

Umsetzung von Semantik-Mappings unter HL7-basierten IHE-Profilen zur Patientenidentifikation

Implementing Semantical Mappings among HL7-based IHE Profiles for Patient Identification

Marek VÁCLAVÍK and Dmytro RUD
*InterComponentWare AG, R&D Professional Gate
Altrottstr. 31, 69190 Walldorf (Baden)
Marek.Vaclavik@icw.de, Dmytro.Rud@icw.de*

Zusammenfassung. Die Interoperabilitätsinitiative „Integrating Healthcare Enterprise“ spezifiziert für Patientenidentifikation mehrere Integrationsprofile mit überlappender Funktionalität und ähnlichen Datenmodellen. Deren gegenseitige semantische Kompatibilität wurde im Zuge einer Schnittstellenerweiterung eines „Master Patient Index“ untersucht. Die Abbildung zwischen den Modellen von HL7 2.x und HL7 Version 3 stößt auf einige Schwierigkeiten. Diese Abbildung ist trotz Einschränkungen jedoch erreichbar und somit der vorteilhafte Bridging-Ansatz anwendbar.

Abstract. The interoperability initiative „Integrating Healthcare Enterprise“ specifies multiple integration profiles related to patient identification, which show overlaps in functionality and similarities in the data model. Their semantic compatibility has been investigated while extending the data exchange interface of a master patient index application. The mapping between HL7 2.x and HL7 Version 3 poses a few difficulties. In spite of limitations this mapping is achievable and the advantageous bridging approach applicable.

Keywords Semantic interoperability, HL7, IHE, PIX, PDQ, XCPD.

Einleitung

Die internationale Initiative „Integrating Healthcare Enterprise“ (IHE) [1] mit dem Ziel, die Interoperabilität von Systemen auf Basis von Standards zu verbessern, trifft insbesondere bei den Softwareherstellern auf außerordentliche Akzeptanz. Dies liegt vor allem in der Praxisnähe des Standardisierungsprozesses begründet [2], in dessen Mittelpunkt die regelmäßigen Interoperabilitätstests (Connectathons) stehen.

Mit den Integrationsprofilen des technischen Rahmenwerks „IT Infrastructure“ (ITI) [3] zeigt die IHE Lösungen beispielsweise für den intersektoralen Datenaustausch oder telemedizinische Anwendungen auf. Solche Szenarien erfordern einen eindeutigen und einrichtungsübergreifenden Patientenbezug. Dessen Herstellung wird durch die Integrationsprofile „Patient Identifier Cross-Reference“ (PIX) und „Patient Demographics Query“ (PDQ) spezifiziert, die auf den Protokollen HL7 v2.3.1 und HL7 v2.5 basieren. Zwei neuere Profile, PIX v3 und PDQ v3 [4], dienen dem gleichen Zweck, jedoch unter Verwendung von HL7 Version 3 (HL7 v3, Normative Edition 2008). Die serverseitigen Akteure PIX/PIX v3 Manager sowie PDQ/PDQ v3 Patient Demographics Supplier werden in der Regel als Master-Patient-Index-Anwendungen (MPI) implementiert.

Die Patientensuche über die Grenzen des lokalen Rechnernetzes (Community) hinweg wird im Profil „Cross-Community Patient Discovery“ (XCPD) [5] spezifiziert, welches ebenfalls auf der oben genannten Edition von HL7 v3 basiert. Eine XCPD-Abfrage ermittelt, in welchen Communities relevante Patientendaten gespeichert sind. Dabei werden die Treffer zuerst bei einzelnen Communities über deren jeweiligen Responding Gateway abgefragt und anschließend durch einen Initiating Gateway Community übergreifend aggregiert.

1. Zielsetzung

In der Ausgangslage unterstützte eine MPI-Anwendung aus dem Hause InterComponentWare AG die serverseitigen Akteure PIX Manager und PDQ Supplier. Anforderungen aus Kundenprojekten machten eine Erweiterung um die Akteure PIXv3 Manager, PDQv3 Supplier und XCPD Responding Gateway erforderlich. Bei der Umsetzung hatten Aufwandsminimierung und Wiederverwendbarkeit eindeutige Priorität.

