



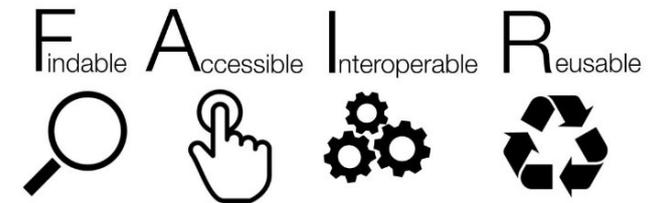
R wie Dokumentation

3. FAIR February Meetup,
5. bis 26. Februar 2025

NFDI Konsortium Text+, Task Area Editions
Sandra König (Leopoldina), Karoline Lemke (BBAW)

Programm FAIR February 2025

- 5. Februar 2025 – **F** wie Qualitätssicherung
- 12. Februar 2025 – **A** wie User Experience
- 19. Februar 2025 – **I** wie BEACON
- 26. Februar 2025 – **R** wie Dokumentation



Graphics: SangyaPundir - Own work, CC BY-SA 4.0,
<https://commons.wikimedia.org/w/index.php?curid=53414062>

Organisationsteam

- Fernanda Alvares Freire (Technische Universität Darmstadt)
- Dr. Philipp Hegel (Akademie der Wissenschaften und der Literatur Mainz)
- Kilian Hensen (Nordrhein-Westfälische Akademie der Wissenschaften und der Künste)
- **Dr. Sandra König (Nationale Akademie der Wissenschaften Leopoldina, Halle (Saale))**
- **Karoline Lemke (Berlin-Brandenburgische Akademie der Wissenschaften)**
- Harald Lordick (Salomon Ludwig Steinheim-Institut für deutsch-jüdische Geschichte)

Ablauf

- Warm-Up
- Hilfreiche Tools
- Praxisübung in zwei Gruppen
- Auswertung
- Abschluss des Workshops
- Wrap-Up FAIR February 2025

Organisatorische Hinweise

- Der Workshop wird nicht aufgezeichnet.
- Der Chat steht jederzeit für Kommentare und Fragen offen.
- In der Diskussion können Sie sich mit dem Handzeichen-Button zu Wort melden.

- Handles: <https://nfdi.social/@Textplus> | [@Textplus_NFDI](https://nfdi.social/@Textplus_NFDI)
- Hashtag: [#FAIRfebruary](https://nfdi.social/#FAIRfebruary)

Warm-up

- „Das gute alte Skript“ - Eine kleine Gruselgeschichte
- Aufgabe:
Was sollte im Rahmen eines digitalen Editionsprojekts dokumentiert werden?
- Link zum Whiteboard: <https://webwhiteboard.com/>

Gefahren einer fehlenden / schlechten Dokumentation

- Sie finden die Daten nicht wieder.
- Sie können die Entstehung der Daten nicht mehr nachvollziehen.
- Sie können die Daten wegen fehlender Kontextinformationen nicht mehr interpretieren.
- Sie könnten die Dateien aufgrund veralteter oder konkurrierender Versionen verwechseln.
- Sie können die Daten nicht mit anderen Personen austauschen oder mit Daten aus anderen Quellen zusammenführen.

Hilfreiche Tools/Schemata - DITA

- Darwin Information Typing Architecture
- Dokumentenformat mit XML-basierter Architektur
- vorwiegend für die technische Dokumentation
- Informationseinheiten werden ihrem Inhalt nach typisiert
- auf Wiederverwendung ausgerichtet
- Output kann auf Zielgruppen angepasst werden
- [DITA im Überblick](#)
- [DITA – Dokumentenformat und Architektur für Wiederverwendung](#)

Hilfreiche Tools/Schemata - DITA

- inhaltlich frei (technisch, Workflow, Editionsrichtlinien)
- sehr gute Planung erforderlich
- integrierte Suche (https://www.oxygenxml.com/xml_editor/webhelp.html)
- Indexierung während der Transformation
- Output in PDF usw. möglich
- ungeeignet für kleinere Projekte
- hohe Einarbeitungszeit
- nachträgliche Anpassung der Ordnerstruktur möglich, aber aufwendig

Hilfreiche Tools/Schemata – TEI-Guidelines

- Kap. 23 „Documentation Elements“ beschreibt die empfohlenen Elemente
- zur Dokumentation von XML-Elementen, inhaltlich frei (technisch, Workflow, Editionsrichtlinien)
- Eingabe in Prosa und mit speziellen Elementen (Deklaration im Schema), die in XML kodiert werden
- XSLT-Stylesheet für die Transformation von XML in z. B. ein PDF
- einfache Handhabung, für komplexere Projekte eher nicht geeignet
- Beispiel: <https://kant-digital.bbaw.de/data/doc/kantTEI.odd.pdf>

