

Datengeneratoren für die Unterstützung von Design-Science-Projekten im Umfeld RFID-basierter Analyseanwendungen

Xuanpu Sun, Henning Baars

*Lehrstuhl für ABWL und Wirtschaftsinformatik I,
Universität Stuttgart*

1 Motivation

Seit einigen Jahren wird verstärkt der Einsatz von Radio Frequency Identification (RFID) im Kontext der Logistik diskutiert (Curtin et al. 2007, S. 87-110; Strüker et al. 2008). Treiber sind hierbei nicht zuletzt die Initiativen größerer Handelskonzerne (Hoffman 2008, S. 16; METRO Group 2004) und die parallel vorangetriebenen Standardisierungsbemühungen durch Gremien wie EPCglobal (EPCglobal 2009). Die Möglichkeiten einer weitgehend automatisierten Objektidentifikation, zur Pulk-Erfassung sowie zur Kopplung von RFID-Systemen mit Sensor- und/oder Tracking-und-Tracing-Technologien (Clauberg 2004, S. 1-6; Günther et al. 2008, S. 7-18; Kreller und Hartmann 2001, S. 1-8) versprechen eine erheblich gesteigerte Transparenz über Güterbewegungen (Brewer und Sloan 1999, S. 245-250; Kambil und Brooks 2002).

Es ist zu konstatieren, dass das Spektrum an Anwendungen aufgrund der zunehmenden Reife und Beherrschung der Technologie stetig zunimmt. Mit der Vielfalt an innovativen Anwendungen wächst auch der Bedarf an Ansätzen für deren Evaluation. Diese erweist sich allerdings bei RFID-Systemen als besonders schwierig. Das gilt insbesondere für Anwendungen, die über eine reine Automatisierung hinausgehen und die auf der im Vergleich zu den Alternativen qualitativ höherwertigen Informationsbereitstellung basieren (Baars et al. 2009). Besonders betroffen hiervon sind konstruktionsorientierte Forschungsvorhaben auf Basis des Design-Science-Ansatzes (Hevner et al. 2004, S. 75-105). Bei entsprechenden Projekten stehen die beteiligten Forscher vor der Herausforderung, bisher noch nicht erprobte Systementwürfe beziehungsweise darin enthaltene Designentscheidungen bewerten zu müssen.

Für die Identifikation und Bewertung des Informationsnutzens von RFID kann mittlerweile auf einen reichen Fundus an Vorschlägen zurückgegriffen werden, etwa auf der Grundlage von Prozessmodellierungen (Laubacher et al. 2005,

S. 1-6) oder von Simulationstechniken (Lee et al. 2004, S. 1146-1151). All diesen Ansätzen entziehen sich jedoch Lösungen, die auf eine *flexible* Informationsbereitstellung und -analyse ausgerichtet sind und die hierfür Datenhaltungssysteme mit historisierten und angereicherten RFID-Daten nutzen (Chawathe et al. 2004, S. 1189-1195; Baars und Sun 2009). Entsprechenden Anwendungen wird in der Forschung in letzter Zeit zunehmend mehr Aufmerksamkeit gewidmet. Insbesondere wird vorgeschlagen, RFID-Systeme mit Business-Intelligence-Infrastrukturen für die integrierte Entscheidungsunterstützung zu koppeln und RFID-Daten in einem Data Warehouse (DWH) bereitzustellen (Ilic et al. 2009; Baars et al. 2008, S. 73-83). Je flexibler derartige Systeme ausgestaltet werden, desto weniger lassen sich mögliche Nutzeffekte vorhersehen.

Obwohl eine exakte Quantifizierung eines solchen „Analysenutzens“ prinzipielle Probleme aufwirft, legen es erste Fallstudien nahe, dass im Bereich der Logistik eine expertenbasierte Abgrenzung und Bewertung grundsätzlicher Einsatzszenarios sehr wohl möglich ist (Baars et al. 2008, S. 73-83). Hierbei kann abgeschätzt werden, welche Daten aus welchen Systemen in welche Arten von Analyseanwendungen einfließen sollten und in welchem Spektrum sich der potentielle Nutzen grundsätzlich bewegt. Eine der Herausforderungen eines solchen Vorgehens besteht darin, mögliche Anwendungsfälle effektiv zu kommunizieren und zu durchdringen. Hierbei bietet es sich an, unterstützend mit Prototypen zu arbeiten. Voraussetzung für den effektiven Einsatz von Prototypen sind jedoch aussagekräftige Demonstrationsdaten. In diesem Beitrag wird vorgeschlagen, für deren Erzeugung einen flexibel konfigurierbaren Datengenerator bereitzustellen. Anders als in einer Simulation sind die Analyseszenarios dabei nicht Ergebnis sondern *Ausgangspunkt* der Datengenerierung.

