



UNIVERSITÄTS-
BIBLIOTHEK



SPECIFY ALS
INSTITUTIONELLER
FORSCHUNGSDATEN-
MANAGEMENTSERVICE

**REPORT 5 DES FORSCHUNGSPROJEKTS
„COMMUNITY-SPEZIFISCHE FORSCHUNGSDATEN-
PUBLIKATION (CS-FDP)“**

KOMPETENZZENTRUM
FORSCHUNGSDATEN
DER
RUPRECHT-KARLS-UNIVERSITÄT
HEIDELBERG

17.03.2020

Ziel des Projekts „Community-spezifische Forschungsdatenpublikation“ ist die Weiterentwicklung des Serviceangebots des Kompetenzzentrums Forschungsdaten der Universität Heidelberg durch den Aufbau eines Pools von generischen Softwarewerkzeugen zur Erstellung fach- und Community-spezifischer Publikationsportale für Forschungsdaten, die Entwicklung eines Konzepts zur nachhaltigen Integration der Heidelberger Forschungsdaten in übergreifende Archivierungskonzepte sowie die Professionalisierung des Datenmanagements an der Universität durch die Verankerung des Themas in Forschung und Lehre.

Kern des Projekts ist die Erprobung der fachspezifischen Erweiterung und projektspezifischen Individualisierung universitärer Repositorienangebote (typischerweise Volltextrepositorien bzw. Dokumentenserver, teilweise mittlerweile auch generische Repositorien für Forschungsdaten) um ein Dienstportfolio, mit dem noch stärker individuelle, fachspezifische Anforderungen erfüllt werden können.

In diesem sowie den weiteren Projektreports werden entsprechende Softwarelösungen vorgestellt, die im Rahmen des Projekts implementiert und evaluiert wurden.

SPECIFY: ANWENDUNGSBEREICHE, FUNKTIONALITÄTEN, TECHNISCHE ANFORDERUNGEN

Die Open-Source-Software Specify wird entwickelt, um biologische, geologische oder andere Sammlungsbestände zu klassifizieren und zu beschreiben. Sie wird daher schwerpunktmäßig im Museumsbereich, insbesondere in naturhistorischen Museen verwendet. Prinzipiell kann sie jedoch für alle Datenbestände genutzt werden, die in hierarchischen Beziehungen zueinanderstehen. Specify deckt damit eine ganze Klasse von Forschungsdaten aus verschiedenen Bereichen ab und bietet daher eine gute Erweiterung eines generischen Forschungsdatenrepositoriums im Sinne der CS-FDP. Die Visualisierung macht die kladistische Datenstruktur deutlich, die in den Lebens- und angrenzenden Wissenschaften Genealogien bzw. klassische Stammbaumdarstellungen abgelöst hat.

Die Software verwaltet Metadaten wie etwa die örtliche und zeitliche Beschreibung von Fundstellen, von beteiligten Wissenschaftlern und der weiteren Bestimmungsgeschichte des jeweiligen Objekts und kann je nach Kontext angepasst werden. Die im Rahmen des Pilotprojekts verwendete Version Specify 6 unterstützt des weiteren Georeferenzierung sowie Import und Export von Daten (<https://www.sustain.specifysoftware.org/>). Nutzer können alle Sammlungen in einer einzigen Datenbank verwalten, um die Verwaltung zu vereinfachen. Specify existiert in je zwei Versionen für die Desktop-Betriebssysteme OS X, MS Windows und Linux, nämlich in einer Vollversion und einer Shareware, wobei die Vollversion von Specify 6 die Installation eines MySQL-Datenbankmanagers und der Java Runtime Environment (JRE) erfordert. Die Nutzung des Specify Webportals benötigt einen Linux Server mit Apache Tomcat, wobei für das Verwenden der Dokumentation die Installation einer Ubuntu-Distribution empfohlen wird. Der Einsatz des Webportals bietet sich insbesondere dann an, wenn Specify nicht nur als Arbeitsplattform, sondern auch für einen Export bzw. eine Veröffentlichung der Daten (unterstützt Darwin Core Standard) dienen soll. Pilotpartner an der Universität Heidelberg ist der Forschungsbereich Biostratigraphie und Paläoökologie des Instituts für Geowissenschaften. Eine zentral gehostete Instanz von Specify wird über den Cloudservice der Universität Heidelberg (<https://heicloud.uni-heidelberg.de/heicLOUD>) für zwei Arbeitsgruppen des Instituts für Geowissenschaften betrieben. Das Kompetenzzentrum für Forschungsdaten (KFD) migrierte hierbei eine zunächst lokale Instanz von Specify, die

