



BONARES

BonaRes Series 2017/1 | DOI: 10.20387/BonaRes-E1AZ-ETD7

SVOBODA, N., HEINRICH, U.

The BonaRes Data Guideline

Soil as a Sustainable Resource
for the Bioeconomy - BonaRes
is a research programme funded
by the German Ministry of
Education and Research (BMBF).

www.bonares.de

SPONSORED BY THE



Federal Ministry
of Education
and Research



BONARES



SPONSORED BY THE

Federal Ministry
of Education
and Research

Imprint

Editor: BonaRes Centre for Soil Research
c/o Helmholtz -Centre for Environmental Science - UFZ
Department of Soil System Science
Theodor-Lieser-Str. 4 | 06120 Halle (Saale), Germany
Phone: (+49) 345 558 5226 | E-Mail: info@bonares.de |
www.bonares.de

Title	The BonaRes Data Guideline
Authors	<p>Svoboda, Nikolai - Leibniz-Centre for Agricultural Landscape Research, Agricultural Landscape Data Centre, Müncheberg;</p> <p>Heinrich, Uwe - Leibniz-Centre for Agricultural Landscape Research, Agricultural Landscape Data Centre, Müncheberg</p> <p><i>In close cooperation with the BonaRes Management Board:</i></p> <p>Bartke, Stephan - Helmholtz Centre for Environmental Research, Halle;</p> <p>Eberhardt, Einar - Federal Institute for Geosciences and Natural Resources, Hannover;</p> <p>Helming, Katharina - Leibniz-Centre for Agricultural Landscape Research, Müncheberg;</p> <p>Kögel-Knabner, Ingrid - Technical University of Munich, München;</p> <p>Russell, David - Senckenberg Museum of Natural History, Görlitz;</p> <p>Vogel, Hans-Jörg - Helmholtz Centre for Environmental Research, Halle;</p> <p>Wollschläger, Ute - Helmholtz Centre for Environmental Research, Halle</p> <p><i>Agreed by the BonaRes Scientific Board:</i></p> <p>Amelung, Wulf - University of Bonn, Bonn;</p> <p>Brüggemann, Nicolas - Jülich Research Centre, Jülich;</p> <p>Brunotte, Joachim - Thünen Institute, Braunschweig;</p> <p>Gebbers, Robin - Leibniz Institute for Agricultural Engineering and Bioeconomy, Potsdam;</p> <p>Geistlinger, Jörg - Anhalt University of Applied Sciences, Bernburg;</p> <p>Grosch, Rita - Leibniz Institute of Vegetable and Ornamental Crops, Großbeeren;</p> <p>Kiese, Ralf - Karlsruhe Institute of Technology, Garmisch-Partenkirchen;</p> <p>Leinweber, Peter - University of Rostock, Rostock;</p> <p>Lorenz, Marco - Thünen Institute, Braunschweig;</p>

Reinhold-Hurek, Barbara - University of Bremen, Bremen;

Sandhage-Hofmann, Alexandra - University of Bonn, Bonn;

Veldkamp, Edzo - Georg-August-Universität Göttingen, Göttingen;

Weiß, Stefan - Leibniz Universität Hannover, Hannover;

Winkelmann, Traud - Leibniz Universität Hannover, Hannover;

Zimmer, Dana - University of Rostock, Rostock

With support from the BonaRes Advisory Aoard:

Bristow, Keith - Commonwealth Scientific and Industrial Research Organisation, Canberra;

Brussaard, Lijbert - Wageningen University, Wageningen;

Folgart, Udo - The German Farmers' Association (Deutscher Bauernverband), Berlin;

Glante, Frank - Federal Environmental Agency, Dessau-Roßlau;

Grandke, Reinhard - German Agricultural Society, Frankfurt;

Hedlund, Katarina - Lund University, Lund;

Kunisch, Martin - Association for Technology and Structures in Agriculture, Darmstadt;

Montanarella, Luca - European Commission, Joint Research Centre, Ispra;

Scholten, Thomas - University of Tübingen, Tübingen;

Scow, Kate - University of California Davis, Davis;

Assisted by the BonaRes Data Centre:

Gärtner, Philipp - Leibniz-Centre for Agricultural Landscape Research, Müncheberg;

Grosse, Meike - Leibniz-Centre for Agricultural Landscape Research, Müncheberg;

Hierold, Wilfried - Leibniz-Centre for Agricultural Landscape Research, Müncheberg;

Hoffmann, Carsten - Leibniz-Centre for Agricultural Landscape Research,

Müncheberg;

Kühnert, Thomas - Leibniz-Centre for Agricultural Landscape Research, Müncheberg;

Lesch, Stephan - Senckenberg Museum of Natural History, Görlitz;

Schulz, Sina - Federal Institute for Geosciences and Natural Resources, Hannover;

Specka, Xenia - Leibniz-Centre for Agricultural Landscape Research, Müncheberg;

Stein, Susanne - Leibniz-Centre for Agricultural Landscape Research, Müncheberg;

Zoarder, Muqit - Leibniz-Centre for Agricultural Landscape Research, Müncheberg;

Correspondence

svoboda@zalf.de

Date

23. January 2017

Abstract

This data guideline provides rules for the handling, digital storage and free re-use of research data generated in the BonaRes program, while respecting the intellectual property right of the owner of the research data and respecting his or her right to the first scientific use, in accordance with the requirements arising from the BonaRes program allocation of funds. The Data Guideline aims at a) a timely submission of quality approved soil research data to the BonaRes Data Centre and its central storage b) facilitate the exchange of data among BonaRes scientists c) the provision of data to the scientific community.

