

## Erfahrungen mit einer interoperablen Datenerfassungsplattform für multizentrische Forschungsnetze basierend auf HL7 CDA

Klein A<sup>1</sup>, Ganslandt T<sup>1</sup>, Brinkmann L<sup>2</sup>, Spitzer M<sup>2</sup>, Ückert F<sup>2</sup>, Prokosch HU<sup>1</sup>

<sup>1</sup>Lehrstuhl für Medizinische Informatik, Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

<sup>2</sup>Institut für Medizinische Informatik und Biomathematik, Westfälische Wilhelms- Universität Münster, Deutschland  
andreas.klein@imi.med.uni-erlangen.de

**Einleitung und Fragestellung:** Für die Forschung an seltenen Erkrankungen werden immer häufiger multizentrische Forschungsnetze etabliert, um durch eine überregionale Zusammenarbeit größere Fallzahlen und somit statistisch aussagekräftigere Schlußfolgerungen zu ermöglichen. Um die individuellen Stärken der beteiligten Einrichtungen optimal einsetzen zu können, bedarf es einer integrierten IT-Plattform einschließlich eines Remote Data Entry (RDE) Moduls für die Erfassung und gemeinsame Nutzung der erhobenen Daten. Dabei muss sichergestellt werden, dass die Plattform eine möglichst weitreichende Interoperabilität mit anderen Systemen und einen wirksamen Schutz der Patientendaten gewährleistet. Unter Verwendung geeigneter Standards für das Speichern von strukturierten Daten soll deshalb ein geeignetes integrierbares RDE-Modul entwickelt werden. In einem multizentrischen Forschungsnetz über Epidermolysis Bullosa, einer seltenen Hautkrankheit, soll ein solches Modul in eine für klinische Forschungsabläufe geeignete Plattform integriert werden.

**Material und Methoden:** In Zusammenarbeit mit zwei weiteren Forschungsnetzen wurde eine bereits existierende, web-basierte, Telemedizin-Anwendung [1] für das Speichern von Bilddaten um ein RDE Modul ergänzt. Anforderungen an die Datenerfassungsmasken des Moduls wurden mit den klinischen Netzwerkpartnern und mit den dortigen bereits vorhandenen papierbasierten Erhebungsbögen abgestimmt. Die Datenelemente wurden hierbei extrahiert und thematisch sowie prozessorientiert zu Eingabefeldern gruppiert.

Um sowohl die langfristige Nutzbarkeit als auch die Interoperabilität zu optimieren, wurde das Speichern und Ändern der Daten auf Grundlage von Standards wie XML, XSLT und CDA (HL7 Clinical Document Architecture) implementiert [2-5]. Ein CDA Release 1 konformes Schema wurde für die dateibasierte Datenhaltung der jeweiligen Datenerfassungsmasken entwickelt. Die Definition und Metadaten jeder Datenerfassungsmaske wurden in Form einer XML-Datei gespeichert und inkl. eines Versionskontrollsystems im Gesamtsystem integriert und zugänglich gemacht. Verschiedene XML Hierarchieebenen in dieser XML-Datei erlauben neben dem Festlegen der Reihenfolge und Ausrichtung der darzustellenden XML-Elemente auch die Definition von Daten- und Legendenelementen. Das Parsen und Rendern der Datenerfassungsmasken wurde mit Hilfe von XSLT-Stylesheets umgesetzt.

Das System wurde auf einer webbasierten LAMPS-Architektur (Linux, Apache, MySQL, Perl, SSL) realisiert. Das zugrunde liegende Datenbankschema basiert teilweise auf dem Entity-Attribute-Value (EAV) Konzept [6] und erlaubt somit eine flexible Speicherung der jeweiligen Forschungsdaten. Durch die Implementierung eines speziellen Indexierungs- und Abfrageverfahrens können unter Verwendung des EAV-Konzepts geschwindigkeitsoptimierte Selektionen auf den gespeicherten XML-Datensätzen durchgeführt werden. Um den deutschen Datenschutzanforderungen gerecht zu werden, wird das generische Datenschutzkonzept der Telematikplattform für Medizinische Forschungsnetze e.V. (TMF) umgesetzt [7].

