



Deidentifizierung freitextlicher Daten

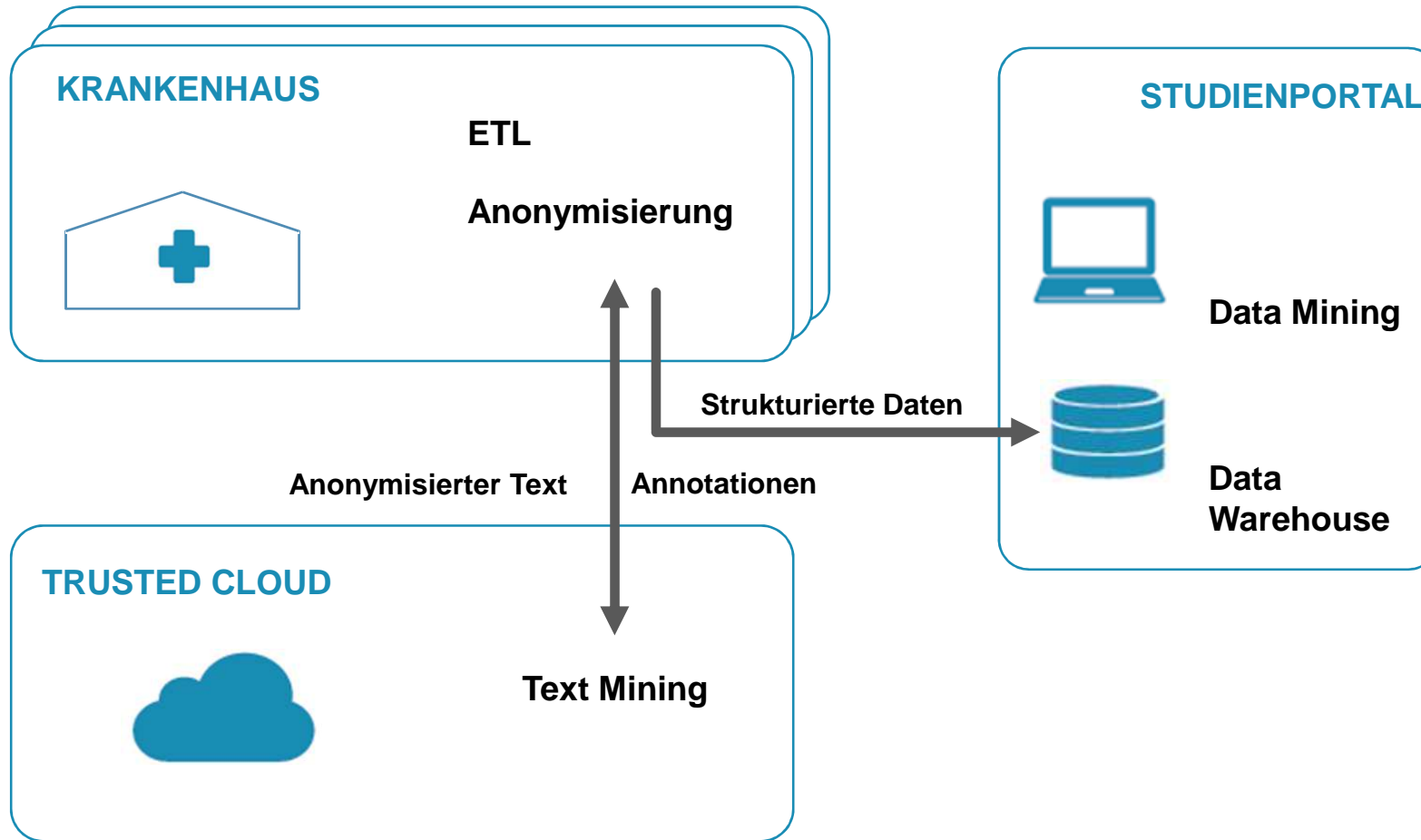
Dr. med. Philipp Daumke

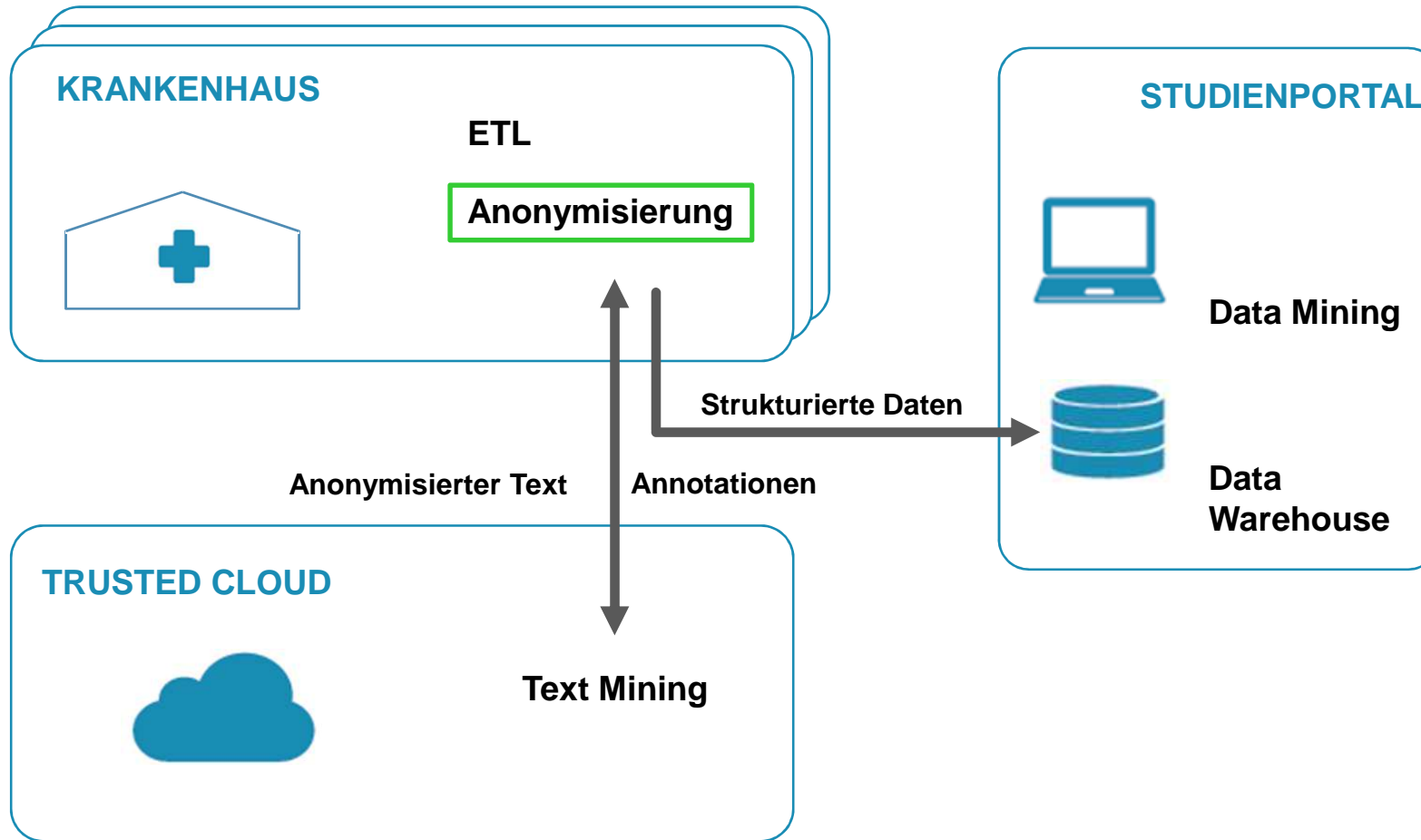
Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages







- Anonymisierung zwingende Voraussetzung für die Freigabe der Daten
- Anforderungen an Anonymisierung
 - Automatisiert
 - Vollständigkeit
 - Manuelle Überprüfbarkeit
 - Dokumentation der Verantwortlichkeit
 - Wiederverwendbarkeit
 - Variable Schutzmechanismen