Key words

BonaRes, policy, provision of research data, intellectual property right, license, embargo, DOI, re-use

Contents

Definitions	5
Purposes and goals of this Data Guideline.....	6
Endorsement	7
Quality management.....	7
Metadata	7
Content of research data	7
Structure of research data	8
Anonymizing of research data.....	8
Provision of research data to the Data Centre.....	8
Ownership and data rights	9
Embargo	9
Research data licensing	9
Review by the Data Centre.....	10
Digital Object Identifier (DOI).....	10
Updated and faulty research data.....	10
Long term data preservation and availability	10
Access to the research data	10
Terms of use and use statistics	11
Liability	11
Validity.....	11
Appendix.....	12
Der BonaRes Datenleitfaden: Deutsche Übersetzung des Dokuments „The BonaRes Data Guideline“	12

Definitions

Data owner: Person or institution with ownership rights to data and metadata. The data owner is not necessarily the data provider but can select the data provider as his/her deputy.

Data provider: Person in charge of providing data and metadata to the BonaRes Data Centre. Ideally, the data provider is assigned deputy for the data owner or his institution. The data provider is the most important point of contact for the BonaRes Data Centre.

Data user: Person using research data stored in and provided by the BonaRes Data Centre.

Embargo: Blocking period for the publication of data by the Data Centre on request by the data owner.

Metadata: Metadata are 'data about data' and contains information on who created the data how, necessary background information to the data, where it is located, characteristics of the file structure, intellectual property right (IPR) information etc. This descriptive information enables users to judge the usability of the data for their own purposes.

Set of data (dataset): At a minimum, the combination of data and associated metadata.

BonaRes decision making body: This body is responsible for decisions concerning the extension of an embargo and is made up of the head of the Data Centre (chairman) and one person from operational Management Board as well as 2 deputies of the Module A-Projects. In case of a stand-off situation, the decision is taken by the chairman.

General provisions




Purposes and goals of this Data Guideline

The purpose of this Data Guideline is to provide rules for the handling, digital storage and free re-use of research data generated in the BonaRes program, while respecting the intellectual property right of the owner of the research data and respecting his or her right to the first scientific use, in accordance with the requirements arising from the BonaRes program allocation of funds. The general directives and rules in this document apply to all people in touch with the BonaRes project related research data and intend to facilitate collaboration among the BonaRes Module A-Projects and the BonaRes Centre. Furthermore, this Data Guideline aims at a) a timely submission of quality approved soil research data to the BonaRes Data Centre (hereafter referred to as Data Centre) and its central storage b) facilitate the exchange of data among BonaRes scientists c) the provision of data to the scientific community in order to guarantee continuity of the research activities within BonaRes and beyond.




Shared research data are the vital backbone of the BonaRes project and science in general. The Data Centre provides a suitable infrastructure to facilitate the provision of BonaRes research data to the scientific community, thus paving the way for BonaRes scientists to increase the scientific impact of their data significantly.

The BonaRes community is encouraged to share all kinds of research data collected within the context of the BonaRes projects.





The BonaRes Data Guideline will


-  grant the right of first use and ensure the intellectual property right of research data,
-  take care of the existing ownership rights to shared research data and
-  specify the procedure of re-use of shared research data.

Further, this document will provide general directives in order to

-  maximize the (inter)national availability/ visibility/ impact of soil research data (by assigning digital object identifier (DOI) to the datasets, thus, making them citable and accessible for the long term),
-  facilitate the collaboration among the BonaRes community and
-  ensure timely submission of data to the Data Centre.

The BonaRes Data Guideline

-  takes the needs of the Module A-Projects into account
-  was discussed and developed together with international experts
-  was evaluated by the BonaRes Advisory Board
-  was discussed and improved with the BonaRes Management Board and the BonaRes Scientific Board

 was endorsed by the BonaRes scientists.

Endorsement

This Data Guideline as presented here specifies the rules of “good scientific practice” with respect to data management as they were pronounced by the “German Research Association”¹, the “Alliance of the German Science Organisations”² and established by the BonaRes Centre.

This Data Guideline is regarded as the implementation and operationalization of the framework of rules as it is described in the funding agreement provided by the Project Management Jülich (Projektträger Jülich: PTJ) and already agreed by the Module A-Projects. As such, it is of mere explicatory character and not a separate legal document.


