzen aus. Diese weisen einen inhaltlichen und einen grammatikalischen Zusammenhang auf. Diesen Textualitätskriterien entsprechend gehören Ansprüche, Zusammenfassung, Beschreibung der Erfindung und der beigefügten Zeichnungen zu den textuellen Patentinformationen. Zwar stellt der Titel eines Patentdokuments im Sinne der strukturell-grammatischen Textauffassung keinen Text an sich dar, wird aber im Zusammenhang mit weiteren textuellen Patentinformationen aufgrund seiner zumeist präzisen Angabe über den Inhalt eines Patentdokuments bei der Durchführung des PatMining berücksichtigt. Die bibliografischen Daten erfüllen die Textualitätskriterien nicht und gelten somit auch nicht als textuelle Patentinformationen. Ebenso handelt es sich auch bei Zeichnungen und Diagrammen nicht um textuelle Patentinformationen.

United States Patent [19]		[11]	4,241,621
Kodama		[45]	Dec. 30, 1980
[54]	TRANSMISSION APPARATUS FOR MOTOR VEHICLE	2,674,136 4/1954 Bryan	74/665 T
		2,714,936 8/1955 Gregory	180/49 X
		2,796,942 6/1957 Hill	74/700 UX
		2,796,943 6/1957 Rolt et al.	74/700 UX
		2,971,595 2/1961 Faberé et al.	74/665 T X
[75]	Inventor: Masayuki Kodama, Tokyo, Japan	Primary Examiner—Roy D. Frazier	
[73]	Assignee: Fuji Jukogyo Kabushiki Kaisha, Tokyo, Japan	Assistant Examiner—Peter A. Aschenbrenner	
[21]	Appl. No.: 918,464	Attorney, Agent, or Firm—Martin A. Farber	
[22]	Filed: Jun. 23, 1978		
[51]	Int. Cl. F16H 37/8	[57] ABSTRACT	
[52]	U.S. Cl. 74/700; 180/247; 74/665 T	A transmission apparatus for motor vehicle in which the engine is positioned in front side of the front axle, a sub-transmission is provided in the space formed above the front axle; a main transmission is adapted to transmit the output of the sub-transmission to the front wheels and to the rear wheels through means for transmitting the output, and the means for transmitting the output includes a clutch means for selectively connecting the transmission to the rear axle.	
[58]	Field of Search 74/700, 701, 745, 665 F, 74/665 G, 665 H, 665 S, 665 T; 180/44, 45-47, 49, 50	2 Claims, 12 Drawing Figures	
[56]	References Cited		
	U.S. PATENT DOCUMENTS		
	2,314,833 3/1943 Keese	180/49 X	
	2,357,781 9/1944 Randol	74/665 T X	

1	4,241,621	2
TRANSMISSION APPARATUS FOR MOTOR VEHICLE		nism will be great value, which makes the manipulation of the sub-transmission heavy.
BACKGROUND OF THE INVENTION		Therefore, it is the object of the present invention to provide a transmission apparatus which need not a special space for the sub-transmission in the transmission housing, which may be manufactured in small size, and of which sub-transmission may be easily manipulated.
The present invention relates to a transmission apparatus for motor vehicle in which the engine is longitudinally mounted in the front position of the vehicle and output of the engine is transmitted to the front and selectively to the rear axle of the vehicle through the transmission. More particularly the present invention relates to a transmission apparatus which transmits the power of the engine both to the front axle and to four wheels of front and rear.		SUMMARY OF THE INVENTION
In recent years, the four-wheel driving passenger car has been preferably and widely used, because it has high		In accordance with the present invention, the engine is positioned in front side of the front axle, a sub-transmission is provided in the space formed above the front axle, a main transmission is adapted to transmit the output of the sub-transmission to the front wheels and to the rear wheels through means for transmitting the output, and the means for transmitting the output includes a clutch means for selectively connecting the

Bibliografische (nicht-textuelle) Informationen (z.B. Erfinder, Anmelder, Patentklassifikation)  
 ≙ strukturierten Daten

Textuelle Informationen (z.B. Beschreibung, Zusammenfassung, Ansprüche)  
 ≙ unstrukturierten Daten

Abbildung 2: Generelle Unterscheidung von textuellen und nicht-textuellen Patentinformationen (Quelle: Möhrle et al. 2012 mit leichten Änderungen)

### 3.3 Text-Mining

Text-Mining zielt auf die Erschließung unstrukturierter, textueller Informationen (Hotho et al. 2005, Zhang, Segall 2010) ab und „helps to dig out the hidden gold from textual information“ (Dörre et al. 1999). Für das Text-Mining wurden unterschiedliche Rahmenmodelle in die Literatur eingeführt (siehe z. B. Hotho et al. 2005, Feldman, Sanger 2007, Zhang, Segall 2010). Die Grundarchitektur von Text-Mining-Systemen nach Feldmann und Sanger (2007) umfasst die Funktionen Vorverarbeitung („Pre-processing Tasks“), Informationserschließung („Core Mining Operations“) und Ergebnispräsentation („Presentation Layer Components“). Eine vergleichbare Architektur stammt von Wise et al. (1995), die die Funktionen Textverarbeitung („Processing text“) und Ergebnisvisualisierung („Visualizing output from processing“) berücksichtigen.

Die Vorverarbeitung hat die Aufgabe, den Rohtext zu säubern und für weitere Analysen zu strukturieren (Feinerer et al. 2008). Dabei werden repräsentative Bestandteile eines natürlich-sprachlichen Texts identifiziert und extrahiert. Hierdurch sollen die unstrukturierten Informationen eines Textes in analysierbare strukturierte bzw. semi-strukturierte Daten überführt werden (Feldman, Sanger 2007). Während der Vorverarbeitung werden beispielsweise Stoppwörter gefiltert und Synonyme ersetzt. Stoppwörter sind solche Begriffe, für die die Wahrscheinlichkeit des Vorkommens in einem relevanten Dokument genauso groß ist wie in einem nicht relevanten Dokument und die so keine Informationen über die thematische Relevanz eines Dokuments liefern (Stock 2007, Manning et al. 2008). Diese Begriffe werden aus dem Dokument entfernt. Synonyme werden zu Gruppen zusammengefasst und durch einen Begriff dieser Gruppe, dem festgelegten Zielbegriff, ersetzt, so dass die Begriffe bei der Auswertung vergleichbar sind. Nach der Filterung erfolgt die Extraktion sogenannter Konzepte. Bei Konzepten handelt es sich allgemein um textuelle Elemente, wie einfache Wörter, Begriffe, bestehend aus einem oder mehreren Wörtern, Satzteilen oder ganzen Sätzen (Carley 1997, Feldman, Sanger 2007, Möhrle 2010).

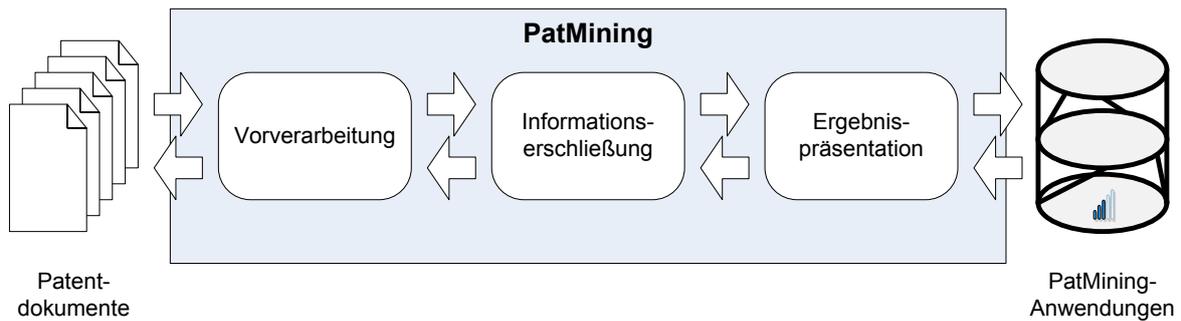
Die Informationserschließung dient dazu, Muster in den vorliegenden Dokumentensammlungen zu erkennen. Typische Aufgaben sind die Auswertung von Verteilungen, Häufigkeiten und Assoziationen (Feldman, Sanger 2007). Aufbauend auf den Assoziationen können dann Klassifikations- und Clusteranalysen (vgl. z. B. Hotho et al. 2005) sowie sog. Mappingverfahren (siehe hierzu z. B. Peters, van Raan 1993, Engelsman, van Raan 1994), die beispielsweise in semantischen Patentlandkarten resultieren (siehe hierzu Dressler 2006, Bergmann 2011), angewendet werden. Die Informationserschließung kann dazu genutzt werden, um neue Relationen zu entdecken, aber auch um bekannte Strukturen zu verfolgen (Dörre et al. 1999).

Die Ergebnispräsentation ermöglicht letztendlich erst die Analyse und Auswertung großzahliger Dokumentenmengen, indem die Ergebnisse in einfach und schnell zu verstehenden Darstellungen repräsentiert werden (Hotho et al. 2005). Dem Nutzer wird so die Arbeit erleichtert, da er für die Erlangung eines thematischen Überblicks davon befreit wird, sämtliche Texte zu lesen, und sich auf die für ihn relevanten Texte oder Textpassagen fokussieren kann (Rohrdantz et al. 2010).

### **3.4 Rahmenmodell des PatMining**

PatMining ist die Anwendung von Text-Mining-Methoden unter patentspezifischen Aspekten, wie die einheitliche Strukturierung von Patentdokumenten und der Möglichkeit der zusätzlichen Verwendung bibliografischer Informationen. Die bibliografischen Angaben werden daher beim PatMining unterstützend hinzugezogen. So kann es von Interesse sein zu wissen, wer ein Patent angemeldet hat, beispielsweise bei einer Wettbewerberanalyse, oder zu wissen, wann ein Patent angemeldet wurde. Dies ist insbesondere bei der Analyse technologischer Veränderungen von entscheidender Bedeutung. Die Zuordnung von Patenten zu bestimmten Zeitpunkten anhand der bibliografischen Daten ermöglicht erst entsprechende Analysen. Außerdem können patentstatistische Auswertungen genutzt werden, um erste Einblicke in ein Technologiefeld zu erhalten (z. B. Comanor, Scherer 1969, Haupt et al. 2007).

Die Grundarchitektur des PatMining umfasst in Anlehnung an die Grundarchitektur von Text-Mining-Systemen die Funktionen der Vorverarbeitung, Informationserschließung und Ergebnispräsentation. Die Anordnung der Funktionen ist von starker Parallelität geprägt. Beispielsweise kann nach der Ergebnispräsentation eine Anpassung der Patentmenge bzw. der Auswahl der Patentdokumente notwendig sein, was zum Aufgreifen aller drei Funktionen führen würde (Abbildung 3).



**Abbildung 3: Rahmenmodell des PatMining für das Technologie-Monitoring (Quelle: Eigene Darstellung basierend auf der Text-Mining Architektur von Feldman, Sanger 2007)**

Das PatMining stellt in Bezug zu den dieser Arbeit unmittelbar vorangegangenen Vorleistungen von Dressler (2006) und Bergmann (2011) zur semantischen Patentanalyse und der Darstellung in semantischen Patentlandkarten eine Verallgemeinerung der Verwendung von Text-Mining-Methoden für die Analyse von Patenten dar. Die semantische Patentanalyse ist somit eine spezifische Form, aber auch Grundlage für die Ausgestaltung weiterer Formen des PatMining.

#### **4 Konkretisierung der Forschungsfragen und Zuordnung der Beiträge**

Nachdem das Technologie-Monitoring als Rahmenkonzept für das PatMining eingeführt und eine methodische Grundlage geschaffen wurde, werden nunmehr die eingangs formulierten Forschungsfragen aufgegriffen und konkretisiert. Sodann wird gezeigt, in welcher Weise die dieser Arbeit zugrundeliegenden Publikationen einen Beitrag zur Beantwortung der Forschungsfragen liefern.