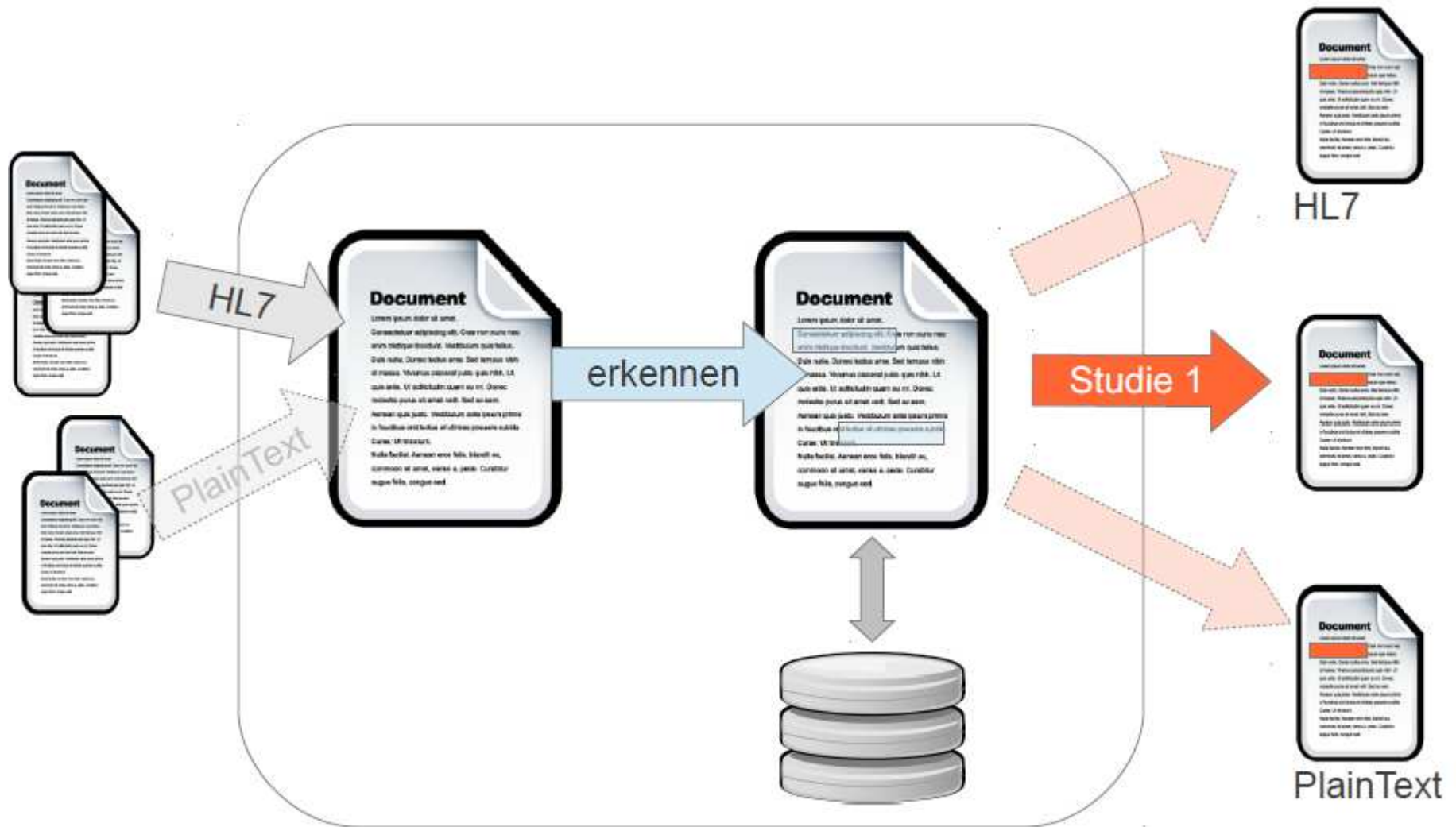


Gemäß HIPAA Safe Harbor Standard:

