

Datenqualität in der Anwendung von Agrosystemdaten

Herzlich willkommen zu der Umfrage **"Datenqualität in der Anwendung von Agrosystemdaten"** im Rahmen von FAIRagro.

Vielen Dank für deine Bereitschaft, an dieser Umfrage teilzunehmen! Diese Umfrage richtet sich an alle, die mit Daten rund um den Agrarbereich arbeiten.

In dieser Umfrage geht es um die Anwendung von Datenqualitätskriterien im Umgang mit Forschungsdaten im Agrarbereich. Nach ein paar einleitenden Fragen zu dir folgt ein Fragenblock zu Datenerhebung und einer zur Datennachnutzung. Am Ende der Umfrage hast du die Möglichkeit, Anmerkungen und/oder Feedback an uns zu hinterlassen.

Ziel der Umfrage ist es, von deinen Erfahrungen und Einschätzung rund um das Thema Datenqualität zu lernen. Beantworte die Fragen am besten aus deinem eigenen Bauchgefühl heraus - es gibt dabei kein richtig oder falsch. Wir freuen uns auf deine Antworten.

Die Umfrage erfolgt freiwillig und anonym. Es werden keine personenbezogenen Informationen verarbeitet. Die Umfrage dient der Weiterentwicklung und Optimierung der FAIRagro Services. Im Anschluss der Umfrage hast du die Möglichkeit, über einen Link deine E-Mail zu hinterlassen, um in Zukunft über unsere weitere Arbeit und weitere Umfragen sowie Workshops rund um das Thema Datenqualität informiert zu werden. Weitere Infos findest du außerdem auf der FAIRagro-Website: <http://www.fairagro.net/>

Viel Spaß bei der Umfrage wünschen,

im Namen des FAIRagro Konsortiums
Anne und Jannes

Bei inhaltlichen Fragen wende die bitte an: Jannes Uhlott (jannes.uhlott@julius-kuehn.de)

Bei technischen Fragen oder Problemen wende dich bitte an: Anne Sennhenn (asennhenn@atb-potsdam.de)



FAIRagro

Die Dauer der Umfrage beträgt 10-12 Minuten.

In dieser Umfrage sind 27 Fragen enthalten.

Herzlich Willkommen

Los geht's! Zu Beginn wollen wir ein paar Informationen zu deinem wissenschaftlichen Hintergrund und Erfahrungen im Umgang mit Daten erfragen.

In deinem alltäglichen Umgang mit Daten, zu welcher Gruppe ordnest du dich am ehesten zu?

❶ Wählen Sie alle zutreffenden Optionen

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Datenproduzierende
- ☐ Datennutzende
- ☐ Infrastrukturdienstleistende (z.B. Datenmanager:innen, Repositoriumsbetreibende, Datenbankentwickelnde und IT-Mitarbeitende)
- ☐ Informationsdienstleistende (z.B. Mitarbeitende in Bibliotheken, Verlagen und Forschungsdatenmanagement-Koordination)

☐ Sonstiges:

Zu welcher der folgenden Gruppen ordnest du dich (primär) zu?

❶ Bitte wählen Sie eine der folgenden Antworten:

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ☐ Studierende
- ☐ wissenschaftliche Mitarbeitende (Promovierende)
- ☐ wissenschaftliche Mitarbeitende (Postdocs)
- ☐ nicht-wissenschaftliche Mitarbeitende
- ☐ Professor:innen/AG-Leitende

☐ Sonstiges

Zu welcher der folgenden Einrichtungen ordnest du dich (primär) zu?

❗ Bitte wählen Sie eine der folgenden Antworten:

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ☐ Universität/Hochschule/Fachhochschule
- ☐ Außeruniversitäre Forschungseinrichtungen
- ☐ Behörden
- ☐ Infrastruktureinrichtungen (z.B. Bibliotheken, Archive)
- ☐ Industrie/Dienstleistung
- ☐ Sonstiges

Zu welcher der folgenden DFG-Fachgruppen lässt sich deine (aktuelle) Arbeit zuordnen?

❗ Wählen Sie alle zutreffenden Optionen

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Bodenwissenschaften
- ☐ Pflanzenzüchtung, Pflanzenpathologie
- ☐ Pflanzenbau, Pflanzenernährung, Agrartechnik
- ☐ Ökologie der Landnutzung
- ☐ Agrarökonomie, Agrarpolitik, Agrarsoziologie,
- ☐ Forstwissenschaften
- ☐ Tierzucht, Tierernährung, Tierhaltung
- ☐ Tiermedizin
- ☐ Zoologie
- ☐ Pflanzenwissenschaften
- ☐ Geophysik und Geodäsie
- ☐ Sonstiges:

Datenerhebung

In diesem Abschnitt folgen Fragen zur Umsetzung einer eigenen Datenerhebung.

