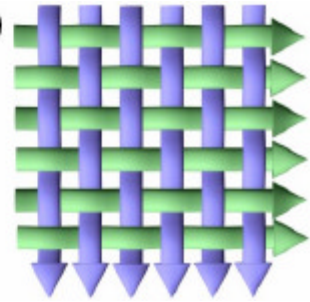


Sonderforschungsbereich 559
**Modellierung großer
Netze in der Logistik**



Technical Report 05004
ISSN 1612-1376

Automatisierte Methoden und Systeme der Datenerhebung

Teilprojekt M9:

Dirk Jodin
Andreas Mayer

Lehrstuhl FLW
Emil-Figge-Straße 73
44221 Dortmund

Dortmund, 23. Mai 2004

Inhaltsverzeichnis

1	Einleitung.....	3
2	Methoden der Datenerhebung.....	3
2.1	Klassifizierung der Datenerhebungsmethoden	4
2.2	Befragung	5
2.2.1	Mündliche Befragung	6
2.2.2	EDV Befragung	7
2.2.3	Schriftliche Befragung.....	8
2.3	Beobachtung	8
2.3.1	Manuelle Beobachtung	9
2.3.2	Automatische Beobachtung	9
2.4	Dokumentenanalyse	17
2.4.1	Optische Dokumente.....	17
2.4.2	Elektronische Dokumente	17
3	Kategorisierung der Methoden.....	18
4	Automatisierte Systeme der Datenerhebung	21
4.1	Maut System Deutschland: Toll Collect	22
4.2	Verkehrsdatenerfassung (Verkehrsvorhersage)	23
4.3	Wetterdatenerfassung	24
5	Ausblick	25

1 Einleitung

Die Methoden der Datenerhebung waren bereits Thema des Technical Report 04002-M9. Mittlerweile ist die Entwicklung insbesondere bei den EDV-gestützten Verfahren weiter gegangen und es haben weitere Untersuchungen stattgefunden, die einen ergänzenden und erweiterten Report erfordern. In der zweiten Untersuchungsphase wurden verstärkt die automatisierten Erhebungsverfahren betrachtet und bewertet. Darüber hinaus wurden anderen Disziplinen, die ebenfalls mit der Erhebung umfangreicher Datenmengen konfrontiert sind, hinsichtlich ihrer verwendeten Methoden oder Verfahren betrachtet. Dementsprechend fokussiert dieser Report auf die automatischen und EDV-gestützten Verfahren sowie die Modifikationen der Klassifizierung.

Beobachtet man die zunehmenden Möglichkeiten und die steigende Nutzung automatisierter Datenerfassung mag man kaum noch über die Probleme der Datenerhebung und -erfassung sprechen, sondern man muss vielmehr zunehmend den Blick auf das Problem der rasch steigenden Datenmengen und ihrer Verarbeitbarkeit richten. In den letzten Jahren hat es hier auf allen relevanten Bereichen entscheidende Entwicklungen gegeben, die das Datenwachstum rapide verstärken. Kosten für Rechnerleistungen und Speicherkapazitäten sind um Größenordnungen kleiner geworden. Fast alle Prozessdaten werden mittlerweile festgehalten. Kundenkarten, Kreditkarten, Mobiltelefone, Käufe im Internet usw. ermöglichen die Erfassung und Speicherung weiterer umfangreicher Daten. Durch den verstärkt zunehmenden Einsatz von RFID (Radio frequency identification) wird das aufkommende Datenvolumen noch einmal sprunghaft ansteigen. Laut Teradata-CEO Mike Köhler in den VDI-Nachrichten vom 13.5.2005 haben Studien ergeben, dass sich alle zwei bis drei Jahre die Datenmenge verdoppelt, mit der die Unternehmen konfrontiert werden [Kel05].

In Zukunft wird man sich also mehr mit Methoden und Verfahren beschäftigen müssen, die vorhandene oder entstehende Daten permanent erkennen, verknüpfen und zielorientiert aufbereiten und zur Auswertung speichern, als die Daten im Nachhinein für den Prozess der Datengewinnung zu erheben. Hier haben beispielsweise die Meteorologie, die Verkehrsprognose aber auch die Mauterfassung Systeme aufgebaut, die permanent Daten erfassen, verknüpfen und auswerten. Aus diesen Systemen zu Lernen ist eine weitere Möglichkeit, den Datengewinnungsprozess im Rahmen der Forschungen des Teilprojekts M9 weiter zu optimieren.

Unabhängig von diesen Entwicklungen ist die Teilaufgabe der Methodenklassifizierung und -kategorisierung zur Datenerhebung und -gewinnung weiterhin von hoher Wichtigkeit und Bedeutung, da insbesondere die systematische Aufbereitung den gewünschten zielorientierten Zugriff auf die richtige Methode im Rahmen des Methodennutzungsmodells [JoH05] erst ermöglicht.

2 Methoden der Datenerhebung

Die Methoden der Datenerhebung werden klassischerweise in primäre und sekundäre Methoden unterteilt. Unterscheidungsmerkmal ist hier das Vorhandensein der benötigten Informationen zum Zeitpunkt der Datenerhebung.

Liegen die benötigten Daten nicht in dokumentierter Form vor und müssen demzufolge gesondert erhoben werden, so sind die zur Informations- und Datenerhebung anzuwendenden Techniken und Methoden den Primärdatenerhebungsmethoden zuzuordnen [Voß04].

Können im Gegensatz dazu die für den jeweiligen Untersuchungszweck benötigten Daten aus bereits für andere unternehmensinterne und -externe Zwecke erstellten Daten extra-

hiert werden, so sind die hierfür notwendigen Erhebungstechniken und -methoden den Sekundärdatenerhebungsmethoden zuzuteilen [Voß04].

In der Literatur wird teilweise noch die Tertiärdatenerhebung angeführt, welche einen Sonderfall der Sekundärerhebung darstellt. Werden für eine Sekundärerhebung keine Rohdaten, sondern vorverarbeitete, wie z.B. aggregierte oder komprimierte Daten genutzt, so ist diese als tertiärstatistische Datenerhebung zu bezeichnen. In der Regel sind Rohdaten gegenüber stark aggregierten oder komprimierten Daten zu bevorzugen, da die Aussagekraft von Schlussfolgerungen aus tertiärstatistischen Erhebungen begrenzt ist [Har99]. Aus diesem Grunde wird die Tertiärdatenerhebung hier nicht weiter berücksichtigt.

Statistisch ist weiterhin zwischen einer Teil- und einer Vollerhebung von Daten zu unterscheiden. Die Datenerhebung für die relevanten Merkmale, Kenngrößen oder Variablen muss je nach Anwendungsfall auf einen Teil der insgesamt verfügbaren realen Sachverhalte, Objekte bzw. Vorgänge beschränkt bleiben. Dieses ist als Stichprobe aus der insgesamt verfügbaren Grundgesamtheit von Sachverhalten/Objekten/Vorgängen bzw. als Teilerhebung zu bezeichnen [Sch03].

Im Gegensatz dazu steht die Vollerhebung, bei der die interessierenden Merkmale für sämtliche Einheiten der Grundgesamtheit beobachtet werden. Im Bereich der Sozialforschungen sind hier Anwendungen wie die Volkszählungen zu vermerken, bei denen trotz des sehr hohen Aufwandes eine Datenerhebung von sämtlichen Einwohnern eines Landes durchgeführt wird [Sch03] und [Har99]. Dabei ist allerdings zu beachten, dass auch Vollerhebungen trotz des hohen Aufwandes nur theoretisch fehlerfrei sind [Har99].

Ein typisches Beispiel für Teilerhebungen ist die Qualitätskontrolle von Produkten, bei der eine Vollerhebung häufig aus Kostengründen bzw. wegen des hohen zeitlichen Aufwandes nicht sinnvoll ist. Zu beachten sind hier allerdings Produkte mit besonderen Sicherheitsanforderungen, z.B. Luft- und Raumfahrttechnologien, bei denen ungeachtet des Aufwandes oftmals auf eine Vollerhebung zurückgegriffen wird.

2.1 Klassifizierung der Datenerhebungsmethoden

Im Hinblick auf die vorgesehene Kategorisierung der Methoden erwies es sich als Zielführender, in der Klassifizierung zunächst möglichst differenziert hinsichtlich der Kriterien und Eigenschaften der Methoden zu systematisieren, als bereits in den oberen Ebenen die konkreten Methoden und Verfahren zu nennen. Demzufolge wurde die bisherige Systematik [HJL04, Abbildung 2] nicht nur erweitert, sondern in ihrem Aufbau modifiziert.

Betrachtet man zunächst lediglich die obersten drei Ebenen der neu entwickelten Klassifizierung, lässt sich erkennen, dass die dort stehenden Kriterien bereits eine gute Kategorisierung beispielsweise in elektronisch/automatische und manuelle Verfahren erwarten lässt.

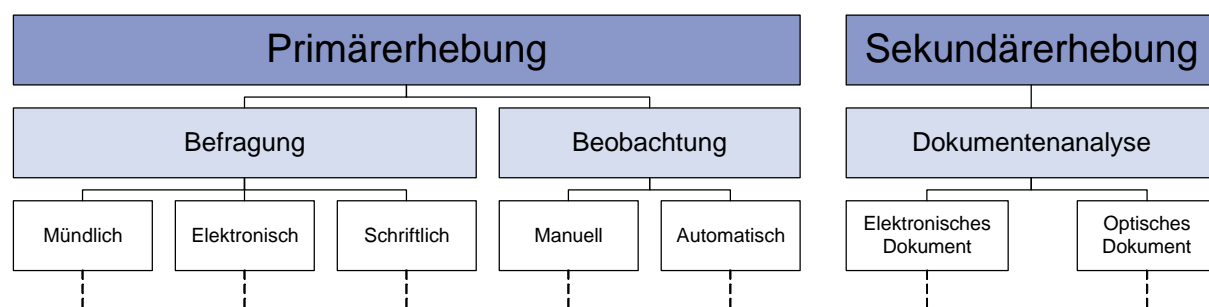


Abbildung 1: Obere Ebenen der Methodenklassifizierung

Entsprechend dieses Leitziels der verbesserten Kategorisierbarkeit sind auch die darunter liegenden Zweige der Methodenklassifizierung aufgebaut. Ein weitere Prämisse war es spezifische Methoden, die sich nicht weiter differenzieren lassen außerhalb der Klassifizierung

aufzuführen und lediglich einem Zweig zuzuordnen um bei einer neu entdeckten oder entwickelten Methode nicht gleich die Klassifizierung ändern zu müssen. Aus Gründen der Übersichtlichkeit werden die Zweige Primär- und Sekundärerhebung getrennt weiter dargestellt.

Die Methoden der *primären Datenerhebung* lassen sich in Methoden der Befragung und der Beobachtung untergliedern. Hauptunterscheidungsmerkmal ist die vorhandene oder nicht vorhandene Interaktion mit dem Untersuchungsobjekt. Diese Unterscheidung ist nicht ohne Bedeutung, da eine Interaktion auf der einen Seite eine Beeinflussung des Befragten bedeuten kann, auf der anderen Seite die Beobachtung ohne Interaktion möglicherweise zu Fehlinterpretationen führen kann. So entscheidet sich hier bereits das Potenzial für Erhebungsfehler.

2.2 Befragung

Ein Hauptanwendungsfeld der Erhebungstechnik „Befragung“ bildet die empirische Sozialforschung, in der Informationen über Meinungen, Einstellungen und Gefühle das Erhebungsziel darstellen. Weitere Anwendung findet diese Technik in der Untersuchung von Organisationen, Strukturen und Abläufen. Mittels einer Befragung lassen sich innerhalb von Organisationen Informationen über Prozesse, Abläufe, Datenflüsse und komplexe organisatorische Sachverhalte erheben.

Die Datenerhebungsmethode „Befragung“ setzt voraus, dass der Befragte über Informationen verfügt, die von Interesse für die Untersuchung sind. Bei der Befragung ist zwischen schriftlichen (Fragebogen), mündlichen (Interview) und EDV-gestützten Befragungen zu unterscheiden. Abbildung 2 zeigt die verschiedenen Befragungsmöglichkeiten.

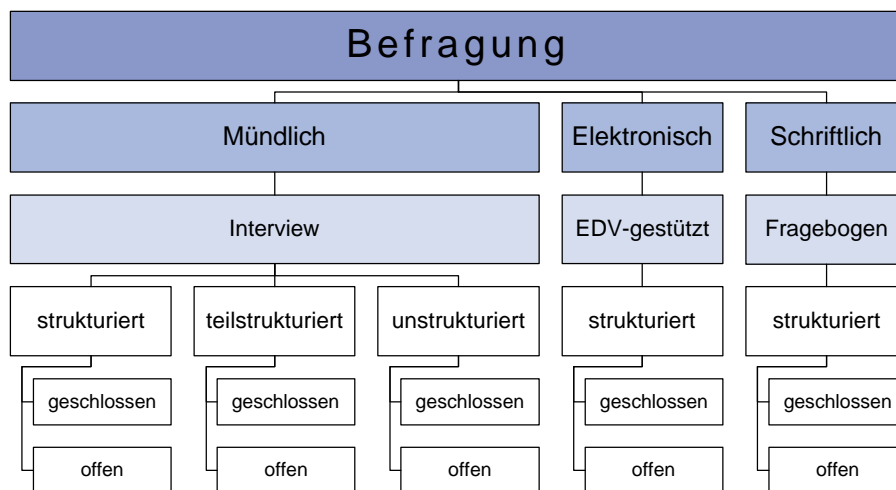


Abbildung 2: Methoden der Befragung

Es ergeben sich drei Varianten der Befragung: mündlich, elektronisch und schriftlich, die jeweils direkt mit einem übergreifenden Verfahren verbunden sind, die allerdings jeweils verschiedene Variationsmöglichkeiten haben und so zu speziellen Methoden der Befragung führen. Die Variationsmöglichkeiten unterscheiden sich zunächst hinsichtlich der Kriterien „strukturiert“, „teilstrukturiert“ und „unstrukturiert“ und jeweils in eine offene und geschlossene Variante. Diese Kriterien finden sich an verschiedenen Stellen der Klassifizierung wieder und haben für die spätere Kategorisierung eine wesentliche Bedeutung. Um nicht an den verschiedenen Stellen diese Begriffe immer wieder neu in dem jeweiligen Sinnzusammenhang definieren zu müssen, sind sie exemplarisch für die mündliche Befragung ausführlich erläutert und können so in den anderen Zweigen der Klassifizierung entsprechend interpretiert werden.

