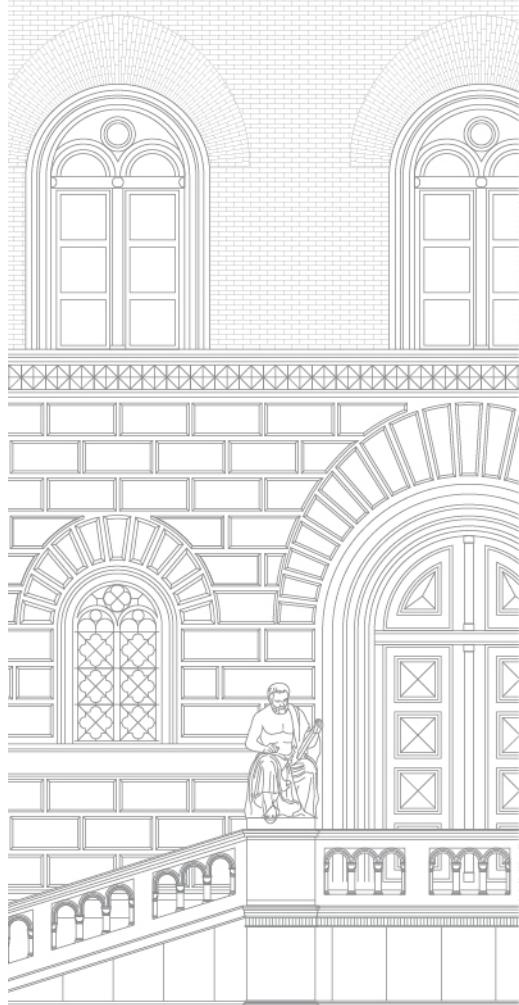


Software-Development at the MDZ

04.04.2019, Marcus Bitzl and Ralf Eichinger, DB/MDZ/IWA



Software developed at the MDZ

- Webapps / Viewer
- Data-Processing / Workflows
- APIs / REST-Servers
- Reusable components / libraries

```
<context: "http://iiif.io/api/presentation/2/context.json",
  @type: "sc:Manifest",
  @id: "https://api.digitale-sammlungen.de/iiif/presentation/v2/bsb00032722/manifest",
  @label: "Ottheinrich-Bibel, Bd. 3: Lk 5,26 - Joh 5,18 - BSB Cgm 8010(3)",
  @metadata: [
    {
      - label: [
        {
          @language: "de",
          @value: "Titel"
        },
        {
          @language: "zh",
          @value: "书名"
        },
        {
          @language: "en",
          @value: "Title"
        }
      ],
      value: "Ottheinrich-Bibel, Bd. 3: Lk 5,26 - Joh 5,18 - BSB Cgm 8010(3)"
    },
    {
      - label: [
        {
          @language: "de",
          @value: "Entstehung"
        },
        {
          @language: "zh",
          @value: "生成"
        },
        {
          @language: "en",
          @value: "Creation"
        }
      ],
      value: "[S.l.] Regensburger Raum, 15./16. Jh."
    },
    {
      - label: [
        {
          @language: "de",
          @value: "Beschreibung"
        },
        {
          @language: "zh",
          @value: "描述"
        },
        {
          @language: "en",
          @value: "Description"
        }
      ],
      value: "[S.l.] Regensburger Raum, 15./16. Jh."
    }
  ]
}
```

The screenshot shows a digital collection interface for the Bavarian State Library (Bayerische Staatsbibliothek). At the top, there's a navigation bar with links for 'Glandlichter', 'Objekte', 'Orte', 'Karten', 'Personen', 'Ausstellungen', and language options 'DE | EN'. A search bar is also present. Below the header, there are several thumbnail images representing different collections: 'Karten' (maps), 'Orte' (places), 'Literatur zu Bayern' (literature about Bavaria), 'Bildahnlichkeitssuche' (image similarity search), 'Glanzlichter' (highlighted items), and 'Zeilungen' (newspapers). On the right side, there's a large logo for 'bavarikon' with the tagline 'Kultur und Wissensschätze Bayerns'. The main content area displays a detailed view of a manuscript page from the 'Ottheinrich-Bibel'. The page features two columns of Latin text in Gothic script and several colorful miniatures depicting scenes from the Bible. Below the text, there are numerous small descriptive labels for each miniature. At the bottom of the page, there are QR codes and numerical identifiers (00006, 00007, 00008, 00009, 00010, 00011, 00012, 00013, 00014, 00015, 00016) corresponding to the individual images. The interface includes standard browser controls like back, forward, and zoom.