Aus diesem Grund bot sich bei der evolutionären Ergänzung des Satzes unterstützter Profile ein Ansatz an, bei dem die Implementierung der fachlichen Logik nur einmal erfolgt, und lediglich die Schnittstellen jedes neuen Akteurs neu zu entwickeln sind, sowie Routinen zur Transformation zwischen seinem Datenformat und dem Datenformat eines bereits vorhandenen Akteurs mit äquivalenter Funktionalität.

Um eine effiziente Implementierung dieses Ansatzes zu ermöglichen, sollten insbesondere semantische Unterschiede zwischen PIX Manager und PIXv3 Manager, PDQ Supplier und PDQv3 Supplier sowie zwischen XCPD Responding Gateway und PDQv3 Supplier betrachtet werden.

Die technische Zulänglichkeit und fachliche Tauglichkeit der Lösung war durch eine Anbindung an kompatible Systeme anderer Hersteller zu verifizieren.

2. Methoden

Unter Berücksichtigung existierender Vorarbeiten wurden Transformationsregeln formuliert und mit Hilfe des Java-basierten Integrationsframeworks IPF [6] der Open eHealth Foundation umgesetzt, wobei der herstellerneutrale Teil der Lösung als Code-Contribution im IPF aufgesetzt wurde.

Beim Design der Lösung wurden Enterprise Integration Patterns [7] eingesetzt. Die ausgewählten Integrationsmuster dienen der Wiederverwendung bestehender Schnittstellen — vollständig nicht-intrusiv („Message Bridge“) oder mit Anpassungsbedarf auf einer Seite der Kommunikation („Message Adapter“).

Die Anwendung dieser Muster erforderte die Abbildung (engl.: mapping) der Semantik zwischen unterschiedlichen Datenmodellen. Die dabei festgestellten Unterschiede wurden zusammen mit den Erfahrungen aus dem praktischen Einsatz der Lösung analysiert.

3. Ergebnisse

Abbildung 1 veranschaulicht die Gesamtarchitektur der Implementierung. Die Transformation zwischen PIX/PDQ und PIXv3/PDQv3 (1) wurde in einer eigenständigen Anwendung gemäß dem Integrationsmuster „Message Bridge“ umgesetzt. Die Transformation zwischen XCPD und PDQv3 (2) übernimmt ein im Responding Gateway eingebauter „Message Adapter“.

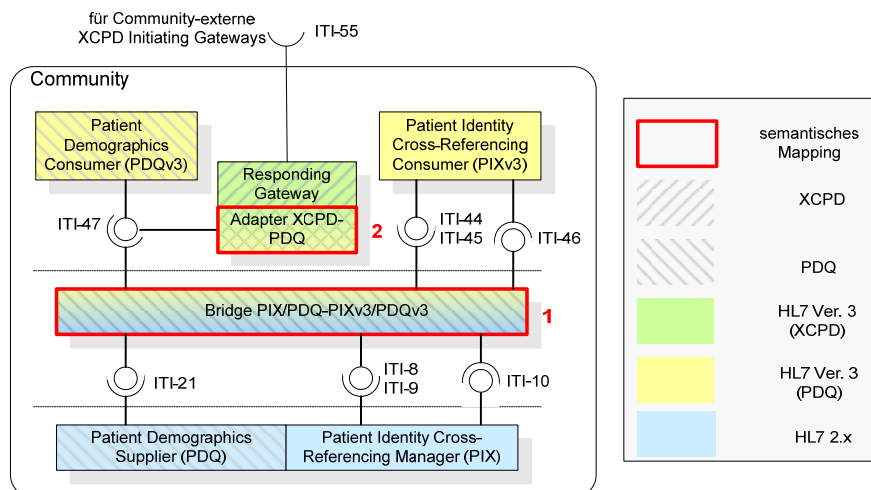


Abbildung 1. Erweiterung der ursprünglichen PIX/PDQ Schnittstelle.

Die Konversion zwischen zwei HL7 v3.0-Strukturen stößt hauptsächlich auf unterschiedliche modellbezogene Einschränkungen (engl.: model constraints) in der PDQv3- und XCPD-Definition. Da sich beide Spezifikationen noch in der Phase „Trial Implementation“ befinden, ist eine künftige Annäherung der Profile vorstellbar.