Wie gezeigt wird, ist die Generierung szenariokonformer, konsistenter und nachvollziehbarer Daten mit einer sehr hohen Komplexität verbunden. Das zur Lösung entworfene und prototypisch realisierte Design des Datengenerators wird in diesem Beitrag skizziert und diskutiert. Hierbei werden jene Entwurfsentscheidungen und Erkenntnisse offengelegt, bei denen davon ausgegangen wird, dass sie auch für die Planung und Strukturierung vergleichbarer Projekte relevant sind. Am Schluss des Beitrags wird die Übertragbarkeit des Ansatzes auf andere Gegenstände und Anwendungsdomänen diskutiert.

2 Einordnung und Forschungsansatz

Der hier vorgestellte Datengenerator wurde zur Unterstützung eines mehrjährigen Forschungsprojektes, das sich am *Design-Science-Ansatz* orientiert, entwickelt. Anders als bei dem *Behavioral-Science-Ansatz* ist es hierbei nicht das Ziel, Theorien zu entwickeln und zu verifizieren, sondern Informationssysteme durch Schaffung innovativer Artefakte zu verbessern. Nach March und Smith (March und Smith 1995, S. 251-266) können dabei die Phasen „Build“ und „Evaluate“ abgegrenzt

werden: In der Build-Phase werden Artefakte zur Lösung definierter Probleme aufgebaut. Mit der sich anschließenden Evaluate-Phase wird versucht, die Ergebnisse methodisch rigoros zu bewerten. Die Evaluate-Phase ist für die Absicherung des Erkenntnisfortschrittes von entscheidender Bedeutung und muss entsprechend sorgfältig angegangen werden (Pries-Heje et al. 2008, S. 255-266). Auf die Unterstützung dieser Phase zielt die hier diskutierte Datengenerator-Lösung.

Im gegebenen Fall werden im Rahmen des übergeordneten Design-Science-Forschungsprojektes DWH-basierte Infrastrukturen zur Entscheidungsunterstützung auf Basis von RFID-Logistikdaten untersucht (Beispiele: Erhöht sich der Ausschuss bei bestimmten Produktmaßen? Wie häufig werden bestimmte, kostspielige Ersatzrouten gewählt? Wie zuverlässig arbeiten die Logistikpartner?).

Der Datengenerator ist dabei die Grundlage für die Evaluation und iterative Weiterentwicklung dieser Lösung: Mit den erzeugten Daten werden Prototypen gespeist, mit denen die konzipierten Anwendungen dargestellt, bestimmte Analyse-szenarios und ihr Nutzen visualisiert sowie daraus abgeleitete, neue Szenarios exploriert werden können. Ohne Zweifel ist die Aussagekraft solcher Prototypen bei einer informationsgetriebenen Technologie wie RFID maßgeblich von der zugrundeliegenden Datenbasis abhängig. Hierin sollten v.a. die wesentlichen Charakteristika der Technologie abgebildet werden können: RFID-Daten versprechen eine umfangreichere, detaillierte und präzisere Datenbereitstellung sowie erweiterte Möglichkeiten, Objekte einzeln zu identifizieren und zu verfolgen (vgl. z. B. Kambil und Brooks 2002).

Der Datengenerator ist für eine möglichst flexible und schnelle Erzeugung entsprechender Datenpools ausgelegt. Das Grunddesign wurde fallunabhängig ausgelegt, sodass es für analoge Evaluationsanforderungen in Forschung und Praxis adaptiert werden kann.

Alternativen zur eigenen Datengenerierung sind v. a. der Einsatz von *Realdaten* und die Durchführung von *Simulationen*. Gegen Realdaten sprechen der begrenzte Zugang (Daten sind vertraulich), die eingeschränkte Verfügbarkeit (Daten liegen aufgrund des Pilotstatus der Projekte noch nicht in brauchbarer Form vor) sowie die erhebliche Menge an benötigten Daten. Simulationen hingegen setzen die Bekanntheit zumindest der Art der eingehenden Parameter voraus. Diese Voraussetzung ist aber in den betrachteten Fällen noch nicht gegeben.

Die Kernanforderungen an den Datengenerator bestanden darin, die iterativ entwickelten Prototypen auf Basis eines zugehörigen Datenmodells mit konsistenten Demonstrationsdaten zu bestücken und dabei bestimmte *Analyseszenarios* zu bedienen. Grundsätzlich wird unter einem Szenario eine Sequenz von Interaktionsschritten verstanden (Pohl 2008, S. 123-124; Balzert 1999, S. 70-78). In dieser Arbeit sind dies die Schritte bei der Auswahl und Aufbereitung von RFID-basierten Logistikdaten, die zur Aufdeckung relevanter Informationen über Entwicklungen (z. B. Zunahme von Schwund) oder Ereignisse (z. B. Lieferverzögerungen auf einer Teilstrecke) in Logistikprozessen sowie zu deren Ergebniskonsequenzen (Zeiten, Kosten, Qualität, Schadstoffe etc.) gewonnen werden sollen. Mit

dem Datengenerator müssen somit für jedes Analyseszenario die erwarteten Informationen in den Datenbestand eingearbeitet werden können. Die Szenarios selbst wurden aus vorgelagerten und parallelen Einzelfallstudien abgeleitet. Weitere Szenarios wurden über Analogiebildung und Generalisierung ergänzt.

