

# Datenqualitätssicherung im Forschungsprozess am Beispiel von Zugversuchen

## *Data quality assurance in the research process using the example of tensile tests*

Laura Müller<sup>1,\*</sup>, Max Leo Wawer<sup>2</sup>, Norman Heimes<sup>3</sup>, Johanna Uhe<sup>3</sup>,  
Oliver Koepler<sup>4</sup>, Sören Auer<sup>4</sup>, Roland Lachmayer<sup>2</sup>, Iryna Mozgova<sup>1</sup>

<sup>1</sup> Data Management in Mechanical Engineering, Paderborn University

<sup>2</sup> Institute of Product Development, Leibniz University Hannover

<sup>3</sup> Institute of Forming Technology and Machines, Leibniz University Hannover

<sup>4</sup> TIB - Leibniz Information Centre for Science and Technology and University Library

\* Korrespondierender Autor:

Laura Müller

Datenmanagement im Maschinenbau (DMB), Universität Paderborn

Warburger Str. 100

33098 Paderborn

☎ 05251 60 2227

✉ [laura.mueller@upb.de](mailto:laura.mueller@upb.de)

---

### Abstract

Progressive digitization throughout the entire product data life cycle requires a more sensitive handling and understanding of data within engineering processes. Regarding engineering research data, the aim is to implement the FAIR data principles (Findable, Accessible, Interoperable, Reusable) to guarantee the post-usability of research data. To ensure the quality of data throughout the entire research process a methodical approach had been developed. Based on the quality categories Intrinsic, Representative, Contextual and Available, the related quality dimensions are considered differentiated along the research data life cycle and presented in a concept. As a use case, this concept is carried out on a tensile test with documentation of results in a research data management system.

---

### Keywords

*data quality assurance, quality dimension, research data management, research data life cycle*

---

---

## 1. Motivation

Die fortschreitende Digitalisierung über den gesamten Produktdatenlebenszyklus ermöglicht eine holistische Betrachtung von Daten innerhalb der ingenieurstechnischen Prozesse und damit ein ganzheitliches Verständnis der Daten. Für Forschende ergeben sich als Folge dessen und durch die Anforderung der Umsetzung der *FAIR-Data*-Prinzipien (*Findable, Accessible, Interoperable, Reusable*) [1] neue Herausforderungen. In großen Verbundprojekten stehen Forschende verschiedener Fachbereiche und Institutionen vor der Herausforderung projektbezogene Forschungsdaten in einer bestimmten Qualität zugänglich zu machen, damit diese entsprechend der eng vernetzten Arbeitsprogrammen von anderen Forschenden nachgenutzt werden können. Hiermit gehen die Anforderungen an ein proaktives und qualitätsgesichertes Forschungsdatenmanagement (FDM) einher, welches die Forschung gemäß einer guten wissenschaftlichen Praxis [2] effizienter und offener durch nachnutzbare Forschungsdaten gestalten soll.

Mit Hilfe von FDM-Systemen können in großen Verbundprojekten Strukturen und Prozesse etabliert werden, um die *FAIR-Data*-Prinzipien umzusetzen und damit die Nachnutzbarkeit von Forschungsdaten zu ermöglichen [3]. Über die *FAIR-Data*-Prinzipien hinaus gilt es die Qualität dieser Strukturen und Prozesse beim Umgang mit Daten sowie der Daten selbst während des gesamten Forschungsprozesses zu sichern.

Für die Datenqualität gibt es unterschiedliche Ansätze diese zu beschreiben. Aufgrund verschiedener Sichtweisen ergeben sich unterschiedlich relevante Eigenschaften, sodass der Prozess der Informationsproduktion die Datenqualität bestimmt [4]. Zur methodischen Spezifikation der Datenqualität können bestimmte Qualitätskategorien mit einhergehenden Qualitätsdimensionen definiert werden [5,6]. Zur Gewährleistung der Datenqualitätssicherung entlang eines Forschungsprozesses gilt es entsprechende Qualitätsdimensionen zu berücksichtigen.

In diesem Zusammenhang lässt sich ein Forschungsprozess entlang des Forschungsdatenlebenszyklus in sechs Phasen unterscheiden [7], wobei relevante Qualitätsdimensionen nachfolgende Phasen beeinflussen. Einhergehende Herausforderungen sind die Identifizierung und Berücksichtigung phasenspezifischer Qualitätsdimensionen, um die Datenqualität zu sichern.

Dieser Beitrag zeigt anhand der Datenerhebung in Zugversuchen, eingebettet in ein Verbundprojekt, wie Qualitätsdimensionen methodisch entlang des Forschungsdatenlebenszyklus identifiziert und in einem FDM-System Berücksichtigung finden können. Hierbei werden sowohl praktische Aspekte der Durchführung von Zugversuchen sowie die Übertragung in ein FDM-System untersucht. Für das methodische Vorgehen wird sich auf die erzielten Ergebnisse aus einem Workshop der Nationalen Forschungsdateninfrastruktur für Ingenieurwissenschaften [8] und auf die Befragung von Forschenden für Zugversuche bezogen. Die Ergebnisse zeigen einen Lösungsansatz zum bewussten Umgang mit qualitätsgesicherten und nachnutzbaren Forschungsdaten durch die Identifizierung relevanter Qualitätsdimensionen im Forschungsprozess.

