



Kommission für
Forschungsinformationen
in Deutschland

KDSF 2.0:

Flexible und qualitätsgesicherte Forschungsberichterstattung nach dem Baukastenprinzip

Grundsatzpapier der Kommission für
Forschungsinformationen in Deutschland

Impressum

Stand

August 2024

Herausgeber

Kommission für Forschungsinformationen in Deutschland (KFiD)

Geschäftsstelle

Schützenstr. 6a

10117 Berlin, Deutschland

Web: www.kfid-online.de

E-Mail: info@kfid-online.de

Redaktion

Dr. Sophie Biesenbender

Gestaltung & Satz

Yasemin Peker

Die Kommission für Forschungsinformationen in Deutschland wird auf der Grundlage einer Verwaltungsvereinbarung zwischen Bund und Ländern gem. Art. 91b Abs. 1 GG finanziert.

Dieses Dokument ist lizenziert unter einer [Creative Commons Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International](#) Lizenz.



Ausgenommen von der oben genannten Lizenz sind Teile, Abbildungen und sonstiges Drittmaterial, wenn anders gekennzeichnet.

Zitationsvorschlag

KFiD – Kommission für Forschungsinformationen in Deutschland (2024):

KDSF 2.0: Flexible und qualitätsgesicherte Forschungsberichterstattung nach dem Baukastenprinzip. Grundsatzpapier der KFiD, Berlin.

<https://doi.org/10.58010/kfid:gp1:2024>

Inhaltsverzeichnis

Inhaltsverzeichnis	3
Zusammenfassung	4
Einleitung	5
1. Ausgangslage: Der KDSF als Aggregatdatenstandard	6
I. Grenzen der Vermittelbarkeit des KDSF als (reiner) Aggregatdatenstandard	7
II. Derzeitige Grenzen der Standardisierung des forschungsbezogenen Berichtswesens ..	7
2. Strukturelle Weiterentwicklung des KDSF	8
I. Wechsel der Standardisierungsebene des KDSF	8
II. Aufbereitung der Berichtsoptionen über KDSF-Module	9
III. Aufbau ergänzender Basisdaten-Services	9
Ausblick	11

Zusammenfassung

Die Kommission für Forschungsinformationen in Deutschland (KFiD) befasst sich im Rahmen ihres aktuellen Arbeitsprogramms¹ mit einer Analyse der Stärken und Schwächen des KDSF-Standards für Forschungsinformationen in Deutschland. Ziel ist es, auf dieser Grundlage wirksame Maßnahmen zur Etablierung des KDSF im Wissenschaftssystem zu entwickeln. Ein zentraler Befund ist, dass die bisherige Fokussierung auf den *einen*, möglichst universell gültigen Berichtsstandard – angesichts der Heterogenität der bestehenden Berichtsanforderungen und Informationsbedarfe verschiedener Zielgruppen – einschränkend auf das Standardisierungspotential und die Nutzbarkeit des KDSF wirkt. Die KFiD sieht demgegenüber einen deutlichen Mehrwert in einer alternativen Herangehensweise: in der Standardisierung der den Berichten zugrunde liegenden Basisdaten. Basisdaten beziehen sich auf konkrete individuelle Merkmale und Eigenschaften der für die Berichterstattung zu erfassenden Informations-

objekte, wie Personen oder Publikationen. Diese sollen einheitlich spezifiziert und definiert werden, um ihre flexible Nutzung für unterschiedliche Berichtszwecke zu ermöglichen. Damit verbindet sich eine Abkehr vom bisherigen Anspruch des KDSF, Berichte in Bezug auf die weiterzugebenden konkreten Berichtsgrößen und Aggregatdaten zu vereinheitlichen. In diesem Sinn zielt der KDSF künftig stärker auf eine Standardisierung bei der Verarbeitung der Basisdaten in Verbindung mit einer größeren Flexibilisierung bei ihrer Nutzung und Aggregation im Rahmen der Berichtslegung ab.

