

FAIR-Prinzipien

COMPASS Arbeitspaket 5_Forschungskompatibilität & Interoperabilität

Die FAIR-Prinzipien (Findable, Accessible, Interoperable, Reusable) wurden im März 2016 in der Zeitschrift Scientific Data von einem Konsortium aus Wissenschaftlern und Organisationen beschrieben und 2018 ergänzt und aktualisiert. Die Hauptidee der FAIR-Prinzipien ist es, die maschinelle Verwertbarkeit von Daten in den Mittelpunkt zu stellen, da das Datenvolumen und die Komplexität der Daten immer weiter zunehmen.

Version 1.0_ 24.08.2021

FAIR-Prinzipien

Wie wir die Findbarkeit, Verfügbarkeit, Interoperabilität und die Wiederverwendbarkeit von Forschungsdaten schon beim Design der Erfassungssapps verbessern können

Deliverable D5.2 im Rahmen des NUM Compass-Projektes

Autoren/Autorinnen:

- Michael Rusongoza Muzoora
- Johannes Oehm
- Rasim Atakan Poyraz
- Sarah Riepenhausen
- Ulrich Sax
- Marco Schaarschmidt

Inhaltsverzeichnis

Autoren/Autorinnen:	2
Was ist FAIR?	4
Wofür steht FAIR	4
Findability 1: (Meta-)Daten werden global eindeutige und dauerhafte Bezeichner zugewiesen	4
Findability 2: Daten werden mit umfangreichen Metadaten beschrieben	4
Findability 3: Metadaten enthalten eindeutig und explizit den Identifikator der Daten, die sie beschreiben	4
Findability 4: (Meta-)Daten sind in einer durchsuchbaren Ressource registriert oder indiziert	5
Accessibility 1: (Meta-)Daten sind anhand ihrer Kennung über ein standardisiertes Kommunikationsprotokoll abrufbar	5
Accessibility 1.1: Das Protokoll ist offen, frei und universell implementierbar	5
Accessibility 1.2: Das Protokoll ermöglicht eine Authentifizierung und Autorisierung, falls erforderlich	5
Accessibility 2: Metadaten sollten auch dann noch zugänglich sein, wenn die Daten nicht mehr verfügbar sind	5
Interoperability 1: (Meta)daten verwenden eine formale, zugängliche, gemeinsame und breit anwendbare Sprache zur Wissensrepräsentation	5
Interoperability 2: (Meta)daten verwenden Vokabulare, die den FAIR-Prinzipien folgen	6
Interoperability 3: (Meta)daten enthalten qualifizierte Verweise auf andere (Meta)daten	6
Reusability 1: (Meta-)Daten sind reichhaltig beschrieben mit einer Vielzahl von genauen und relevanten Attributen	7
Reusability 1.1: (Meta)daten werden mit einer klaren und zugänglichen Datennutzungslizenz freigegeben	7
Reusability 1.2: (Meta)daten sind mit einer detaillierten Provenienz verbunden	8
Reusability 1.3: (Meta)daten erfüllen domänenrelevante Community-Standards	8
Warum ist was wichtig?	9
Aktuelle Herausforderungen	9
Übersicht über relevante existierende – FAIR Spezifikationen für Gesundheitsapps	9
Übersicht über relevante existierende – medizinische Informationssysteme (Forschung und Krankenversorgung)	10
Existierende Beispiele der Umsetzung	10
Referenzen	10
Anhang	13

Was ist FAIR?

Die FAIR-Prinzipien (Findable, Accessible, Interoperable, Reusable) wurden im März 2016 in der Zeitschrift Scientific Data von einem Konsortium aus Wissenschaftlern und Organisationen beschrieben[1] und 2018 ergänzt und aktualisiert[2,3]. Die Hauptidee der FAIR-Prinzipien ist es, die maschinelle Verwertbarkeit von Daten in den Mittelpunkt zu stellen, da das Datenvolumen und die Komplexität der Daten immer weiter zunehmen. Um mit dieser Menge und Komplexität umgehen zu können, sollten Daten sorgfältig nach FAIR-Prinzipien betrachtet werden, damit sie aussagekräftig und leichter zu handhaben sind.

Wofür steht FAIR

Findable: Laut FAIR Guiding Principles ist der erste Schritt bei der (Wieder-)verwendung von Daten, sie zu finden. Metadaten und Daten sollten sowohl für Menschen als auch für Maschinen leicht zu finden sein.

Findability 1: (Meta-)Daten werden global eindeutige und dauerhafte Bezeichner zugewiesen

F1: Das wohl wichtigste Prinzip F1 fokussiert auf global eindeutige und persistente Bezeichner. Einer der meistgenutzten medizinischen Datenstandards HL7 FHIR unterstützt dieses Argument, da die kanonische URL, also die Ursprungsquelle der Terminologien, und der Identifier entsprechend abgebildet werden müssen, um FHIR-Profilen erstellen zu können.