2.2.1 Mündliche Befragung

Die mündliche Befragung ist fest mit dem Begriff des Interviews verbunden, welches bereits in vielen Varianten bekannt ist. Bei der Erhebungstechnik „Interview“ handelt es sich um eine gezielte mündliche Befragung von Personen. Das Interview wird durch persönlichen Kontakt mit dem Befragten durchgeführt. Dabei versucht der Interviewer durch ein planmäßiges Vorgehen relevante Informationen für den Untersuchungszweck zu erheben. Anhand der Freiheiten des Interviewers bei der Fragestellung sind das strukturierte Interview, das teilstrukturierte Interview und das unstrukturierte Interview zu differenzieren.

Beim *strukturierten Interview* ist ein Fragebogen vorhanden, in dem Fragen vor der Befragung vollständig ausgearbeitet und in der Reihenfolge festgelegt werden. Die Fragen werden vom Interviewer in der vorgegebenen Reihenfolge wörtlich vorgelesen. Erklärungen bei der Durchführung des Interviews sind nicht zulässig. Anpassungen der Fragestellung an den Gesprächsverlauf, an den Befragten oder an die Funktion des Befragten darf der Interviewer nicht vornehmen. Das strukturierte Interview entspricht somit weitgehend einer Fragebogentechnik; das Unterscheidungsmerkmal liegt im Vortragen der Fragen und dem Protokollieren der Antworten durch den Interviewer. Als Nutzungsvoraussetzung beider Befragungstechniken gilt, dass die Formulierungen, der Inhalt und die Reihenfolge der Fragestellung für alle Interviewten identisch sind. Wesentlicher Vorteil des strukturierten Interviews ist die Erleichterung der Arbeit des Interviews durch präzise Vorgabe von Fragen. Dabei sind explizite Kenntnisse über den Untersuchungsgegenstand seitens des Interviewers nicht notwendig. Weiter wird durch fest vorgeschriebene Fragen eine bewusste oder unbewusste Manipulation durch suggestive Fragestellungen, Auslassen und Übertreiben verhindert. Anwendungsmöglichkeiten für standardisierte Interviews sind hauptsächlich die Erfassung von quantitativen, bereits bekannten organisatorischen Aspekten.

Dem *teilstrukturierten Interview* liegt ein flexibel aufgebautes Fragenraster zugrunde. Diesem folgt der Interviewer nach eigenem Ermessen mit individuellen Formulierungen. Der jeweilige Gesprächsverlauf bestimmt dabei die Reihenfolge der Fragen. Der Fragende muss mehr mit der Materie betraut sein als beim strukturierten Interview, da er die Antworten und daraus resultierende logische Verzweigungen selbständig im Interviewverlauf interpretieren muss.

Bei einem *unstrukturierten Interview* benutzt der Interviewer lediglich einen Interviewleitfaden. Dieser enthält stichwortartige Merkhilfen für den Interviewer. Die Reihenfolge und die Formulierung der Fragen liegen im Ermessen des Interviewers und werden gegebenenfalls dem Verlauf des Interviews angepasst. Aufgrund der nicht vorgegebenen Struktur erfordert diese Befragungsform einen entsprechend erfahrenen und möglichst mit der zu untersuchenden Thematik vertrauten Interviewer. Unklarheiten werden durch zusätzliche Fragen und Erklärungen beseitigt. Dabei entstehen jedoch Möglichkeiten zur Manipulation [Büh99].

Der Technical Report 04002 gibt in Tabelle 2 [HJL04, S. 10] einen Überblick über die verschiedenen Interviewformen. In der betrieblichen Praxis wird oftmals das unstrukturierte Interview auf der Basis eines Interviewleitfadens angewandt. Das teilstrukturierte Interview wird nur gelegentlich durchgeführt. In Fällen mit vergleichsweise einfach aufgebauten Fragenkatalogen wie zum Beispiel der Marktforschung in der Konsumgüterindustrie, Umfragen zur Kundenzufriedenheit, politischen Umfragen und Ähnlichem findet das strukturierte Interview Anwendung.

Des Weiteren ist zwischen einer offenen und einer geschlossenen Interviewtechnik zu unterscheiden. Während die offene Technik eine freie Antwort erlaubt, sind bei einem geschlossenen Interview die Äußerungen des Befragten auf vorgegebene Antworten beschränkt. Dabei wird oftmals eine Skala („Wie würden Sie unsere Leistung auf der Schulnotenskala bewerten?“) oder ein Bewertungsschema („Sind Sie mit unserer Leistung: sehr zufrieden; zufrieden oder unzufrieden?“) vorgegeben. Im Gegensatz dazu zwingt die offene Interviewtechnik den Befragten, eigenständig über die Antwort nachzudenken und in eige-

nen Worten zu formulieren. Dies ist insbesondere bei komplexeren Themengebieten sinnvoll, wogegen sich für die Erhebung von quantitativen Größen die geschlossene Interviewtechnik anbietet.

Die aus der Befragung gewonnenen Informationen werden in der Regel sofort vom Interviewer protokolliert. Dieses Protokoll sollte dem Interviewpartner vorgelegt und bei eventuellen Unstimmigkeiten oder Missverständnissen entsprechend berichtigt werden. Ansonsten wurde die Interviewmethode in dem Technical Report 04002 hinreichend beschrieben [HJL04, S.7-11], so dass hier auf eine weitere Vertiefung verzichtet wird.

2.2.2 EDV Befragung

Die EDV-Befragung stellt eine durch die technische Entwicklung der vergangenen Jahre mittlerweile etablierte Befragung dar. Insbesondere die im Gegensatz zur konventionellen schriftlichen Befragung gegebenen Multimedia- und Hypertextmöglichkeiten sowie die Möglichkeiten einer flexiblen interaktiven Programmierung, ergeben weit reichende Möglichkeiten dieser Datenerhebungsform.

Prinzipiell stellt jede Eingabe in ein EDV-Formular, sei es für eine Bestellung im Internet, für den Zugang zu geschützten Webseiten oder zur Anforderung von Informationsmaterialien bereits eine elektronische Befragung dar. Aber auch die vielfältigen Fragen zu Finanzierungs- oder Aufnahmeanträgen können über ein strukturiertes Formular in der EDV erfasst werden. Ein weiteres Beispiel ist die elektronische Steuererklärung (ELSTER). Wer sich hiermit noch nicht auseinandergesetzt hat darf sich über die auf der Internetseite präsentierten Zahlen wundern, wonach bislang bereits 5.447.000 Einkommensteuererklärungen und 21.980.000 Umsatzsteuer-Voranmeldungen sowie 24.180.000 Lohnsteuer-Anmeldungen elektronisch abgegeben wurden [Els05].

Die Befragung muss nicht im Internet erfolgen, sie kann darüber hinaus auf Datenträgern oder per Datenübertragung übermittelt und retourniert werden, weiterhin werden beispielsweise bei der Messemarktforschung spezielle Rechner aufgestellt, an denen die Befragten ihre Antworten interaktiv eingeben können.

Die Techniken beschränken sich hierbei nicht auf mit Feldern vordefinierte Word- oder PDF-Dokumente, sondern es handelt sich in der Regel um intelligente Web-Schnittstellen komplexer Datenbanken, mit direkter Plausibilitätsprüfung und kontextbezogener Verzweigung. Ein weiteres komplexes Feld der elektronischen Befragung sind Prüfungen. Beispielsweise existiert der allgemeine englische Sprachtest TOEFL in eine paper based Version und einer computer based Version, bei der die verschiedenen Testteile per EDV abgeprüft werden [Man05].

Prinzipiell gelten die gleichen Voraussetzungen und Bedingungen wie für die schriftliche Befragung. Die Befragung ist jedoch nur strukturiert möglich. Neuere Entwicklungen sind so genannte adaptive Befragungen, bei denen entsprechend der gegebenen Antworten die folgenden Fragen ausgewählt werden. Der angesprochene TOEFL-Test ist ein adaptiver Test, er passt sich automatisch dem Niveau des Kandidaten an. „Das funktioniert so: Der Test beginnt mit einer Frage von durchschnittlicher Schwierigkeit. Wenn der Kandidat die Frage richtig beantwortet, wird seine Punktezah erhöht und der Computer wählt als nächste Frage eine etwas schwerere aus seinem Pool aus. Wenn eine Frage falsch beantwortet wird, wird die Punktezah reduziert und die nächste Frage ist etwas leichter.“ [Man05]. Ein weiteres Beispiel ist der GRE (Graduate Record Examination) Test der amerikanischen Universitäten. Hier beeinflusst die Antwort auf eine Frage ebenfalls direkt den Schwierigkeitsgrad der folgenden Fragen und damit die erreichbare Punktzahl. Trotz dieser Flexibilität sind die Befragungen derzeit noch streng strukturiert.

2.2.3 Schriftliche Befragung

Die schriftliche Befragung wird generell in Form eines Fragebogens durchgeführt. Fragebögen werden zur Erhebung in der Situationsanalyse und eventuell zur Zielformulierung eingesetzt [Sch99]. Im Gegensatz zum Interview findet hier keine Interaktion zwischen dem Befragten und einem Interviewer statt. Die Fragen werden dem Befragten vielmehr schriftlich gestellt, worauf dieser seine Antworten formuliert und auf dem Fragebogen festhält. Diese Form der Datenbeschaffung wird in der Literatur als kostengünstigste Form der Datenbeschaffung bezeichnet, allerdings bedingen die oftmals niedrigen Rücklaufquoten eine deutlich geringere Antwortquote als bei der mündlichen Befragung [Voß04].

Die Frageform gleicht den Fragen im strukturierten Interview, welches im vorherigen Kapitel dargestellt wurde. Unterscheidungsmerkmal ist hier das eigenständige Bearbeiten der Fragen ohne mit einem Interviewer zu interagieren. Sobald eine Person den Befragten bei der Beantwortung unterstützt, ist von einem strukturierten Interview zu sprechen, da eine Manipulation auch ungewollt nicht ausgeschlossen werden kann.

Da Rückfragen in der Regel nicht möglich sind, sind die Anforderungen an die Präzision der Fragen höher als im Interview. Die Eindeutigkeit und Verständlichkeit der Fragestellung sowie die klare und vollständige Möglichkeit der Informationsgewinnung aus den Antworten sollte deshalb mit Hilfe einer Probebefragung (Pretest) überprüft werden.

Die Unterscheidung in offene und geschlossene Fragestellung ist sinngemäß auch für die schriftliche Befragung von Bedeutung. Für weitere Details sei ebenfalls auf den Technical Report 04002 verwiesen [HJL04, S. 11-14].

2.3 Beobachtung

Die Aufnahme und Interpretation sinnlich wahrnehmbarer Tatbestände im Rahmen einer Situationsanalyse werden in der Erhebungstechnik der Beobachtung durchgeführt. Im Gegensatz zur bisher dargestellten Befragung findet bei der Beobachtung keine Interaktion zwischen dem beobachteten Gegenstand, Vorgang oder Menschen und dem beobachtenden Mittel statt. Somit hat die Beobachtung als Form der Datenbeschaffung für den Analytiker den Vorteil, nicht von Rücklaufquoten und der korrekten Arbeitsweise der Interviewer abhängig zu sein [Voß04].

Sie findet manuell als Selbstbeobachtung oder Fremdbeobachtung, letztere direkt oder indirekt, ihre Anwendung. Hinzu kommen automatische Beobachtungsverfahren, die vor allem in den letzten Jahren an Bedeutung gewonnen haben, da die technische Entwicklung nutzbarer Systeme weit vorangeschritten ist. Abbildung 3 zeigt die gefundene Klassifizierung, die den Schwerpunkt stärker auf die automatischen Beobachtungsverfahren legt und die manuellen weniger strukturiert.

Innerhalb der unterhalb einzelner Zweige dargestellten gestrichelten Rechtecke werden verschiedene Verfahren oder Systeme aufgelistet, um die entsprechende Methode zu realisieren. Die Verfahren sind bewusst nicht mit der Klassifizierung verbunden, da zum einen eine eindeutige Zuordnung nicht immer möglich ist und zum anderen die technische Entwicklung hier zu einer raschen Überarbeitungsnotwendigkeit führen könnte.

