

Open Life Science Publication Database (OLSPub)

Drei aufeinander aufbauende Entwicklungsphasen

Phase 1: Basis-Dienste

Zeitraum: Monate 1-6

Budget: 600.000 Euro

Hauptergebnisse: Rechtsgrundlagen, Submission Site, Datenintegration, Automatische Qualitätskontrolle, 10 Verlage integriert



Phase 2: Infrastruktur

Zeitraum: Monate 7-12

Budget: 800.000 Euro

Hauptergebnisse: Routinebetrieb Submission, MeSH-Tagging und Metadaten, externer Zugriff auf die indizierten Daten



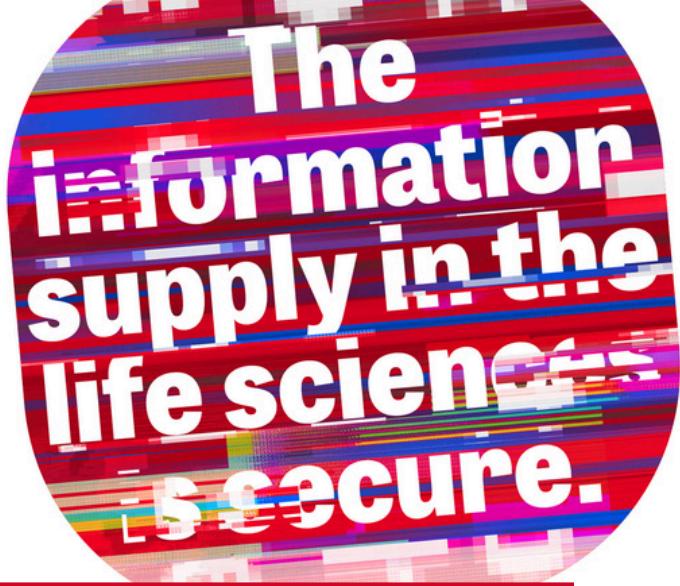
Phase 3: Finale App

Zeitraum: Monate 13-18

Budget: 600.000 Euro

Hauptergebnisse: finale App, API für Nutzende, Skalierung und Load Balancing, Echtzeitbetrieb





ZB MED

Phase 1: „Basis-Dienste“

(Monate 1–6) – 600.000 Euro

Ziel: Aufbau der technischen und rechtlichen Grundlagen für das Repository

Detaillierte Arbeitsschritte

1.1 Vertragliche Grundlagen (120.000 Euro)

Vertragsverhandlungen mit Verlagen (Open Access, kleinere Verlage), Lizenzvereinbarungen für automatische Verfahren, um Metadaten zu generieren und integrieren, Rechtsgutachten zu Urheberrecht und Datenschutz (EU-DSGVO-Konformität), Governance-Struktur etablieren (Beirat, Steering Committee)

Personal: Lizenzmanager, Jurist

1.2 Entwicklung Datenpipeline (240.000 Euro)

Technische Infrastruktur (Server, Cloud-Speicher, Datenbanken, ca. 40 kB), Entwicklung automatisierter Harvesting-Prozesse für Open-Access-Repositorien, Einrichtung ETL-Pipeline (Extract, Transform, Load) für strukturierte Metadaten, Aufbau eines Indexes, Integration von PubMed Central (PMC) Open-Access-Subset, Europe PMC-Daten, LIVIVO-Bestände, Aufbau des Repositorys (ca. 5–10 Millionen Publikationen), Qualitätssicherung und Datenväldierung, Datenbereinigung und Deduplizierung

Personal: DevOps Engineer, Data Engineer, Data Scientist, Senior Developer (50%)

1.3 Datenzugang (120.000 Euro)

Erfassung von Requirements und Spezifikation des User Interfaces (s. 1.4, Community-Arbeit), ebenso wie für weitere Access Points (Download site), Anpassung Pubservatory an das Repository, Repository verfügbar stellen

Personal: 2 Software Developer

1.4 Projekt- und Community-Koordination (120.000 Euro)

Stakeholder-Kommunikation, Aufbau Advisory Board, Aufbau eines Editoren-Boards (Auswahl von Journals), Abstimmung von Aktivitäten, reguläres Projektmanagement, Organisation von Workshops, Meetings zu Datenstandards, zur Erfassung von Requirements, Erstellung von Dokumentation, Vorbereiten der Open Access und Open Source Initiativen

Personal: Projektmanager, Community-Manager

Meilenstein Phase 1 („Basis-Dienste“):

Early-Alpha Release (Pubservatory, Download Site) mit 5–10 Millionen Publikationen, 10 Verlagspartnerschaften, Community-Workshops.

The information supply in the life sciences is secure.



ZB MED

Phase 2: „Infrastruktur“ (Monate 7–12) – 800.000 Euro

Ziel: Zugriffe auf das Repository (Alpha-Release), substantielle Erweiterung des Content.