Die Übersetzung zwischen HL7 v2.x und HL7 v3.0 birgt größere Herausforderungen. Spronk ([8], Regel 230) sieht die größte Hürde für eine Umsetzung ausschließlich

mittels Integrationsmiddleware in der unterschiedlichen Granularität der Nachrichten auf beiden Seiten. Im vorliegenden Fall ist die Implementierung aufgrund der Übereinstimmung der Transaktionen möglich (siehe Tabelle 1). Die im Supplement ([4], Appendix R.2) veröffentlichte Mapping-Richtlinie wurde im Zuge der Arbeiten implementierungsreif verfeinert und um zusätzliche Interaktionen erweitert.

IHE- Transaktion PIXv3/PDQv3	HL7 Ver. 3 Interaction	IHE- Transaktion PIX/PDQ	HL7 2.x Nachricht
ITI-44 (PIXv3 Identity Feed)	PRPA_IN201301UV02 PRPA_IN201302UV02 PRPA_IN201304UV02 MCCI_IN000002UV01	ITI-8 (PIX Identity Feed)	ADT^A01^ADT_A01 ADT^A08^ADT_A01 ADT^A40^ADT_A39 ACK^Axx
ITI-45 (PIXv3 Identity Query)	PRPA_IN201309UV02 PRPA_IN201310UV02	ITI-9 (PIX Query)	QBP^K23^QBP_K21 RSP^K23^RSP_K21
ITI-46 (PIXv3 Update Notification)	PRPA_IN201302UV02 MCCI_IN000002UV01	ITI-10 (PIX Update Notification)	ADT^A31^ADT_A05 ACK
ITI-47 (PDQv3 Query)	PRPA_IN201305UV02 PRPA_IN201306UV02	ITI-21 (PDQ Query)	QBP^K22^QBP_K22 RSP^K22^RSP_K22

Tabelle 1. Abbildung der PIX/PDQ-Events zwischen HL7 2.x und HL7 Version 3.

Trotz weitgehender Überschneidungen verwendet jedes der beteiligten Nachrichtenformate (HL7 2.x PIX, HL7 2.x PDQ, HL7 v3 PIX/PDQ, HL7 v3 XCPD) ein spezifisches Objektmodell und stellt somit ein eigenes semantisches System dar. Die Schwierigkeiten, welche bei der Abbildung von Inhalten zwischen den Nachrichtenformaten auftreten, lassen sich wie folgt kategorisieren:

1. Semantische Elemente werden im Zielformat strukturell anders zugeordnet als im

Quellformat: $x \rightarrow y$.

Beispiel: Private Telefonnummern eines Patienten werden in HL7v2 im Feld PID-13 spezifiziert, dienstliche im Feld PID-14. Eine präzisere Zuordnung (z.B. Haupt-/ Nebenwohnsitz) kann im Feld PID-13-2 bzw. PID-14-2 vorgenommen werden. Darüber hinaus wird im Feld PID-13-3 bzw. PID-14-3 der jeweilige Gerätetyp (z. B. „CP“ für Mobiltelefon oder „Internet“ für E-Mail) gekennzeichnet.

In HL7v3 können diese Merkmale durch das Attribut `//telecom/@use` sowie das Präfix des Attributes `//telecom/@value` ausgedrückt werden. Das erstere kann einen einzelnen Code oder eine Liste von Codes enthalten, z. B. „MC HP“ für

„Mobiltelefon am Hauptwohnsitz“. Das Präfix des letzteren Attributs kann z. B. „tel:“ für Telefon oder „mailto:“ für E-Mail-Adresse sein.

Die Inkompatibilität besteht unter anderem darin, dass die Eigenschaft Mobiltelefon in HL7v2 als Gerätetyp, im HL7v3 aber als Adresstyp betrachtet wird.

Lösung/Workaround: Zwischen den Sätzen der Telekom-Merkmale in HL7v2 (Feldnummer 13/14, Kontakt- und Gerätetyp) und HL7v3 (Attribut @use und Präfix des Attributs @value) lässt sich eine bidirektionale Abbildung erstellen. Dabei ist zu beachten, dass nicht alle Merkmalkombinationen im jeweiligen Satz erlaubt sind: so stellt beispielsweise der Wert „PRN“ im Feld PID-14-2 („dienstliche Hauptwohnsitznummer“) einen Widerspruch dar. Dennoch ist die Erstellung der oben genannten Abbildung ziemlich aufwändig, weil auch korrekte Merkmalkombinationen in jeder HL7-Version jeweils sehr zahlreich sind und komplizierte Varianten beinhalten, wie z. B. „WP DIR EC MC“ („direkte Mobilfunknummer am Arbeitsplatz, welche im Notfall verwendet werden soll“) im HL7v3-Attribut //telecom/@use.

