



Semantic Data Platform for Healthcare

---

Kurzpräsentation für Vollversammlung  
MUG-IMI, 25.6.2014

---

Stefan Schulz, Markus Kreuzthaler



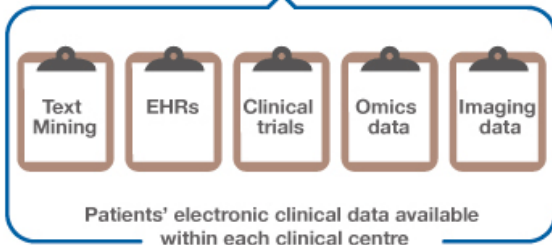
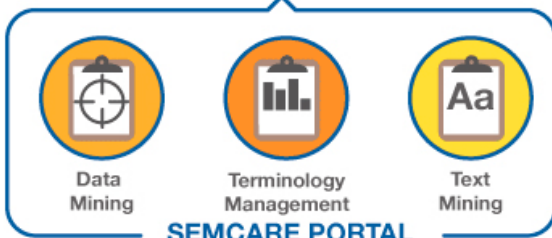
© Copyright 2014-2015 SEMCARE Consortium

- Inhalte der elektronischen Krankenakte in großen Teilen unstrukturiert, heterogen
- Primärnutzung: Krankenversorgung
- Zunehmendes Interesse an so genannter „Sekundärnutzung“
- Bedarf nach strukturierten Abfragen
  - Kohortenbildung für klinische Forschung
  - Unterstützung der Diagnostik, Schwerpunkt „Seltene Erkrankungen“
  - Controlling, Qualitätssicherung etc.
- Anteil an strukturierten und kodierten Inhalten weiterhin unzureichend, nicht operabel, unzuverlässig
- Ähnliche Situation in verschiedenen europäischen Ländern

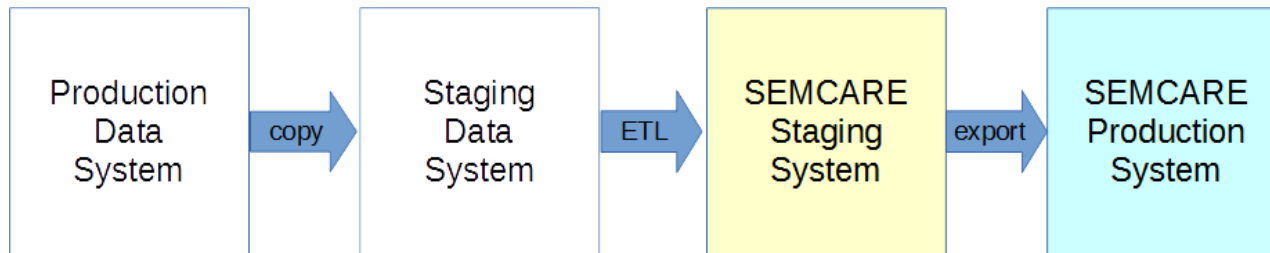
- EU-Projekt:
  - Kurze Laufzeit (2 Jahre)
  - SME Call: konkrete Verwertungsperspektive
- Partner
  - Averbis GmbH – Freiburg (Projektkoordinator)
  - Erasmus Universitair – Rotterdam
  - Medizinische Universität Graz – Graz
  - St. George’s Hospital Medical School – London
  - Synapse Research Management Partners SL – Barcelona
- [www.semcare.eu](http://www.semcare.eu)



Patient selection for clinical studies



- Methoden
  - Fortgeschrittene natürlichsprachliche Technologien für Dokumentenretrieval und Informationsextraktion
  - Standardisierte und benutzerdefinierte Terminologien
  - In Englisch, Deutsch, Niederländisch
- Anwendungsdomäne: Kardiologie
  - Use Cases: Passagerer Bewusstseinsverlust, plötzlicher Herzstillstand, anhaltende ventrikuläre Arrhythmie, Kardiomyopathie, Ischämische Herzerkrankung, Anfallsleiden
- Ziele
  - Vereinfachte Kohortenbildung aus Routinedaten
  - Fernziel: marktfähiges Produkt

