

Ontologie und Medizin

Warum benutzen Ärzte keine Computer?

Die Akzeptanz elektronischer Aufzeichnungen muss wesentlich erhöht werden. Dazu gilt es, deren Nutzen deutlich zu machen. Das Institut für formale Ontologie und medizinische Informationswissenschaft an der Universität Leipzig erarbeitet Grundlagen, um diese Ziele zu erreichen.

ie elektronische Patientenakte (EPA) hat sich in der ärztlichen Praxis bislang nicht durchsetzen können, obwohl in den Industrieländern kaum noch Arztpraxen komplett auf die Verwendung von Computern verzichten und die jetzt ins Berufsleben startende Ärztegeneration mit dem Computer aufgewachsen ist. Der Einsatz konzentriert sich jedoch auf die Patientenverwaltung, die Abrechnung von Leistungen und die Praxisorganisation. Dazu kommen Untersuchungsverfahren, die ohne Computer nicht oder nicht mehr rationell realisierbar sind. Doch nur in den skandinavischen Ländern ist die EPA - aufgrund gesetzlicher Regelungen - ärztlicher Alltag.

Die Probleme sind vielfältig - oft lassen sie sich auf die empfundene Nutzbarkeit der Systeme zurückführen: Während des Gesprächs mit dem Patienten benötigen Ärzte leicht zu handhabende Geräte. PCs sind - wenn überhaupt - nur im Sprechstundenbetrieb verwendbar, Laptops und Tablet-PCs für die Visite immer noch zu schwer, die Bildschirme der PDAs für viele Anwendungen zu klein. Hinzu kommen wichtige (sozio-) psychologische Fragen. Beispielsweise erfordern Eingaben in ein elektronisches System häufig mehr Aufmerksamkeit als handschriftliche Notizen und stören so stärker die Kommunikation zwischen Arzt und Patient. Gleichzeitig bleiben die Darstellungsmöglichkeiten der meisten elektronischen Systeme erheblich hinter Papier und Stift zurück und tragen so zum Eindruck einer geringeren Ausdrucksstärke dieser Systeme bei. Deshalb besteht nach wie vor eine weit verbreitete Präferenz für handschriftliche gegenüber elektronischen Aufzeichnungen sowie für freien Text gegenüber restriktiven Eingabemasken.

Eine kürzlich veröffentlichte Studie (Hippisley-Cox u. a.: The Electronic Patient Record in Primary Care. In: BMJ 2003; 326: 1439–1443) weist dagegen die Überlegenheit elektronisch geführter Patientenakten gegenüber herkömmlichen nach. Die Vorteile beschränken sich nicht auf die Lesbarkeit (und auf die durch bessere Lesbarkeit vermeidbaren Irrtümer). Ärzte, die elektronisch dokumentieren, konnten auch eine größere Anzahl detaillierter Informationen rekonstruieren.

Schwachstellen

Auch in der medizinischen Forschung steht die automatisierte Informationsverarbeitung vor erheblichen Schwierig-

| Description of Terror (Control of the Control of

Abbildung: Fehlerhafte Zuordnung in der Systematisierten Nomenklatur der Medizin (SNOMED): "Struktur der Vena labialis" ist sowohl "Schamvene" als auch "Struktur einer Vene des Kopfes" direkt untergeordnet. Dieser Fehler beruht wahrscheinlich darauf, dass "labia" sowohl für die Lippen des Mundes als auch für die Schamlippen verwendet wird und in SNOMED die Konsistenz nicht über eine Ontologie räumlicher Beziehungen gesichert ist (siehe http://ontology.buffalo.edu/medo/SNOMED.pdf).

keiten: Obwohl die Computer aus der Forschung nicht mehr wegzudenken sind, wird das volle Potenzial noch nicht genutzt. Abgesehen von Standard-Büroanwendungen wie Textverarbeitung und Tabellenkalkulation unterstützt die Informationsverarbeitung vorzugsweise Einsatzfälle, bei denen das schiere Volumen gut strukturierter Daten das vorrangige Problem darstellt – beispielsweise das Untersuchen quantitativer Korrelationen mittels statistisch basierter Mustererkennungsmethoden. Sobald die Daten jedoch nicht wohlstrukturiert sind, zeigt sich auch hier schnell, dass geeignete qualitative Grundlagen fehlen, die dem Datenchaos vorbeugen oder das Verknüpfen von Daten aus heterogenen Quellen erlauben würden. So werden im Rahmen evidenzbasierter Verfahren sowohl in der Entscheidungsunterstützung für Ärzte als auch bei retrospektiven Fallkontrollstudien in der klinischen Praxis die vorhandenen Datenbestände bislang nicht ausreichend genutzt, weil sie unter Verwendung unterschiedlicher, nicht kompatibler Terminologie- und Klassifikationsschemata angelegt wurden. Der fehlende terminologische