2. Im Quellformat darstellbare semantische Elemente sind im Zielformat nicht

darstellbar: $x \rightarrow \emptyset$.

Beispiel: Ein XCPD Responding Gateway bekommt eine ITI-55 Anfrage, in welcher die Telefonnummer des Patienten als Suchparameter benutzt wird. In der PDQ- bzw. PDQv3-Spezifikation ist die Suche nach Telefonnummern dagegen ausgeschlossen. Obwohl sowohl XCPD- als auch PDQv3-Anfragen dieselbe HL7v3-Interaktion PRPA_IN201305UV02 verwenden, unterscheiden sich die Sätze zusätzlicher Bedingungen (engl.: constraints).

Lösung/Workaround: Es ist durch organisatorische bzw. technische Maßnahmen sicherzustellen, dass die im Zielformat nicht darstellbaren Elemente auch in Nachrichten des Quellformats nicht verwendet werden, auch wenn dadurch die Schärfe der Suche eventuell beeinträchtigt wird.

3. Semantische Elemente, welche im Quellformat nicht darstellbar oder optional sind,

sind im Zielformat obligatorisch: $\emptyset \rightarrow x$.

Beispiel: Die ID einer PDQv3-Antwortnachricht entspricht dem HL7v3-Datentyp „II“ und besteht somit aus zwei Komponenten — einer obligatorischen „Root-OID“ und einer optionalen „Extension“. Bei der Umwandlung einer HL7v2-basierten PDQ-Antwortnachricht in das HL7v3-Format entspricht ihr Feld MSH-10 der „Extension“, während sich die „Root-OID“ aus der HL7v2-Nachricht nicht ermitteln lässt.

Lösung/Workaround: Meistens lässt sich das beschriebene Problem durch sinnvolle Standardwerte lösen. Ist die Festlegung eines Standardwertes nicht möglich, so ist die Transformation nicht durchführbar, und die Nachricht muss abgewiesen werden.

4. Es lässt sich nicht eindeutig bestimmen, welches semantische Element im Quellformat dem gegebenen Element im Zielformat entspricht: $? \rightarrow \mathbf{x}$.

Beispiel: Im Feld EVN-2 einer PIX Feed-Nachricht muss der Zeitpunkt der Erfassung des jeweiligen Ereignisses im System spezifiziert werden. Im Nachrichtenmodell der PIX Feed v3-Transaktion gibt es gleichzeitig drei Elemente, welche den gewünschten Zeitstempel enthalten können:

```
//controlActProcess/subject/registrationEvent/effectiveTime,  
//controlActProcess/effectiveTime und //creationTime.
```

Lösung/Workaround: Ein generischer Ansatz besteht in der Festsetzung der Auswahlkriterien sowie der Reihenfolge, in welcher die potenziellen Quellelemente geprüft werden. Sobald ein Quellelement gefunden wird, welches den Kriterien entspricht, ist die Mehrdeutigkeit gelöst. Da die zwei ersten Elemente im obigen PIX Feed-Beispiel optional sind, kann das Auswahlkriterium im einfachsten Fall so lauten: „das jeweilige Element ist vorhanden und beinhaltet einen korrekten Wert“, und die Reihenfolge der Anwendung dieses Kriteriums kann mit der Folge der Aufzählung übereinstimmen.

5. Das Zielformat schränkt die Kardinalität eines semantischen Elements stärker ein, als das Quellformat: $\mathbf{n} \cdot \mathbf{x} \rightarrow (\mathbf{n} - \mathbf{m}) \cdot \mathbf{x}$.

Beispiel: Die XCPD-Spezifikation erlaubt es, in Anfragenachrichten mehrere Unterelemente `<value>` pro Suchparameterelement `<livingSubjectId>` (Patient-IDs) zu spezifizieren, z. B.:

```
<livingSubjectId>  
  <value root="root oid 1" extension="ext 1" />  
  <value root="root oid 2" extension="ext 2" />  
</livingSubjectId>
```

Im Gegensatz dazu erlaubt PDQv3 nur je ein Unterelement `<value>` pro Element `<livingSubjectId>`.

