

# SGDb: Semantic Video Game Database

## Technical Report

Anastasia Chernysheva   Jakob Götz   Ardian Imeraj   Patrice Korinth   Philipp Stangl  
*a.chernysheva@oth-aw.de   j.goetz2@oth-aw.de   a.imeraj@oth-aw.de   p.korinth@oth-aw.de   p.stangl1@oth-aw.de*

**Zusammenfassung—Dieser Technical Report beschreibt die Architektur von SGDb – eine webbasierte Anwendung mit einer Graphen-basierten Suche von Videospielen.**

### I. EINFÜHRUNG UND ZIELE

In den weiteren Abschnitten des Technical Reports wird zuerst auf die Lösungsstrategie in Abschnitt II eingegangen. Im nächsten Abschnitt IV wird das Gesamtsystem aus Bausteinsicht beschrieben. Anschließend wird in Abschnitt V die Verteilungssicht der Anwendung beschrieben. In Abschnitt VI werden die angewandten Werkzeuge zur Entwicklung der Anwendung vorgestellt. Abschließend wird ein Fazit und Ausblick in Abschnitt VII gegeben.

### II. LÖSUNGSSTRATEGIE

Dieser Abschnitt enthält einen stark verdichteten Architekturüberblick. Eine Gegenüberstellung der wichtigsten Ziele und Lösungsansätze.

### III. DATENBESCHAFFUNG

Für die Datenbeschaffung wird eine Client-Anfrage an die Endpunkte der IGDB-API **igdb-api** gesendet. Als Antwort wird im Anschluss der Bearbeitung ein Datensatz zurückgegeben, der wesentliche Informationen zu einem Spiel beinhaltet. Zu den Informationen zählen: Spiele-ID, Spielertitel, Beschreibung, Veröffentlichungsdatum, Genre, Spieleplattform, Bewertung, Entwicklername- und Land sowie eine URL des Coverbildes.

Der Datensatz wird im JSON-Format gespeichert und auf 500 Spiele, die eine hohe Bewertung erzielt haben (>85), reduziert. Begründet wird diese Bedingung durch das Bestreben, einen Datensatz zu erzeugen, der über viele Informationen zu den einzelnen Spielen verfügt. Der Datensatz weist an mancher Stelle numerische Werte auf, die jedoch nicht informativ wären für einen Durchschnittsnutzer. So ist das Veröffentlichungsdatum im Unix-Zeitstempel angegeben. Dieser ist auch als "The Epoch" bekannt und zählt die Anzahl der vergangenen Sekunden seit 1. Januar 1970 **unix**. Der Standort des Entwicklerunternehmens wird als dreistelliger Country-Code definiert, der von der International Organization for Standardization (ISO) entwickelt und im ISO-3166 publiziert wurde **iso**. Es wurde eine Funktion implementiert, die die Unix-Zeit in ein reguläres Zeitformat DD-MM-YYYY und den ISO-Country-Code in Ländernamen konvertiert.

Der Datensatz wird anschließend mit „Refine“ **refine** strukturiert und in RDF-Triples (Graphs) umgewandelt.

### IV. BAUSTEINSICHT

Diese Sicht zeigt die statische Zerlegung des Systems in Bausteine sowie deren Beziehungen.

#### A. Gesamtsystem

#### B. Datenbank

Für die Datenverwaltung und Vernetzung der Informationen wird die Ontotext GraphDB **graphdb** verwendet. In Ontotext werden die Daten in Form von RDF-Triplen gespeichert, die aus einem Subjekt, einem Prädikat und einem Objekt bestehen. Diese Triples werden in sogenannten Repositories gespeichert, die ähnlich wie Tabellen in relationalen Datenbanken organisiert sind.

#### C. Backend

Das Backend ist mit der Programmiersprache Python und dem Framework „FastAPI“ **fastapi** realisiert. Es dient als Schnittstelle zwischen Datenbank und Frontend. Zusätzlich ist es für die Verarbeitung von Ressourcen zuständig. Hierunter fällt die Aufgabe, Daten vom Frontend so aufzubereiten, dass diese für eine Datenbank-Abfrage genutzt werden können. Des Weiteren werden Ergebnisse aus einer Datenbank-Abfrage vom Backend aufbereitet und über einen Endpunkt dem Frontend zur Verfügung gestellt. Das Format der Rückgaben ist bei jedem Endpunkt JSON. Das Frontend kommuniziert mit dem Backend über eine RESTful-API, die mithilfe des Frameworks implementiert ist.

