



## **UIID - Universal Institute Identifier**

Eindeutige Identifizierung von Datensätzen und Proben als Voraussetzung für standortübergreifende nationale und internationale Biobanken

Universitätsklinikum Köln  
Institut für Pathologie  
Centrum für Integrierte  
Onkologie Köln/Bonn  
Kerpener Straße 62  
50937 Köln

# Ziel

Der **UIID (Universal Institute Identifier)** wurde entwickelt, um

**Datenbestände,**

**Dokumente** und

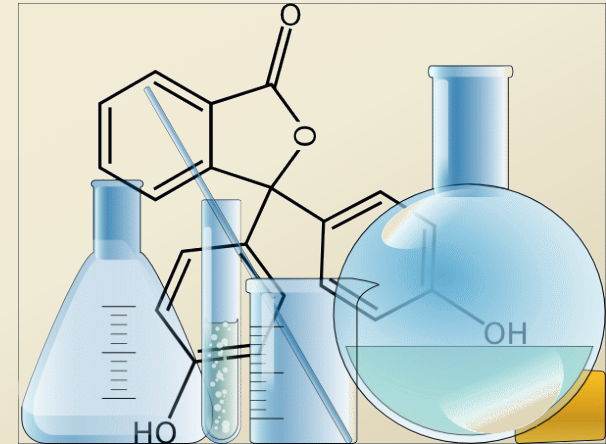
**Materialien**

institutsübergreifend eindeutig identifizieren und

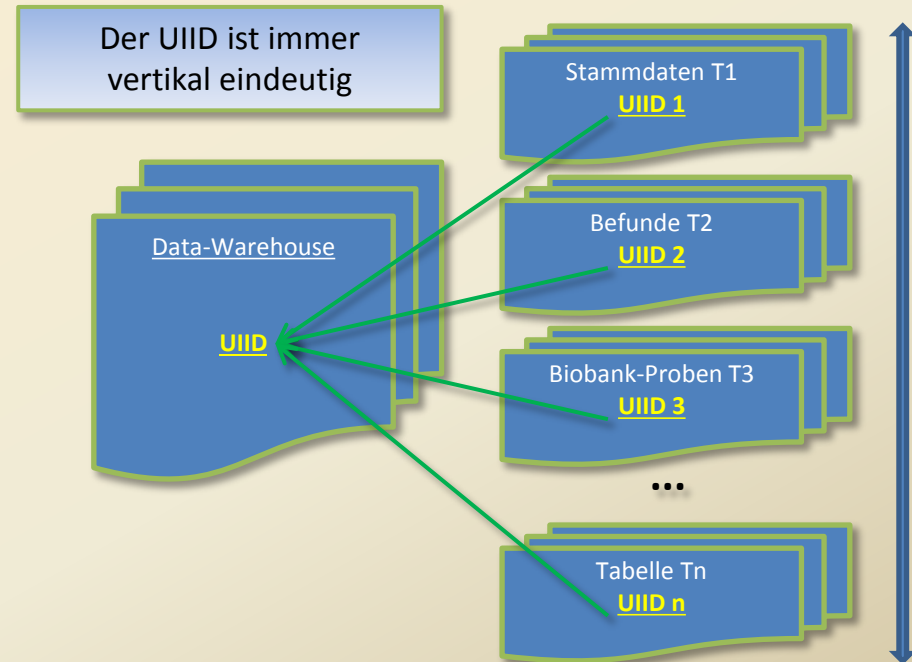
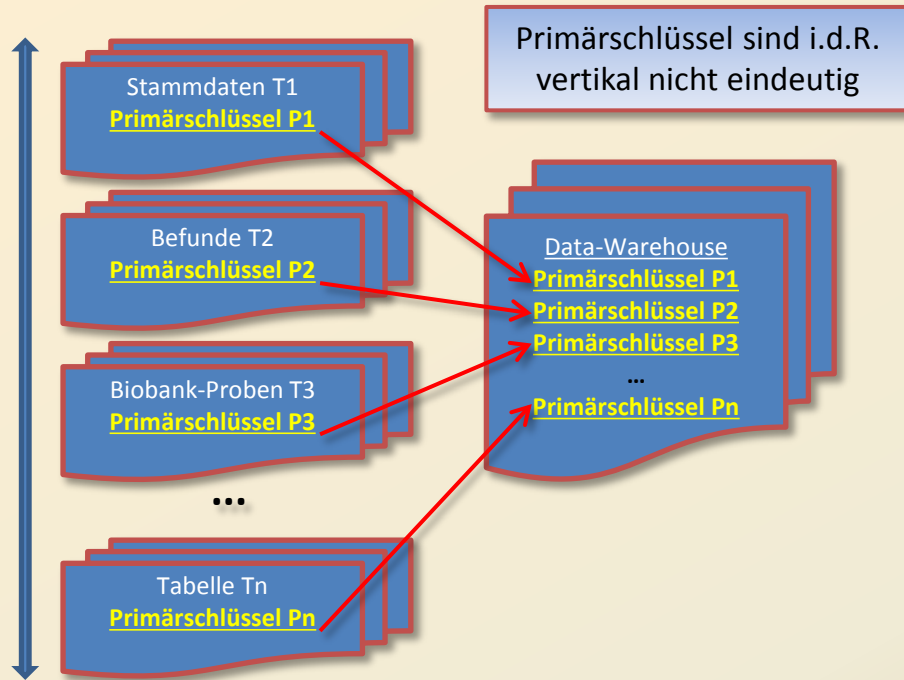
**universitären Einrichtungen,** sowie