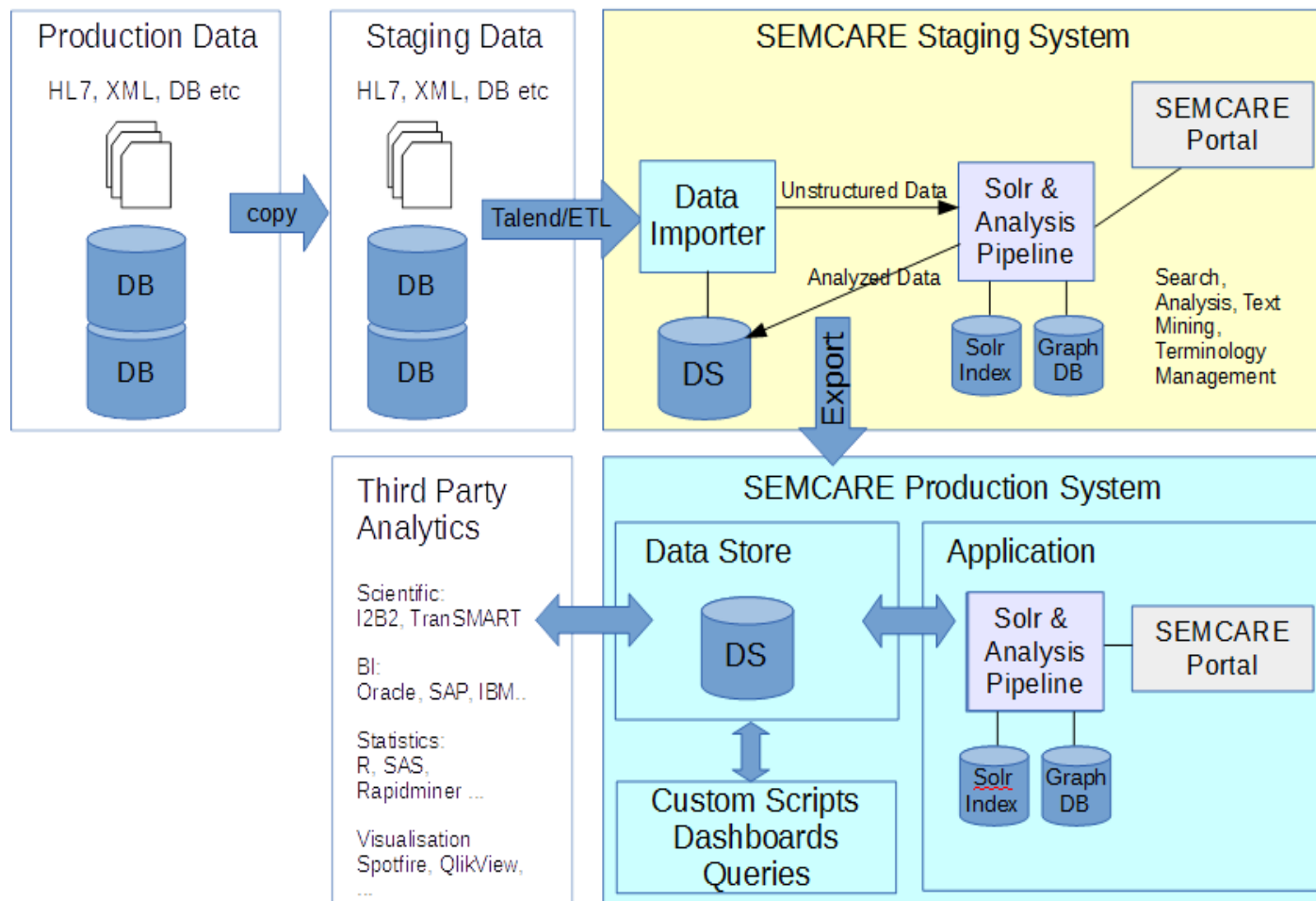


*Datenbanken  
Dateien  
HL7-  
Nachrichten*

*Datenbanken  
Dateien  
HL7-  
Nachrichten*

*Semantische  
Dokumenten-  
Indexierung  
Informations-  
Extraktion →  
Graphdatenbank  
Testumgebung*