→ **Take-Away:** Beseitigen Sie Mehrdeutigkeiten in der Bedeutung Ihrer veröffentlichten Daten, indem Sie jedem Element der Metadaten und jedem Konzept/Messung in Ihrem Datensatz einen eindeutigen Bezeichner zuweisen.

Chronic lung disease present
Example condition that is confirmed present

```
{
  "resourceType": "Condition",
  "id": "4f7ba670-119e-4964-9cd5-c191f1c00ce1",
  "meta": {
    "profile": [
      "https://www.netzwerk-universitaetsmedizin.de/fhir/StructureDefini
    ]
  }
}
```

Abbildung 1: Beispiel für den Fakt „eine chronische Lungenerkrankung vorhanden“

Findability 2: Daten werden mit umfangreichen Metadaten beschrieben

Die Beschreibung von Daten mit reichhaltigen Metadaten ist ebenfalls entscheidend. Diese ermöglichen es, routinemäßige, oft langwierige Sortier- und Priorisierungsaufgaben automatisiert von einem Computer übernehmen zu lassen. Der Forscher kann seine Aufmerksamkeit auf andere Dinge konzentrieren
→ Jemand sollte in der Lage sein, Daten auf der Grundlage der durch Ihre Metadaten bereitgestellten Informationen zu finden, auch ohne den Bezeichner der Daten.

Findability 3: Metadaten enthalten eindeutig und explizit den Identifikator der Daten, die sie beschreiben

Hierzu ist eine eindeutige Nummer vorzusehen (Digital Object Identifier), der auch über einen längeren Zeitraum im entsprechenden Verzeichnis referenziert bzw. aufgelöst werden kann. Beispiel: Die GECCO-Profilen enthalten Identifier-Datenelemente, die Metadaten und Daten verbinden: (a) Informationsmodell[4] und (b) im MDM Portal[5].

Findability 4: (Meta-)Daten sind in einer durchsuchbaren Ressource registriert oder indiziert

Wenn die Verfügbarkeit einer digitalen Ressource wie eines Datensatzes, eines Dienstes oder eines Repositories nicht bekannt ist, kann niemand (auch keine Maschine) sie entdecken. GECCO-Datensätze sind auf der NUM-Plattform registriert.

Accessibility 1: (Meta-)Daten sind anhand ihrer Kennung über ein standardisiertes Kommunikationsprotokoll abrufbar

Sobald die gewünschten Daten gefunden sind, muss der Zugriff auf diese klar sein, gegebenenfalls einschließlich Authentifizierung und Autorisierung.

Accessibility 1.1: Das Protokoll ist offen, frei und universell implementierbar

Jede Person mit einem Computer und Internetzugang, hat Zugriff auf NUM-COMPASS, da es ein open-source Projekt ist, dessen Repository auf Github (<https://github.com/NUMde/compass-numapp>) erreichbar ist (HTTPS Protokoll).

Die Metadaten des GECCO Referenzdatensatz [6,7] können auf der Webseite cocos.team [8] für eine detaillierte Erläuterung abgerufen werden.

Accessibility 1.2: Das Protokoll ermöglicht eine Authentifizierung und Autorisierung, falls erforderlich

Die Metadaten sind frei verfügbar (open-source) und können im Simplifier [7] ohne Autorisierung abgerufen werden.

Accessibility 2: Metadaten sollten auch dann noch zugänglich sein, wenn die Daten nicht mehr verfügbar sind

Die Metadaten (=GECCO FHIR Profiles) sind im Simplifier[7], Art-Decor [6] und MDM [5] online verfügbar. Die CODEX Central Data Platform, mit der COMPASS verbunden wird, speichert die Instanzdaten.

Interoperability 1: (Meta)daten verwenden eine formale, zugängliche, gemeinsame und breit anwendbare Sprache zur Wissensrepräsentation

- Menschen sollten in der Lage sein, ihre Daten untereinander auszutauschen und zu interpretieren (also möglichst keine toten Sprachen verwenden).
- Dies gilt aber auch für Computer, d. h. Daten, die für Maschinen lesbar sein sollten, ohne dass spezielle oder Ad-hoc-Algorithmen, Übersetzer oder Mappings erforderlich sind. Interoperabilität bedeutet typischerweise, dass jedes Computersystem zumindest Kenntnis von den Datenaustauschformaten des anderen Systems hat.
- Der FHIR-Standard und die FHIR-Profile definieren das Datenmodell, d.h. GECCO fügt sich in existierende Datenmodelle/Frameworks ein [7].
- GECCO verwendet gängige kontrollierte Vokabulare, Ontologien, Thesauri (LOINC, SNOMED, ICD-10 GM)[9,10,11] und ein gutes Datenmodell (ein wohldefinierter Rahmen zur Beschreibung und Strukturierung von (Meta-)Daten).