Mit dieser neuen Akzentuierung geht einher, dass der KDSF zukünftig um Basisdaten-Services ergänzt wird, die mit der *KDSF-Community* laufend weiterentwickelt werden. Die KFiD wird einen solchen Service beispielhaft erproben anhand einer kuratierten Liste von Forschungspreisen und Auszeichnungen.

¹ KFiD (2023): Arbeitsprogramm der Kommission für Forschungsinformationen in Deutschland: 2022 – 2024, Berlin: KFiD, <https://doi.org/10.58010/kfid:ap1:2023>

Einleitung

Um eine Grundlage für eine aktivere und zielgruppenorientierte Kommunikation, Beratung und Information zum KDSF und zum Thema Forschungsinformationsmanagement zu schaffen setzt die KFiD einen Schwerpunkt auf die strukturelle Weiterentwicklung des KDSF. Die Mitglieder der KFiD eint die Überzeugung, dass die gezielte Weiterentwicklung die Passfähigkeit des Standards mit den tatsächlichen Anforderungen des Forschungsinformationsmanagements wesentlich erhöhen und damit seine Nutzungsmöglichkeiten erweitern kann. Eine Weiterentwicklung entlang der Informationsbedarfe und Anforderungen der Nutzer:innen stellt somit die Basis für eine nachhaltige, erfolgreiche und überzeugende Kommunikation zum KDSF-Standard dar.

Vor diesem Hintergrund hat die KFiD die **strukturelle Weiterentwicklung des KDSF von einem Berichts- bzw. Aggregatdatenstandard**

hin zu einem Basisdatenstandard beschlossen, welche in dem vorliegenden Papier erläutert wird. In Ergänzung dazu fassen die auf dem Basisdatenmodell aufbauenden, sogenannten **KDSF-Module** den Standard entlang thematischer Schwerpunkte zusammen, um einen inhaltlichen und schrittweisen Zugang zum KDSF-Standard zu ermöglichen.

Das KDSF-Basisdatenmodell baut auf den Datendefinitionen auf, die mit Blick auf das bisherige KDSF-Aggregatdatenmodell (Version 1.3) entwickelt wurden, und ergänzt diese um weitere Hinweise für die Datenverarbeitung und -aggregation.

Das vorliegende Grundsatzpapier umfasst eine Analyse der Ausgangslage in Bezug auf die gegenwärtigen Stärken und Schwächen des KDSF, dokumentiert Maßnahmen der strukturellen Weiterentwicklung des KDSF zur Erhöhung des Standardisierungspotentials des KDSF und schließt mit einem Ausblick.

Motivation: Gründe für die strukturelle Überarbeitung des KDSF

Die KFiD erachtet es als notwendig, über eine strukturelle Weiterentwicklung des KDSF

- das Potential des Standards und seiner **Inhalte niedrigschwelliger zu demonstrieren** und zielgruppenadäquat zu vermitteln,
- die **Anforderungen des Standards an die Datenhaltung und Datenverarbeitung** in institutionellen Prozessen und mit den an Institutionen vorhandenen Systemen transparent und verständlich aufzuzeigen,
- den KDSF **näher an bestehende Informationsbedarfe und Nutzungsanforderungen** der datenabfragenden Akteure zu bewegen sowie die vielfältigen bestehenden Anforderungen zusammenzuführen und dabei
- das **Standardisierungspotential durch Abgleiche** mit bestehenden Berichtspflichten und Referenznutzungen transparent zu illustrieren und zu erhöhen.

1. Ausgangslage: Der KDSF als Aggregatdatenstandard

Im Jahr 2016 veröffentlichte der Wissenschaftsrat die Spezifikation des KDSF – ein **Set von Definitionen für Aggregatdaten**, die sich für die Berichtslegung über Forschung heranziehen lassen.² Der Wissenschaftsrat empfahl sowohl Hochschulen und außeruniversitären Forschungseinrichtungen als auch datenabfragenden Stellen (wie z. B. Ministerien oder Forschungsförderern) den KDSF als Standard für ihre Berichtslegung und Berichtsabfragen einzusetzen.

