



Vortrag zur Bachelorarbeit

Universität Heidelberg
Fakultät für Mathematik & Informatik

Mario Bijeljanin

**Definition der Anforderungs- und Risikoanalyse sowie
Implementierung einer relationalen Datenbank mit einer
Data Ingest Pipeline für ein Krankenhaus-Triage-System**



Überblick

- TEDIAS
- PROBLEM
- LÖSUNG
- User Interface
- Requirements Analyse
- Risiko Analyse
- Datenbank
- Data Ingestion Pipeline
- Unit Test

TEDIAS

- Digitalisierung der Patientenaufnahme
- Digitale Krankenhaus der Zukunft





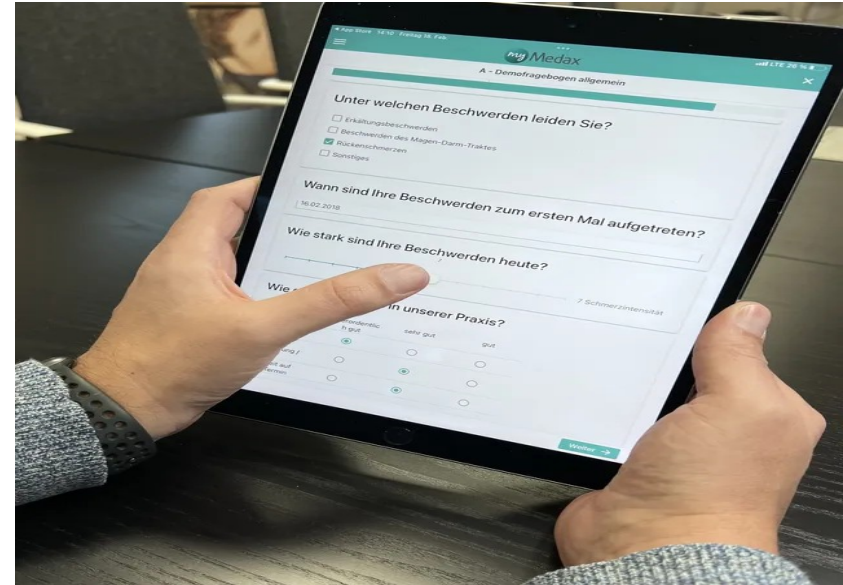
Problem

- Patientenregistrierung in der Notaufnahme (Triage)
- Manuelle Erfassung von Patientendaten
- Lange Wartezeiten für Patienten
- Dringende Fälle mit Wartezeit verbunden

Platzierung	Ranking 2016	Ranking 2018
1	Dänemark, Schweden	Dänemark
2	Estland, Finnland, Slowakei	Finnland, Schweden
3	Portugal	Estland, Spanien
4	Spanien	Schweiz
5	Österreich	Slowakei, Vereinigtes Königreich
6	Schweiz	Portugal
7	Belgien	Frankreich
8	Deutschland, Litauen, Niederlande	Niederlande, Österreich
9	Vereinigtes Königreich	Belgien, Deutschland, Litauen, Polen
10	Italien	Tschechische Republik
11	Frankreich, Slowenien	Italien, Slowenien
12	Polen	Irland
13	Tschechische Republik	
14	Irland	

Lösung

- Wartezeiten für dringende Fälle verkürzen
- Digitaler Tablet Fragebogen
- Priorität des Patienten ermitteln und durch Dashboard darstellen





User Interface

ID	Vorname	Nachname	Besuchgrund	Wartezeit			
e6f93635-cbd6-468b-9fb6-1fabde7b879d	Desire	Holland	Missempfindung (Kribbeln)/Taubheit der Haut	26		Nach Hause geschickt	Weitergeleitet
e6f93635-cbd6-468b-9fb6-1fabde7b879a	Emine	Schmidt	Schwindel	18		Nach Hause geschickt	Weitergeleitet
Seit wann besteht der Schwindel? Länger als 6 Tage ohne kürzliche Verschlimmerung Leiden Sie zusätzlich unter Übelkeit/Erbrechen? Nein Ist der Schwindel dauerhaft vorhanden? Ja Tritt der Schwindel nur bei bestimmten Bewegungen auf (z. B. Umdrehen im Bett, Bücken, Kopfnegung nach hinten)? Ja Haben Sie zusätzliche Beschwerden seit dem Auftreten des Schwindels? Nein							
e6f93635-cbd6-468b-9fb6-1fabde7b879c	Paul	Müller	Krampfanfall	2		Nach Hause geschickt	Weitergeleitet
e6f93635-cbd6-468b-9fb6-1fabde7b879b	Jialu	Wang	Kopfschmerzen	12		Nach Hause geschickt	Weitergeleitet