Falls du momentan keine Daten selbst erhebst, kannst du dich vielleicht an eine frühere Datenerhebung zurück erinnern.

Wenn du bereits mehrere Daten erhoben hast oder gerade mehrere Daten gleichzeitig erhebst, lege dich bitte auf **ein Beispiel** fest. Beantworte die folgenden Fragen bitte bezogen auf dein gewähltes Beispiel.

Erhebst du gerade oder hast du schon mal Daten selbst erhoben? *

Bitte wählen Sie nur eine der folgenden Antworten aus:

☐ Ja

☐ Nein

Für welche Art von Daten fand die Datenerhebung deines Beispiels (primär) statt?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

B1 == "A1" or B1 == "A2"

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ☐ Schlagdaten
- ☐ Omics Daten (z.B.: Genomics, Transcriptomics, Proteomics, Metabolomics)
- ☐ Fachinformationsdaten (z.B.: gesetzliche Vorgaben, Zulassungsdaten zu Pflanzenschutzmitteln und Saatgut, Viehverkehrsdaten)
- ☐ Marktinformationsdaten (z.B.: Ertragsschätzungen, Ernteprognose, Planungsrichtwerte, Zuchtwertschätzungen)
- ☐ Prognose-/Modell-/Simulationsdaten
- ☐ Wetterdaten
- ☐ Vorkommens-/Befalls-/Verteilungsdaten
- ☐ Bodendaten
- ☐ Boniturdaten
- ☐ Labordaten (z.B.: Chemische Analysedaten)
- ☐ Fernerkundungsdaten (z.B.: Drohnendaten)
- ☐ Technische Gerätedaten
- ☐ Phenotypisierungsdaten
- ☐ Leistungsprüfungsdaten (z.B.: Milchleistungsprüfung bei Rind oder Fleischleistungsprüfung Schwein)
- ☐ Fragebögen
- ☐ Statistiken
- ☐ Sonstiges

Für welche Art von Datentypen fand die Datenerhebung deines Beispiels statt?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

Antwort war 'Ja' bei Frage ' [B1]' (Erhebst du gerade oder hast du schon mal Daten selbst erhoben?)

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ☐ Numerische Daten
- ☐ Textdaten
- ☐ Audiodaten
- ☐ Fotodaten
- ☐ Videodaten
- ☐ Tabellarische Daten
- ☐ Räumliche Daten
- ☐ Source Codes
- ☐ Nicht digitale Daten (z.B. Bodenproben)
- ☐ Sonstiges

Für welchen Anwendungsbereich fand die Datenerhebung deines Beispiels (primär) statt?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

B1 == "A1" or B1 == "A2"

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ☐ Bodenkunde
- ☐ Pflanzenzüchtung, Pflanzenpathologie
- ☐ Pflanzenbau, Pflanzenernährung, Agrartechnik
- ☐ Ökologie der Landnutzung
- ☐ Agrarökonomie, Agrarpolitik, Agrarsoziologie,
- ☐ Forstwissenschaften
- ☐ Tierzucht, Tierernährung, Tierhaltung
- ☐ Tiermedizin
- ☐ Zoologie
- ☐ Pflanzenwissenschaften
- ☐ Geophysik und Geodäsie
- ☐ Sonstiges

Was wären für dich die drei wichtigsten Kriterien, um die Qualität deiner erhobenen Daten zu beschreiben?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

B1 == "A1" or B1 == "A2"

Bitte nummerieren Sie jede Box in der Reihenfolge Ihrer Präferenz, beginnen mit 1 bis 10

Bitte wähle Sie nicht mehr als 3 Einträge aus.

Auflösung (z.B. räumlich, inhaltlich, zeitlich)

Vollständigkeit (z.B. räumlich, inhaltlich, zeitlich)

Informationen über statistischen Sicherheiten (z.B. Angabe von Unsicherheiten, geprüfte inhaltliche Konsistenz)

Maschinenlesbarkeit

Verwendung von Standards / Dateninteroperabilität

Informationen zur Datenvorbehandlung (z.B. Anwendung von Filtern, Ausreißerkorrektur)

Umfangreiche Metadaten

Informationen zur Methodik (z.B. Feldprotokolle)

Umfangreiche Beschreibung der Daten (z.B. Beschriftung von Spalten, Einheiten)

Aktualität des Datensatzes

Hat bei der vorherigen Frage ein für dich relevantes Datenqualitätskriterium gefehlt? Wenn ja, welches?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

Antwort war 'Ja' bei Frage ' [B1]' (Erhebst du gerade oder hast du schon mal Daten selbst erhoben?)

