



**Herzlich Willkommen!**

Rheiner Archivtage der GMDS AG AKU, Rheine, 08.12.2006

## **XML-Archivierung in der medizinischen Forschung – Ergebnisse des Industrie-Workshops „Perspektiven der XML- Archivierung“**

Prof. Dr. Anke Häber<sup>1</sup>, Sebastian Claudius Semler<sup>2</sup>,  
Dr. Wolfgang Kuchinke<sup>3</sup>

<sup>1</sup> Westsächsische Hochschule Zwickau

<sup>2</sup> Telematikplattform für Medizinische Forschungsnetze (TMF) e.V., Berlin

<sup>3</sup> KKS Düsseldorf



# Wer ist die TMF ?

## Telematikplattform für Medizinische Forschungsnetze

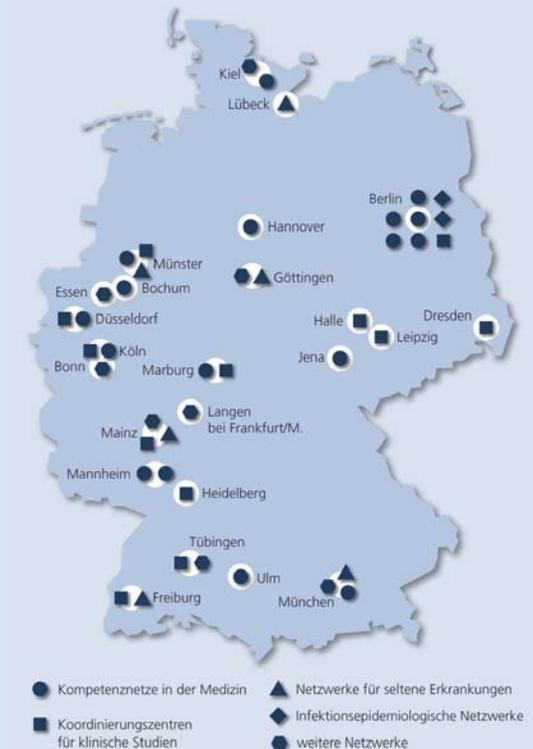
- ↪ Dachorganisation der medizinischen Forschungsverbünde
- ↪ gefördert vom BMBF
- ↪ 1999 parallel gegründet zur Etablierung der
  - ↪ Kompetenznetze in der Medizin (KN)
  - ↪ Koordinierungszentren für klinische Studien (KKS)
- ↪ Parallelinitiative zur Gesundheitstelematik im Versorgungsbereich (ATG, bit4health-Projekte etc.)
- ↪ nationale Zentralinstanz für die vernetzte medizinische Forschung – **eine der wenigen zentralen Telematik-Institutionen im deutschen Gesundheitswesen**
- ↪ Ende 2003 als e.V. ausgegründet



## Mitglieder

### ↪ Nationale Netzwerke und Institutionen

- ↪ Kompetenznetze in der Medizin
- ↪ Koordinierungszentren für Klinische Studien
- ↪ Netzwerke Seltene Erkrankungen
- ↪ Nationales Genomforschungsnetz
- ↪ Fraunhofer Institut für Toxikologie und experimentelle Medizin
- ↪ Paul-Ehrlich-Institut
- ↪ und andere



↪ 2006: 45 Mitglieder (2003: 30 Mitglieder)

## Arbeitsgruppen

- ↪ IT Infrastruktur und Qualitätsmanagement
- ↪ Datenschutz
- ↪ Biomaterialbanken
- ↪ Management Klinischer Studien
- ↪ Molekulare Medizin  
(in Gründung)

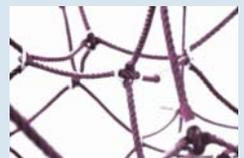
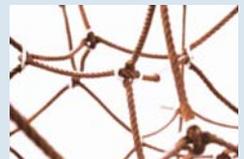
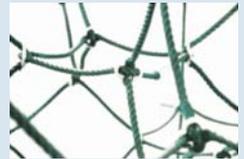