Requirements Analyse - Überblick

- Ein Dokument mit allen Anforderungen für das Projekt.
- Teil eines strukturierten Projektmanagement Prozesses.
- Ohne Anforderungen können sich große Probleme im Projekt etablieren. Fehlende Ziele /KPIs (Key Performance Indicators) führen zu ineffizienten Projektmanagement/verfehlten Projektzielen.



Anforderungen ermitteln

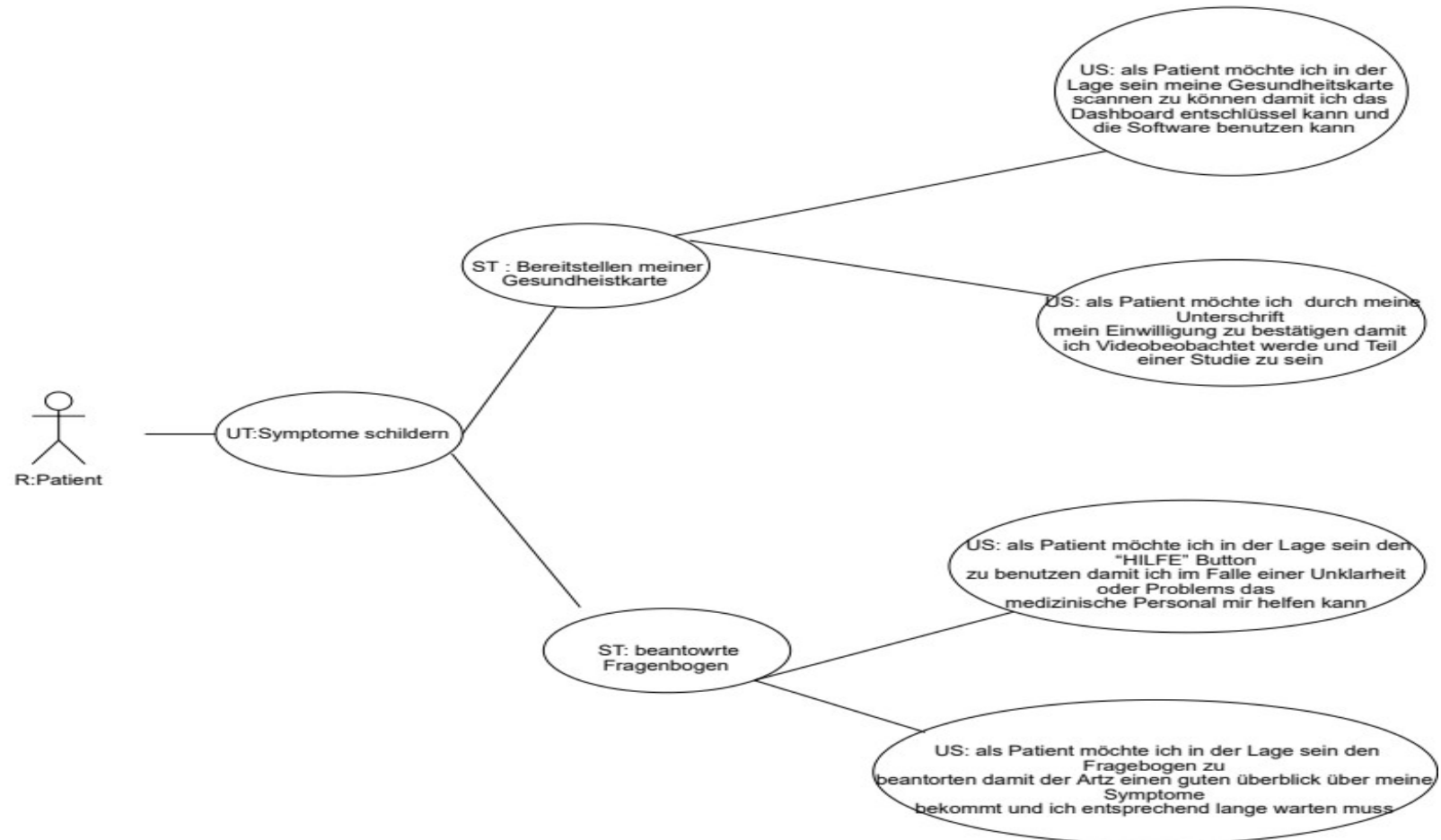
- Interviews mit Projekt Stakeholdern/Product Owner
- Ausführliche Beschreibung des Interviews (User Journeys) für alle Stakeholder (Rollen)
- Ermittlung der Requirements
- Validierung durch Stakeholder
- Dokumentation
- Verwaltung
- Kontrolle