## 2. Stand der Technik

In diesem Kapitel werden die grundlegenden Inhalte für das in einem Verbundprojekt entwickelte FDM-System vorgestellt. Diese Inhalte werden zusätzlich in die Umsetzung des FDM, orientiert an den Phasen des Datenlebenszyklus, eingebettet. Erweitert wird das Kapitel um die Hintergründe zur Qualitätssicherung im Kontext des FDM. Der Fokus liegt dabei auf den bereits erwähnten *FAIR-Data*-Prinzipien [1] und den Datenqualitätsdimensionen nach Wang und Strong [6], welche den Bezugsrahmen zur Datenqualitätssicherung darstellen.

### 2.1. Forschungsdatenmanagement-System

Das FDM umfasst alle Methoden und Verfahren zum Umgang mit anfallenden Digital Objects in der Forschung. Diese sind vielfältig und abhängig von der Forschungsmethode, wie beispielsweise entwickelte Software oder Datenartefakte aus Experimenten. Die relevanten

---

Inhalte zur Umsetzung des FDM werden oftmals an Modellen des Datenlebenszyklus dargestellt [9]. Dieser umfasst die grundlegenden Phasen: Planen, Erheben, Analysieren, Archivieren, Zugriff und Nachnutzung. In den einzelnen Phasen spielen unterschiedliche Aspekte eine wichtige Rolle. So gilt es im Rahmen der Datenerhebung und Datenanalyse diese mittels einer Datendokumentation ausreichend zu beschreiben, so dass eine Nachvollziehbarkeit und Rekonstruktion erfolgen kann. Zur Umsetzung des FDM werden technische Lösungen in Form von FDM-Systemen entwickelt, welche Forschende unterstützen. Dies können Datenrepositorien zur Zugänglichkeit von Daten oder Werkzeuge zur Erstellung von Datenmanagementplänen sein [10].

Innerhalb des Sonderforschungsbereichs (SFB) 1153 wird die Prozesskette zur Herstellung hybrider Hochleistungsbauteile durch *Tailored Forming* erforscht, um ressourceneffizienten Leichtbau mittels lokal angepasster Eigenschaften von Massivbauteilen umzusetzen [11]. So bilden die anfallenden Forschungsdaten und zugehörigen Prozessbeschreibungen die Grundlage für die relevante Prozesskettenentwicklung und im Falle der Zugversuche für die Entwicklung von Materialmodellen der hybriden Bauteile. Diese dienen als Inputdaten für die Berücksichtigung der mechanischen Kennwerte und einhergehenden Materialeigenschaften innerhalb der im SFB erforschten Simulationen. Durch die Beteiligung verschiedener Institutionen und die räumliche Diskrepanz unterschiedlicher Teilprojekte ist ein gemeinschaftliches FDM bei der Höhe der Informationsweitergabe unerlässlich. Im Rahmen des SFB 1153 wird dahingehend ein FDM-System für die Forschenden entwickelt, welches eine technische Lösung für das FDM in großen Verbundprojekten darstellt [3]. Dabei setzt sich das FDM-System aus zwei eigenständigen Lösungen zusammen. Zum einen besteht es aus einem Wissensmanagementsystem zur Dokumentation der vorherrschenden Forschungsaktivitäten und der vornehmlichen Protokollierung der Experimente sowie Prozessschritte in dem Verbundprojekt, zum anderen aus einem Datenrepositorium zum Archivieren und Zugänglichmachen von Forschungsdaten innerhalb des SFB [12]. Anfallende Forschungsdaten und alle weiteren Artefakte sowie Informationen, die zur Kontextualisierung der Datenerhebung und Analysen relevant sind, werden durch das FDM-System erfasst. Diese beiden Lösungen sind essenziell bei der Prozesskettenentwicklung der *Tailored-Forming*-Technologie und weisen wesentliche Schnittstellen zur Verknüpfung der Inhalte zu einem gesamtheitlichen FDM-System auf [13].

## 2.2. Qualitätssicherung

Die Datenqualität wird im Allgemeinen als die Eignung von Daten für unterschiedliche Verwendungszwecke in Hinblick auf den Grad ihrer Erfassung und Generierung beschrieben, kurz „*fitness for use*“ und ist somit abhängig vom Kontext und Anwendungsfall. Die Qualitätssicherung ist Teil des Qualitätsmanagements und betrifft Umsetzungen zur Erfüllung der Qualitätsanforderungen [14]. Im Rahmen des FDM spielt die Qualität der Digital Objects eine wesentliche Rolle für die wissenschaftliche Nachnutzung und Reproduzierbarkeit der Forschungsergebnisse, welche durch die *FAIR-Data*-Prinzipien [1] beschrieben werden können. Die Beachtung der Prinzipien hat Auswirkungen auf die Qualität der Datenartefakte im Rahmen der Forschungsdateninfrastruktur. Der Fokus liegt hierbei auf der Maschinenlesbarkeit von Daten im Kontext einer Weiterverarbeitung großer und komplexer Datenmengen. Die Anwendung und Messung der Prinzipien spielt über den gesamten Lebenszyklus der Daten eine wichtige Rolle, um eine Weiterverarbeitung der Digital Objects kontextuell zu sichern [15]. Die *FAIR-Data*-Prinzipien beziehen sich primär auf die Qualität der Inhalte in einem Datenrepositorium als auf die inhaltliche Qualität der Digital Objects. So wurden von Zaveri et al. [5] auf Basis einer Literaturrecherche, in welcher vorherrschende Qualitätsdimensionen verschiedener Veröffentlichungen identifiziert wurden, 18 Qualitätsdimensionen definiert, welche nach den von Wang et al. eingeführten

Datenqualitätskategorien (*Intrinsisch, Verfügbar, Repräsentativ, Kontextuell*) [6] klassifiziert wurden (vgl. Bild 1).