## Foren

- ↪ Netzsprecher/Koordinatoren
- ↪ Geschäftsführer
- ↪ PR Verantwortliche

## Themen

- ↪ IT- Infrastruktur für klinische Forschung  
(z. B. Pseudonymisierungsdienst, SAE-Management,...)
- ↪ Verzahnung zwischen Forschung und Versorgung  
(z. B. eGK/HPC, Leitlinien, Schnittstellen)
- ↪ Standards und Terminologie  
(z. B. CDISC, SAS Makros, SDTM-Wandler, ATC-Kodierung)
- ↪ Rechtliche und ethische Rahmenbedingungen  
(z. B. Datenschutzkonzept, BMB, Sicherheitskonzept, AMG)
- ↪ Qualitätssicherung  
(z. B. Systemvalidierung, SOPs, Checkliste, IIT Monitoring)
- ↪ Technologiebewertung  
(z. B. Mobile Computing, eArchivierung, CME)
- ↪ Öffentlichkeitsarbeit  
(z. B. Parlamentarische Abende, Infostand DGIM)



## ↳ CDISC

- ↳ TMF-Projekt V009-01 CDISC I (2004/05)
  - ↳ Evaluation, Implementationsstand in RDE-Systemen in D
- ↳ TMF-Projekt V033-01 CDISC II (2005)
  - ↳ Use Cases Spezifikation zum Einsatz von CDISC
- ↳ TMF-Projekt V021-01 & -02 (2004/05 & 2005/06)
  - ↳ SDTM-basierte SAS-Makros zur statist. Auswertung
- ↳ TMF-Projekt V034-01 SDTM-Wandler (2006)
  - ↳ Datenkonverter CDISC ODM → SDTM
- ↳ CDISC-Unterstützung im TMF-Pseudonymisierungsdienst
  - ↳ Nutzung von CDISC ODM zum Aufbau v. Forschungsregistern
- ↳ Nationales CDISC-Symposium 2005
  - ↳ u.a. Nutzung von LOINC in CDISC LAB

## ↳ Archivierung

- ↳ TMF-Projekt V042-01 eArchivierung (2006/07)

- ↪ Vorbereitung der Verbünde auf eine einheitliche, Standardbasierte Archivierung ihrer Forschungsdokumente
- ↪ Erarbeitung einer nachhaltigen und langfristigen Lösung unter Berücksichtigung der verschiedenen Aspekte der Archivierungsproblematik

AP 0: Arbeitsprozesse der Archivierung

AP 1: Rechtliche Rahmenbedingungen

AP 2: XML als neues Archivformat

AP 3: CDISC-Format für die Archivierung

AP 4: Konventionelle Formate

AP 5: Wirtschaftlichkeitsanalyse

AP 6: Beurteilung und Handlungsempfehlungen

APXs: begleitende Arbeitspakete

↪ Laufzeit: bis 08/2007

↪ Zielsetzung: Handlungsempfehlung für die Forschungsverbände  
(Schwerpunkt: KKS / klinische Studien)

↪ langfristiges Ziel: Etablierung/Vermittlung von BMBF-  
förderfähigen Archivierungsservices für die medizinische  
Forschung über die TMF

XML wird in verschiedenen Bereichen zukünftig eine wesentliche Rolle spielen

- a) als Dokumentationsformat
- b) als Austauschformat
- c) als Archivierungsformat !

Bsp.:

- ↳ HL7 CDA → Patientenversorgung
- ↳ CDISC (ODM) → klinische Forschung
- ↳ MAGE-ML, RNAML, BIOML etc. → Bioinformatik

- auch außerhalb des Gesundheitswesens (eGovernment, Bibliotheken etc.)



## Extensible Markup Language (XML)

- ↪ XML dient der Formatierung und Strukturierung von Daten
- ↪ XML ist textbasiert
- ↪ Trennung von Formatierung und Inhalt, zum Erfassen des Inhalts sind Verarbeitungsprogramme notwendig
- ↪ XML Markups bestehen **immer** aus einer öffnenden und einer schließenden Marke (<marke></marke>).
- ↪ Dieser Marken dürfen verschachtelt werden

und:

- ↪ XML wird bereits genutzt für die Erfassung und den Datenaustausch klinischer Studien ! (CDISC)

↳ XML-Dokument

↳ Tagdefinitionen

↳ Metadaten

↳ Dokumentinhalt

↳ Zusätzlich ggf. Style Sheet (XSL) u.a. zur Datenpräsentation



Eignung zur Langzeitarchivierung



Metadatenauswertbarkeit



Signaturfähigkeit



↳ Wie sieht die zukünftige Unterstützung der XML-Archivierung bezüglich

↳ Import/Export

↳ Retrieval

↳ Interpretation/Präsentation

↳ Unveränderlichkeit

durch digitale Archivsysteme aus?