Lösung/Workaround: Bei der Umwandlung einer XCPD-Anfragenachricht in das PDQv3-Format ist das Element `<livingSubjectId>` zu spalten, so dass die folgende Konstruktion entsteht:

```
<livingSubjectId>  
  <value root="root oid 1" extension="ext 1" />  
</livingSubjectId>  
<livingSubjectId>  
  <value root="root oid 2" extension="ext 2" />  
</livingSubjectId>
```

6. Regeln der Kombinierbarkeit semantischer Elemente im Quellformat und im Ziel-

format sind miteinander nicht kompatibel: $\emptyset \rightarrow \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_n)$.

Beispiel: In PDQv3 ist es möglich, den Patientennamen als den einzigen Suchparameter in einer Anfrage anzugeben. Empfängt der XCPD Initiating Gateway eine solche

Anfrage, so wird er keine gültige Transformation durchführen können, weil in der entsprechenden Form der XCPD-Anfrage das Patientengeburtsdatum erforderlich ist.

Lösung/Workaround: Sofern das Quellformat es zulässt, ist durch organisatorische bzw. technische Maßnahmen sicherzustellen, dass die Nachrichten des Quellformats genug Daten für die Einhaltung der Kombinierbarkeitsregeln des Zielformats enthalten bzw. eine Ableitung zulassen. Ggf. sind mehrere Nachrichten des Quellformats zu einer Nachricht des Zielformats zusammenzufassen, passende Standardwerte zu definieren, usw.

7. Beziehungen zwischen semantischen Elementen lassen sich nicht wiedergeben:

$$f(x_1, \dots, x_n) \rightarrow \emptyset.$$

Beispiel: Es soll nach einem Patienten namens „Jan“ gesucht werden, wobei nicht klar ist, ob es sich um den Vornamen oder den Nachnamen handelt. Bei einem anderen Patienten können Teile des Namens problemlos zugeordnet werden — er heißt „John Smith“. Die entsprechenden Fragmente von PDQv3-Anfragen zu diesen Patienten würden wie folgt aussehen:

Patient „Jan“	Patient „John Smith“
<name> <family>Jan</family> </name> <name> <given>Jan</given> </name>	<name> <family>John</family> <given>Smith</given> </name>

Tabelle 2. Unterschiedliche logische Verbindungen von Suchkriterien in PDQv3-Anfragen

Im ersten Fall (Patient „Jan“) handelt es sich um eine ODER-Verbindung, und im zweiten um eine UND-Verbindung. Nach der Umwandlung der oben angeführten PDQv3-Anfragen in das HL7 v2-Format sieht das Feld QPD-3 der resultierenden QBP^Q22-Nachricht folgendermaßen aus:

Patient „Jan“	Patient „John Smith“
@PID.5.1^Jan~@PID.5.2^Jan	@PID.5.1^Smith~@PID.5.2^John

Tabelle 3. Ununterscheidbarkeit logischer Verbindungen von Suchkriterien in PDQ-Anfragen

Die ODER-Semantik im Falle des Patienten „Jan“ geht unvermeidlich verloren, da in PDQv2 eine UND-Verknüpfung aller Parameter angenommen wird.

Lösung/Workaround: Das Problem aus dem oben angeführten Beispiel lässt sich umgehen, indem für den Patienten „Jan“ auf der Basis einer einzelnen PDQv3-Anfrage zwei PDQv2-Anfragen generiert werden, von denen jede jeweils eine Variante (Vor- oder Nachname) in QPD-3 enthält:

Anfrage Nr. 1	Anfrage Nr. 2
@PID.5.1^Jan	@PID.5.2^Jan

Tabelle 4. Ausdruck der logischen ODER-Verbindung mit Hilfe mehrerer PDQ-Anfragen

Da die ursprüngliche PDQv3-Anfrage mit genau einer Nachricht beantwortet werden muss, sind die Ergebnisse der v2-Anfragen auf entsprechende Weise zu aggregieren.

4. Diskussion

Zu den Vorteilen der Bridging-Lösung zählen hohe Wartbarkeit, Aufwandsminimierung und eine kurze Time-to-Market, welche sich aus der intensiven Wiederverwendung von Softwarekomponenten ergeben. Ein eventueller Nachteil des Ansatzes ist die Performance: die mehrstufige Transformation ist flexibler, aber i. d. R. nicht schneller als eine direkte Übersetzung. Jedoch bieten Integrationsprofile auf Basis von HL7v3 und Webservices die Möglichkeit, den Durchsatz durch die Verwendung asynchroner Aufrufe zu steigern.