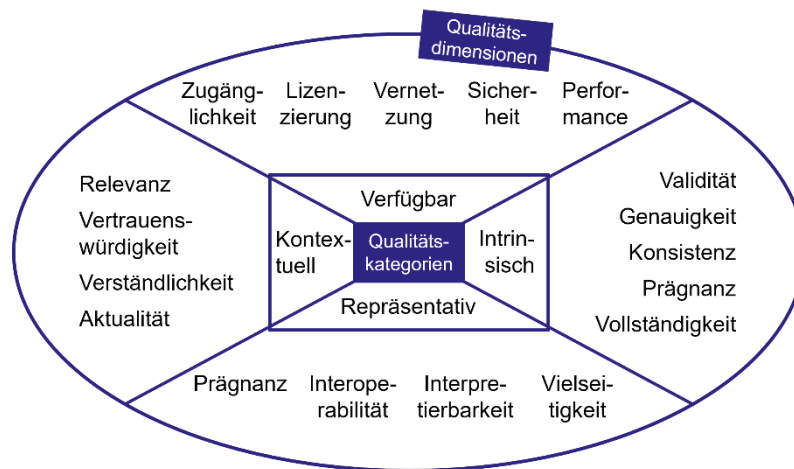


Bild 1: Qualitätskategorien und Qualitätsdimensionen nach Zaveri und Wang [5,6]

Die Qualitätskategorien wurden von Kindling [16] in den Kontext des FDM gesetzt. So basieren die Qualitätskategorien *Intrinsisch* und *Kontextuell* auf Kriterien, die von den Fachbereichen bereits während der Vorbereitung der Datenerhebung und -aufbereitung berücksichtigt werden müssen. *Repräsentativ* bezieht sich auf die Anreicherung der Digital Objects mit Metadaten und die Verwendung spezifischer Dateiformate. Die Kategorie *Verfügbar* legt Bezug zu der Forschungsdateninfrastruktur und der Bereitstellung der Digital Objects. Daraus ergeben sich zwei untrennbare Ebenen der Qualität, welche unmittelbar den Digital Objects zugeschrieben werden kann und die der Einbettung in eine Forschungsdateninfrastruktur zur Archivierung und Bereitstellung.

### 3. Forschungsproblem und -ziel

Damit die Transparenz sowie Nachverfolgung von Daten aus der SFB spezifischen *Tailored-Forming*-Prozesskette gewährleistet sind, wird wie in 2.1. beschrieben ein FDM-System entwickelt, um die Forschenden im FAIRen Umgang ihrer Prozessdaten zu unterstützen. Die Konsolidierung und Harmonisierung dieser Daten sowie die Dokumentation der Datengenerierung stellen in diesem Zusammenhang für die Rückverfolgbarkeit sowie Analyse einen kritischen Erfolgsfaktor für einen SFB dar [13].

Innerhalb des Forschungsprozesses spielen während der Datenerhebung Qualitätsmaßnahmen eine relevante Rolle, um letztendlich qualitative, inhaltlich nachnutzbare Forschungsdaten gewährleisten zu können. Zur Erreichung dieser nachnutzbaren Forschungsdaten, sind sowohl während der Forschungsprozesse als auch innerhalb des FDM-Systems relevante Qualitätsdimensionen in Hinblick auf die Datenqualität zu sichern [6]. In der Praxis führen Forschende der jeweiligen Teilprozesse zunächst ihre Versuche durch und tragen im Anschluss ihre Ergebnisse in das FDM-System ein. Hieran lassen sich bereits zwei verschiedene Perspektiven in Bezug auf die Datenqualität erkennen. Aus Sicht der Forschenden ist für die Nachnutzung der Forschungsdaten von Relevanz, welche Versuchsparameter bei der praktischen Durchführung erhoben wurden, um diese im Nachgang vergleichen und analysieren zu können. Demgegenüber ist es z. B. in Bezug auf das FDM-System qualitativ wichtig, möglichst forschungsspezifische und kontextbezogene Metadaten zu nutzen, um die Forschungsdaten inhaltlich nachnutzbar zu gestalten [17]. Daher besteht in Bezug auf die *Tailored-Forming*-Prozesskette die große Herausforderung beide Perspektiven zusammenzubringen, sodass insgesamt eine hohe Datenqualität prozessübergreifend erzielt werden kann. In diesem Beitrag wird erörtert, welche

---

Qualitätsdimensionen innerhalb der verschiedenen Phasen des Forschungsdatenlebenszyklus für die Nachnutzung von Forschungsdaten am Beispiel von Zugversuchen von besonderer Wichtigkeit sind. Dahingehend sollen im weiteren Verlauf folgende Forschungsfragen diskutiert werden:

- Welche Qualitätsdimensionen sind in Anbetracht der Datendokumentation und Datenpublikation für die Nachnutzbarkeit der Forschungsdaten relevant?
- Wie müssen Versuchsplanung und -durchführung im Sinne einer nachhaltigen Datennutzung dokumentiert werden?
- Wie lassen sich Maßnahmen zur Qualitätssicherung in FDM-Systeme integrieren?

