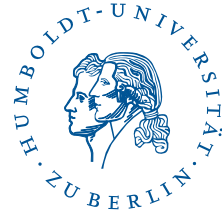


HUMBOLDT-UNIVERSITÄT ZU BERLIN



Exposé einer Diplomarbeit zur Enterprise Application Integration in Krankenhausinformationssystemen

Daniel Neumann¹
September 2009

Betreuung:

Jens Zaedow, Helios IT²

Prof. Dr. Ulf Leser, Institut für Informatik, Humboldt-Universität zu
Berlin

¹dneumann (at) informatik.hu-berlin.de

²jens.zaedow (at) helios-kliniken.de

1 Hintergrund

Ein Krankenhausinformationssystem (KIS) ist ein an seinen Einsatzort angepasstes aus einzelnen Informationssystemen (IS) aufgebautes, einem evolutionären Prozess unterworfenen [Lan03], heterogenes Gesamtsystem in der Domäne eines Krankenhauses [WKKP05]. Die Aufgaben des KIS sind die Patientendatenverwaltung, die Administration des Krankenhausbetriebs sowie die medizinische Dokumentation und Organisation. Ein voll integriertes KIS existiert nicht [Pro01], daher müssen die einzelnen Informationssysteme miteinander kommunizieren, um Informationen zu erhalten und weiterzugeben. Zu diesem Zweck wurden medizinische Kommunikationsprotokolle entwickelt und standardisiert: HL7³ [HL788] [HL790] [HL794] [HL797] hat sich durchgesetzt, da es für alle Aufgaben im KIS eingesetzt werden kann. Im Speziellen für die Diagnostik ist DICOM [BHPVS97] zu nennen. Ein weiteres in der Praxis eingesetztes Nachrichtenformat ist SAP-HCM. Es ist Bestandteil des für SAPs Branchenlösung Krankenhaus IS-H entwickelten IS-H Communication Module (IS-HCM) [SAP06].

In einem KIS existieren ältere neben neu angeschafften Informationssystemen. Dies bedeutet eine ständige Schnittstellenpflege, um die Kommunikation zwischen ihnen zu gewährleisten. Dem Konzept der Enterprise Application Integration (EAI) [CHKT06] folgend, unter anderem die Schnittstellenwartung und -einrichtung zu erleichtern, wird ein Kommunikationsserver (KS) eingesetzt. Es soll kein bereits existierender Datenbestand (nach der Erhebung der Daten) integriert werden, sondern der KS ist in den Prozess der Datenproduktion integriert [LN07] und verteilt die ihm zugeführten Informationen an die jeweiligen Informationssysteme.

Eine Daten- bzw. Informationsintegration der Patientendaten findet nicht statt, ebenso keine funktionale Integration, Präsentationsintegration oder Ablaufintegration [LS03], da ein KS nicht die semantische Heterogenität der kommunizierenden Teilsysteme auflöst. Somit entsteht durch ein KS kein Krankenhausinformationssystem, sondern „lediglich“ ein Krankenhauskommunikationssystem [Pro01].

³Health Level 7; Das erklärte Ziel von HL7 ist es, „einen Kommunikationsstandard für das Gesundheitswesen anzubieten, der es ermöglicht, die sonst nötigen Arbeiten an eigenen Schnittstellenentwicklungen und Programmwartungen zu eliminieren oder diese wesentlich zu verringern.“ (Vgl. [HL788][HL790][HL794][HL797])

2 Problemstellung

Der KS stellt die „Kommunikationsdrehscheibe“ [WKKP05] für die Kommunikation der einzelnen verteilten IS im KIS dar. Seine Aufgabe ist die Transformation von Nachrichten zwischen HL7 und dessen verschiedenen Versionen sowie zwischen proprietären Nachrichtenformaten wie SAP-HCM. Transformation bedeutet, eingehende Nachrichten in jenes Nachrichtenformat zu übersetzen, welches das eigentliche Zielsystem der Nachricht versteht. Außerdem kann zwischen socket- und dateibasierter Kommunikation [Lan03] gewandelt werden. Wer Empfänger einer Nachricht ist, ist in einem KIS meist eindeutig: Es existiert zu jedem Sender genau ein Empfänger. Einzig das führende IS verteilt ADT -Daten⁴ [HL798] in einer 1:n-Beziehung. Die Kommunikation ist immer ereignisgesteuert: Tritt eine neue Situation ein, so kann diese interrogativ, indikativisch oder imperativ behandelt werden⁵. Somit sind typische Nachrichten, die ein Kommunikationsserver verarbeiten muss, der Art:

- ADT
- Aufträge
- Anfragen
- Befundberichterstattung
- Leistungsdatenübermittlung