3 Datenmodell

Das Design und die Entwicklung sowohl der DWH-basierten Lösung als auch des parallel realisierten Datengenerators erfolgten in mehreren Iterationsschritten zwischen den Monaten Februar 2008 und September 2009. Der verfolgte Ansatz besteht darin, aufbereitete RFID-Leseereignisse mit Daten aus heterogenen Systemen/Datenquellen zusammenzuführen und diese dann mithilfe gängiger Analyse-Tools aufzubereiten. Der Kern der entwickelten Lösung ist ein integriertes, relationales Datenmodell, das konsequent den Anforderungen eines *Star-Schemas* genügt, bei dem relevante Analysedimensionen über (nicht vollnormalisierte) Relationen abgebildet werden, die um eine Faktenrelation mit den Ergebniskennzahlen gruppiert werden (Kimball und Ross 2002, S. 13). Das Datenmodell ist dabei gleichsam eine Komponente des Lösungsentwurfs als auch die Zieldatenstruktur für den Datengenerator. Wie in Kapitel 4 ausgeführt wird, liefert sein Aufbau darüber hinaus Ansatzpunkte für die Datengenerierung.

Besondere Herausforderungen bei der Entwicklung des Datenmodells lagen darin, in einem Star-Schema Mengen-, Zeiten-, Sensor- und Kostendaten konkreten Ist- oder Soll-Lieferprozessen zuzuordnen, eine Aggregation auf mehreren Granularitätsebenen (u.a. Strecken, Organisationseinheiten) zu ermöglichen, mehrstufige Verpackungsvorgänge darzustellen sowie eine flexible Analyse anhand einer Vielzahl von Einflussgrößen abzubilden. Darüber hinaus sollte das Datenmodell sowohl für Near-Time- (z. B. wie viele Objekte sind *momentan* auf einer bestimmten Teilstrecke unterwegs) als auch für historische Analysen (z. B. an welcher Stelle häuft sich Schwund) nutzbar sein. Einen Ausschnitt aus dem entwickelten Modell stellt Abbildung 1 dar.

Die granulare Analyseeinheit im Modell ist die Bewegung eines Objektes („Item“) zwischen zwei Readern – der *Item Transfer*. Ein Item Transfer kann neben einem reinen Transport auch Aktivitäten wie z. B. Verpackung oder Lagerung beinhalten. Die Zuordnung zur physischen Strecke erfolgt über die *Track Section*. Die Art der durchgeführten Aktivität wird dafür über die *Mode of Transformation*-Dimension charakterisiert. Das Modell spiegelt dabei Charakteristika der Technologie RFID wider – insbesondere hinsichtlich der Rückverfolgbarkeit einzelner Objekte, der Auflösung und den Möglichkeiten zur punktgenauen Anreicherung objektbezogener Ereignisse. Zur weiteren Diskussion des Datenmodells und seiner Ableitung vgl. Baars und Sun (2009).

