

Neues zur ExpoDB

Vortrag auf dem 21. MusIS-Nutzertreffen online am 25.03.2021.

Von Christof Mainberger

Das „Projekt“ ExpoDB wurde formal im November 2020 abgeschlossen. Eindrückliche Ergebnisse auf Basis der ExpoDB sind u.a. die Online Sammlungen der Städtischen Museen Freiburg und des Linden-Museums Stuttgart sowie der zukünftige Webauftritt zum Thronsaal des Badischen Landesmuseums, der dann parallel mit der Ausstellungseröffnung aufgeschaltet wird. Welche Ziele hatte das Projekt im Rückblick?

- Die digitale Sammlungsdokumentation der Museen sollte für vielfältige, innovative Zwecke geöffnet werden: Für Digitale Kataloge, Medienstationen, Kulturportale, weitere digitale Geschäftsgänge, für die Nachnutzung als Linked Open Data bzw. als Forschungsdaten.
- Es sollte sich nicht jedes Museum seine „eigene“ ExpoDB entwickeln; die Entwicklung sollte vielmehr einheitlich und standardisiert erfolgen. Vielfältige und innovative Zwecke stehen in einem Spannungsverhältnis mit einer solchen Einheitlichkeit. Dies löst die ExpoDB durch eine standardisierte Plattform, die flexibel konfiguriert wird. Standardisiert werden auch die Datenstrukturen und das Vokabular der primären Sammlungsdaten in imdas pro, z.B. um diese zu internationalisieren sowie verlässlich maschinell weiterzuverarbeiten.
- Die ExpoDB musste unmittelbar für die Projekte der Museen zur Verfügung stehen. Vorhaben, z.B. im Rahmen der Förderlinie Digitale ins Museum II, konnten nicht das Ende einer Entwicklung abwarten, sondern benötigten, nicht nur prototypisch, sofort Daten und Schnittstellen. Umgekehrt ist die Entwicklung der ExpoDB mit Ende des Projekts nicht abgeschlossen: Kreative Ideen in den Museen erzeugen einen permanenten Strom weiterer Anforderungen, die sich stetig in einem Ausbau der Funktionalitäten niederschlagen.