Webapps / Viewer

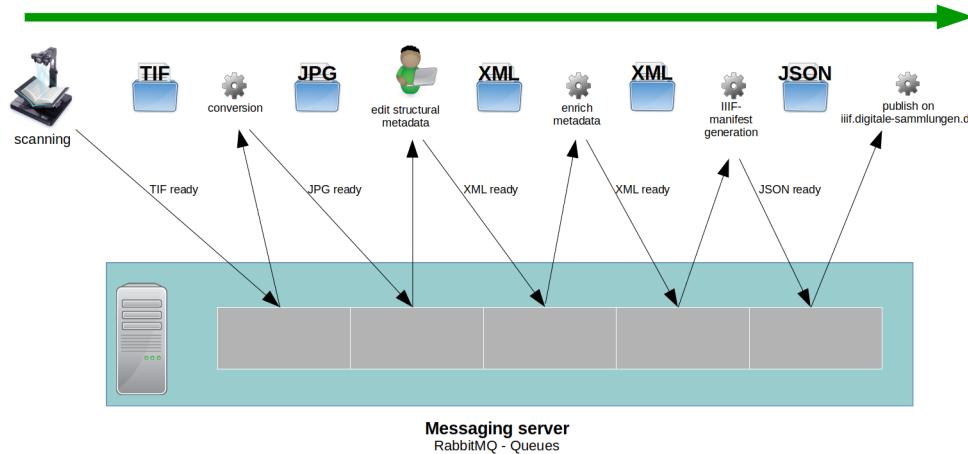
- Technologies:
Java (Spring Boot), HTML/CSS/Bootstrap 4, JavaScript, React
Nginx-LoadBalancer/-Caching,
Debian-Linux-Server
- Viewer: Mirador (<http://projectmirador.org/>)
- cudami Object Management and CMS
(OpenSource at <https://github.com/dbmdz/cudami>)



A detailed view of a page from the Ottheinrich-Bibel. The page contains handwritten text in Gothic script and several illuminated miniatures. One prominent miniature shows Jesus standing among a group of people. The left side of the page has a sidebar with a list of related items and their descriptions, such as 'Vorderer Spiegel' and 'Das Zwinglings zu Nazareth'. The footer of the page includes navigation icons and a footer bar with links to 'Bayerische Staatsbibliothek', 'Münchener Digitalisierungszentrum', 'Über', 'Impressum', 'Datenschutzerklärung', and 'Kontakt'.

Data-Processing / Workflows

- Technologies:
Java, Python
Rabbit MQ (Message Queue)
Debian-Linux-Server
- Flusswerk-Workflow-Engine (OpenSource at GitHub: <https://github.com/dbmdz/flusswerk>)



APIs / REST-Servers

- Technologies:
Java (Spring Boot)
neo4j (Graph database), Solr (Search Server), PostgreSQL (Relational database)
Debian-Linux-Server
- Hymir IIIF-Server (OpenSource at GitHub: <https://github.com/dbmdz/iiif-server-hymir>)
- Euphoria Streaming Server (OpenSource at <https://github.com/dbmdz/streaming-server-euphoria>)
- DZP (Data Access Platform) (Inhouse)

You can access the IIIF API endpoints of the Bavarian State Library at the following endpoints:

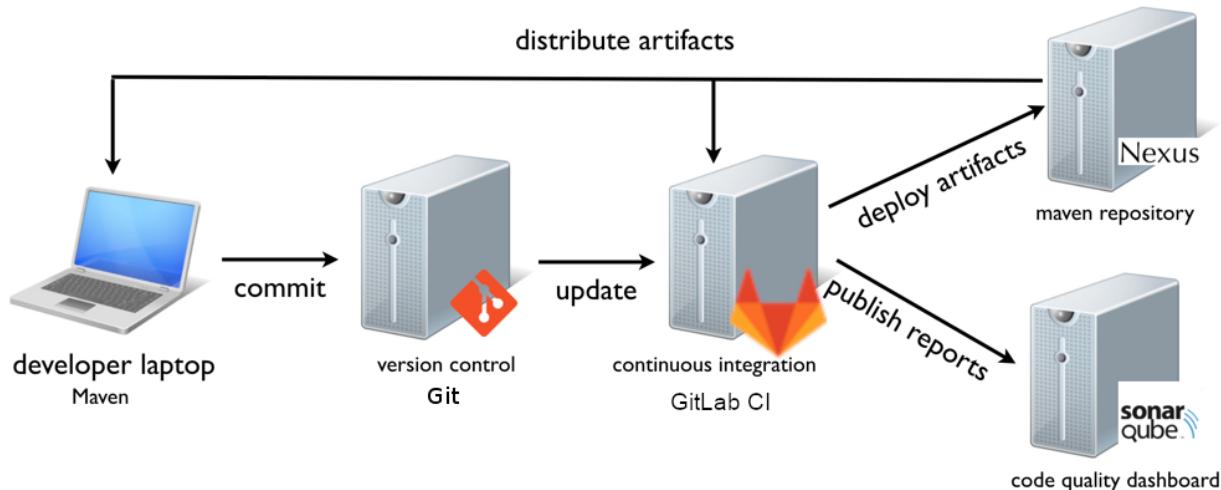
- Presentation API: https://api.digitale-sammlungen.de/iiif/presentation/v2/{object_id}/manifest
- Image API: https://api.digitale-sammlungen.de/iiif/image/v2/{image_id}