Auch wenn die Wege der Nachrichten klar definiert sind und damit auch die Transformationsaufgaben des KS, bedarf es bei der Programmierung der Transformationsagenten einer besonderen Sorgfalt: Falsche, fehlerhafte oder fehlende Informationen können lebensbedrohend sein. Ein Verfahren wie die (semi-)automatisierte Suche nach Übereinstimmungen von Attributwerten in 2 Nachrichten verschiedenen

⁴Stamm- und Bewegungsdaten zur Aufnahme („**A**dmission“), Entlassung („**D**ischarge“) und Verlegung („**T**ransfer“) von Patienten

⁵Interrogativ bedeutet, dass eine Situation Information benötigt, die über eine Anfrage erhalten werden kann (engl. query). Indikativ steht für ein eingetretenes Ereignis, dessen Information an anderer Stelle verfügbar gemacht werden soll (engl. unsolicited update). Der Imperativ fordert auf, mit der Übermittlung der Information eine bestimmte Handlung auszuführen (engl. order) [Sch00]

Nachrichtenformats zur Unterstützung bei der Programmierung des Transformationsagenten muss „bei Zweifel“ dem Entwickler dies anzeigen. Diese Suche wird auch instanzbasiertes Schema Matching genannt.

3 Zielstellung

Wie kann dem Entwickler durch Schema Matching [Nau04] [Nau06] ein Teil der manuellen Zuordnungsarbeit abgenommen werden? Welche Matching-Verfahren bieten sich hier an [RB01]? Ist eine automatisierte Zuordnung tatsächlich in dieser Domäne praktikabel?

Schema-Matching-Methoden werden klassifiziert und im Detail jene betrachtet, die geeignet scheinen. Als guter Kandidat für die automatisierte Suche nach Übereinstimmungen soll instanz-basiertes Schema Matching anhand von Duplikaten [LN07] [BN05] mit einer Implementierung überprüft werden. Diese Methode ist geeignet, weil sie die nötige Sorgfalt prinzipbedingt erfüllt, wenn Beispielnachrichten vorliegen und diese vorab manipuliert werden können: Nur Übereinstimmungen in den Attributwerten führen zu vorgeschlagenen Korrespondenzen. Welche Möglichkeiten existieren, wenn die Beispielnachrichten fehlen, soll ebenfalls diskutiert werden.

4 Vorgehen

Die Diplomarbeit wird zunächst das Thema EAI behandeln und dabei die später in dieser Arbeit genutzte EAI-Lösung Sun Java Composite Application Platform Suite (Java CAPS) vorstellen. Da Sun Java CAPS Nachfolger des eGate Integrators ist, wird sowohl die von eGate verwendete Repository-basierende Entwicklung betrachtet, als auch die von Sun entwickelte Java Business Integration (JBI). Das Produkt eGate begann als Kommunikationsserver in Krankenhäusern, spätestens seit der Übernahme durch Sun ist aber die Domäne Gesundheitswesen nicht mehr alleiniges Einsatzgebiet. Welches Werkzeug, ob Repository-based oder JBI, im Hinblick auf die Zielstellung in Frage kommt, wird in einem Vergleich am Ende dieses Kapitels dargestellt.

Daran anschließend wird der Einsatz von EAI im Gesundheitswesen, speziell bei der Ausbildung von Krankenhausinformationssystemen, diskutiert:

- Anpassung an klinische Prozesse
- Der Kommunikationsserver als Enterprise Service Bus (ESB)
- Benutzung von medizinischen Nachrichtenformaten.

Außerdem werden Alternativen zur EAI, die den Gedanken der Integration weiter als bis zur Prozessintegration tragen, aufgeführt (siehe dazu Abschnitt 1).

Wie in der Problemstellung formuliert, werden Mappings, also Transformationsregeln, zwischen verschiedenen Nachrichtenformaten entweder per Hand erstellt oder semiautomatisch [LN07] per Schema Matching ermittelt. Semiautomatisch bedeutet, dass das Schema Matching automatisch abläuft und Vorschläge für Korrespondenzen generiert, die dann vom Domänenexperten verifiziert werden müssen. Methoden des Schema Matching werden vorgestellt: Kombinierte Ansätze sowie individuelle Ansätze, im Hinblick auf die Zielstellung werden schemabasierte und instanzbasierte individuelle Ansätze verglichen.