Vier Jahre später im Jahr 2020 analysierte der Wissenschaftsrat erstmals den Stand der Implementation und die Nutzung des KDSF im Wissenschaftssystem.³ Auf Grundlage der Evaluationsergebnisse empfahl er – neben der Einrichtung eines für den KDSF zukünftig verantwortlichen Gremiums (d. h. der Einrichtung der KFiD) – auch die Prüfung und Initiierung tiefgreifender Weiterentwicklungsprozesse nach einer Phase der Konsolidierung des

KDSF. Dabei verwies er insbesondere auf das große, bisher jedoch weitgehend ungenutzte Potenzial des KDSF, das in der einrichtungsinternen Verknüpfung verschiedener Datenbestände und dem Einsatz von Identifikatoren liege.

Vor diesem Hintergrund hat sich die KFiD in ihrer ersten Amtsphase (2022 – 2024) mit einer Analyse der Gründe für die geringe Verbreitung des KDSF bei datenbereitstellenden und datenabfragenden Akteuren befasst. Aus Sicht der KFiD liegen die Grenzen der Vermittelbarkeit des KDSF sowie der zielgruppenübergreifenden Standardisierung von Berichten und Datenabfragen in erster Linie in der gegenwärtigen Struktur des KDSF als Aggregatdatenstandard.⁴ Ein stärkerer Fokus auf die Basisdaten und die institutionelle Forschungsinformationsverarbeitung erscheint der KFiD zielführend für die zukünftige Etablierung des KDSF-Standards im Wissenschaftssystem.

Begriffsbestimmung: Anspruch und Eigenschaften KDSF

Der KDSF fokussiert bisher vor allem auf Definitionen für **Aggregatdaten**, d. h. gruppierte Daten in Summen- oder Listenform, welche in der Spezifikationstabelle des KDSF abgebildet werden. Die Aggregatdaten sind nach Bereichen (Beschäftigte, Drittmittel, Publikationen etc.) gegliedert. Ausdifferenzierungen der Aggregatdaten sind verschiedene Eigenschaften, die je nach Informationsbedarf flexibel kombiniert werden können. Ein Beispiel ist die Anzahl von Personen (etwa von Beschäftigten oder Promovierten) nach Geschlecht (Ausdifferenzierung) und Fach oder Forschungsfeld (Aggregationsniveau).

Zur Erzeugung der Aggregatdaten ist eine einrichtungsinterne Verarbeitung von **Basisdaten** notwendig. Diese umfassen Informationen, die in den meisten Einrichtungen bereits für datenunterstützte Berichte verarbeitet werden. Bezogen auf das Personal sind dies etwa Angaben zu Vertragsumfang und -laufzeit; bezogen auf Drittmittel umfassen sie Angaben zum Einwerbedatum und Namen des Mittelgebers.

Basisdaten lassen sich mit weiteren **Attributen** versehen. Arbeitsverträge von Personen etwa können mit Hilfe von Angaben zur Laufzeit mit dem Attribut „befristet/unbefristet beschäftigt“ versehen werden, Drittmittel erfahren eine Zuordnung zu den Kategorien einer Geber-Klassifikation (Bund, DFG, EU etc.). Statt von Attributen wird verschiedentlich auch von Metadaten oder von Merkmalen gesprochen.

² Wissenschaftsrat (2016): Empfehlungen zur Spezifikation des Kerndatensatz Forschung, Köln, Dr. 5066-16, https://www.wissenschaftsrat.de/download/archiv/5066-16.pdf?__blob=publicationFile&v=1

³ Wissenschaftsrat (2020): Stellungnahme zur Einführung des Kerndatensatz Forschung, Köln, Dr. 8652-20, https://www.wissenschaftsrat.de/download/2020/8652-20.pdf?__blob=publicationFile&v=5

⁴ Siehe <https://doi.org/10.58010/kdsf:spez:1.3:2022> bzw. https://kerndatensatz-forschung.de/version1/spez_versionen/Spezifikationstabelle_KDSF_v1_3.html

I. Grenzen der Vermittelbarkeit des KDSF als (reiner) Aggregatdatenstandard

Der KDSF ist bisher maßgeblich ein konfigurierbarer Aggregatdatenstandard. In der bisher empfohlenen Anwendung des KDSF lassen sich z. B. die Kerndaten „Vollzeitäquivalente von Beschäftigten“ und „Anzahl Beschäftigter“ jeweils mit einer oder mehreren der zugeordneten Ausdifferenzierungen (Geschlecht, Finanzierungsform, Staatsangehörigkeit etc.) sowie einem Aggregationsniveau (Organisationseinheiten, Fächer, Forschungsfelder etc.) verknüpfen.