Quality management

Metadata

All research data generated within the BonaRes projects must be described with metadata. The BonaRes metadata schema integrates the two most widespread standards INSPIRE³ and DataCite⁴. It is defined by the BonaRes Centre and can be supplemented with additional information by the data owner if necessary. The metadata contains, among other things, references to the author, the license, the location, the file structure, the reference to the BMBF as a source of funding, the quality inspections carried out and the correct citation of the data following the DataCite prescriptions⁵. Metadata are entered online using software provided by the Data Centre. Access to metadata is always unrestricted and provided with the Creative Commons license CC-0⁶.

Content of research data

The data owner is responsible for the quality⁷ of the data. When providing data to the BonaRes Data Centre, potential interests of third parties who could also work with the data in the future should be considered. For free re-use it is necessary to fully provide

 quality controlled research data (data that has passed quality assurance procedures) and

¹http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf

²http://www.allianzinitiative.de/fileadmin/user_upload/www.allianzinitiative.de/Grundsaeetze_Forschungsdaten_2010.pdf

³ Infrastructure for Spatial Information in the European Community.

⁴ An international non-for-profit organization with the aim of improving data citation.

⁵ Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., ... Bourne, P. E. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data*, 3, 160018. <http://doi.org/10.1038/sdata.2016.18>

⁶ Dedicated to public domain. <https://creativecommons.org/>

⁷ The rules of good scientific practice as published by the German Research Foundation (DFG), the German Association of University Professors and Lecturers (DHV), the Leibniz and the Helmholtz Association apply.

- 🌐 derived research data in terms of data products generated from raw, processed and quality controlled data through scientific interpretation.

Structure of research data

For storage, integration, provision, security and quality purposes research data must meet the basic requirements of information management like described formats, structured data and well defined data types⁸. The use of non-proprietary data formats is strongly recommended.

The Data Centre

- 🌐 may support with basic transformation procedures in order to provide (international) standard-compliant data (e.g. SI-units).
- 🌐 will change the syntax of data if necessary. This does not need acknowledgement of the data owner. The content of the data remains unchanged.
- 🌐 will notify the data provider if semantical changes are necessary, as far as they are discovered during the data processing at the Data Centre.

Anonymizing of research data

The data owner is responsible to adequately anonymize sensitive data before providing it to the Data Centre and to define suitable licenses for using the data. To not infringe on the rights of third parties, further steps of anonymization may be necessary. The BonaRes Centre can advise on questions of anonymization and offer technical solutions in close cooperation with the data owner.

Provision of research data to the Data Centre

Research data generated within the BonaRes projects are provided to and stored in the Data Centre. In principle, research data must be made available to the general public as early as possible.

The BonaRes Centre will make every effort to facilitate the successful provision. To provide the data owner with sufficient time to perform the quality control of the research data as well as to analyze dependent research data, such as time series, a delay of research data submission up to 12 months after research data generation is granted. In the case of journal publication, the research data linked to the publication are provided within two months to the Data Centre to be published with a separate DOI to allow re-use and to ensure an explicit reference to the author of the research data. The latest submission date for any data should be 6 months before the end of the relating project.

In case of specific research data (e.g., genomic data) that must/should be stored and provided for free re-use in external specialized repositories metadata for this research data must be provided to the Data Centre as well. These freely available research data are searchable and findable via the metadata provided by the Data Centre.

⁸ https://en.wikipedia.org/wiki/Data_type





Ownership and data rights

Intellectual property right⁹ of the data owner remains intact. The intellectual property right is part of the metadata and entered by the data owner or provider during metadata submission. This is an important prerequisite to permit this research data to be accessible through the Data Centre.

Embargo

The data owner can temporarily restrict the free access to their research data to secure the scientific right of first use. The embargo period starts with the provision of the research data to the Data Centre and is generally limited to 24 months, but may be extended if necessary.

Important reasons to justify an extension of the embargo are

-  dependent data, such as time series, must be analyzed coherently for a proper quality management,
-  applied patents (where research data are part of the algorithm),
-  ongoing publication processes (where research data are part of a paper under review) or
-  theses, which take longer to complete than planned.

In order to apply for an extension, it is sufficient to mark the data accordingly and to enter a brief justification and the required extension time. Subsequently this request will be reviewed by the Decision Making Body and the embargo will accordingly be extended for another 12 months as standard or for a period of time individually adapted to the respective needs. Further requests are possible.

With expiration of the embargo period, research data are made available following the specific Creative Commons License as selected by the data owner (see below). One month prior, the data provider is informed.

The data owner can always release embargoed research data earlier (i.e., to individual data users) for use. This requires a bilateral agreement with the data owner and data user. Corresponding arrangements are communicated to the Data Centre with the purpose of proper user management.

Research data licensing

The data owner states his intellectual property right for research data while selecting a proper Creative Commons (CC¹⁰) license model in the metadata. The standard license for all research data stored in the Data Centre is CC-BY¹¹ what is strongly recommended by the BonaRes Centre.