➔ Fragen im Workshop vom 07.12.2006 an die Anbieter und Hersteller von digitalen Archivsystemen im Gesundheitswesen



## Tagesordnung:

- ↳ Begrüßung (Schmücker, Semler)
- ↳ Kurzvorstellung TMF (Semler)
- ↳ Vorstellung des Projekts  
eArchivierung der TMF  
(Kuchinke)
- ↳ Fragen an die Softwarehersteller  
digitaler Archivsysteme (Häber)  
- Frage- & Diskussionsrunde  
in 8 Fragekomplexen –
- ↳ Abstimmung des weiteren  
Vorgehens  
(Moderation: Semler)

## Teilnehmer:

### a) Industrie

- ↳ **d-velop** (Kuhlemann, Bönning,  
Sprengard)
- ↳ **Heydt** (Paul)
- ↳ **Marabu** (Winkler, Tünge)
- ↳ **Optimal Systems** (Versteeg)
- ↳ **Siemens** (Priesmeyer, Gaiser)
- ↳ **Windream** (Schlieper)

### b) GMDS AG AKU & TMF

- ↳ Schmücker, Dujat, Häber
- ↳ Semler, Kuchinke

- ↪ Haben Sie in Ihrem Produkt bereits die Archivierung von XML-Dokumenten (z.B. CDA-Dokumente) implementiert bzw. in welcher Weise ist dies in Ihrem Produkt geplant?
- ↪ Haben Sie bereits Projekterfahrungen mit der Archivierung von XML-Dokumenten (z.B. CDA-Dokumenten)?
  
- Ein Archivsystem kann prinzipiell jedes Format speichern.
- Der CDA-Arztbrief als Kommunikationsmedium ist weitgehend umgesetzt
- Die Archivierung erfolgt neben XML zum Teil zusätzlich in pdf/tiff wegen Revisionssicherheit (Problematik Dokumententransformation ist ungelöst)
- XML dient der Auswertung, ist zur Zeit aber kein alleiniges Archivierungsformat. Insbesondere unklar sind die Punkte Beweissicherheit und Langzeitstabilität von XML.

- ↪ Gibt es in Ihrem Produkt die Möglichkeit, XML-Dokumente zu importieren? Werden dazu Konvertierungen in andere Formate notwendig?
  - ↪ Gibt es in Ihrem Produkt Möglichkeiten des Exports von XML-Dokumenten?
- XML-Dokumente können importiert und exportiert werden.
  - Für den Meta-/Indexdatenimport scheint XML sich heute durchgesetzt zu haben.
  - In der Regel werden Metadaten (nicht komplette XML-Dokumente) in einer Datenbank abgelegt, die Archivierung erfolgt zum Teil unter Konvertierung in pdf.
  - Problematisch erscheint die Stabilität bzgl. Viewing.

- ↪ Werden Tag-Inhalte der XML-Objekte zur Indexierung verwendet? Welche?
- ↪ Sind Mechanismen zur Konsistenzsicherung von Tag-Inhalten und Indexdaten denkbar? Wenn ja, welche können Sie sich vorstellen?
- Verwendet werden klassische Importmethoden durch Nutzung von „mitgelieferten“ Daten (z.B. HL7-Message) für die Indizierung des Dokuments.
- Es wird derzeit nicht ins Dokument selber gegriffen (kein Auslesen von Indexdaten aus dem XML-Dokument selbst), auch wenn dies technisch (z.B. über XSLT) möglich wäre.
- Forderung nach klarer Verlässlichkeit von Standards (feste Schemata)

- ↪ Sind Tag-Inhalte der XML-Objekte im Retrieval recherchierbar? Welche Strategien erscheinen Ihnen hierzu denkbar und sinnvoll? (XPath, Nutzung von XML-Datenbanken?)
  - Retrieval nur über Nutzung von strukturierten Indexdaten (aus der Datenbank des Archivsystems).
  - Volltextrecherche über die Objekte selbst ist möglich.
  - Die Frage danach, wo recherchiert wird, ist abhängig von der KIS-Architektur des Hauses.
  - Komplette XML-Tag-Archivierung ist perspektivisch bzgl. Retrieval sinnvoll, aber Technik und Kosten/Nutzen sind zur Zeit noch problematisch.
  - Erste technische „Fingerübungen“ haben zudem Fragen hinsichtlich der Abfrageperformance ergeben.

- ↪ Gibt es bei Ihnen Vorüberlegungen oder Erfahrungen zur digitalen Signatur von XML-Objekten im Archivsystem?
  - ↪ Wie werden XML-Objekte bei Ihnen signiert bzw. welche Mechanismen könnten Sie sich vorstellen?
- Signaturen von XML-Objekten sind nicht im Einsatz, da nicht geklärt ist, was signiert wird und wie die Revisionssicherheit erreicht wird.