So kann beispielsweise eine strukturierte Fremdbeobachtung durch eine Zeitaufnahme, eine Multimomentstudie, sowie durch Zählen oder Messen erfolgen. Sollten sich zukünftig weitere Verfahren ergeben oder entwickelt werden, können diese in die entsprechenden Zweige der Klassifizierung eingeordnet werden. Die Klassifizierung an sich bleibt aber erhalten.

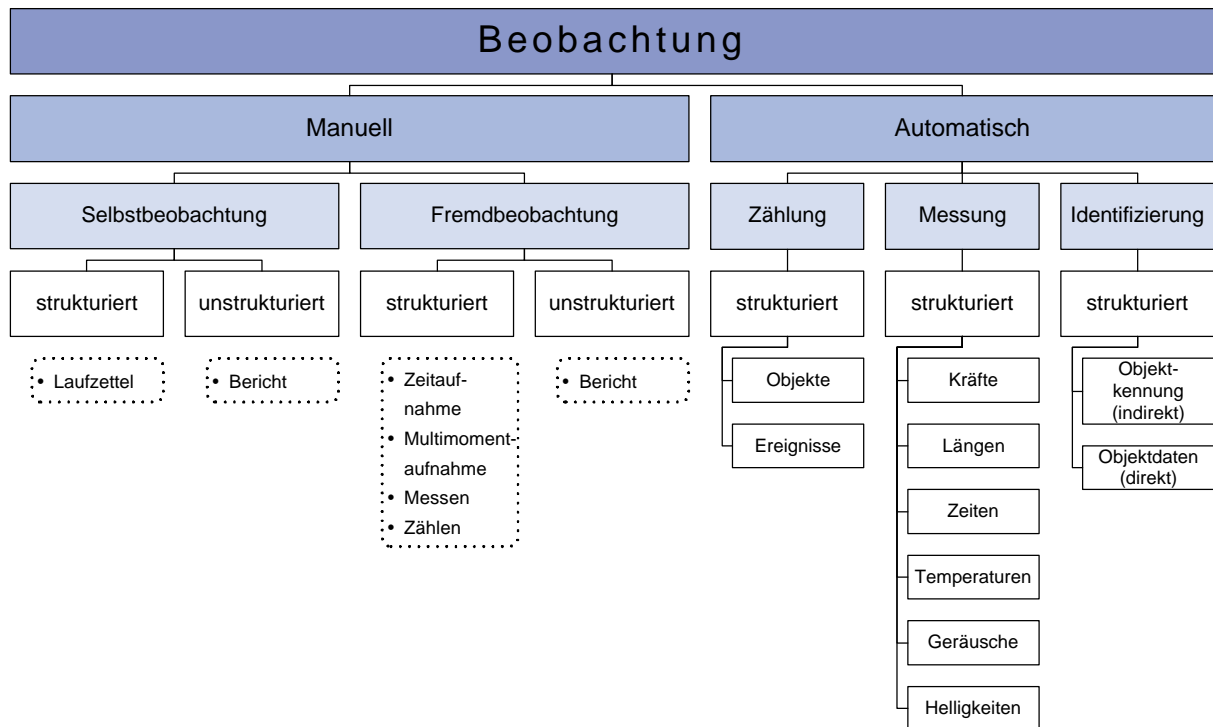


Abbildung 3: Methoden der Beobachtung

2.3.1 Manuelle Beobachtung

Die Verfahren der manuellen Beobachtung wurden im Technical Report 02004 umfassend beschrieben [HJL05, S.14-22], so dass hier auf eine weitere Darstellung verzichtet wird.

2.3.2 Automatische Beobachtung

Die Automatische Beobachtung lässt sich in die Bereiche Zählen, Messen und Identifizieren unterscheiden.

2.3.2.1 Automatisches Zählen

Automatisch zählbar sind entweder *Objekte* wie beispielsweise Paletten, Fahrzeuge und Personen oder es können *Ereignisse* erfasst und gezählt werden, wie z. B. Messwertüberschreitungen, Auslösen von Endschaltern, Füllung von Puffern, usw. Der Vielzahl der zählbaren Objekte und Ereignisse sind kaum Grenzen gesetzt. Die hierzu notwendigen Sensoren und Messsysteme werden immer kleiner, preiswerter leistungsfähiger und intelligenter, so dass hier eine weitere starke Zunahme erfasster Daten zu erwarten ist. Netzwerktechnologien und Funktechniken wie Wireless Lan oder Bluetooth ermöglichen die Positionierung der Sensoren an nahezu beliebigen Stellen und deren Einbindung in übergeordnete Mess- und Erfassungssysteme.

Als Sensoren werden alle Bauelemente bezeichnet, die eine von einem Ereignis oder einem Zustand beeinflusste physikalische Messgröße aufnehmen und in eine verarbeitbare Signalgröße umwandeln, die in einem direkten Zusammenhang zur Ursache steht [JüB98]. Ein Kriterium für den Einsatz eines Sensors ist die Messgrößendimension. Sie reicht von 0-dimensionalen Sensoren zur Anwesenheitskontrolle bis zu 3-dimensionalen Sensoren, die eine räumliche Situation erfassen können. Für die Zählfunktion sind somit prinzipiell die 0-dimensionalen Sensoren ausreichend. Die folgende Systematik zeigt auszugsweise typische Sensoren für diese Messaufgaben.

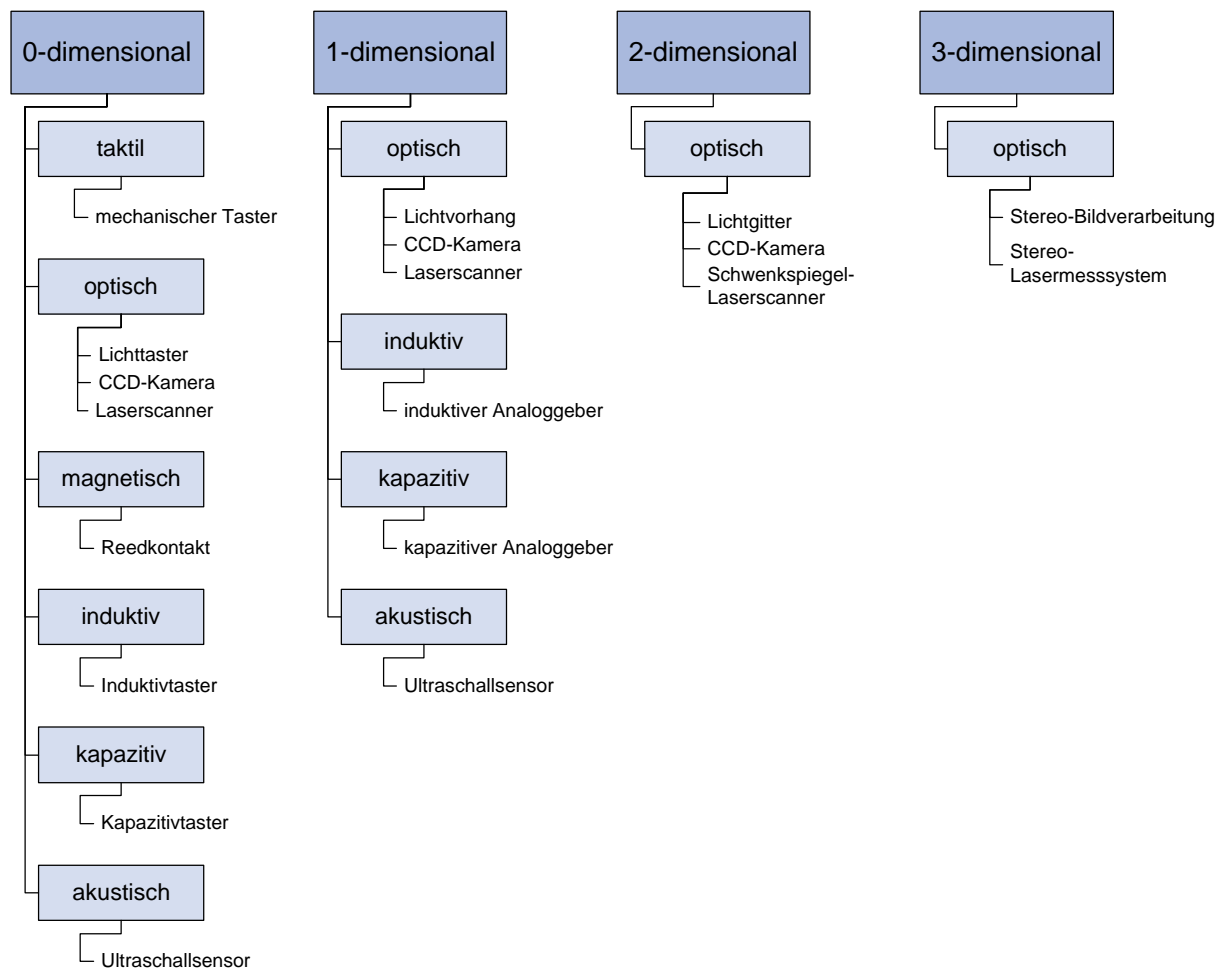


Abbildung 4: Sensoren zur Lagedetektion (nach [JüB98, Bild 8.2-2])

Für eine weitergehende und detaillierte technische Funktionsbeschreibung der aufgelisteten Sensoren sei auf die spezielle Fachliteratur [JüB98] verwiesen.

Unabhängig vom eingesetzten Sensor reicht für das Zählen als elektrisches Signal ein Bit aus, das die Anwesenheit oder Abwesenheit des Objektes oder Ereignisses repräsentiert.

2.3.2.2 Automatisches Messen

Beim Messen wird als Signal ein Messwert erwartet, der ja nach gewünschter Auflösung in der Regel mehrere Bit Datenlänge hat.

Man unterscheidet das Messen diskreter und kontinuierlicher sowie analoger und digitaler Signale. Diskrete Signale werden zu einem gewünschten Zeitpunkt als Einzelwert, kontinuierliche Signale werden über einen längeren Zeitraum in der Regel mit Schwankungen und Veränderungen erfasst.

Analoge Signale bilden die Werte einer Größe stetig ab. Die Amplitude ist ein Maß für den Wert der Größe, der zeitliche Verlauf zeigt die Änderungen. Die Werte lassen sich eigentlich nur optisch speichern, beispielsweise mit einem Lichtschreiber, einem Plotter oder einem Drucker. Zur digitalen Speicherung müssen die Daten digitalisiert werden. Das Analogsignal wird mit einer möglichst hohen Frequenz und Auflösung abgetastet und in einen diskreten digitalen Datenstrom umgewandelt. Die technisch mathematischen Grundlagen gehen auf die Arbeiten des Mathematikers Claude E. Shannon zurück, in denen das so genannte Abtasttheorem formuliert wurde, welches unter anderem zwei wesentliche Bedingungen beinhaltet.

- Die Abtastfrequenz, in der die Abtastung des Analogsignals erfolgt, muss mindestens den doppelten Wert der höchsten im abzutastenden Signal enthaltenen Frequenz haben.
- Im Originalsignal darf keine Frequenz vorhanden sein, die größer ist als die Hälfte der Abtastfrequenz.

In der Praxis haben sich Abtastfrequenzen von der 2,2-fachen bis zur 2,56-fachen Originalfrequenz etabliert. So werden Audio-CDs mit einer Abtastfrequenz von 44,1 kHz abgetastet, so dass Frequenzen bis 20 kHz digitalisiert werden können [Wik05].

Digitale Signale bilden die Werte einer stetigen oder unstetigen Messgröße entsprechend als eine Folge von diskreten Ganzzahlwerten ab.

Messbar ist mittlerweile nahezu alles. Demnach ist die dargestellte Klassifizierung an dieser Stelle nur rudimentär, und soll primär die für die Messung von Logistikdaten relevanten Messgrößen auflisten. Sofern für einen Logistikprozess beispielsweise das Messen von Radioaktivität von Bedeutung sein sollte, ist dieses Verfahren ebenso unschädlich an dieser Stelle hinzuzufügen, sofern es sich um eine automatische Messung handelt, die ein Mehrbitdatum liefert. Häufig benötigte Messgrößen in der Logistik sind Gewichte (Kräfte), Abmessungen (Längen), Zeiten und Geschwindigkeiten. Weitere häufig benötigte Größen sind Farben (z.B. als Sortierkriterium) und Temperaturen (z.B. Überwachung der Kühlkette).

Der dritte Bereich der automatischen Beobachtung ist die automatische Identifizierung, die auch als Auto-ID bezeichnet wird. Zum einen werden entweder Objektkennungen identifiziert, vom eindimensionalen Barcode über mehrdimensionale Stapelcodes bis hin zu elektronischen Etiketten. Zum anderen werden aber auch Objektdaten wie Form, Größe oder nichtmaschinenlesbare Beschriftungen usw. automatisch identifiziert.

2.3.2.3 Automatische Identifizierung

Die Identifizierung ist ein Vorgang, der zum eindeutigen Erkennen eines Lebewesens oder eines Objekts dient. Sie erfolgt anhand klassifizierender Merkmale wie Größe, Haarfarbe, Stimme, Aussehen usw.. Das Gehirn teilt die Wahrnehmungen auf der Basis komplexer Informationen ein, die im Gedächtnis verfügbar sind.

Verfügen die zu identifizierenden Lebewesen oder Objekte nicht über genügend unterscheidbare oder bekannte Kriterien, dann benötigen wir zusätzliche Informationen, die in der Regel künstlich hinzugefügt wurden. Einen Personalausweis, eine Hundemarke oder ein Etikett.