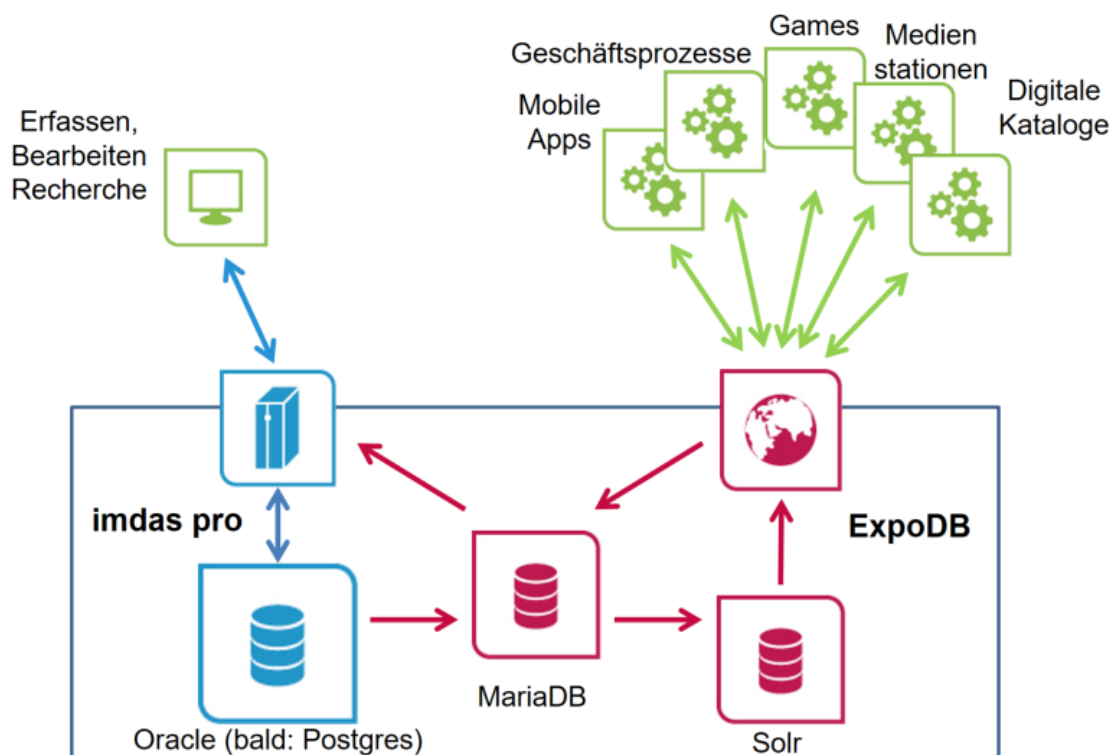


Abb1: Die ExpoDB (rot) verbindet als Infrastruktur aus Datenbank, Suchmaschine und Schnittstelle die Sammlungsdokumentation mit imdas pro mit neuen Anwendungen in den Museen

Die ExpoDB umfasst mehr als ihre technische Infrastruktur aus Datenbank, Suchmaschine und Webservices, die die hermetische Datenhaltung von imdas pro mit den diversen Applikationen der Museen verbindet. Ein zweiter Aspekt betrifft die Ablösung obsoleter Datenstrukturen in imdas pro, die Standardisierung zusätzlicher Datenstrukturen, z.B. zur Aufnahme publikumsspezifischer Texte, sowie die Etablierung eines Systems von Markierungen, um die Publikation der Daten differenziert zu steuern. Um für die Museen die massenhafte Anreicherung ihrer Daten durch Normdaten, z.B. Geokoordinaten, vorzunehmen, wurden neue, automatisierte Verfahren ausgearbeitet und zusätzliche Online-Schnittstellen für imdas pro in Auftrag gegeben.

Aus der Zusammenarbeit mit den Museen, sowie der von jenen beauftragten Agenturen, ergibt sich drittens als entscheidende Aufgabe, wie das Verständnis dafür geschaffen werden kann, wie die Daten aus imdas pro innerhalb der ExpoDB verarbeitet und schließlich in deren Schnittstellen repräsentiert werden. Denn schließlich muss das Museum inhaltlich selbst steuern können, wie seine Sammlungsdaten in den gestalterischen Konzepten der Agenturen in Erscheinung treten.

Da sowohl die URL-Anfrage-Syntax als auch die maschinenlesbare XML- bzw. JSON-basierte Ergebnisdarstellung der Expo-Schnittstelle für Nicht-Techniker*innen unübersichtlich und schwer verständlich sind, wurde die „Viewer“-Applikation, eine browserbasierte Rechercheoberfläche bereitgestellt. Das Webformular enthält eine einfache und eine erweiterte Suche ergänzt um Facetten zum Explorieren der Indexlisten sowie zum Verfeinern von Suchergebnissen. Auf der