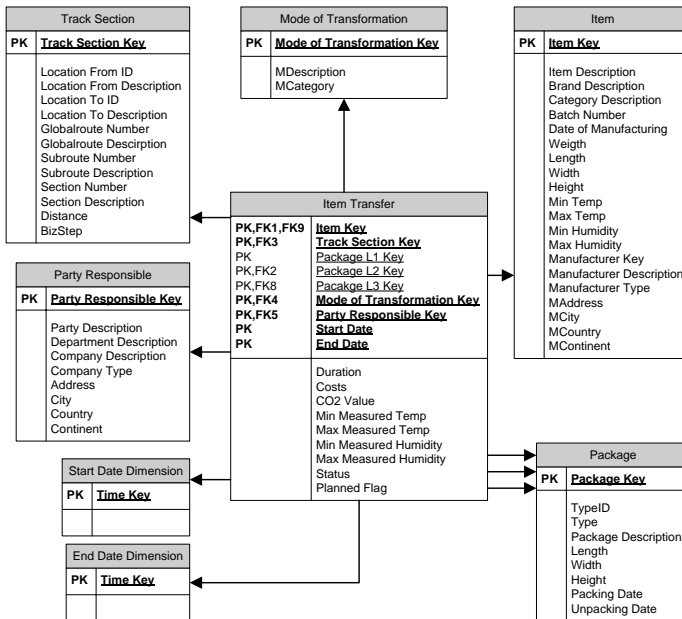


Abbildung 1: Struktur des Datenmodells

4 Struktur des Datengenerators

Die Grundlogik des Datengenerators beruht auf einer Manipulation der Kennzahlenwerte in der Item-Transfer-Faktentabelle. Verändert werden die erzeugten Ausgangsdaten jeweils immer dann, wenn bei einem Item Transfer vorgegebene, ein Analyseszenario repräsentierende Attributwerte in den Dimensionen auftreten (z. B. „Teilstrecke x, Produkt y und Transportmittel z“). Die Szenarios werden somit gewissermaßen in einem „Reverse Engineering“ in eine Grunddatenbasis eingearbeitet, wobei die Dimensionsattribute als potentielle Kosten- oder Nutzentreiber verstanden werden. Dieses Vorgehen ist generell für Datengeneratoren im Star-Schema geeignet und wurde auch bereits entsprechend eingesetzt (vgl. z. B. Kemper und Sopp 2003, S. 226-228). Anders als in diesen Fällen ist eine unmittelbare Umsetzung in Form einer einfachen Bearbeitung der Faktentabelle allerdings unzureichend, da ansonsten elementare Konsistenzanforderungen aus dem Prozesskontext sowie Abhängigkeiten zwischen den diversen Logistikparametern verletzt werden und damit einer der definierten Grundanforderungen an die Datengenerierung nicht genügt wird.

Im Rahmen der weiteren Analyse dieser Problematik wurden insgesamt drei Grundtypen von Analyseszenarios identifiziert, die beliebig kombiniert werden können und logisch wechselseitig aufeinander einwirken. Jeder Szenariotyp be-

schreibt eine spezifische Klasse von Problemsituationen im Logistiknetzwerk und benötigt eine separate Herangehensweise für die Datengenerierung:

1. Der erste Szenariotyp ergibt sich absolut oder stochastisch über eine mathematische Verknüpfung der Dimensions- und Faktenattributwerte – separat für jeden einzelnen Item Transfer.
Beispiele: Auf einem gegebenen Streckenabschnitt besteht bei einem definierten Transportmittel eine Wahrscheinlichkeit von $x\%$, dass ein bestimmter Temperaturmesswert erreicht wird. Bei bestimmten Lieferanten ist in einer bestimmten geographischen Region die Schwundwahrscheinlichkeit $y\%$.
2. Der zweite Szenariotyp resultiert aus der Grundkonfiguration des Liefernetzwerkes. Beispiele: Routingszenarios, bei denen ein bestimmter Weg in absoluter oder stochastischer Abhängigkeit der Attributwerte gewählt wird.
3. Der dritte Szenariotyp wird über die Ausgangsmengen beschrieben. Beispiele: Produktionsengpässe beim Hersteller oder fehlende Ladungsträger am Ort des Bedarfs.

Diese Unterscheidung ist ein unmittelbares Resultat der Prozessbetrachtung und auf jede Form von prozessorientierten Zusammenhängen übertragbar.

Das letztlich entwickelte Lösungsdesign für den Datengenerator ist in Abbildung 2 abgebildet. Hierbei wurde nicht nur darauf geachtet, dass es unmittelbar möglich ist, die Rahmenbedingungen bestimmter Szenarios zu konfigurieren (Strecken, Beteiligte, Art der transportierten Güter, Transportmittel, Soll- und Ist-Strecken, Packvorgänge etc. – die Grundlage für Szenarios des Typs 2), über Inputfaktoren den Einfluss bestimmter Parameterkombinationen auf die Ergebnisse festzulegen (für Typ 1) sowie die Inputmengen zu variieren (für Typ 3). Über die gewählte komponentenorientierte Architektur können weitergehende Modifikationen eingearbeitet werden, etwa zur Ergänzung neuer Kennzahlen, Dimensionsattribute, mathematischer Verknüpfungen oder gar ganzer Dimensionen.

Für die konsistente Abbildung der prozessualen Zusammenhänge wurde ein „simulativer Kern“ entwickelt, der die eigentlichen Materialflüsse nachvollzieht und dabei die Faktendaten generiert. Dieser wurde in der zentralen Komponente *Hauptprozess* gekapselt. Der „Hauptprozess“ berücksichtigt dabei zwei Arten von Inputdaten, die über XML-Dateien bereitgestellt werden:

1. Die Rahmendaten der abzubildenden *Input-Konfiguration*. Dies beinhaltet die Definition und Verknüpfung aller logistischen Strecken, die verantwortlichen Organisationseinheiten, die Art der betrachteten Item Transfers, deren Rahmenbedingungen (z. B. Transportmittel) sowie die Produkt- und Verpackungstypen. Außerdem werden hier die Ausgangsmengen vorgegeben. Auf Basis der Input-Konfiguration wird die Grunddatenstruktur generiert.
2. *Input-Einflussfaktoren*, auf deren Basis die Kennzahlenberechnung erfolgt. So kann beispielsweise angegeben werden, wie sich die Transportkosten berechnen, d.h. welchen Einfluss konkrete Distanzen, Transportmittel und Transportunternehmen usw. auf die Kosten haben.