Profile - Chronic Lung Diseases

Canonical: <https://www.netzwerk-universitaetsmedizin.de/fhir/StructureDefinition/chronic-lung-diseases>

Snapshot

Element	Cardinality	Type	Binding
Condition		Condition	
extension	0..*	Extension	
identifier	Σ 0..*	Identifier	
clinicalStatus	S Σ ?! 0..1	CodeableConcept	Binding
verificationStatus	S Σ ?! 0..1	CodeableConcept	Binding
category	S 1..*	CodeableConcept	Binding
severity	0..1	CodeableConcept	Binding
code	S Σ 1..1	CodeableConcept	
coding	S Σ 1..*	Coding	
icd10-gm	S Σ 0..1	Coding	Binding
extension	0..*	Extension	
system	S Σ 1..1	uri	
version	S Σ 1..1	string	
code	S Σ 1..1	code	
display	Σ 0..1	string	
userSelected	Σ 0..1	boolean	
alpha-id	S Σ 0..1	Coding	Binding
sct	S Σ 0..1	Coding	Binding
orphanet	S Σ 0..1	Coding	
text	Σ 0..1	string	

Element Id ↔

Condition.co

Short description
Code defined by a terminology system

Definition
A reference to a code defined by a terminology system.

Requirements
Allows for alternative encodings within a code system, and translations to other code systems.

Comments
Codes may be defined very casually in enumerations. or

Beispiel der möglichen Kodierung von Chronologischen Lungenerkrankungen durch verschiedene international anerkannte medizinischen Standards, wie ICD-10 GM [7,11].

Interoperability 2: (Meta)data verwenden Vokabulare, die den FAIR-Prinzipien folgen

Eine Sammlung von FAIR Data Points (FDPs) zielt darauf ab, dieses Interoperabilitätsproblem zu lösen, indem sie Dateneigentümern die Möglichkeit gibt, ihre Daten auf FAIR-Basis zu teilen und somit die Auffindbarkeit, Zugänglichkeit, Interoperabilität und Wiederverwendbarkeit zu fördern.

Datenstandards spielen eine entscheidende Rolle, wenn es darum geht, die Daten an die FAIR-Prinzipien anzupassen. Medizinische Vokabularen wie SNOMED

CT[10], LOINC[9,12] und ICD-10 GM [11] helfen, dass Daten verständlich, auffindbar und durchsuchbar sind. In GECCO werden die Daten nach den Terminologien SNOMED CT, LOINC, und ICD-10 GM kodiert.

Es stellt sich die Frage aber, "Wie FAIR ist SNOMED-CT eigentlich?" In Deutschland wird über die Medizin Informatik Initiative (MII) die Lizenz von SNOMED CT für die Krankenversorgung und der Forschung zur Verfügung gestellt. Siehe "Reusability 1.1".

Interoperability 3: (Meta)data enthalten qualifizierte Verweise auf andere (Meta)data

Die GECCO Profile, die vom MII Kerndatensatz erben, verlinken bereits auf die Terminologien.

Zu sehen In der MII-Referenz [13, 14] (PACKAGE de.medizininformatik initiative.kerndatensatz.meta 1.0.3 /sind Links zu SNOMED CT, FHIR, LOINC, ICD-10/11 enthalten.

Die MII setzt die Abstimmung der gesammelten Kerndatensätze mit Verweisen auf weitere Metadaten wie folgt um:

“Nach § 301 und § 295 SGB V ist die Kodierung bei der Abrechnung im deutschen Gesundheitswesen mit der Internationalen statistischen Klassifikation (ICD-10-GM) gesetzlich vorgeschrieben.

Da die ICD-10-GM aber für viele Fragestellungen nicht aussagekräftig genug ist (siehe Beschreibung Modul), soll im Kerndatensatz die Kodierung auch zusätzlich mit anderen Terminologien und Klassifikationen (wie Alpha-ID, SNOMED CT und ORPHA-Kennnummern) ermöglicht werden.” - (Sozialgesetzbuch SGB V - Gesetzliche Krankenversicherung - Stand: Zuletzt geändert durch Art. 4 G v. 12.5.2021 I 1087) [15]
Guide für internationale Terminologien (IG MII KDS Modul Diagnose): [16]

Reusability 1: (Meta-)Daten sind reichhaltig beschrieben mit einer Vielzahl von genauen und relevanten Attributen

Durch die Umsetzung des Late Mapping Scheme (LMS) werden die gegebenenfalls nicht GECCO konformen Daten nachhaltig wiederverwendbar.

Allerdings ist R1 nicht komplett erfüllt, weil in den FHIR-Profilen keine Metadaten dokumentiert sind, die den Kontext, unter dem die Daten erzeugt wurden, umfassend beschreiben.