TEI text body

```

1 <TEI xmlns="http://www.tei-c.org/ns/1.0" xmlns:tei="http://www.tei-c.org/ns/1.0" xmlns:rng="http://relaxng.org/ns/structure/1.0" xml:lang="de">
2   <teiHeader>
3     <fileDesc>
4       <titleStmt> [4 lines]
9       <editionStmt> [2 lines]
12      <publicationStmt> [6 lines]
19      <sourceDesc> [3 lines]
23    </fileDesc>
24    <revisionDesc status="draft"> [6 lines]
31  </teiHeader>
32  <text>
33    <body>
34      <divGen type="toc" xml:id="toc"/>
35      <div xml:id="about">
36        <head>Überblick Kant-Online TEI P5 Encoding Guidelines</head>
37        <p>Diese Guidelines (TEI P5 ODD Customization) geben einen Überblick über die [2 lines]
40        <p>Die Guidelines sind im Rahmen des DFG-Projektes "Neuedition der Abteilung I, [3 lines]
44        <p>An den Guidelines haben mitgewirkt: Gerald Neumann, Edna Fricke.</p>
45        <p>Die Guidelines sind unter der Lizenz <ref type="newWindow" target="https://creativecommons.org/licenses/by-sa/4.0/">Creative Common:
46          Attribution-ShareAlike 4.0 International</ref> verfügbar.</p>
47      </div>
48      <!-- <div xml:id="overview"> [103 lines]
152     <div xml:id="header">
153       <head>Der TEI Header</head>
154       <p>Jedes TEI-Dokument hat ein <gi>teiHeader</gi>-Element am Anfang des Dokumentes, [4 lines]
159       <egXML xmlns="http://www.tei-c.org/ns/Examples"> [7 lines]
167       <div xml:id="file-description"> [207 lines]
375       <div xml:id="encoding-description"> [10 lines]
386       <div xml:id="profile-description"> [12 lines]
399       <div xml:id="revision-description"> [15 lines]
415     </div>
416     <div xml:id="body"> [454 lines]
871     <div xml:id="editorial-apparatus"> [83 lines]
955     <div xml:id="highlighting"> [103 lines]
1059    <div xml:id="links"> [131 lines]
1191    <div xml:id="content"> [74 lines]
1266    <div xml:id="editor"> [21 lines]
1288  </body>
1289  </back> [639 lines]

```



1. Überblick Kant-Online TEI P5 Encoding Guidelines

Diese Guidelines (TEI P5 ODD Customization) geben einen Überblick über die Auszeichnungsrichtlinien des Projektes Kant-Online in TEI P5.

Die Guidelines sind im Rahmen des DFG-Projekts "Neuedition der Abteilung I, Werke, der Akademie-Ausgabe von Kant's gesammelten Schriften" entstanden. Sie sind in vielfacher Hinsicht angeregt durch die Dokumentation des Projektes Livingstone Online.

An den Guidelines haben mitgewirkt: Gerald Neumann, Edna Fricke.

Die Guidelines sind unter der Lizenz Creative Commons Attribution-ShareAlike 4.0 International verfügbar.

2. Der TEI Header

Jedes TEI-Dokument hat ein `<teiHeader>`-Element am Anfang des Dokumentes, in welchem alle notwendigen und verfügbaren Metadaten enthalten sind, die Struktur wurde für vom DTABf übernommen. Der Header ist unterteilt in die `<fileDesc>`, die `<encodingDesc>`, die `<profileDesc>` und die `<revisionDesc>`:

```
<teiHeader>
  <fileDesc>[...]/</fileDesc>
  <encodingDesc>[...]/</encodingDesc>
  <profileDesc>[...]/</profileDesc>
  <revisionDesc>[...]/</revisionDesc>
</teiHeader>
```

2.1. Die file description

Hilfreiche Tools/Schemata – Anderes

- KI zur vollautomatisierten Kommentierung von Code
- API-Entwicklung: [Swagger](#), ein Toolset, das z. B. automatisierte Dokumentation unterstützt
- Dokumentation von Code: [Pdoc](#) für Python, [Javadoc](#) für Java, [Jsdoc](#) für Javascript
 - teilautomatisiert
 - eigene Kommentare notwendig
 - Output in Form von HTML-Dokumente
- [SMPe](#), [DokuWiki](#), [MediaWiki](#)

Praxisübung

Aufgabe:

Erarbeiten Sie einen Dokumentationsworkflow für das gewählte Setting im Rahmen des Projektabschnittes, der Ihrer Gruppe zugewiesen ist.

- Gruppe 1 „Vorbereitung und Erfassung“
- Gruppe 2 „Durchführung und Anreicherung“

Praxisübung

Setting:

- a) Wählen Sie ein aktuelles Beispielprojekt aus Ihrer Gruppe
- b) Sie arbeiten in einem Projekt für die Erarbeitung einer digitalen historisch-kritischen Edition zu kunsttheoretischen Schriften einer 7-köpfigen Künstlergruppe (um 1900). Die Materialien liegen in diversen deutschsprachigen Archiven in Form von Handschriften (Kurrent) und Maschinenschriften in deutscher, englischer und französischer Sprache vor. Es sind Skizzen und Reproduktionen (Kunstwerke) zu bedenken.

