Privacy and Big Data

GNUmed

Antoine Janssens[1], Maria Karegla[1], Michiel Creemers[2], and Yosra Ben Taarit[2]

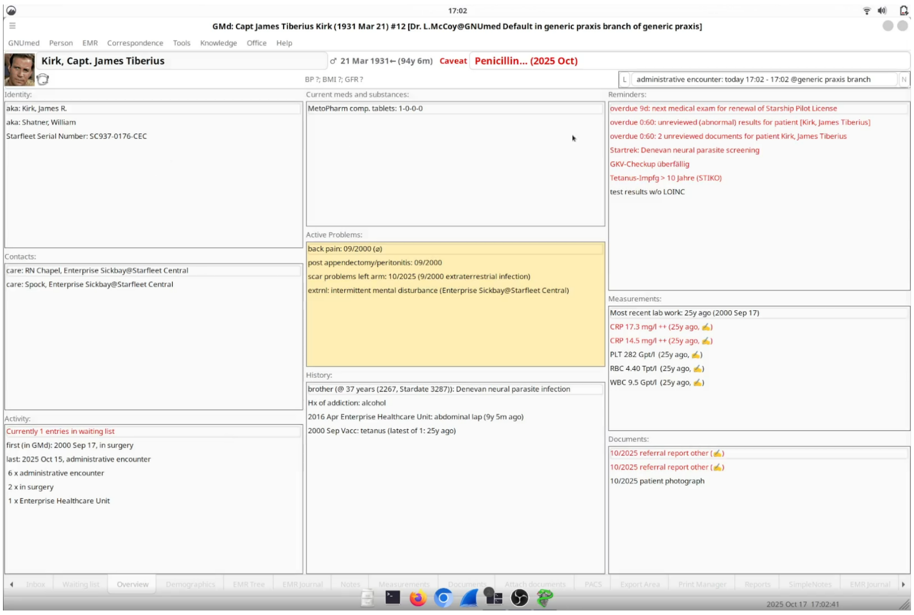
[1]Advanced Master's in Cybersecurity
[2]Master of Artificial Intelligence

October 19, 2025

Abstract

Keywords
Big Data, Health, Privacy, GNU, open-source

1 Introduction



GMd: Capt James Tiberius Kirk (1931 Mar 21) #12 [Dr. L.McCoy@GNUmed Default in generic praxis branch of generic praxis]

Kirk, Capt. James Tiberius

♂ 21 Mar 1931← (94y 6m) Caveat Penicillin... (2025 Oct)

BP ?; BMI ?; GFR ?

administrative encounter: today 17:02 - 17:02 @generic praxis branch

Identity:
aka: Kirk, James R.
aka: Shatner, William
Starfleet Serial Number: SC937-0176-CEC

Contacts:
care: RN Chapel, Enterprise Sickbay@Starfleet Central
care: Spock, Enterprise Sickbay@Starfleet Central

Activity:
Currently 1 entries in waiting list
first (in GMd): 2000 Sep 17, in surgery
last: 2025 Oct 15, administrative encounter
6 x administrative encounter
2 x in surgery
1 x Enterprise Healthcare Unit

Current meds and substances:
MetoPharm comp. tablets: 1-0-0-0

Active Problems:
back pain: 09/2000 (⌀)
post appendectomy/peritonitis: 09/2000
scar problems left arm: 10/2025 (9/2000 extraterrestrial infection)
extrnl: intermittent mental disturbance (Enterprise Sickbay@Starfleet Central)

History:
brother (@ 37 years (2267, Stardate 3287)): Denevan neural parasite infection
Hx of addiction: alcohol
2016 Apr Enterprise Healthcare Unit: abdominal lap (9y 5m ago)
2000 Sep Vacc: tetanus (latest of 1: 25y ago)

Reminders:
overdue 9d: next medical exam for renewal of Starship Pilot License
overdue 0:60: unreviewed (abnormal) results for patient [Kirk, James Tiberius]
overdue 0:60: 2 unreviewed documents for patient Kirk, James Tiberius
Startrek: Denevan neural parasite screening
GKV-Checkup überfällig
Tetanus-Impfg > 10 Jahre (STIKO)
test results w/o LOINC

Measurements:
Most recent lab work: 25y ago (2000 Sep 17)
CRP 17.3 mg/l ++ (25y ago, ✍)
CRP 14.5 mg/l ++ (25y ago, ✍)
PLT 282 Gpt/l (25y ago, ✍)
RBC 4.40 Tpt/l (25y ago, ✍)
WBC 9.5 Gpt/l (25y ago, ✍)

Documents:
10/2025 referral report other (✍)
10/2025 referral report other (✍)
10/2025 patient photograph

Figure 1: GNUmed Screenshot of patient file.