> Kontext behindert die Übersicht über und das Navigieren zwischen den Daten. Die medizinische Forschung ist aber kaum denkbar ohne einen schnellen, koordinierten und vor allem konsistenten Zugang zu allen relevanten Daten. Und dies in einem Rahmen, der diese Daten dem Nutzer in verständlicher, möglichst intuitiver Form präsentiert. Diese Probleme werden umso deutlicher, je zahlreicher die zu integrierenden



Datenquellen sind. Auch die Ärzte sind davon unmittelbar betroffen, wenn sie sich beispielsweise auf die evidenzbasierte Medizin stützen wollen.

Woran mangelt es den Computern, um den Ärzten diejenige Unterstützung zu bieten, die ihnen die (tägliche) Arbeit erleichtert? Das Arbeitsleben ist durch eine Explosion der zur Verfügung stehenden und mehr oder weniger relevanten Informationsmenge gekennzeichnet. Den für die eigene Arbeit wichtigen Teil der Informationen herauszufiltern. erfordert erheblichen zeitlichen Aufwand. Warum spielt die Informationstechnik hier ihre Stärken nicht aus, sondern offenbart vielmehr erhebliche Schwächen? Warum liefert die Suche in großen Freitext-Datenbeständen oft unbefriedigende Ergebnisse beziehungsweise ist langwierig und erfordert entsprechendes Training? Warum führt die Verknüpfung von Datenbeständen zu Inkonsistenz und Widersprüchen?

Fehlertoleranz

Diese Fragen lassen sich zum Teil auf bestimmte Aspekte der in den Anwendungen verwendeten Terminologien zurückführen: Menschen zeigen in ihrer Kommunikation eine hohe Fehlertoleranz. In die Interpretation der empfangenen Signale beziehen sie implizit oder explizit einen Kontext mit ein und können so Auslassungen und Fehler ausgleichen. Unabdingbare Voraussetzung dafür sind sowohl die Beherrschung der zugrunde liegenden natürlichen Sprache als auch der verwendeten fachspezifischen Terminologie. Allerdings sind diese beiden Bedingungen nicht hinreichend: Medizinstudenten sind ein gutes Beispiel dafür, dass eine große Menge isolierter Fakten meilenweit von anwendungsbereitem Wissen entfernt ist. Viele angehende Ärzte berichten davon, dass die isolierten Fakten erst in der Auseinandersetzung mit den Patienten eine sinnvolle Ordnung erhalten. Zu dieser Integration, einem weiteren Lernprozess, den Medizinstudenten durchlaufen und dem die neue Approbationsordnung Rechnung trägt, fehlt den Informationsverarbeitungssystemen das Pendant. Letztlich sind die Computerpro-

gramme nicht fehlertolerant. Die Integration bloßer Daten zu echtem Wissen, die Menschen durch aktive Auseinandersetzung mit Fakten, gestützt auf Erfahrung und auf Kommunikation mit anderen leisten, kann im Computer nur teilweise und unter Rückgriff auf exakte Definitionen und Strukturen nachgebildet werden. Darüber hinaus ist nur ein kleiner Teil der elektronisch gespeicherten klinischen Daten überhaupt im softwaretechnologischen beziehungsweise datenbanktechnischen Sinne strukturiert; das Gros liegt dagegen als unstrukturierter Freitext vor. Schließlich spiegeln die vorhandenen strukturierten Datenbestände die unterschiedlichsten disziplinspezifischen Blickwinkel der Nutzergruppen wider - von Synonymen, Ungenauigkeiten und Widersprüchen in der Terminologie bis hin zu unterschiedlichen Maßeinheiten und Bezugsgrößen für Messwerte. Diese Diversität stellt für die automatisierte Integration eine hohe Hürde dar. Die Tabelle gibt einen Überblick über einige der zurzeit in der Medizin verwendeten Terminologiesysteme.

Zentrale Referenztaxonomie

Ursprünglich hat man versucht, diese Probleme für jeden einzelnen Fall ad hoc zu lösen. Spätestens mit dem Aufkommen des PC und dem in der Folge exponentiell wachsenden Datenvolumen erschien dieser Ansatz jedoch nicht mehr realistisch: Mit jeder neuen Software musste, im ungünstigsten Falle, je eine Schnittstelle zu allen anderen bereits eingesetzten Programmen erarbeitet werden, um den Datenaustausch zu ermöglichen. Da mit der wachsenden Zahl der Anwendungssysteme auch die der bereitzustellenden Schnittstellen explodierte, kam die Idee einer einzigen integrierenden Referenztaxonomie auf, einer Standardterminologie, die als eine Art Esperanto dienen würde. All die verschiedenen Klassifikationen wären dann nur einmal in diese Taxonomie zu übersetzen, und all die auf dieser Grundlage abgestimmten Datenbestände wären automatisch zueinander kompatibel.