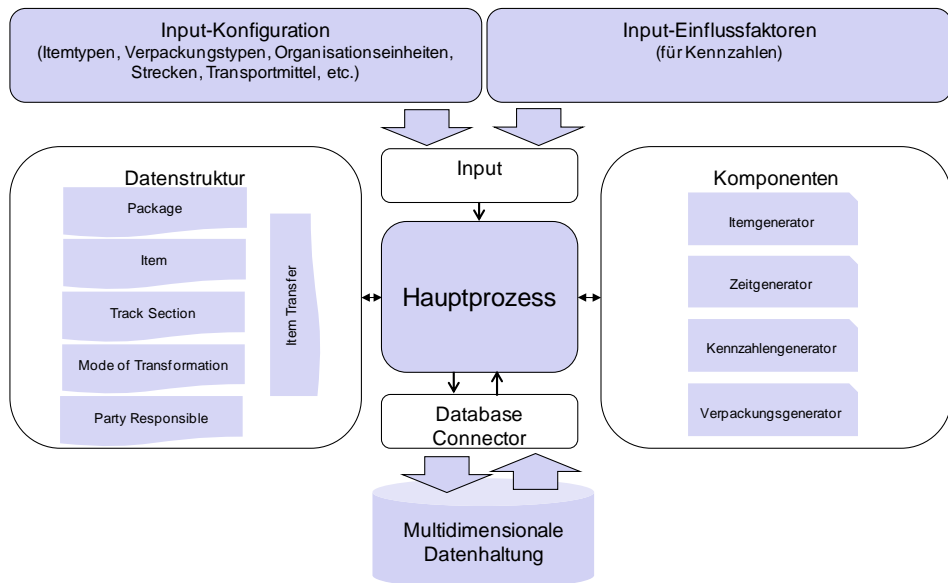


Abbildung 2: Struktur des Datengenerators

Die größte Herausforderung bei der Realisierung des Hauptprozesses ist es, die Prozesse möglichst realitätsnah abzubilden. Die Berücksichtigung der folgenden, aus den Analyseszenarios abgeleiteten Anforderungen hat sich dabei als besonders aufwändig erwiesen:

- Die einzelnen Liefervorgänge müssen sich *zeitlich und inhaltlich korrekt in die Streckenlogik* einfügen. Diese werden über die Input-Konfiguration sowie die Hierarchie in der Dimension „Track Section“ vorgegeben. Die Struktur wird vom Datengenerator erkannt und beim Generieren von Prozessen verfolgt.
- In der Realität sind üblicherweise eher *logistische Netzwerke* als lineare Lieferketten vorzufinden (Pfohl 2004, S. 106-138). Der Datengenerator muss in der Lage sein, in den Daten das Geschehen logistischer Netzwerke abzubilden.
- *Verpackungsprozesse sowie das Verpackungsmanagement* sind die Basis einer Reihe wichtiger RFID-Anwendungen (z. B. Ladungsträgermanagement, Kanban-Prozesse, Kommissionierung). Der Datengenerator kann Ver- und Entpackprozesse erzeugen und unterschiedliche Gegenstände (Produkte oder Verpackungen) auf insgesamt drei Verpackungsebenen verfolgen.

Im Hauptprozess werden die Inputdaten eingelesen und die einzelnen Liefervorgänge nachvollzogen, um so die Faktentabelle aufzubauen. Die eigentliche Berechnung erfolgt mit dem ausgegliederten „*Kennzahlengenerator*“, der für jede Kennzahl auf individuelle mathematische Verknüpfungen der Einflussfaktoren zurückgreift.

Es werden drei Typen von Kennzahlen unterschieden, die jeweils mit getrennten Ansätzen erzeugt werden müssen:

- *Typ A* – Additive Kennzahlen, die kontinuierlich für jeden Item Transfer auf einer Strecke dokumentiert werden, beispielsweise die Dauer, die Kosten, der CO₂-Ausstoß oder der Kraftstoffverbrauch.
- *Typ B* – Nichtadditive Kennzahlen, die nur punktuell festgehalten werden. Hierzu zählen v.a. Sensordaten, bei denen nur Minimal- oder Maximalausprägungen pro Streckenabschnitt gespeichert werden, etwa „Min Measured Temp“, „Max Measured Humidity“ etc.
- *Typ C* – Alphanumerische Kennzahlen, die nominale oder ordinale Sachverhalte messen. Typ C beinhaltet insbesondere den *Status*, der im entwickelten Prototyp die vier möglichen Zustände „completed“, „ongoing“, „lost“, und „rejected“ annehmen kann.