Requirement Typen

- Metriken/KPIs/Purpose
- Business/Regulatorik
- Nicht-Funktional
- Funktional
- Risiken

Patient User Journey





Ermittelte Requirements Zusammenfassung

Functional, Non-Functional, Latenz

Es gibt single choice Fragen , Multiple Choise Fragen , Patient kann sich mehrmals in die Warteschlange einreihen und den Fragebogen ausfüllen, Patienteneinträge können gelöscht werden, Wenn der Patient ein Symptom auswählt, werden nur Fragen zu diesem Symptom angezeigt, einzigartige IDs erstellen, Duplikate vermeiden

NF Der Löschvorgang sollte in einer Zeit von nicht mehr als 200ms erfolgen, Öffnung der Fenster dauert nicht länger als 200ms , Scalability das System muss die größe der eingaben gerecht werden, Reliability das System muss immer bereit sein



Risiko Analyse

Was ist eine Risikoanalyse?

- Potentielle Risikoquellen werden analysiert
- Risiken werden identifiziert und beschrieben
 - User Stories Teil des Prozesses
- Risiko Score wird berechnet
- Projektrisiken können nun besser eingeschätzt und mitigiert werden
- Risiko-Maßnahmen werden ergriffen

Arten von Risiken





Overall Risk Ranking Score

Overall Risk ranking Score = Weighted Risk Threat Score (WRTS) +
Weighted Data Importance Score (WDIS)

Weighted Risk Threat Score (WRTS) = [Risk Probability (RP) + Risk Impact (RI)] * Confidence Level (CL)

Weighted Data Importance Score (WDIS) = [Data Confidentiality (DC) + Data Integrity (DI) + Data Availability (DA)]
* Confidence Level (CL)

Risk	Category	Description	Overall Risk Rank (ORR)	Mitigation Strategy
Implementation	Software funktionalität	Red-Flag-Klassifizierungs- und Ranking-System – Red-Flag-Stapelung und/oder -Schweregrad (angezeigt in roter Farbtintensität) verbessert das Ranking der Warteliste.	32,4	In der nächsten Version sollte eine bessere Red Flag Klassifikation eingefügt werden.
Datenschutz	Datenschutz	Der Benutzer lässt das Dashboard geöffnet und belässt Patienteninformationen auf dem Bildschirm.	26,4	Sperrung nach einem Satz Verspätung (d.h. 5 Minuten) von keiner Aktivität, nach die der Benutzer benötigt um ihre Passwort.
Datenschutz	Daten Display	Patientendaten werden falsch angezeigt.	21,6	Daten sind in PyMySQL geschrieben wo es möglich ist die zu ändern
Dashboard speichert nicht richtig	Datenspeicherung	Das Dashboard ermöglicht die Bearbeitung und das Hinzufügen von Informationen zu einer Patientenakte. Dadurch besteht die Gefahr, dass Daten verloren gehen können, wenn das Dashboard geschlossen oder abgestürzt ist und nicht ordnungsgemäß gespeichert wurde.	21	Das Dashboard soll automatisch alle 60 Sekunden oder wenn Informationen irgendwo eingegeben werden, speichern und ein Backup der Daten auf dem geschützten Krankenhausserver speichern.