2 Technical description of the system

2.1 System Architecture

Based on the GNUmed introduction document[?], we deduct that GNUmed makes use of a flexible client-server model. Thigh high-level component diagram details their key responsibilities.

The **GNUmed Client** consists of the main GNUmed application. A Database Service Broker is used to abstract the backend service distribution. This provides a unified data source that the rest of the application can make use of, providing transparency to the physical layout for user and developer alike. In adherence with security requirements, a client side cryptography engine handles the encryption of sensitive data, ensuring that even before being transmitted, the data of this nature cannot be accessed.

The **GNUmed Backend** is built on a PostgreSQL database and is not monolithic. It is made up of a collection of distributed services. Data that is related is regarded as a service and hence can be treated as a logical component. Services are not necessarily present on the same physical server. Within the database, business logic if enforced through the use of triggers. These triggers also maintain an Audit Trail, as required legally, by automatically capturing all changes, by keeping copies of the unchanged instances. The tables are highly normalized in order to optimize storage. For the client, data access is simplified through an abstraction layer of updateable views, which sit on top of the tables. Access to these data is managed through a user hierarchy that regulates permissions.

Finally, the **Communication** between the client and server is secured using industry standard protocols, with examples being SSL for encryption and Kerberos for authentication. The architecture also allows for **Optional Middleware**, to be placed between client and backend, allowing for more complex integrations or business logic external to the database.

2.2 The Logical Data Model

The system's UML Class Diagram as shown in Figure **??**, represents the logical data model and details the structure of the data located in the PostgreSQL backend. The figure is outdated and is lacking information. It provides a blueprint for some of the services, as previously described, with each prefix corresponding to a specific schema and hence a service. In this case we have the `Demographic Service`, represented by `dem.`, `Clinical Service` by `clin.`, `Document Storage Service` represented by `blobs.` and `German Specific Schema` represented by `de_de.`

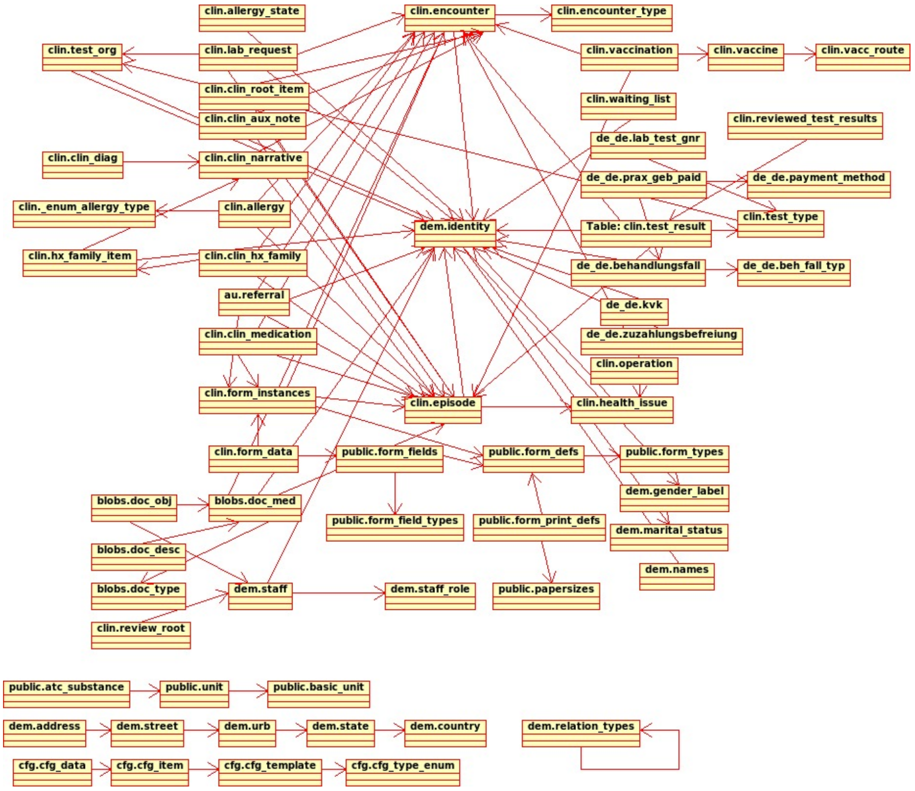


Figure 2: GNUmed UML Class Diagram. Note: outdated (2002) and missing elements [**?**]

3 technical privacy issues

4 technological recommendations

References

[1] Res Medicinae. domain.png (927×798).

[2] GNUmed Development Team. GNUmed-introduction_v4.pdf.