*Portal für  
End-User  
Dokumenten-  
und Fakten-  
Retrieval  
Terminologie-  
management  
“Dashboard”*



## Semantische Indexierung

- Terminologiebasierte Dokumenten-Suchmaschine
- Abbildung von Synonymen und Varianten auf sprachunabhängige Deskriptoren
- Myokarditis = Myocarditis = Herzmuskelentzündung → SNOMED-ID: 50920009
- Abstraktion des Dokumenteninhalts, aber keine reale Inhaltsrepräsentation
- Weitgehend automatisierbar
- **Dokumentenretrieval**

## Informationsextraktion

- Füllung definierter «Schablonen»
- Erkennung von Relationierungen → Graph-DB, Triple stores
- #123 isA BloodChemistryResult  
#123 hasTimeStamp 201405061320  
#123 hasPart #3824932  
#3824932 isA Serum-K  
#3824932 hasValue 7.0  
#3824932 hasUnit mval/l

Anfrage: Hyperkaliämie ?

- Erfordert manuelle Feinarbeit
- **Faktenretrieval**

- Terminologie für Echokardiogramm

|                |   |   |   |
|----------------|---|---|---|
| Echocardiogram | echocardiogr*<br>echo<br>TTE<br>heart scan<br>ejection fraction<br>ventricular function<br>ventricular dysfunction<br><br>(...) | normal<br>impaired<br>poor<br>cardiomyopathy<br>aortic stenosis<br>mitral stenosis<br>pulmonary hypertension<br><br>(...) | E |
|                | echokardiogr*<br>echocardiogr*<br>Herz*schall*<br>ejection fraction<br>Auswurf fraktion<br>Kammerfunktion*<br>LV<br>(...)       | normal*<br>beeinträchtigt*<br>gering*<br>Kardiomyopath*<br>Cardiomyopath*<br>*dysfunkt*<br>*fehlfunkt*<br>(...)           | D |



**Text**

Thorax im Liegen auf Intensiv


VA vom selben Tag, 05:56 h. Nach Bronchoskopie nun Wiederentfaltung der rechten Lunge mit nur noch vereinzelte verschattete Herz. Hiliregelrecht. anliegend. Spitze ober

**Manual Annotations**

Annotations

|                          |                |                    |          |
|--------------------------|----------------|--------------------|----------|
| <input type="checkbox"/> | <b>CONCEPT</b> | "Wiederentfaltung" | ORIGINAL |
|--------------------------|----------------|--------------------|----------|

**"Wiederentfaltung"**

**CONCEPT:** Wiederentfaltung 

---

New Annotation

**CONCEPT**

**DIAGNOSIS**

- Anzeigen von nicht lexikalisierten Wörtern / Termen in Beispieltexten
- Entscheidung, ob relevant
- Wenn ja, hinzufügen zu Terminologie

- Architektur
  - semantische Strukturierung
- Medizinische Terminologien – semiautomatische Term-Akquisition
  - Englisch, Deutsch, Niederländisch
- Beispieldaten
  - Kardiologie (Prof. D. Scherr)
- NLP-Pipeline
  - Satzerkennung, Abkürzungserkennung, Kontexterkenkung, Term-Mapping etc.
  - Synergien mit epaNavi - Projekt

- Freitextliche Dokumentation in der med. Dokumentation
  - Bedeutung, Qualität
- Rolle internationaler Terminologien (SNOMED CT)
- Out-of-the-box Lösung – realistischer Ansatz?
  - Arbeitsteilung Ärzte / Spezialkräfte?
  - Vorformulierte Query-Spezifikationen?
- Portierbarkeit
  - Regionale und fachspezifische Unterschiede der Medizinsprache
- Akzeptanz / Transparenz / Zugang zu Daten
  - z.B. undiagnostizierte kardologische Krankheiten bei Patienten in anderen Fachabteilungen
- Tragfähigkeit des Geschäftskonzepts
  - Rolle der forschenden Pharmaunternehmen als potentielle Sponsoren