**privatwirtschaftlichen Forschungsinstituten**

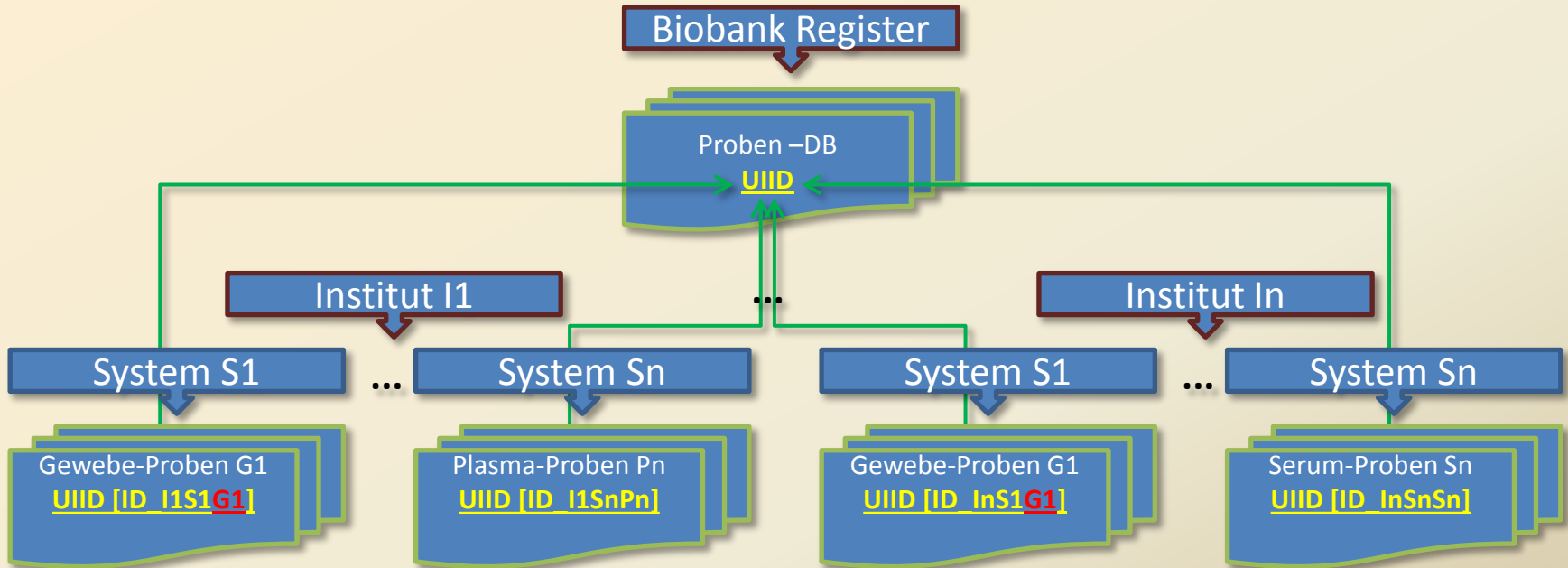
zuordnen zu können.



