
**Biotechnology — Specification on
data management and publication in
microbial resource centers**

*Biotechnologie — Spécifications relatives à la gestion et à
la publication des données au sein des centres de ressources
microbiennes*

STANDARDSISO.COM : Click to view the full PDF of ISO 21710:2020



STANDARDSISO.COM : Click to view the full PDF of ISO 21710:2020



COPYRIGHT PROTECTED DOCUMENT

© ISO 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

	Page
Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Data publication in MRCs	5
4.1 General requirements and recommendations for data publication.....	5
4.2 Data fields of catalogue.....	6
5 Data management and data quality control	6
5.1 General.....	6
5.2 Requirements for in-house data management.....	6
5.3 Data management for the patent deposit of microbial material and associated data.....	7
5.4 Data management requirements for the transfer and traceability of microbial material and associated data.....	7
5.5 Unique identifier of microbial material and associated data.....	8
5.6 Data quality control.....	9
5.6.1 General procedures and principles.....	9
5.6.2 Personnel responsibility and authority of data management.....	10
5.6.3 Data content and format checking.....	10
5.6.4 Recommendations for improving data interoperability.....	10
Annex A (informative) Recommended data sets	12
Bibliography	24

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 276, *Biotechnology*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

The application of varied data forms in different microbial resource centres (MRCs) for in-house data management and publication is a major problem, which severely impedes efficient data exchange globally. It can also hinder depositors/users and potential users of MRCs from accessing their holdings and related information, and therefore impedes exploitation of microbial material and associated data in academia and bio-industries in the long run.

This document provides a set of data fields for data publication, aiming to improve data exchange of online catalogues among MRCs, by applying unique identifiers and a uniform data form. It also helps to facilitate subsequent access and benefit sharing based on microbial material and associated data.

This document also specifies requirements for data management and internal data quality control to improve the overall accuracy and reliability of data and information documented in MRCs, which is the basis of efficient data sharing and exchange.

STANDARDSISO.COM : Click to view the full PDF of ISO 21710:2020

[STANDARDSISO.COM](https://standardsiso.com) : Click to view the full PDF of ISO 21710:2020

Biotechnology — Specification on data management and publication in microbial resource centers

1 Scope

This document specifies requirements for data management and publication in microbial resource centres (MRCs) to enable consistent formatting, and a quality control workflow to improve the overall quality of data. It also provides recommendations for MRCs to improve data sharing and integration of microbial material and associated data.

This document is intended to facilitate procedures such as accessioning, acquisition, authentication, preservation, storage, and distribution, and can be used by MRCs, regulatory authorities, organizations, and schemes using peer-assessment to confirm or recognize the competence of MRCs in data management and publication

NOTE International, national and/or regional regulations or requirements can also apply to specific topics covered in this document.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 20387:2018, *Biotechnology — Biobanking — General requirements for biobanking*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia, available at <http://www.electropedia.org/>

3.1

accessioning

documenting the addition of a new *microbial material* (3.17) and/or associated data to an *MRC* (3.18)

[SOURCE: ISO 20387:2018, 3.1, modified — Replaced “biological” by “microbial”; “a biobank” was replaced by “an MRC”.]

3.2

acquisition

act of obtaining possession and/or custody of *microbial material* (3.17) and/or associated data

[SOURCE: ISO 20387:2018, 3.2, modified — Replaced “biological” by “microbial”.]

3.3

associated data

any information affiliated with *microbial material* (3.17) including but not limited to collection, taxonomic, deposit history and provider data

[SOURCE: ISO 20387:2018, 3.3, modified — Replaced “biological” by “microbial”, the coverage in the inclusion was modified to data management related categories.]

3.4

authentication

process by which *microbial material* (3.17) is characterized to a defined level of specification using appropriate technology/documentation to establish a conclusive basis for accepting the material as genuine

[SOURCE: ISO 20387:2018, 3.4, modified — Replaced “biological” by “microbial”.]

3.5

catalogue

systematically arranged list or record often including descriptive information

[SOURCE: ISO 20387:2018, 3.10]

Note 1 to entry: It can include printed and online catalogues.

3.6

deposit

act of transferring possession and/or custody of *microbial material* (3.17) and/or associated data from *depositor* (3.7) to *MRC* (3.18)

3.7

depositor provider

person or entity from whom/which the *microbial materials* (3.17) and/or associated data is received or acquired

[SOURCE: ISO 20387:2018, 3.41, modified — Replaced “biological” by “microbial” and deleted “for biobanking” at the end.]