##### **4.1 Umfang und Frühzeitigkeit von Patenten**

Die Vorteile von Patentinformationen werden zwar häufig betont, aber selten überprüft. Dies ist sicherlich auch in Bezug auf die Objektivität, die Klassifikation, den Marktbezug, die Verfügbarkeit, die Vollständigkeit, die Standardisierung und die daraus abgeleitete Analysierbarkeit von untergeordneter Bedeutung, da diese Vortei-

le in der Regel durch die Patentämter bzw. das Patentgesetz garantiert bzw. sichergestellt werden. Umfang und Frühzeitigkeit, die sich aus der Aktualität von Patentdokumenten ableiten, unterliegen jedoch stark den Einflüssen unternehmensstrategischer Entscheidungen, die unmittelbar die Nützlichkeit von Patentinformationen für das Technologie-Monitoring beeinflussen.

Patentinformationen sind aufgrund der hohen Anmeldezahlen als umfangreich anzusehen. Gleiches gilt im Hinblick auf die unterschiedlichen Technologiefelder, für die Patentinformationen bereitstehen. Jedoch ist hierbei zu berücksichtigen, dass in der Regel nur ein bestimmter Anteil sämtlicher Erfindungen zum Patent angemeldet wird. Um die Nützlichkeit von Patentinformationen für das Technologie-Monitoring aber beurteilen zu können, bedarf es eines relativen Begriffsverständnisses des Umfangs. So sind Patentinformationen nur als „umfangreich“ für das Technologie-Monitoring zu bezeichnen, wenn diese tatsächlich einen relativ großen Teil sämtlicher Erfindungen eines Technologiefelds abbilden. Da aber aufgrund patentstrategischer Entscheidungen der in einem Technologiefeld tätigen Unternehmen nicht sämtliche Erfindungen zu Patenten angemeldet werden, ist der Umfang von Patentinformationen für das Technologie-Monitoring zu hinterfragen.

Patentinformationen wird eine hohe Aktualität zugeschrieben. Dies stimmt insofern, da Patentanmeldungen spätestens 18 Monate nach deren Einreichung veröffentlicht werden. Für die Verwendung von Patentinformationen im Technologie-Monitoring stellt sich jedoch ergänzend die Frage nach der Frühzeitigkeit der Patentinformationen. Frühzeitig meint dabei, dass Patentinformationen derart früh veröffentlicht werden, dass Unternehmen technologische Veränderungen wahrnehmen und darauf reagieren können. Dies hängt maßgeblich von patentstrategischen Entscheidungen Patent anmeldender Unternehmen ab: Meldet ein Unternehmen in einer frühen Pha-

se des Innovationsprozesses seine Erfindung zum Patent an, ist davon auszugehen, dass auch die Veröffentlichung der Patentschrift relativ frühzeitig erfolgt. Entscheidet sich ein Unternehmen hingegen für eine späte Patentanmeldung mit geringer Aktualität, ist auch die Frühzeitigkeit geringer. Nur wenn Erfindungen tatsächlich frühzeitig in der Patentliteratur zu erkennen sind, stellen Patentdokumente für das Technologie-Monitoring eine adäquate Informationsquelle dar.

Ausgehend von der Bedeutung des Umfangs und der Frühzeitigkeit von Patentinformationen, werden die Forschungsfragen in Bezug auf die Nützlichkeit von Patentinformationen für das Technologie-Monitoring wie folgt formuliert:

*F1.1: Wie umfangreich sind Patentinformationen für das Technologie-Monitoring verfügbar?*

*F1.2: Wie frühzeitig sind Patentinformationen für das Technologie-Monitoring verfügbar?*

#### **4.2 Methodische Umsetzung des PatMining für das Technologie-Monitoring**

Die methodische Umsetzung des PatMining für das Technologie-Monitoring setzt die Anpassung der vom Text-Mining bekannten Funktionen Vorverarbeitung und Informationserschließung voraus. Die Ergebnispräsentation wird maßgeblich durch die Informationserschließung bedingt und stellt vor allem ein programmiertechnisches Problem dar. Sie wird daher im Rahmen dieser Arbeit nicht explizit betrachtet. Die Funktion Vorverarbeitung umfasst Aufgaben wie die Entfernung von Satzzeichen oder Stoppwörtern und - von besonderer Bedeutung - die Extraktion von Konzepten. Entwurfsentscheidungen haben hierbei eine besondere Bedeutung, da diese grundlegend für die weiteren Funktionen der Informationserschließung und Ergebnispräsentation sind (Feldman, Sanger 2007). Im Gegensatz zu früheren, schlagwortba-

sierten Arbeiten (z. B. Yoon, Park 2005, Kim et al. 2008, Yoon 2008, Lee et al. 2009, Lee et al. 2011) werden im Rahmen dieser Arbeit kombinierte Konzepte aus dem gesamten Text extrahiert und gebildet. Dieser Besonderheit wird durch die Forschungsfrage F2.1 Rechnung getragen:

*F2.1: Welche Auswirkungen haben die Entwurfsentscheidungen bzgl. der Extraktion und Bildung kombinierter Konzepte bei der Durchführung des PatMining auf die Ergebnisse des Technologie-Monitorings?*

Die Methoden zur Informationserschließung im Rahmen des PatMining zielen grundsätzlich auf die Erfüllung der Aufgaben des Technologie-Monitorings ab: Neue Technologien sowie Strukturen von und zwischen Technologiefeldern sollen anhand von Technologieindikatoren entdeckt und überwacht werden. Hierauf zielen die Forschungsfragen F2.2 bis F2.4 dieser Arbeit ab:

*F2.2: Wie können (neue) Strukturen zwischen zwei Technologiefeldern durch die Auswertung textueller Patentinformationen mit Hilfe von Text-Mining-Methoden überwacht und entdeckt werden?*

*F2.3: Wie können (neue) Strukturen innerhalb eines Technologiefelds durch die Auswertung textueller Patentinformationen mit Hilfe von Text-Mining-Methoden überwacht und entdeckt werden?*

*F2.4: Wie können (neue) Technologien innerhalb eines Technologiefelds durch die Auswertung textueller Patentinformationen mit Hilfe von Text-Mining-Methoden überwacht und entdeckt werden?*

### 4.3 Zuordnung der berücksichtigten Publikationen zu den Forschungsfragen F1.1, F1.2, F2.1 bis F2.4

Unter Berücksichtigung der Aufgaben des Technologie-Monitorings und der Überwachung von Technologien auf verschiedenen Ebenen wurde das PatMining mit den Funktionen Vorverarbeitung, Informationserschließung und Ergebnispräsentation im Rahmen dieser Arbeit angepasst. Grundlage hierfür stellte die Überprüfung der Nützlichkeit von Patentinformationen für das Technologie-Monitoring dar. Entsprechend wurden Forschungsfragen in Bezug auf die Qualität von Patentinformationen sowie in Bezug auf die methodische Anpassung der Funktionen Vorverarbeitung und Informationserschließung abgeleitet. Diese Forschungsfragen wurden in insgesamt fünf Publikationen beantwortet (Tabelle 1).

**Tabelle 1: Zuordnung der in dieser Arbeit berücksichtigten Publikationen in das Rahmenmodell des PatMining unter Berücksichtigung der drei Ebenen des Technologie-Monitorings**

		Einzeltechnologie bezogen	Technologiefeld bezogen	Technologiefeld übergreifend
Patentinformationen		F1.1 und F1.2: Gerken, Möhrle, Walter 2012		
PatMining	Vorverarbeitung	F2.1: Möhrle, Gerken 2012		
	Informationserschließung	F2.4: Gerken, Möhrle 2012	F2.3: Gerken, Walter, Möhrle 2010a, 2010b	F2.2: Möhrle, Gerken 2012
	Ergebnispräsentation	—	—	—

Die Forschungsfragen F1.1 und F1.2 in Bezug auf die Qualität von Patentinformationen für das Technologie-Monitoring werden in Gerken et al. (2012) beantwortet. In dieser Publikation wurde für das Technologiefeld der Allradtechnik überprüft, ob Patentinformationen nützlich in Bezug auf die Frühzeitigkeit und den Umfang sind.

Darüber hinaus werden Technologiefeld übergreifende Einflussfaktoren auf die Nützlichkeit von Patentinformationen für das Technologie-Monitoring eingeführt.

Die Forschungsfrage F2.1 wird beantwortet, indem, basierend auf einer Literaturrecherche in Möhrle und Gerken (2012), verschiedene Entwurfsentscheidungen bei der Extraktion komplexer Konzepte aufgezeigt werden. Grundlegende Entwurfsentscheidungen sind dabei die Konzeptgröße und die Fenstergröße. Die Einflüsse von Konzept- und Fenstergröße werden anhand eines Fallbeispiels zur Beantwortung der Forschungsfrage F2.1 veranschaulicht.

Forschungsfrage F2.2 wird ebenfalls in Möhrle und Gerken (2012) beantwortet. Hierzu wird ein bekanntes Fallbeispiel zur Konvergenz untersucht. Es wurde hier die semantische Patentanalyse zur Erkennung konvergierender Technologien angepasst und erprobt.

In Gerken et al. (2010a, 2010b) wird auf die Forschungsfrage F2.3 eingegangen. Hier wird gezeigt, dass die Auswertung textueller Patentinformationen geeignet ist, um technologische Strukturen und Veränderungen in einem Technologiefeld abzubilden. Die Kombination verschiedener Methoden zur Informationserschließung ermöglicht zudem die einfache inhaltliche Erschließung der semantischen Patentlandkarten.

Zur Beantwortung der Forschungsfragen F2.4 wird in Gerken und Möhrle (2012) ein Technologieindikator eingeführt und erprobt. Der Technologieindikator ermöglicht die Messung des Grades der Neuheit einer Erfindung. Neuheit wurde exemplarische ausgewählt, da dies ein Merkmal radikaler Innovationen (z. B. Dahlin, Behrens 2005) sowie grundlegend für das Konzept der schwachen Signale ist (Zeller 2003).

#### 4.4 Darstellung der berücksichtigten Publikationen

##### 4.4.1 F1.1 und F1.2: Gerken, Möhrle und Walter 2012 (under review): *One Year Ahead! Investigating the Time Lag between Patent Publication and Market Launch: Insights from a Longitudinal Study in the Automotive Industry*

Das Technologie-Monitoring und die rechtzeitige Erkennung besonders neuartiger Technologien, die zu einem späteren Zeitpunkt in Produkten münden könnten, sind vor dem Hintergrund der Bedeutung technologischer Veränderungen eine wichtige Aufgabe für Unternehmen. In diesem Zusammenhang wird häufig die Nützlichkeit von Patentinformationen betont. Entscheidend für die Nützlichkeit von Patentinformationen zur rechtzeitigen Erkennung neuer Technologien sind jedoch (i) deren Frühzeitigkeit und (ii) deren relativer Umfang.

- (i) Patente sind für das Technologie-Monitoring und dem damit einhergehenden Erkennen technologischer Veränderungen nur nützlich, wenn technologische Veränderungen in Patentdokumenten vor deren Markteinführung zu erkennen sind. Daher kommt dem Zeitabstand zwischen der Patentveröffentlichung und der Einführung einer Technologie am Markt eine besondere Bedeutung zu. Der Zeitpunkt der Patentveröffentlichung ist der früheste Zeitpunkt, zu dem eine Erfindung für die Öffentlichkeit und damit auch für das betriebliche Technologie-Monitoring eingesehen werden kann. Ab dem Zeitpunkt der Markteinführung wird der Marktpionier versuchen, Pioniervorteile aufzubauen und so mögliche Wettbewerber aus dem neuen Markt auszuschließen. Die Herausforderung für Unternehmen besteht daher nach Veröffentlichung eines Patentdokuments in der Wahrnehmung einer technologischen Veränderung, der Entscheidung über eine Reaktion auf eine technologische Veränderung und ggf. der Nachahmung der Technologie. Der Zeitabstand zwischen Patentveröffentlichung und

Markteinführung bzw. erfolgreicher Imitation umfasst daher eine Wahrnehmungsverzögerung („*awareness lag*“), eine Entscheidungsverzögerung („*decision lag*“) und eine Nachahmungszeit („*imitation time*“) (Granstrand 2000).

