

Corona-Datenspende - Teildatensatz Erleben und Verhalten in der Pandemie

[Robert Koch-Institut](#) | RKI

Nordufer 20

13353 Berlin

[Marc Wiedermann](#) & [Dirk Brockmann](#)

P4 | Epidemiologische Modellierung von Infektionskrankheiten | Robert Koch-Institut

[Research on Complex Systems](#) | Institut für theoretische Biologie | Humboldt-Universität zu Berlin

[Philipp Sprengholz](#)

Seminar für Medien- und Kommunikationswissenschaft | Universität Erfurt

Wiedermann M, Sprengholz P und Brockmann D (2022): Corona-Datenspende - Teildatensatz Erleben und Verhalten in der Pandemie, Berlin: Zenodo. DOI:[10.5281/zenodo.7185199](https://doi.org/10.5281/zenodo.7185199).

Informationen zum Datensatz und Entstehungskontext

Die Daten von Fitnessarmbändern und Smartwatches, sogenannten Wearables, können Hinweise auf Symptome einer Infektion mit COVID-19 liefern. Mit Hilfe der [Corona-Datenspende-App](#) (CDA) können Bürger:innen dem Robert Koch-Institut diese Daten zur wissenschaftlichen Auswertung zur Verfügung stellen. Zusammen mit Informationen aus anderen Quellen, z.B. offiziellen Meldedaten zu Fallzahlen, helfen diese Daten den Wissenschaftler:innen, die Ausbreitung des Coronavirus besser zu erfassen und zu verstehen.

In ihrer ersten, seit April 2020 verfügbaren, Version erhob die CDA nur sogenannte Vitaldaten der Nutzer:innen, insbesondere den Ruhepuls, körperliche Aktivität und Schlafverhalten. In einem Update im Oktober 2021 wurden der App Umfragemodule zu verschiedenen, für die Pandemieforschung relevanten Fragestellungen, hinzugefügt. Die hier bereitgestellten Daten entstammen der Teilstudie "Erleben und Verhalten in der Pandemie" die momentan als eines von drei Befragungsmodulen innerhalb der CDA durchgeführt wird. Sie basiert auf dem [COVID-19 Snapshot Monitoring \(COSMO\)](#), einem sich wiederholenden querschnittlichen Monitoring von Wissen, Risikowahrnehmung, Schutzverhalten und Vertrauen während des aktuellen COVID-19 Ausbruchsgeschehens (Betsch et al., 2022). In dieser Studie wollen das Robert Koch-Institut sowie führende Verhaltensforscher:innen erfahren, wie sich die Bevölkerung unter pandemischen Bedingungen verhält. Durch die Beantwortung regelmäßiger Fragebögen können Teilnehmer:innen dabei helfen, Strategien zur Bekämpfung des Coronavirus zu optimieren. Die Forscher:innen wollen hierzu mehr über den Arbeitsalltag, die persönliche Belastung und die Risikowahrnehmung lernen.

Die hier bereitgestellten Daten dienen der Reproduktion aller Ergebnisse in der Studie "From Delta to Omicron: The role of individual factors and social context in compliance with pandemic regulations and recommendations" (Sprengholz et al., 2022, in Begutachtung). Es werden hier daher jene Datenpunkte aus der Studie „Erleben und Verhalten in der Pandemie“ bereitgestellt, die zu diesem Zweck benötigt werden.

Das heißt auch, dass nicht alle im Rahmen dieser Studie erhobenen Datenpunkte, beispielsweise Informationen zu COVID-19-Testergebnissen, sowie alle Zeitpunkte der Erhebung im vorliegenden Datensatz enthalten sind. Ein Link zur oben genannten Publikation folgt nach der Veröffentlichung in einer Fachzeitschrift.

Projektbeteiligte und Rollenbesetzung

Die CDA wird vom Robert Koch-Institut herausgegeben. Die Projektleitung und -koordination liegt bei der Projektgruppe [P4 | Modellierung von Infektionskrankheiten](#). Die App wurde in Zusammenarbeit mit [Thryve \(mHealth Pioneers GmbH\)](#) entwickelt, einem auf Digital Health spezialisierten Unternehmen. Dieses Unternehmen ist der technologische Dienstleister. Fragen bezüglich der Corona-Datenspende können an corona-datenspende@rki.de gerichtet werden.