Reusability 1.1: (Meta)daten werden mit einer klaren und zugänglichen Datennutzungslizenz freigegeben

Folgende Systeme, Plattformen und Tools benötigen lizenzierungen zur Nutzung:

- **SNOMED CT** [17] Die UMLS ist eine globale Lizenz und kann für den Zugriff auf SNOMED CT verwendet werden. Wenn Sie planen, SNOMED CT in einem Mitgliedsgebiet zu verwenden, einzusetzen oder zu vertreiben, kann die Lizenz von diesem Mitglied erworben werden.

“Seit dem 01.01.2021 ist Deutschland Mitglied von SNOMED International. Die medizinische Terminologie ermöglicht es, medizinische Begriffe in verschiedenen Computersystemen international eindeutig abzubilden. Damit schafft SNOMED CT die Voraussetzung für den elektronischen Austausch von Gesundheitsdaten. Seit dem 01.01.2021 ist das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) als National Release Center für die Bereitstellung und Weiterentwicklung von SNOMED CT in Deutschland zuständig. Der Erhalt einer Sublizenz für SNOMED CT ist für Nutzer in Deutschland kostenfrei.”[18]

- **LOINC** [9] Die LOINC-Lizenz erlaubt sowohl die kommerzielle als auch die nicht-kommerzielle Nutzung, mit einigen Bestimmungen zur Namensnennung und dem Verbot, LOINC-Inhalte zu ändern oder zu verwenden, um einen anderen Vokabularstandard zu schaffen.
- **CODEX+ Plattform** [19] (Stand 22.08.2020)

Use & Access Verfahren

Ein Use & Access-Committee (UAC) prüft alle Daten- und Probennutzungsanträge. Das UAC beschließt mit Zwei-Drittel-Mehrheit seiner Mitglieder.

Das UAC kann zu spezifischen Themen die Meinung externer Fachgutachtern einholen. Antragsteller, deren Daten- und/oder Probennutzungsantrag vom UAC abgelehnt wurde, können sich an die Nationale Task Force oder ein zukünftig ggf. noch entstehendes anderes zentrales Steuerungsgremium des NUM wenden (Revisionsinstanz). Dieses entscheidet final.

Es werden geeignete Sicherheitsvorkehrungen getroffen, um die Anonymität der Teilnehmer/innen und die Vertraulichkeit ihrer Daten und Bioproben bei Weitergabe für Forschungsvorhaben zu gewährleisten. (Quelle: Codex Antrag 2020) Weitere Informationen befinden sich in Anhang 1.

- Das Deutsche Zentrum für Herz- Kreislauf- Forschung e.v. (DZHK) bietet lizenzierte Datenmanagement Lösungen an, woraus SOPs und die UMLS-Annotation vorhanden sind. Die detaillierte Beschreibung der in der wissenschaftlichen Infrastruktur erfassten Daten und Biomaterialien (was wird erfasst) inklusive der dazugehörigen Metadaten (wie wird erfasst) wird im DZHK-Data Catalogue in Übereinstimmung mit den international anerkannten FAIR-Prinzipien veröffentlicht, was den Wert der DZHK-Sammlung für die wissenschaftliche Gemeinschaft fördert (<https://dzhk.de/ressourcen/data-catalogue/>) [20]. Dies ist eine deutliche Zeitersparnis für Entwickler*innen. Diese Komponenten wurden auch im NUM-COMPASS Projekt angewandt. Der DZHK Datacatalogue ist weltweit zugänglich, mit verfügbaren Inhalten und den zugehörigen SOP Links, wie zum Beispiel MDM mit den UMLS-Annotation <https://medical-data-models.org/search?query=dzhk> [21].
 - Nutzungsanträge werden einem Evaluationsprozess im Hinblick auf ihr Ziel und ihren wissenschaftlichen Nutzen unterzogen. Das neutrale Usage-Access-*Committee* übernimmt diese Aufgaben nach strikten und transparenten Kriterien gemäß der DZHK-Nutzungsordnung.
- **Apache Lizenz 2.0:** [22] Die Version 2.0 der Apache-Lizenz, die 2004 von der ASF genehmigt wurde, hilft uns, unser Ziel zu erreichen, zuverlässige und langlebige Softwareprodukte durch gemeinschaftliche, quelloffene Softwareentwicklung anzubieten.

Reusability 1.2: (Meta)daten sind mit einer detaillierten Provenienz verbunden

Die GECCO-Profile wurden von verschiedenen Universitäten und Forschungsgruppen entwickelt. Die detaillierten Beschreibungen dieser Datenelemente und Profile sind auf den jeweiligen Webseiten zu finden[5,6,7].

Reusability 1.3: (Meta)daten erfüllen domänenrelevante Community-Standards

Der GECCO-Datensatz besteht aus HL7 FHIR-Profilen, SNOMED CT, ICD-10.