Detaillierte Arbeitsschritte

2.1 Verlagsvereinbarungen ausweiten (150.000 Euro)

Verhandlungen mit mittelgroßen bis großen Verlagen (Springer Nature, Wiley, Elsevier), Content-Deals für Metadaten, Abstracts, evtl. Volltext-Lizenzen, Integration von ca. 15–20 Millionen zusätzlichen Publikationen

Personal: Lizenzmanager, Jurist

2.2 MeSH-Tagging implementieren (120.000 Euro)

Entwicklung automatisches MeSH-Tagging mit Machine Learning, Training auf existierenden PubMed-MeSH-Zuordnungen, 85%+ Genauigkeit bei automatischer Verschlagwortung, Qualitätskontrolle über die Community (s. 2.4), Integration in die automatische Indexierung und Qualitätskontrolle

Personal: Zwei Data Scientists

2.3 Suchoberfläche entwickeln (300.000 Euro)

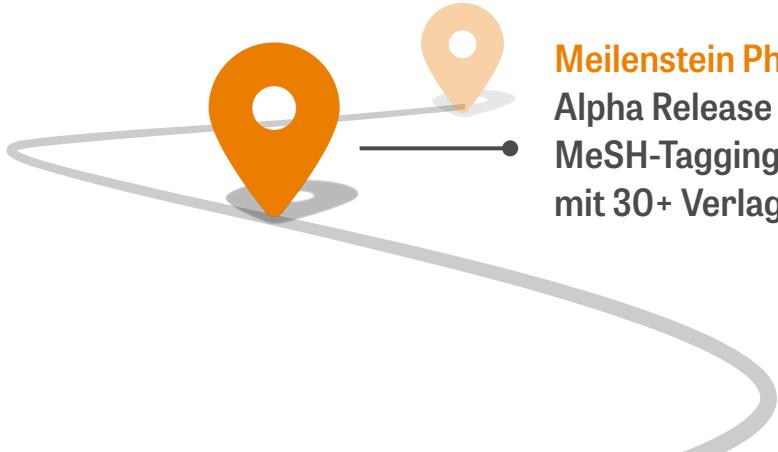
Bereitstellen eines APIs für den Suchindex, externer Zugriff auf das Repository, Entwicklung des Prototypen für eine benutzerfreundliche Weboberfläche, Aufbau des Development-, Test- und Produktivsystems, Virtualisierung, Allgemeine Suchmaske und Expertensuche auf spezifische Felder des Indexes, beides mit Boolean Operators, Proximity, und weiteren Index-Operatoren (s. PubMed/Medline), Interaktives und responsives Design für die Nutzung mit mobilen Endgeräten

Personal: DevOps Engineer, zwei Software Engineers, Web Designer, Senior Developer (50%)

2.4 Projektarbeit, Community-Testing & -Feedback (120.000 Euro)

Projektmanagement, Organisation User-Tests mit Bibliotheken, Ärzten, Forschern, Organisation von Workshops, Meetings zur Datennutzung, Integration in externe Systeme, Qualitätskontrolle (MeSH), Teilnahme an Community-Meetings, Planung und Umsetzung der Feedback-Integration und iterative Verbesserung

Personal: Projektmanager, Community-Manager



Meilenstein Phase 2 („Infrastruktur“):
Alpha Release mit 20–30 Millionen Publikationen,
MeSH-Tagging, Verlagspartnerschaften
mit 30+ Verlagen.

The information supply in the life sciences is secure.



ZB MED

Phase 3: „Finale App“

(Monate 13–18) – 600.000 Euro

Ziel: Finalisierung der App zur Produktionsreife (Beta-Release, inkl. User-API), Finalisierung der europäischen Governance.

Detaillierte Arbeitsschritte

3.1 Verlagsangebote vollumfänglich integrieren (120.000 Euro)

Verhandlungen europäischen und amerikanischen Verlagen und Fachgesellschaften, Content-Deals für Metadaten, Abstracts, evtl. Volltext-Lizenzen, Abdeckung Europe: >80% des PubMed-Content für europäische Verlage, >50% für amerikanische Verlage, automatische Integration von 1 Million Abstracts plus Metadaten jährlich Abstimmung der Inhalte mit dem Editoren-Board (Qualitätssicherung)

Personal: Lizenzmanager, Jurist

3.2 Beta-Release Finale App (330.000 Euro)

Abschluss der Entwicklung: letzte Requirements und volle Funktionalität der App (User-API), Bug fixing, Umsetzung der Anforderungen aus den User-Tests, Durchführen von A-/B-Tests, Integration der Lösung auf Test- und Produktivsystem, Entwicklung RESTful API (ähnlich PubMed E-Utilities), Implementierung Authentifizierung, Rate-Limiting, Last-Test, Performance-Tuning, Load balancing, hardening, Integration in bestehende Dienste (z.B. Zotero, Endnote), Umsetzung der Backup- und Disaster-Recovery-Strategie

Personal: DevOps Engineer, zwei Software Engineers, Web Designer, Senior Developer (50%)

3.3 Nachhaltiges europäisches Netzwerk (60.000 Euro)

Organisation europäischer Partnernetzwerk-Workshop (20+ Institutionen), Memoranda of Understanding mit europäischen Bibliotheken, Governance-Modell für langfristigen Betrieb (EMBL-EBI, de.NBI, evtl. NLM)

Personal: Projektmanager, Community-Manager

3.4 Projektarbeit, Beta-Testing, Launch-Umsetzung (120.000 Euro)

Organisation von Workshops, Meetings zum Beta-Testing, Teilnahme an Community-Meetings, Editoren- und Advisory-Board Meetings, Externe User-Tests und Bug-Fixing, A-/B-Testing, Sicherheits-Audit über eine externe Firma, Marketing-Kampagne für Beta-Launch, Planung und Umsetzung des institutionellen Betriebs, Erarbeitung, Abstimmung, Abschluss der europäischen Governance

Personal: Projektmanager, Community-Manager

Meilenstein Phase 3 („Finale App“):

Beta Release mit 38+ Millionen

Publikationen (ca. 1 Million

Publikationen Zuwachs/per anno),

öffentliche API, europäisches

Governance-Modell etabliert.