Zwar ist der Nutzen eines solchen Ansatzes offensichtlich. Die sich allmählich abzeichnende Vielzahl zu berücksichtigender Gesichtspunkte ließ die Entwicklung dieser Referenztaxonomie jedoch zunehmend schwieriger erscheinen als zunächst angenommen: Jeder aus irgendeiner anderen Klassifikation stammende Ausdruck muss in all seinen Facetten präzise dargestellt werden. Beispielsweise unterscheiden sich

Tabelle: In der Medizin verwendete Taxonomien		
System	Inhalt/Zweck	Kritikpunkte
SNOMED www.snomed.org	Kodieren medizinischer Aufzeichnungen	klare ontologische Struktur fehlt aus der Zielsetzung heraus für das Sammeln forschungsrelevanter Daten ungeeignet
GALEN www.opengalen.org	sprachlich neutrale Ontologie medizinischer Phänomene	verwendet für menschliche Kommunikation kaum geeignete Bezeichnungen
UMLS Metathesaurus http://umlsinfo.nlm.nih.gov	Standardkompendium verschiedener großer medizinischer Sprachsysteme	Inkonsistenz zwischen den verschiedenen Teilen
Diagnosis Related Groups – DRG www.g-drg.de	Kategoriensystem für Kalkulationen ("Fallpauschale")	ontologisch naiv aus der Zielsetzung heraus für das Sammeln forschungsrelevanter Daten ungeeignet
MeSH (Medical Subject Headings) www.nlm.nih.gov/mesh	Kategoriensystem für Literaturrecherche	fokussiert auf das Indizieren von Literatur ontologisch naiv
Digital Anatomist http://sig.biostr.washington.edu/projects/fm	Referenzontologie der Anatomie des menschlichen Körpers	auf Anatomie beschränkt



die über einen Patienten in einem medizinischen Informationssystem gespeicherten Daten stark in Abhängigkeit davon, ob der individuelle Krankheitsverlauf des einzelnen Patienten oder die Abrechnung medizinischer Leistungen im Mittelpunkt des Interesses steht. Somit ist eine große inhaltliche Ausdrucksstärke der Referenztaxonomie erforderlich. Gleichzeitig sind für Computerprogramme logisch einfache Ausdrucksmittel notwendig. Gerade das führte zu "modell"-artigen Vereinfachungen der Eingangsdaten und in der Folge zu vielen verschiedenen, jeweils durch vermeintliche Optimierung für den jeweiligen Anwendungsfall geprägten Terminologie- und Klassifikationssystemen, die somit letztlich das Ausgangsproblem der Dateninkompatibilität nur auf eine andere Ebene verlagerten. Das "Unified Medical Language System" der National Library of Medicine der USA (www.nlm.nih.gov) ist aus der Sicht seiner Autoren kein einheitliches System, sondern ein gigantisches Konglomerat verschiedener Terminologien.

Das Institut für Formale Ontologie und Informationswissenschaften an der medizinischen Fakultät der Universität Leipzig (IFOMIS, *Kasten*) geht das Projekt einer zentralen Referenztaxonomie grundsätzlich an. Der Blick geht dabei weit über die Softwareentwicklung hinaus, denn das IFOMIS stützt sich in seiner Arbeit auf 2500 Jahre philosophischer Forschung.

Bemerkenswerterweise haben die Informatiker die Aufgabe, zentrale Klassifikationssysteme zu entwickeln, als "Ontologie" bezeichnet - mit einem Wort, das aus der Philosophie stammt und dort mit dem klassischen Terminus "Metaphysik" nahezu gleichbedeutend ist. Die Arbeit an einer Referenztaxonomie wirft in der Tat viele in der Metaphysik seit langem diskutierte Fragen neu auf: Fragen nach den für eine Referenztaxonomie adäquaten Konstruktionsprinzipien sind originär philosophischer Natur. Es geht um so generelle Probleme wie die Auswahl der Kategorien auf der obersten Ebene einer Taxonomie oder bezogen auf die Medizin - um das Verständnis der Beziehungen zwischen den anatomischen und den physiologischen