Bei der automatischen Identifikation (Auto-ID) funktioniert der Prozess des Erkennens vorrangig auf Basis anderer Merkmale als der direkten Wahrnehmung von Form und Farbe eines Gegenstandes, da die für eine schnelle Erkennung benötigten, natürlichen Merkmale meist nicht ausreichend eindeutig sind und zu viel Rechen- und Zeitaufwand bedeuten.

So existieren für die automatische Identifizierung besondere Verfahren zur künstlichen Kennung wie beispielsweise:

- Chipkarten für Personen
- Barcodes und Smart-Label für Objekte
- Implantierbare elektronische Etiketten für Tiere
- Barcodes für Dokumente

Generell differenzieren die automatischen Identifizierungssysteme in die *direkte Identifizierung* anhand natürlicher Merkmale und die *indirekte Identifizierung* anhand künstlicher Merkmale.

Die verfügbaren Identifizierungssysteme sind leistungsfähig und vielfältig und sind aufgrund ihrer spezifischen Eigenschaften jeweils für individuelle Einsatzfälle prädestiniert. Es lassen

sich Systeme unterscheiden, die in der Lage sind, mechanische, magnetische, optische oder elektronische Codierungen zu lesen.

Mechanische Codierungen sind in der Regel preiswert, einfach zu handhaben und robust. Das Lesen kann beispielsweise mit einem mechanischen oder induktiven Schalter sowie mit einer optischen Lichtschranke erfolgen

Magnetische Codierungen sind in der Regel unempfindlich gegenüber Verschmutzungen und besitzen eine hohe Lesesicherheit. Die Magnetkarte kann eine große Datenmenge aufnehmen. Nachteilig ist der relativ hohe Preis.

Optische Codierungen, speziell Barcodes, besitzen gegenwärtig durch die preiswerten Datenträger eine hohe Lesesicherheit der Codierungen und die größte Verbreitung.

Elektronische Codierungen benötigen einen technisch aufwändigen und kostenintensiven Datenträger. Aufgrund seiner gegenüber den anderen Codes erweiterten Möglichkeiten nimmt seine Verbreitung als RFID aber stetig zu.

Obwohl in der einschlägigen Fachliteratur ausführlich und detailliert über Barcode- und RFID-Technologien berichtet wird, sollen im Folgenden beide Technologien hier kurz dargestellt werden.

2.3.2.4 Barcode

Barcode-Identifizierung ist die heute am weitesten verbreitete Methode der automatischen Identifizierung. 1974 wurde erstmals ein Strichcode von einem Scanner in einem Supermarkt in Ohio, USA eingelesen: Auf einer Packung Kaugummi der Marke Wrigleys [Wik05a].

Der Strich- oder Barcode ist eine maschinenlesbare Schrift, die aus verschiedenen breiten Strichen und Lücken besteht. Sie kann über optische Lesegeräte (Barcodelesegerät oder *Scanner*) maschinell gelesen und weiterverarbeitet werden.

Konventionelle Barcodes sind eindimensional aufgebaut. Sie werden von links nach rechts oder umgekehrt gelesen. Die Höhe des Symbols hat keinen Bezug auf den Inhalt. Über eine Umsetztabelle wird der Dateninhalt in einen vom Menschen lesbaren Zeichenvorrat umgewandelt.

Der Barcode enthält keine beschreibenden Daten, sondern je nach Barcodeart eine unterschiedliche Anzahl von Ziffern oder Zeichen. Zum Barcodefeld gehören zwei Ruhezone vor und nach der Strichcodierung sowie eine Klarschriftzeile darunter. Nachfolgend eine Übersicht der heute am meisten verbreiteten Barcodearten und deren Einsatzbereiche.

Codegruppen	Produktion	Paketdienste	Lebensmittel Branche	Einzelhandel	Elektronik Industrie	Automobile Industrie	Pharma Industrie	Transport Logistik
1D Codes	Ja	Ja	Ja	Ja	Ja	Odette	Ja	Ja
2/5 Interleaved	v	v	v	-	v	-	v	v
Code 39	v	v	-	-	v	v	-	v
Code 128	v	v	-	-	v	-	v	v
EAN 128	v	v	v	v	v	-	-	v
EAN	-	-	v	v	v	-	-	v
RSS	-	-	v	v	v	-	v	-
2D Codes	Ja	Ja	Nein	Nein	Ja	Nein	Ja	Ja
Stapelcode	Nein	Nein	-	-	Nein	-	Ja	Ja
Code 16K	-	-	-	-	-	-	-	-
Code 49	-	-	-	-	-	-	-	-
Codablock	-	-	-	-	-	-	v	-
PDF 417	-	-	-	-	-	-	-	v
Matrix Codes	Ja	Ja	-	-	Ja	-	Ja	Ja
Data Matrix	v	-	-	-	v	-	v	-
Maxi Code	-	v	-	-	-	-	-	v

Abbildung 5: Verbreitete Barcodearten (nach [Dat00])

Weitere Informationen zu den Barcodes sind in einschlägigen Veröffentlichungen [Len00], [Len04], [HoS03], [Dat00] oder im Internet [Act05] zu finden.

Bei den Barcodetypen wird zwischen numerischen und alphanumerischen Typen unterschieden. Je nach Anwendung kann ein Nummernsystem ausreichen oder es können Klarschriftinformationen notwendig sein.

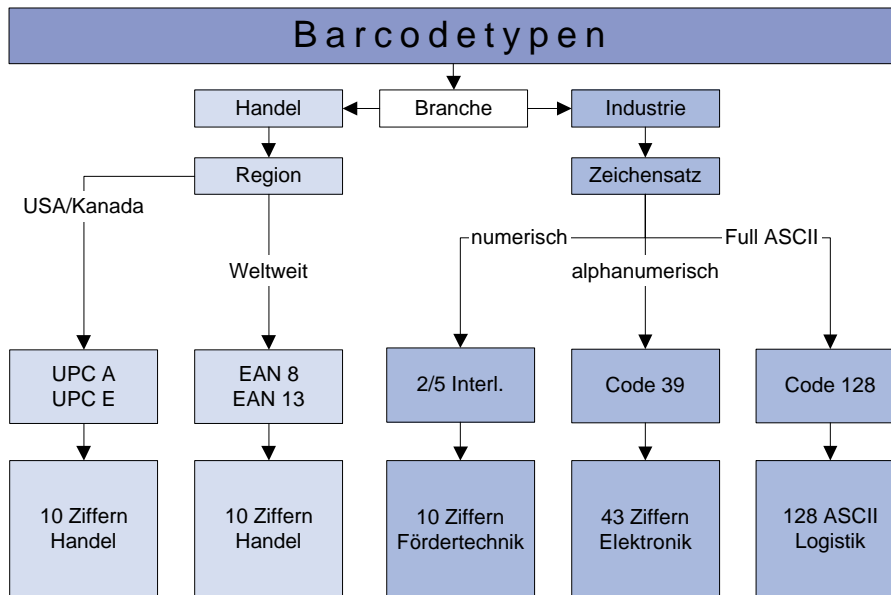


Abbildung 6: Barcodetypen (nach [Len00, S. 113])

UPC steht für **U**niversal **P**roduct **C**ode, EAN für **E**uropäische **A**rtikel **N**ummer und Interl. ist die Abkürzung für Interleaved, was bedeutet, dass nicht nur die schwarzen Balken, sondern auch die weißen Lücken einen Code beinhalten.

Gelesen wird der Barcode durch optische Lesegeräte. Prinzipiell können folgende Arten von Lesegeräten unterschieden werden:

Lesestift

Der Lesestift wird von Hand unmittelbar über den Barcode bewegt. Das durch eine LED erzeugte Licht wird am Barcode reflektiert und als Hell/Dunkel-Signal durch eine Photodiode aufgenommen und ausgewertet.

CCD-Scanner oder CCD-Zeilenkamera

Dem CCD-Scanner oder Zeilen-Sensor liegt eine CCD-Zeile mit 4.000, 6.000 oder 8.000 Pixeln (einzelne Photodioden) zugrunde. Der Barcode wird durch die LEDs beleuchtet und der Barcode reflektiert je nach Helligkeit oder Dunkelheit auf die Photodioden zurück. Handelsübliche Flachbettscanner sind in der Regel mit einer CCD-Zeile ausgestattet.

CCD-Matrixkamera

Eine CCD-Matrixkamera besteht aus einer rechteckigen Matrix photoempfindlicher Halbleiterelemente. Durch das flächenförmige Erfassen eines Bildes kann auf eine Relativbewegung zwischen Kamera und Barcode verzichtet werden. Typische Auflösungen sind 640x480 Pixel. Hochauflösende Kameras erreichen auch 2.000 x 2.000 Pixel Auflösung,

Laserscanner

Der Laserscanner wertet ebenfalls von einem Barcode reflektiertes Licht zum Lesen aus. Als Lichtquelle dient hier allerdings ein Laser. Im Gegensatz zu den anderen Systemen kann der Scanner mit einem bewegten Lichtstrahl arbeiten, so dass eine Relativbewegung zwischen Barcode und Scanner nicht notwendig ist. Durch den scharf gebündelten Lichtstrahl sind auch größere Leseentfernungen möglich. Je nach Bewegung des Laserstrahls werden drei Abtastprinzipien unterschieden:

- Fixed Beam Scanner (Punktförmige Abtastung)
- Moving Beam Scanner (Linienförmige Abtastung)
- Fächer-Scanner (Flächenförmige Abtastung)

Der Laserscanner ist das in der Logistik am häufigsten eingesetzte automatische Lesesystem. Der Scannerstrahl muss den Barcode komplett durchlaufen und keine Lücken oder Fehlerstellen ertasten. Da dieses in der Logistik kaum realisierbar ist, werden durch komplexe Korrekturverfahren, Mehrfachanordnungen oder so genannte omnidirektionale Scanner die Leseraten auf fast 100% gesteigert.

2.3.2.5 RFID-Technologie

Radio Frequency Identification (RFID) ist ein Verfahren, um Daten berührungslos und ohne Sichtkontakt lesen und speichern zu können. RFID wird als Oberbegriff für die komplette technische Infrastruktur verwendet. Sie umfasst

- den Transponder (auch *RFID-Etikett*, *-Chip*, *-Tag*, *-Label*, *Funketikett* oder *-chip* genannt)¹,
- die Sende-Empfangs-Einheit (auch *Reader* genannt) und
- die Integration mit Servern, Diensten und sonstigen Systemen wie z.B. Kassensystemen oder Warenwirtschaftssystemen.

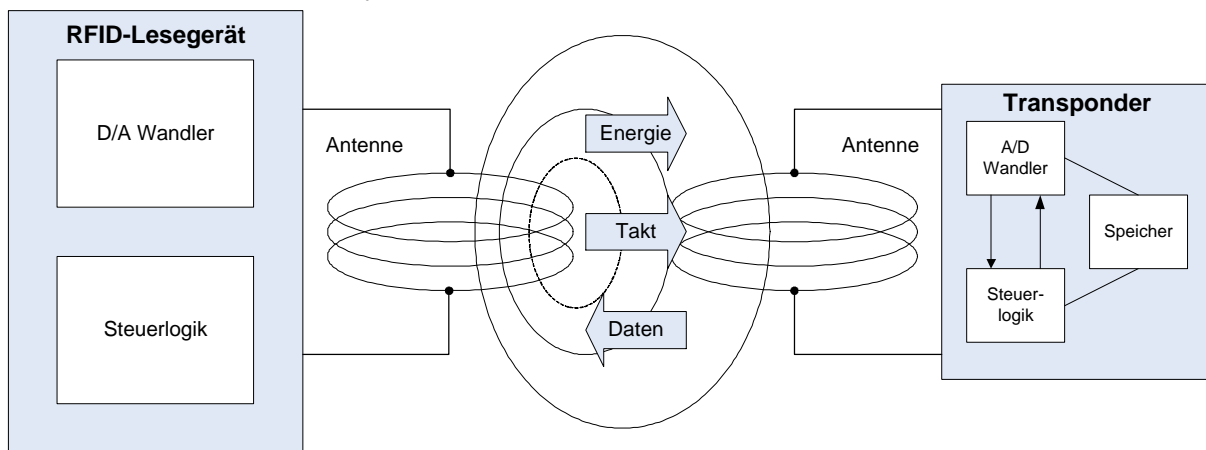


Abbildung 7: Aufbau eines RFID-Systems (nach [HoS03, Abb. 6.30])

Die Daten werden auf den Transpondern gespeichert. Das Auslesen bzw. Schreiben der Informationen geschieht per Radiowellen. Bei niedrigen Frequenzen erfolgt die Kommunikation induktiv über ein Nahfeld, bei höheren Frequenzen über ein elektromagnetisches Fernfeld. Die Entfernung, über die ein RFID-Transponder ausgelesen werden kann, schwankt aufgrund der Ausführung (aktiv/passiv), benutztem Frequenzband, Sendestärke und Umwelteinflüssen zwischen wenigen Zentimetern bis zu ca. einigen Metern.