- (ii) Weitere Voraussetzung der Nützlichkeit von Patentinformationen für das Technologie-Monitoring ist ein hoher relativer Umfang von Patentinformationen, denn nicht alle Erfindungen werden auch zum Patent angemeldet (Fleming 2001). Der Anteil der Erfindungen, die tatsächlich zum Patent angemeldet werden, variiert von Industrie zu Industrie (Arundel, Kabla 1998). Damit aber Patente in einem Technologiefeld für das Technologie-Monitoring geeignet sind, muss der Anteil der Erfindungen, die auch zum Patent angemeldet werden, hoch sein.

Gerken et al. (2012) setzt sich vorrangig mit dem Zeitabstand zwischen Patentveröffentlichung und Markteinführung auseinander. Darüber hinaus werden aber auch Einblicke bezüglich des relativen Umfangs von Patentinformationen gewährleistet. So trägt Gerken et al. (2012) zur Beantwortung der Forschungsfragen F1.1 und F1.2 bei:

*F1.1: Wie umfangreich sind Patentinformationen für das Technologie-Monitoring verfügbar?*

*F1.2: Wie frühzeitig sind Patentinformationen für das Technologie-Monitoring verfügbar?*

Zur Beantwortung dieser Fragen wurden zunächst (i) bestehende Studien zu dieser Thematik recherchiert und anhand eines Rahmenmodells von Granstrand (2000) systematisiert. Aufbauend auf den recherchierten Studien wurde dann (ii) ein Rahmenmodell, das Einflussfaktoren auf die Zeitabstände von Patenten und Produkten

darstellt, eingeführt. Anschließend (iii) wurde der Zeitabstand zwischen Patentveröffentlichung und Markteinführung anhand einer Fallstudie in der Antriebstechnik untersucht.

- (i) Bisherige Studien, die Einblicke in Zeitabstände von Patenten liefern, untersuchten zumeist den Zeitabstand zwischen Patentanmeldung und Erfolgswirksamkeit von Patenten. Dieser Zeitabstand ist zwar für die Erfolgswirksamkeit von Forschung und Entwicklung sowie Patenten von hoher Wichtigkeit, weniger aber für das Technologie-Monitoring. Dies ist damit begründet, dass die Patentveröffentlichung in der Regel erst 18 Monate nach der Patentanmeldung erfolgt. In diesem Zeitraum ist die Patentanmeldung nicht öffentlich zugänglich und damit für das Technologie-Monitoring nicht verwertbar. Die Erfolgswirksamkeit einer Technologie kann hinter der Markteinführung liegen. Der Zeitraum zwischen Markteinführung und Erfolgswirksamkeit kann jedoch durch den Pionier zum Aufbau von Pioniervorteilen, wie der Positionierung an zentralen geographischen Orten (Lieberman, Montgomery 1998), genutzt werden.
- (ii) Ein zweites zentrales Ergebnis der Literaturübersicht ist, dass die Zeitabstände, die in den verschiedenen Studien identifiziert wurden, voneinander abweichen. Ein systematischer Vergleich der bisherigen Studien ermöglichte die Identifizierung von Einflussfaktoren auf die Zeitabstände sowie die Darstellung dieser Einflussfaktoren in einem Rahmenmodell. Das Rahmenmodell diente zur Profilierung der durchgeführten Fallstudie.
- (iii) Für die Untersuchung des Zeitabstands zwischen der Patentveröffentlichung und der Markteinführung wurden zunächst Patente für zwei aufeinander aufbauende Patentsets recherchiert. Während sich das erste Patentset sehr stark auf die Allradtechnologie, insbesondere Differentiale, fokussiert, enthält das zweite Patentset allgemein Patente mit Bezug zu Allradfahrzeugen. Das erste

Patentset ist Untermenge des zweiten Patentsets. Für beide Patentsets wurden Patente identifiziert, die erstmals bestimmte Aspekte der SUBARU-Allradtechnologie beschreiben. Als Maßstab diente hier eine Historie der SUBARU-Allradtechnologie. Grundlage für die Zeitpunkte der Markteinführung war eine Pressemappe von SUBARU (2005) und eine ergänzende Internetrecherche. Die Berechnung der Zeitabstände erfolgte in Monaten.

Die Ergebnisse der Fallstudie zeigen, dass (i) Patente durchaus umfangreich genug sind, um produktrelevante Technologien zu identifizieren, und dass (ii) Unternehmen in der Antriebstechnik ca. ein einziges Jahr haben, um auf neue Technologien zu reagieren.

- (i) Der Umfang von Patenten wird in der Literatur häufig hervorgehoben. Letztendlich gibt es aber keine umfassenden Informationen darüber, wie groß dieser Schatz technologischer Informationen tatsächlich ist. In der hier durchgeführten Fallstudie konnten zahlreiche Technologien, die später im Markt eingeführt wurden, in Patenten erkannt werden. Diese Erkenntnis stärkt zunächst die Bedeutung von Patenten als Informationsquelle im Technologie-Monitoring. Zugleich muss aber einschränkend festgestellt werden, dass der relative Umfang stark von der Anzahl analysierter Patentdokumente abhängt. So wurde im Rahmen der Untersuchung ein zweistufiges Verfahren gewählt, wobei zunächst ein fokussiertes und dann in einem zweiten Schritt ein größeres Patentset untersucht wurde. Wie zu erwarten war, wurden in dem größeren Patentset eine höhere Anzahl der Technologien identifiziert, die später am Markt eingeführt wurden.
- (ii) Technologien konnten in Patentdokumenten zwischen 0 und 38 Monaten vor ihrer Einführung am Markt erkannt werden. Der durchschnittliche Zeitabstand

für das erste Patentset betrug 8,1 Monate, für das zweite Patentset 12,5 Monate. Vergleichbare Ergebnisse wurden bereits durch Ernst (2001) im Maschinenbau mit einem anderen methodischen Ansatz erzielt. Analog zu den Ergebnissen in Bezug auf den Umfang von Patentinformationen ist auch in Bezug auf die Frühzeitigkeit von Patentinformationen festzustellen, dass durch die Auswertung eines größeren Patentsets die Qualität der Information in Hinblick auf deren frühzeitige Verfügbarkeit erhöht wird. Obwohl einige Technologien erst zu einem eher späten Zeitpunkt identifiziert werden konnten, eröffnen Patentdokumente die Möglichkeit, frühzeitig Technologien zu erkennen, die später am Markt eingeführt werden.

Die Ergebnisse lassen die Schlussfolgerung zu, dass (i) Patentdokumente durchaus eine geeignete Quelle für die Erkennung neuer Technologien vor deren Markteinführung sind. (ii) Jedoch bedarf es effizienter interner Prozesse, um auf eine neue Technologie reagieren zu können. (iii) Das Technologie-Monitoring kann hier einen Beitrag durch neue, computergestützte Verfahren, die Patentdokumente schnell und im großen Umfang auswerten können, leisten.

- (i) Der hohe Umfang von Patentinformationen konnte durch die Fallstudie, zumindest für die Allradtechnik, bestätigt werden. Ebenso deuten die Ergebnisse der Fallstudie auf die Nützlichkeit von Patentinformationen in Bezug auf deren Frühzeitigkeit hin. Es wurde festgestellt, dass Technologien in der betrachteten Fallstudie ca. ein Jahr vor ihrer Markteinführung in Patentdokumenten veröffentlicht werden. Die durchschnittliche time-to-market beträgt im deutschen Maschinenbau 12,7 Monate (Voigt, Wettengl 1999). Dies bedeutet, dass Unternehmen sehr schnell entsprechende Signale wahrnehmen, über die Signale entscheiden und ggf. Entscheidungen umsetzen können und müssen. Für diese

schnellen Reaktionszeiten sollten Ressourcen durch das strategische Management bereitgehalten werden.

- (ii) Die notwendigen schnellen Reaktionen auf neue Technologien machen effiziente interne Prozesse notwendig, die Erkennung, Entscheidung und Umsetzung unterstützen. Die Handlungsfreiheit von Unternehmen wird letztendlich durch eine schnelle Erkennung von neuen Technologien erhöht. So hat das strategische Management beispielsweise die Möglichkeit, eine Technologiegeneration zu überspringen und sich auf die nächste Generation zu konzentrieren. Eine weitere Option stellt die Fokussierung auf die Markt- nicht aber auf die Technologieführerschaft dar.
- (iii) Strategische Handlungsfreiheit wird durch die frühzeitige, aber auch umfangreiche Erkennung neuer Technologien ermöglicht. Frühzeitigkeit und Umfang werden durch die Ausweitung des Patentsets erhöht. Dies setzt zeitgleich einen erhöhten Ressourcenbedarf bei der Überwachung von Patenten voraus. Die manuelle, inhaltliche Erschließung ist in der Regel zeitaufwendig und damit auch kostspielig. Entsprechend besteht der Bedarf computergestützter Methoden, die die gezielte technologische Überwachung ermöglichen und die Wahrnehmungsverzögerung verringern.

#### *4.4.2 F2.1 und F2.2: Möhrle und Gerken 2012: Measuring Textual Patent Similarity on the Basis of Combined Concepts: Design Decisions and their Consequences*

Methoden zur Auswertung textueller Patentinformationen beruhen zumeist auf der Auswertung einzelner sprachlicher Konzepte. Für die textuelle Analyse von Patentdokumenten ist jedoch ein einfaches Konzept in der Regel nicht ausreichend, da die Bedeutung eines einzelnen Konzepts meist nicht eindeutig und damit nicht präzise genug ist, um den Inhalt eines Patentdokuments zu repräsentieren (Tseng et al.

2007). Grund hierfür ist die Mehrdeutigkeit von Begriffen. So ist das Konzept „Bank“ erst in Verbindung mit weiteren Konzepten wie „Geld“ oder „Park“ richtig zu deuten. Die Bedeutung eines Konzepts ergibt sich also erst durch seinen Kontext. Der Kontext eines Konzepts wird durch seine Nähe zu weiteren Konzepten bestimmt. Daher ist auch für die Analyse textueller Patentinformationen die Auswertung kombinierter Konzepte statt einzelner Konzepte notwendig. Möhrle und Gerken (2012) setzen hier an und beantworten zunächst die Forschungsfrage F2.1. Aufbauend auf den Erkenntnissen der beantworteten Forschungsfrage F2.1 wird die Methode der semantischen Patentanalyse für die Erkennung konvergierender Industrien (F2.2) angepasst:

*F2.1: Welche Auswirkungen haben die Entwurfsentscheidungen bzgl. der Extraktion und Bildung kombinierter Konzepte bei der Durchführung des PatMining auf die Ergebnisse des Technologie-Monitorings?*

*F2.2: Wie können (neue) Strukturen zwischen zwei Technologiefeldern durch die Auswertung textueller Patentinformationen mit Hilfe von Text-Mining-Methoden überwacht und entdeckt werden?*

Aufbauend auf der Darstellung des generellen Prozesses der textuellen Patentanalyse werden zunächst zentrale Entwurfsentscheidungen bei der Bildung komplexer Konzepte erläutert. Es werden Möglichkeiten (i) der Extraktion und Bildung komplexer Konzepte aufgezeigt. Grundlage hierfür ist die Extraktion einzelner Konzepte. Sodann werden (ii) Methoden der Ähnlichkeitsberechnung für komplexe Konzepte erläutert. Die Auswirkungen der zuvor eingeführten Entwurfsentscheidungen werden anhand einer (iii) Patentpriorisierung untersucht. Abschließend wird die semantische Patentanalyse für (iv) die Analyse konvergierender Technologien angepasst und angewendet.

