



# Evidence Based Telemedicine? – Sinnvolle Nutzung von Daten als Mittel der Qualitätssicherung

A. Parasta<sup>1</sup>, R. Metzger<sup>2</sup>, M. Pabst<sup>3</sup>, D. Nikmond<sup>1</sup>, A. M. Parasta<sup>4</sup>

<sup>1</sup> Epitop GmbH – Technology and Service Group München

<sup>2</sup> Dr. von Haunersches Kinderspital – Ludwig Maximilian Universität München

<sup>3</sup> Chirurgische Klinik, Kantonsspital Schaffhausen, Schweiz

<sup>4</sup> Augenklinik des Klinikum Rechts der Isar, Technische Universität München

Die Anzahl heterogener, nebeneinander existierender Datenerfassungssysteme (OP-, Behandlungs-, Bild-, Labordaten etc.) ist heute schon nahezu unüberschaubar. Obwohl evidenzbasierte medizinische Daten aus der täglichen Behandlungsroutine ein enormes Potential bieten, bleibt die objektive Verarbeitung dieser heterogen erfassten Daten (z.B. zur Fortbildung, Forschung, Qualitätssicherung, etc.) aufgrund der unterschiedlichen Datenstrukturen (s. u.) eine technische und inhaltliche Herausforderung. Die semantische Analyse von heterogenen Daten ist eine bislang weitestgehend dem Menschen vorbehaltene Fähigkeit. Das Ziel unserer Arbeitsgruppe in dem hier dargestellten Projekt war es, diese Fähigkeit in eine Software basierte Lösung umzusetzen.

Bereits ein großer Teil der Behandlungsdaten werden elektronisch erfasst. Betrachtet man die Menge an Erfahrungswerten, die in diesen täglich anfallenden Daten enthalten sind, ergibt sich hieraus ein enormes wissenschaftlich nutzbares Potential. Im Rahmen der zunehmenden Forderungen nach Evidenz basierter Medizin muss dieses Potential besser ausgeschöpft und zugänglicher gestaltet werden.

Betrachtet man die bestehenden Möglichkeiten der Datenverwaltung, lassen sich leicht die großen Probleme, die bei der Aufarbeitung der Behandlungsdaten entstehen, ableiten.

## 1 Die Untersuchung bestehender Datenmodelle

Laut der offiziellen Statistik der Kassenärztlichen Bundesvereinigung (KBV) ergibt sich folgende Aufstellung von Systemen zur Verwaltung von Patientendaten (Abrechnung bzw. Behandlungsdaten): Bei 91.995 untersuchten Installationen werden 197 unterschiedliche Systeme<sup>1</sup> eingesetzt. Rund 100 dieser Systeme können bei einer Installationsfrequenz von bundesweit unter 50 Installationen als Individuallösungen angesehen werden<sup>2</sup>. Ein gemeinsamer Standard zur Strukturierung und Speicherung von Patientendaten existiert derzeit nicht<sup>3</sup>. Zwar bestehen unterschiedliche Datenmodelle, die vor mehr als 8 Jahren als Standardisierungsversuch entwickelt wurden und die Vorgaben zur

Strukturierung von Behandlungsdaten bieten (z.B. HL7, ADT, BDT etc.), diese bieten aber auch bei einer strikten Einhaltung nicht die notwendige Detailtiefe in der Datenstruktur, um detaillierte Aspekte einer Behandlung auslesen zu können. So wird z.B. die Eingangsanamnese eines Patienten meist als fortlaufender Text abgespeichert. Die Recherche in einem größeren Patientenkollektiv unter einer bestimmten Fragestellung (z.B. Nach Symptomen gegliedert, Form und Ausmaß einer beobachteten Nebenwirkungen

einer Medikation etc.) ist bislang die Domäne der manuellen Bearbeitung durch einen Experten.

## 2 Technischer Hintergrund für genannte Defizite

Grundsätzlich lassen sich bislang zwei Datenstrukturen unterscheiden, in der medizinische Daten durch unterschiedliche Softwareprogramme exportiert werden. Der Export von Daten ist Voraussetzung für eine zentrale Datenanalyse.

### 2.1 Strukturmodell A: Defined Dataset

Hierbei bestimmt allein der Aufbau eines Datensatzes mit festgelegten Positionen die Definition der einzelnen Informationseinheiten (z.B.: Position 3 = Befund). Die Informationseinheiten werden hierbei in einer festgelegten Reihenfolge hintereinander geschrieben. Die genaue Position der einzelnen Einheiten wird entweder durch den herangezogenen Standard (z.B. HL7) oder individuell durch die genutzte Software bestimmt. Vor allem die Hersteller abhängigen individuellen Datenstrukturen stellen eine große Herausforderung bei der automatisierten Datenverarbeitung dar. So müssen für jede Softwarelösung individuelle Algorithmen erstellt werden, die Daten in eine standardisierte und vergleichbare Struktur übersetzen. Zusätzlich