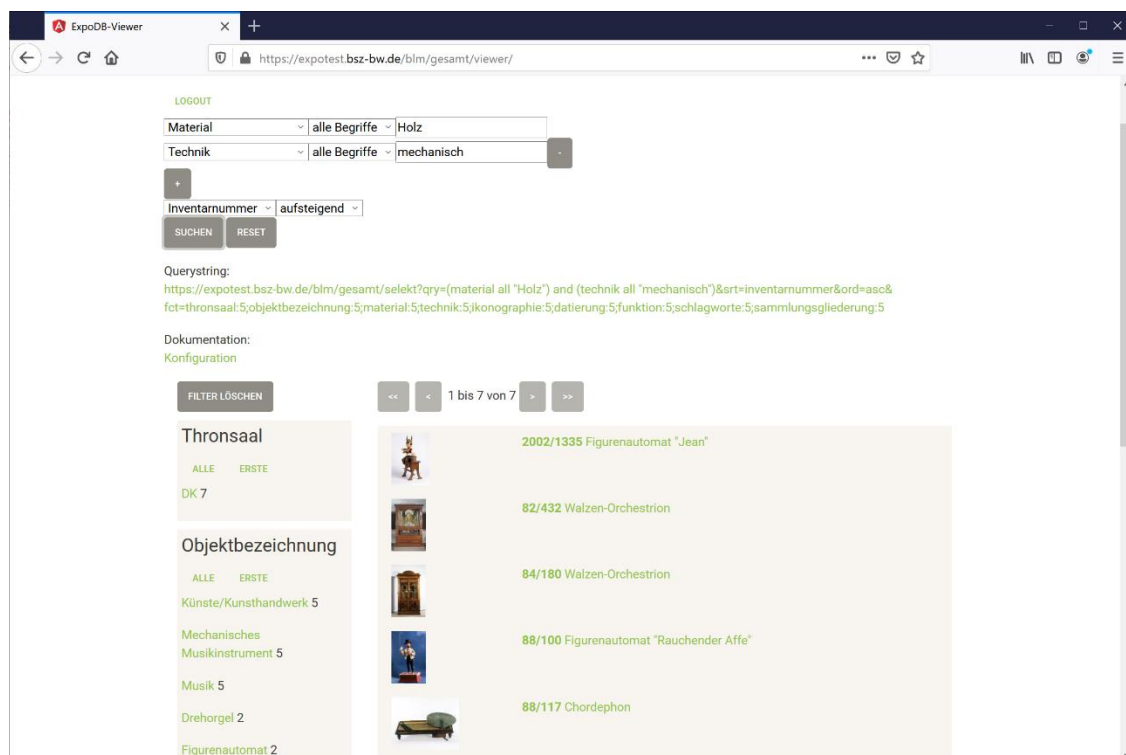


Abb 2: Die Viewer-Applikation enthält als Referenzimplementierung für die ExpoDB-Schnittstelle eine erweiterte und eine einfache Suche und Facetten zur Verfeinerung von Suchergebnissen.

Oberfläche werden die URLs dokumentiert, die der Viewer im Hintergrund verwendet und die seitens der Agenturen implementiert werden müssen, um die gleichen Suchergebnisse zu erzielen. Als Novum kann diese Anwendung mit einem Passwort geschützt werden, so dass neben bereits publizierten Daten nun auch Daten eingestellt werden könnten, die für eine Veröffentlichung erst noch zu überarbeiten sind. Alternativ zu imdas pro wird damit auch eine zugangsgeschützte, leicht handhabbare Recherche in allen Sammlungsdaten eines Museums bereitgestellt.

In der Detailansicht ermöglicht die „Viewer“-Anwendung die Anzeige aller publizierten Bilder zum Objekt, ggf. Links auf Audio- und Videodateien, alle beschreibenden Sammlungsdaten – soweit diese über die Schnittstelle abrufbar sind -, sowie wiederum die URLs zu den Objektdaten in einem JSON- und in einem XML-Format. Das JSON-Format entspricht dem, das die Agenturen zu einem Treffer erhalten und weiterverarbeiten müssen; das XML ist der Export der Sammlungsdaten aus imdas pro. Bei Defiziten in der Trefferpräsentation kann so sehr einfach ermittelt werden, ob der vermisste Inhalt gar nicht aus imdas pro exportiert wurde, oder ob er erst bei der Verarbeitung innerhalb der ExpoDB auf der Strecke blieb. Schon eine solche Möglichkeit, Fehler leicht identifizieren und lokalisieren zu können, beschleunigt deren Behebung enorm und damit den gesamten Prozess der Schnittstellenkonfiguration.