- (i) Grundlage für die Bildung komplexer Konzepte ist die Extraktion einzelner Konzepte. Diese kann auf verschiedenen sprachlichen Ebenen erfolgen. Beispiele hierfür sind Wörter, Satzteile oder ganze Sätze. Für Patentanalysen stellen Wörter, syntaktische Funktionen, wie Subjekt, Prädikat und Objekt, und technische Komponenten typische einfache Konzepte dar.

Die Bildung komplexer Konzepte schließt an die Extraktion einfacher Konzepte an. Entwurfsentscheidungen, die bei der Auswertung kombinierter Konzepte gegenüber der Auswertung einzelner Konzepte von besonderer Bedeutung sind, sind die Konzeptgröße und die Fenstergröße.

Die Größe eines kombinierten Konzepts wird durch die Anzahl einzelner Konzepte, die in einem kombinierten Konzept zusammengefasst werden, bestimmt. Ein kombiniertes Konzept besteht aus mindestens zwei einzelnen Konzepten. Die maximale Konzeptgröße ist grundsätzlich nicht beschränkt. Jedoch scheint es in der Regel nicht sinnvoll, eine Konzeptgröße zu wählen, die die Kombination von vier einfachen Konzepten übersteigt.

Kombinierte Konzepte werden aus einem Bereich zusammenhängender Konzepte gebildet. Dieser Bereich wird als Fenster bezeichnet. Die Größe dieses Fensters, welche durch die Vorgabe der Anzahl zusammenhängender Konzepte bestimmt wird, ist die Fenstergröße. Die Fenstergröße muss mindestens so groß sein wie die gewählte Konzeptgröße. Wird eine Fenstergröße gewählt, die die Konzeptgröße übersteigt, so ist es möglich mehr als ein kombiniertes Konzept aus einem Fenster zu extrahieren. Beispielsweise werden bei einer Konzeptgröße von zwei und eine Fenstergröße von drei je Fenster drei kombinierte Konzepte gebildet.

- (ii) Die Ähnlichkeitsbestimmung besteht aus der Variablendefinition und -messung sowie der Ähnlichkeitsberechnung (Möhrle 2010). Während die Ähnlichkeitsbe-

rechnung weitestgehend analog zu den bei Möhrle (2010) vorgestellten Formeln für die Ähnlichkeitsberechnung bei einfachen Konzepten erfolgen kann, sind bei der Variablendefinition und -messung weitere Entwurfsentscheidungen zu beachten. So besteht die Möglichkeit des Vergleichs zweier kombinierter Konzepte hinsichtlich ihrer vollständigen Übereinstimmung, partieller Übereinstimmung und einer Kombination von vollständiger und partieller Übereinstimmung. Die Kombination von partieller und vollständiger Übereinstimmung eröffnet perspektivisch die Möglichkeit des Vergleichs kombinierter Konzepte unterschiedlicher Größe.

- (iii) Die Patentpriorisierung stellt eine typische Aufgabe im Technologie-Monitoring dar. Die Priorisierung ermöglicht grundsätzlich die Identifizierung von Patenten, die eine hohe inhaltliche Nähe zu einem vorgegebenen Referenzpatent aufweisen. Sie kann beispielsweise die Patentrecherche unterstützen und dazu dienen, Patentverletzungen zu erkennen (siehe hierzu z. B. Gerken et al. 2010c, Bergmann 2011). Für das Technologie-Monitoring kann die Priorisierung genutzt werden, um neue Wettbewerber in einem technologischen Gebiet zu identifizieren oder aber ein technologisches Gebiet unabhängig von Patentklassifikation und Schlagwortrecherche abzugrenzen. Anhand eines Fallbeispiels wurde die Auswirkung von Entwurfsentscheidungen auf die Ergebnisse der Patentpriorisierung überprüft. Im Mittelpunkt standen dabei die Konzeptgröße und die Fenstergröße. Konzept- und Fenstergröße wurden variiert, so dass für ein Anwendungsbeispiel insgesamt drei Priorisierungen mit unterschiedlichen Parametern vorgenommen wurden. Als einfache Konzepte wurden technische Komponenten von Erfindungen extrahiert. Die Ergebnisse der Patentpriorisierungen wurden dann mit einer Rangkorrelation und deskriptiven Statistiken ausgewertet und verglichen.

- (iv) Aufbauend auf den Erkenntnissen bzgl. der Entwurfsentscheidungen wurde die semantische Patentanalyse für die Messung und die Erkennung von Konvergenz zwischen Industrien angepasst und erprobt. Die Erprobung erfolgt anhand eines bekannten Beispiels zur Konvergenz der Pharma- und Chemieindustrie (siehe hierzu Curran, Leker 2011). Zunächst wurden Patente recherchiert. Die recherchierten Patente wurden in drei unterschiedlichen Auswertungsläufen hinsichtlich ihrer inhaltlichen Nähe mit der semantischen Patentanalyse untersucht. Hierbei wurden Konzept- und Fenstergröße variiert. Die Ergebnisse der unterschiedlichen Auswertungsläufe wurden mit deskriptiven Statistiken und in Bezug auf ihre Signale hinsichtlich der Konvergenz von Pharma- und Chemieindustrie ausgewertet.

Als Ergebnisse dieser Arbeit für das PatMining im Technologie-Monitoring wurden (i) die Auswirkungen von Entwurfsentscheidungen bei der Bildung kombinierter Konzepte aufgezeigt. (ii) Darauf aufbauend wurde die semantische Patentanalyse für die Erkennung konvergierender Industrien angepasst.

- (i) Die Wahl von Konzept- und Fenstergröße ist von zentraler Bedeutung für die Bildung kombinierter Konzepte, da diese die Ergebnisse der semantischen Patentanalyse maßgeblich beeinflussen. Wird die semantische Patentanalyse mit großen kombinierten Konzepten durchgeführt, ist ein sehr präziser inhaltlicher Vergleich zwischen Patenten möglich. Auftretende Ähnlichkeiten weisen auf Übereinstimmungen längerer textueller Elemente hin. Entsprechend ist auf eine hohe inhaltliche Ähnlichkeit zu schließen. Es besteht aber hierbei die Gefahr des Übersehens von geringen Ähnlichkeiten zwischen Patenten. Die Nutzung kleiner Konzepte ermöglicht die Erkennung der Übereinstimmungen kleiner tex-

tueller Elemente. So wird zwar eine Vielzahl ähnlicher Patente identifiziert. Die inhaltliche Ähnlichkeit kann jedoch gering sein.

Die Fenstergröße wirkt entgegengesetzt zur Konzeptgröße. Eine Erhöhung der Fenstergröße erhöht die Anzahl der Ähnlichkeiten mit einem Wert größer als null in einem Patentset. Mit der Erhöhung der Fenstergröße wird jedoch auch der Kontext, aus dem die kombinierten Konzepte gebildet werden, erweitert. Daher weisen die kombinierten Konzepte einen geringeren kontextuellen Zusammenhang aus, so dass davon auszugehen ist, dass auftretende Ähnlichkeiten bei einer hohen Fenstergröße nicht zwangsläufig auf eine große inhaltliche Nähe der verglichenen Patente schließen lässt. Entscheidet sich der PatMining-Nutzer jedoch für eine kleine Fenstergröße, weisen auftretende Ähnlichkeiten sehr wahrscheinlich auf eine inhaltliche Nähe verglichener Patenten hin.

- (ii) Aufbauend auf diesen Erkenntnissen wurde die semantische Patentanalyse für die Erkennung konvergierender Industrien angepasst und erprobt. In den drei unterschiedlichen Auswertungsläufen konnte jeweils die zunehmende Ähnlichkeit zwischen den Industrien beobachtet werden. Daher kann auf ein Konvergieren dieser Industrien geschlossen werden. Jedoch war das Signal für die Konvergenz bei den drei Auswertungsläufen unterschiedlich stark. Eine Fenstergröße von zwei, in Kombination mit einer Konzeptgröße von zwei, führte zu dem deutlichsten Signal von Konvergenz. Die Ergebnisse der beiden übrigen Auswertungsläufe waren miteinander vergleichbar. Eine kleine Konzept- und Fenstergröße stellt offensichtlich die Anzeichen der Konvergenz verstärkt heraus. So scheint es aus praktischer Sicht angebracht, zunächst mit kleiner Konzept- und Fenstergröße zu überprüfen, ob Konvergenz vorliegt. Wird anhand einer kleinen Konzept- und einer kleinen Fenstergröße Konvergenz beobachtet, können Fenster- und Konzeptgröße erhöht werden, so dass auch eine Ein-

schätzung der Stärke der Konvergenz möglich ist. Bleibt bei einer hohen Fenster- und Konzeptgröße das Signal für Konvergenz erhalten, ist dies ein Anzeichen dafür, dass bereits größere zusammenhängende Beschreibungen von Technologien in Patentdokumenten beider Industrien verwendet werden. Entsprechend ist auf eine große inhaltliche Nähe und damit auf eine starke Konvergenz zu schließen.

*4.4.3 F2.3: Gerken, Möhrle und Walter, 2010a,b: Patents as an Information Source for Product Forecasting: Insights from a Longitudinal Study in the Automotive Industry & Semantische Patentlandkarten zur Analyse technologischen Wandels: Eine Längsschnittstudie aus der Allradtechnik*

Ausgehend von der Relevanz von Technologien und technologischen Veränderungen einerseits und der Informationsflut andererseits, erproben Gerken et al. (2010a, 2010b) die semantischen Patentanalyse für das Technologie-Monitoring auf Ebene eines Technologiefelds. Hierbei werden exemplarisch Strukturen innerhalb eines Technologiefelds anhand einer semantischen Patentlandkarte aufgezeigt und so die Forschungsfrage F2.3 beantwortet:

*F2.3: Wie können (neue) Strukturen innerhalb eines Technologiefelds durch die Auswertung textueller Patentinformationen mit Hilfe von Text-Mining-Methoden überwacht und entdeckt werden?*

Die Beantwortung der Forschungsfrage F2.3 erfolgt in Gerken et al. (2010a) und Gerken et al. (2010b).

Gerken et al. (2010a) passen dabei zunächst (i) die semantischen Patentanalyse situations- und domänenspezifischen an und (ii) überprüfen, ob sich technologische Veränderungen in semantischen Patentlandkarten abbilden.

- (i) Im Rahmen eines Forschungsvorhabens mit der Forschungsvereinigung Antriebstechnik (FVA) wurden Sprachfilter für die semantische Patentanalyse im Technologiefeld Antriebstechnik entwickelt. Es wurden Synonymfilter und Stoppwortlisten erstellt. Darauf aufbauend wurden die Sprachfilter für die Fallstudie in der Allradtechnik adaptiert. Dabei wurde weniger das Gesamtfahrzeug betrachtet, sondern vielmehr charakteristische Aspekte eines Allradantriebs, beispielsweise das Zentraldifferential. Der Sprachfilter wurde um diese charakteristischen Aspekte und deren Synonyme erweitert. Entscheidend ist hierbei insbesondere das Verhältnis zwischen Ober- und Unterbegriffen. So kann es für eine bestimmte Anwendungssituation notwendig sein, Ober- und Unterbegriffe zu differenzieren. In einer anderen Anwendungssituation aber können gleiche Begriffspaare wieder vereinheitlicht werden. Diese Entscheidung hängt maßgeblich von dem Ziel und dem Detaillierungsgrad der Untersuchung ab. Grundlegend für die Anpassung der Sprachfilter war in dieser Publikation die Systematisierung von Allradantrieben von Naunheimer et al. (2007).
- (ii) Als Ergebnisse dieser Publikation ist neben der erfolgreichen domänen- und situationsspezifischen Adaption der semantischen Patentanalyse die Erkenntnis, dass sich technologische Veränderungen auf der Ebene eines Technologiefelds in der Tat in semantischen Patentlandkarten darstellen lassen. Patente, deren Anmeldedatum dicht beieinander liegen, weisen auch in der semantischen Patentlandkarte eine räumliche Nähe auf. Die Plausibilität der semantischen Patentlandkarte wurde durch manuelle inhaltliche Erschließung der Patente sowie der daran anschließenden formalen Klassifikation überprüft. Inhaltliche Überschneidungen nahe beieinanderliegender Patente konnten identifiziert und formal klassifiziert werden, so dass die Ergebnisse der semantischen Patentanalyse plausibel erscheinen.