Das grundsätzliche Vorgehen beim Generieren von Kennzahlen kann am Beispiel der Typ-A-Kennzahl Dauer („Duration“) erläutert werden: Zunächst hängt die Berechnung der Dauer von der Art eines betrachteten Vorgangs ab (Mode of Transformation). So unterscheidet sich das mathematische Modell zur Kennzahlenberechnung je nachdem, ob ein Transport, ein Kommissionierungs- oder ein Lagerungsprozess betrachtet wird. Bei einem Transport wird die Dauer von dem betrachteten Streckenabschnitt (Ort, Distanz, etc.), den Transportmitteln (z. B. Bahn oder LKW) und den Organisationseinheiten (z. B. ein bestimmter Lieferant) bestimmt. Bei einem Verpackungsvorgang hingegen ist u.a. auch die Verfügbarkeit von Verpackungsmaterial entscheidend. Schließlich kann die Kennzahl „Duration“ in bestimmten Fällen selbst zur Einflussgröße für andere Kennzahlen werden, z. B. für die Kosten und den CO₂-Ausstoß.

Zur Sicherstellung der Anpassbarkeit wurden neben der Kennzahlengenerierung noch weitere Aufgaben im Datengenerierungsprozess in separate und austauschbare Komponenten ausgegliedert (*Itemgenerator*, *Verpackungsgenerator*, *Zeitgenerator*). Der komponentenorientierte Ansatz erlaubt eine unkomplizierte Erweiterung des Generators, etwa für zusätzliche Attribute oder Verpackungsebenen.

5 Anwendungsbeispiel

Der Einsatz des Datengenerators wird im Folgenden anhand eines ausgewählten Anwendungsfalls demonstriert. Hierbei wird die Lieferung von Hemden chinesischer Hersteller zu deutschen Filialen verfolgt. Die Lieferstrecke kann auf drei Hierarchieebenen betrachtet werden (vgl. Abbildung 3). Zwei Produkte („Markenhemd“ und „Noname Hemd“) werden an zwei Standorten produziert, zusammen verpackt und wahlweise per Schiff oder per Flugzeug nach Deutschland transportiert. Nach dem Entpacken in einem Warenverteilzentrum werden die Hemden ausgeliefert. In den erzeugten Daten wird somit ein logistisches Netzwerk mit ei-

nem indirektem Güterfluss und zwei Verpackungsebenen sowie entsprechenden Auflösungspunkten („L3 Tianjing Hafen Einpacken“ und „L3 Frankfurt Entpacken“) und Konzentrationspunkten („L3 Tianjing Hafen Einpacken“ und „L3 Hamburg Entpacken“) (Pfohl 2004, S. 106-138) abgebildet.

