

Virtuelle Forschungsdatenumgebungen

Struktur, Erkenntnis, Nachnutzung

Rosenkranz, Vinzenz

vinzenz.rosenkranz[at]uni-tuebingen.de
eScience-Center Universität Tübingen, Deutschland
ORCID-ID: 0000-0002-5387-425X

Schmidt, Karsten

karsten.schmidt[at]uni-tuebingen.de
eScience-Center Universität Tübingen, Deutschland
ORCID-ID: 0000-0003-0337-3024

Zusammenfassung. Das Forschungsdatenmanagement steht am Beginn eines jeden Forschungsvorhabens und sollte bis zum Ende mitgedacht werden. Eine Forschungsdatenumgebung unterstützt die Forschenden bei der Strukturierung und Analyse ihrer Daten. Diese Systeme werden meist erst während des Projektzeitraums entwickelt. Das „Wie“ und „Was“ muss also bei jedem einzelnen Projektbeginn entweder neu entworfen oder in ähnliche vergangene Formen eingefügt werden. Das am eScience-Center der Universität Tübingen entwickelte System **Spacialist**, verlagert die Forschungsumgebung in den virtuellen Raum. Dadurch ermöglicht es ein ortsungebundenes und kollaboratives Arbeiten. Die Datenaufnahme in das generisch definierte Datenmodell standardisiert die Datenstruktur und garantiert dadurch eine qualitative Analyse der Daten. Die Forschungsdaten können im System kommentiert, validiert und gegebenenfalls korrigiert werden. Dabei unterstützen ein integriertes Moderationssystem mit Rollenvergabe, eine Kommentarspalte sowie Möglichkeiten zur Referenzierung von Literatur oder auf in der integrierten Dateiverwaltung hinterlegten Dateien. Ebenfalls können Geodaten und benutzerdefinierte Kartenlayer verwendet werden, um raumbezogene Daten zu erfassen und einzubeziehen. Die Entwicklung von *Spacialist* ist von Anfang an transparent gestaltet, um das System auf die Workflows der Fachbereiche auszurichten. Die dabei einbezogenen Fachcommunities der Digital Humanities haben darüber hinaus die Möglichkeit *Spacialist* durch Plugins zu erweitern und an spezifische Anforderungen anzupassen. Zudem können die Forschungsergebnisse unkompliziert in bestehende (universitäre) Archivsysteme integriert werden.

Eine virtuelle Forschungsdatenumgebung dient der systematischen Erfassung von Forschungsdaten, im besten Fall über den gesamten Lebenszyklus der Daten hinweg. Angereichert mit einer Bandbreite unterschiedlicher Informationen und Formaten, können diese interdisziplinär, individuell, aber auch generisch, eingesetzt werden. Dabei findet die Integration in den Forschungsprozess bereits direkt bei der Aufnahme statt und ermöglicht dort eine Anreicherung mit essenziellen (administrativen, inhaltsbeschreibenden) Metainformationen sowie die Verknüpfung zu wissenschaftlichen Identifikationssystemen (wie beispielsweise ORCID). Dabei stehen die Anforderungen des Forschenden selbst, seine Organisation und individuelle Strukturierung, die gerade bei großen Datenmengen und unterschiedlichen Formaten wichtig ist, im Zentrum unserer entwickelten virtuellen Forschungsumgebung **Spacialist**¹ (Abb. 1). Zusätzlich ermöglicht es diese Forschungsumgebung direkt bei der Eingabe erste Analysen, Korrekturen und Validierungen vorzunehmen. Dabei kombiniert **Spacialist** sowohl datenbank-technische Aspekte für Erfassung und Quellennachweise als auch Visualisierungskomponenten, wie z.B. Raumdaten und 3D-Objekte. Für die technische Umsetzung haben wir großen Wert auf Flexibilität gelegt und damit den Fokus auf Software gesetzt, die dem Entwickler diese Möglichkeiten bietet. Deshalb verwenden wir als Grundlage für **Spacialist** das etablierte Datenbanksystem **PostgreSQL**², welches mit der Erweiterung **PostGIS** bestens für das Verarbeiten von Raumdaten gerüstet ist. Bei der Entwicklung des Frontends, stützen wir uns auf etablierte Tools und Frameworks mit großer Community. Für das Erscheinungsbild nutzen wir **Bootstrap**³, welches eine schnelle und einheitliche Entwicklung von UI-Komponenten erlaubt. Als Kommunikationselement mit der Datenbank setzt unsere Umgebung auf das **PHP-Webframework Laravel**⁴, welches eine Vielzahl an Funktionen anbietet, die über die einfache Kommunikation mit der Datenbank hinaus gehen. Dessen große Community erweitert Laravel stetig durch Plugins, die offen zur Verfügung gestellt werden. Auf der Benutzerseite setzen wir das **JavaScript-Framework Vue.js**⁵ ein, welches sich neben den großen Konkurrenten **React** (Facebook) und **Angular**