Hintergrund IFOMIS

Das Institut für Formale Ontologie und Medizinische Informationswissenschaften wurde im April 2002 als Teil der medizinischen Fakultät der Universität Leipzig gegründet. Es stützt sich auf zwei Millionen Euro Preisgeld des Wolfgang Paul-Preises, gestiftet vom Bundesministerium für Bildung und Forschung, und sieht sich als Zentrum für theoretische Forschung in der formalen und angewandten Ontologie. IFOMIS ist eine interdisziplinäre Forschergruppe mit Mitgliedern aus Philosophie. Medizin und biomedizinischer Informatik. 2003 wurde die Arbeit des IFOMIS im Bereich der medizinischen Ontologie mit rund 400000 Euro aus dem 6. Rahmenprogramm für Forschung und technologische Entwicklung (FP6) der Europäischen Union gefördert. Damit ist IFOMIS in das europäische Network of Excellence zum Thema Medical Informatics and Semantic Data Mining (demnächst unter www.semanticmining.org erreichbar) aufgenommen, das eine Reihe europäischer Institutionen umfasst. Das FP6 wurde entwickelt, um sowohl auf nationaler als auch auf EU-Ebene Forschung und innovative Entwicklung strukturieren, koordinieren und integrieren zu können. Die Informatik ist eines der sieben Forschungsgebiete, die durch das Rahmenprogramm als besonders wichtig eingestuft wurden.

Kategorien untereinander und jeweils zu den auf Krankheiten bezogenen Kategorien. Viele der Schwierigkeiten, mit denen sich die Informatiker heute auseinander setzen müssen, gleichen denen, die die Philosophen seit Aristoteles beschäftigen: Wie sind Funktionen und Relationen, Ereignisse und Prozesse, Identität und Differenz zu definieren? Entstehen Klassen und Taxa durch menschliches Kategorisieren, oder spiegeln sie grundlegende Unterschiede auf der Seite der Sachen selbst wider?

Interdisziplinär

Nicht die Konstruktion vereinfachender begrifflicher Modelle, sondern – wie ehrgeizig oder unwahrscheinlich das auch klingen mag – die adäquate Widerspiegelung der (biologischen und medizinischen) Realität als Ganzes ist die Richtschnur für die Ontologie des IFOMIS. Daher arbeiten dort Philosophen mit Informatikern, Medizinern und Biologen eng zusammen; zum Beispiel mit den Anatomie-Spezialisten der University of Washington in Seattle, die für das Digital Anatomist System

verantwortlich sind; einen der bisher erfolgreichsten Versuche, eine Referenzontologie für die biomedizinische Forschung bereitzustellen (http://sig.biostr.washington.edu/projects/fm).

Die Arbeit des Instituts basiert darauf, dass der Entwicklung von Informationssystemen eine gründliche und nicht reduzierende Analyse des biomedizinischen Anwendungsgebietes voranzustellen ist, um den Gehalt, das Wesen der Dinge selbst nicht zu verfälschen. Während in der Vergangenheit vorhandene Datenbanken oder in der Praxis gebräuchliche Terminologien als Ausgangspunkt für die Entwicklung von Ontologien für Informationssysteme dienten, verfolgt IFOMIS die Ausarbeitung einer umfassenden Theorie über die Struktur und die Wechselbeziehungen der für die Medizin relevanten Entitäten der realen Welt. Diese Theorie enthält zum Beispiel ein Modul, das die räumlichen Beziehungen innerhalb des Körpers darstellt und damit Fehler in der Zuordnung zu Körperregionen verhindert (Abbildung).

Dahinter steht die Idee, einen Rahmen für die Datenintegration in Computersystemen zu schaffen, der mit der aus der Behandlung von Patienten erwachsenen Erfahrung des Arztes Schritt halten kann. Diese Theorie muss Raum bieten für unterschiedliche Sichten auf dieselbe Realität, sowohl für die substanzielle als auch für die prozessuale Seite der Realität sowie für Sichten auf unterschiedlichen Granularitätsebenen und aus verschiedenen Perspektiven. Im Rahmen dieser Theorie sollen die gängigen für Ärzte entwickelten Terminologien mit neuen biomedizinischen Terminologien mittels der zentralen Referenztaxonomie so verknüpft werden, dass sowohl die für Computer benötigte formale Präzision als auch die Grundlagen für schnelles und effizientes Navigieren durch die Welt biomedizinischer Daten gewährleistet sind - in einer Weise, die auch für den praktizierenden Arzt relevant und nachvollziehbar ist.

Barry Smith, Dirk Siebert

Kontaktadresse: Institute for Formal Ontology and Medical Information Science (IFOMIS), Härtelstraße 14–16, 04107 Leipzig; Internet: http://ifomis.de