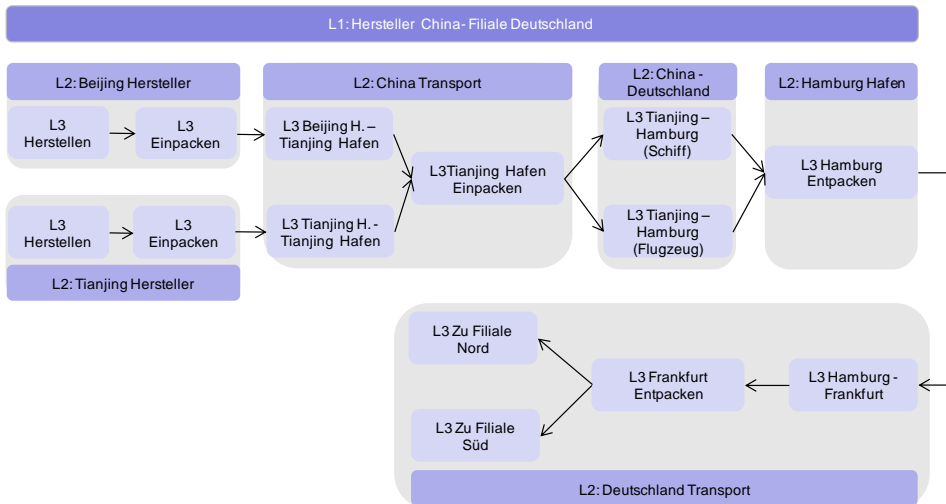


Abbildung 3: Beschreibung des Anwendungsfalls

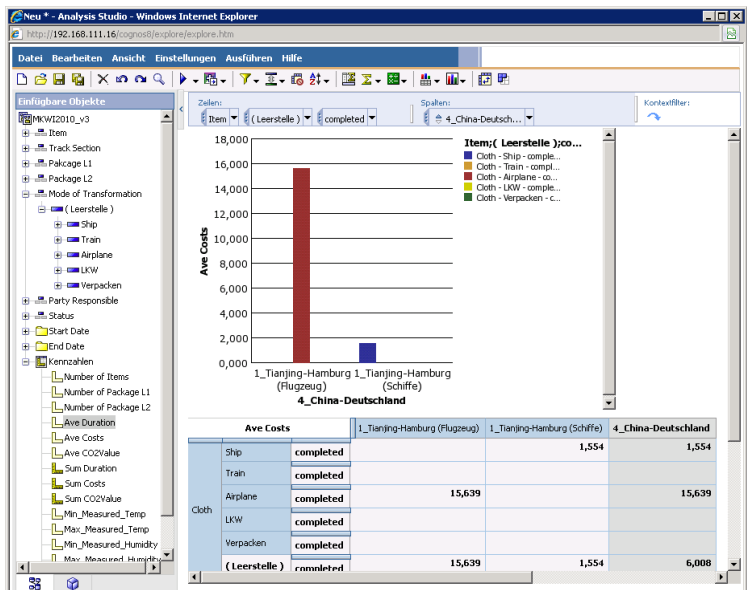


Abbildung 4: Kosten für die Strecke "China-Deutschland" auf dritter Ebene

Für den beschriebenen Fall wurde eine Testdatenbasis aus 20.200 Items für den Zeitraum vom 01.01.2009 bis zum 28.02.2009 erstellt. Insgesamt wurden ca. 150.000 Datensätze und ein Datenvolumen von über 80 MB generiert. Die Daten wurden anschließend in das OLAP-Analyse-Frontend Cognos 8.3 importiert. Vgl. Abbildung 4: In der linken Spalte befinden sich die selektierbaren Dimensionen und Kennzahlen, in der rechten wird das Analyseergebnis graphisch und tabellarisch aufbereitet. Im Screenshot werden die durchschnittlichen Ist-Transportkosten für die Lieferalternativen „Tianjing-Hamburg (Flugzeugtransport)“ und „Tianjing-Hamburg (Schifftransport)“ verglichen.

Number of Items		1_Beijing Hersteller	2_Tianjing Hersteller	3_China Transport	4_China-Deutschland	5_Hamburg Hafen	6_Deutschland Transport	China-Germany
Markenhemd	completed	9306	0	7405	5479	4570	3985	9306
	ongoing	1532	0	1213	510	467	1019	4741
	rejected	0	0	0	0	215	0	215
	Weitere & nicht sichtbar							
Status		10100	0	8568	6653	5479	4570	10100
Noname Hemd	completed	0	9199	6587	4738	3833	3275	9199
	ongoing	0	1776	1101	470	349	885	4581
	rejected	0	0	0	0	303	0	303
	Weitere & nicht sichtbar							
Status		0	10100	7750	5817	4738	3833	10100
Cloth	completed	9306	9199	13992	10217	8403	7260	18505
	ongoing	1532	1776	2314	980	816	1904	9322
	rejected	0	0	0	0	518	0	518
	Weitere & nicht sichtbar							
Status		10100	10100	16318	12470	10217	8403	20200

Abbildung 5: Anzahl der Items für die Produkte auf zweiter Detaillierungsebene

Einen anderen Ausschnitt derselben Datenbasis zeigt Abbildung 5. Hier wird die Kennzahl „Number of Items“ auf der zweiten Detaillierungsebene aufbereitet, wobei der unterschiedliche Lieferstatus der Items im Liefernetzwerk dargestellt wird. In der Tabelle ist deutlich zu sehen, dass die Zahl an Items im Hamburger Hafen deutlich reduziert wird. Dies ist die Folge eines Input-Szenarios, das bei dem Noname-Produkt aufgrund von Qualitätsmängeln zu einer hohen „rejected“ Quote führt – die Hemden bestehen die Eingangsqualitätsprüfung nicht.

6 Diskussion

Wie die Ausführungen verdeutlichen, lagen wesentliche Herausforderungen beim Design des Datengenerators in der Gewährleistung einer hohen Flexibilität bei einer gleichzeitig möglichst weitgehenden Aufrechterhaltung von Konsistenz und Plausibilität. In dem übergeordneten Forschungsvorhaben ist als nächster Schritt vorgesehen, mit dem Generator erzeugte Daten im Kontext von Anwenderinterviews zu RFID-basierten BI-Systemen einzusetzen und damit den Evaluations-schritt im übergeordneten Design-Science-Projekt einzuleiten. Dies führt auch zu einer indirekten Evaluation des Generators selbst. Hierbei ist insbesondere syste-

matisch zu prüfen, ob im Generator eine Berücksichtigung zusätzlicher Interdependenzen, Strukturen und Szenarios erforderlich ist

Weiterer Untersuchungsbedarf besteht bezüglich der Frage, ob der Datengenerator auch in einem Praxisumfeld als Werkzeug für die Informationsbedarfsanalyse eingesetzt werden kann. Daneben sind auch Optionen zu prüfen, das entwickelte Werkzeug in grundsätzlich anderen *Domänen* einzusetzen, etwa in der Lehre und Weiterbildung (z. B. zur Vermittlung von Logistik- und Analysekenntnissen). Forschungspotentiale ergeben sich schließlich auch hinsichtlich der Adaptierbarkeit des Ansatzes auf andere *Anwendungsgegenstände*, wobei mit dem aufgezeigten Design vor allem solche Bereiche in Frage kommen, in denen komplexe Prozessstrukturen im Mittelpunkt stehen. Grundsätzlich sollte der Idee des Einsatzes von Datengeneratoren im speziellen und von Generatoren im Allgemeinen im Kontext des Design Science stärkere Aufmerksamkeit gewidmet werden.