Die API ist wie folgt aufgebaut::

Method	Route	Description
GET	/	Startpage
POST	/	Filter
GET	/search/{search}	Search
GET	/detail/{game}	Detailpage

Die Startpage Route gibt als key die Jahre beginnend von 1985 bis 2022 zurück. Jeder key enthält als value die Titel, welche in diesem Jahr veröffentlicht wurden. Die Filter Route gibt, nachdem sie Daten im JSON Format mit den entsprechenden Filtern erhalten hat, einen aktualisierten Graphen zurück, der die angegebenen Filter berücksichtigt. Das Format und die Struktur des Graphen sind identisch mit demjenigen, der von der Startpage zur Verfügung gestellt wird. Die Search Route verarbeitet eine Suchanfrage und gibt drei Listen zurück. Die erste Liste enthält alle Spielertitel, welche auf die gestellte Suchanfrage zutreffen. Die zweite Liste enthält alle Spielertitel,

welche auf die gestellte Suchanfrage zutreffen und im aktuell angezeigten Graphen zu finden sind. Die dritte Liste enthält alle Spieletitel, welche auf die gestellte Suchanfrage zutreffen und nicht im aktuell angezeigten Graphen zu finden sind. Denn durch die Einstellungen der Filter Route kann es passieren, dass gesuchte Spiele nicht im angezeigten Graphen enthalten sind. Die Detailpage Route gibt alle in der Datenbank gespeicherten Details zu einem Spiel zurück.

Zur Interaktion mit der Datenbank (RDF Triplestore) wird die Python-Bibliothek „SPARQLWrapper“ **sparqlwrapper** genutzt. Diese dient als Python-Wrapper für SPARQL und ermöglicht so die Ausführung von Queries in der SPARQL-Syntax. Dazu bietet es die Möglichkeit, das Ergebnis in das gewünschte Format zu formatieren, sodass diese Daten leichter in Python weiterverarbeitet werden können. SPARQL selbst ist eine graphenbasierte Query-Sprache, die mit Daten im RDF-Format arbeitet.

Zuerst wird ein Graphenobjekt definiert, über welches der Zugriff auf die Datenbank realisiert wird. Der Rückgabotyp ist auf das JSON-Format festgelegt. Über das definierte Graphenobjekt können nun Abfragen direkt in der Python Datei vollzogen werden. Die Antwort der Abfragen ist eine geschachtelte JSON-Ausgabe, sodass diese noch weiter formatiert wird.

#### D. Frontend

Das Frontend ist für die Benutzerschnittstelle verantwortlich. Hauptbestandteile des Frontends sind die Suchmaske, der Graph und die Detailseite zu einem Videospiel. Für die Darstellung des Graphen und die Interaktion mit diesem wird „Sigma.js“ **sigma** verwendet. Sigma.js rendert Graphen mit WebGL. Damit lassen sich größere Graphen schneller zeichnen als mit Canvas- oder SVG-basierten Lösungen. Das Graphenmodell wird in einer separaten Bibliothek namens „Graphology“ **graphology** verwaltet. Dies ist eine Standardbibliothek mit Algorithmen aus der Graphentheorie und allgemeinen Hilfsprogrammen wie z.B. Graphengeneratoren.

Generierung Für das Graphen Layout wird der ForceAtlas2 Algorithmus **forceatlas2** verwendet. Vor dem Start von ForceAtlas 2 Layout muss die Startposition jedes Knotens festgelegt werden. Daher müssen zwei Attribute namens x und y für alle Knoten des Diagramms definiert werden.

Dazu wird ein Graph Objekt mit dem Random layout erzeugt. Positionierung jedes Knotens, indem die Koordinaten gleichmäßig nach dem Zufallsprinzip auf dem Intervall [0, 1) ausgewählt.

Die Interaktion mit dem Backend erfolgt über die REST-Schnittstelle.

### V. VERTEILUNGSSICHT

Zentraler Bestandteil der Verteilungsstruktur sind Docker-Container. Jeder Baustein von SGDb ist für sich isoliert in einem eigenen Docker-Container untergebracht. Damit das gesamte System mit allen verteilten Komponenten in der korrekten Reihenfolge gestartet wird, werden die Docker-Container mit Docker Compose orchestriert. Docker Compose übernimmt die Netzwerkkonfiguration und die Vergabe von

Host-Namen an die jeweiligen Docker-Container. Darüber hinaus können Umgebungsvariablen verwaltet werden.

### VI. ENTWICKLUNGSWERKZEUGE

#### A. Backend

Im Backend wird als Test-Framework „Pytest“ **pytest** genutzt, welches bereits in FastAPI integriert ist.

#### B. Frontend

Im Frontend wird als Test-Framework Vitest in Kombination mit Svelte Testing Library verwendet.

### VII. FAZIT UND AUSBLICK

Im Rahmen der Arbeit wurde mit einem begrenzten Datensatz von 500 Spielen eine erste Version der Anwendung entwickelt. Diese ist voll funktionsfähig und erfüllt alle Anforderungen des Fachkonzepts. Zusätzlich konnten einige Optionale-Anforderungen verwirklicht werden. Somit konnte eine bis dato einmalige Anwendung umgesetzt werden. Die Anwendung ist darauf ausgelegt, auch ohne Komplikationen mit einem größeren Datensatz zu arbeiten.