Für die Kommunikation werden verschiedene Frequenzen verwendet, aus denen sich auch entsprechende Eigenschaften ergeben:

¹ Das Wort Transponder setzt sich zusammen aus den englischen Begriffen *transceive* und *respond*

	Niedrigfrequenz (NF)	Hochfrequenz (HF)	Ultrahochfrequenz (UHF)	Mikrowelle (MW)
Frequenz	125 – 134 kHz	13,56 MHz	868 MHz	2,45 bzw. 5,8 GHz
Leseabstand	bis 1,2 m	bis 1,2 m	bis 4 m	bis zu 15 m (in Einzelfällen bis zu 1km)
Lesege- schwindigkeit	langsam	mittel	schnell	sehr schnell (aktive Transponder)
Feuchtigkeit*	kein Einfluss	kein Einfluss	negativer Einfluss	negativer Einfluss
Metall*	negativer Einfluss	negativer Einfluss	kein Einfluss	kein Einfluss
Typische Transponder-Bautypen	Glasröhrchen-Transponder, Transponder im Plastikgehäuse, Chipkarten Smart Label, Chipkarten	Smart Label, Industrie-transponder	Smart Label, Industrie-transponder	Großformatige Transponder
Beispielhafte Anwendungen	Zutritts- und Routenkontrolle Wegfahrsperrern Wäschereinigung Energieablesung	Wäschereinigung Asset Management Ticketing Tracking & Tracing Pulkerfassung	Palettenerfassung Container-Tracking	Straßenmaut, Container-Tracking

* Die Werte variieren je nach Anbieter

Abbildung 8: Frequenzbereiche der RFID-Technologie

Die Frequenzen in der rechten Spalte (Mikrowelle) befinden sich derzeit im Versuchs- und Entwicklungsstadium und sind noch nicht in der Praxis eingesetzt. Die Werte gelten für Europa, in den USA werden teilweise andere Frequenzen innerhalb der Frequenzbänder verwendet.

RFID-Transponder gibt es als aktive und passive Ausführungen

- Aktive RFID-Transponder sind batteriebetrieben und können typischerweise sowohl gelesen, als auch beschrieben werden. Aktive RFID-Transponder befinden sich normalerweise im Ruhezustand, d.h. sie senden keine Informationen aus. Nur wenn ein spezielles Aktivierungssignal empfangen wird, aktiviert sich der Sender. Der interne Speicher kann, je nach Modell, bis zu 1 Million Bytes aufnehmen. Aktive RFID-Transponder sind im Vergleich zu passiven RFID-Transpondern meist größer, besitzen eine höhere Sendereichweite, haben eine geringere Lebensdauer und sind deutlich teurer.
- Passive RFID-Transponder beziehen ihre Energie zur Übertragung der Informationen aus den empfangenen Funkwellen. Die Menge der gespeicherten Daten ist wesentlich geringer als bei aktiven RFID-Transpondern. In ihrem Speicher wird üblicherweise eine eindeutige Identifikationsnummer hinterlegt. Manche passive Transponder sind mit einem wieder beschreibbaren Speicher ausgerüstet. Passive RFID-Transponder sind im Vergleich zu aktiven RFID-Transpondern kleiner und leichter, haben eine geringe Reichweite, eine nahezu unbegrenzte Lebensdauer und sind vergleichsweise günstig.

„Der Preis von (passiven) RFID-Transpondern bewegt sich, laut einem ZDNet-Artikel, bei einer Auflage von 1 bis 10 Milliarden zwischen 5 und 10 Cent pro Stück. Bei einer Auflage von ca. 10.000 RFID-Transpondern bewegen sich die Preise je nach Größe zwischen 0,50 € bis 1 €. Derzeit kosten aktuelle, passive RFID-Transponder etwa 0,30 €, was noch viel zu

teuer für einen Breitbandeinsatz ist. Man rechnet aber damit, dass die Preise bis 2008 auf wenige Cent pro Einheit fallen werden" [Wik05b].

RFID-Systeme wurden in der Vergangenheit oft als Nachfolger der Barcodes gesehen [Mey04] und sind schon geraume Zeit unter anderem in Lagern und Bibliotheken, beim Paketdienst und der Abfallentsorgung sowie als Zutrittskontrolle im öffentlichen Nahverkehr oder als Wegfahrsperre in Kraftfahrzeugen im Einsatz [Ebe02].

Obwohl viele Vorteile gegenüber der Barcodetechnik existieren, wie

- Kontaktlose Identifikation (auch ohne Sichtkontakt möglich)
- Durchdringung verschiedener Materialien wie Karton, Holz etc.
- Beliebiges Lesen und Beschreiben des Speichers
- Schnelle Identifizierung in weniger als einer Sekunde
- Gleichzeitiges Erfassen vieler Transponder (Pulkerfassung)
- Resistent gegen Umwelteinflüsse
- Form und Größe des Transponders sind nahezu beliebig anpassbar
- Transponder können komplett in das Produkt integriert werden
- Hohe Sicherheit durch Kopierschutz/Verschlüsselung
- Der Chip ist ein Datenspeicher, auf dem Produktdaten hinterlegt werden können. Es ist keine redundante Datenbank notwendig, um erste Informationen gewinnen zu können
- Das Auslesen eines RFID-Tags ist selbst bei größter Verschmutzung möglich
- Die Platzierung des zu erfassenden Objekts ist gegenüber dem Barcode weniger problematisch. Es genügt, wenn sich das Objekt innerhalb des Leseabstands der Erfassungseinheit befindet.

behindern die weitere Verbreitung und Durchsetzung gegenüber dem Barcode bis heute noch die zu hohen Kosten für die einzelnen Transponder, fehlende oder international nicht abgeglichene Standards, Datenschutzprobleme [BSI04] und zum Teil heftiger Widerstand von Verbraucherorganisationen.

Die zum Teil mangelnde öffentliche Akzeptanz von RFID-Systemen hat dazu geführt, dass im Internet Bauanleitungen für aktive Störsender kursieren und auch Firmen beginnen, Blockadetechniken zu vermarkten [Sch04]. Auch werden komplette Baupläne für Detektoren von renommierten und seriösen Quellen angeboten [BaA04].

Dennoch weiten sich die Einsatzbereiche stetig aus. In einer Studie des BSI wird unter folgenden Anwendungsgebieten eine Vielzahl von Einsatzbeispielen der RFID-Technologie aufgeführt [BSI04]

- Kennzeichnung von Objekten
- Echtheitsprüfung von Dokumenten
- Instandhaltung und Reparatur, Rückrufaktionen
- Zutritts- und Routenkontrolle
- Diebstahlsicherung und Reduktion von Verlustmengen
- Umweltmonitoring und Sensorik
- Supply-Chain-Management: Automatisierung, Steuerung und Prozessoptimierung

Der Einzelhandel angeführt von Metro-Group, Wal-Mart und Tesco treiben die RFID - Verbreitung ebenfalls intensiv durch Forschung und Umsetzung voran [Met05], [Met05a].

2.4 Dokumentenanalyse

Die Dokumentenanalyse ist eine Sekundärerhebung und wird auf Grundlage von unternehmensexternen oder für andere unternehmensinterne Zwecke erhobenen Daten durchgeführt.

In Anlehnung an die VDI-Richtlinie 2689 zählen zu den zu analysierenden Unterlagen Organisationspläne, Unterlagen über die Anzahl der Beschäftigten im Untersuchungsbereich, Lage- und Bbauungspläne, Bauzeichnungen, Einrichtungszeichnungen, Produktionsprogramm, Stücklisten, Arbeits- und Fertigungspläne, Terminpläne und Kapazitätsunterlagen, Materialbezugs- und Lieferscheine, Lagerkarten, Inventurbestandslisten, Lohnbelege, Aufschreibungen von automatischen und selbst registrierenden Zeit-Messgeräten, Anlagenkartei, Instandhaltungskartei, Betriebsabrechnungsbogen [VDI2689].

Auf Grund des heutigen Standes der Technik sind oder sollten zumindest die meisten dieser Daten in elektronischen Dokumenten und rechnergestützten Datenbanken abgespeichert und prinzipiell leicht abrufbar sein. Sind die Daten nicht in elektronischer Form verfügbar, so wird im Regelfall ein optisches Dokument vorliegen. Abbildung 9 zeigt die möglichen Formen der Analyse.

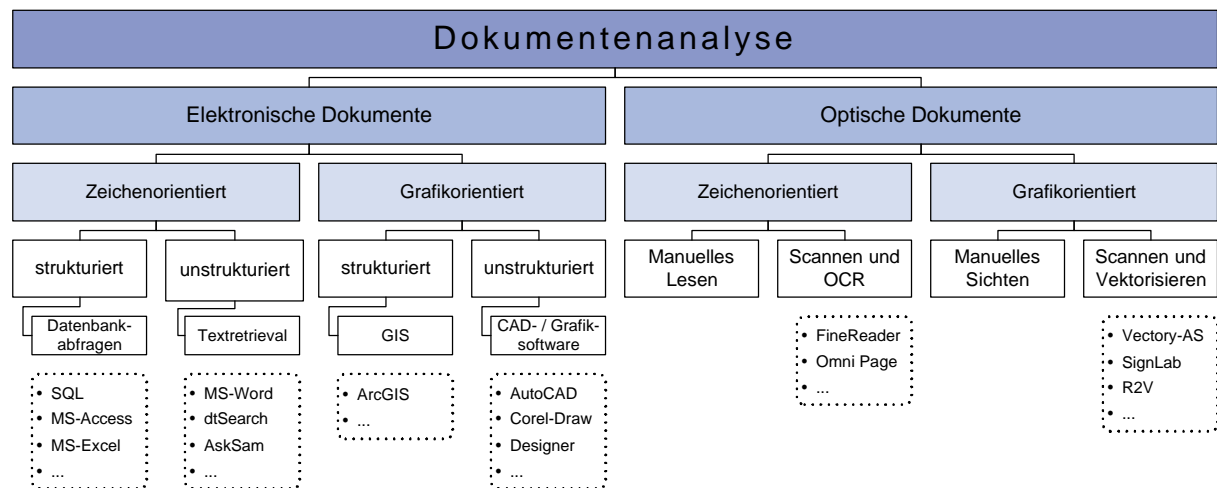


Abbildung 9: Methoden der Dokumentenanalyse

2.4.1 Optische Dokumente

Optische Dokumente können heutzutage ebenfalls elektronisch ausgewertet werden. Durch das Einscannen und Bearbeiten mit einer geeigneten OCR-Software (**O**ptical **C**haracter **R**ecognition) können zeichenorientierte Dokumente ob Fließtexte oder Tabellen digitalisiert und in einen ASCII-Zeichensatz überführt werden. Anschließend entsprechen sie einem elektronischen Dokument und können entsprechend weiterverarbeitet werden. Sind lediglich Teilinhalte oder Auszüge umfangreicher Texte von Bedeutung, wird das Lesen und gegebenenfalls manuelle Erfassen durch Mitschrift jedoch effizienter sein. Grafikorientierte Dokumente können ebenfalls eingescannt werden und liegen dann als so genannte Pixelgrafik vor, die dann mit einer Vektorisierungssoftware in eine vektororientierte Grafik umgewandelt werden kann.

2.4.2 Elektronische Dokumente

Elektronische Dokumente sind Tabellen, Datenbanken, Textdateien, Grafikdateien usw. Sie lassen sich per Drucker oder Plotter in optische Dokumente verwandeln. In zeichenorientierten Dokumenten sind die Daten und Informationen in der einfachsten Form als ASCII-Zeichen gespeichert. Strukturierte Dokumente sind Datenbanken oder Tabellen, die mit ent-

sprechender Software (Abfrage und Analyseprogramme) ausgewertet werden können. Diese Programme verfügen zumeist ebenfalls über Statistik- und Datenvisualisierungshilfsmittel, so dass unter Umständen der gesamte Datenerhebungsprozess mit einem derartigen Programm durchführbar ist. Auf jeden Fall können die Daten nicht nur gefiltert oder sortiert erhoben, sondern zugleich für die Weiterverarbeitung verdichtet und aufbereitet werden.

Aber auch in unstrukturierten zeichenorientierten elektronischen Dokumenten lassen sich mit einfachen Textsuchefunktionen bis hin zu ausgefeilten Retrieval-Programmen Schlüsselworte, Zahlen oder Namen suchen.

Grafikorientierte strukturierte Dokumente sind unter dem Begriff Grafische Informationssysteme (GIS) oder häufig auch als Geografische Informationssysteme bezeichnete Systeme. Der Begriff wird hier aber ein wenig weiter gefasst. Kennzeichnend für diese Systeme ist eine Verknüpfung grafischer Daten mit strukturierten positionsgebundenen Zusatzinformationen. Beispielsweise wird die Verlegung unterirdischer Leitungen für Telekommunikation, Energie, Wasser, Abwasser in speziellen GIS-Karten hinterlegt, so dass für jeden geografischen Ort die dort liegenden Leitungen genannt werden können oder der Verlauf einer bestimmten Leitung oder eines Leitungsnetzes in der Kartenansicht dargestellt werden kann. Es wird nicht nur die Leitung und deren Verlauf, sondern alle damit verbundenen Informationen gespeichert.

Ein weiteres praktisches Beispiel was hier unter GIS subsumiert wird, sind Routenplanungsprogramme, bei denen nicht nur eine Karte, sondern sämtliche zum Verkehrsnetz gehörenden Daten wie Entfernungen, Hindernisse, Geschwindigkeitsprofile usw. gespeichert werden, so dass aus einem solchen System eine „optimale Fahrtroute“ mit Entfernung und Fahrtzeit ermittelt werden kann. In Verbindung mit einem GPS-System wird hieraus ein Navigationssystem, was weiterhin über in der Nähe der aktuellen Position liegende Tankstellen, Restaurant, Hotels und Sehenswürdigkeiten informiert.