Autoren: A. Parasta, R. Metzger, M. Pabst, D. Nikmond, A. M. Parasta  
Titel: Evidence Based Telemedicine? – Sinnvolle Nutzung von Daten als Mittel der Qualitätssicherung  
In: Jäckel (Hrsg.) Telemedizinführer Deutschland, Ober-Mörlen, Ausgabe 2004  
Seite: 163-165



erschwert wird die Datenübersetzung durch die zusätzliche Individualisierung bestehender Softwarelösungen durch den Anwender selbst (Definition eigener Datenfelder innerhalb der Software). Entsprechend sind über die Hersteller bezogenen Übersetzungsalgorithmen hinaus zusätzliche anwenderabhängige Algorithmen notwendig. Aufgrund der Vielzahl an vorhandenen Systemen erscheint dies als eine nahezu unrealisierbare Aufgabenstellung.

## 2.2 Strukturmodell B: Labeled Dataset

Hierbei wird jeder Informationseinheit im Datensatz stets dessen Definition angehängt (z.B.: *Diagnosis = Appendizitis*). Die bekannteste und z.Zt. modernste Form der gelabelten Datenstrukturierung ist die Struktursprache XML. Hierdurch können Daten auch in individuellen Strukturen exportiert und mit einem geringen Aufwand in eine standardisierte Struktur übersetzt werden. Leider ist die Verbreitung dieses Strukturmodells derzeit nicht ausreichend, so dass die Anwendung der Defined Datasets bei der automatisierten Datenverarbeitung für Patientendaten deutlich dominiert.

Als weitere Schwierigkeit zeigt sich der differierende Sprachschatz, der von Arzt zu Arzt bei der Dokumentation seiner Behandlungen zu finden ist. So kann z.B. die *Diagnose „Appendizitis“* als „Blinddarmentzündung“ oder „Appendizitis“ dokumentiert werden. Diese zwar für

Pat.Nr.	Datum	Befund	True	Datum_1	Diagnose	Datum_2
XXXXXX	22.01.02	OM	EBA	22.01.02	EXO	
XXXXXX	22.01.02	PA	----	22.01.02	AUNH	
XXXXXX	22.01.02	NH	----	06.06.00	MYAS	
XXXXXX	22.01.02	OM	TITR	06.06.00	EXO	
XXXXXX	06.06.00	LI	----	06.06.00	AUNH	
XXXXXX	06.06.00	OM	EBA	06.06.00	SIC	
XXXXXX	06.06.00	NH	----	06.06.00	CLTS	
XXXXXX	06.06.00	BH	BHR			
XXXXXX	06.06.00	HH	----			
XXXXXX	06.06.00	SO	BUT			
YYYYYY	10.12.01	OM	EBA	10.12.01	EXO	
YYYYYY	10.12.01	NH	----	10.12.01	AUNH	
YYYYYY	23.02.01	NH	NHD-	08.02.01	EXO	
YYYYYY	23.02.01	PE	----	08.02.01	AUNH	
YYYYYY	08.02.01	OM	EBA	08.02.01	DIA4	
YYYYYY	08.02.01	NH	----	08.02.01	HYPTY	
YYYYYY	08.02.01	HYPT				

(c) www.epitop.com

Abbildung 1: Muster eines Datensatzes mit Defined Datastructure

den Experten eindeutige Zuordnung der unterschiedlichen Begriffe stellt für die Technik eine besondere Herausforderung dar.

## 3 Lösungsansatz

Unsere gemeinsame Arbeitsgruppe bestehend aus zwei Universitätskliniken, mehreren Praxen und einem Technologiepartner entwickelte und pilotierte einen Lösungsansatz, der eine neue Generation der automatisierten Verarbeitung anonymisierter Daten darstellt. Als Ergebnis konnte eine große Zahl an echten Behandlungsdaten für die Entwicklung eines Expertensystems für die Augenklinik der technischen Universität München

(OPHTEXPERT) herangezogen werden. Der sonst notwendige hohe Aufwand bisheriger Autoren basierter Expertensysteme entfiel hier nahezu vollständig und reduzierte sich auf die Endkontrolle der gefilterten Daten. Über die Nutzung der Daten für Expertensysteme sind nun auch eine Vielzahl weiterer Einsatzmöglichkeiten (Forschung, Produktentwicklung etc.) denkbar.

## 4 Semantische Datenmodelle als Ausweg

Ziel hierbei war es die Kontext basierten Denkprozesse eines Experten bei der Analyse heterogen strukturierter Daten nachzubilden. Die wichtigsten Erkenntnisse auf dem Weg zu unserer Zielsetzung entstammen aus den Sprachwissenschaften sowie den Vorgaben der semantischen Logik. Dabei ist es uns gelungen die Verarbeitung der Daten unabhängig von ihrer zugrunde liegenden Struktur durchzuführen. Maßgeblich für die Verarbeitung ist allein der Inhalt der Datensätze, die es zu analysieren gilt. Eine unabdingbare Voraussetzung ist stets die Anonymisierung der Patientendaten ohne die eine Nutzung der Daten außerhalb der Behandlung datenschutzrechtlich nicht möglich ist.