⁹ Intellectual property right (IPR) clarify the ownership of data and how others may use it. These rights guarantee that all legal, ethical, and professional obligations that data users may have to the owner of the data are met. See also <http://data.research.cornell.edu/content/intellectual-property>

¹⁰ <https://creativecommons.org/licenses/>

¹¹ The CC-BY license promotes the maximum dissemination of research data and thus increases their visibility and impact.

Review by the Data Centre

After a data set has been transferred to the Data Centre, the metadata are first checked for its completeness. Subsequently, the research data are checked for completeness and other quality features (consistency, gaps) to exclude possible transmission errors. Within the scope of the possibilities of the BonaRes Centre, the content is also checked. Distinctions are communicated to the data owner.



Digital Object Identifier (DOI)

The Data Centre offers to provide all stored research data with a DOI. As a precondition, the data has to be stored in the Data Centre. After a DOI has been assigned to research data, modifications are no longer permitted with the exception of adding further research data from current time series to existing research data. During the embargo period, a provisional URL of the DOI Landing Page is generated with which the research data can be made accessible to selected persons before the actual publication. If research data has already a DOI from a different repository, it is referred to this DOI in the metadata.

Updated and faulty research data

Research data that have been allocated a DOI name must not be altered.

In case of errors occur in the research data, data owners and data providers are obligated to report faulty research data to the Data Centre immediately. Depending on the severity of the errors, the following procedure is discussed with the data owner:




-  If possible, an updated or new version of the research data are provided and a new DOI is registered. New and old versions of the metadata are interlinked.
-  In the case that the whole research data associated with the DOI is faulty, and no update is possible, the research data must be deleted from the Data Centre. In this case the “old” DOI resolves to a so-called “tombstone landing page”. This page informs about the formerly available data and the reason of its deletion.

Long term data preservation and availability

To enable the maximum value from the “investment” of public funds made in the collection and provision of research data, it is to ensure that research data stored in the Data Centre is preserved and available over the long term, but at least for 10 years.

Access to the research data

Research data provided to the Data Centre is generally free for external use (CC-BY). Thus, any restrictions for the access to the research data arise from the chosen license (unlimited in time) or embargo (limited in time). Entry points to all data stored in the Data Centre are

-  unspecified search engines (like google),
-  the BonaRes portal (main gateway: www.bonares.de) or
-  specialized external portals through the standardized international research data infrastructure provided by the Data Centre.

Terms of use and use statistics

Prior to each use of the research data stored in the Data Centre, the data user must agree to the terms of use (i.a. license, IPR, proper citation). Depending on the access path, various options, such as downloading or online reading the usage agreement and activation of a check box, are implemented for this purpose.

Every access to the data stored in the Data Centre is recorded and communicated on request to the data provider in terms of a user statistic.

Liability

The BonaRes Centre acts exclusively as a manager and provider of stored research data and cannot be held responsible for the content or the availability of the research data.

Although every effort is made to ensure the quality of the data and metadata, neither the data providers nor the Data Centre can guarantee that no errors are contained. Therefore, research data provided through the Data Centre is made available without any warranty.

Validity

The BonaRes Data Guideline was endorsed on 23. January 2017.

Acknowledgement

This work was funded by the German Federal Ministry of Education and Research (BMBF) in the framework of the funding measure “Soil as a Sustainable Resource for the Bioeconomy – BonaRes”, project “BonaRes (Module B): BonaRes Centre for Soil Research, subproject B” (grant 031A608B).

License

This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) License. To view a copy of the license, visit <https://creativecommons.org/licenses/by/4.0/legalcode>.

Appendix

Der BonaRes Datenleitfaden: Deutsche Übersetzung des Dokuments „The BonaRes Data Guideline“

Definitionen

Dateneigentümer: Person oder Institution mit Eigentumsrechten an den Forschungsdaten. Der Dateneigentümer kann den Datenbereitsteller als seinen Stellvertreter auswählen. Der Dateneigentümer kann, muss aber nicht der Datenbereitsteller sein.

Datenbereitsteller: Die zur Bereitstellung von Forschungs- und Metadaten autorisierte Person oder Institution, welche nicht mit dem Dateneigentümer identisch sein muss. Der Datenbereitsteller ist für das Datenzentrum der wichtigste Ansprechpartner.

Datennutzer: Eine Person die Forschungsdaten, die im Datenzentrum gespeichert und veröffentlicht sind, verwendet.

Embargo: Sperrfrist für die Veröffentlichung von Daten durch das Datenzentrum auf Veranlassung des Dateneigentümers.

Metadaten: Metadaten sind beschreibende Daten, die zu allen Forschungsdaten gehören. In den Metadaten sind zum Beispiel Informationen über den Dateneigentümer, die Herkunft, die Methoden der Erhebung, die Datenstruktur und Angaben zu Lizenz und Nachnutzung enthalten. Metadaten sind für die wissenschaftliche Nachnutzung essenziell.

Datensatz: „Zusammenfassung von Daten, die in einer direkten Beziehung zueinander stehen oder gemeinsame Merkmale haben“ (itwissen.info); hier Forschungsdaten und Metadaten.