Die Veröffentlichung der Daten sowie das Qualitätsmanagement der (Meta-)Daten erfolgen durch das Fachgebiet [MF 4 | Informations- und Forschungsdatenmanagement](#). Fragen zum Datenmanagement und zur Publikationsinfrastruktur können an das Open Data Team des Fachgebiets MF4 unter OpenData@rki.de gerichtet werden.

Durch sorgfältige Auswahl und regelmäßige Kontrolle stellt das Robert Koch-Institut sicher, dass der Dienstleister alle organisatorischen und technischen Maßnahmen trifft, die zum Schutz der Daten erforderlich sind. Alle Maßnahmen stehen im Einklang mit dem geltenden Datenschutzrecht.

Datenerhebung

Zunächst laden sich Nutzer:innen die CDA in einem der gängigen App-Stores ([iOS](#), [Android](#), [Huawei](#)) herunter. Wenn sich neue Nutzer:innen anmelden, werden zunächst Geburtsjahr, Geschlecht und die ersten drei Stellen der Postleitzahl erfasst. Alle Nutzer:innen nehmen automatisch am Teilprojekt "Fiebermonitor" teil, in dessen Kontext die oben angesprochenen Vitaldaten erhoben und ausgewertet werden. Die Teilnahme an den Befragungsstudien ist optional und eine Zustimmung erfolgt nach vorheriger Aufklärung innerhalb der App. Hierbei ist eine Teilnahme auch ohne Wearable möglich.

In der Befragungsstudie „Erleben und Verhalten in der Pandemie“ werden in einem einmaligen Basisfragebogen zunächst Persönlichkeitsmerkmale gemäß der "Big Five" abgefragt. Danach folgen wöchentliche und monatliche Erhebungen unter anderem zum persönlichen Wohlbefinden, COVID-19 Testergebnissen sowie Symptomen im Verlauf einer Erkrankung, dem eigenen Schutzverhalten zur Vermeidung von Infektionen oder einer Selbsteinschätzung des persönlichen Infektionsrisikos. Die Nutzer:innen werden per Push-Notification auf Ihrem Smartphone über das Vorhandensein neuer Fragebögen informiert.

Die meisten der in diesem Teildatensatz bereitgestellten Informationen wurden mittels der monatlichen Fragebögen erhoben. Da die Teilnehmer:innen die Fragebögen zu beliebigen Zeitpunkten in der jeweiligen Woche oder im jeweiligen Monat beantworten konnten, liegen die Daten kontinuierlich und nicht notwendigerweise für die Tage der Aussteuerung vor. Die Teilnahme ist freiwillig und unentgeltlich, alle Teilnehmer:innen erklärten sich mit der anonymisierten Verwendung und Weitergabe ihrer Daten für wissenschaftliche Zwecke einverstanden.

Technische Voraussetzungen

Die Datenerhebung erfolgt über die Corona-Datenspende-App (CDA). Diese ist für [iOS](#), [Android](#) und [Huawei](#) in den jeweiligen Anwendungsplattformen verfügbar und zur Teilnahme notwendig.

Neben einem Smartphone mit installierter CDA wird für die Teilnahme am "Fiebermonitor" ein Fitnessarmband oder eine Smartwatch benötigt. Die CDA unterstützt aktuell Geräte von Apple, Samsung, Fitbit, Garmin, Amazfit, Oura, Polar und Withings. Eine Teilnahme an Befragungsstudien ist auch ohne Wearable möglich. Der hier bereitgestellte Datensatz speist sich ausschließlich aus Befragungen.

Teilnehmer:innen

Ein- und Ausschlusskriterien

Eine Teilnahme ist allen Personen ab 16 Jahren möglich die Zugriff auf einen deutschen App-Store haben. Eine Veröffentlichung der CDA in App Stores außerhalb von Deutschland ist im Moment aus rechtlichen Gründen nicht möglich. Die CDA wird vom Robert Koch-Institut herausgegeben, einer deutschen Bundesbehörde im Geschäftsbereich des Bundesministeriums für Gesundheit.

Grundgesamtheit

Insgesamt haben bereits mehr als eine Million Menschen die CDA installiert. Mehr als 500,000 Menschen haben schlussendlich mindestens ein Wearable verbunden und somit mindestens einen Datenpunkt übermittelt. An den Befragungsstudien nehmen regelmäßig bis zu 30,000 Menschen teil.

In diesem Datensatz werden ein Teil der im Rahmen der Studie "Erleben & Verhalten in der Pandemie" vom Oktober 2021 bis Februar 2022 erhobenen Befragungsdaten bereitgestellt. Dies beinhaltet die Antworten von 29,355 Teilnehmenden.