Die Publikation Gerken et al. (2010b) stellt eine Fortführung der zuvor vorgestellten Publikation Gerken et al. (2010a) dar. Im Mittelpunkt steht die Weiterentwicklung der semantischen Patentlandkarten im Hinblick auf die Ergebnisqualität und die Interpretierbarkeit der semantischen Patentlandkarten.

Ausgangspunkt war hierfür (i) die Fehleranfälligkeit der Multidimensionalen Skalierung und (ii) der hohe zeitliche Aufwand bei der Interpretation bisheriger semantischer Patentlandkarten.

- (i) Die Multidimensionale Skalierung ist ein Verfahren zur Dimensionsreduktion (Wollschläger 2010) und ermöglicht so die Darstellung hochdimensionaler Daten in einem zweidimensionalen Raum. Grundsätzlich sollen dabei Objekte, die eine hohe Ähnlichkeit aufweisen, dicht beieinander im nieder-, zumeist zweidimensionalen Raum platziert werden (Kruskal 1964). Objekte mit einer geringeren Ähnlichkeit sollen im zweidimensionalen Raum eine entsprechend höhere Distanz zueinander aufweisen. Aufgrund der zahlreichen Abhängigkeiten zwischen den Objekten ist dies nicht immer eindeutig möglich. Die Darstellung der Objekte im zweidimensionalen Raum ist somit kein exaktes Abbild der Ähnlichkeiten der Objekte und folglich mit Ungenauigkeiten behaftet. Diese Ungenauigkeiten werden mit dem sog. Stresswert berechnet. Ein geringer Stresswert deutet auf eine hohe Anpassungsgüte hin, ein hoher Stresswert hingegen ist ein Hinweis auf eine geringe Anpassungsgüte. Obwohl ein Stresswert kleiner als 0,2 als akzeptabel angesehen wird (Kruskal 1964), unterliegt die resultierende zweidimensionale Darstellung doch gewissen Ungenauigkeiten. Ein typischer Fehler der Multidimensionalen Skalierung ist der sog. „horseshoe“-Effekt (van Eck, Waltman 2007). Hiermit ist gemeint, dass das Verfahren der Multidimensionalen Skalierung dazu neigt, Objekte im Halbkreis anzuordnen. So kann

zwar davon ausgegangen werden, dass Objekte, die weit auseinander liegen, eine geringe Ähnlichkeit haben. Jedoch bedeutet dies noch nicht, dass Objekte, die dicht beieinander liegen, tatsächlich auch sehr ähnlich sind.

- (ii) Die Ungenauigkeit der Multidimensionalen Skalierung sowie die fehlende Abgrenzung unähnlicher Objekte in der zweidimensionalen Darstellung bedingt die schwere Interpretierbarkeit der semantischen Patentlandkarten. So ist, wie bereits oben dargestellt, eine Ähnlichkeit zwischen dicht beieinander liegenden Patenten nicht zwangsläufig gegeben. Ebenso können aber auch Patente, die eine hohe Ähnlichkeit aufweisen, aufgrund einer geringen Anpassungsgüte im zweidimensionalen Raum relativ weit voneinander entfernt platziert werden. Die Ungenauigkeit der Multidimensionalen Skalierung erschwert die Klassifikation von Patenten in einer semantischen Patentlandkarte ungemein, da Grenzen zwischen verschiedenen Klassen nicht intuitiv gefunden werden können. Die formale Klassifikation setzt die inhaltliche Erschließung zahlreicher Patente, wenn nicht sogar aller Patente, in einer Patentlandkarte voraus, so dass die Klassifizierung zeitaufwendig und damit auch ressourcenintensiv ist.

Der Fehleranfälligkeit und dem damit einhergehenden hohen Aufwand bei der Interpretation semantischer Patentlandkarten wurde mit der Kombination von Multidimensionaler Skalierung und Clusteranalyse begegnet. Die verwendete hierarchische Clusteranalyse ermöglicht das Auffinden von Clustern auf verschiedenen Clusterlevel. Diese werden in der semantischen Patentlandkarte farblich unterschieden. Der Nutzer der semantischen Patentlandkarte erkennt so (i) Ungenauigkeiten der Multidimensionalen Skalierung. Wichtiger ist jedoch, dass (ii) Cluster in der semantischen Patentlandkarte auf verschiedenen Ebenen intuitiv erkannt und interpretiert werden können.

- (i) Die Clusteranalyse basiert auf den berechneten Ähnlichkeitswerten zwischen den Patenten. Sie ist somit unabhängig von der Multidimensionalen Skalierung. Ungenauigkeiten der Multidimensionalen Skalierung werden deutlich, da anhand der Ergebnisse der Clusteranalyse zu erkennen ist, ob nahe beieinander liegende Patente tatsächlich einem Cluster und ob weit entfernte Patente evtl. doch einem gemeinsamen Cluster angehören.
- (ii) Während bei den bisherigen semantischen Patentlandkarten das Aufdecken von Clustern eine zeitaufwendige Aufgabe darstellte, sind diese durch die Kombination von Multidimensionaler Skalierung und Clusteranalyse intuitiv auf verschiedenen hierarchischen Ebenen zu erkennen. So kann der Nutzer einer semantischen Patentlandkarte sich vollständig auf die Herausarbeitung der Gemeinsamkeiten der Patente eines Clusters konzentrieren. Die Abgrenzung zu anderen Clustern erfolgt durch die Clusteranalyse. Dies vereinfacht die inhaltliche Interpretation.

#### 4.4.4 F2.4: Gerken und Möhrle 2012: A New Instrument for Technology Monitoring:

##### *Novelty in Patents Measured by Semantic Patent Analysis*

Ausgehend von der Tatsache, dass technologische Veränderungen maßgeblich durch Technologien mit einem hohen Neuheitsgrad getrieben werden, ist die Identifikation dieser Technologien eine zentrale Aufgabe des Technologie-Monitorings. Diese Aufgabe steht bei Gerken und Möhrle (2012) zur Beantwortung der Forschungsfrage F2.4 im Mittelpunkt:

*F2.4: Wie können (neue) Technologien innerhalb eines Technologiefelds durch die Auswertung textueller Patentinformationen mit Hilfe von Text-Mining-Methoden überwacht und entdeckt werden?*

Zunächst wurden (i) bestehende Methoden zur Bestimmung der Neuheit von Patenten bzw. der in Patentdokumente beschriebenen Technologien vorgestellt. (ii) Die Methode der semantischen Patentanalyse wurde angepasst. (iii) Die angepasste Methode der semantischen Patentanalyse wurde dann mit bestehenden Methoden zur Messung der Neuheit in einer zweistufigen Fallstudie verglichen.

- (i) Die Darstellung bestehender Methoden zur Neuheitsmessung zeigt, dass die meisten Methoden auf der Auswertung bibliografischer Informationen, zumeist Zitationen, beruhen (z. B. Trajtenberg et al. 1997, Dahlin, Behrens 2005, von Wartburg et al. 2005). Textuelle Patentinformationen fanden bisher wenig Berücksichtigung. Methoden, die textuelle Patentinformationen nutzen, sind auf die Auswertung von Schlüsselbegriffen begrenzt (z. B. Kim et al. 2008, Lee et al. 2009) oder lassen keine Auswertung auf Ebene einzelner Patente zu (z. B. Yoon, Kim 2011).
- (ii) Der Prozess der semantischen Patentanalyse wurde daher für die Messung von Neuheit auf Ebene einzelner Technologien angepasst. Er umfasst die Schritte der Extraktion semantischer Strukturen, Anwendung domänen- und situationspezifischer Sprachfilter, Berechnung der Ähnlichkeiten sowie die Berechnung der Neuheit. Während die Schritte der Extraktion semantischer Strukturen und der Anwendung domänen- und situationsspezifischer Sprachfilter als Aufgabe der Vorverarbeitung unverändert genutzt blieben, mussten die Schritte der Ähnlichkeitsberechnung und der Berechnung der Neuheit situationsadäquat als neue Methode der Informationserschließung ergänzt werden. Die Anpassung der semantischen Patentanalyse erfolgte unter der Annahme, dass Neuheit nur vor dem Hintergrund von Bestehendem bewertet werden kann. Daher wurde ein Ähnlichkeitsmaß genutzt, das den Bestandteil eines Patentdokuments berechnet, der bereits in einem anderen Patentdokument beschrieben wurde. Die

Ähnlichkeitsberechnung zwischen Patentdokumenten erfolgt dabei paarweise unter Berücksichtigung des Anmeldedatums. Der resultierende Ähnlichkeitswert liegt in den Grenzen von null bis eins. Basierend auf den Ergebnissen der Ähnlichkeitsberechnung erfolgt die Berechnung der Neuheit eines Patents. Die maximale Ähnlichkeit des betrachteten Patents zu den zuvor angemeldeten Patenten wird von eins subtrahiert. Die gemessene Neuheit eines Patents ist somit der Anteil eines Patentdokuments, der nicht in dem vorangegangenen Patentdokument mit der maximalen Ähnlichkeit beschrieben wird.

- (iii) Für die Erprobung wurde das Fallbeispiel, das bereits im Zusammenhang mit Gerken und Möhrle (2012) vorgestellt wurde, aufgegriffen und die Patentsets übernommen. Ziel war es nun, in den zwei aufeinander aufbauenden Patentsets Patente zu identifizieren, die Innovationen der SUBARU-Allradtechnologie mit herausragender Neuheit beschreiben. Für beide Patentsets wurden Patente identifiziert, die erstmals bestimmte Aspekte der SUBARU-Allradtechnologie beschreiben. Als Maßstab diente hier eine Historie der SUBARU-Allradtechnologie. Die Patente wurden mit der semantischen Patentanalyse sowie weiteren Verfahren, die auf der Auswertung von Zitationen beruhen, hinsichtlich ihrer Neuheit bewertet. Die Ergebnisse wurden mit den Maßen Precision und Recall ausgewertet und verglichen. Precision und Recall sind Qualitätsmaße aus der Informationswissenschaft. Precision stellt dabei den Anteil der relevanten Dokumente an der Gesamtmenge der gefundenen Dokumente dar. Recall ist hingegen der Anteil der gefunden relevanten Dokumente an der Gesamtmenge der relevanten Dokumente (Manning et al. 2008). Recall spiegelt somit die Vollständigkeit, Precision die Genauigkeit der Recherche wider (Stock 2007). Zusätzlich wurden ein Mann-Whitney U-Test sowie eine Rangkorrelation durchgeführt (siehe hierzu Mann, Whitney 1947, Bühl 2010).