- ↪ Gibt es die Möglichkeit, XML-dokumentspezifische XSLs / StyleSheets zu verwalten und für die Anzeige zu nutzen?
  - ↪ Welche Mechanismen werden angeboten zur Überprüfung von digitalen Signaturen aus XML-Dokumenten bzw. XML/XSL/...-Containern?
- XML selber beinhaltet selbst keine Präsentationsschicht. Zur Präsentation bedarf es eines Style Sheets (XSL).
  - Style Sheets müssen zeitpunktorientiert zum Dokument zugeordnet werden können. Dieses ist logisch grundsätzlich möglich (durch Verweis auf zugehöriges XSL im XML), ein dokumentenspezifisches Viewing mit XSL ist derzeit nicht implementiert.
  - Grundsätzlich müssen verbundene Dokumente verwaltet und zur Anzeige gemeinsam verwendet werden können (XML+XSL); dies ist derzeit nicht in der Routine implementiert.
  - Ein solches Viewing-Konzept erscheint für die Langzeitarchivierung als fragwürdig, solange die Rechtssicherheit nicht geklärt ist.
  - Viewing wird in der Regel über konventionelle Formate realisiert.

- ↪ Ist Ihnen die Anforderung, CDISC als (XML-basierten) Standard zur Archivierung zu unterstützen, schon einmal begegnet?
  - ↪ Wie schätzen Sie die Marktrelevanz einer solchen Forderung ein?
- Anforderungen nach CDISC-Archivierung sind nur in Einzelfällen aufgetreten, bietet hohes Entwicklungspotential.

- ↪ Nutzung von XML als Import/Export-Format für **Metatdaten** in der Archivierung ist bereits heute Standard.
  - ↪ Archivierung von XML-**Dokumenten** ist mit herkömmlichen Methoden möglich und verfügbar, wird aber bis dato noch kaum nachgefragt (tendenziell eher außerhalb des Gesundheitswesens).
    - ↪ Grund: unklare Revisionsicherheit von XML-Dokumenten
  - ↪ Spezifischere Methoden sind derzeit nicht implementiert:
    - ↪ standardisierte Nutzung von XML-Tags zur Indizierung der Dokumente – nur projektspezifisch (mit entspr. Aufwand)
    - ↪ Nutzung von XPath zur Abfrage von Dokumenten - nicht
- Gründe:
- ↪ mangelnde Standardisierung und Stabilität von Standards
  - ↪ Fragezeichen hinter Performance solcher Abfragen

- ↳ Signatur von XML-Dokumenten derzeit nicht implementiert.  
Grund:
  - ↳ unklare Signaturstandards von XML-Dokumenten
  - ↳ Frage der Präsentationsschicht (Mitsignieren von XSL, Container)
- ↳ CDISC ist derzeit nur vereinzelt als Anforderung relevant geworden, wird aber als Beitrag zur Standardisierung bei XML-Formaten eingeschätzt.
- ↳ Perspektivisch wird XML-basierte Dokumentation und XML-Archivierung als relevant erachtet (v.a. CDA).
- ↳ Perspektive, über „selbstbeschreibende Dokumente“ (XML) den Mehrwert eines DMS-Systems in service-orientierten Architekturen zu steigern (→ „echtes Content Management“), wird unterschiedlich beurteilt.

XML-Archivierung steht noch am Anfang ! – Offen Fragen:

↳ Reversionssicherheit von XML-Dokumenten

↳ auch: Konvertierung (Transformation) von Dokumentenformaten  
(↔ siehe Projekt TransiDoc)

↳ verbindliche, stabile Standardisierung von XML-Dokumenten  
(CDA, CDISC, ...) –

↳ Reduktion der Freiheitsgrade, Berücksichtigung der Semantik

↳ technische Exploration der Performance von kombinierten SQL-  
/XPath-Abfragen auf Dokumente in einem Archiv/DMS-System.

↳ evtl. Standardisierung von Export/Import-Schnittstellen im  
Archiv/DMS-Bereich

*Herzlichen Dank an alle teilnehmenden Unternehmen !*

Über das Projekt eArchivierung wird weiter berichtet.



Die Ems in Rheine - Panoramablick

*„Jetzt muss man auch mal einen Schlusspunkt ziehen.“  
[Peter Neururer]*

Vielen Dank für Ihre Aufmerksamkeit – und gute Heimfahrt !

Mehr Information:

<http://www.tmf-ev.de/>