Fragebogen

Zur besseren Übersicht, werden im Folgenden zunächst alle im Datensatz enthaltenen Abfragen und Antwortausprägungen vorgestellt. Eine Aufschlüsselung, wie Fragen und Ausprägungen codiert werden, findet sich in einer zweiten Tabelle im Anschluss. Über die Spalte **Varname** ist eine Verknüpfung der entsprechenden Zeilen in beiden Tabellen möglich.

Frage	Ausprägungen	Varname
Das Coronavirus ist für mich...	1. nicht besorgniserregend 2. kaum besorgniserregend 3. wenig besorgniserregend 4. eher besorgniserregend 5. besorgniserregend	AFFRISK
Bitte schätzen Sie sich selbst ein:	1. sehr barsch (kühl, kritisch, wird leicht ärgerlich, misstrauisch) 2. eher barsch 3. teils/teils 4. eher umgänglich 5. sehr umgänglich (kooperativ, herzlich, nett, nachgiebig)	BF_AGREE

Frage	Ausprägungen	Varname
Bitte schätzen Sie sich selbst ein:	1. sehr gewissenhaft (zuverlässig, sorgfältig, ordentlich, pflichtbewusst, ehrgeizig) 2. eher gewissenhaft 3. teils/teils 4. eher nachlässig 5. nachlässig (salopp, unordentlich, unpünktlich, chaotisch, nonchalant)	BF_CONSC
Bitte schätzen Sie sich selbst ein:	1. sehr extrovertiert (gesprächig, gesellig, voller Tatendrang, kontaktfreudig) 2. eher extrovertiert 3. teils/teils 4. eher introvertiert 5. sehr introvertiert (still, schweigsam, zurückgezogen, eher ein Einzelgänger)	BF_EXTRA
Bitte schätzen Sie sich selbst ein:	1. sehr emotional (nervös, ängstlich, regt sich leicht auf, angespannt, empfindlich) 2. eher emotional 3. teils/teils 4. eher ausgeglichen 5. sehr ausgeglichen (gelassen, entspannt, selbstzufrieden, robust)	BF_NEURO
Bitte schätzen Sie sich selbst ein:	1. sehr aufgeschlossen (offen für Neues, wissbegierig, kultiviert, fantasievoll) 2. eher aufgeschlossen 3. teils/teils 4. eher festgelegt (feste Ansichten und Meinungen, an Neuem eher wenig interessiert, traditionell, mag Routinen und feste Regeln) 5. sehr festgelegt (feste Ansichten und Meinungen, an Neuem eher wenig interessiert, traditionell, mag Routinen und feste Regeln)	BF_OPEN

Frage	Ausprägungen	Varname
Haben Sie Kinder unter 18 Jahren?	Ja/nein	CHILDREN
Ich finde die Maßnahmen, die derzeit zur Pandemiebekämpfung ergriffen werden, stark übertrieben.	1. stimme überhaupt nicht zu 2. stimme eher nicht zu 3. teils teils 4. stimme eher zu 5. stimme voll und ganz zu	EXAGG
Wie hoch schätzen Sie Ihre Wahrscheinlichkeit ein, dass Sie sich mit dem Coronavirus infizieren?	1. extrem unwahrscheinlich 2. eher unwahrscheinlich 3. vielleicht 4. eher wahrscheinlich 5. extrem wahrscheinlich	PROB
Bei Ihnen selbst - Wie sehr haben Sie sich in der letzten Woche privat an die geltenden Vorschriften und Empfehlungen zur Pandemiebekämpfung gehalten?	1. überhaupt nicht 2. kaum 3. etwas 4. überwiegend 5. sehr	PROTECT_7D_PRIVATE
Wie sehr haben sich Ihre Kollegen in der letzten Woche am Arbeitsplatz an die geltenden Vorschriften und Empfehlungen zur Pandemiebekämpfung gehalten?	1. überhaupt nicht 2. kaum 3. etwas 4. überwiegend 5. sehr	PROTECT_7D_WORK_COLLEAGUES
Wie sehr haben Sie sich in der letzten Woche am Arbeitsplatz an die geltenden Vorschriften und Empfehlungen zur Pandemiebekämpfung gehalten?	1. überhaupt nicht 2. kaum 3. etwas 4. überwiegend 5. sehr	PROTECT_7D_WORK_OWN
Wie schätzen Sie eine Infektion mit dem Coronavirus für sich selbst ein?	1. völlig harmlos 2. eher harmlos 3. mäßig 4. eher gefährlich 5. extrem gefährlich	SEVERITY
Wie sehr vertrauen Sie der Bundesregierung, dass sie mit dem Coronavirus gut und richtig umgeht?	1. überhaupt nicht 2. eher nicht 3. teilweise 4. überwiegend 5. ganz und gar	TRUST_GOV
Sind Sie berufstätig?	Ja/nein	WORK