#### 4. Vorgehensweise zur Identifizierung relevanter Qualitätsdimensionen in Forschungsprozessen

Zur Rekonstruktion des Forschungsprozesses und der Identifikation relevanter Datenqualitätskriterien wird auf Methoden aus dem Bereich der Datenqualitätssicherung und -verbesserung zurückgegriffen [18]. Mit dem Ziel der Identifikation qualitätsrelevanter Inhalte über datenbezogene Forschungsinhalte während des Forschungsprozesses wird die Durchführung von Zugversuchen und Übertragung der Digital Objects in das FDM-System als *Use Case* herangezogen. Dabei spielt die Durchführung von Zugversuchen bei der Prozesskettenentwicklung innerhalb des SFB eine wichtige Rolle für die Auslegung nachgelagerter Prozessschritte und der Prozesskettenentwicklung. So werden die Analysen der Zugversuche mit der Identifizierung mechanischer Kennwerte z. B. in Prozesssimulationen eingebettet, dessen Materialmodelle auf den Forschungsergebnissen der Zugversuche beruhen und die Daten dahingehend nachgenutzt werden. Dabei werden innerhalb des SFB stranggepresste hybride Bauteile untersucht und Materialeigenschaften bestimmt [19]. Deswegen spielt die Datenqualität eine wichtige Rolle bei der Prozesskettenentwicklung innerhalb des SFB.

Die Rekonstruktion des Zustands zielt darauf ab, kontextbezogene Informationen über die datenbezogenen Forschungsaktivitäten und damit verbundene qualitätsbezogene Fragestellungen zu sammeln. Dafür wird in einem ersten Schritt der Anwendungsfall beschrieben. Orientiert am Forschungsdatenlebenszyklus wird die Durchführung von Zugversuchen phasendifferenziert von der Planung bis zum Zugriff mit einem Forschenden beschrieben. Die Qualitätssicherung der Aktivitäten in dieser Phase soll eine qualitativ gesicherte Nachnutzung der Daten im Rahmen des Forschungsdatenlebenszyklus ermöglichen. Zur Definition der qualitätsbezogenen Anforderungen werden zwei Workshops durchgeführt, um die Anforderungen an das FDM und der Forschenden aus dem Bereich der Zugversuche zu definieren. Hierfür werden in einem Workshop mit Forschenden der Nationalen Forschungsdateninfrastruktur für Ingenieurwissenschaften (NFDI4Ing) aus der *Special Interest Group quality assurance & metrics for FAIR data*<sup>1</sup>, welche als Spezialisten des FDM gelten, qualitätsbezogene Anforderungen aus Sicht des FDM definiert. In einem zweiten Workshop werden unter den gleichen Umständen die qualitätsbezogenen Anforderungen der Forschenden aus dem Bereich der Zugversuche definiert. Auf diese Weise ergeben sich sowohl von Anwenderseite als auch von den FDM-Spezialisten qualitätsbezogene Anforderungen, die in einem nächsten Schritt den vorherrschenden Datenqualitätsdimensionen (vgl. Bild 1) zugeordnet werden. Auf diese Weise ergeben sich differenzierte qualitätsrelevante Anforderungen entlang des Forschungsprozesses, welche es mit dem FDM-System zu sichern gilt.

---

<sup>1</sup> <https://nfdi4ing.de/special-interest-groups-sig/qa-metrics/>

## 5. Ergebnis

In diesem Kapitel werden die Ergebnisse zur Rekonstruktion des Anwendungsfalles und die Identifikation relevanter Qualitätsdimensionen nach der vorgestellten Methode dargestellt. Weiter soll eine gesamtheitliche Qualitätssicherung bei der Versuchsdurchführung und der Nutzung des FDM-Systems in Hinblick auf eine qualitätsgesicherte Nachnutzung der Daten adressiert werden.

### 5.1. Identifikation von Qualitätsdimensionen über den Forschungsprozess

Zur Identifizierung relevanter Qualitätsdimensionen wurde in einem ersten Schritt der Forschungsprozess mit dem Fokus auf die Durchführung und Umsetzung des Forschungsdatenmanagements phasenorientiert am Datenlebenszyklus mit einem Forschenden aus diesem Forschungsfeld des SFB aufgenommen. So haben sich datenintensive Aktivitäten der Forschenden in Abhängigkeit der Phasen ergeben. Der identifizierte Anwendungsfall dient als Grundlage zur weiteren Identifikation qualitätsrelevanter Anforderungen. In Abbildung 2 ist das Ergebnis der Aufnahme des Anwendungsfalles dargestellt. In diesem Kontext muss die Qualität der Daten und die der Forschungsdateninfrastruktur für eine adäquate Nachnutzung sichergestellt werden. Darum wird die Phase der Nachnutzung im Nachgang nicht genauer betrachtet, sondern stellt die Anforderung, dass in dieser Phase qualitativ gesicherte Forschungsdaten vorliegen sollen.