Warum ist was wichtig?

- Die Nachnutzung der Daten wird durch die Beachtung der FAIR Prinzipien ermöglicht.
- Verbesserte Forschung durch Datenzugang mit umfassenden Metadatenmanagement. [23]
- Ein konsentiertes Metadatenmodell auf Metadaten Ebene ist notwendig, um die gesammelten Daten zusammenzuführen und gemeinsam verarbeiten zu können.
- Eine heterogene Datensammlung auf verschiedenen Endgeräten erschwert die Forschung. Aufgrund dessen, dass unterschiedliche Endgeräte verschiedene Formen der Daten erfassen können, kann es zu Fehleingaben der Daten kommen. Zum Beispiel wird die Sensorik bei unterschiedlichen Endgeräten unterschiedlich dargestellt, was wiederum die wissenschaftliche Bewertung durch Fehleingaben erschwert.
- Internationale Kollaborationen werden möglich, da die Metadatenmodelle sich ähneln und die Daten daher kompatibler sind. Hier werden regionale, nationale und internationale Profile berücksichtigt, die unabhängig von Standort und Gerät der Forschung zur Verfügung stehen [24].
- Seit 2017 hat die Europäische Union die Dringlichkeit von der Nutzung und Nachnutzung medizinischer Daten verstanden. "GO FAIR zielt auf die schrittweise Öffnung vorhandener Forschungsdaten an wissenschaftlichen und akademischen Einrichtungen in allen Forschungsbereichen und über Ländergrenzen hinweg - und ist damit ein Baustein zur Realisierung der European Open Science Cloud" [25]

Aktuelle Herausforderungen

- Derzeit benötigt die Umsetzung der FAIR Prinzipien einen zusätzlichen Aufwand/Expertise, der sich vor allem bei kleineren Forschungsprojekten finanziell nicht rentiert und deswegen eine untergeordnete Rolle spielt. Kurzfristige Vorteile der FAIR Prinzipien werden nicht erkannt oder umgesetzt [26].
- Die wenigsten Apps verwenden konsentierete Datenmodelle wie z.B. GECCO [7]. Diese Tendenz ist auch bei vielen Pandemie-Apps zu beobachten. Es existieren keine Richtlinien oder Implementierungshilfen, wie die FAIR Guiding Principles auf mit Apps erhobene Daten angewendet werden können.[1,27] Dies führt zu einem erhöhten Bearbeitungsaufwand die per App erhobenen Daten nachträglich FAIR zu machen, um sie anschließend in ein Datenmanagementsystem einzuspielen.
- Die FAIR Guiding Principles sind dabei kein Standard, den man implementieren kann, sondern Richtlinien, die für den jeweiligen Anwendungsfall spezifiziert werden müssen.

Übersicht über relevante existierende – FAIR Spezifikationen für Gesundheitsapps

- Das Medizinprodukte-Durchführungsgesetz (MPDG) setzt FAIR Spezifikationen für Gesundheitsapps nicht als Voraussetzung.[28]
- Daher werden die FAIR Prinzipien auch nicht großflächig umgesetzt.
- Das Infektionsschutzgesetz (IFSG) verpflichtet auch keine FAIR Prinzipien in der Umsetzung von Gesundheitsapps.[29]

Übersicht über relevante existierende – medizinische Informationssysteme (Forschung und Krankenversorgung)

- Da FAIR per se kein Standard ist, wird es auch keine per se „FAIRen“ Systeme geben können.
- Neben den datenliefernden Systemen sind auch die Schnittstellen, die Transformationsschritte und die Ladeschritte (ETL) von höchster Bedeutung.
- Somit kann die FAIRifizierung einer Forschungsinfrastruktur beschrieben werden[30,31].
- Aus dem FAIRdom Umfeld wäre FAIRdom SEEK[32,33] als prominentes Beispiel eines FAIR betreibbares System zu nennen.

Existierende Beispiele der Umsetzung

- CovApp basierend auf FHIR [34]
 - CovApp enthält einen Fragenkatalog zu möglichen Symptomen, Kontakt mit infizierten Personen, Meldungen der Corona-Warn-App oder Ergebnissen von (Schnell-)Tests. Basierend auf den Antworten werden Handlungsempfehlungen und Ansprechpartner präsentiert.
 - Der Fragebogen ist anonym und nur auf dem lokalen Endgerät ausgewertet.
 - Falls ein Arztgespräch notwendig ist, wird ein QR-Code mit den “Questionnaire Responses” generiert.
 - Dynamischer und geräteunabhängiger Zugriff im Web-Browser (Web-App): <https://covapp.charite.de>
 - **Accessibility/Reusibility:** Open Source mit MIT-Lizenz, Quellcode ist frei zugänglich auf GitHub. [35]
 - Quellcode kann angepasst werden: Farbe, Logo, Sprache kann verändert werden.
 - **Interoperability:** Zusätzlich zur XML Repräsentation der “Questionnaire Responses” ist es auch möglich eine FHIR Repräsentation zu generieren.