BonaRes Entscheidungsgremium: Dieses Gremium trifft Entscheidungen hinsichtlich der Verlängerung des Embargos und besteht aus dem Leiter des Datenzentrums, der als Vorsitzender fungiert, sowie einem weiteren Mitglied des operationellen Management Bords und zwei Vertretern der Modul A-Projekte. In einer Pattsituation wird die anstehende Entscheidung vom Vorsitzenden getroffen.

Allgemeine Bestimmungen




Ziele und Zweck dieses Datenleitfadens

Ziel dieses Datenleitfadens ist es, eine den Zuwendungsaufgaben für das BonaRes-Projekt gemäße, digitale Speicherung und freie Nachnutzung der im Projekt erzeugten Forschungsdaten unter Wahrung des Rechts des Dateneigentümers am geistigen Eigentum an seinen Daten und des Rechts auf ihre erste wissenschaftliche Nutzung sicherzustellen. Die Prinzipien in diesem Dokument gelten für alle Mitarbeiter des BonaRes Projektes und sollen die Zusammenarbeit zwischen den Modul-A Projekten und dem BonaRes Zentrum vereinfachen. Darüber hinaus soll dieser Datenleitfaden a) die zeitige Bereitstellung von qualitätsgesicherten Bodenforschungsdaten an das BonaRes Datenzentrum (im Folgenden: Datenzentrum) absichern, b) den Austausch von Daten zwischen den BonaRes Wissenschaftlern erleichtern und c) die Bereitstellung von Forschungsdaten für die wissenschaftliche Gemeinschaft sicherstellen, um eine Kontinuität der Forschung im BonaRes Projekt und darüber hinaus zu erreichen.




Freie und zugängliche Forschungsdaten sind nicht nur für das BonaRes Projekt eine wichtige Komponente, sondern befördern die gesamte Wissenschaft. Das Datenzentrum hält eine geeignete Infrastruktur vor, um BonaRes Forschungsdaten der Forschungsgemeinschaft zur Verfügung zu stellen und damit den wissenschaftlichen „Impact“ wesentlich zu erhöhen.

BonaRes Mitarbeiter stellen ihre Forschungsdaten, die im Rahmen des Projektes erzeugt werden, über das Datenzentrum zur allgemeinen Nachnutzung bereit.





Der BonaRes Datenleitfaden soll

-  das Erstverwertungs- und Urheberrecht an Forschungsdaten sicherstellen,
-  bestehende Rechte an Forschungsdaten, auch nach der Veröffentlichung durch das Datenzentrum, gewährleisten und
-  die allgemeine Nachnutzung der Forschungsdaten regeln.

Darüber hinaus werden in diesem Dokument grundsätzliche Regeln aufgestellt, um

-  die (inter)nationale Verfügbarkeit, Sichtbarkeit und den „Impact“ von Forschungsdaten zu erhöhen (durch die Vergabe eines „Digital Object Identifier“ (DOI) wodurch die Daten zitierbar und langfristig verfügbar werden),
-  die Zusammenarbeit zwischen den BonaRes Mitarbeitern zu erleichtern und
-  die zeitige Bereitstellung der Forschungsdaten an das Datenzentrum sicherzustellen.

Der BonaRes Datenleitfaden

-  berücksichtigt die Anforderungen der Modul A-Projekte,
-  wurde zusammen mit internationalen Experten diskutiert und entwickelt,
-  wurde vom BonaRes Advisory Board inhaltlich begutachtet,
-  wurde vom Management Board und Scientific Board diskutiert und vervollkommen und

 wurde von den beteiligten Wissenschaftlern gebilligt.

Billigung

Dieser Leitfaden spezifiziert die Regeln der „guten wissenschaftlichen Praxis“ mit Bezug auf das Datenmanagement wie sie von der Deutschen Forschungsgemeinschaft¹² (DFG) und der Deutschen Forschungsallianz¹³ beschrieben und vom BonaRes Zentrum implementiert wurden.

Dieser Leitfaden konkretisiert die im Zuwendungsbescheid des PTJ beschriebenen Regeln zum Umgang mit im BonaRes Projekt erhobenen Forschungsdaten; eine weitere Unterschrift ist nicht erforderlich.

Qualitätssicherung

Metadaten

Alle Forschungsdaten, die im Rahmen des BonaRes Projektes entstehen, müssen mit Metadaten beschrieben werden. Das BonaRes Metadatenschema vereint die wichtigen international gültigen Standards INSPIRE¹⁴ und DataCite¹⁵. Es wird vom Datenzentrum vorgegeben und kann durch den Datenbereitsteller, wenn notwendig, um weitere Einträge ergänzt werden. Die Metadaten enthalten, unter anderem, Angaben zu Urheber und Lizenz, zu Speicherort und Dateistruktur, den Hinweis auf das BMBF als Mittelgeber, die durchgeführten Qualitätsprüfungen und auf die richtige Zitierweise der Daten¹⁶. Metadaten werden durch eine vom Datenzentrum bereitgestellte Software online eingegeben. Alle Metadaten sind zur allgemeinen Verwendung unter der Lizenz CC-0¹⁷ freigegeben.