Auswertung Praxisübung

Bericht aus den Kleingruppen

- Wie wurde die Dokumentation geplant?
(Bestandteile, Ressourcen, Verantwortlichkeiten, Zeitplan)
- Welche Fragen kamen auf?
- Link zum Projekt aus Gruppe 2
<https://www.simpk.de/forschung/themen/interpretationsforschung/archiv-des-konzertlebens/topographie-des-berliner-konzertlebens-1880-1945.html>

Bestandteile einer Projektdokumentation

WAS ist zu dokumentieren?

- Datenmanagementplan
- Softwaremanagementplan
- Editionsrichtlinien
- Datenmodell der Edition
- Redaktionsworkflows
- Verwaltung digitaler Ressourcen
- Ordnerstruktur / Dateibenennung, Dokumentenablage
- Übersicht zu / Anwendung von Tools zum kollaborativen Arbeiten
- Verwendung von Credits

Vorgehensweise in der Projektdokumentation

Die Bestandteile der Projektdokumentation können sowohl manuell als auch automatisiert erstellt/aktualisiert und in unterschiedlichsten Formaten abgespeichert werden. Dabei ist wichtig:

- Nutzung von standardisierten und offenen
 - Vokabularen (des jeweiligen Fachbereiches)
 - Sprachen (z. B. allgemein verwendete Programmiersprachen)
 - Protokollen (z. B. ftp und HTTP)
- Maschinenlesbarkeit (z. B. strukturierte Metadaten / Ablage in einem Repository)
- eindeutige Zuordnung Metadaten >< beschriebener Datensatz, z. B. mittels PID

<https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/datendokumentation/>

<https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/persistente-identifikatoren/>

Formate einer Projektdokumentation

Womit kann man dokumentieren?

- in einer begleitenden ReadMe-Datei
- sehr strukturiert in einer Metadatenbank
- in einem projektinternen Wiki
- in der Datei selber bzw. in den Metainformationen der Datei
- direkt in einem DMP

Text+ bietet z.B. einen RDMO-Fragenkatalog über die eResearch Alliance der SUB Göttingen an (<https://plan.goettingen-research-online.de/>)

- in einem elektronischen Laborbuch
- innerhalb der Ordnerstruktur und Dateibenennung

<https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/datendokumentation/>

Erleichterung der Projektdokumentation

- bei Projektbeginn Aufsetzen der benötigten einheitlichen Strukturen und Hilfsmittel
- Nutzung von Tools, die bereits im Projekt verwendet werden
- Dokumentation kontinuierlich fortschreiben, im Arbeitsalltag regelmäßig einplanen
- Versionierung, Benennung von Verantwortlichkeiten für die Dokumentationsteile

Wrap-Up FAIR February 2025

- Themen: Qualität, Usability Experience, BEACON, Dokumentation
- Fazit: User Feedback integrieren, Austauschgruppe bilden, Folgeworkshop konzipieren, fleißig dokumentieren

Nützliche Hinweise

- FAIR-February Blogbeiträge:
 - Let's Talk FAIR – Die Findability digitaler Editionen <https://textplus.hypotheses.org/5736>
 - Let's talk FAIR – Die Zugänglichkeit digitaler Editionen <https://textplus.hypotheses.org/5904>
 - Let's talk FAIR – Die Interoperabilität von Editionen <https://textplus.hypotheses.org/8355>
 - Let's talk FAIR – Zur Nachnutzbarkeit digitaler Editionen <https://textplus.hypotheses.org/9338>
 - Let's talk FAIR again. Nachbetrachtung zum zweiten FAIR February <https://textplus.hypotheses.org/10982>
- Helpdesk: <https://text-plus.org/helpdesk/>
- Registry: <https://registry.text-plus.org>
- Textplus-Portal: <https://text-plus.org>
- Guidelines: https://gitlab.gwdg.de/textplus/textplus-editions/guidelines_sde
- RIDE (FAIR Criteria) <https://ride.i-d-e.de/issues/issue-16/> ; <https://ride.i-d-e.de/issues/issue-17/>
- Text+ RDMO: <https://text-plus.org/themen-dokumentation/forschungsdatenmanagement/#rdmo>

Veranstaltungsankündigung

- Wir geben einen Workshop zum Thema „Qualitativ hochwertige Metadaten in digitalen Editionen“ an der DHd 25 in Bielefeld am 4. März 2025.
[Programm](#), [Anmeldung](#)

Vielen Dank für Ihre Aufmerksamkeit!

Die vorliegende Präsentation wurde im Rahmen des Konsortiums Text+ im Kontext der Arbeit des Vereins Nationale Forschungsdateninfrastruktur (NFDI) e.V. verfasst. NFDI wird von der Bundesrepublik Deutschland und den 16 Bundesländern finanziert, und das Konsortium Text+ wird gefordert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 460033370. Die Autor:innen bedanken sich für die Förderung sowie Unterstützung. Ein Dank geht außerdem an alle Einrichtungen und Akteur:innen, die sich für den Verein und dessen Ziele engagieren.