Unstrukturiert vorliegende grafikorientierte Dokumente wie CAD-Zeichnungen, Grafiken und Diagramme können mit entsprechender Software ebenfalls ausgewertet werden. Es können Längen, Durchmesser oder Abstände analysiert werden. Bei hochwertigen Grafikprogrammen werden die Grafikdaten darüber hinaus ohnehin als Datentabelle abgelegt, aus der an der grafischen Oberfläche vorbei sehr komfortabel über entsprechende Schnittstellen alle grafikrelevanten Daten ausgelesen werden können.

Genereller Vorteil der Dokumentenanalyse ist der relativ geringe Zeitaufwand für die Informationsbeschaffung, der wegfallende Dokumentationsaufwand, die Objektivität und dass der Untersuchungsbereich nicht in seinen Abläufen gestört wird. Problematisch ist die Möglichkeit der Fehlinterpretation oder das Übersehen von falschen Daten in der Vorlage.

3 Kategorisierung der Methoden

Eine Kategorisierung teilt eine Menge von Objekten, aufgrund gemeinsamer Merkmale, in Abgrenzung zu anderen Objekten, einer Gruppe zu. Die Kategorisierung ist häufig ein Ergebnis der Klassifizierung.

Die Kategorisierung der Datenerhebungsmethoden im Rahmen des Teilprojektes M9 hat das Ziel, die aufgaben- und zielorientierte Methodenauswahl zu ermöglichen und zu unterstützen. Der Anwender, beziehungsweise das Methodennutzungsmodell sucht beispielsweise Methoden mit geringem Zeitaufwand für die Vorbereitung, die für die Erfassung großer Datenmengen in normaler Qualität. Der Filter oder der Suchalgorithmus sucht jetzt Methoden, die gleichzeitig in den Kategorien „niedriger Zeitaufwand zur Vorbereitung“, „Erfassung großer Datenmengen“ und „normale Datenqualität“ enthalten sind. Diesem Ziel folgend wurden die Kategorien entwickelt.

Gleichzeitig wurde auch für die Kategorien eine Systematik oder Klassifizierung verwendet, so dass sich die Kategorisierung letztendlich als Matrix darstellt. Die Zeilen der Matrix werden durch die Methodenklassifikation gebildet, die Spalten durch die Kategorienklassifikation. In den Schnittpunkten wird dann die Zugehörigkeit der Methode zu der jeweiligen Kategorie markiert wobei auch hier noch eine Bewertung der Zugehörigkeitsintensität erfolgen kann.

				Aufwand der Informations- und Datenerhebung										Nutzbarkeit der erhobenen Daten									
				Zeitverbrauch		Personalaufwand	Kostenaufwand		Beeinflussung betrieblicher Abläufe		Komplexität des Verfahrens	Fehleranfälligkeit	Quantität	Qualität	Granularität	Komplexität	Zusatzinformationsgewinn						
				Vorbereitung	Durchführung		Vorbereitung	Durchführung	Vorbereitung	Durchführung													
Sekundärerhebung	Dokumentenanalyse	Optische Dokumente	Grafikorientiert	Scannen und Vektorisieren																			
			Manuelles Sichten																				
		Zeichenorientiert	Scannen und OCR																				
			Manuelles Lesen																				
	Elektronische Dokumente	Grafikorientiert	unstrukturiert	CAD- / Grafiksoftware																			
			strukturiert	GIS																			
		Zeichenorientiert	unstrukturiert	Textretrieval																			
			strukturiert	Datenbankabfrage																			
Primärerhebung	Befragung	Mündlich	Interview	strukturiert	geschlossen																		
				offen																			
				teilstrukturiert	geschlossen																		
				offen																			
		Elektronisch	EDV gestützt	strukturiert	geschlossen																		
			offen																				
	Schriftlich	Fragebogen	strukturiert	geschlossen																			
			offen																				
	Beobachtung	Manuell	Selbstbeobachtung	strukturiert																			
				unstrukturiert																			
			Fremdbeobachtung	strukturiert																			
				unstrukturiert																			
Automatisch		Zählung	strukturiert	Objekte																			
			Ereignisse																				
Messung	strukturiert	Kräfte																					
	Längen																						
	Zeiten																						
	Temperaturen																						
Identifizierung	strukturiert	Geräusche																					
	Helligkeiten																						
Objektkennung																							
Objektdaten																							

Abbildung 10: Kategorisierungsmatrix der Datenerhebungsmethoden

Die Klassifizierung der Kategorien erfolgt in der obersten Ebene durch eine Unterscheidung in Aufwand und Nutzen. Ohne direkt auf eine Nutzwertanalyse ab zu zielen, ergibt sich so auf jeden Fall die Möglichkeit einer guten Abwägung bei der Methodenauswahl.

Der **Aufwand der Informations- und Datenerhebung** wird durch folgende Kategorien dargestellt:

- Zeitverbrauch**
 Der Zeitverbrauch entsteht zwischen dem Tag des Erhebungsbeginns bis zum Tag der Ergebnisfertigstellung unabhängig ob gearbeitet wird oder nicht. Im Vorgehensmodell entspricht er genau der Prozessdauer der Erhebung und Datenerfassung, wie sie im Technical Report 04002 in Abbildung 1 markiert wurde [HJL04, S.3]. Wenn beispielsweise bei einer Befragung mittels eines Fragebogens nach dem Versand vier Wochen für die Beantwortung eingeplant wurden, zählt dieser Zeitraum zum Zeitverbrauch.
- Personalaufwand**
 Der Personalaufwand beinhaltet alle eigenen an der Erhebung beteiligten Personen mit ihren anteiligen Kosten inklusive aller Personalnebenkosten. Die erfassungsbedingten

Personalkosten, die bei dem Unternehmen in dem die Erhebung stattfindet entstehen, fallen unter den Punkt „Beeinflussung betrieblicher Abläufe“.

- **Sachaufwand**
Im Sachaufwand sind alle für die Erhebung anfallenden Kosten abzüglich der Personalkosten, ansonsten gelten die gleichen wie dort genannten Bedingungen
- **Beeinflussung betrieblicher Abläufe**
Hier werden alle Aufwendungen seitens des untersuchten Unternehmens zusammengefasst. Sei es, dass durch fremdes Personal die Abläufe behindert werden oder unternehmenseigenes Personal durch die Datenerfassung zeitlich gebunden wird. Durch Selbstaufschreibung, durch Befragung oder durch das Bereitstellen von Dokumenten.
- **Komplexität / Kompliziertheit des Verfahrens**
Je komplizierter und komplexer das Erhebungsverfahren, desto höher der Aufwand der Vorbereitung und die Gefahr von Störungen oder Akzeptanzproblemen bei Erhebungs- und Unternehmensmitarbeitern.
- **Fehleranfälligkeit**
Die Fehleranfälligkeit kategorisiert das Verhalten der Methode, bei Durchführungsfehlern. Das bedeutet für ein fehleranfälliges Verfahren, dass z.B. besser qualifiziertes Erhebungspersonal eingesetzt werden muss, als bei einem robusteren Verfahren. Beispiel: Bei einem unstrukturierten Interview muss sich der Interviewer viel besser in der Materie auskennen, also besser qualifiziert sein, als bei einem strukturierten Interview.

In diesem Zweig der Klassifizierung folgt bei einigen Kategorien eine weitere Gliederungsebene, die den jeweiligen Prozess der Datenerhebung in die Vorbereitung, die Durchführung und die Aufbereitung unterteilt. Insbesondere bei Verfahren die häufig wiederholt werden kann lohnt es sich unter Umständen, mehr Zeit, Personal und Kosten in die Vorbereitung zu investieren, wenn anschließend die mehrfach durchzuführende Durchführung einen geringeren Aufwand verursacht. Daher ist die Unterscheidung, sofern möglich, hilfreich für die Auswahl des richtigen Verfahrens und eine eigenen Kategorie wert.

Der **Nutzen der erhobenen Daten und Informationen** wird durch folgende Kategorien geprägt:

- Die **Quantität** der Daten bezeichnet eine Kategorie, die Umfang der erhaltenen Daten in Bezug auf die Menge unterschiedlicher Daten und die Anzahl von Aussagen zu einem Datum umfasst. Beispielsweise können mit einem Fragebogen relativ viele Fragen gestellt und damit viel Daten und Informationen gesammelt werden. Gleichzeitig kann der Fragebogen weit gestreut werden und trotz einer geringen Rücklaufquote eine Vielzahl an Antworten zu den Fragen, also viele Daten um beispielsweise Trends zu erkennen, gesammelt werden. Eine Datenbankabfrage einer Artikeldatei von 100.000 Artikeln kann ebenfalls viele Daten erzeugen, also ein eine hohe Quantität. Es wird hier immer ein begrenzter überschaubarer Zeitraum oder Aufwand impliziert.
- Die **Qualität** der Daten ist ein weit reichender Begriff, der im Rahmen dieses Teilprojektes noch Forschungsgegenstand ist. Sie ist auf jeden Fall eine wichtige Kategorie, die mit dem Vertrauen zur Datenquelle aber auch mit dem Erhebungsverfahren zu tun hat und von wesentlicher Bedeutung für die Methodenauswahl ist.
- Die **Granularität** der Daten hängt insbesondere mit der Detaillierung der Daten und der Feingliedrigkeit zusammen. Die Umsätze beispielsweise können Jahresweise, Quartals- oder Monatsweise aber auch für das Gesamtsortiment, für bestimmte Artikelgruppen oder einzelne Artikel von Interesse sein. Die Jahresweisen Umsätze haben dann eine geringere Granularität als die Artikelbezogenen für jeden einzelnen Tag.
- Mit der **Komplexität** der Daten steigt die Vernetzung der Daten untereinander. Einzelne Datenfelder haben eine geringe Komplexität. Eine im Rahmen eines Interviews erhaltene Erläuterung eines Produktionssystems hingegen, hat eine hohe Komplexität. Gleich-

sam unterscheidet sich die Kurzvita einer Person in ihrer Komplexität erheblich von den auf 300 Seiten niedergeschriebenen Memoiren dieser Person.

- Der Nutzen einer möglichen **Erfassung zusätzlicher Informationen** ist sicherlich schwierig zu bewerten. Im Rahmen der vorgesehenen zielorientierten Datenerfassung sind zusätzliche Informationen eigentlich von keiner Bedeutung. Aber ähnlich einer Prüfziffer im Barcode können derartige Informationen aber helfen, die Plausibilität anderer Daten und Informationen besser zu bewerten oder auch Lücken bei der Planung der Datenerhebung aufdecken oder zusätzliche für die Auswertung interessante Bereiche detektieren. Daher ist die mögliche Erfassung weiterer Informationen eine weitere Kategorie der Nutzbarkeit.

Die entstehenden Schnittpunkte der Methoden und Kategorienklassifizierungen dienen der Zuordnung zueinander. An dieser Stelle ist es sinnvoll sich zurückzulehnen und die Angelegenheit mit ein wenig Abstand zu betrachten:

- Das Ziel der Kategorisierung wurde weiter oben genannt und beschrieben. Diesem Ziel sollte sich die Zuordnung anpassen. Das heißt, die Zuordnung dient der Auswahl und dem Vorschlagen geeigneter Methoden um unter vorgegebenen Randbedingungen eine bestimmte Art von Daten und Informationen zu erheben. Es ist anschließend kein automatisierter Prozess vorgesehen, der nun eine bestimmte Erhebungsmethode initialisiert und startet, sondern der Anwender ob Simulationsexperte und/oder Planer erhält eine Auswahl geeigneter Methoden, die bereits einer Vorbewertung unterzogen wurden die mit dem Auswahlvorschlag offen liegen. Nach Ansicht der Autoren werden dem Anwender in der Regel mehr Methoden angeboten, als die er aufgrund seiner Erfahrung zurückgreifen würde.
- Die Zuordnung in die Kategorien ist disjunkt eigentlich nicht möglich. Der Zeitverbrauch eines Verfahrens ist nicht nur vom Verfahren abhängig, sondern von vielen weiteren Randbedingungen. Natürlich vom Umfang der Datenerhebung, von den betrieblichen Rahmenbedingungen usw. Man müsste also komplexe Verknüpfungen einfügen, um hier zu Einzelfallrepräsentativen Einordnungen zu kommen. Für die manuelle Zählung eingehender Paletten in den Wareneingang mit 10 Toren sind zehn Personen zum Zählen viel. Für eine manuelle Verkehrszählung vor den Dortmunder Innenstadtparkhäusern sind zehn Personen hingegen wenig. Dennoch wird man Einig werden, dass bei der manuellen Zählung der Personalbedarf bei der Durchführung prinzipiell höher ist als bei einer automatischen Zählung.

Vor diesem Hintergrund findet derzeit die weitere Diskussion in der Arbeitsgruppe hinsichtlich der Kategorisierung statt. Tenor ist, dass eine komplexe feingliedrige Zuordnung wenig hilfreich ist und man mit maximal drei Stufen auskommt, beispielsweise: hoch-mittel-niedrig. Weiterhin sind einige Klassifikationen noch Bestandteil der laufenden oder zukünftigen Arbeitspunkte wie z.B. die Datenqualität, so dass die Kategorienzuordnung erst zu einem späteren Zeitpunkt erfolgt und Inhalt eines weiteren Technical Reports in 2006 sein wird.