Inhalt und Umfang der Forschungsdaten

Der Dateneigentümer ist für die Qualität seiner Forschungsdaten¹⁸ verantwortlich. Bei der Veröffentlichung sollen auch Interessen Dritter berücksichtigt werden, die in der Zukunft mit den Daten arbeiten könnten. Für die freie Nachnutzung ist es notwendig,

¹²http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf

¹³

http://www.allianzinitiative.de/fileadmin/user_upload/www.allianzinitiative.de/Grundsaeetze_Forschungsdaten_2010.pdf

¹⁴ Infrastructure for Spatial Information in the European Community

¹⁵ Internationales Konsortium, das sich u.a. zum Ziel gemacht hat, den Zugang zu Forschungsdaten zu vereinfachen und deren Zitierung zu verbessern

¹⁶ Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., ... Bourne, P. E. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data*, 3, 160018.

¹⁷ Gemeinfrei. <https://creativecommons.org/>

¹⁸ Es gelten die Regeln der guten Wissenschaftlichen Praxis, wie sie unter anderem von der Deutschen Forschungsgemeinschaft (DFG), des Deutschen Hochschulverbandes (DHV), der Leibniz- und der Helmholtzgemeinschaft herausgegeben werden.

- 🌐 qualitätsgesicherte Forschungsdaten (Daten die einen Qualitätssicherungsprozess durchlaufen haben) und
- 🌐 prozessierte Forschungsdaten (Produkte aus Rohdaten, verarbeiteten und qualitätsgeprüften Daten als Ergebnis eines wissenschaftlichen Prozesses)

vollständig bereitzustellen.

Form der Forschungsdaten

Das Forschungsdatenmanagement im Datenzentrum (u.a. Speicherung, Integration, Bereitstellung, Sicherheit und Qualität) stellt an die Form der Forschungsdaten grundsätzliche Anforderungen (z.B. beschriebene Formate, Struktur und definierter Typ¹⁹). Die gewählten Datenformate sollten eine leichte Nachnutzung ermöglichen.

Das Datenzentrum

- 🌐 kann bei der Transformation von Daten unterstützen, damit diese internationalen Standards entsprechen (z.B. SI-Einheiten).
- 🌐 kann bei Bedarf die Syntax der Daten anpassen. Dies bedarf nicht der Zustimmung des Dateneigentümers. Der Inhalt der Daten bleibt dabei unverändert.
- 🌐 wird den Datenbereiter benachrichtigen, wenn inhaltliche Änderungen notwendig sind, soweit diese bei der Durchsicht der Daten auffallen.

Anonymisierung der Forschungsdaten

Es liegt in der Verantwortung des Dateneigentümers, sensible Daten vor der Bereitstellung an das Datenzentrum hinreichend zu anonymisieren und passende Lizenzen in den Metadaten zu definieren. Um nicht die Rechte Dritter zu verletzen, können weitere Schritte der Anonymisierung notwendig sein. Das BonaRes Zentrum kann bei Fragen der Anonymisierung beraten und in enger Kooperation mit dem Dateneigentümer technische Lösungen erarbeiten.

Bereitstellung der Forschungsdaten an das Datenzentrum

Forschungsdaten, die im Rahmen des BonaRes Projektes entstehen, werden dem Datenzentrum bereitgestellt und dort gespeichert. Grundsätzlich sollen im BonaRes Projekt erzeugte Forschungsdaten zugänglich der Öffentlichkeit zur Verfügung gestellt werden.

Das BonaRes Zentrum wird die erfolgreiche Bereitstellung von Forschungsdaten nach Kräften unterstützen. Abhängig von der jeweiligen Fragestellung, müssen zusammenhängende Forschungsdaten, wie zum Beispiel Zeitreihen, durch den Projektpartner vor der Bereitstellung an das Datenzentrum kohärent ausgewertet und in ihrer Qualität geprüft werden. Dafür ist ein Zeitraum von 12 Monaten vorgesehen. Wenn die Forschungsdaten Teil einer Publikation (Erstverwertungsrecht) sind, sollen sie innerhalb von 2 Monaten nach der Publikation dem Datenzentrum bereitgestellt werden, um mit einem DOI versehen und veröffentlicht zu werden. Auf diesem Wege wird die Nachnutzung ermög-

¹⁹ <https://de.wikipedia.org/wiki/Datentyp>

licht und ein fachgemäßes Referenzieren sichergestellt. Der späteste Zeitpunkt für die Forschungsdatenbereitstellung, sollte 6 Monate vor dem Ende des betreffenden Projektes sein.

Sollten/müssen bestimmte Forschungsdaten verpflichtend in speziellen Repositorien gespeichert und zur freien Nachnutzung bereitgestellt werden (z.B. Genomsequenzen), werden die betreffenden Forschungsdaten im Datenzentrum mit Metadaten beschrieben. Diese frei verfügbaren Forschungsdaten sind über die im Datenzentrum verfügbaren Metadaten such- und findbar.