Bitte geben Sie Ihre Antwort hier ein:

Wie stellst du die Qualität deiner erhobenen Daten sicher?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

B1 == "A1" or B1 == "A2"

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Kalibrierung von Instrumenten (Genauigkeit/Skala)
- ☐ Mehrfachmessungen
- ☐ Überprüfung durch Expert:innen
- ☐ Verwendung von standardisierten Methoden und Protokollen
- ☐ Minimierung der manuellen Dateneingabe (z.B. kontrollierte Vokabulare, Codelisten, Auswahllisten)
- ☐ Manuelle Überprüfung (z.B. Vollständigkeit, Doppeleingaben)
- ☐ Statistische Analysen (z.B. Häufigkeiten, Mittelwerte, Streuung, Ausreißerwerte)
- ☐ Vergleiche mit Referenzdaten
- ☐ Diskussion der Daten mit Fachkolleg:innen
- ☐ Berücksichtigung der Randbedingungen (z.B. Bewölkung, Temperatur, Bodenfeuchte...)

☐ Sonstiges:

Was sind für dich die größten Herausforderungen in der Datenerhebung im Hinblick auf Datenqualität?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

B1 == "A1" or B1 == "A2"

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Keine klaren Metadatenstandards
- ☐ Keine klaren Datenqualitätsstandards
- ☐ Fehlende Informationen zu Datenqualität (Was ist wichtig? Was benötigen Andere?)
- ☐ Fehlende Ansprechpersonen (für Fragen/Herausforderungen)
- ☐ Fehlende Zeit
- ☐ Fehlende Standards bei der Datenerhebung (z.B. einheitliche Protokolle)
- ☐ Unzureichende Sensorgenauigkeit
- ☐ Einflüsse von Randbedingungen (z.B. Witterungen)
- ☐ Fehlender Ort für die Dokumentation von Metadaten/Datenqualitätsinformationen
- ☐ Technische Herausforderungen (keine Metadatenfelder vorhanden)
- ☐ Ich finde es nicht notwendig Metadaten/Datenqualitätsinformationen zu dokumentieren
- ☐ Mir sind keine Qualitätsanforderungen an meine Daten bekannt (z.B. die erste Datenerhebung in einem spezifischen Bereich)
- ☐ Sonstiges:

Datennutzung

In diesem Abschnitt folgen Fragen zur Datennutzung bzw. Datennachnutzung, also konkret um die Nutzung von bereits existierenden Daten. Falls du zur Zeit keine Daten nutzt, die du nicht selbst erhoben hast, kannst du dich vielleicht an eine frühere Verwendung von Datensätzen erinnern und diese nutzen. Wenn du mehrere Datensätze nutzt / genutzt hast, dann lege dich bitte in den folgenden Fragen auf **ein Beispiel** fest. Beantworte die folgenden Fragen bitte bezogen auf dein gewähltes Beispiel.

Nutzt du oder hast du schon mal Daten genutzt, die du nicht selbst erhoben hast? *

Bitte wählen Sie nur eine der folgenden Antworten aus:

☐ Ja

☐ Nein

Welche Art von Daten hast du in deinem Beispiel (primär) genutzt?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

C1 == "A1" or C1 == "A2"

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ☐ Schlagdaten
- ☐ Omics Daten (z.B.: Genomics, Transcriptomics, Proteomics, Metabolomics)
- ☐ Fachinformationsdaten (z.B.: gesetzliche Vorgaben, Zulassungsdaten zu Pflanzenschutzmitteln und Saatgut, Viehverkehrsdaten)
- ☐ Marktinformationsdaten (z.B.: Ertragsschätzungen, Ernteprognose, Planungsrichtwerte, Zuchtwertschätzungen)
- ☐ Prognose-/Modell-/Simulationsdaten
- ☐ Wetterdaten
- ☐ Vorkommens-/Befalls-/Verteilungsdaten
- ☐ Bodendaten
- ☐ Boniturdaten
- ☐ Labordaten (z.B.: Chemische Analysedaten)
- ☐ Fernerkundungsdaten (z.B.: Drohnendaten)
- ☐ Technische Gerätedaten
- ☐ Phenotypisierungsdaten
- ☐ Leistungsprüfungsdaten (z.B.: Milchleistungsprüfung bei Rind oder Fleischleistungsprüfung Schwein)
- ☐ Fragebögen
- ☐ Statistiken
- ☐ Sonstiges

Welche Art von Datentypen hast du in deinem Beispiel (primär) genutzt?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

Antwort war 'Ja' bei Frage ' [C1]' (Nutzt du oder hast du schon mal Daten genutzt, die du nicht selbst erhoben hast?)