Daraus können verschiedene Anwendungsmöglichkeiten entstehen. Diese sind jedoch aus der derzeitigen Dokumentation und Darstellung des KDSF nicht unmittelbar ersichtlich. Der KDSF umfasst aktuell keinerlei (beispielhafte) Einordnung, für welche Fragen oder in welchen Kontexten die vielfältigen Verknüpfungen zur Anwendung kommen könnten. Eine Erschließung der Nutzbarkeit sowie eine Relevanzbewertung des KDSF in seiner gegenwärtigen Form sind somit voraussetzungsvoll.

Ferner lassen sich aus dem KDSF in seiner aktuellen Darstellung nur indirekt die Anforderungen an die zugrundeliegende Datenhaltung

und -verarbeitung ableiten. Es fehlt eine geeignete und verbindliche Übersicht dazu, welche Informationen in welcher Form in Hochschulen und Forschungseinrichtungen erfasst und verarbeitet werden sollten, um eine KDSF-konforme Berichterstattung zu ermöglichen. Die bisherige Aufbereitung des Standards schafft hohe Anforderungen für datenbereitstellende Akteure beim Aufbau KDSF-konformer institutioneller Forschungsinformationssysteme und Prozesse.

Damit ist die Anwendung des KDSF durch Nutzer:innen – sowohl in der Datenhaltung als auch in der Datenabfrage – mit hohen Hürden verbunden. Eine niedrigschwelligere und akteursadäquatere Darstellung des KDSF (z. B. für Hochschulen) umfasst aus Sicht der KFiD daher angepasste Zugänge, orientiert an den Informationsbedarfen von datenbereitstellenden und datenabfragenden Akteur:innen, die umfangreich und heterogen sind. Sie reichen von Wissenschaftsministerien über Forschungsförderer bis hin zu Leitungen von Hochschulen und Forschungseinrichtungen.

II. Derzeitige Grenzen der Standardisierung des forschungsbezogenen Berichtswesens

Über den KDSF wird der Anspruch verfolgt, bestehende Berichtslegungen und Übermittlungen von (aggregierten) Forschungsinformationen zu standardisieren. Es werden keine neuen Berichtsanlässe geschaffen. Die Grenzen der Harmonisierung werden jedoch dort erreicht, wo gegenläufige Vorgaben und Informationsinteressen sich nicht vereinen lassen. Dies lässt sich beispielsweise bei ausgewählten *Kategorisierungen* beobachten. So ist etwa der aktuelle Ansatz des KDSF, der Drittmittelgeber nach „Bund“, „Bundesländer“, „EU“,

„DFG“, „Gewerbliche Wirtschaft und sonstige Bereiche“ unterscheidet (und damit den Bedarf der amtlichen Hochschulstatistik bedient), zu grob für die Anforderungen, die beim Monitoring des Pakts für Forschung und Innovation (PFI) für die außeruniversitäre Forschung bestehen – wo etwa ein konkretes Interesse an Zahlen zu Drittmitteln aus Wirtschaft und Industrie auf der Agenda steht.⁵ Mit dem Wechsel hin zu einem Basisdatenmodell soll dieses Defizit ausgeglichen werden (s. u.).

⁵ Vorgaben des Indikatorenkatalogs der Gemeinsamen Wissenschaftskonferenz (GWK) als Grundlage für die Berichtslegung im Rahmen des Monitorings des Pakts für Forschung und Innovation (siehe unter https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/PFI-Monitoring-Bericht_2023_Bd._II.pdf).