Rechte und Eigentum

Das Recht am geistigen Eigentum²⁰ bleibt jederzeit unberührt. Die Beschreibung des Rechts auf geistiges Eigentum ist als Teil der Metadaten untrennbar mit den Forschungsdaten verknüpft. Auf diesem Weg ist das Recht auf geistiges Eigentum auch bei der späteren Nachnutzung immer für den Nutzer einsehbar. Eine entsprechende Kennzeichnung ist Grundvoraussetzung für eine spätere Nachnutzung der Forschungsdaten.

Embargo

Der Dateneigentümer kann den freien Zugang zu seinen Forschungsdaten zeitweise beschränken (Embargo), um das wissenschaftliche Erstverwertungsrecht auszuüben. Das Embargo kann mit der Bereitstellung der Forschungsdaten an das Datenzentrum beginnen und ist zunächst auf 24 Monate beschränkt; es kann bei Bedarf verlängert werden.

Mögliche Gründe für ein Verlängern des Embargos sind

-  lange kohärente Datenreihen, die während des Erhebungsprozesses nicht qualitätsgeprüft bereitgestellt werden können,
-  angemeldete Patente bei denen die Daten ein Teil des Patentes darstellen,
-  langwierige Publikationsprozesse bei denen die Daten ein Teil der Publikation sind oder
-  Abschlussarbeiten, die mehr Zeit in Anspruch nehmen, als vorgesehen.

Für die Verlängerung der Embargozeit wird unter Nennung des betreffenden Datensatzes, einer Begründung für die Verlängerung und der geplanten Dauer eine Anfrage an das Datenzentrum gestellt. Diese Anfrage wird anschließend vom BonaRes Entscheidungsgremium bearbeitet und das Embargo um 12 Monate (Standard) oder um einen an die individuellen Anforderungen angepassten Zeitraum verlängert. Weitere Anträge sind möglich.

Mit Ablauf des Embargos werden die betreffenden Forschungsdaten unter Beachtung der vergebenen „Creative Commons“ Lizenzen zur Nachnutzung freigegeben (siehe unten). Einen Monat vorher wird der Datenbereinsteller informiert.

Forschungsdaten, die mit einem Embargo belegt sind, können auch während des Embargos zur Nutzung (durch einzelne Datennutzer) freigegeben werden. Dies bedarf bilateraler Absprachen zwischen

²⁰ Das Recht auf geistiges Eigentum klärt die Eigentümerschaft an Forschungsdaten und wie diese durch andere Nutzer verwendet werden dürfen Siehe auch: <http://data.research.cornell.edu/content/intellectual-property>

Datennutzer und Eigentümer. Entsprechende Absprachen werden dem Datenzentrum mitgeteilt, um die Nutzer-Zugangsverwaltung dementsprechend zu konfigurieren.

Lizensierung der Forschungsdaten

Der Dateneigentümer erklärt in den Metadaten sein Recht auf geistiges Eigentum an den Forschungsdaten und wählt ein passendes Lizenzmodell aus. Es stehen verschiedene CC-Lizenzen²¹ zur Verfügung, von denen das Datenzentrum CC-BY²² als Standardlizenzmodell vorsieht. Diese Lizenz wird vom BonaRes Zentrum ausdrücklich empfohlen.

Review durch das Datenzentrum

Nachdem ein Datensatz dem Datenzentrum bereitgestellt wurde, werden zunächst die Metadaten auf ihre Vollständigkeit geprüft. Anschließend werden die Forschungsdaten auf eventuelle Übertragungsfehler überprüft. Das BonaRes Zentrum führt im Rahmen seiner Möglichkeiten weitere Qualitätsprüfungen (z.B. Konsistenz, Datenlücken, Inhalt) an den Forschungsdaten durch. Auffälligkeiten werden an den Datenbereinsteller kommuniziert.



Digital Object Identifier (DOI)

Das Datenzentrum bietet an, die über das Datenzentrum bereitgestellten Forschungsdaten mit einem DOI zu versehen. Als Grundvoraussetzung müssen die Daten im Datenzentrum gespeichert sein. Nachdem die Forschungsdaten mit einem DOI versehen wurden, sind Änderungen i.d.R. nicht mehr möglich, mit der Ausnahme, bestehende Forschungsdatensätze um weitere Forschungsdaten aus laufenden Zeitreihen zu ergänzen. Während der Embargozeit wird eine provisorische URL generiert, mit der ausgewählte Personen vor der eigentlichen Veröffentlichung auf die DOI „Landing Page“ geleitet werden können. Wenn bereits ein DOI aus anderen Repositorien vorhanden ist, wird dieser in den betreffenden Metadaten benannt.

Veränderte und fehlerhafte Forschungsdaten

Grundsätzlich sollen Forschungsdaten, die mit einem DOI versehen wurden, nicht mehr verändert werden.