Grundlegendes Ergebnis bei Gerken und Möhrle (2012) für das Technologie-Monitoring ist (i) die Anpassung der semantischen Patentanalyse für die Messung von Neuheit und die Erkennung von Technologien mit herausragender Neuheit. Die Ergebnisse der Fallstudie liefern außerdem Erkenntnisse in Hinblick auf (ii) die Nützlichkeit der semantischen Patentanalyse für die Messung von Neuheit. (iii) Die Nützlichkeit dieser wie auch anderer Methoden hängt jedoch stark von dem Umfang des Patentsets ab. (iv) Die verschiedenen Verfahren zur Messung der Neuheit unterscheiden sich nicht nur hinsichtlich der Ergebnisgüte, sondern auch in Hinblick auf die Ergebnisse an sich.

- (i) Im Rahmen dieser Publikation wurde ein Verfahren entwickelt, dass aus textuellen Patentinformationen sog. Subject-Action-Object-Strukturen (SAO-Strukturen) extrahiert und bzgl. der Neuheit von Erfindungen ausgewertet. Dieses Verfahren kann das Technologie-Monitoring unterstützen. Durch die Priorisierung der Patente anhand ihrer Neuheit können Unternehmen gezielt ihre Aufmerksamkeit auf Technologien mit hoher Neuheit richten, Ressourcen schonen und Risiken senken. Hierbei wurde jedoch mit den SAO-Strukturen ein spezifischer Typ sprachlicher Konzepte, die funktionelle Zusammenhänge repräsentieren, verwendet. Möhrle und Gerken (2012) haben jedoch gezeigt, dass zahlreiche weitere Konzepte aus Patentdokumenten extrahiert werden können, beispielsweise Bauteile. So wurde eine erste Möglichkeit zur Messung von Neuheit vorgestellt und eine Grundlage für die Messung von Neuheit mit textuellen Patentinformationen geschaffen. Die Erweiterung der Methode auf weitere Konzepte eröffnet die Möglichkeit der Differenzierung zwischen verschiedenen Neuheitsarten.
- (ii) Der Vergleich der semantischen Patentanalyse mit anderen Methoden in Hinblick auf Precision und Recall deutet darauf hin, dass die semantische Pa-

tentanalyse durchaus eine wertvolle Unterstützung für die Identifikation von Technologien mit einem hohen Neuheitsgrad sein kann. In dem hier untersuchten Fallbeispiel war die semantische Patentanalyse den übrigen Methoden zu meist überlegen. Dies wurde auch durch die Rangkorrelation bestätigt. Außerdem wurde die Neuheit von Patenten mit der semantischen Patentanalyse differenzierter beurteilt als mit anderen Verfahren.

- (iii) Die Auswertung der Rangkorrelation sowie von Precision und Recall weisen jedoch darauf hin, dass die verschiedenen Methoden zur Messung von Neuheit in dem kleineren, klar fokussierten Patentset die besseren Resultate erzielten.
- (iv) Außerdem konnte festgestellt werden, dass die Nützlichkeit der semantischen Patentanalyse und der übrigen Methoden von den Rahmenbedingungen abhängig sind. Es wurde erkannt, dass das Maß TechB (siehe hierzu Trajtenberg et al. 1997) dem hier verwendeten Technologieindikator, basierend auf der semantischen Patentanalyse, überlegen ist, wenn Patente in neue, bisher weniger beachtete Patentklassen eingeordnet werden. Entsprechend scheint es zukünftig sinnvoll, eine Kombination der semantischen Patentanalyse mit weiteren Methoden anzustreben.

## **5 Zusammenfassung und Ausblick**

Mit dieser kumulativen Dissertation wurde der Frage nachgegangen, wie textuelle Patentinformationen mittels Text-Mining-Methoden für das Technologie-Monitoring genutzt werden können. Zunächst wurden hierzu die Aufgaben des Technologie-Monitorings und Grundlagen sowie Rahmenmodell des Text-Mining erläutert. Vor dem Hintergrund der besonderen Anforderungen des Technologie-Monitorings wurden dann Forschungsfragen in Bezug auf den Informationscharakter von Patentinformationen und die methodische Nutzung von Text-Mining formuliert. Die For-

schungsfragen wurden in insgesamt fünf Publikationen beantwortet. Ergebnisse in Bezug auf den Informationscharakter sind, dass Patente (i) umfangreich und (ii) frühzeitig für das Technologie-Monitoring verfügbar sind. (iii) Die Frühzeitigkeit von Patentinformationen variiert jedoch und unterliegt verschiedenen Einflussfaktoren:

- (i) Patentinformationen werden in der Regel zahlreiche Vorteile zugesprochen, u.a. auch ein hoher Umfang. Jedoch ist für das Technologie-Monitoring nicht die absolute Zahl von Patentdokumenten relevant, sondern vielmehr, dass Erfindungen, die zu einem späteren Zeitpunkt auch am Markt eingeführt werden, umfangreich in Patentdokumenten zu erkennen sind. Gerken et al. (2012) weisen darauf hin, dass tatsächlich zahlreiche produktrelevante Erfindungen in Patentdokumenten zu erkennen sind und Patente damit einen hohen relativen Umfang aufweisen.
- (ii) Der Zeitabstand zwischen Patentveröffentlichung und Markteinführung der geschützten Erfindung wurde bisher nur wenig untersucht. Im Rahmen dieser Arbeit wurde aufgezeigt, dass der Zeitabstand zwischen Patentveröffentlichung und Markteinführung ca. ein Jahr beträgt. Dies stellt eine große Herausforderung an Unternehmen bzgl. des Erkennens technologischer Veränderungen, des Entscheidens über Reaktionsmaßnahmen sowie des Durchführens der beschlossenen Maßnahmen dar. Daher hat das Technologie-Monitoring die Aufgabe, das Erkennen technologischer Veränderungen zu unterstützen und so die Wahrnehmungsverzögerung zu verringern.
- (iii) Zwar lag der Zeitabstand zwischen Patentveröffentlichung und Markteinführung im Mittelwert bei ca. einem Jahr, jedoch variierte dieser Zeitabstand in der durchgeführten Fallstudie, aber auch zwischen vorangegangenen Studien anderer Autoren, die zumeist den Zeitabstand zwischen Patentanmeldung und Erfolgswirksamkeit von Patenten analysierten, stark. Daher wurden Einflussfaktoren

ren auf den Zeitabstand zwischen Patentveröffentlichung und Markteinführung durch den Vergleich früherer Studien ermittelt.

Aus methodischer Sicht wurden (i) Auswirkungen von Entwurfsentscheidungen bei der Bildung kombinierter Konzepte untersucht und Text-Mining basierte Methoden für die Auswertung textueller Patentinformationen derart angepasst, dass (ii) Strukturen zwischen Technologiefeldern, (iii) Strukturen in Technologiefeldern und (iv) einzelne Technologien überwacht werden können.

- (i) Zentrale Entwurfsentscheidungen bei der Bildung kombinierter Konzepte sind die Konzept- und die Fenstergröße. Während eine Vergrößerung der Konzeptgröße die Wahrscheinlichkeit des Auftretens von Ähnlichkeiten sowie die Ähnlichkeitswerte zwischen zwei Dokumenten verringert, wird durch die Vergrößerung der Fenstergröße die Wahrscheinlichkeit des Auftretens von Ähnlichkeiten sowie die Ähnlichkeitswerte zwischen zwei Dokumenten erhöht. Wird die Konzeptgröße hingegen verringert, ist auch mit höheren Ähnlichkeitswerten sowie einer erhöhten Wahrscheinlichkeit des Auftretens von Ähnlichkeiten zu rechnen. Eine kleine Fenstergröße verringert hingegen die Wahrscheinlichkeit des Auftretens von Ähnlichkeiten sowie die Ähnlichkeitswerte zwischen zwei Dokumenten. Die Wahl von Konzept- und Fenstergröße hängt maßgeblich von dem Ziel der Untersuchung ab: Wird mit hoher Präzision, in anderen Worten mit einer großen Suchtiefe, gesucht, ist die Patentanalyse mit einer hohen Konzept- und niedrigen Fenstergröße durchzuführen. Wird hingegen ein hoher Recall bzw. eine hohe Suchbreite angestrebt, um geringe Ähnlichkeiten aufzudecken, sollte die Patentanalyse mit einer kleinen Konzept- und einer höheren Fenstergröße durchgeführt werden.

- (ii) Die Erkenntnisse zur Konzept- und Fenstergröße stellen die Grundlage zur Anpassung der semantischen Patentanalyse für die Erkennung konvergierender Industrien dar. Die semantische Patentanalyse wurde für diesen Zweck angepasst und anhand eines bekannten Beispiels erprobt. Hierbei wurde deutlich, dass die semantische Patentanalyse durchaus dazu geeignet ist, konvergierende Industrien zu erkennen.
- (iii) Gerken et al. (2010a, 2010b) passten das bestehenden Verfahren der semantischen Patentanalyse für das Technologie-Monitoring an und erprobten dies. Sprachfilter wurden hierzu erstellt und die semantische Patentanalyse in Hinblick auf die Ergebnisvalidität und die intuitive Interpretierbarkeit weiterentwickelt. Dies wurde durch die Kombination der Multidimensionale Skalierung und Clusteranalyse erreicht. Die Anwendung des Verfahrens zeigt, dass technologische Strukturen und Veränderungen in einem Technologiefeld durch die semantische Patentanalyse aufgezeigt werden können.
- (iv) Zur Bewertung von einzelnen Erfindungen wurde ein Technologieindikator zur Messung von Neuheit, basierend auf der Auswertung textueller Patentinformationen, eingeführt und evaluiert. Der Vergleich mit anderen Methoden zur Messung von Neuheit weist daraufhin, dass der entwickelte Technologieindikator durchaus ein wertvolles Instrument für das Technologie-Monitoring sein kann.

Die vorliegende Arbeit unterliegt verschiedenen Limitierungen, aus denen sich weiterer Forschungsbedarf ergibt. Die Grenzen sind (i) in der Validität der Einflussfaktoren auf die Nützlichkeit von Patentinformationen, (ii) in dem Automatisierungsgrad bei der Durchführung der semantischen Patentanalyse und (iii) in der Anwendbarkeit des hier vorliegenden Ansatzes für weitere technologieorientierte Analysen zu sehen. Damit geht einher, dass (iv) die geschickte Abgrenzung von Technologiefeldern hier nicht thematisiert wurde. Dies ist aber Voraussetzung, um (v) die hier vorgestellten

Methoden der semantischen Patentanalyse auf weitere Technologiefelder zu übertragen.

- (i) Im Hinblick auf die Nützlichkeit von Patentinformationen wurde ein Rahmenmodell entwickelt, das Einflussfaktoren auf die Nützlichkeit von Patentinformationen, insbesondere in Bezug auf den Zeitabstand zwischen Patentveröffentlichung und Markteinführung, darstellt. Dieses Rahmenmodell wurde nicht im Detail überprüft. Trotzdem kann es zukünftig ein wertvolles Instrument zur Einschätzung der Nützlichkeit von Patentinformationen in unterschiedlichen Branchen und für unterschiedliche Technologien darstellen. Gleichzeitig bietet es sich an, weitere Informationsquellen, beispielsweise wissenschaftliche Publikationen, Blogs und Foren, anhand dieses Rahmenmodells zu bewerten.
- (ii) Mit der Kombination von Multidimensionaler Skalierung und Clusteranalyse sowie der Einführung des Technologieindikators der Neuheit wurde die Ergebnisinterpretation vereinfacht. Jedoch setzt die Durchführung der Patentanalyse nach wie vor die starke Interaktion mit dem Nutzer voraus. Zwar werden Experten für die Bewertung der Ergebnisse nie ganz verzichtbar sein (Bürgel et al. 2000, Zeller 2003), jedoch besteht bei der Vorverarbeitung, der Informationserschließung und der Ergebnispräsentation die Möglichkeit, den Nutzeraufwand weiter zu mindern. So können während der Vorverarbeitung beispielsweise statistische Methoden wie das Latente Semantische Analysieren zur Erkennung von Synonymen genutzt werden (siehe hierzu z. B. Foltz et al. 1998, Landauer et al. 1998, Turney 2001). Die Interpretation der semantischen Patentlandkarten kann durch die Visualisierung von Schlüsselbegriffen in den Patentlandkarten weiter vereinfacht werden.
- (iii) Das Technologie-Monitoring stellt nur eine von zahlreichen Technologie-Analysen dar. Andere Technologie-Analysen haben die Aufgabe zukünftige

Entwicklungen zu antizipieren, sozioökonomische Einflussfaktoren zu berücksichtigen, Chancen und Risiken abzuleiten und Stakeholder am Analyse- und Planungsprozess zu beteiligen (siehe hierzu Porter, Cunningham 2005). Es ist zu überprüfen, ob das PatMining über das Technologie-Monitoring hinaus einen Beitrag für andere Technologie-Analysen leisten kann. So sind beispielsweise neue Methoden der Informationserschließung für die Zukunfts- und Trendforschung zu entwickeln.