¹ eScience-Center Homepage, "Spacialist".

² PostgreSQL Homepage, "PostgreSQL: The world's most advanced open source database".

³ Bootstrap Homepage, "Bootstrap · The most popular HTML, CSS, and JS library in the world.".

⁴ Laravel Homepage, "Laravel - The PHP Framework For Web Artisans".

⁵ Vue.js Homepage, "Vue.js".

(Google) etabliert⁶ hat und bereits bei großen Anwendungen, wie *Nextcloud*, zum Einsatz kommt. Auch *Vue.js* lässt sich durch Community-Plugins oder eigene Entwicklungen erweitern.

Durch diese Kombination ermöglichen wir nicht nur eine agile Entwicklung, sondern stellen den Forschenden eine Anwendung zur Verfügung, die sie nach Ihren Ansprüchen anpassen können. Dabei setzt unsere Forschungsumgebung auf ein – in relationalen Datenbankstrukturen abgebildetes – Datenmodell, welches von den Benutzern selbst konfiguriert wird. Damit ist die Zuordnung und Ausgestaltung der Attribute der zu beschreibenden Objekttypen frei wählbar. Lediglich die Auswahl an Datentypen (Text, Zahl, ...) für die Attribuierung ist dabei von der Umgebung vorgegeben.

Die Beschreibung der einzelnen Attribute und Auswahlmöglichkeiten (beispielsweise für Dropdown-Auswahllisten), können von den Projektbenutzern frei vergeben werden. Darüber hinaus besteht die Möglichkeit diese Zuordnungen und Ausgestaltungen im Nachhinein zu ändern, zu erweitern oder gänzlich zu entfernen. Dieser Teil wird über ein parallel entwickeltes und darauf zugeschnittenes – auf *SKOS*-Standard basierendes – Thesaurus-Tool realisiert. Neue Sprachen können jederzeit hinzugefügt oder entfernt werden. Auch lassen sich Einträge ändern, Definitionen hinzufügen sowie Notizen ergänzen.

Neben der Eingabe neuer Daten können ebenfalls vorhandene Daten in Form von verschiedenen Dateiformaten oder Literaturnachweisen zum System hinzugefügt werden, um die erhobenen Daten mit weiteren (Meta-)daten anzureichern. Die direkte Unterstützung verschiedener Dateiformate ist dabei ein wichtiger Schritt, um alle Bereiche des Forschungsdatenmanagements zu bündeln. So kann *Spacialist* mit einfachen Textdaten über Bilder und Archivdaten bis zu 3D-Dateien umgehen und diese zum Teil weiterverarbeiten.

Durch die oben genannten Möglichkeiten und die Verlagerung in den virtuellen Raum ergeben sich vielzählige Möglichkeiten zur ortsunabhängigen Kollaboration von Wissenschaftler*innen. Über Zugriffskontrollen können verschiedene Rollenelemente vergeben werden, um z.B. im Moderator-Modus Datenbearbeitungen erst nach entsprechender Diskussion aufzunehmen. So lassen sich gesicherte Informationen und eine Versionierung von einzelnen Datenbeständen realisieren. Gleichzeitig können Änderungen im Verlauf nachvollzogen und Fehler korrigiert werden. Dieses Kontrollsystem schützt vor

⁶ Daityari, "Angular vs React vs Vue".

unbefugten Zugriffen, ungewollten Änderungen und ermöglicht die Abfrage ältere Bestände. Ein weiterer Vorteil der Nutzung der hier vorgestellten Forschungsdatenumgebung *Spacialist* ist die durch die Realisierung als Webanwendung einfach zu realisierende Sicherung der Datenbestände in den (universitären) Serverstrukturen. Hier sind die Daten selbst gegen einen Serverausfall oder einen Angriff geschützt. Ein Verlust der Daten durch einen Ausfall der EDV des Bearbeiters an den jeweiligen Standorten ist somit ausgeschlossen und die Erhaltung der zusammengetragenen Forschungsdaten gewährleistet.