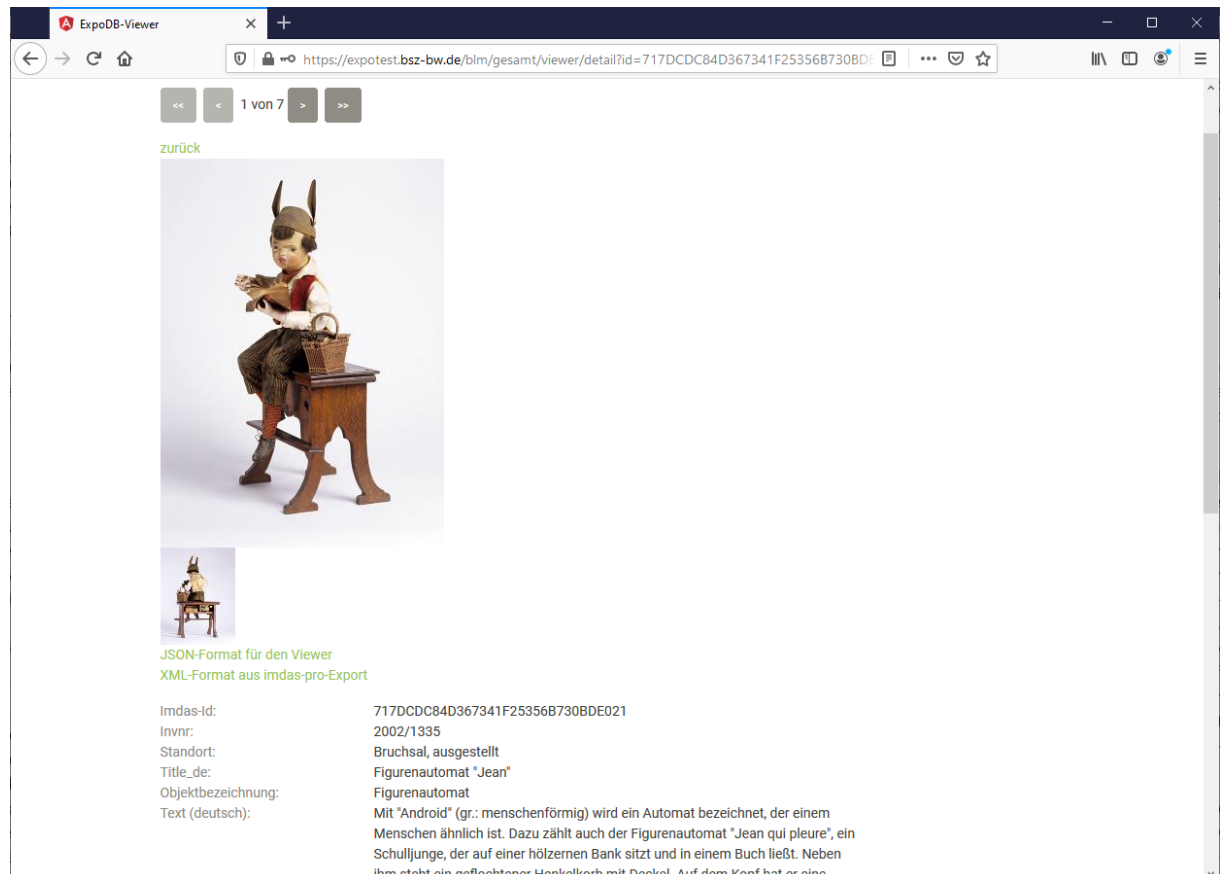


Abb 3: Die Detailansicht des Viewers zeigt Bilder und Metadaten. Über Links sind die JSON- sowie XML-Daten der Schnittstelle zu erreichen.

Da der Viewer ausschließlich mit denselben Schnittstellen der ExpoDB arbeitet, die auch die Applikationen der Museen einsetzen und – nach und nach – die diversen Features dieser Applikationen umgekehrt in den Viewer integriert werden, dient er als Referenzimplementierung für die ExpoDB-Schnittstellen: Alles was im Viewer funktioniert, können auch die Museen in ihren Vorhaben einsetzen; was im Viewer (noch) nicht funktioniert, muss spezifisch vereinbart werden. Das der Viewer bereits zu Beginn eines Entwicklungsvorhabens zur Verfügung steht, hilft er den Museen bei der parallel laufenden Aufbereitung ihrer Daten zur Publikation: im Viewer können sie unmittelbar überprüfen, wie die Daten später „raus kommen“.