### 4.1 Beispiel

Wird in einem Datensatz der Eintrag „Appendizitis“ gefunden, wird ein Mengenvergleich in den anderen Dokumenten des gleichen Nutzers durchgeführt und

```

- <Examination>
  <PatGUID>7602C3A8-13EC-4202-8151-54F9FEAF22E2</PatGUID>
  <ExaminationDate>27.07.2003</ExaminationDate>
- <ExaminationResult>
  <Diagnosis>Appendizitis</Diagnosis>
  <Diagnosis>HerzInsuffizienz</Diagnosis>
  <Anamnese>Seit 3 Tagen über Schmerzen im linken Bein</Anamnese>
  <systemicDisease>Diabetes Mellitus</systemicDisease>
  <systemicDisease>Asthma</systemicDisease>
  <systemicDisease>Herzinsuffizienz</systemicDisease>
</ExaminationResult>
- <Treatment>
  - <Medication>
    <Drug>Metformin</Drug>
    <Dosis attribute="percent">0,005</Dosis>
    <TakingPerDay>1</TakingPerDay>

```

(c) www.epitop.com

Abbildung 2: Muster eines Datensatzes mit Labeled Datastructure

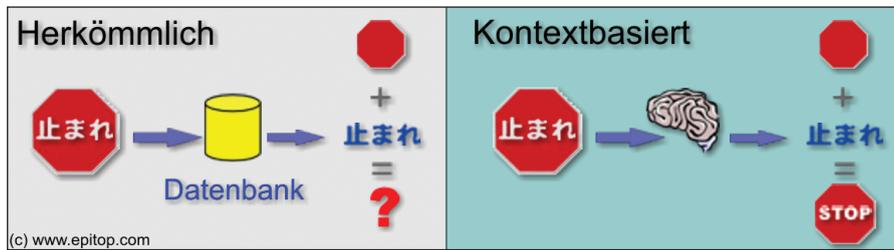


Abbildung 3: Schematische Prozessdarstellung semantischer Erkennung

ein strukturelles Schema der Kontextverhältnisse für „Appendizitis“ erstellt und gespeichert. Befindet sich nun der Begriff „Blinddarmentzündung“ im gleichen strukturellen Schema wird es als Synonym für „Appendizitis“ erkannt und entsprechend gewertet.

Die Stärken dieser neuen Generation der Datenverarbeitung zeigen sich v. a. bei großen Datenmengen, da nur so die notwendigen Kontextanalysen mit einer ausreichenden Zuverlässigkeit durchgeführt werden können.

Um die Flexibilität zu steigern, werden in einem nächsten Schritt die so verarbeiteten Daten in einer einheitlichen Struktur in eine statische Datenbank geschrieben, die dann für die unterschiedlichsten Softwaresysteme (Wissensmanagement, Statistiken etc.) verfügbar ist und nach Belieben weiter verarbeitet werden kann. So werden beliebige Daten miteinander vergleichbar und können flexibel analysiert werden.

## 5 Zusammenfassung

Das schnelle Wachstum der Informationstechnologie führte bislang zu einer divergierenden Entwicklung zahlreicher heterogener Datenmodelle. Diese Heterogenität stellt jedoch eine wachsende Belastung dar, wenn man die bereits während einer Behandlung erfassten Daten zu Wissensmanagement und Qualitätssicherung heranziehen will. Die Ergebnisse unserer Arbeitsgruppe zeigen, dass eine effektive Datennutzung, auch bei der bestehenden Heterogenität durch eine kontextbezogene Analyse der Dateninhalte mit der vorhandenen Technologie möglich ist. Diese hochinnovative Generation der Datenverarbeitung – wie in unserem Modell verwirklicht – eröffnet neue Mög-

lichkeiten für Forschung, Fortbildung und Kommunikation in der Medizin. Ist die Problematik der Datenkompatibilität einmal gelöst, so können Daten ungeachtet der verwendeten Plattform gesammelt und verglichen werden. Diese neue Generation an Wissensmanagement- und Decisionsupport-Programme eröffnet nahezu unbegrenzte Möglichkeiten der technischen Unterstützung evidenzbasierter Behandlungsverfahren.

## Fußnoten

- 1 EDV Statistik der Kassenärztlichen Vereinigung, Installationsstatistik – Systeme, 31.12.2002 <http://www.kbv.de>
- 2 EDV Statistik der Kassenärztlichen Vereinigung, Installationsstatistik – Anbieter, 31.12.2002 <http://www.kbv.de>
- 3 Parasta A, Metzger R, Till H, Kampmann P, Parasta AM. Digitale Patientenakten: Potential ausgeschöpft? Deutsche Gesellschaft für Chirurgie, J.R. Siewert/W.Hartel (Hrsg.) Kongressband 2002 – Springer Verlag; S 906, Deutscher Chirurgenkongress, Berlin, 7.-10.5.2003