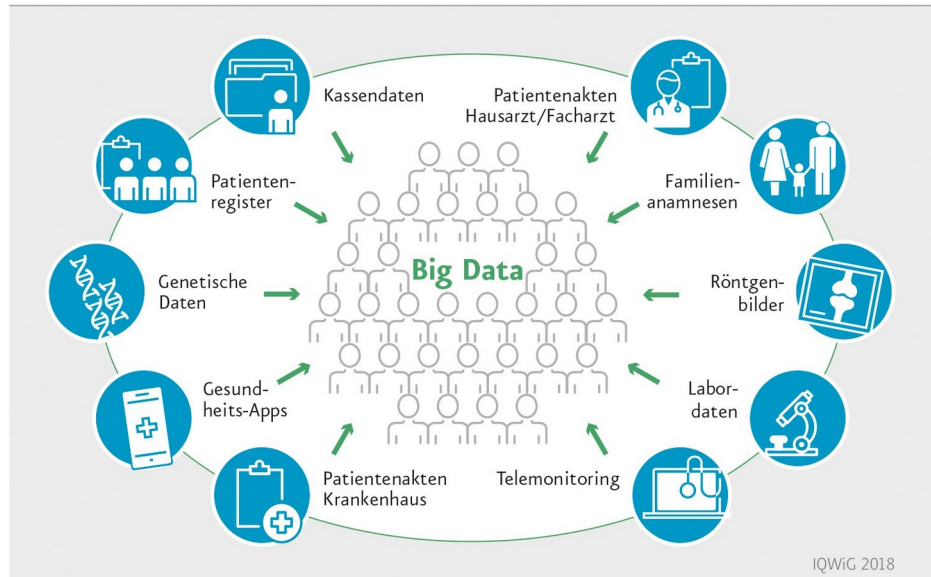
Risk	Category	Description	Overall Risk Rank (ORR)	Mitigation Strategy
Datenschutzrisiko durch die Speicherung von Patienten-Ids	Datenschutz	Dazu gehört die Speicherung von Patientenkennungen in der Software und wie einfach es für einen unbefugten Benutzer ist, Patienteninformationen zu durchsuchen.	20	Die Suche nach Patienteninformationen muss für den Benutzer einfach sein. Es ist nicht ratsam, die Benutzerfreundlichkeit des Dashboards zu beeinträchtigen, um dieses Risiko zu mindern.
Patienten werden vertauscht	Daten Display	Patientendaten werden inkorrekt angezeigt oder gespeichert	19,8	UNDO button in der nächsten Version der Software. Falls aber eine fehlerhafte Eingabe in dem Programm existieren sollte, gibt es eine beliebige Möglichkeit, Kehren Sie die Option(UNDO) um oder löschen Sie die Fehler ohne den Patient aus der Queue zu löschen. Dies sind die Maßnahmen die getroffen werden sollte um eine fehlerhafte Patienten Eingabe zu beheben.
Datenschutzrisiko durch Diebstahl	Datenschutz	Wenn der Laptop oder das Dashboard gestohlen werden, auf dem sich die Software befindet.	18	1) Stellen Sie sicher, dass das Dashboard-Programm zusätzlich zum Laptop selbst ordnungsgemäß kennwortgeschützt ist. Das Dashboard ist browserbasiert, daher sollte ein zusätzliches Passwort für den Zugriff auf das VPN des Krankenhauses enthalten sein. 2) Einen Sticker auf die Rückseite kleben mit der Aufschrift dies ist das Eigentum von der Klinik bitte geben sie es zurück.

Risk	Category	Description	Overall Risk Rank (ORR)	Mitigation Strategy
Datenschutzrisiko durch Subtilität	Datenschutz	Dazu gehört die Möglichkeit, dass jemand dem Benutzer über die Schulter schaut und Patienteninformationen sammelt, auf die er keinen Zugriff haben sollte.	15,5	1)Vermeiden Sie große Anzeigen von Patienteninformationen mit Identifikatoren auf demselben Bildschirm. Vermeiden Sie nach Möglichkeit die Verwendung von Patientenbildern. 2)Der Patient sollte drauf aufmerksam gemacht werden dass andere Patienten eventuell seine Ergebnisse mitlesen können deshalb sollte der Patient vorsichtig sein beim Ausfüllen.
User Eingabefehler	Software funktionalität	Ist ein Eingabefehler schwer zu korrigieren, erhöht sich dieses Risiko.	15,4	Jede Eingabe im Programm sollte eine Möglichkeit haben, den Fehler rückgängig zu machen oder zu löschen.
Implementation	Software funktionalität	Dringende Pop-Up Fenster bei wichtigen Informationen wie Allergien oder anderen Vorerkrankungen	13	In der nächsten Version sollte eine Option eingefügt werden sobald der Arzt auf einen Patienten drückt sollte wichtige Pop-Up Fenster erscheinen