Mit dem Viewer kann also leicht beobachtet werden, welche Daten eine Schnittstellenkonfiguration ergibt; der Viewer erklärt allerdings nicht, wie diese Daten genau zustande kommen. Technisch geschieht dies über eine Abfolge von Datenverarbeitungen, die jeweils durch spezielle Kon-

figurationen, inklusive der von imdas pro, gesteuert werden. Deren Verständnis erfordert Kenntnisse in einer Reihe von Technologien und auch dann ist es nicht leicht, die verstreuten Informationen im Überblick zu behalten. Aus diesem Grund wird eine Referenz der Konfigurationen erstellt, die alle Einstellungen zusammenführt und für menschliche Leser*innen formuliert.

Die Referenz ist in die unterschiedlichen Objekttypen, Museumsobjekte, Personen, Medien, Thesauruseinträge etc. gegliedert, die über die ExpoDB-Schnittstelle angeboten werden. Zu jedem Objekttyp beschreibt die Referenz, welche Datenelemente in den Detaildatensätzen enthalten sind, auf welchen Eingabefeldern in welchen imdas-pro-Masken sich diese beziehen und wie die Inhalte ggf. innerhalb der ExpoDB bearbeitet wurden. Die Referenz enthält die Positiv- bzw. Negativlisten zu Rollen- oder Typangaben, anhand derer der Export einzelner Datenelemente gesteuert wird. Schließlich werden zu jedem Objekttyp in der Referenz die Indexe erläutert, über die diese Angaben gesucht, sortiert bzw. gefiltert werden können.

Die Referenz wird in der ExpoDB nicht als eigenständiges Dokument separat gepflegt, sondern die Referenz wird aus Kommentaren in den Konfigurationsfiles zu den Datenverarbeitungen in der ExpoDB jeweils on-the-fly kompiliert. Insofern jede Verarbeitung somit vor Ort dokumentiert wird, wird vermieden, dass die Konfigurationen und ihre Dokumentation auseinanderlaufen. Für die Abstimmungsprozesse mit den Museen und Agenturen dient die Referenz als Unterlage, anhand der Datenumfänge und Recherchefunktionalitäten der Schnittstelle systematisch und auf dem aktuellen Stand durchgesprochen werden können.

Bezeichnung	Erläuterung
Titel (deutsch)	Im Suchindex titel_de wird ggf. der Feldinhalt 'titel_de', sonst ggf. 'eigenname', sonst ggf. 'indigenebez' und andernfalls '.o.T' indiziert.
Standort	Je nachdem, wie der aktuelle Standort beginnt, wird 'Staufen', 'Bruchsal' oder 'Karlsruhe' indiziert.

Abb 4: Die Referenz dokumentiert Objekttypen, Feldkataloge und Indexbeschreibungen und wird aus den verteilten Konfigurationsdateien der ExpoDB generiert.

Da die Entwicklung der ExpoDB unmittelbar Anforderungen aus den Projekten in den Museen bedienen muss, werden für Funktionalitäten jeweils zügig Prototypen zur Verfügung gestellt. Diese werden dann im Rahmen der Zusammenarbeit mit Museen und Agenturen optimiert und

schließlich für die Nachnutzung in anderen Projekten und Museen konsolidiert und generalisiert. Auch diese agile Vorgehensweise adressiert zentral den Aspekt der Kommunikation in der Zusammenarbeit. Sie kann auch angewandt werden, um proaktiv neue Funktionsbereiche zur Diskussion zu stellen. Für jede Idee kann früh und einfach ausprobiert werden, ob sie trägt, bevor große Aufwände evtl. doch in eine falsche Richtung investiert werden. Zwei Themengebiete, die beim 21. MusIS-Nutzertreffen angesprochen wurden, sollen in diesem Sinne untersucht werden: Die Bereitstellung von Linked-Open-Data über die ExpoDB sowie die Verwaltung von sammlungsbezogenen Content in der ExpoDB.