**Ergebnisse:** Das entwickelte System ist seit 09/2005 im Produktivbetrieb und erlaubt das Erfassen von klinischen Daten in 6 unterschiedlichen Kategorien wie z.B. Patientenanamnese, klinischer Befund, Laborergebnisse etc. Zusätzlich wurde ein Modul integriert, das die graphische Erfassung von Stammbäumen erlaubt und diese entsprechend der Newick-Notation abspeichern kann [8]. Um den beteiligten Forschungseinrichtungen erste einfache Datenauswertungen zu ermöglichen, wurde ein Exportverfahren im CSV- und XLS-Format implementiert. Aktuell existieren 13 Benutzer aus 3 verschiedenen klinischen Einrichtungen, die mehr als 101 Sitzungen mit einer maximalen Sitzungsdauer von 3 Stunden durchgeführt haben. Im System sind mehr als 100 Patienten und mehr als 300 strukturierte Erhebungsbögen erfasst. Am häufigsten wurden dabei die Bögen zum klinischen Befund (108), zur Patientenanamnese (93) und zum Antigenmapping (67) eingegeben. Das System erlaubt die Implementierung komplexer Bögen mit einer Vielzahl von Merkmalen sowie 1:n-zugeordneten Unterelementen und flexibel erweiterbaren Merkmalslisten. Aus dem HL7 CDA Release 1-Standard werden die strukturiert definierten Header-Elemente für die Ablage der Patientenstammdaten genutzt, während der CDA-Body im Release 1 noch nicht ausreichend strukturiert ist, so daß die klinischen Dokumentationsinhalte zwischen den <local\_markup>-Segmenten in einer speziellen selbstdefinierten Art gespeichert werden. Die durch das System generierten CDA Dokumente sind gegenüber dem CDA Release 1 Schema valide.

**Diskussion:** Eine webbasierte Datenerfassungsplattform wurde erfolgreich in den Produktivbetrieb in einem multizentrischen Forschungsnetz integriert. Der Gebrauch von Standards und daher die Trennung von Logik, Darstellung und Inhalt ermöglichen ein sehr flexibles Dokumentationssystem für klinische Daten. Zudem erlauben die verwendeten Standards eine leichte Erweiterbarkeit des Systems und eine leichte Austauschbarkeit der Daten mit anderen Systemen und garantieren zusätzlich eine langfristige Zugänglichkeit der Daten. Um die klinischen Daten für den Austausch mit anderen Systemen noch besser vereinheitlichen zu können, sollen zukünftig weitere strukturierte Elemente des CDA Release 2 [9] unterstützt werden. Zudem sollen auch komplexe Datenauswertungen möglich sein, weshalb das vorhandene Exportverfahren um einen grundlegenden SQL-Datenbankgenerierungsexport ergänzt werden soll.

Das Projekt wird im Rahmen des Netzwerks Epidermolysis Bullosa durch das BMBF gefördert (Förderkennzeichen 01GM0301).

### Literatur

- [1] Ückert F, Fischer G, Brinkmann L, Jürgens H, Paulussen M. Flexible and secure Internet-based discussion of variable medical image data. *CARS* 2004; 12(6): 274-278
- [2] Dolin R, Alschuler L, Beebe C, Biron P, Boyer S L, Essin D, Kimber E, Lincoln T, and Mattison J. The HL7 Clinical Document Architecture. *J Am Med Inform Assoc* 2001; 8(6): 552-569
- [3] Heitmann K U, Schweiger R, Dudeck J. Discharge and referral data exchange using global standards – the SCIPHOX project in Germany. *Int J Med Inform* 2003; 70(2-3): 195-203
- [4] Schweiger R, Brumhard M, Hoelzer S, Dudeck J. Implementing health care systems using XML standards. *Int J Med Inform* 2005; 74(2-4): 267-277
- [5] Müller M, Ückert F, Bürkle T, Prokosch H-U. Cross-institutional data exchange using the clinical document architecture (CDA). *Med Inform* 2005; 74(2-4): 245-256
- [6] Nadkarni P, Marenco L, Chen R, Skoufos E, Shepherd G, Miller P. Organization of Heterogeneous Scientific Data Using the EAV/CR Representation. *J Am Med Inform Assoc* 1999; 6(6): 478-493
- [7] Pommerening, K, Reng, M. Secondary use of the EHR via pseudonymisation. *Stud Health Technol Inform* 2004; 103: 441-446
- [8] Lachmund P, Nebel IT, Führer D, Paschke R. The Pedigree Tool: Web-Based Visualization of a Family Tree. *Hum Mutat* 2004; 23(2): 103-105
- [9] Dolin RH, Alschuler L, Boyer S, Beebe C, Behlen FM, Biron PV, Shabo Shvo A. HL7 Clinical Document Architecture, Release 2. *J Am Med Inform Assoc* 2006; 13(1): 30-39