Risk	Category	Description	Overall Risk Rank (ORR)	Mitigation Strategy
Dashboard crashing	Software funktionalität	Eine hohe Priorität wird auch der Internet und Server Verbindung Problemen gewidmet. Das Dashboard ist so konzipiert, dass Ärzte kann aktualisiert werden Informationen zu Patienten bei alle Zeiten. Wenn das Internet oder Backend-Server sind runter, einige Informationen kann nicht aktualisiert werden, und eine Arzt könnte machen Entscheidungen im Gesundheitswesen basierend auf veraltet Information.	11	Um die Risiken zu vermeiden können wir die folgenden Schritte Einfügen um die normale Funktionalität des Krankenhauses zu gewährleisten Eine Warnung auf die Homepage, die zeigt, dass die das Dashboard ist derzeit offline und ein Datum/eine Uhrzeit, zu der es wurde zuletzt aktualisiert. So dass Krankenhaus personal möglichst leicht wieder den Überblick bekommt und die Patient bestmöglich versorgt sind.
Datenschutzerklärung	Datenspeicherung	Patienten sind nicht bereit ihre persönliche Daten preiszugeben	2,5	Bei Datenschutzbedenken kann ein Arzt ein persönliches Gespräch mit dem Patienten zur Thematik führen.

Implementierung SQL Datenbank für Dashboard

Quellen für Big Data im Gesundheitswesen – einige Beispiele



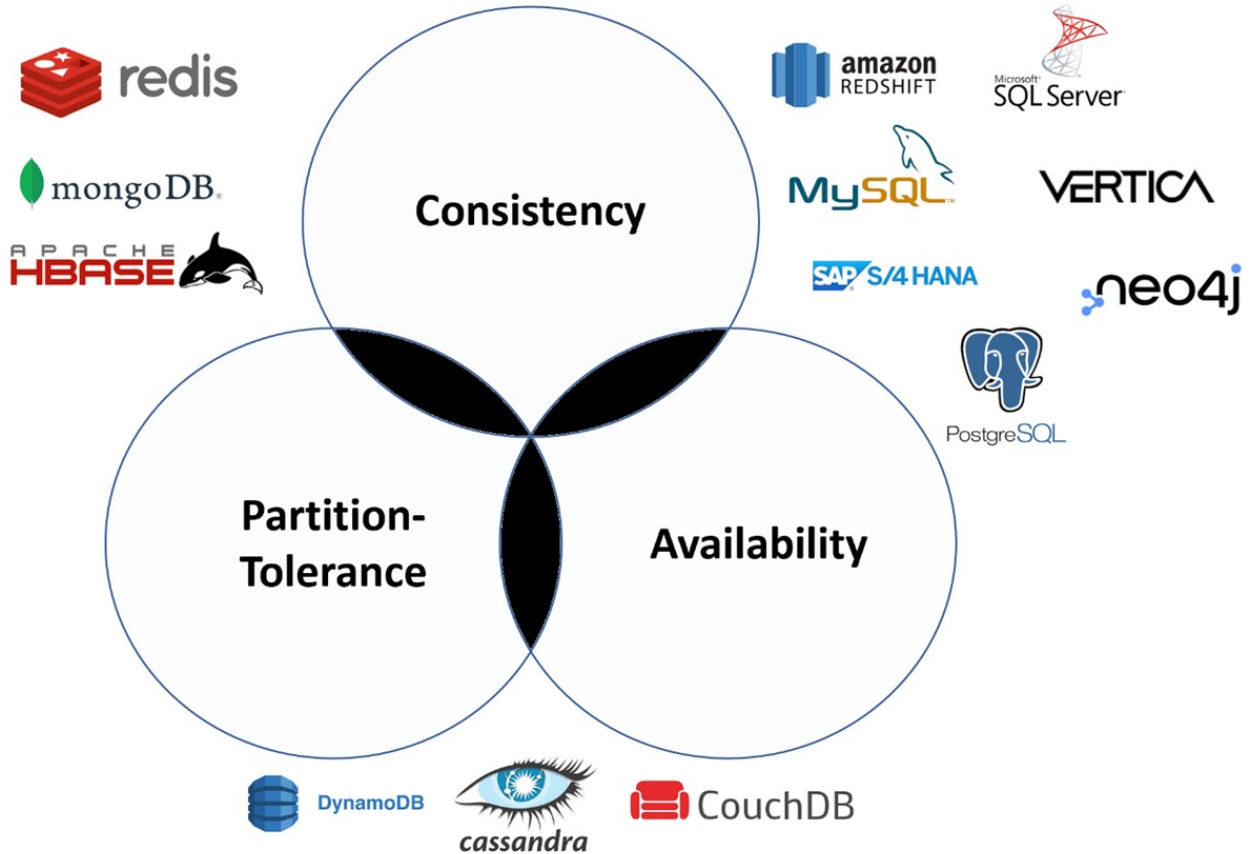


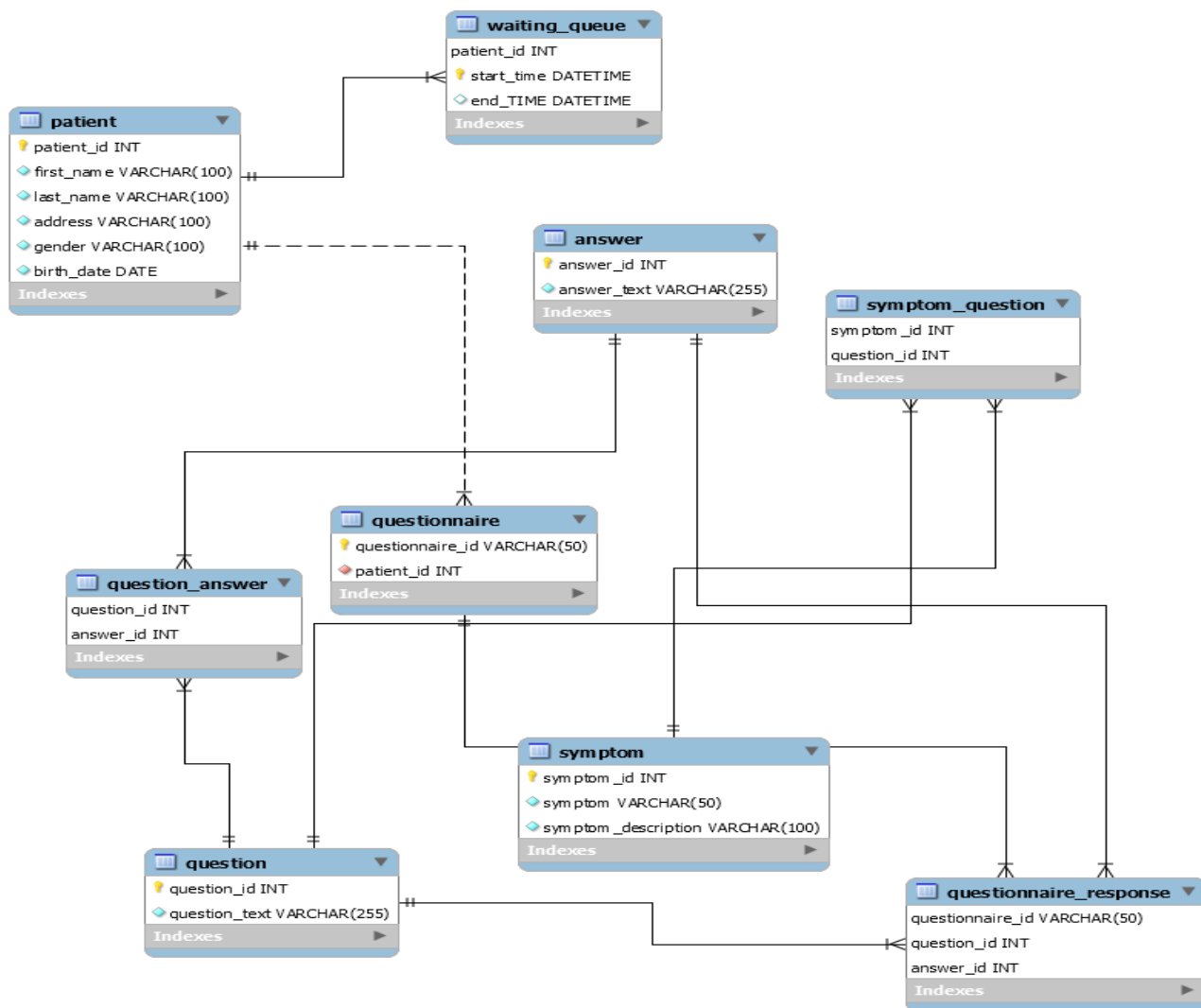
Datenbank

- SQL
- CAP Theorem
- Schema on read vs Schema on write
- Statistiken
- 3 normal norm (Normalisierung)