Die in dieser Arbeit untersuchten Nachrichtenstandards sind HL7 und SAP-HCM. HL7 ermöglicht durch die große Anzahl an Ereignistypen das Erhalten des Kontexts, in dem die Nachricht gesendet wurde, aufgrund von standardisierten Kodierungsregeln eine exakte Syntax und durch die ereignisgesteuerte Ausführung ein Erfüllen der Pragmatik [Neu08]. Dagegen ist mit SAP-HCM [SAP99] eine kundenspezifische Anpassung von Nachrichtensegmenten möglich⁶. Der Vorteil ist eine individuelle Gestaltung der Nachrichten, der Nachteil besteht darin, dass der Empfänger (im SAP-Jargon als Transceiver bezeichnet) die Syntax und Semantik erst „erlernen“ muss.

Die verwendeten Versionen sind die derzeit am häufigsten eingesetzten: HL7 2.x sowie SAP-HCM 1.2. Die Schachtelungstiefe einer Nachricht ist 3, die Anzahl der Attribute, Wiederholungsfelder unbeachtet, ist abhängig vom Ereignis und liegt im dreistelligen Bereich. Für ADT-Daten unterscheidet HL7 in Version 2.4 allein 62 Ereignisse [Lan03], SAP-HCM fasst diese in 4 Ereignisse zusammen [SAP99].

⁶Im HL7-Standard als sogenannte Z-Segmente realisiert. Sie erlauben die Definition eigener Nachrichtensegmente, angehängt am Ende einer Standard-HL7-Nachricht

Um die Schwierigkeiten beim Finden von Korrespondenzen aufzuzeigen, werden Beispielnachrichten herangezogen. Eine HL7-Nachricht besteht aus einzelnen Nachrichtensegmenten. Unabhängig vom Ereignistyp beginnt die Nachricht mit dem Nachrichtenkopf. Darin enthalten ist eine Kodierung für das Ereignis (zum Beispiel A04 für Registrierung Patient). Diese Kodierung definiert, welches Feld in der Nachricht welchen Inhalt besitzt. Eine hierarchische Struktur ist zu erkennen, die einzelnen Nachrichtensegmente beginnen mit einem 3-Buchstaben-Kürzel (zum Beispiel MSH für Message Header, Nachrichtenkopf). Attributnamen sind dagegen nur in der Dokumentation [HL798] zu finden⁷. Allein der Attributwert gibt somit Aufschluss über die Art der Nachricht.

Davon ausgehend, dass die Kodierung eindeutig ist und in der Dokumentation zum Standard somit wiederzufinden ist, lassen sich mit einem HL7-Editor, der diesen Standard enthält, zu den in der Nachricht enthaltenen Attributwerten die passenden Attribute anzeigen. Für SAP-HCM existiert ebenfalls eine Dokumentation [SAP99], die für HL7 genannten Schwierigkeiten gelten auch hier. Somit gilt: Nur mit detaillierten Kenntnissen der Nachrichtentypen und der Domäne selbst, lässt sich ein Matcher erstellen.

Sollen Nachrichten zwischen HL7 und SAP-HCM transformiert werden und dem Domänenexperten/Entwickler dafür Vorschläge für Korrespondenzen unterbreitet werden, so ist der Weg über Beispielnachrichten zu nehmen: Attributwerte der eingehenden und ausgehenden Nachricht sollten möglichst übereinstimmen, damit ein Matcher eine Korrespondenz erkennen kann. Instanzbasiertes Schema-Matching anhand Duplikaten erfüllt diesen Anspruch.

Daher soll diese Methode implementiert und in der Praxis auf Tauglichkeit überprüft werden. Als Entwicklungsumgebung wird Sun Netbeans verwendet, eine Entwicklungsplattform für den Sun Java CAPS. Die Schema-Matching-Methode selbst wird als ein Skript für Netbeans programmiert.

Evaluert wird anhand dreier Nachrichtentypen (ADT - Patientenstammdaten, ORM - Aufträge, ORU - Befunde). Für bereits implementierte Kommunikationswege kann auf Realweltnachrichten aus dem Krankenhausbetrieb zurückgegriffen werden. Bei

⁷HL7 Version 3 benutzt XML, um u.a. Nachteile in Version 2.x wie fehlende Attributnamen oder schwere Lesbarkeit zu beseitigen

neu zu implementierenden Wegen sind wenige Beispielnachrichten der Hersteller der beteiligten Informationssysteme vorhanden. Die von der Methode gefundenen Mappings werden von einem Domänenexperten auf Korrektheit überprüft. Maßzahlen für die Prüfung sind richtig oder falsch gemappt sowie richtiges oder falsches fehlendes Mapping.

Das Fazit wird zusammenfassen, ob die Idee, den Experten bei der Programmierung der Transformationsagenten mit Schema-Matching-Techniken zu unterstützen, im Umfeld des Gesundheitswesens praktikabel ist. Welches sind die wesentlichen Faktoren, von denen ein Gelingen abhängt? Sind in der Praxis Probleme entstanden, die nicht vorhersehbar waren? Mit einem Ausblick soll die Arbeit schließen.