Mit Initiativen wie Coding DaVinci und dem Forschungsdatenmanagement in den Digital Humanities, hat auch in den Museen die Datenveröffentlichung als Linked-Open-Data bzw. nach dem FAIR-Prinzip an Fahrt aufgenommen. Neben technischen Aspekten werden dabei lizenzrechtliche Gesichtspunkte als auch die Dauerhaftigkeit und Verlässlichkeit stark betont. MusIS ist hierbei als Partner sehr gut aufgestellt; wir können sowohl die technische Infrastruktur wie auch die institutionelle Nachhaltigkeit gewährleisten: Z.B. kann die Imdas-Id technisch und organisatorisch als persistenter Identifier garantiert werden. Entscheidend ist hier der Beitrag der Museen: sie müssen sich nicht nur – auch für die Metadaten – zu einer geeigneten Lizenz verständigen. Sie müssen auch den semantischen Gehalt ihrer Datenveröffentlichung genau definieren. Nur dies macht eine sinnvolle Nachnutzung der Daten wirklich möglich; hier etwa einfach nur auf den LIDO-Standard zu referenzieren, reicht nicht aus.

Im Zusammenspiel von ExpoDB und Digitalen Katalogen wird die Frage aufgeworfen, wo welche Inhalte verwaltet werden sollen. Oft wird ein Web-Content-Management-System (CMS) eingesetzt, um z.B. Themenseiten oder Virtuelle Ausstellungen zu realisieren. Die integrierten Redaktionswerkzeuge versprechen eine flexible Gestaltung im Rahmen des beauftragten Webdesigns. Allerdings entsteht so auch eine separate Datenhaltung, die bei einem Wechsel des CMS oder der Agentur migriert werden muss. Oft werden auch Sammlungsdaten oder Bilder in den Content-Management-Systemen „gecached“, was bei Datenaktualisierungen zu Irritationen führt. Insgesamt besteht die Gefahr, dass nicht nur pro Museum insgeheim im CMS eine „eigene“ ExpoDB entsteht, sondern sogar pro einzelner Anwendung. Die ExpoDB versucht daher jedenfalls für sammlungsbezogene Daten möglichst zentrale Datenstrukturen und Schnittstellen anzubieten. Die virtuelle Ausstellung zum Thronsaal des Badischen Landesmuseum ist ein Beispiel, wie mit Bordmitteln von imdas pro Content-Management unabhängig von der aktuellen Frontend-Lösung realisiert werden kann.

Wie ist nun das konkrete Vorgehen bei einem neuen Vorhaben? Für ein neues Vorhaben sollten möglichst früh die bereits vorhandenen Ideen mit dem MusIS-Ansprechpartner besprochen werden. Dieser Ideen-Katalog muss noch nicht vollständig sein. Als langwierigste Arbeit stellt sich regelmäßig die Vorbereitung der Daten in imdas pro heraus, so dass erforderliche Datenstrukturen oder Normdateneinspielungen unmittelbar in Angriff genommen werden sollten.

Ein weiterer Schritt ist die Vorbereitung einer kleinen Menge von Testdatensätze; sehr hilfreich ist die Einrichtung eines künstlichen „Dummy-Datensatzes“, der alle denkbaren Datenelemente und zugehörige Medien umfasst und natürlich vor dem Launch wieder entfernt wird. In der ExpoDB und im Viewer, die zu Projektbeginn eingerichtet werden, können so die Fortschritte in der Datenaufbereitung und der Optimierung der Retrievalfunktionalitäten leicht verfolgt werden.

Während der Implementierung werden vierzehntägig Videokonferenzen verabredet. Es ist wichtig, dass an diesen Terminen aus den Museen die imdas-Koordinator*innen, Mediengestalter sowie Digitalmanager beteiligt werden und auch die beauftragte Agentur hinzugebeten wird.

Auch MusIS wird bei Bedarf Spezialisten hinzuziehen. Nur so können die gestalterischen Erwartungen mit der Datenlage und den technischen Möglichkeiten effizient in Einklang gebracht werden. Das BSZ richtet in seinem Wiki-System eine Dateiablage ein, über die Dokumente von den Beteiligten gemeinsam bearbeitet werden können.

Gewöhnlich dauert die Realisierung eines solchen Vorhabens mehrere Monate. Bei jedem Projekt war kurz vor dem Ausrollen noch eine Kraftanstrengung erforderlich, bis letzte Datenelemente richtig aufbereitet waren und jede Schnittstelle richtig konfiguriert war. Noch jedes Mal gelang es allerdings, schließlich rechtzeitig fertigzuwerden.