CAP Theorem



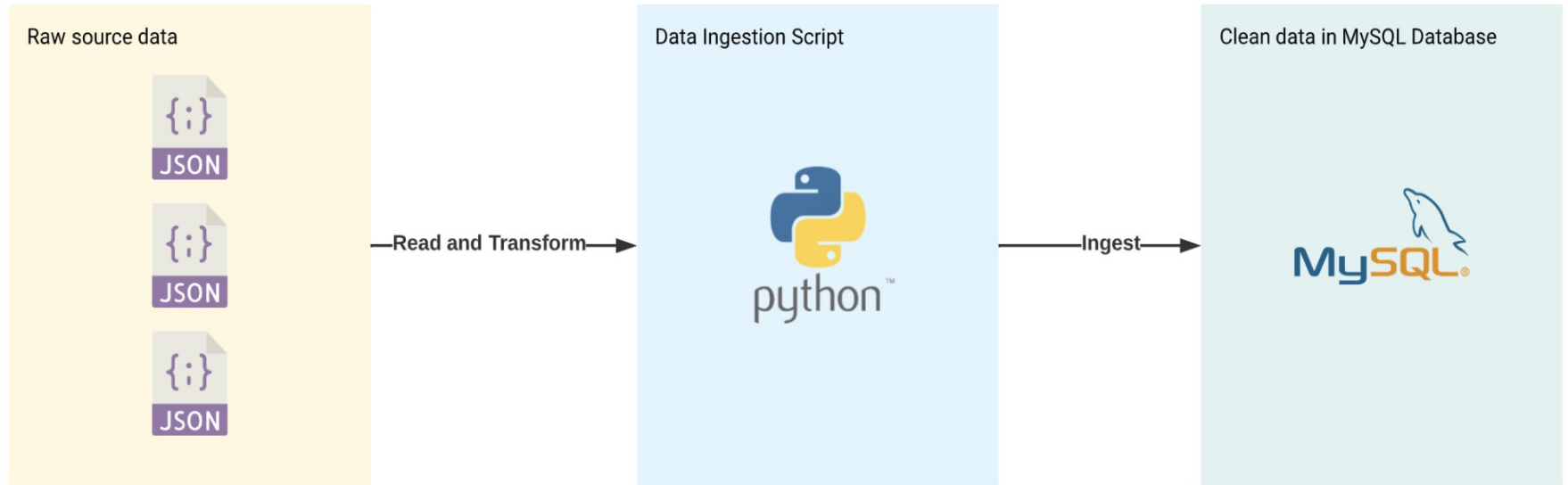




Python Data Ingest Pipeline

- Die rohen Quelldaten werden in mehreren JSON-Dateien abgespeichert (Questionnaire Template & Submitted Patient Questionnaire).
- Das Python Skript liest beide Dateien ein (Extract), kombiniert die Felder und transformiert die Daten in das passende Schema (Transform) und lädt diese in die MySQL-Datenbank (Load) ein.
- Die MySQL-Datenbank speichert die Daten im vorgegebenen Schema ein (Clean Data) - Datenintegrität forciert
- Das Python-Datenerfassungsskript ist idempotent, was bedeutet, dass wir bei mehrmaliger Ausführung des Skripts keine doppelten Zeilen in der Datensenke (MySQL DB) erhalten.

Python Data Ingest Diagramm





Unit Tests

- ist ein Softwareentwicklungsprozess bei dem Code Elemente auf ihre ordnungsgemäße Funktionalität hin untersucht werden
- tragen dazu bei, frühzeitig Fehler im Code aufzudecken.

Was haben wir getestet

1. Testet das Verhalten der Frageparserfunktion, wenn die Eingabe leer ist
2. Testet, ob alle Tabellen aus der ddl.sql Datei erfolgreich erstellt wurden
3. Testet, ob Frage- und Antwortdaten in die DB eingefügt werden



Danke für die Aufmerksamkeit



Back Up Slides