4 Automatisierte Systeme der Datenerhebung

Um die in anderen Disziplinen entwickelten Verfahren zur automatisierten Datenerhebung kennen zu lernen und eventuell adaptieren zu können wurden verschiedene Realisierungen untersucht, von denen hier einige Ausgewählte im Überblick dargestellt werden. Bei allen Systemen werden Daten im Moment ihrer Entstehung erfasst, mit anderen Daten und Informationen verknüpft, um so zu den gewünschten Informationen wie Stauprognose, Wetterprognose usw. zu gelangen.

4.1 Maut System Deutschland: Toll Collect

Das unter großer öffentlicher Beobachtung stehende und mehrfach verschobene Maut-System stellt eine Kombination aus verschiedenen Erfassungstechniken dar.

Das deutsche System wird als GNSS/CN-Maut (Global Navigation Satellite System mit Rückkanal Cellular Network) bezeichnet.

Herzstück des Systems ist eine On Board Unit (OBU) mit einem Empfänger für das US-amerikanische GPS-Ortungssignal, einer GSM-Verbindung zur Mautzentrale, einem Infrarot-Modul und einem Speicher, in dem fahrzeugspezifische Angaben wie Kfz-Kennzeichen, Abgasschadensklasse und die Zahl der Achsen sowie die Positionsdaten der deutschen Autobahnen hinterlegt sind. Für die GSM-Verbindung ist in jeder OBU ein Mobiltelefon eingebaut, das die Fahrdaten und die per GPS ermittelten Positionen an die Toll Collect Zentrale weitergibt. Dieser Teil des Mautsystems funktioniert somit ohne flächendeckende Installationen weiterer Hardware [Bor03].

Problematisch ist die Tatsache, dass eine Pflicht zur Installation einer OBU nicht mit EU-Recht zu vereinbaren ist, welches ausländischen LKW-Fahrern einen „diskriminierungsfreien Zugang“ zum deutschen Mautsystem sichern will [BuM03].

Um auch LKWs ohne installierter oder funktionsuntüchtiger OBU korrekt abrechnen zu können, war daher eine Möglichkeit zu schaffen, die anfallenden Mautgebühren manuell bezahlen zu können. Dieses ist im Internet, an Terminals an den Grenzen, vor Autobahnauffahrten, an Parkplätzen, Tankstellen und Raststätten rund um die Uhr möglich. Dem Fahrer hilft dabei eine Karte, auf der bereits alle Fahrzeugdaten gespeichert sind, die zur Mauterhebung benötigt werden. Abgerechnet wird per Kredit- oder Tankkarte, zudem gibt es die Möglichkeit, die Maut am Automaten bar und in mehreren europäischen Währungen zu bezahlen [BuM03].

Um die korrekte Zahlung der Maut entweder per funktionierender OBU oder Terminalbuchung kontrollieren zu können, wurden zusätzlich etwa 300 Kontrollbrücken, sog. „Tollchecker“ aufgestellt. Diese erfassen jedes Fahrzeug auf jeder Fahrspur, fotografieren es und führen eine dreidimensionale Vermessung durch um PKW von LKW zu unterscheiden. Handelt es sich um einen LKW, so werden zwei Bilder von der Frontpartie und dem Kfz-Kennzeichen geschossen. Parallel dazu gibt sich die OBU per Signal zu erkennen. Gleichzeitig liest eine Gegenstelle per DSRC-Kommunikation (**D**edicated **S**hort **R**ange **C**ommunication) die in der OBU gespeicherten Fahrzeugdaten aus und vergleicht sie mit den Werten der Lasermessung. Stimmen die Angaben überein, werden alle abgefragten Daten sowie die Digitalbilder umgehend gelöscht. Ist keine OBU vorhanden, so wird das Kfz-Kennzeichen mittels OCR-Schrifterkennung in einen Datensatz umgewandelt und eine Abfrage in der Zentrale von Toll Collect durchgeführt, ob für dieses Fahrzeug die Maut entrichtet wurde. Ist dieses der Fall, werden die Daten vom Tollchecker gelöscht, andernfalls gespeichert um Nachforderungen und Bußgeldbescheide rechtlich abzusichern.

Die installierte Technologie kann noch für weitere Anwendungen genutzt werden, welche allerdings noch juristisch umstritten sind.

So bedarf eine Erweiterung der Tollchecker um erkenntnisdienliche Methoden nur eines Software-Updates, um etwa zur Fahndung ausgeschriebene Fahrzeuge zu identifizieren. Zwei entsprechende Brücken mit modifizierter Software wurden zur Kontrolle des einreisenden Verkehrs an zwei deutsch-tschechischen Grenzübergängen aufgebaut und bis Ende 2002 erprobt. Sie erfassten die Kfz-Kennzeichen aller Einreisenden und glichen die Daten mit dem Fahndungscomputer der bayerischen Polizei ab.

Eine Ausweitung der Maut auf PKW ist nach den Vorgaben der EU ohnehin angedacht, denkbar sind auch Geschwindigkeitskontrollen auf Autobahn-Teilstrecken mit einer durchgehenden Geschwindigkeitsbegrenzung. Dabei meldet die Brücke die Kfz-Daten der jeweils

nachfolgenden Kontrollbrücke: Kommt das Fahrzeug zu früh an, kann auf eine Geschwindigkeitsüberschreitung geschlossen werden.

4.2 Verkehrsdatenerfassung (Verkehrsvorhersage)

Um Störungen im Verkehrsfluss auf den bundesdeutschen Autobahnen schnell zu identifizieren und nach Möglichkeit potenzielle Störungen bereits im Voraus zu prognostizieren, hat die Gesellschaft für Verkehrsdaten (DDG) ein bundesweites Netz von stationären und mobilen Sensoren installiert. Diese dienen als Basis für ein intelligentes Verkehrsinformationssystem.

Das Rückrat dieses Systems bilden etwa 4.000 an den Autobahnen installierte Passiv-Infrarot Sensoren. Entgegen zuerst von Automobilclubs geäußerten Befürchtungen können diese Sensoren die Fahrzeuge lediglich zählen und deren Geschwindigkeit und Länge messen, nicht aber die Fahrzeuge identifizieren und zur Geschwindigkeitsüberwachung eingesetzt werden. Die Sensoren arbeiten vollkommen autonom mit einer Energieversorgung mittels Solar-Panel und Pufferbatterie während die Kommunikation des Sensors mit der Zentrale über ein GSM-Modul per Short Messages (SMS) geschieht.

Das physikalische Messprinzip der Infrarot-Stausensoren ist einfach: Durch zwei "aktive Flächen" auf der Fahrbahnoberfläche ist die Funktion einer doppelten Lichtschranke realisiert, die Geschwindigkeitsmessungen erlaubt. Aus einer Geschwindigkeit und einer Belegungsdauer lässt sich die Fahrzeuglänge und damit näherungsweise der Fahrzeugtyp ermitteln. Aus der Zählung von Impulsen ergibt sich die Verkehrsmenge aufgeschlüsselt auf Individualfahrzeuge. Dieses bedingt eine vergleichsweise hohe Datenmenge, die zwar aufgrund der Genauigkeit für Verkehrsforschungszwecke von höchstem Interesse wäre, allerdings auch einen erheblichen Aufwand zur Übertragung nach sich ziehen würde.

Würde man dem Detektor gestatten, in kurzen zeitlichen Abständen sämtliche Messwerte zu übertragen, wären zu hohe Leistungsaufnahmen und Datentransferkosten die Folge. Andererseits ist es aber gerade die Aufgabe des Systems, so aktuell wie möglich, den Verkehrszustand in der DDG-Zentrale verfügbar zu machen und auch den Tagesgang von Verkehrsaufkommen und mittleren Geschwindigkeiten vollständig für statistische Auswertungen und zu Prognosezwecken aufzuzeichnen. Datenlücken sind daher nicht gestattet.

Der Ausweg aus diesem Zielkonflikt besteht darin, bereits im Detektor eine Vorverarbeitung der Messdaten durchzuführen und zeitkritische Daten sofort, andere jedoch erst später zu senden. Ein offensichtlicher Grund, sofort eine Meldung abzusetzen, ist etwa ein plötzlicher Rückgang der mittleren Fließgeschwindigkeit: Der Verkehr unmittelbar unter dem Sensor ist ins Stocken geraten. Ein etwas weniger offensichtlicher Grund ist ein plötzlicher Rückgang der Verkehrsmenge bei gleich bleibend hoher Geschwindigkeit. Dies könnte ein Hinweis auf das Eintreten einer Störung weiter stromaufwärts sein (zum Beispiel eines Unfalls) und muss daher auch sofort gemeldet werden.

Mit den Daten der 4.000 fest installierten Sensoren ist allerdings nur das Primärnetz der Autobahnen abgedeckt. Informationen über das aus Bundesstraßen und Umleitungsstrecken bestehende Sekundärnetz liegen somit noch nicht vor. Da im Falle einer Störung im Primärnetz der Verkehr über das Sekundärnetz geroutet werden soll, sind diese Daten unabdingbar. Aufgrund der vierfachen Länge des Sekundärnetzes gegenüber dem Primärnetz scheidet eine flächendeckende Installation von Sensoren aus. Es werden vielmehr mobile Sensoren in Form von Floating Cars eingesetzt. Dabei handelt es sich um private Fahrzeuge, die mittels eines GPS-Moduls und einer Kommunikationseinrichtung, GSM-Telefon oder Modem, seinen Aufenthaltsort und seine Geschwindigkeit übermitteln kann. Da derzeit etwa 40.000 Floating Cars im Einsatz sind und als flächendeckende Zielgröße etwa das achtfache anzunehmen ist, verbietet sich ähnlich wie bei den stationären Sensoren eine permanente Übertragung der Daten in die Zentrale. Auch hier liegt die Lösung darin, den Sensor intelligent zu machen: Ein spezieller Algorithmus analysiert ständig das Geschwindigkeits-

profil seit Fahrtbeginn, um von Fahrzeugtyp und Fahrstil abstrahieren und wesentliche, verkehrsbedingte Veränderungen des Verkehrsflusses detektieren zu können. So muss nicht jeder Messwert für Ort, Zeit und Geschwindigkeit fein aufgelöst und übertragen werden, sondern die Fahrt kann in längere Abschnitte unterteilt werden, die sich wiederum durch wenige charakteristische Kennzahlen beschreiben und an die Zentrale übertragen lassen. Auch lassen sich so genau die Ereignisse definieren, die eine Übertragung erst auslösen, zum Beispiel die Einfahrt in einen Stau oder das Verlassen eines solchen.

Durch diese Beschränkung der Erfassungsaufgabe der Floating Cars auf das Erkennen von Verkehrsstörungen ist es möglich, dass ein durchschnittliches Floating Car nicht mehr als ein Telegramm pro Tag sendet.

Eine detaillierte Beschreibung der Technik und insbesondere die zugrunde liegenden Datenverarbeitungsalgorithmen finden sich in der Literatur, z. B. [Fas04].

4.3 Wetterdatenerfassung

Die Gesamtheit der meteorologischen Parameter, die den Zustand der Atmosphäre an einem Ort, und damit das Wetter, vollständig beschreiben, kann nur durch sehr großen Aufwand erfasst werden.

Ein nachträgliches Einholen von wetterdienstlichen Informationen und Messwerten führt in fast allen Fällen zu unbefriedigenden oder falschen Ergebnissen, da die zeitliche und räumliche Repräsentanz der mit den Routine-Stationen eines Wetterdienstes erfassten Daten im allgemeinen nicht ausreicht. Deshalb hat man z.B. in der Bundesrepublik Deutschland in die Luftüberwachungsmessnetze der Länder die automatische Erfassung meteorologischer Parameter integriert. Außer diesen in Ausstattung und Betrieb sehr aufwändigen Systemen werden viele Messprogramme mit mobilen Anlagen oder an temporären Stationen und mit unterschiedlicher Dauer durchgeführt [VDI3786].

Die Richtlinienreihe VDI 3786 legt dabei im Detail fest, welche Daten wie zu erheben, zu übermitteln und zu verarbeiten sind.

Die beobachteten Daten sind im Allgemeinen die folgenden meteorologischen Größen:

- Windgeschwindigkeit und –richtung
- Lufttemperatur
- Luftfeuchte
- Luftdruck
- Niederschlag
- Solare und terrestrische Strahlung
- Lufttrübung
- Luftinhaltsstoffe

Aufgrund der sehr großen Anzahl an notwendigen Beobachtungsstationen wurden im Bereich der Wetterdatenerfassung automatische oder halbautomatische Beobachtungssysteme entwickelt, die der jeweiligen Messaufgabe angepasst wurden. Diese beinhalten einen Sensor (u.a. Temperatur, Druck, Feuchte) und eine Speichereinheit, welche die erhobenen Daten zum Teil noch analog auf Registrierpapier schreibt oder in modernen Ausführungen im Digitalformat in Festkörperspeichern ablegt oder direkt an eine übergeordnete Stelle übermittelt.

Die einzusetzenden Geräte lassen sich in drei Gruppen einteilen:

- Mechanisch angetriebene Registriergeräte
- Elektrisch betriebene Registriergeräte, entweder batteriebetrieben oder netzabhängig

- Festkörperspeicher, entweder batteriebetrieben oder netzabhängig

Ein weiterer wesentlicher Aspekt der Wetterbeobachtung ist das Problem der Datenübermittlung. Unter diesem Begriff ist der Transport der Messwerte mit allen Zusatzinformationen wie Messort, Messzeit, Komponente u.a. vom Messort zum Ort der Auswertung zu verstehen.