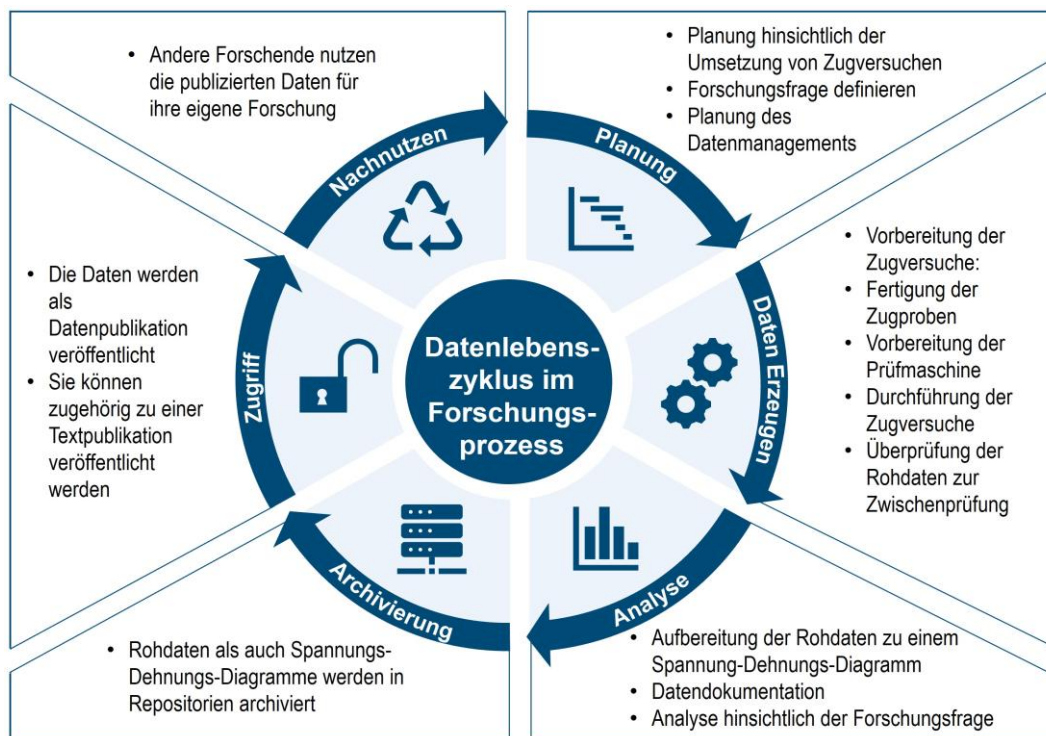


Bild 2: Phasenorientierter Forschungsprozess für Zugversuche entlang des Forschungsdatenlebenszyklus

Auf Grundlage dessen wurden Anforderungen in Hinblick auf die Qualitätssicherung im Kontext des FDM definiert. Hierfür wurde ein Workshop ausgearbeitet, der sowohl mit FDM-Spezialisten aus dem Konsortium der NFDI4Ing als auch mit Forschenden aus dem SFB zur Durchführung von Zugversuchen durchgeführt wurde. Das Ergebnis dieser Workshops ist die phasenorientierte Definition von Anforderungen in der Umsetzung des FDM aus Sicht der Anwender als auch FDM-Spezialisten. Diese Anforderungen wurden im Nachgang den aus der Literatur bekannten Datenqualitätsdimensionen zugeordnet, um relevante Kriterien in

Abhängigkeit des Forschungsprozesses zu identifizieren. In Tabelle 1 sind die Ergebnisse der Workshops zur Definition qualitätsbezogener Anforderungen mit den zugehörigen Datenqualitätsdimensionen entlang des beschriebenen Anwendungsfalles dargestellt. Die identifizierten Anforderungen, welche explizit von den FDM-Spezialisten genannt wurden, sind in der Farbe Grau hervorgehoben.

Tabelle 1: Relevante Qualitätsdimensionen im Kontext des FDM

Phase	Identifizierte Anforderungen	Qualitätsdimension
Planung	Anpassung der Versuche bei möglichen Fehlern und Änderungen	Aktualität
	Strukturierte und einheitliche Versuchsplanung	Verständlichkeit, Vollständigkeit
	Versuchsplanung nach DIN Normen auslegen (Probenanzahl)	Validität
	Zugversuche hinsichtlich der Forschungsfrage planen	Relevanz, Verständlichkeit
Erhebung	Genormte Fertigung der Versuchsproben (DIN) und Prüfung	Validität
	Kalibrierung der Prüfmaschine, Leerfahrten	Genauigkeit, Validität
	Prozessspezifische Dokumentation des Erhebungsprozesses (Prozessparameter, Maschineneinsatz)	Verständlichkeit, Vollständig, Relevanz, Vertrauenswürdigkeit
	Reproduzierbarkeit der Daten gewährleisten	Vertrauenswürdigkeit, Konsistenz,
	Vorscreening der Daten der einzelnen Daten hinsichtlich möglicher Fehler	Prägnanz, Genauigkeit, Validität, Konsistenz
Analyse	Überprüfung der erhobenen Daten aus den einzelnen Zugversuchen im Gesamtkontext	Konsistenz
	Dokumentation der Auswertungsprozeduren und Datentransformationen	Interpretierbarkeit, Verständlichkeit, Vertrauenswürdigkeit
Archivierung	Auswahl und Format der zu archivierenden Daten zur Reproduzierbarkeit und Nachnutzbarkeit der Daten	Interoperabilität, Relevanz
	Verknüpfung der Daten mit der Datendokumentation	Verständlichkeit, Vertrauenswürdigkeit, Interpretierbarkeit
	Sicherstellung eines Archivierungsmediums entsprechend der Projektbedingungen	Sicherheit, Zugänglichkeit
Zugriff	Sicherstellung eines Zugangs zu den Daten und Dokumentationen entsprechend der Projektbedingungen	Lizenzierung, Sicherheit, Zugänglichkeit