- MII AG Interoperabilität [36]

Ziele und Aufgaben:

- Ist eine Plattform für den Abstimmungsprozess der Konsortien im Hinblick auf die Interoperabilität zwischen den aufzubauenden Datenintegrationszentren zu schaffen
- Konkrete Schritte zur Erreichung von Interoperabilität sind zu planen.
 - z.B ein Eckpunktepapier zur Interoperabilität ist Meilenstein 4.5.3 : Konzept für Metadaten liegt vor (Verfügbarkeit/Qualität/Provenance)

“Die Metadaten zu den verfügbaren Daten sollen transparent gemacht werden (FAIRPrinzipien).” ¹

- Entsprechende Mindestanforderungen zu konsentieren: diese werden aktiv bei der verwendung der FHIR Basisprofile der HL7 Deutschland e.V., insbesondere Basisprofil DE (R4), erfüllt. [37]

- FAIRDOM-SEEK[32,33,38]

- SEEK-Plattform ist eine webbasierte Ressource für den Austausch von heterogenen wissenschaftlichen Forschungsdatensätzen, Modellen oder Simulationen, Prozessen und Forschungsergebnissen.
- Sie bewahrt Assoziationen zwischen diesen, zusammen mit Informationen über die beteiligten Personen und Organisationen.
- Die Forschung kann basierend auf dem ISA-Tools-Format in Untersuchungen, Studien und Assays (Experimente) organisiert werden.