2. Strukturelle Weiterentwicklung des KDSF

Weiterentwicklungsansatz für den KDSF

Die KFiD identifiziert drei Handlungsdesiderate zur Stärkung des Standardisierungspotentials des KDSF:

- I. Fokus der Standardisierung auf die den Berichten zugrunde liegenden **Basisdaten**
- II. Aufbereitung der Basisdaten über **nach Themen strukturierte KDSF-Module** (wie Beschäftigte, Promotionen, Projekte etc.) sowie **Abgleiche** mit bekannten Berichtslegungen und Informationsabfragen
- III. Erweiterung des Angebots von Klassifikationsschemata und **Bereitstellung von kuratierten Metadaten**

I. Wechsel der Standardisierungsebene des KDSF

Die KFiD beschließt einen **Wechsel in der Ebene der Standardisierung von einem Aggregatdatenmodell hin zu einem Basisdatenmodell**, um eine eindeutige Referenz für datenbereitstellende Einrichtungen und eine verbesserte Grundlage für die Kommunikation, Beratung und Information der KFiD zum KDSF zu bieten. Dieses Basisdatenmodell wird neben dem Gebot der Datensparsamkeit die Vorgaben und Definitionen des bestehenden KDSF-Berichtstandards berücksichtigen. Auf dieser Grundlage kann der KDSF perspektivisch seine größtmögliche Standardisierungswirkung entfalten. Das Basisdatenmodell wird ferner angereichert um Standards für die Datenhal-

tung und -verknüpfung sowie die Nutzung von (persistenten) Identifikatoren (siehe [Abbildung 1](#)).

Das Basisdatenmodell bündelt somit künftig verschiedene Anforderungen an die Berichterstattung und bildet eine Grundlage für die Konfiguration von institutionellen Forschungsinformationssystemen und Prozessen. Für die Umsetzung und Darstellung dieses Basisdatenmodells werden auch Möglichkeiten der Orientierung an bestehenden Technologiestacks (wie RDF(S) mit OWL-Elementen oder Schema.org) geprüft, um eine möglichst einfache Nutzung im Rahmen vorhandener Systeme und Prozesse zu unterstützen.

→ **Basisdaten verbleiben stets bei den datenbereitstellenden Einrichtungen** und werden nicht an die datenabfragenden Akteure weitergegeben. Ihre systematische und an transparenten Vorgaben orientierte Verarbeitung stellt eine wichtige Voraussetzung für die Stärkung der Datenqualität der darauf basierenden Aggregate und Berichte dar.

II. Aufbereitung der Berichtsoptionen über KDSF-Module

Darüber hinaus initiiert die KFiD die **Entwicklung** von **KDSF-Modulen**, anhand derer verschiedene und vielfältige (beispielhafte sowie an realen Berichtsabfragen orientierte) KDSF-konforme Aggregate und Ausgaben dargestellt werden.

KDSF-Module fassen das Basisdatenmodell entlang von Themen (z. B. Beschäftigte, Promotionen, Projekte etc.) zusammen und sind die Grundlage für Abgleiche mit bestehenden Anwendungsfällen. Ein Modul bezieht sich jeweils auf ausgewählte Erfassungsobjekte der Basisdaten (z. B. Förderungen, Qualifizierungsverfahren, Publikationen) und ihre Attribute. Ein Beispiel ist etwa das Modul „Beschäftigte“, welches darstellt, wie und welche Informationen über Personen, ihre

Arbeitsverträge, ihre Verortung in der Organisationsstruktur oder ihre Rollen in Drittmittelprojekten (z. B. als Koordinator/-in) KDSF-konform verarbeitet werden und wie diese sich in der Praxis anwenden lassen. Das Modul „Projekte“ als weiteres Beispiel zeigt die KDSF-konforme Erfassung von Forschungsprojekten sowie ihrer Förderungs- und ggf. Kooperationskontexte auf.

Die KDSF-Module sind der Ausgangspunkt für die Dokumentation von Anwendungshinweisen zur Heranziehung des KDSF in bestimmten Nutzungskontexten (z. B. PFI-Monitoring, Erhebungen des Wissenschaftsrats im Rahmen der Exzellenzstrategie, die amtliche Hochschulstatistik oder die bisherigen KDSF-Aggregatdaten).