Als Ergebnis hat sich eine differenzierte Relevanz der Dimensionen in Abhängigkeit der Phasen ergeben. In dem Bereich der Qualitätssicherung spielt die Sichtweise und der Zweck der Forschenden eine relevante Rolle. So liegen die Anforderungen während der Planung, Datenerhebung und Datenanalyse zumeist im Bereich der intrinsischen und kontextuellen Qualitätskategorie (vgl. Bild 1), da hier die Erhebung der Daten geplant und durchgeführt wird, welche in einem unmittelbaren Zusammenhang mit der Qualität der Daten an sich und für den

Forschungszweck qualitativ gesichert sein sollen. Hierbei ist die Orientierung an eine genormte Durchführung von Zugversuchen nach DIN 6892 [20] und die genormte Fertigung der Zugproben nach DIN 50125 [21] relevant hinsichtlich der Datenqualität durch qualitativ gesicherte Versuchsdurchführungen. Auch steht in diesem Kontext die Kalibrierung der eingesetzten Maschinen und Messsysteme mit der Durchführung von Leerfahrten und die Prüfung der gefertigten Zugproben im Fokus. Zudem wird immer ein gewisser Anteil an Zugproben zurückgehalten, um auf mögliche Fehler Einfluss nehmen zu können und dahingehend die Versuchsplanung anzupassen. In Anbetracht der vorgesehenen Datenarchivierung und Datenzugänglichkeit muss auch eine Dokumentation der Versuchsdurchführung und Analysen in diesen Phasen erfolgen, um die Interpretierbarkeit und Verständlichkeit der Daten durch Dritte gewährleisten zu können.

In den Phasen der Archivierung und des Zugriffs liegt der Fokus auf den Qualitätskategorien *Repräsentativ* und *Verfügbar* und bezieht sich zunehmend auf die Interaktion und Nutzung von Forschungsdateninfrastrukturen. Hier gilt es die Strukturen nach den *FAIR-Data*-Prinzipien auszulegen. Jedoch müssen Forschende schon während der Planung und Erhebung sicherstellen, dass eine Interoperabilität durch die Auswahl entsprechender Dateiformate und die Verwendung bestimmter Metadatenschemata gesichert ist. Insgesamt ergibt sich eine ganzheitliche Relevanz der Qualitätskategorien im Forschungsprozess zur Sicherung der Qualität der Digital Objects in der Forschung im Kontext des FDM. Die Ergebnisse decken sich dabei mit den Erkenntnissen von Kindling [16] und konnten in dieser Arbeit forschungsspezifisch herausgearbeitet werden und dienen als Grundlage für die Qualitätssicherung in dem vorherrschenden FDM-System.

## 5.2. Integration der Qualitätssicherung in FDM-Systeme

Nach der praktischen Durchführung der Zugversuche erfolgt die Datendokumentation und Datenbereitstellung im FDM-System, welches mit dessen Hauptfunktionen in Abbildung 3 ersichtlich ist. Dabei wird dargelegt inwieweit die in Tabelle 1 ermittelten Qualitätsdimensionen mit Hilfe des SFB-spezifischen FDM-Systems erfüllt werden können.