Referenzen

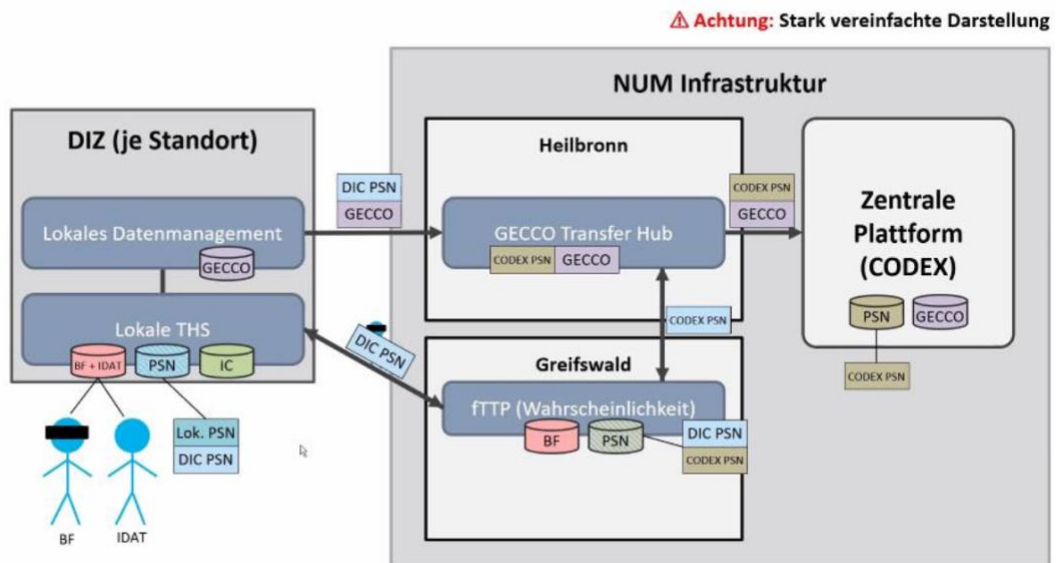
1. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016 Mar 15;3:160018.
2. Wilkinson MD, Dumontier M, Sansone S-A, Santos, Luiz Olavo Bonino da Silva, Prieto M, McQuilton P, et al. Evaluating FAIR-Compliance Through an Objective, Automated, Community-Governed Framework. *bioRxiv*. 2018;418376.
3. Wilkinson MD, Sansone S-A, Schultes E, Doorn P, Bonino da Silva Santos LO, Dumontier M. A design framework and exemplar metrics for FAIRness. *Sci Data*. 2018;5:180118.
4. Sass J, Bartschke A, Lehne M, Essenwanger A, Rinaldi E, Rudolph S, et al. The German Corona Consensus Dataset (GECCO): a standardized dataset for COVID-19 research in university medicine and beyond. *BMC Med Inform Decis Mak*. 2020 Dec;20(1):341.
5. GECCO - German Corona Consensus - Covid-19 Research-Dataset [Internet]. Available from: <https://medical-data-models.org/41855>
6. Art-Decor.org: Covid-19 Research-Dataset - Datasets [Internet]. Available from: <https://art-decor.org/art-decor/decor-datasets--covid19f->
7. Simplifier.net: Forschungsnetz Covid-19 - SIMPLIFIER.NET [Internet]. Available from: <https://simplifier.net/forschungsnetz-covid-19>
8. Bundesministerium für Gesundheit (BMG) Pressestelle. (2020). cocos – Corona Component Standards. COCOS– Corona Component Standards. <http://cocos.team/>
9. Logical Observation Identifiers Names and Codes (LOINC®) — LOINC [Internet]. [cited 2016 Nov 17]. Available from: <https://loinc.org/>
10. SNOMED CT. (n.d.). SNOMED CT. <https://www.Snomed.Org/Snomed-Ct/Use-SNOMED-CT>. Retrieved June 2021, from <https://www.snomed.org/snomed-ct/Use-SNOMED-CT>
11. Bundesinstitut für Arzneimittel und Medizinprodukte. (2021). ICD-10-GM Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, German Modification. [www.Bfarm.De. https://www.bfarm.de/DE/Kodiersysteme/Klassifikationen/ICD/ICD-10-GM/_node.html](https://www.bfarm.de/DE/Kodiersysteme/Klassifikationen/ICD/ICD-10-GM/_node.html)
12. Chen ES, Melton GB, Engelstad ME, Sarkar IN. Standardizing Clinical Document Names Using the HL7/LOINC Document Ontology and LOINC Codes. *AMIA Annu Symp Proc AMIA Symp*. 2010 Nov 13;2010:101–5.
13. Zautke, Essenwanger, Wulff, Böhnisch, Kamann, Büschges, Ammon, Kraska. (2021, May 27). Medizininformatik Initiative - Modul Diagnose - ImplementationGuide. Medizininformatik Initiative - Modul Diagnose - ImplementationGuide. <https://simplifier.net/guide/medizininformatikinitiative-moduldiagnosen-implementationguide/igmiikdsmoduldiagnose>
14. Ammon D, Bietenbeck A, Boeker M., Ganslandt T, Heitmann K, Sax, U, Schepers, J, Semler SC, Thun S, Zautke A. Der Kerndatensatz der Medizininformatik-Initiative – Interoperable Spezifikation am Beispiel der Laborbefunde mittels LOINC und FHIR. *Mdi – Forum Med Med*. 2019;(4):113–7.
15. Bundesministerium der Justiz und für Verbraucherschutz. (1988, December). SGB 5 - Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - (Artikel 1 des Gesetzes v. 20. Dezember 1988, BGBl. I S. 2477). *Gesetze Im Internet*. https://www.gesetze-im-internet.de/sgb_5/BJNR024820988.html
16. Medizininformatik Initiative. (2020, May 1). Medizininformatik Initiative - Modul Diagnose - ImplementationGuide. Medizininformatik Initiative - Modul Diagnose - ImplementationGuide. <https://simplifier.net/guide/MedizininformatikInitiative-ModulDiagnosen-ImplementationGuide/Terminologien>
17. Licensing SNOMED CT. (2021, February 16). National Library of Medicine. https://www.nlm.nih.gov/healthit/snomedct/snomed_licensing.html#:~:text=The%20UMLS%20is%20a%20global,be%20obtained%20from%20that%20Member
18. National Release Center | Medizininformatik-Initiative. (2021, January 1). National Release Center. <https://www.medizininformatik-initiative.de/de/zusammenarbeit/national-release-center>