Kenntnisse von Fehlern müssen umgehend vom Dateneigentümer oder Datenbereinsteller dem Datenzentrum gemeldet werden. In Abhängigkeit der Schwere der aufgetretenen Fehler, wird zusammen mit dem Dateneigentümer das weitere Vorgehen besprochen:

-  Werden Veränderungen notwendig, werden diese in einer neuen Version mit neuen Metadaten abgespeichert. Für die neue, korrigierte Version wird ein neuer DOI vergeben. Alter und neuer DOI werden über die Metadaten miteinander verbunden.
-  Wenn die Forschungsdaten fehlerhaft sind und keine Korrektur möglich ist, werden sie gelöscht. In diesem Fall verweist der „alte“ DOI auf eine so genannte „Tombstone Landing Pa-

²¹ <https://creativecommons.org/licenses/>

²² Die CC-BY Lizenz fördert die maximale Verbreitung der Forschungsdaten und erhöht damit deren Sichtbarkeit und Bedeutung.




ge“. Diese gibt Auskunft über die vormals vorhandenen Forschungsdaten und den Grund der Löschung.

Langzeitarchivierung und -verfügbarkeit

Um die maximale Wertschöpfung der Daten zu gewährleisten, die mit öffentlichen Geldern erhoben wurden, soll sichergestellt werden, dass alle Daten, die im Datenzentrum gespeichert werden langfristig gesichert und verfügbar sind, mindestens jedoch für 10 Jahre.

Zugang zu den Forschungsdaten

Grundsätzlich sind alle im Datenzentrum gespeicherten Forschungsdaten frei verfügbar (CC-BY). Die einzigen Einschränkungen für die Forschungsdaten ergeben sich aus dem gewählten Lizenzmodell (zeitlich unbefristet) und einem möglichen Embargo (zeitlich befristet). Zugang zu den im Datenzentrum gespeicherten Forschungsdaten ist möglich über

-  Suchmaschinen (z.B. Google),
-  das BonaRes Portal (Hauptzugang: www.bonares.de) oder
-  spezialisierte externe Portale, die über verschiedene international standardisierte Forschungsdateninfrastrukturen vom Datenzentrum bedient werden.

Nutzungsbedingungen und Nutzungsstatistik

Vor jeder Nutzung der im Datenzentrum gespeicherten Forschungsdaten, muss der Datennutzer den Nutzungsbedingungen (z.B. Lizenz, Recht auf geistiges Eigentum, Zitierweise) zustimmen. Für die Zustimmung sind verschiedene Möglichkeiten, wie zum Beispiel Download oder online Lesen der Nutzungsvereinbarung mit anschließender Bestätigung, vorgesehen.

Jede Nutzung der Forschungsdaten, die über das Datenzentrum bereitgestellt werden, wird protokolliert und dem Datenbereiter auf Anfrage in Form einer Statistik mitgeteilt.

Haftung

Das BonaRes Zentrum fungiert ausschließlich als Verwalter und Bereitsteller gespeicherter Forschungsdaten und kann weder für Inhalte noch für die Verfügbarkeit der Forschungsdaten verantwortlich gemacht werden.

Zwar werden jegliche Anstrengungen unternommen, um die Qualität der Daten und Metadaten zu gewährleisten, jedoch können weder die Datenbereiter, noch das Datenzentrum garantieren, dass keine Fehler enthalten sind. Daher werden über das Datenzentrum bereitgestellte Forschungsdaten ohne jegliche Gewährleistung zur Verfügung gestellt.

Gültigkeit

Dem Dokument „**The BonaRes Data Guideline**“ wurde am 23. Januar 2017 zugestimmt.

Danksagung

Diese Arbeit wurde gefördert vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen der Förderinitiative „Boden als nachhaltige Ressource für die Bioökonomie - BonaRes“, Projekt „BonaRes (Modul B): BonaRes-Zentrum für Bodenforschung, Teilprojekt B“ (Förderkennzeichen 031A608B).

Lizenz

Diese Arbeit unterliegt der Creative Commons Attribution 4.0 International (CC BY 4.0) Lizenz. Um eine Kopie der Lizenz anzuzeigen, besuchen Sie

<https://creativecommons.org/licenses/by/4.0/legalcode>.



BONARES

SPONSORED BY THE



Federal Ministry
of Education
and Research

BonaRes Series

The BonaRes Series publishes various formats related to the “BonaRes” funding initiative of the German Federal Ministry for Education and Research (BMBF). The BonaRes Series received funding in the framework of the funding measure “Soil as a Sustainable Resource for the Bioeconomy - BonaRes”, project “BonaRes (Module B): BonaRes Centre for Soil Research, subproject A” (grant 031A608A).

BonaRes is short for “Soil as a Sustainable Resource for the Bioeconomy”. The focus lies on the sustainable use of soils as a limited resource. BonaRes extends the evidence base for scientists and decision-makers regarding the soil systems to improve the productivity of soils and other soil functions while developing new strategies for a sustainable management of soils.

Learn more about BonaRes: www.bonares.de

Submit your manuscript for publication to: info@bonares.de



BONARES



SPONSORED BY THE

Federal Ministry
of Education
and Research