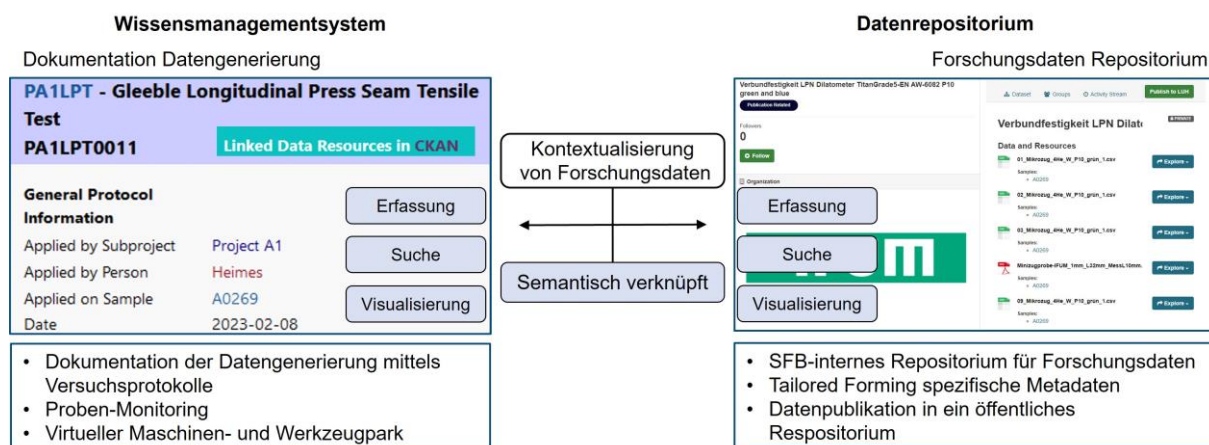


Bild 3: Systemarchitektur und Hauptaufgaben des FDM-Systems im SFB 1153

In dem Wissensmanagementsystem können mit versuchsspezifischen Protokolltemplates hinsichtlich der Relevanz Versuchsparameter und Materialeigenschaften im Kontext der Prozesskette dokumentiert werden. Hierfür wurde ein *Tailored Forming* spezifisches Vokabular für die Interoperabilität, Interpretierbarkeit und Verständlichkeit der Forschungsinhalte im Sinne einer nachhaltigen Datendokumentation aufgebaut [21]. Durch die Vorgabe definierter einheitlicher Templates werden Forschende bezüglich ihrer Eingabe hinsichtlich der Nutzbarkeit ihrer Forschungsdaten beeinflusst. Demnach können mit Hilfe



---

des Wissensmanagementsystems Qualitätsdimensionen wie Vollständigkeit, Verständlichkeit und Vertrauenswürdigkeit erfüllt werden. Die qualitätsgesicherte Versuchsdurchführung wird hierbei von den Forschenden durch normorientierte Versuchsplanungen in Hinblick auf die Genauigkeit, Validität und Reproduzierbarkeit der Versuche gesichert. Darüber hinaus werden Sichtprüfungen von gefertigten Proben dokumentiert [22]. Als ableitbares Erkenntnis unterstützt das Wissensmanagementsystem die Forschenden vor allem bei den Qualitätskategorien *Kontextuell* und *Intrinsisch*.

In dem Datenrepositorium können die Forschenden ihre Forschungsdaten archivieren und neben allgemeinen Beschreibungen mit vorgegebenen projektspezifischen kontextuellen Metadaten annotieren [23]. Zudem lassen sich einzelne Forschungsdaten als gesammelte Datensets darstellen. Das Datenrepositorium ist als geschlossenes internes System aufgebaut, ermöglicht aber eine direkte Überführung in das öffentliche Repositorium der Universität, um eine öffentliche Nachnutzung der Daten zu ermöglichen. Hierbei werden die Dimensionen Sicherheit, Lizenzierung und Zugänglichkeit gesichert. Außerdem wird innerhalb des SFB die Datenauswahl der Datenbereitstellung forschungsspezifisch eruiert und die Auswahl der Dateiformate bestimmt, um die Relevanz und Interoperabilität der zugänglich gemachten Daten zu sichern. Somit trägt das Datenrepositorium primär zu den Qualitätskategorien *Verfügbar* und *Repräsentativ* bei.

Die beiden Systeme weisen Schnittstellen auf, um die Dokumentation und die Daten miteinander verknüpfen zu können und somit ein gesamtes FDM-System abzubilden. Durch die Verknüpfung der Forschungsdaten im Datenrepositorium mit den Versuchsprotokollen aus dem Wissensmanagementsystem wird die Kontextualisierung der Daten weiter erhöht und sichert die Interpretierbarkeit, Verständlichkeit und Vertrauenswürdigkeit.

Somit kann durch den forschungsspezifischen Aufbau des FDM-Systems dazu beigetragen werden die Datenqualität zu beeinflussen und zu sichern. Hierbei können auch Vollständigkeitsüberprüfungen relevanter Felder, Reviews durch FDM-Experten und Schulungen der Nutzenden die Qualität sichern und fördern.

## 6. Zusammenfassung und Ausblick

In diesem Beitrag konnten anhand des Forschungsdatenlebenszyklus relevante Qualitätsdimensionen und Qualitätskriterien am Beispiel von Zugversuchen methodisch abgeleitet werden. Der als Use Case herangezogene Anwendungsfall lässt sich durch das methodische Vorgehen und die generische Beschreibung sowie Identifikation der qualitätsrelevanten Anforderungen auch auf andere Forschungsfelder in der Umsetzung des FDM übertragen. Als kritischer Erfolgsfaktor wurde die Zusammenführung zweier Perspektiven erkannt, welcher für die umfassende Nachnutzung von Forschungsdaten nur im Zusammenspiel möglich ist. Zum einen müssen für Forschende relevante Qualitätskriterien hinsichtlich der Forschungsergebnisse erfüllt werden und auf der anderen Seite Forschungsdateninfrastruktur-spezifische Qualitätskriterien. Durch die Hinzunahme eines FDM-Systems konnte belegt werden, dass dieses ein Hilfsmittel zur Datenqualitätssicherung darstellt. Gleichzeitig bestehen Fehleranfälligkeiten, da es zum Teil von der subjektiven Eintragung der Forschenden abhängig ist. Damit weitere Qualitätsaspekte zukünftig gesichert werden, können Aspekte der Qualitätsmetrik aufgenommen werden, um die Qualität weitergehend messen zu können.

Weiter sollen die Inhalte für die Entwicklung von Reifegradmodellen zur qualitativen Bewertung von Datenmanagementprozessen in Forschungsprojekten genutzt werden, da die Reife dieser Prozesse eine qualitativ gesicherte Umsetzung vorsieht (vgl. [24]).

## Danksagung

Die Autoren möchten sich bei Bund, Ländern, bei der Gemeinsamen Wissenschaftskonferenz (GWK), dem Ministerium für Wissenschaft und Kultur (MWK) und

---

Deutsche Forschungsgemeinschaft (DFG) für die Förderung und Unterstützung im Rahmen des Konsortiums NFDI4Ing und des Sonderforschungsbereichs 1153 „Prozesskette zur Herstellung hybrider Hochleistungsbauteile durch *Tailored Forming*“, Projektnummer 442146713 und Projektnummer 252662854, bedanken.