Frage	Ausprägungen	Varname
Wie haben Sie in der letzten Woche ihren Arbeitsalltag gestaltet?	1. Ich habe ausschließlich im Home Office gearbeitet 2. Ich habe teilweise im Home Office gearbeitet 3. Ich habe ausschließlich vor Ort bei meinem Arbeitgeber gearbeitet	WORKSITUATION

Aufbau und Inhalt des Datensatzes

Zentraler Inhalt des Datensatzes sind die Umfrage-Rohdaten der Studie "Erleben und Verhalten in der Pandemie" die innerhalb der CDA erhoben werden. Im Datensatz insgesamt enthalten sind:

- Umfrage-Rohdaten der Studie "Erleben und Verhalten in der Pandemie" von Oktober 2021 bis Februar 2022
- Kontextmaterialien zur Studie
- Lizenz-Dateien mit der Nutzungslizenz des Datensatzes in deutsch und englisch
- Datensatzdokumentation in deutscher Sprache
- Metadaten zum Bereitstellung

Umfrage-Rohdaten

Formatierung der Daten

Die Daten der Studie sind im Datensatz als kommaseparierte `.csv` Datei enthalten. Der verwendete Zeichensatz der `.csv` Datei ist UTF-8. Trennzeichen der einzelnen Werte ist ein Komma. Datumsangaben entsprechen dem Standard ISO-8601. Jede Zeile der `.csv` Datei enthält die abgegebenen Antworten einer Personen auf jeweils einen Fragebogen. Einträge die mit "NA" gekennzeichnet sind wurden im entsprechenden Fragebogen nicht abgefragt.

Variablen und Variablenausprägungen

Insgesamt werden Antworten der Nutzer:innen auf 16 sich teils regelmäßig wiederholende Fragen bereit gestellt (siehe oben). Zusätzlich wird eine konstant bleibende User-ID der antwortgebenden Person, sowie Geschlecht und Geburtsjahr (in 5-Jahres Schritten) angegeben. Diese Daten werden bei Installation der CDA erhoben. Ebenso enthalten ist die Frequenz/Art des Fragebogens (einmalig, wöchentlich, monatlich) sowie Meta-Informationen zum Zeitpunkt der Antwortabgabe.

Varname	Format	Ausprägungen	Kurzbeschreibung
userid	ganzzahlig	Fortlaufend	User ID der antwortgebenden Person

Varname	Format	Ausprägungen	Kurzbeschreibung
questionnaire	kategorisch	[1 , 2 , 3]	Art des Fragebogens (1. Baseline, 2. Monatlich, 3. Wöchentlich)
gender	dichotom	male , female	Geschlecht
birthDate	ganzzahlig	5-Jahres Kategorien	Geburtsjahr
AFFRISK	Ordinal	[1 , 2 , 3 , 4 , 5]	Selbsteinschätzung des affektiven Risikos, Formulierung und Ausprägungen siehe oben
BF_AGREE	Ordinal	[1 , 2 , 3 , 4 , 5]	Selbsteinschätzung der eigenen Vertraeglichkeit, Formulierung und Ausprägungen siehe oben
BF_CONSC	Ordinal	[1 , 2 , 3 , 4 , 5]	Selbsteinschätzung der eigenen Gewissenhaftigkeit, Formulierung und Ausprägungen siehe oben
BF_EXTRA	Ordinal	[1 , 2 , 3 , 4 , 5]	Selbsteinschätzung der eigenen Extraversion, Formulierung und Ausprägungen siehe oben
BF_NEURO	Ordinal	[1 , 2 , 3 , 4 , 5]	Selbsteinschätzung des eigenen Neurotizismus, Formulierung und Ausprägungen siehe oben
BF_OPEN	Ordinal	[1 , 2 , 3 , 4 , 5]	Selbsteinschätzung der eigenen Offenheit, Formulierung und Ausprägungen siehe oben
CHILDREN	dichotom	ja , nein	Vorhandensein von Kindern, Formulierung siehe oben
EXAGG	Ordinal	[1 , 2 , 3 , 4 , 5]	Individuelle Einschätzung von Schutzmaßnahmen, Formulierung und Ausprägungen siehe oben
PROB	Ordinal	[1 , 2 , 3 , 4 , 5]	Einschätzung der eigenen Erkrankungswahrscheinlichkeit, Formulierung und Ausprägungen siehe oben
PROTECT_7D_PRIVATE	Ordinal	[1 , 2 , 3 , 4 , 5]	Privates Schutzverhalten, Formulierung und Ausprägungen siehe oben