- (A) **Names**
- (B) All **geographic** subdivisions smaller than a State
- (C) All elements of **dates** (except year)
- (D) **Telephone** numbers
- (E) **Fax** numbers
- (F) Electronic **mail** addresses
- (G) Social security **numbers**
- (H) Medical record **numbers**
- (I) Health plan beneficiary **numbers**
- (J) Account **numbers**
- (K) Certificate/license **numbers**
- (L) Vehicle identifiers and serial **numbers**, including license plate **numbers**
- (M) Device identifiers and serial **numbers**
- (N) Web Universal Resource **Locators** (URLs)
- (O) Internet Protocol (IP) **address** numbers
- (P) **Biometric identifiers**, including finger and voice prints
- (Q) Full face **photographic images** and any comparable images
- (R) Any other **unique identifying number**, characteristic, or code

Merkmale in Kategorien zusammengefasst:

NAME	Namensangaben und Titel
ALTER	Altersangaben und altersbezogene Ereignisse
DATUM	Datumsangaben, Therapiezyklen, Zeiträume
LOCATION	Adressangaben
KONTAKT	Telefon-, Handy- und Faxnummern, E-Mail-Adressen, URLs
DIVISION	Krankenhäuser, Abteilungen, Stationen, Ambulanzen, ...
ID	Fall-, Patienten-, Versicherungsnummern, etc.
BIOMETRIE	Gewicht, Haarfarbe, Rasse, Tattoos, Größe
ANDERE	Versicherungsnamen, Rechtsformen, Religionszugehörigkeit



Metadaten

Reguläre Ausdrücke

Lexikonbasierte Verfahren

Lernende Verfahren

Document | PHI List | Error List

Stadisches Krankenhaus Braunschweig
Abteilung Innere 3
Station DaVinci

Ärztlicher Direktor: Frau Prof. Marta Müller

Arztzimmer Tel. 0121/123 Fax 01217223

Buxtehude, den 17. März 2012
Unser Zeichen: 5k7

Herrn
Med. Martin Müller
Brandt-Allee 12
Berlin

Luise Kisselbach, geb. 17.07.1921
Kollege Müller,

wir berichten Ihnen nachfolgend über o.g. Patienten, der sich am 21.04.2003 in unserer Behandlung befand.

Document | PHI List | Error List

XXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXXX
XXXXXXXXXX XXXXXX X
XXXXXXXXX XXXXXXXX

Ärztlicher Direktor: Frau XXXXX XXXXX XXXXXX

Arztzimmer Tel. XXXXXXXX Fax. XXXXXXXX

XXXXXXXXXX, den XX. XX. XXXX
Unser Zeichen: XXXXX

XXXXX
XXX XXXX XXXXXXX XXXXXXX
XXXXXXXXXXXXXXXXXXXXXX XX
XXXXX XXXXXXX

Betrifft Patienten XXXXX XXXXXXXXXXXX, geb. XX. XX. XXXX

Sehr geehrter Kollege XXXXXX,
wir berichten Ihnen nachfolgend über o.g. Patienten, der sich am XX. XX. XXXX in unserer Behandlung befand.



The image shows a screenshot of the deID tool interface. At the top, there is a process flow diagram with three icons in rounded square boxes connected by right-pointing chevrons. The first icon shows a document with a folder icon, the second shows a document with a red prohibition sign, and the third shows a document with a globe icon. Below this diagram is the text "deID" in a bold, sans-serif font, with the "e" in red. Underneath "deID" is the text "averbis deidentification tool" in a smaller, bold, sans-serif font. Below the text is a login form consisting of three vertically stacked rounded rectangular boxes. The top box contains the text "c4h". The middle box contains ten black dots representing a password field. The bottom box is a solid blue button with the text "Sign in" in white.