- (iv) Die geschickte Suchraumabgrenzung bei der Patentrecherche ist nicht Bestandteil des PatMining. Die Ergebnisse in Bezug auf die Nützlichkeit von Patentinformationen haben aber gezeigt, dass die Wahl des Patentsets entscheidenden Einfluss auf die Qualität der Ergebnisse der Patentanalyse hat. Entsprechend besteht hier Nachholbedarf.
- (v) Die geschickte Suchraumabgrenzung sowie die geschickte Abgrenzung des Patentsets ist Voraussetzung für die Erprobung der hier vorgestellten Methoden und der Validierung der Ergebnisse und Einflussfaktoren in weiteren Technologiefeldern. Dies ist notwendig, da die Erprobung der Methoden sowie die Überprüfung der Nützlichkeit von Patentinformationen für das Technologie-Monitoring bisher hauptsächlich im Bereich des Maschinenbaus erfolgten.

## Literaturverzeichnis

- Ahuja, G. und Lampert, C. M. (2001): Entrepreneurship in the Large Corporation: A Longitudinal Study of how Established Firms Create Breakthrough Inventions. *Strategic Management Journal*, 22(6-7), 521-544.
- Anderson, P. und Tushman, M. L. (1990): Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change. *Administrative Science Quarterly*, 35(4), 604-633.
- Arundel, A. und Kabla, I. (1998): What Percentage of Innovations are Patented? Empirical Estimates for European Firms. *Research Policy*, 27(2), 127-141.
- Ashton, W. B. und Stacey, G. S. (1995): Technical Intelligence in Business: Understanding Technology Threats and Opportunities. *International Journal of Technology Management*, 10(1), 79-104.
- Barney, J. (1991): Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99-120.
- Behrmann, N. (1998): *Technisches Wissen aus Patenten: Eine empirische Untersuchung auf der Grundlage von Ansätzen des Wissensmanagements*. Hochschule St. Gallen für Wirtschafts-, Rechts- und Sozialwissenschaften.
- Benkard, G. (2006): *Patentgesetz, Gebrauchsmustergesetz*. 10., neubearbeitete Auflage, München: Beck.
- Bergmann, I. (2011): *Patentverletzungen in der Biotechnologie: Einsatz semantischer Patentanalysen*. Wiesbaden: Gabler Verlag.
- Bergmann, I.; Möhrle, M. G.; Erdmann, V. A.; Walter, L. und Butzke, D. (2007): The use of Semantic Maps for Recognition of Patent Infringements: A Case Study in Biotechnology, in: Ernst, H.; Gemünden, H. G.: *Open Innovation between and within Organizations*, *Zeitschrift für Betriebswirtschaft*, Special Issue 4/2007, 69-86.
- Bühl, A. (2010): *PASW 18: Einführung in die moderne Datenanalyse*. 12., aktualisierte Auflage, München u.a.: Pearson Studium.
- Bürgel, H. D.; Ackel-Zakour, R. und Zeller, A. (2000): IV-Unterstützung im Rahmen der Technologiefrühaufklärung, in: Foschiani, S.; Habenicht, W. und Schmid, U. (Hrsg.): *Strategisches Management im Zeichen von Umbruch und Wandel. Festschrift für Prof. Dr. Erich Zahn zum 60. Geburtstag*. Stuttgart: Schäffer-Poeschel, S. 269-292.
- Carley, K. M. (1997): Extracting Team Mental Models through Textual Analysis. *Journal of Organizational Behavior*, 18(S1), 533-558.
- Chandy, R. K. und Tellis, G. J. (1998): Organizing for Radical Product Innovation: The Overlooked Role of Willingness to Cannibalize. *Journal of Marketing Research*, 35(4), 474-487.

Christensen, C. M. und Overdorf, M. (2000): Meeting the Challenge of Disruptive Change. *Harvard Business Review*, 78(2), 66-77.

Comanor, W. S. und Scherer, F. M. (1969): Patent Statistics as a Measure of Technical Change. *Journal of Political Economy*, 77(3), 392-398.

Curran, C. und Leker, J. (2009): Seeing the Next Iphone Coming Your Way: How to Anticipate Converging Industries. *PICMET - Portland International Conference on Management of Engineering & Technology 2009*. Portland, Oregon.

Curran, C. und Leker, J. (2011): Patent Indicators for Monitoring Convergence – Examples from NFF and ICT. *Technological Forecasting and Social Change*, 78(2), 256-273.

Dahlin, K. B. und Behrens, D. M. (2005): When is an Invention really Radical? Defining and Measuring Technological Radicalness. *Research Policy*, 34(5), 717-737.

Debackere, K.; Verbeek, A.; Luwel, M. und Zimmermann, E. (2002): Measuring the Progress and Evolution in Science and Technology - II: The Multiple Uses of Technometric Indicators. *International Journal of Management Reviews*, 4(3), 213-231.

Dörre, J.; Gerstl, P. und Seiffert, R. (1999): Text Mining: Finding Nuggets in Mountains of Textual Data. *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data mining*. San Diego, Kalifornien.

Dressler, A. (2006): *Patente in technologieorientierten Mergers & Acquisitions: Nutzen, Prozessmodell, Entwicklung und Interpretation semantischer Patentlandkarten*. Wiesbaden: Deutscher Universitäts-Verlag.

Engelsman, E. C. und van Raan, A. F. J. (1994): A Patent-Based Cartography of Technology. *Research Policy*, 23(1), 1-26.

Ernst, H. (2003): Patent Information for Strategic Technology Management. *World Patent Information*, 25(3), 233-242.

Ernst, H. (2001): Patent Applications and Subsequent Changes of Performance: Evidence from Time-Series Cross-Section Analyses on the Firm Level. *Research Policy*, 30(1), 143-157.

European Patent Office (Hrsg.) (2012): *Annual Report 2011*. Online-Ressource. <http://www.epo.org/about-us/office/annual-report/2011.html>. Abruf: 22.04.2012.

European Patent Office (Hrsg.) (2005). *Annual Report 2004*. München.

Feinerer, I.; Hornik, K. und Meyer, D. (2008): Text Mining Infrastructure in R. *Journal of Statistical Software*, 25(5), 1-54.

Feldman, R. und Sanger, J. (2007): *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge u.a.: Cambridge University Press.

- Fleming, L. (2001): Recombinant Uncertainty in Technological Search. *Management Science*, (1), 117-132.
- Foltz, P. W.; Kintsch, W. und Landauer, T. K. (1998): The Measurement of Textual Coherence with Latent Semantic Analysis. *Discourse Processes*, 25(2-3), 285-308.
- Foster, R. N. (1986): *Innovation: The Attacker's Advantage*. London: Macmillan.
- Gansel, C. (2009): *Textlinguistik und Textgrammatik: Eine Einführung*. 3., unveränderte Auflage, Göttingen: Vandenhoeck & Ruprecht.
- Gassmann, O. und Bader, M. A. (2007): *Patentmanagement: Innovationen erfolgreich nutzen und schützen*. 2. Auflage, Berlin, Heidelberg: Springer.
- Gerken, J. M. und Möhrle, M. G. (2012): A New Instrument for Technology Monitoring: Novelty in Patents Measured by Semantic Patent Analysis. *Scientometrics*, 91(3), 645-670.
- Gerken, J. M.; Möhrle, M. G. und Walter, L. (2012): One Year Ahead! Investigating the Time Lag between Patent Publication and Market Launch: Insights from a Longitudinal Study in the Automotive Industry. *R&D Management*, under review.
- Gerken, J. M.; Möhrle, M. G. und Walter, L. (2010a): Patents as an Information Source for Product Forecasting: Insights from a Longitudinal Study in the Automotive Industry. *R&D Management Conference 2010 Proceedings*. Manchester, UK.
- Gerken, J. M.; Möhrle, M. G. und Walter, L. (2010b): Semantische Patentlandkarten zur Analyse technologischen Wandels: Eine Längsschnittstudie aus der Allradtechnik, in: Gausemeier, J. (Hrsg.): *6. Symposium für Vorausschau und Technologieplanung*. Paderborn: Heinz-Nixdorf-Institut, S. 325-349.
- Gerken, J. M.; Walter, L. und Möhrle, M. G. (2010c): *Semantische Patentlandkarten. Einsatz semantischer Patentlandkarten im Anwendungsfeld der Antriebstechnik – Eine explorative Analyse am Beispiel der Planetengetriebe*. Frankfurt/Main: VDMA.
- Geschka, H.; Schaufele, J. und Zimmer, C. (2008): Explorative Technologie-Roadmaps - Eine Methodik zur Erkundung technologischer Entwicklungslinien und Potenziale, in: Möhrle, M. G. und Isenmann, R. (Hrsg.): *Technologie-Roadmapping. Zukunftsstrategien für Technologieunternehmen*. 3., neu bearbeitete und erweiterte Auflage, Berlin, Heidelberg: Springer, S. 165-188.
- Granstrand, O. (2000): *The Economics and Management of Intellectual Property: Towards Intellectual Capitalism*. Cheltenham u.a.: Edward Elgar.
- Grawatsch, M. (2005): *TRIZ-basierte Technologiefrüherkennung*. Aachen: Shaker.
- Griliches, Z. (1990): Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28(4), 1661-1707.
- Grupp, H. (1997): *Messung und Erklärung des technischen Wandels: Grundzüge einer empirischen Innovationsökonomik*. Berlin u.a.: Springer.

- Gupta, V. K. und Pangannaya, N. B. (2000): Carbon Nanotubes: Bibliometric Analysis of Patents. *World Patent Information*, 22(3), 185-189.
- Harmann, B. (2003): *Patente als strategisches Instrument zum Management technologischer Diskontinuitäten*. Universität St. Gallen Hochschule für Wirtschafts-, Rechts- und Sozialwissenschaften.
- Haupt, R.; Kloyer, M. und Lange, M. (2007): Patent Indicators for the Technology Life Cycle Development. *Research Policy*, 36(3), 387-398.
- Hotho, A.; Nürnberger, A. und Paaß, G. (2005): A Brief Survey of Text Mining. *Machine Learning*, 20(1), 19-62.
- Kim, Y. G.; Suh, J. H. und Park, S. C. (2008): Visualization of Patent Analysis for Emerging Technology. *Expert Systems with Applications*, 34(3), 1804-1812.
- Knight, H. J. (2004): *Patent Strategy for Researchers and Research Managers*. 2. Auflage, Chichester u.a.: Wiley.
- Kruskal, J. B. (1964): Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. *Psychometrika*, 29(1), 1-27.
- Landauer, T. K.; Foltz, P. W. und Laham, D. (1998): An Introduction to Latent Semantic Analysis. *Discourse Processes*, 25(2), 259-284.
- Lee, C.; Jeon, J. und Park, Y. (2011): Monitoring Trends of Technological Changes Based on the Dynamic Patent Lattice: A Modified Formal Concept Analysis Approach. *Technological Forecasting and Social Change*, 78(4), 690-702.
- Lee, S.; Yoon, B. und Park, Y. (2009): An Approach to Discovering New Technology Opportunities: Keyword-Based Patent Map Approach. *Technovation*, 29(6-7), 481-497.
- Lichtenthaler, E. (2008): Methoden der Technologiefrüherkennung und Kriterien zu ihrer Auswahl, in: Möhrle, M. G. und Isenmann, R. (Hrsg.): *Technologie-Roadmapping. Zukunftsstrategien für Technologieunternehmen*. 3., neu bearbeitete und erweiterte Auflage, Berlin, Heidelberg: Springer, S. 59-84.
- Lichtenthaler, E. (2002): *Organisation der Technology Intelligence. Eine empirische Untersuchung der Technologiefrühaufklärung in technologieintensiven Großunternehmen*. Zürich: Verlag Industrielle Organisation.
- Lieberman, M. B. und Montgomery, D. B. (1998): First-Mover (Dis)Advantages: Retrospective and Link with the Resource-Based View. *Strategic Management Journal*, 19(12), 1111-1125.
- Mann, H. B. und Whitney, D. R. (1947): On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other. *Annals of Mathematical Statistics*, 18(1), 50-60.