Reusable components / libraries

- Technologies:
Java, C, Python
- OpenSource at GitHub: <https://github.com/dbmdz>
- Examples
 - efficient OCR-Storage Solr-Plugin
 - JPEG2000, TurboJPEG image processing
 - File access libraries
 - IIIF-presentation and -image libraries

Software Development Process

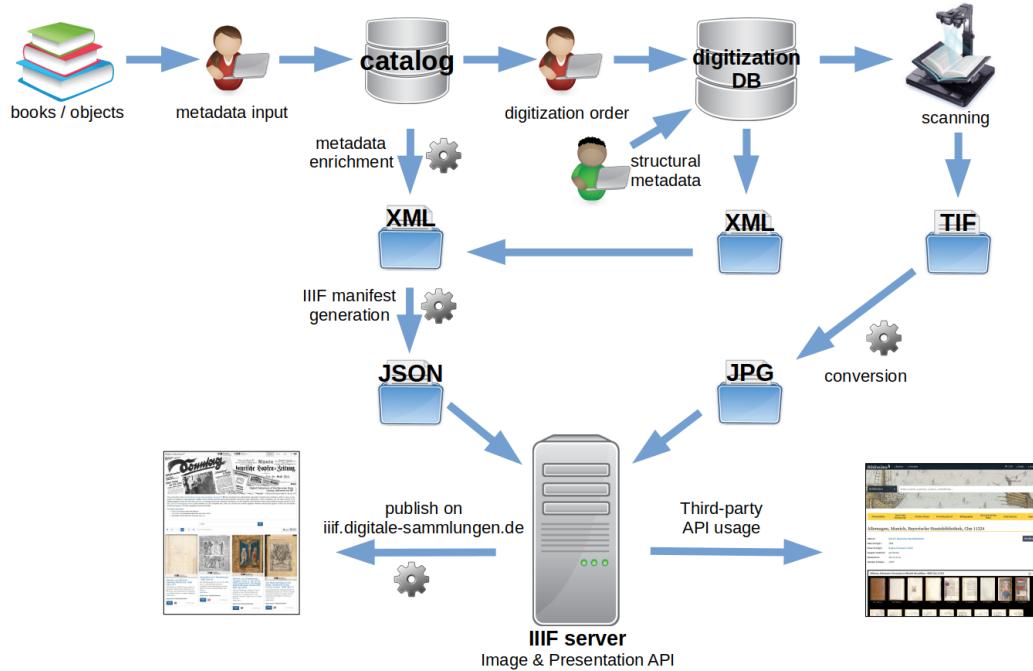
- Continuous Integration:
 - Local Java development with OpenJDK 11, Apache Maven Buildtool and IntelliJ/Netbeans IDEs
 - Code in Git (GitLab/GitHub)
 - CI-Builds: GitLab CI
 - Quality-checks with Maven plugins (e.g. checkstyle, spotbugs) and Quality tools: dependabot, SonarQube
 - Built software-artifacts in Nexus Repository



IIIF Production Process

- Images from scanning to Web

IIIF Production Workflow (simplified)



Lessons Learned

- **Legacy** architecture dictates a lot of design decisions:
Store everything on the NAS, JPEG as the source format
- Performance bottleneck is often not the image transformations, but the transfer from network storage → **Caching** is important at every layer, not just the frontend
- Generating new formats for hundreds of thousands of objects can lead to some nasty surprises: **Corrupted/Invalid/Duplicate legacy data...**
- A common, **standardized API for serving digitized objects** is a game changer
→ “IIIFify all the things”, reduced maintenance costs, better monitoring, easier sharing, collaboration with the community

IIIF Roadmap

- Provide IIIF access to all copyright-free objects (around 2 Mio.)
- Define and add machine-readable IIIF core metadata to make filtering (e.g. in Bookshelf webapp) and discovery easier
- Expose IIIF Content Search API endpoints for all OCRed objects
- Introduce collections to group all manifests semantically
- Offer support for annotating content, both privately and publicly
- Improve and maintain OpenSource products: Hymir IIIF Server, Java IIIF Image and Presentation APIs, IIIF Bookshelf
- Continue with participation in development of Mirador

Links

Web Services

- BSB IIIF Portal: <https://iiif.digitale-sammlungen.de>
- IIIF Presentation API endpoint:
https://api.digitale-sammlungen.de/iiif/presentation/v2/{object_id}/manifest
- IIIF Image API endpoint:
https://api.digitale-sammlungen.de/iiif/image/v2/{image_id}

BSB Open Source Products

- Java-Library IIIF Image API: <https://github.com/dbmdz/iiif-image-api>
- Java-Library IIIF Presentation API: <https://github.com/dbmdz/iiif-presentation-api>
- IIIF Server Demo: <https://github.com/dbmdz/iiif-server-demo>
- IIIF Server “Hymir”: <https://github.com/dbmdz/iiif-server-hymir>
- IIIF Bookshelf Webapp: <https://github.com/dbmdz/iiif-bookshelf-webapp>
- BSB Mirador-Plugins: <https://github.com/dbmdz/mirador-plugins>