

Open Data Sandbox

[AKTIN-Notaufnahmeregister](#), [Robert Koch-Institut](#)

Beitragende

Susanne Drynda¹ (ProjectLeader), Ronny Otto¹ (DataManager), Wiebke Schirrmeister¹ (ProjectLeader), Jonas Bienzeisler² (DataManager), Alexander Kombeiz² (DataCurator), [Robert Koch-Institut | Fachgebiet 32](#), [Madlen Schranz](#)³ (Researcher), [Theresa Kocher](#)³ (Researcher)

¹[AKTIN-Notaufnahmeregister](#) | [AKTIN-Geschäftsstelle](#)

²[AKTIN-Notaufnahmeregister](#) | [AKTIN-IT](#)

³[Robert Koch-Institut](#) | [Fachgebiet 32](#)

Zitieren

Drynda, S., Otto, R., Schirrmeister, W., Bienzeisler, J., Kombeiz, A., Schranz, M., & Kocher, T. (2025). Open Data Sandbox [Data set]. Zenodo. <https://doi.org/10.5072/zenodo.171784>

Zusammenfassung

Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum Lorem Ipsum
Lorem Ipsum.

Inhaltsverzeichnis

- [Open Data Sandbox](#)
 - [Beispieltext](#)
 - [Hinweise zur Nachnutzung der Daten](#)
 - [Appendix](#)

Beispieltext

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem.

Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a,

Variablen und Werte

Die Datei [Sandbox_Data.tsv](#) enthält die in der folgenden Tabelle abgebildeten Variablen und deren Ausprägungen. Ein maschinenlesbares Datenschema ist im [Data Package Standard](#) in [tableschema_Sandbox_Data.json](#) hinterlegt:

[tableschema_Sandbox_Data.json](#)

| Variable | Typ | Ausprägungen | Beschreibung |
|-----------------------|--------|-------------------|---|
| LINEAGE | string | Beispiel: BA.2 | Zugewiesene Pangolin Lineage |
| WHO_LABEL | string | Beispiel: Omikron | Name der Virusvariante, der von der World Health Organisation vergeben wurde |
| CONTRIBUTING_LINEAGES | string | Beispiel: JN.13.1 | Pangolin Lineages, die von der Lineage abstammen |
| COLOR | any | | Veraltete Variable. Ist nicht mehr relevant und wird persepektivisch entfernt. |
| variant_category | string | Werte: VOC , VOI | WHO Einstufung der Variante als VOC (variant of concern) oder VOI (variant of interest) |

Die Datei [Sandbox_Data_lfs.tsv](#) enthält die in der folgenden Tabelle abgebildeten Variablen und deren Ausprägungen. Ein maschinenlesbares Datenschema ist im [Data Package Standard](#) in [tableschema_Sandbox_Data_lfs.json](#) hinterlegt:

[tableschema_Sandbox_Data_lfs.json](#)

| Variable | Typ | Ausprägungen | Beschreibung |
|---------------------|--------|---|--|
| igs_id | string | Beispiel: IGS-10099-CVDP-01A2C74B-54A8-47B1-B7E4-6562C6231234 | Ein eindeutiger Identifikator der Sequenzdaten und Metadaten zusammenführt. Dieser Identifikator wird als Teil der FASTA ID in den Sequenzdaten genutzt. |
| date_of_sampling | date | Format: YYYY-MM-DDTHH:MM:SS | Datum der Probeentnahme im ISO 8601 Format ohne Zeitzone |
| sequencing_platform | string | Beispiel: ILLUMINA | Die verwendete Sequenzierungs-Plattform auf Basis der von ENA zugelassenen Ontologie (siehe ena). |

| | | | |
|-----------------------------------|---------|---|---|
| sequencing_reason | string | Werte: random , requested , clinical , other | Grund für die Durchführung der Sequenzierung random : Die Probe wurde randomisiert genommen. requested : Die Probe wurde aufgrund von Bedenken/Verdacht auf eine neue Variante oder Vergleichbares genommen. clinical : Die Probe kommt aus einem klinischem Umfeld. other : Der Grund ist keiner der oben genannten. |
| isolation_source | string | Beispiel: Nasopharyngeal swab (specimen) | DEMIS Vokabular |
| lab_sequence_id | string | Beispiel: 873a7cc28d29e3f17b0544ea6e9e8436defe32f6d60649159ee8ac78d4147ac9 | Vom Labor genutzte FASTA ID in verschlüsselter Form |
| date_of_submission | date | Format: YYYY-MM-DDTHH:MM:SS | Datum des Eingangs des Genoms am RKI im ISO 8601 Format ohne Zeitzone |
| version | integer | Werte: ≥0 | Version der Sequenz startend mit 0 |
| prime_diagnostic_lab.demis_lab_id | string | Beispiel: DEMIS-10099 | Identifikationsnummer des primärdiagnostischen Labors |
| prime_diagnostic_lab.postal_code | string | Beispiel: 50858 | Postleitzahl des primärdiagnostischen Labors |
| sequencing_lab.demis_lab_id | string | Beispiel: DEMIS-10099 | Identifikationsnummer des sequenzierenden Labors |
| sequencing_lab.postal_code | string | Beispiel: 50858 | Postleitzahl des sequenzierenden Labors |
| | | Beispiele: [{ 'method': 'PANGOLIN_LATEST', 'classification_version': | |

| | | | |
|----------|--------|--|-----------------------------------|
| lineages | string | <pre>'PUSHER-v1.28.1', 'tool_version': '4.3', 'lineage': 'BA.2', '@qc_notes': 'Ambiguous_content:0.02', '@is_designated': False, '@qc_status': 'pass', '@conflict': 0.0, '@note': 'Usher placements: BA.2(1/1)'}]]</pre> | Pangolin Zuordnung im JSON-Format |
|----------|--------|--|-----------------------------------|

Metadaten

Zur Erhöhung der Auffindbarkeit sind die bereitgestellten Daten mit Metadaten beschrieben. Über GitHub Actions werden Metadaten an die entsprechenden Plattformen verteilt. Für jede Plattform existiert eine spezifische Metadatendatei, diese sind im Metadatenordner hinterlegt:

Metadaten/

Versionierung und DOI-Vergabe erfolgt über [Zenodo.org](https://zenodo.org). Die für den Import in Zenodo bereitgestellten Metadaten sind in der [zenodo.json](#) hinterlegt. Die Dokumentation der einzelnen Metadatenvariablen ist unter <https://developers.zenodo.org/#representation> nachlesbar.

Metadaten/zenodo.json

In der zenodo.json ist neben dem Publikationsdatum ("publication_date") auch der Datenstand in folgendem Format enthalten (Beispiel):

```
"dates": [
  {
    "start": "2023-09-11T15:00:21+02:00",
    "end": "2023-09-11T15:00:21+02:00",
    "type": "Collected",
    "description": "Date when the Dataset was created"
  }
],
```

Hinweise zur Nachnutzung der Daten

Offene Forschungsdaten des RKI werden auf [Zenodo.org](https://zenodo.org), [GitHub.com](https://github.com), [OpenCoDE](https://opencode.de) und [Edoc.rki.de](https://edoc.rki.de) bereitgestellt:


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

Lizenz

Der Datensatz "Open Data Sandbox" 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 Person 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.

Appendix

Dies ist ein Beispiel-Appendix .