- Manning, C. D.; Raghavan, P. und Schütze, H. (2008): *Introduction to Information Retrieval*. Cambridge u.a.: Cambridge University Press.
- Möhrle, M. G. und Gerken, J. M., (2012): Measuring Textual Patent Similarity on the Basis of Combined Concepts: Design Decisions and their Consequences. *Scientometrics*, doi: 10.1007/s11192-012-0682-0.
- Möhrle, M. G.; Gerken, J. M. und Walter, L. (2012): Textbasierte Patentinformationen. *Wirtschaftsstudium*, 41(1), 91-96.
- Möhrle, M. G. (2010): Measures for Textual Patent Similarities: A Guided Way to Select Appropriate Approaches. *Scientometrics*, 85(1), 95-109.
- Naunheimer, H.; Novak, W. und Ryborz, J. (2007): *Fahrzeuggetriebe: Grundlagen, Auswahl, Auslegung und Konstruktion*. Berlin u.a.: Springer.
- Peters, H. P. F. und van Raan, A. F. J. (1993): Co-Word-Based Science Maps of Chemical Engineering. Part I: Representations by Direct Multidimensional Scaling. *Research Policy*, 22(1), 23-45.
- Porter, A. L.; Roper, A. T.; Mason, T. W.; Rossini, F. A.; Banks, J. und Wiederholt, B. J. (1991): *Forecasting and Management of Technology*. New York u.a.: Wiley-Interscience.
- Porter, A. L. und Cunningham, S. W. (2005): *Tech Mining: Exploiting New Technologies for Competitive Advantage*. Hoboken, NJ: Wiley.
- Porter, M. E. (1998): *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York u.a.: Free Press.
- Reger, G. (2001): Technology Foresight in Companies: From an Indicator to a Network and Process Perspective. *Technology Analysis & Strategic Management*, 13(4), 533-554.
- Rohrdantz C. et al. ( 2010): Visuelle Textanalyse. Interaktive Exploration von semantischen Inhalten. *Informatik-Spektrum*, 33(6), 1-611.
- Schilling, M. A. (1998): Technological Lockout: An Integrative Model of the Economic and Strategic Factors Driving Technology Success and Failure. *Academy of Management Review*, 23(2), 267-284.
- Schoenmakers, W. und Duysters, G. (2010): The Technological Origins of Radical Inventions. *Research Policy*, 39(8), 1051-1059.
- Schumpeter, J. A. (1927): The Explanation of the Business Cycle. *Economica*, 21(Dec., 1927), 286-311.
- Schumpeter, J. A. (1996): *Capitalism, Socialism and Democracy*. London u.a.: Routledge.

- Schumpeter, J. A. (1939): *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*. New York, NY u.a.: McGraw-Hill.
- Shih, M.; Liu, D. und Hsu, M. (2010): Discovering Competitive Intelligence by Mining Changes in Patent Trends. *Expert Systems with Applications*, 37(4), 2882-2890.
- Sternitzke, C. (2008): *Betriebswirtschaftliche Patentportfoliobewertung: Eine informationswissenschaftliche Perspektive*. Bremen: Universität Bremen.
- Stock, W. G. (2007): *Information Retrieval: Informationen suchen und finden*. München: Oldenbourg.
- SUBARU Deutschland GmbH (Hrsg.) (2005): *33 Jahre SUBARU-Allradantrieb*. Online-Ressource. [http://www.subaru-presse.de/fileadmin/templates/downloads/awd/PressemappeSubaru\\_Allrad-Technologie\\_04-2005\\_SUB.doc](http://www.subaru-presse.de/fileadmin/templates/downloads/awd/PressemappeSubaru_Allrad-Technologie_04-2005_SUB.doc). Abruf: 15.05.2009.
- Technology Futures Analysis Methods Working Group (2004): Technology Futures Analysis: Toward Integration of the Field and New Methods. *Technological Forecasting and Social Change*, 71(3), 287-303.
- Teece, D. J.; Pisano, G. und Shuen, A. (1997): Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18(7), 509-533.
- Teichert, T. und Mittermayer, M. A. (2002): Text Mining for Technology Monitoring. *Engineering Management Conference*. Cambridge, UK.
- Trajtenberg, M.; Henderson, R. und Jaffe, A. (1997): University Versus Corporate Patents: A Window on the Basicness of Invention. *Economics of Innovation and New Technology*, 5(1), 19-50.
- Tseng, Y.; Lin, C. und Lin, Y. (2007): Text Mining Techniques for Patent Analysis. *Information Processing & Management*, 43(5), 1216-1247.
- Turney, P. D. (2001): Mining the Web for Synonyms: PMI-IR Versus LSA on TOEFL. *Lecture Notes in Computer Science*, 2167, 491-502.
- U.S. Patent and Trademark Office (Hrsg.) (2012): *U.S. Patent Statistics Chart Calendar Year 1963 - 2011*. Online-Ressource. [http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us\\_stat.htm](http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm). Abruf: 22.04.2012.
- van Eck, N. J. und Waltman, L. (2007): VOS: A New Method for Visualizing Similarities between Objects. *Advances in Data Analysis: Proceedings of the 30th Annual Conference of the German Classification Society*. Berlin.
- van Zeebroeck, N.; van Pottelsberghe de la Potterie, B. und Guellec, D. (2009): Claiming More: The Increased Voluminosity of Patent Applications and its Determinants. *Research Policy*, 38(6), 1006-1020.
- Voigt, K. und Wettengl, S. (1999): Innovationskooperation im Zeitwettbewerb, in: Engelhard, J. und Sinz, E. J. (Hrsg.): *Kooperationen im Wettbewerb*. Wiesbaden: Gabler, S. 411-443.

von Wartburg, I.; Teichert, T. und Rost, K. (2005): Inventive Progress Measured by Multi-Stage Patent Citation Analysis. *Research Policy*, 34(10), 1591-1607.

Watts, R. J. und Porter, A. L. (1997): Innovation Forecasting. *Technological Forecasting and Social Change*, 56(1), 25-47.

Wellensiek, M.; Schuh, G.; Hacker, P. A. und Saxler, J. (2011): Technologieführerkennung, in: Schuh, G. und Klappert, S. (Hrsg.): *Technologiemanagement*. Berlin, Heidelberg: Springer, S. 89-169.

Wernerfelt, B. (1984): A Resource-Based View of the Firm. *Strategic Management Journal*, 5(2), 171-180.

Willke, H. (2007): *Einführung in das systemische Wissensmanagement*. 2. Auflage, Heidelberg: Carl-Auer Verlag.

Wise, J. A. et al. (1995): Visualizing the Non-Visual: Spatial Analysis and Interaction with Information from Text Documents. *Proceedings on Information Visualization (INFOVIS '95)*. Atlanta, Georgia.

Wolf, J. (2008): *Organisation, Management, Unternehmensführung: Theorien, Praxisbeispiele und Kritik*. Wiesbaden: Gabler.

Wollschläger, D. (2010): *Grundlagen der Datenanalyse mit R: Eine anwendungsorientierte Einführung*. Heidelberg u. a.: Springer.

Yoon, B. (2008): On the Development of a Technology Intelligence Tool for Identifying Technology Opportunity. *Expert Systems with Applications*, 35(1), 124-135.

Yoon, J. und Kim, K. (2011): Identifying Rapidly Evolving Technological Trends for R&D Planning using SAO-Based Semantic Patent Networks. *Scientometrics*, 88(1), 213-228.

Yoon, B. und Park, Y. (2005): A Systematic Approach for Identifying Technology Opportunities: Keyword-Based Morphology Analysis. *Technological Forecasting and Social Change*, 72(2), 145-160.

Yoon, B. und Park, Y. (2004): A Text-Mining-Based Patent Network: Analytical Tool for High-Technology Trend. *The Journal of High Technology Management Research*, 15(1), 37-50.

Zeller, A. (2003): *Technologieführererkennung mit Data Mining: Informationsprozess-orientierter Ansatz zur Identifikation schwacher Signale*. Wiesbaden: Gabler|Deutscher Universitäts-Verlag.

Zhang, Q. und Segall, R. S. (2010): Review of Data, Text and Web Mining Software. *Kybernetes*, 39(4), 625-655.

## Anhang: Erklärung über Eigenanteil

Die hier eingereichten Publikationen stellen einen wesentlichen Bestandteil meiner wissenschaftlichen Forschungsleistung dar. Die Publikationen wurden so ausgewählt, dass sie dem Themenkomplex des Technologie-Monitorings zuzuordnen sind und ein substantieller eigener Beitrag durchgängig gegeben ist. Dies schlägt sich formal in der überwiegenden Erst-Autorenschaft nieder. Die einzige Ausnahme stellt die Publikation Möhrle und Gerken (2012) dar. Diese Publikation wurde ausgewählt, da u.a. die Anpassung der semantischen Patentanalyse für die Konvergenzanalyse und die Durchführung der Konvergenzanalyse, also die typische Aufgabe des Technologie-Monitorings, meine Aufgabe war. Die hier eingereichten Publikationen basieren maßgeblich auf den Forschungsvorhaben 560 der Forschungsvereinigung Antriebstechnik, das ich bearbeitet habe. Meine Mitwirkung bei der Erstellung der hier berücksichtigten Publikationen ist schematisch in Tabelle 2 dargestellt.

**Tabelle 2: Darstellung meiner Mitwirkung bei der Erstellung der hier berücksichtigten Publikationen**

	<b>Gerken et al. 2010a</b>	<b>Gerken et al. 2010b</b>	<b>Gerken, Möhrle 2012</b>	<b>Möhrle, Gerken 2012</b>	<b>Gerken et al. 2012</b>
<b>Ideenfindung</b>	Walter, Gerken	Gerken	Gerken	Möhrle	Möhrle
<b>Konzept-erstellung</b>	Gerken, Walter	Gerken	Gerken, Möhrle	Möhrle, Gerken	Gerken, Möhrle, Walter
<b>Theorie-entwicklung</b>	Gerken, Möhrle, Walter	Gerken, Möhrle, Walter	Gerken, Möhrle	Möhrle	Möhrle, Gerken
<b>Methoden-entwicklung</b>	Gerken, Möhrle, Walter	Gerken	Gerken, Möhrle	Möhrle, Gerken	Gerken, Möhrle, Walter
<b>Datenauswahl und -recherche</b>	Gerken	Gerken	Gerken	Möhrle, Gerken	Gerken
<b>Auswertung</b>	Gerken, Möhrle	Gerken, Möhrle	Gerken, Möhrle	Möhrle, Gerken	Gerken, Möhrle