Die hier beschriebene Methodik für semantische Transformation erhebt keinerlei Anspruch auf Verallgemeinerung oder Vereinheitlichung. Die Einbeziehung von Semantik-Mappern [9] oder Ontologien [10] wirken sich auf die Systemkomplexität, den Entwicklungsaufwand sowie auf nichtfunktionale Eigenschaften der Software aus. Die Vorteile solcher Ansätze erscheinen angesichts des Datenmodellumfangs fraglich.

Trotz der angesprochenen Einschränkungen erfüllt die Bridge für PIX/PDQ ihre Aufgabe. Sie wurde 2010 in zwei IHE-Connectathons erfolgreich getestet und kommt in den ersten Kundenprojekten zum Einsatz. Die Transformationslogik steht der IPF-Community gemäß Apache License 2.0 frei zur Verfügung. Einem PIX- bzw. PDQ-Server, der zusätzlich eine HL7 v3-Schnittstelle zu exponieren wünscht, bietet IPF erstmalig eine herstellerneutrale Open-Source-Lösung.

Der Brückenschlag zwischen einzelnen technischen Standards gehört aktuell zu den wichtigsten Herausforderungen in der medizinischen Informatik [11]. Die vorgestellte Lösung leistet dazu einen Beitrag und treibt in ihrem Teilbereich die Konvergenz von HL7 v2.x und HL7 v3 voran.

Referenzen

- [1] Integrating Healthcare Enterprise (IHE), Homepage. Online (09.06.2010): <http://www.ihe.net/>.
- [2] Wozak F, Ammenwerth E, Hörbst A, Sögner P, Mair R, Schabetsberger T. IHE based interoperability – benefits and challenges. *Stud Health Technol Inform.* 2008;136:771-6. Online (09.06.2010, abstract): <http://www.ncbi.nlm.nih.gov/pubmed/18487825>.
- [3] IHE IT Infrastructure Technical Framework. Revision 6.0, Volumes 1, 2a, 2b, 2x, 3. 2009. Online (09.06.2010): <http://www.ihe.net/Technical%5FFramework/index.cfm#IT>
- [4] Patient Identifier Cross-Reference HL7 V3 (PIXV3) and Patient Demographic Query HL7 V3 (PDQV3). *Trial Implementation Supplement.* August 10, 2009. Online (09.06.2010): http://wiki.icw.int/confluence/download/attachments/19731782/IHE_ITI_TF_Supplement_PIX_PDQ_HL7v3_TI_2009-08-10.pdf?version=1
- [5] Cross-Community Patient Discovery (XCPD). *Trial Implementation Supplement.* August 10, 2009. Online (09.06.2010): http://wiki.icw.int/confluence/download/attachments/19731782/IHE_ITI_TF_Supplement_XCPD_PC_2009-08-10.pdf?version=1
- [6] Das IPF-Projekt. Online (09.06.2010): <http://gforge.openehealth.org/gf/project/ipf/>

- [7] Hohpe G, Woolf B. Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions. Addison-Wesley. 2003. ISBN 0-321-20068-3.
- [8] Spronk R. Navigating the Pitfalls: Implementing HL7 version 3. 2006. Online (09.06.2010): http://www.ringholm.de/docs/04100_en.htm.
- [9] V2 & V3 Mapping Tools. Projekt bei HL7-Forge. Online (09.06.2010): <http://gforge.hl7.org/gf/project/v2v3-mapping/>.
- [10] Oemig F, Blobel B. Semantic Interoperability between Health Communication Standards through Formal Ontologies. Medical Informatics in a United and Healthy Europe. K.-P. Adlassnig et al. (Eds.). IOS Press, 2009. Online (09.06.2010): <http://person.hst.aau.dk/ska/MIE2009/papers/MIE2009p0200.pdf>.
- [11] Ohmann C, Kuchinke W. Future developments of medical informatics from the viewpoint of networked clinical research. Interoperability and integration. Methods Inf Med. 2009;48(1):45-54. Online (09.06.2010, abstract): <http://www.ncbi.nlm.nih.gov/pubmed/19151883>.