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ☐ Numerische Daten
- ☐ Textdaten
- ☐ Audiodaten
- ☐ Fotodaten
- ☐ Videodaten
- ☐ Tabellarische Daten
- ☐ Räumliche Daten
- ☐ Source Codes
- ☐ Nicht digitale Daten (z.B. Bodenproben)
- ☐ Sonstiges

Welches Ziel hat deine Datennachnutzung?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

C1 == "A1" or C1 == "A2"

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Generierung Datenprodukte (z.B. fachspezifische Karten)
- ☐ Kommerzielle Nutzung/Produktentwicklung (z.B. Dünger)
- ☐ Analysen (z.B. Qualitätsanalysen)
- ☐ Eingangsdaten für Modelle und Algorithmen (z.B. Ertragsschätzung)
- ☐ Planung/ Entscheidungsfindung (z.B. Standortbestimmung)
- ☐ Sonstiges:

Woher stammen diese Daten? Bitte spezifiziere die von dir genutzten Datenquellen.

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

C1 == "A1" or C1 == "A2"

Bitte wählen Sie die zutreffenden Punkte aus und schreiben Sie einen Kommentar dazu:

☐ Öffentlich zugängliche Daten (z.B. Repositorien, Suchmaschinen, Datenbanken)

☐ Interne Daten (z.B. Kolleg:innen innerhalb der AG, Institution oder Projekten)

Sonstiges:

Welche drei Datenqualitätskriterien sollten mindestens erfüllt sein, damit du Daten gut nachnutzen könntest?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

C1 == "A1" or C1 == "A2"

Bitte nummerieren Sie jede Box in der Reihenfolge Ihrer Präferenz, beginnen mit 1 bis 12

Bitte wähle Sie nicht mehr als 3 Einträge aus.

Auflösung (z.B. räumlich, inhaltlich, zeitlich)

Vollständigkeit (z.B. räumlich, inhaltlich, zeitlich)

Informationen über statistische Sicherheiten (z.B. Angabe von Unsicherheiten, Geprüfte inhaltliche Konsistenz)

Maschinenlesbarkeit

Verwendung von Standards / Dateninteroperabilität

Informationen zur Datenvorbehandlung (z.B. Anwendung von Filtern, Ausreißerkorrektur)

Umfangreiche Metadaten

Informationen zur Methodik (z.B. Feldprotokolle)

Umfangreiche Beschreibung der Daten (z.B. Beschriftung von Spalten, Einheiten)

Aktualität des Datensatzes

Sichere Datenquelle (z.B. offizielles Repositorium, veröffentlichte Daten)

Sichere Nutzungsrechte/Open Access

Was wären Gründe, weshalb du Daten nicht verwenden könntest?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

C1 == "A1" or C1 == "A2"

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Unvollständige Daten
- ☐ Inkorrekte Daten
- ☐ Datenaufbereitung nicht dokumentiert
- ☐ Fehlende Material- und Methodenbeschreibung
- ☐ Kein Open Access/Open Data
- ☐ Kosten
- ☐ Mangelnde Infrastruktur (z.B. Big Data, fehlende Rechenkapazität)
- ☐ Sonstiges:

Welche Schritte würdest du unternehmen, um Datensätze trotz unzureichender Qualität für die gewünschten Zwecke nutzbar zu machen?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

Antwort war 'Ja' bei Frage ' [C1]' (Nutzt du oder hast du schon mal Daten genutzt, die du nicht selbst erhoben hast?)

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Interpolation (z.B. räumlich, zeitlich, inhaltlich)
- ☐ Extrapolation (z.B. räumlich, zeitlich, inhaltlich)
- ☐ Datenreduktion
- ☐ Datenbereinigung
- ☐ Datenergänzung
- ☐ Normalisierung/Standardisierung
- ☐ Expert:innenwissen nutzen

☐ Sonstiges:

Was sind für dich die größten Herausforderungen in der Datennachnutzung im Hinblick auf Datenqualität?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

C1 == "A1" or C1 == "A2"

Bitte geben Sie Ihre Antwort hier ein:

Datennachnutzung in der Anwendung

Bei der Verwendung von externen Daten kann es vorkommen, dass mehrere Datensätze vorliegen, die für den

Anwendungsfall genutzt werden könnten. In diesem Abschnitt folgen Fragen, die die Herausforderungen in diesem Fall abbilden sollen.