ursprünglich am Institut für Geowissenschaften angesiedelt war, in den universitären Cloudservice (heiCLOUD); dies bedeutet zunächst eine Steigerung der Effizienz hinsichtlich der Ressourcennutzung und ermöglicht ein langfristig sicheres Vorhalten der Daten. Darüber hinaus soll die Migration längerfristig zu einer Vereinfachung von Datenveröffentlichungen führen (Apel et al., 2018).

BEWERTUNG HINSICHTLICH LANGFRISTIGER NUTZUNG

Die an dem Pilotprojekt beteiligten Wissenschaftler äußern sich sehr positiv über das Nutzungsprofil und Benutzerfreundlichkeit der Software. Insbesondere die breite Palette an Möglichkeiten, einzelne Sammlungsstücke mit Metadaten zu beschreiben, sowie die Funktion der Georeferenzierbarkeit erleichtert das Beschreiben von geowissenschaftlichen und paläontologischen Daten erheblich und stellt einen großen Fortschritt in der Beschreibbarkeit von wissenschaftlichen Sammlungen dar. Die Einfachheit der verwendeten Software erlaubt eine weitestgehend autonome Nutzung durch den Wissenschaftler, d.h. der zentrale Support durch das KFD wird nur in Ausnahmefällen in Anspruch genommen. Für die Verwendung von Specify spricht ferner die vergleichsweise hohe internationale Verbreitung, was eine auch zukünftig rege Entwicklung und Verbesserung hoffen lässt (Coetzer, 2012). Aktuell nutzen mehr als 450 naturhistorische Sammlungen in mehr als 38 Ländern die Software. Für eine längerfristige Nutzung der Specify-Software an der Universität Heidelberg durch das Kompetenzzentrum für Forschungsdaten ist wünschenswert, noch weitere Arbeitsgruppen an der Universität Heidelberg sowie an anderen Universitäten für die Nutzung von Specify zu gewinnen und so weitere Synergien für ein zentrales Hosting zu generieren. Eine Befragung der am Projekt beteiligten Wissenschaftler ergab, dass das Potential für die gemeinschaftliche Nutzung der Software in ihrer Forschungscommunity vorhanden ist.

Der Betreuungsaufwand, sowohl im Hinblick auf Systemadministration als auch auf erforderliche Nutzerbetreuung, ist verhältnismäßig gering. Spezifische Anpassungen an lokale Gegebenheiten waren nicht erforderlich.

WEITERFÜHRENDE LITERATUR

Apel, J., Gebhart, F., Maylein, L., Wlotzka, M., 2018. Offene Forschungsdaten an der Universität Heidelberg: von generischen institutionellen Repositorien zu fach- und projektspezifischen Diensten. O-Bib Offene Bibl. Herausgeber VDB 5, 61–71. <https://doi.org/10.5282/o-bib/2018H2S61-71>

Coetzer, W., 2012. A new era for specimen databases and biodiversity information management in South Africa. Biodivers. Inform. 8. <https://doi.org/10.17161/bi.v8i1.4263>

Specify Software Project | Sustain Specify [WWW Document], n.d. URL <https://www.sustain.specifysoftware.org/> (besucht am 17.03.20).

WEITERFÜHRENDE WEBLINKS

<https://www.sustain.specifysoftware.org/>