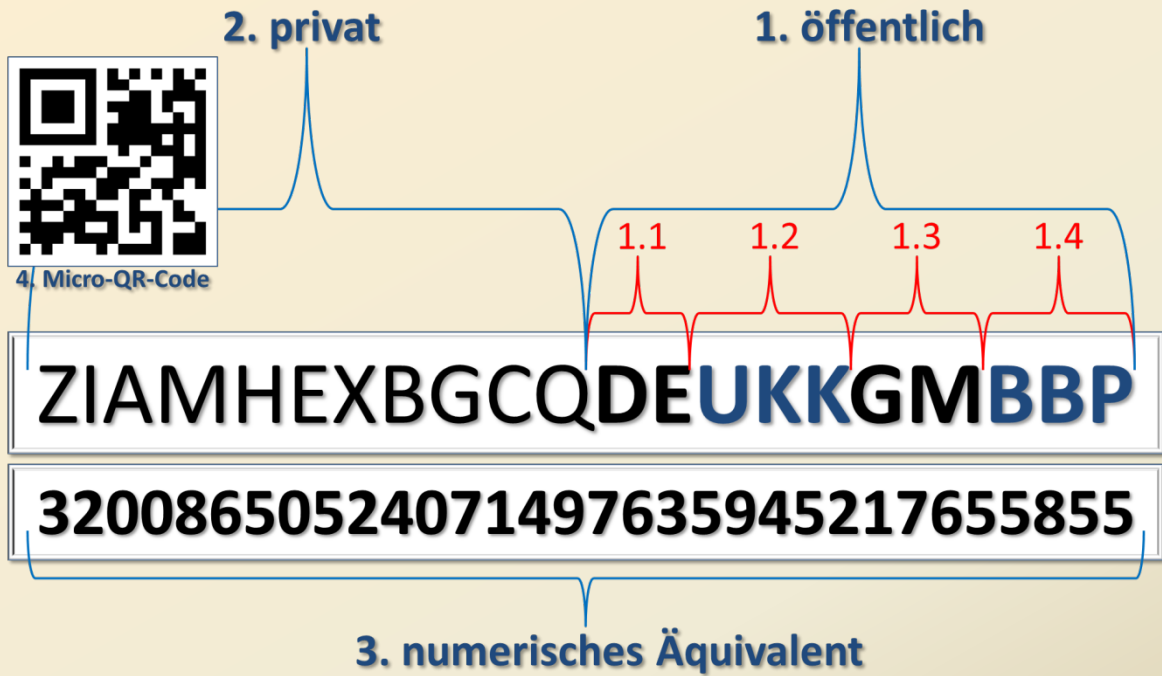
# Vertikale Eindeutigkeit



# Horizontale Eindeutigkeit

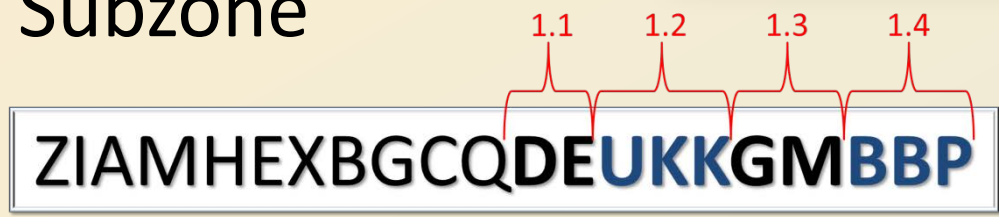


# Aufbau des UIID-Schlüssels



1. Öffentliche Subzone
  - 1.1 – Nationalität
  - 1.2 – Institut
  - 1.3 – System
  - 1.4 – Tabelle
2. Private Subzone
3. Numerisches Äquivalent
4. Micro-QR-Code-fähig

# Die öffentliche Subzone



Element	Beschreibung
1. Länder-Code nach ISO-3166	Da es bereits hinreichend normierte Länder-Codes gibt, sollte man diese nutzen, wie den zweistelligen Länder-Code nach ISO-3166, der für den UUID-Schlüssel Verwendung findet.
2. Der Instituts-Code	Der Instituts-Code sollte pro Land zentral vergeben werden, da er eindeutig sein muss. Der dreistellige Code lässt bis zu 32.768 Institute pro Nation zu, welche den UUID-Schlüssel nutzen können.

Element	Beschreibung
3. Der System-Code	Der System-Code bezeichnet das IT-System, welches den UUID-Schlüssel nutzt. Der zweistellige Code lässt bis zu 1.024 verschiedene Systeme pro Institut zu, welche den UUID-Schlüssel nutzen können.
4. Der Tabellen-Code	Der Tabellen-Code bezeichnet die System-Tabelle. Der dreistellige Code lässt bis zu 32.768 Tabellen pro System zu, welche den UUID-Schlüssel nutzen können.



# Die private Subzone

2. privat

ZIAMHEXBGCQDEU**KKGM**BBP

**Die private Subzone beinhaltet den  
eigentlichen Tabellenindex in  
verschlüsselter Form:**

- So kann z. B. ein bisheriger sequentieller Index 1:1 in die private Subzone übernommen werden.