Stell dir vor, du würdest einen neuen Datensatz nutzen wollen. In der Recherche findest du neben den Metadaten auch Informationen, wie gut (oder schlecht) der Datensatz bereits in einem ähnlichen Anwendungsfall nutzbar war. Diese Information finde ich:

Bitte wählen Sie nur eine der folgenden Antworten aus:

☐ Hilfreich

☐ Neutral

☐ Unwichtig

☐ Sonstiges

Würdest du selbst Zeit investieren, um abgeschlossene Anwendungen eines Datensatzes in dessen Metadaten zu dokumentieren?

Bitte wählen Sie nur eine der folgenden Antworten aus:

☐ Ja

☐ Nein

Wie viel Zeit (in Minuten) wärst du bereit zu investieren, um die abgeschlossene Anwendung eines Datensatzes in dessen Metadaten zu dokumentieren?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

Antwort war 'Ja' bei Frage ' [D4]' (Würdest du selbst Zeit investieren, um abgeschlossene Anwendungen eines Datensatzes in dessen Metadaten zu dokumentieren?)

Bitte geben Sie Ihre Antwort hier ein:

Hast du schon mal für einen Anwendungsfall vor der Berechnung oder für ein Modell (o.ä.) verschiedene Eingangsdaten miteinander verglichen, um den "besten" Datensatz zu nutzen?

Bitte wählen Sie nur eine der folgenden Antworten aus:

☐

Ja

☐

Nein

Wonach hast du den "passenden" Datensatz ausgewählt?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:

D2 == "A1"

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Höchste Auflösung (z.B. räumlich, inhaltlich, zeitlich)
- ☐ Ausführlichste Metadaten
- ☐ Ausführlichste Beschreibung der Datenaufbereitung
- ☐ Beste Referenzen (z.B. Empfehlungen durch Andere)
- ☐ Dichteste Zeitreihe
- ☐ Wenigsten Ausreißer
- ☐ Höchste Güte (z.B. geringste Fehlerquote)
- ☐ Höchste Wiederholrate (z.B. jedes Jahr, statt alle zwei Jahre)
- ☐ Größte Abdeckung (z.B. ganz Deutschland statt einzelnes Bundesland)
- ☐ Beste Ergebnisse in der Anwendung
- ☐ Passenste Semantik (z.B. Klassen)
- ☐ Geringste Kosten
- ☐ Open Access
- ☐ Geringste Dateigröße
- ☐ Sonstiges:

Weiteres zum Thema Datenqualität

Dieser letzte Abschnitt gibt dir Raum, um deine persönlichen Gedanken zum Thema Datenqualität auszuführen.

Gibt es weitere Informationen oder Gedanken zum Thema Datenqualität, die du mit uns teilen möchtest?

Bitte geben Sie Ihre Antwort hier ein:

Herzlichen Dank für deine Teilnahme an der Umfrage zum Thema "**Datenqualität in der Anwendung von Agrosystemdaten**" und deinem wertvollen Beitrag auf dem Weg zu einer FAIRen Datenzukunft in der Agrarsystemforschung!

Bist du explizit am Thema Datenqualität und an weiteren Möglichkeiten der Zusammenarbeit im Rahmen von FAIRagro interessiert? Dann kannst du uns in einem [separaten Kontaktformular](#) deine Kontaktinformationen hinterlassen!

Die Ergebnisse der Umfrage werden beim FAIRagro-Workshop am 13.+ 14.03.2024 in Braunschweig vorgestellt. Mehr Infos dazu folgen im FAIRagro Newsletter. Trag dich gern in die [Mailingliste für den FAIRagro Newsletter](#) ein und erhalte weitere Informationen auch zu aktuellen Veranstaltungsangeboten und Services.

Weitere Informationen über die NFDI Initiative FAIRagro sowie FAIRes Datenmanagement findest du auch auf den [FAIRagro Webseiten](#).

Für aktuelle Einblicke in die Aktivitäten von FAIRagro folge FAIRagro auf X (formerly known as Twitter; [@FAIR_agro](#)) und/oder Mastodon (<http://nfdi.social/@FAIRagro>)!

Bei Fragen oder Anmerkungen zu der Umfrage: Jannes.uhlott@julius-kuehn.de



FAIRagro

20.01.2024 – 23:59

Senden Sie Ihre Umfrage ein.

Vielen Dank für die Beantwortung des Fragebogens.