→ Eine wichtige **Änderung gegenüber dem bisherigen Modell des einen Berichtsstandards** resultiert dabei aus dem Umstand, dass KDSF-nutzende Einrichtungen sich auf solche Module (und Regeln für die zugrundeliegenden Basisdaten) fokussieren können, die für ihre Berichtszwecke geeignet sind. Auf diese Weise eröffnen sich Möglichkeiten der schrittweisen sowie selektiven Nutzung genau jener Module und Basisdaten, die für das eigene Berichtswesen als relevant eingeschätzt werden.

III. Aufbau ergänzender Basisdaten-Services

Zur Unterstützung der KDSF-konformen Datenverarbeitung in Hochschulen und außeruniversitären Forschungseinrichtungen konzipiert die KFiD verschiedene Angebote und Services. Ein Fokus besteht dabei in der Entwicklung von Klassifikationen

für den KDSF (wie etwa der Forschungsfeldklassifikation) sowie in der gebündelten Dokumentation unterschiedlicher, für verschiedene Berichtskontexte relevanter Klassifikationen (z. B. unterschiedliche Kategorisierungsschemata für Drittmittelgeber).

→ Ein besonderer Fokus wird auf den **Aufbau von Stammdaten-Services** gesetzt, deren Mehrwert darin besteht, dass relevante Basisdaten inklusive Merkmale (Metadaten) – z. B. über Listen von Forschungspreisen und Auszeichnungen – durch die KFiD kuratiert und qualitätsgesichert sowie zur Nachnutzung öffentlich zur Verfügung gestellt werden, um Einrichtungen beim Aufbau einer qualitätsgesicherten Informationsverarbeitung zu Forschungspreisen und Auszeichnungen zu unterstützen.

Abbildung 1 gibt einen Überblick zu den Komponenten der strukturellen Weiterentwicklung des KDSF-Standards.