Außerdem erlaubt sie eine einfachere Archivierung der Daten in bestehende Systeme, wie beispielsweise Universitäts-Repositoryn. Diese zeitintensive Aufgabe wird einfacher, wenn schon während des Forschungsprozesses notwendige Dokumentationen vom System erfolgen.

Spacialist ist seit Beginn als Open-Source-Projekt umgesetzt⁷, um die gesamte interessierte Community einzubinden und am Entwicklungsprozess teilhaben zu lassen. Dadurch war es möglich in kürzester Zeit eine Vielzahl an Rückmeldungen zu erhalten, die in die weitere Ausrichtung und Stabilität der Forschungsumgebung eingeflossen sind. So haben wir es in den letzten Jahren geschafft, eine zuverlässige, intuitive und umfängliche Umgebung zu schaffen, die die Forschenden aus den verschiedensten Disziplinen bei ihrer Arbeit unterstützt.

⁷ GitHub, „Projektseite Spacialist“.

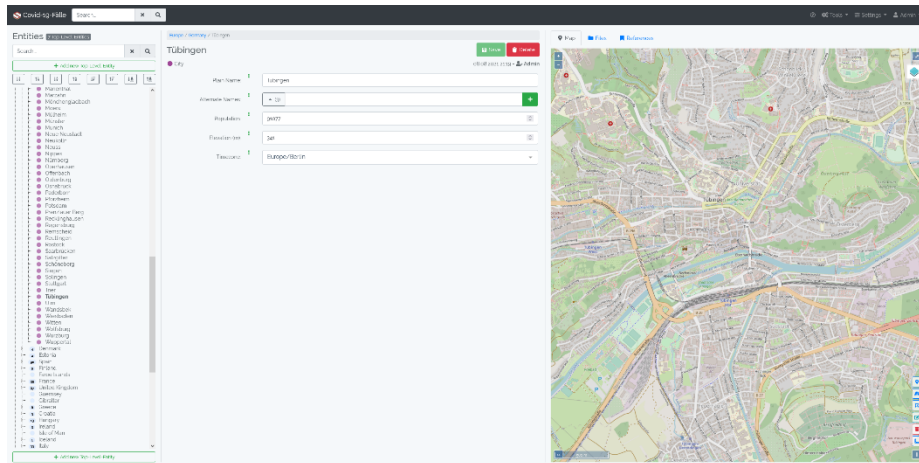


Abb. 1. Spatialist – Eintrag zu Tübingen.

Bibliografie

“Bootstrap · The most popular HTML, CSS, and JS library in the world.”, Bootstrap Homepage, aufgerufen am 08.08.2021, <https://getbootstrap.com>.

Daityari, Shaumik. “Angular vs React vs Vue: Which Framework to Choose in 2021”, veröffentlicht am 15.03.2021 auf [codeinwp.com](https://www.codeinwp.com), <https://www.codeinwp.com/blog/angular-vs-vue-vs-react/#popularity>.

“Laravel - The PHP Framework For Web Artisans”, Laravel Homepage, aufgerufen am 08.08.2021, <https://laravel.com>.

PostgreSQL: The world's most advanced open source database”, PostgreSQL Homepage, aufgerufen am 08.08.2021, <https://www.postgresql.org>.

“Spatialist – Die virtuelle Forschungsumgebung für die Spatial Humanities”, eScience-Center Uni Tübingen, aufgerufen am 08.08.2021, <https://uni-tuebingen.de/forschung/forschungsinfrastruktur/escience-center/spacialist-1>.

"Spacialist", GitHub Projektseite, Stand 08.08.2021,
<https://github.com/eScienceCenter/Spacialist>.

"Vue.js", Vue.js Homepage, aufgerufen am 08.08.2021,
<https://v3.vuejs.org>.