Literatur

- Baars H, Gille D, Strüker J (2009) Evaluation of RFID applications for logistics: a framework for identifying, forecasting and assessing benefits. *European Journal of Information Systems (EJIS)* 18(6). *In Druck*.
- Baars H, Kemper H-G, Lasi H, Siegel M (2008) Combining RFID Technology and Business Intelligence for Supply Chain Optimization – Scenarios for Retail Logistics. In: *Proceedings of the 41. Hawaii International Conference on System-Sciences (HICCS-41)*, Hawaii.
- Baars H, Sun X (2009) Multidimensional Analysis of RFID Data in Logistics. In: *Proceedings of the 42. Hawaii International Conference on System Sciences (HICCS-42)*, Hawaii.
- Balzert H (1999) *Lehrbuch der Objektmodellierung. Analyse und Entwurf*. Spektrum Akademischer Verlag, Heidelberg, Berlin.
- Brewer A, Sloan N (1999) Intelligent Tracking in Manufacturing. *Journal of Intelligent Manufacturing* 10(3-4):245-250.
- Chawathe SS, Krishnamurthy V, Ramachandran S, Sarma S (2004) Managing RFID Data. In: *Proceedings of the 30th VLDB Conference*, Toronto, Canada.
- Clauberg R (2004) RFID and Sensor Networks. RFID Workshop, University of St. Gallen, Switzerland.
- Curtin J, Kauffman RJ, Riggins FJ (2007) Making the ‘MOST’ out of RFID technology: a research agenda for the study of the adoption, usage and impact of RFID. *Information Technology and Management*. 8(2):87-110.

- EPCglobal (2009) EPCglobal-Standards. <http://www.epcglobalinc.org/standards/>
Abruf am 2009-10-01.
- Günther O, Kletti W und Kubach U (2008) RFID in Manufacturing. Springer, Berlin, Heidelberg.
- Hevner AR, March ST, Park J (2004) Design Science in Information Systems Research. *MIS Quarterly*, 28(1):75-105.
- Hoffman W (2008) RFID: Play or Pay. *Traffic World*, 272(5):16.
- Ilic A, Anderson T, Michahelles F (2009) EPCIS-based Supply Chain Visualization Tool. White Paper, Auto-ID Labs, St. Gallen.
- Kambil A, Brooks JD (2002) Auto-ID Across the Value Chain: From Dramatic Potential to Greater Efficiency & Profit. White Paper, Auto-ID Labs St. Gallen.
- Kemper, Sopp (2003) Computergestützte Entwicklung betriebswirtschaftlicher Fallstudien. In: *WISU*, 32(2), S. 224-229.
- Kimball R, Ross M (2002) *The Data Warehouse Toolkit*. Wiley, USA.
- Kreller B, Hartmann J (2001) The Field Trial Scenario Of An Inter-Modal, End-To-End And Real-Time Tracking And Tracing System. In: 8th World Congress on Intelligent Transport Systems, Sydney, Australia.
- Laubacher R, Kothari S, Malone T, Subirana B (2005) What is RFID Worth to Your Company? MIT Center for eBusiness Research Brief, 7(2):1-6.
- Lee, Y.M., Cheng, F., Leung, Y.T. (2004) Exploring the impact of RFID on Supply Chain Dynamics. In: *Proceedings of the 2004 Winter Simulation Conference (WSC '04)*, S. 1145-1152, Washington DC (USA).
- March ST, Smith G (1995) Design and Natural Science Research on Information Technology. *Decision Support Systems*, 15(4):251-266.
- METRO Group (2004) RFID Uncovering the Value Applying RFID within the Retail and Consumer Package Goods Value Chain. Arbeitsbericht.
- Pfohl H-C (2004) *Logistikmanagement - Konzeption und Funktionen*. Springer, Berlin
- Pohl, K. (2008). *Requirements Engineering - Grundlagen, Prinzipien, Techniken*. dpunkt. Verlag, Heidelberg.
- Pries-Heje J, Baskerville R, Venable J (2008) Strategies for design science research evaluation. In: *Proceedings of the 16th European Conference on Information Systems (ECIS)*, Galway, Ireland.
- Strüker J, Gille D, Faupel T (2008) RFID Report 2008. *VDI Nachrichten*.