Varname	Format	Ausprägungen	Kurzbeschreibung
PROTECT_7D_WORK_COLLEAGUES	Ordinal	[1 , 2 , 3 , 4 , 5]	Schutzverhalten der Kollegen am Arbeitsplatz, Formulierung und Ausprägungen siehe oben
PROTECT_7D_WORK_OWN	Ordinal	[1 , 2 , 3 , 4 , 5]	Eigenes Schutzverhalten am Arbeitsplatz, Formulierung und Ausprägungen siehe oben
SEVERITY	Ordinal	[1 , 2 , 3 , 4 , 5]	Selbsteinschätzung des Schweregrads einer COVID-19 Erkrankung, Formulierung und Ausprägungen siehe oben
TRUST_GOV	Ordinal	[1 , 2 , 3 , 4 , 5]	Vertrauen in die Regierung, Formulierung Ausprägungen siehe oben
WORK	dichotom	ja , nein	Berufstätigkeit, Formulierung siehe oben
WORKSITUATION	kategorisch	[1 , 2 , 3]	Gestaltung des Arbeitsalltags, Formulierung und Ausprägungen siehe oben
DATE	YYYY-MM-DD	fortlaufend	Antwortdatum im ISO-8601 Format
sessionIteration	Ganzzahlig	fortlaufend	Wiederholung des Fragebogens. Gibt an, wie häufig der entsprechende Fragebogen in der Vergangenheit bereits an die betreffende Person ausgesteuert wurde
sessionTimestamp	Ganzzahliger Zeitstempel	fortlaufend	Zeitstempel der Antwortübermittlung in Unixzeit . Die Zeitstempel bezieht sich auf die vergangenen Millisekunden seit dem 1. Januar 1970, 00:00 Uhr UTC.

Kontextmaterialien

Umfassende weitere Informationen zur Corona-Datenspende sowie aktuelle Auswertungsergebnisse finden sich auf <https://corona-datenspende.de/science/>.

Eine Publikation der Analysen auf Basis des hier bereit gestellten Datensatzes folgt in Kürze. Skripte zur Aufbereitung der Daten und Reproduktion der Ergebnisse in vorgenannter Publikation folgen ebenfalls in Kürze auf dem [Open Science Framework](#). Die entsprechenden Links werden dann hier angegeben.

Metadaten

Die bereitgestellten Daten sind mit Metadaten beschrieben und wissenschaftlich zitierbar, u.a. durch die Vergabe einer DOI durch Zenodo.org. Die für den Import in Zenodo bereitgestellten Metadaten sind in folgender Datei hinterlegt:

- [Metadaten/zenodo.json](#)

Die Dokumentation der einzelnen Metadatenvariablen ist unter <https://developers.zenodo.org/#representation> nachlesbar.

Hinweise zur Nachnutzung der Daten

Offene Forschungsdaten des RKI werden auf github.com, zenodo.org und edoc.rki.de bereitgestellt:

- <https://github.com/robert-koch-institut>
- <https://zenodo.org/communities/robertkochinstitut>
- <https://edoc.rki.de/>

Lizenz

Der Datensatz "Corona-Datenspende --- Teildatensatz Erleben und Verhalten in der Pandemie" ist lizenziert unter der [Creative Commons Namensnennung 4.0 International Public License | CC-BY 4.0 International](#)

Die im Datensatz bereitgestellten Daten sind, unter Bedingung der Namensnennung des Robert Koch-Instituts als Quelle, frei verfügbar. Das bedeutet, jede:r hat das Recht die Daten zu verarbeiten und zu verändern, Derivate des Datensatzes zu erstellen und sie für kommerzielle und nicht kommerzielle Zwecke zu nutzen. Weitere Informationen zur Lizenz finden sich in der [LICENSE](#) bzw. [LIZENZ](#) Datei des Datensatzes.