19. Bundesministerium für Bildung und Forschung & Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (2020). CODEX | COVID-19 Data Exchange Platform | Medizininformatik-Initiative. Medizin Informatik Initiative. <https://www.medizininformatik-initiative.de/de/use-cases/codex-covid-19-data-exchange-platform>
20. Deutsches Zentrum für Herz-Kreislauf-Forschung e. V. (2018, August 24). Data Catalogue: DZHK. <https://dzhk.de/ressourcen/data-catalogue/>
21. Dugas, M. (2021). Suchergebnisse - Portal für Medizinische Datenmodelle (MDM-Portal). Institute of Medical Informatics Münster. <https://medical-data-models.org/search?query=dzhk>
22. The Apache Software Foundation. (2021). Apache License, Version 2.0. Apache.Org. <https://www.apache.org/licenses/LICENSE-2.0>
23. Das Bundesministerium für Gesundheit. (2020, March). „Datenspende“ – Bedarf für die Forschung, ethische Bewertung, rechtliche, informationstechnologische und organisatorische Rahmenbedingungen.
24. Datenstrategie der Bundesregierung. (2021, January). Die Bundesregierung. <https://www.bundesregierung.de/resource/blob/992814/1845634/45aee6da9554115398cc6a722aba08cb/datenstrategie-der-bundesregierung-download-bpa-data.pdf?download=1>
25. Ministerie van Onderwijs, Cultuur en Wetenschap. (2017, December 1). Progress towards the European Open Science Cloud. News Item | Government.NL. <https://www.government.nl/latest/news/2017/12/01/progress-towards-the-european-open-science-cloud>
26. Lehne, M., Sass, J., Essenwanger, A., Schepers, J., & Thun, S. (2019). Why digital medicine depends on interoperability. NPJ digital medicine, 2, 79. <https://doi.org/10.1038/s41746-019-0158-1>
27. GO FAIR initiative. (2021b, May 27). FAIR Principles. GO FAIR. <https://www.go-fair.org/fair-principles/>
28. Bundesministeriums der Justiz und für Verbraucherschutz sowie des Bundesamts für Justiz –. (2020, April 28). Gesetz zur Durchführung unionsrechtlicher Vorschriften betreffend Medizinprodukte (Medizinprodukte-Durchführungsgesetz - MPDG). Gesetze Im Internet. <https://www.gesetze-im-internet.de/mpdg/MPDG.pdf>
29. Bundesministerium für Gesundheit. (2000, July 20). Gesetz zur Verhütung und Bekämpfung von Infektionskrankheiten beim Menschen (Infektionsschutzgesetz - IfSG). Gesetze Im Internet. <https://www.gesetze-im-internet.de/ifsg/ifsg.pdf>
30. Parciak M, Bender T, Sax U, Bauer CR. Applying FAIRness: Redesigning a Biomedical Informatics Research Data Management Pipeline. Methods Inf Med. 2019 Dec;58(06):229–34.
31. Parciak M, Bauer C, Bender T, Lodahl R, Schreiweis B, Tute E, et al. Provenance Solutions for Medical Research in Heterogeneous IT-Infrastructure: An Implementation Roadmap. Stud Health Technol Inform. 2019 Aug 21;264:298–302.
32. Wolstencroft K, Owen S, du Preez F, Krebs O, Mueller W, Goble C, et al. The SEEK: a platform for sharing data and models in systems biology. Methods Enzymol. 2011;500:629–55.
33. Wolstencroft K, Owen S, Krebs O, Nguyen Q, Stanford NJ, Golebiewski M, et al. SEEK: a systems biology data and model management platform. BMC Syst Biol. 2015 Jul 11;9:33.
34. CovApp. (n.d.). Covapp Charite. Retrieved August 17, 2021, from <https://covapp.charite.de/>
35. C. (n.d.). GitHub - CovOpen/CovApp-2.0: The app lets everyone assess their symptoms using a questionnaire. The app also informs users about next steps, for example, precautionary measures or contacting healthcare providers and health authorities. GitHub. Retrieved August 17, 2021, from <https://github.com/CovOpen/CovApp-2.0>
36. Arbeitsgruppe Interoperabilität | Medizininformatik-Initiative. (n.d.). Medizininformatik-Initiative. Retrieved August 17, 2021, from <https://www.medizininformatik-initiative.de/index.php/de/zusammenarbeit/arbeitsgruppe-interoperabilitaet>
37. Basisprofil DE (R4) - SIMPLIFIER.NET. (2018). Simplifier. <https://simplifier.net/basisprofil-de-r4>
38. SEEK | FAIRDOME. (n.d.). Faidom. Retrieved August 17, 2021, from <https://fair-dom.org/platform/seek/>

Anhang

Anhang 1

Pseudonymisierung / GECCO Datenübermittlung



Hierzu gehört eine projektspezifische Zweitpseudonymisierung durch die Treuhandstelle. Personenidentifizierende Daten (Namen, Adressen) werden nicht an Forschungsprojekte übergeben. Im Rahmen des Nutzungsvertrags verpflichten sich die Wissenschaftler/innen, keinen Versuch zu unternehmen, Personen zu reidentifizieren, deren Daten sie erhalten haben, und keine Daten zu veröffentlichen oder an Dritte weiterzugeben, die es Dritten ermöglichen könnten, einzelne Personen zu reidentifizieren. Einzelne Forschungsvorhaben können die Re-Kontaktierung von Teilnehmern/innen erfordern, z. B. um einen weitergehenden Informed Consent zu erheben, zusätzliche Daten zu erheben oder zusätzliche Bioproben zu gewinnen. (Quelle: Codex Antrag 2020)

**Folgende Universitätskliniken des
Netzwerks Universitätsmedizin
nehmen am COMPASS-Projekt teil:**

Charité – Universitätsmedizin Berlin
Universitätsmedizin Göttingen
Universitätsmedizin Mainz
Universitätsklinikum Würzburg
Uniklinik Köln
Universitätsklinikum Münster
Universitätsklinikum Regensburg
Universitätsklinikum Ulm
Universitätsklinikum Erlangen

Ansprechpartner für weitere Fragen:

COMPASS Koordinierungsstelle
compass@unimedizin-mainz.de



<https://num-compass.science>



@CompassNum