Nach VDI 3786 bieten sich hier insbesondere folgende Möglichkeiten an:

- Erfassung auf Registrierpapier, dann manueller Transport und Weiterverarbeitung
- Erfassung auf magnetischen und elektronischem Datenträger (Magnetband, Diskette, Festkörperspeicher), dann manueller Transport und rechnergestützte Weiterverarbeitung
- Erfassung auf magnetischen und elektronischem Datenträger, rechnergestützter Datenabruf (passiv) über IT-Systeme (z.B. Telefon, Telex oder Funk) und rechnergestützter Weiterverarbeitung
- Erfassung und mikroprozessorgesteuerter Datenübermittlung (aktiv) über Telefon, Telex oder Funk und rechnergestützter Weiterverarbeitung

5 Ausblick

Nach Erweiterung und Modifikation der Methodenklassifizierung steht nunmehr eine allgemeingültige Systematik zur Verfügung, die als Basis verschiedener Prozesse dienen kann. Zum einen verschafft sie einen Überblick über die Vielzahl der Methoden zur Datenerhebung in der Logistik. Sie hilft, die Methoden zueinander in Relation zu setzen, gegenseitig abzugrenzen oder Gemeinsamkeiten zu finden.

Zum anderen ist sie Grundlage einer schlüssigen Kategorisierung und unterstützt durch ihre Struktur bereits den Methodenfindungsprozess im Sinne der Ziele im Teilprojekt M9.

Weiterhin wurden Kategorien gefunden und definiert, die hinsichtlich der zielorientierten Auswahl der Methoden klare und eindeutige Kriterien beinhalten. Wie beschrieben bleibt es aber problematisch die Methoden eindeutig innerhalb einer Kategorie zuzuordnen, was im Rahmen der hier verfolgten Ziele bisher unkritisch erscheint. Dennoch wird hier weiterhin nach Verbesserungen gesucht, da ohnehin einige Kategorien noch Forschungsgegenstand anstehender Arbeitspakete sind.

Im Verlauf der nächsten zwölf Monate sollen daher parallel zu den anderen laufenden Arbeitspunkten die Methoden sukzessive den Kategorien zugeordnet werden. Wobei dieser Prozess durch evaluierende Abstimmungen mit anderen Teilprojekten sowie Brainstorming-Sitzungen und Expertenmethoden innerhalb des Teilprojektes und mit externen Experten begleitet wird.

Die im letzten Gliederungspunkt vorgestellten automatisierten Datenerhebungsmethoden aus anderen Disziplinen sind nur eine Auswahl der bisher betrachteten Systeme. Hier werden weitere Beispiele gesammelt und untersucht werden. Der Nutzen dieser externen Betrachtung liegt in zwei Bereichen. Zum einen können Verfahren und Methoden für die vorliegenden Fragestellungen abgeleitet werden, zum anderen können einige Systeme aber auch als Datenbasis für benötigte Logistikdaten dienen.

Literatur

- [Act05] Active Barcode (Hrsg.): *Barcodetypen*. <<http://www.activebarcode.de/codes/>>, Abrufdatum 16.5.2005.
- [BaA04] Bartels, Oliver; Ahlers, Ernst: *Gegenspionage – RFID-Detektor im Taschenformat*. In: c't 9/2004, Heise Verlag, Hannover, 2004.
- [Bor03] Borchers, Detlef: *Verursacherbedingt verspätet – Das „fortschrittlichste Maut-System der Welt“ und die Realität*. In: c't 22/2003, Heise Verlag, Hannover, 2003.
- [BSI04] Bundesamt für Sicherheit in der Informationstechnik (Hrsg.): *Risiken und Chancen des Einsatzes von RFID-Systemen - Trends und Entwicklungen in Technologien, Anwendungen und Sicherheit*. Studie des Bundesamts für Sicherheit in der Informationstechnik, SecuMedia Verlagsgemeinschaft, Ingelheim, 2004, ISBN 3-922746-56-X.
- [Büh99] Bühner, Rolf: *Betriebswirtschaftliche Organisationslehre*, Verlag Oldenbourg, München, 1999.
- [Dat00] Datalogic (Hrsg.): *Strichcodebibel*. Firmenschrift 9C0003160, Rel. 5.0, 2000.
- [Ebe02] Ebeling, Adolf: *Etikettierungen – Vom Barcode zum Smart-Label*. In: c't 9/2002, Heise Verlag, Hannover, 2002.
- [Els05] Oberfinanzdirektion München (Hrsg.): *ElsterWeb*, <<http://www.Elster.de>> Abrufdatum: 15.5.05.
- [Fas04] Fastenrath, Ulrich: *Stau-Schau - Auf dem Weg zum intelligenten Informationssystem*. In: c't 9/2004, Heise Verlag, Hannover, 2004.
- [Fin02] Finkenzeller, Klaus: *RFID-Handbuch - Grundlagen und praktische Anwendungen induktiver Funkanlagen, Transponder und kontaktloser Chipkarten*. 3. Auflage, Hanser Verlag, München, 2002.
- [Har99] Hartung, Joachim: *Statistik – Lehr- und Handbuch der angewandten Statistik*. Verlag Oldenbourg, München, 1999.
- [HJL04] Hömberg, Kay; Jodin, Dirk; Leppin, Maren: *Methoden der Informations- und Datenerhebung*. Technical Report – Sonderforschungsbereich 559 „Modellierung großer Netze der Logistik“ 04002, 2003, ISSN 1612-1376.
- [HoS03] ten Hompel, Michael; Schmidt, Thorsten: *Warehouse Management*. Springer-Verlag, Berlin et al., 2003.
- [JBW01] Jodin, Dirk; Becker, Claudia; Wenzel Siegrid (2001): *Methoden der Informationsgewinnung zur Bestimmung des Eingangsdatenraums für Simulationsmodelle in der Logistik*. In: Panreck, K.; Dörscheidt, F. (Hrsg.): *Simulationstechnik*, Tagungsband zum 15. Symposium ASIM 2001, Paderborn, 11.-14. September 2001. Reihe: Fortschrittsberichte Simulation Frontiers in Simulation, SCS-Europe BVBA, Ghent, 2001, S. 417-422.
- [JoH05] Jodin, Dirk; Hömberg, Kay (2005): *Datenbeschaffung für die Simulation eines Logistikzentrums*. In: Wolf-Kluthausen, Hanne (Hrsg.): *Jahrbuch der Logistik 2005*, S. 184 – 187. 2005.
- [JüB98] Jünemann, Reinhardt; Beyer, Andreas: *Steuerung von Materialfluß- und Logistiksystemen: Informations- und Steuerungssysteme, Automatisierungstechnik*; 2. Aufl., Springer-Verlag, Berlin et al., 1998.

- [Kel05] Kellerhoff, Peter: *Chaos durch steigende Datenmengen*. In: VDI-Nachrichten Nr. 19. 13. Mai 2005, 2005, S. 12.
- [Kus04] Kuschner, David (2004): *RFID-Technik geht unter die Haut*, Technology Review, Heise Verlag, Hannover, 2004.
<http://www.heise.de/tr/aktuell/meldung/print/53785>
- [Len00] Lenk, Bernhard: *ID-Techniken, 1D-Codes, 2D-Codes, 3D-Codes*, Handbuch der automatischen Identifikation, Band 1, Monika Lenk Fachbuchverlag, 2000.
- [Len04] Lenk, Bernhard: *Strichcode Praxis*. Handbuch der automatischen Identifikation, Band 3, Monika Lenk Fachbuchverlag, 2004
- [Man05] Manhattan Review (Hrsg.): *TOEFL-Information*. <<http://www.review.de/de/toefl/>>, Abrufdatum 15.5.2005.
- [Met04] Metro-Group (2004): *Einsatz von RFID im Future Store in Rheinberg*. <http://www.future-store.org/servlet/PB/-s/3iwvcj1a2il4gyq7py0zs1ujk108ufs5/menu/1002376_l1/index.html>, Erscheinungsdatum: 20.04.2004, Abrufdatum: 15.03.2005.
- [Met05] Metro AG (Hrsg.): *Schulterschluss im Handel*. RFID Newsletter 2/2005, S. 2-4. 2005.
- [Met05a] Metro AG (Hrsg.): *Lösungen nach Maß*. RFID Newsletter 2/2005, S. 2-4, 2005.
- [Mey04] Meyer, Angela: *Lückenlos dokumentiert - RFID Technik löst Barcode Etiketten ab*. In: c't 3/2004, Heise Verlag, Hannover, 2004.
- [Sie04] Sietmann, Richard: *Chipkarten im Aufwind – vom elektronischen WM-Ticket zur digitalen Patientenkarte*. In: c't 3/2004, Heise Verlag, Hannover, 2004.
- [Sch04] Schüler, Peter: *Schnüffeltechnik ausgetrickst - Ein Störsender verwirrt RFID-Lesegeräte*. In: c't 6/2004, Heise Verlag, Hannover, 2004.
- [Sch04] Schulzki-Haddouti, Christiane: *Elektronischer Pass – „Biometrische Reisepässe“ mit RFID-Chips in der Einführungsphase*. In: c't 9/2004, Heise Verlag, Hannover, 2004.
- [Sch03] Schmidt, Hendrik: Datenerhebung, <<http://www.mathematik.uni-ulm.de/stochastik/lehre/ss03/wirtschaftsstatistik/skript9/node6.html>>, Abrufdatum: 12.3.2005.
- [Sch99] Schulte-Zurhausen, Manfred: *Organisation*; Verlag: Vahlen; München 1999.
- [VDI2492] VDI-Richtlinie 2492 (1968): *Multimomentaufnahmen im Materialfluß*. VDI-Handbuch Förderwesen, VDI-Verlag GmbH, Düsseldorf, 1968.
- [VDI2689] VDI-Richtlinie 2689 (1974): *Leitfaden für Materialflußuntersuchungen*. VDI-Handbuch Materialfluß und Fördertechnik, VDI-Verlag GmbH, Düsseldorf, 1974.
- [VDI3786.9] VDI-Richtlinie 3312 Blatt 9 (2003): *Visuelle Wetterbeobachtungen*. Beuth Verlag GmbH, Berlin, 2003.
- [Voß04] Voß, Werner: *Taschenbuch der Statistik*, Fachbuchverlag Leipzig im Carl Hanser Verlag, München, 2004.
- [WiT93] Wittlage, Helmut: *Methoden und Techniken praktischer Organisationsarbeit*, Verlag Neue Wirtschafts-Briefe, Herne, 1993.

- [Wik05] Wikipedia (Hrsg.): *Nyquist-Shannon-Abtasttheorem*.
<<http://de.wikipedia.org/wiki/Nyquist-Shannon-Abtasttheorem>>, Abrufdatum:
16.5.2005.
- [Wik05a] Wikipedia (Hrsg.), <<http://de.wikipedia.org/wiki/Barcode#Geschichtliches>>,
Abrufdatum 16.5.2005
- [Wik05b] Wikipedia (Hrsg.): *Radio Frequency Identification*.
<<http://de.wikipedia.org/wiki/RFID>>, Aktualisierungsdatum: 14.5.2004, Abruf-
datum: 16.5.2005.

Sonderforschungsbereich 559

Bisher erschienene Technical Reports

- 03020 Michael Kaczmarek, Marcus Völker: Entwicklung von Simulationsmodellen für die Analyse von Supply Chain-Strategien und -Strukturen im ProC/B-Paradigma
- 03021 Michael Kaczmarek: Beschreibung ausgewählter Strategien zur Steuerung der Austauschprozesse in der Supply Chain
- 03022 Michael Kaczmarek: Organisation der Planung und Steuerung in Supply Chains
- 03024 Anne Schulze im Hove, Frank Stüllenberg, Stefan Weidt: Inhaltliche Ausgestaltung der Netzwerk-Balanced-Scorecard für Beschaffungsketten
- 03029 Hilmar Heinrichmeyer, Andreas Reinholz: Entwicklung eines Bewertungsmodells für die Depotstandortoptimierung bei Servicenetzen
- 03032 Marco Motta, Iwo Riha, Stefan Weidt: Simulation eines Regionallagerkonzeptes
- 03034 Frank Laakmann, Iwo Riha, Niklas Stracke, Stefan Weidt: Workbenchgestützte Konstruktion von Beschaffungsketten
- 03035 Iwo Riha, Stefan Weidt: Entwicklung einer Bewertungssystematik für Beschaffungsketten
- 04001 André Alberti, Bernd Hellingrath, Stefan Weidt, Markus Witthaut: Ergebnisse und Schlussfolgerungen der Simulationsexperimente im Szenario Automobilindustrie
- 04002 Kay Hömberg, Dirk Jodin, Maren Leppin: Methoden der Informations- und Datenerhebung
- 04003 Carsten Tepper: Prozessablauf-Visualisierung von ProC/B-Modellen
- 05001 Jochen Bernhard, Miroslav Dragan, Sigrid Wenzel: Evaluation und Erweiterung der Kriterien zur Klassifizierung von Visualisierungsverfahren für GNL
- 05002 Entwicklung eines Analyse Rahmens für die Untersuchung organisatorischer Aspekte in der Supply Chain
- 5003 Einsatz der Response Surface Methode zur Optimierung komplexer Simulationsmodelle
- 05004 Automatisierte Methoden und Systeme der Datenerhebung

Alle Technical Reports können im Internet unter
<http://www.sfb559.uni-dortmund.de/>
abgerufen werden. Für eine Druckversion wenden Sie
sich bitte an die SFB-Geschäftsstelle
e-mail: grosseca@iml.fhg.de