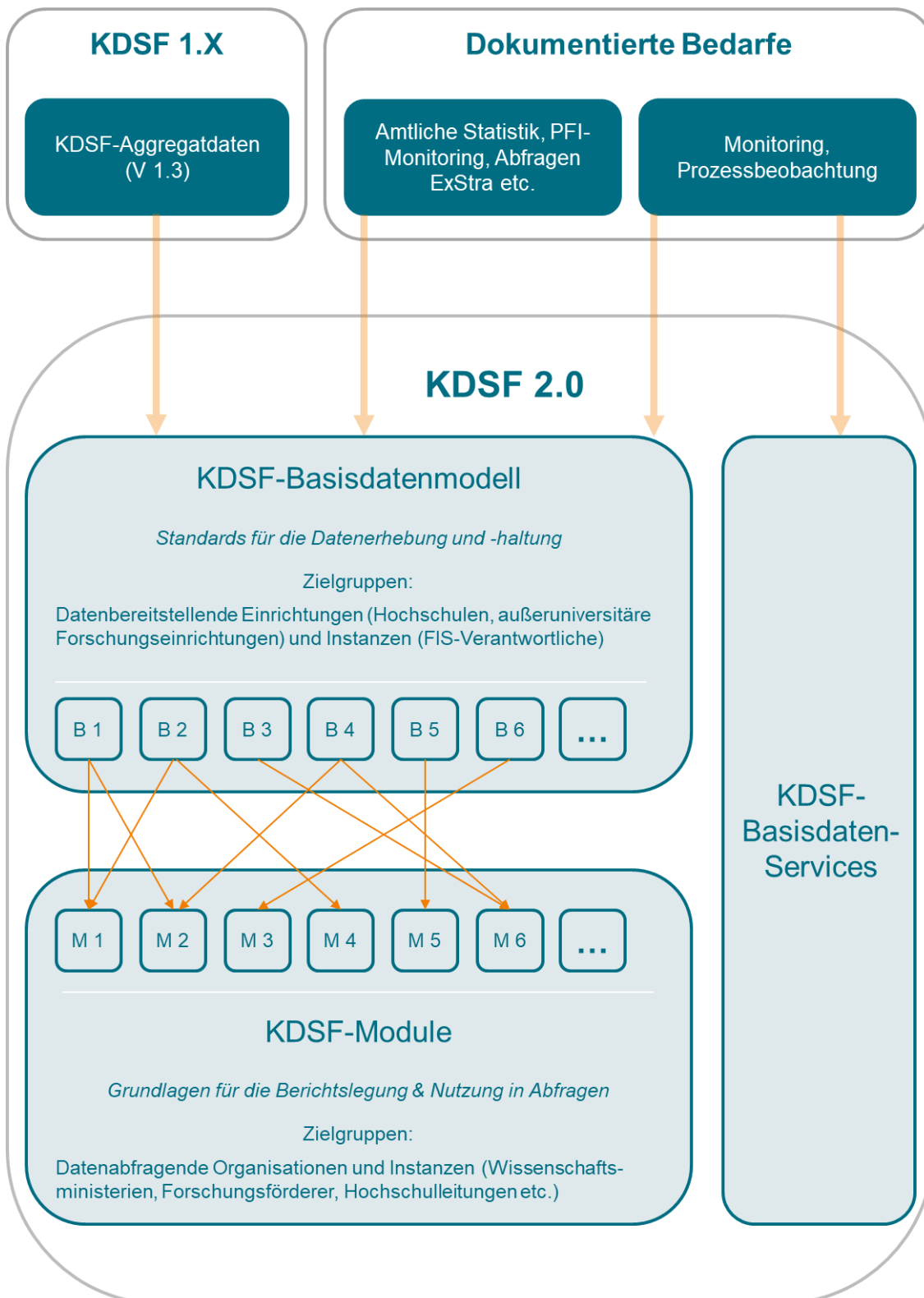


Abbildung 1: Komponenten des weiterentwickelten KDSF [B: Basisdatenobjekt; M: Modul]

Ausblick

Als ein Schwerpunkt der laufenden ersten Amtszeit (bis Ende 2024) liegt der Fokus der Kommission für Forschungsinformationen in Deutschland gegenwärtig auf der schrittweisen Erarbeitung des KDSF-Basisdatenmodells und von KDSF-Modulen inklusive Verarbeitungshinweisen für bestehende Berichte sowie auf dem Aufbau ergänzender Basisdaten-Services. Die Angebote werden parallel mit der *Community* validiert und sollen mit dem Übergang in die zweite Amtszeit der KFiD (2025 – 2027) öffentlich bereitgestellt werden.

Damit verbunden ist das Ziel, den KDSF weiter als Standard zu etablieren, der für alle Anwender:innen eine Vereinfachung und mehr Transparenz bei der Nutzung von Forschungsinformationen schafft. Die Modellierung von Basisdaten und die Spezifizierung von Aggregierungsregeln sowie die Darstellung des Nutzungspotentials des KDSF in konkreten Berichtskontexten sollen verschiedenen Zielgruppen und Nutzer:innen den Zugang zum KDSF erleichtern und einen Beitrag zur Qualitäts-

sicherung im Umgang mit Forschungsinformationen leisten.

Ein Standard lebt von seiner Verbreitung und Anwendung. Der zukünftige KDSF soll deshalb so gut wie möglich die Erwartungen und Rahmenbedingungen seiner Nutzer:innen berücksichtigen. Die KFiD sucht vor diesem Hintergrund den kontinuierlichen Austausch mit der *Community*. Sie wird eine Einladung zur gemeinsamen Weiterentwicklung und Gestaltung des KDSF aussprechen – etwa im Hinblick auf die ergänzenden Basisdaten-Services wie den Stammdaten-Service Forschungspreise und Auszeichnungen, dessen Qualität und Erfolg maßgeblich von der aktiven Partizipation der Nutzer:innen abhängen werden.

Aus Sicht der KFiD schaffen diese Maßnahmen die Grundlage für eine Stärkung des freiwilligen KDSF-Standards insgesamt, eine Erhöhung der Akzeptanz des Instruments unter den Nutzer:innen sowie die volle Entfaltung seines Standardisierungspotentials.