- Vorgesehen ist ein UTC-Zeitstempel auf Millisekundenbasis nach POSIX-Standard (UNIX-konform) – mit einem Gültigkeitsbereich vom 01.01.1980 00:00:00,000 Uhr bis zum 31.12.2999 23:59:59,999 Uhr.
- Der Gültigkeitsbereich vor dem 01.01.1980 00:00:00,000 Uhr ist für Migrations-Projekte reserviert, so dass zusätzlich bis zu 3 Milliarden Datensätze ohne zeitliche Zuordnung übernommen werden können.

# Sicherheit

Die private Subzone des UUID-Schlüssels beinhaltet eine einfache aber effektive Integritätsprüfung über den gesamten Schlüssel. Diese ermöglicht es Übertragungsfehler aber auch Manipulationsversuche zu erkennen.

Wird z. B. über ein Interface eine Abfrage der UUID-Schlüssel ermöglicht, kann mittels einfacher Mechanismen nach mehrmaliger

Falscheingabe ein Delay automatische Systeme hinreichend ‚ausbremsen‘, wenn sie z.B. per Brute-Force-Methode einen gültigen Schlüssel zu erzwingen versuchen.

Weiterhin ist die private Subzone des UUID-Schlüssels verschlüsselt, so dass keinerlei Rückschlüsse auf den Inhalt des privaten Teils möglich sind, wie Datum und Uhrzeit der Erstellung.



## Informationen zum Aufbau

- Der UUID-Barcode wird im Base32-Format nach RFC 4648 dargestellt. Dieses Format beinhaltet lediglich das Alphabet in Großbuchstaben A-Z und die Zahlen 2-7 (die Zahlen ,0' und ,1' fehlen, da sie je nach verwendetem Zeichensatz leicht mit den Buchstaben ,O' und ,L' verwechselt werden können). Somit beansprucht jedes darstellbare Zeichen 5 Bit. Als Barcode dargestellt, ergibt der 105-bit UUID-Schlüssel genau 21 Zeichen. Das Base32-Format ist das größtmögliche Format, welches uneingeschränkt transparent ist.
- Das Base32-Format garantiert so eine leichte Lesbarkeit um z. B. den Schlüssel bei defektem Barcode manuell zu übernehmen.
- Der UUID-Schlüssel ist ein 105-bit Schlüssel. 105 Bit, um den Schlüssel per Micro-QR-Code darstellen zu können. Da der Micro-QR-Code wesentlich kleiner sein kann, als z. B. ein DataMatrix-Code, sind die Ansprüche an einen Plate-/Rack-Scanner wesentlich geringer, als bei einem DataMatrix-Code.

# Das numerische Äquivalent

**32008650524071497635945217655855**

Neben dem Barcode wird der UUID-Schlüssel auch als numerisches Äquivalent angeboten, um die Effizienz der Systeme zu gewährleisten. In der Regel sind numerischen Schlüsseln zur Indizierung wesentlich leistungsfähiger als alphanumerische Schlüssel.

# Kapazität

Adressraum	Potenz	Kapazität
Gesamt	$2^{97}$	158.456.325.028.528.675.187.087.900.672
pro Nation	$2^{87}$	154.742.504.910.672.534.362.390.528
pro Institut	$2^{72}$	4.722.366.482.869.645.213.696
pro System	$2^{62}$	4.611.686.018.427.387.904
pro Tabelle	$2^{47}$	140.737.488.355.328

Von den 105 Bit des UUID-Schlüssels stehen insgesamt 97 Bit zur Adressierung zur Verfügung. Ein Start-Bit garantiert eine Länge von 21 Zeichen, respektive 32 Dezimalzahlen beim numerischen Äquivalent. 7 Bit sind für die Prüfsumme reserviert, um den Schlüssel auf Integrität zu prüfen.

# Implementierung

## Funktionen:

- UIID\_Get**  
Erzeugung des UIID-Schlüssels
- UIID\_Check**  
Prüfung des UIID-Schlüssels
- UIID\_Convert**  
Integer nach Base32-Barcode  
und vice versa
- UIID\_Info**  
Auslesen der Teilinformationen

## Implementierungs-Varianten:

- Implementierung auf Systemebene**  
Library oder Plug-In
- Externer Service**  
z. B. als Web-Service

Sollte ein entsprechender Schlüssel universell eingesetzt werden, muss das System für verschiedene Plattformen zur Verfügung gestellt werden. Weiterhin sollte dem Qualitätsmanagement besonderes Augenmerk geschenkt werden; bis hin zur ISO-9001 Zertifizierung.

Vielen Dank für Ihre Aufmerksamkeit!