Literatur

- [BHPVS97] BIDGOOD, W.D. ; HORII, S.C. ; PRIOR, F.W. ; VAN SYCKLE, D.E.: *Understanding and using DICOM, the data interchange standard for biomedical imaging*. <http://www.j-amia.org/cgi/content/full/4/3/199>. Version: 1997
- [BN05] BILKE, A. ; NAUMANN, F.: Schema matching using duplicates. In: *Proceedings of the 21st International Conference on Data Engineering (ICDE 2005)*, 2005, 69–80
- [CHKT06] CONRAD, S. ; HASSELBRING, W. ; KOSCHEL, A. ; TRITSCH, R.: *Enterprise Application Integration: Grundlagen, Konzepte, Entwurfsmuster, Praxisbeispiele*. Elsevier, Spektrum, Akad. Verl., 2006
- [HL788] HL7: *Health Level Seven Standards, Version 2.0: An Application Protocol for Electronic Data Exchange in Health Care Environment*. Philadelphia : Health Level-7, Inc, 1988
- [HL790] HL7: *Health Industry Level Seven Interface Standards, Version 2.1*. Chicago : Health Level-7, Inc, 1990
- [HL794] HL7: *Health Level Seven: an application protocol for electronic data exchange in health care environments, version 2.2*. Ann Arbor : Health Level-7, Inc, 1994
- [HL797] HL7: *Health Level Seven, Version 2.3*. Ann Arbor : Health Level-7, Inc, 1997
- [HL798] HL7: *Health Level Seven Implementation Support Guide for HL7 Standard Version 2.3*. Ann Arbor : Health Level-7, Inc, 1998 <http://www.hl7.org/Special/IG/final.pdf>
- [Lan03] LANGE, M.: *Semantische und syntaktische Integration von konkurrierenden Informationssystemen in der Medizin*, Münster (Westfalen), Diss., 2003

- [LN07] LESER, U. ; NAUMANN, F.: *Informationsintegration Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen*. dpunkt-Verl., 2007
- [LS03] LENZ, R. ; SCHMÜCKER, P: Stete Weiterentwicklung nötig - Integration heterogener IT-Systeme im Krankenhaus. In: *Der Klinikarzt* 32 (2003), November, Nr. 11, 385-390. <http://www.thieme-connect.de/ejournals/pdf/klinikarzt/doi/10.1055/s-2003-44532.pdf>
- [Nau04] NAUMANN, Felix: *Informationsintegration Antrittsvorlesung*. <http://edoc.hu-berlin.de/docviews/abstract.php?id=25188>.
Version: Mai 2004. – [Online: Stand 28.06.2009 16:58Uhr]
- [Nau06] NAUMANN, Felix: *Vorlesung Informationsintegration: Schema Matching*. Januar 2006
- [Neu08] NEUMANN, Daniel: *Studienarbeit zur Integration von Krankenhausinformationssystemen am Beispiel von SAP und FileMaker*. Dezember 2008
- [Pro01] PROKOSCH, Hans U.: KAS, KIS, EKA, EPA, EGA, E-Health: Ein Plädoyer gegen die babylonische Begriffsverwirrung in der Medizinischen Informatik. In: *Informatik, Biometrie und Epidemiologie in Medizin und Biologie* 32 (2001), Nr. 4, S. 371–382
- [RB01] RAHM, E. ; BERNSTEIN, P.A.: A survey of approaches to automatic schema matching. In: *The VLDB Journal The International Journal on Very Large Data Bases* 10 (2001), Dezember, Nr. 4, 334–350. <http://dx.doi.org/10.1007/s007780100057>. – DOI 10.1007/s007780100057
- [SAP99] SAP®: *IS-HCM-Handbuch für externe Systempartner*. Walldorf : SAP®, 1999. – 127 S.
- [SAP06] SAP®: *IS-H/IS-H*MED*. http://help.sap.com/saphelp_ish472/helpdata/DE/59/1c86353e13d558e10000009b38f889/frameset.htm.
Version: April 2006

- [Sch00] SCHADOW, Gunther: Krankenhauskommunikation mit HL7. In: *Berichte aus der medizinischen Informatik und Bioinformatik*. Aachen : Shaker, 2000, S. 360. – Zugl.: Berlin, Freie Univ., Diss., 1999
- [WKKP05] WENTZ, B. ; KRASKA, D. ; KNISPEL, S. ; PROKOSCH, HU: 10 Jahre Erlanger Kommunikationsdrehscheibe - der Weg zu einer zukunftssicheren Integrationsplattform. In: *50. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (gmds)*. Freiburg im Breisgau : gmds, September 2005