## Literaturverzeichnis

- [1] Wilkinson, Mark D. et al.: The FAIR Guiding Principles for scientific data management and stewardship. In: *Scientific data* 3 (2016), S. 160018.
- [2] Deutsche Forschungsgemeinschaft E.V.: Leitlinien zur Sicherung guter wissenschaftlicher Praxis (September 2019). URL [https://www.dfg.de/download/pdf/foerderung/rechtliche\\_rahmenbedingungen/gute\\_wissenschaftliche\\_praxis/kodex\\_gwp.pdf](https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf) – Überprüfungsdatum 2023-03-30.
- [3] Mozgova, Iryna et al.: Research Data Management System for a large Collaborative Project. In: *Balancing Innovation and operation: The Design Society*, 2020.
- [4] Helfert, Markus: Massnahmen und Konzepte zur Sicherung der Datenqualität. In: Jung, Reinhard; Winter, Robert (Hrsg.): *Data Warehousing Strategie*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000 (Business Engineering), S. 61–77.
- [5] Zaveri, Amrapali et al.: Quality assessment for Linked Data: A Survey. In: *Semantic Web* 7 (2015), Nr. 1, S. 63–93.
- [6] Wang, Richard Y.; Strong, Diane M.: Beyond Accuracy: What Data Quality Means to Data Consumers. In: *Journal of Management Information Systems* 12 (1996), Nr. 4, S. 5–33.
- [7] Liebig, Stefan; Matiaske, Wenzel; Rosenbohm, Sophie: *Handbuch Empirische Organisationsforschung*. Wiesbaden: Springer Fachmedien Wiesbaden, 2017.
- [8] Schmitt, Robert H. et al.: NFDI4Ing - the National Research Data Infrastructure for Engineering Sciences. 2020.
- [9] Wolf, A. Harry; Leppla, Cindy: Harmonisierung von Datenlebenszyklus-Modellen: Nutzung von Synergien für optimierte Anwendungen im FDM. In: *Bausteine Forschungsdatenmanagement* (2020), Nr. 2, S. 1–19. URL <https://bausteine-fdm.de/article/view/8281>.
- [10] Mozgova, Iryna et al.: Product Life Cycle Oriented Data Management Planning with RDMO at the Example of Research Field Data. 2022 (E-Science-Tage 2021).
- [11] Behrens, Bernd-A.; Uhe, Johanna.: Introduction to tailored forming. In: *Production Engineering* 15 (2021), Nr. 2, S. 133–136.
- [12] Sheveleva, Tatyana et al.: Creation of a Knowledge Space by Semantically Linking Data Repository and Knowledge Management System - a Use Case from Production Engineering. In: *IFAC-PapersOnLine* 55 (2022), Nr. 10, S. 2030–2035.
- [13] Wawer, Max L. et al.: Parametrization of a Hybrid Component Production Process Chain Based on Semantically Annotated Data: Novi Sad: University of Novi Sad, 2022.
- [14] Norm DIN EN ISO 9000, 2015. Qualitätsmanagementsysteme - Grundlagen und Begriffe.
- [15] Devaraju, Anusuriya et al.: From Conceptualization to Implementation: FAIR Assessment of Research Data Objects. In: *Data Science Journal* 20 (2021).
- [16] Kindling, Maxi: Qualitätssicherung im Umgang mit digitalen Forschungsdaten / Quality assurance of digital research data / La garantie de la qualité des données numériques de recherche. In: *Information – Wissenschaft & Praxis* 64 (2013), 2-3.
- [17] Batini, Carlo et al.: Methodologies for data quality assessment and improvement. In: *ACM Computing Surveys* 41 (2009), Nr. 3, S. 1–52.
- [18] Behrens, Bernd-A. ; Chugreev, Alexander; Matthias, Tim: Characterisation of the joining zone of serially arranged hybrid semi-finished components. In: *AIP Conference Proceedings* 1960 (2018), Nr. 1, S. 40002.
- [19] Norm DIN EN ISO 6892-2, 2018. Zugversuch - Teil 2: Prüfverfahren bei erhöhter Temperatur.
- [20] Norm DIN 50125, 2022. Zugproben - Prüfung metallischer Werkstoffe.
- [21] Altun, Osman et al.: Enhanced Findability and Reusability of Engineering Data by Contextual Metadata. In: *Proceedings of the Design Society* 3 (2023), S. 1635–1644.
- [22] Sheveleva, Tatyana et al.: Development of a Domain-Specific Ontology to Support Research Data Management for the Tailored Forming Technology. In: *Procedia Manufacturing* 52 (2020), S. 107–112.
- [23] Sheveleva, Tatyana et al.: Ontology-Based Documentation of Quality Assurance Measures Using the Example of a Visual Inspection. In: Valle, Lehmhus et al. (Hg.) 2023 – *Advances in System-Integrated Intelligence*, Bd. 546. S. 415–424.
- [24] Wawer, Max L.; Lachmayer, Roland: Reifegradmodell für die Verwaltung des Datenzugriffs. 2023 (E-Science-Tage 2023) (In Press).