My projects

<input type="checkbox"/> ▾	Id	Project name	Gold standard	Description	Documents	Edit	Delete	Import	Export
<input type="checkbox"/>	3	test hl7		Synthetisierte HL7-Briefe.	8				
<input type="checkbox"/>	4	test	5: gold		5				
<input type="checkbox"/>	5	gold			1				
<input type="checkbox"/>	8	test_label			1				

deID
averbis deidentification

Help ▾ c4h ▾

Home » test ⚙️ ▾

“test” documents

Page 1 / 1 Show 10 ▾ documents per page

first prev **1** next last Status filter ▾

<input type="checkbox"/> ▾	☆	Id	Document name	Status	Confidence	Modified by	Last modified
<input type="checkbox"/>	☆	19	OP_Bericht_11.txt	Original	0	c4h	16.12.2014 12:49
<input type="checkbox"/>	☆	20	OP_Bericht_68.txt	Approved	0.889	c4h	20.11.2014 10:44
<input type="checkbox"/>	☆	21	OP_Bericht_51.txt	Finalized	0.820	c4h	20.11.2014 13:32
<input type="checkbox"/>	☆	22	OP_Bericht_45.txt	Auto-processed	0.919	c4h	16.12.2014 12:48
<input type="checkbox"/>	☆	23	OP_Bericht_102.txt	Finalized	0.740	c4h	20.11.2014 13:46

Showing documents 1 - 5 of 5

Select batch process ▾

*** VORLÄUFIGER BERICHT ***

Eingriffs-Nummer: **1105720**
 OP-Datum: **22.11.2013** 18:55:00
 Abteilung: **UC**
 1. Operateur: **Prof. Dr. Gustav Ganz**
 Operateur:
 Assistent: **Dr. Brad Pitt**; . PJ
 Anästhesist: **Maria Furtwängler**; OAV **Dr. Joachim Voigt**
 Anästhesiepflege: **Daisy Ducky**; **Johanna Vernandes**
 Instrumentation: **Sylvester Stallone**
 Springer: **Meg Ryan**
 Lagerung:

Präoperative Diagnose: Luxation einer Hüftgelenktotalendoprothese [Hüft-TEP] (T84.0)
 Postoperative Diagnose: Luxation einer Hüftgelenktotalendoprothese [Hüft-TEP] (T84.0)
 Therapie: Geschlossene Reposition Endoprothese Hüftgelenk ohne Osteosynthese (8-201.g R)

Annotations:		
<input type="checkbox"/>	ID	"1105720"
<input type="checkbox"/>	RE, ML Date	"22.11.2013"
<input type="checkbox"/>	Div.	"UC"
<input type="checkbox"/>	RE Name	"Prof. Dr. Gustav Ganz"
<input type="checkbox"/>	RE Name	"Dr. Brad Pitt"
<input type="checkbox"/>	Name	"Maria Furtwängler"
<input type="checkbox"/>	RE, LI Name	"Dr. Joachim Voigt"
<input type="checkbox"/>	Name	"Daisy Ducky"
<input type="checkbox"/>	Name	"Johanna Vernandes"
<input type="checkbox"/>	Name	"Sylvester Stallone"
<input type="checkbox"/>	Name	"Meg Ryan"

- **Umfassende Software zur Deidentifizierung von Freitexten**
 - Datenhaltung, Prozessunterstützung, ...
 - verschiedene Erkennungsmechanismen
 - unterstützt halbautomatische und automatische Anonymisierung
- **Zahlen & Fakten**
 - Manuelle Nachbearbeitung und Kontrolle: ca. 45 sec (500-600 Briefe pro Tag)
 - sind Metadaten bekannt, nahezu 100%ige Erkennung von Patientennamen
 - Tool übergeneriert bei der Erkennung
 - insg. kein 100%iger Schutz
 - weitere Schutzmaßnahmen erforderlich
- **Bewertung der Datenschützer**
 - Insgesamt sehr positiv, insb. in Verbindung mit zusätzlichen Schutzmaßnahmen
 - Im Detail uneinheitliche Bewertung
- **Adressiert offensichtlich einen großen Bedarf**

FÜR WEITERE INFORMATIONEN



Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

Ansprechpartner:
Dr. med. Philipp Daumke
☎ 0761 - 203 97690
✉ philipp.daumke@averbis.com



Universitätsklinikum
Erlangen