3.8

distribution

process of providing selected *microbial material* (3.17) and/or associated data to *recipient(s)* (3.24)

[SOURCE: ISO 20387:2018, 3.20, modified — Replaced “biological” by “microbial”; “/user(s)” was deleted.]

3.9

interoperability

capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units

[SOURCE: ISO/IEC 20944-1:2013, 3.6.1.24]

3.10

IRCC

international recognized certificate of compliance

permit or equivalent document issued at the time of access as the evidence that the *microbial material* (3.17) and associated data has been accessed in accordance with the decision to grant *PIC* (3.22), and that *MTAs* (3.19) have been established for the user and the utilization

Note 1 to entry: When available on the ABS Clearing-House of the Nagoya Protocol, an IRCC number can be attributed, which can be used as a reference number.

Note 2 to entry: This term is based on the Nagoya Protocol.

3.11

legacy microbial material and associated data

microbial material (3.17) and associated data acquired or received by the *MRC* (3.18) before the *MRC* has implemented this document

3.12

life cycle of data

all stages in the process of data usage from initial establishment to its discontinuation

3.13

MAA

material accession agreement

material acquisition agreement

contractual document between a *depositor* (3.7) and an *MRC* (3.18) in a material transfer process

Note 1 to entry: It documents basic data such as place and date of sampling in a standardized form, and specifies the role, rights, and duties of the depositor and *MRC*.

Note 2 to entry: *MAA* is synonymous to material deposit agreement (*MDA*). It is normally put in place by an *MRC*.

3.14

MAT

mutually agreed terms

contractual arrangements between a *provider* (3.7) of *microbial material* (3.17) and/or associated data, and a *recipient* (3.24)

Note 1 to entry: It sets out specific conditions for the fair and equitable sharing of benefits arising from the utilization of microbial materials and/or associated data, and that can also include further conditions and terms for such utilization as well as subsequent applications and commercialization.

3.15

MDS

minimum dataset

collection of technical and scientific data digitized in specific fields of a database, which is necessary to distinguish unambiguously a particular *microbial material* (3.17) and provides a minimum amount of information available for each accession in an *MRC* (3.18)

Note 1 to entry: Microbial materials, except for legacy microbial materials, cannot be inserted into the catalogue, if this information is not available.

3.16

metadata

data that defines and describes other data

[SOURCE: ISO/IEC 11179-1:2015, 3.2.16]

3.17

microbial material

any substance(s) derived or part obtained from a microorganism, and any complexes or association between microorganisms

Note 1 to entry: It comprises all prokaryotes (archaea and bacteria), some eukaryotic organisms (fungi, yeasts, algae, protozoa), any association between the latter (e.g. lichens), non-cellular entities (e.g. viruses), their replicable parts and other derived materials (e.g. genomes, plasmids, cDNA). It also includes some viable but not yet culturable microorganisms.

3.18

MRC

microbial resource centre

microbial biobank

microbial BRC

legal entity or part of a legal entity that performs biobanking of *microbial material* (3.17) and associated data

3.19

MTA

material transfer agreement

contractual documents between an *MRC* (3.18) and a *recipient* (3.24) in a material transfer process

Note 1 to entry: All the documents can be designated as MTA as long as they contain information about the *in situ* origin or the source of the microbial material and associated data, information about the provider and recipient, and information that defines the limits of the use of the microbial material and associated data

Note 2 to entry: An MTA can also be associated with a microbial material being deposited to meet the need of its depositor country/country of origin, particularly those that are the parties of the Convention of Biological Diversity (CBD) and Nagoya Protocol (NP).

3.20

patent deposit

service of MRCs for long-term preservation of *microbial materials* (3.17) for the purposes of patent procedure

Note 1 to entry: It is a special service of an MRC recognized as the International Depositary Authority (IDA) by the World Intellectual Property Organization (WIPO) for the purposes of patent procedures under the Budapest Treaty.

3.21

persistent identifier

unique identifier (3.27) that ensures permanent access for a digital object by providing access to it independently of its physical location or current ownership

EXAMPLE [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8) is an example of a persistent identifier.

Note 1 to entry: It is used in the context of digital objects that are accessible over the Internet. Typically, such an identifier is not only persistent but also actionable (e.g. URI).

[SOURCE: ISO 24619:2011, 3.2.4, modified — note 1 to entry was entirely replaced to describe the usage of this term.]

3.22

PIC

prior informed consent

permission of the country of origin to a user to access *microbial materials* (3.17) and/or associated traditional knowledge

Note 1 to entry: This term is based on the Nagoya Protocol.

3.23

preservation

act or activity to prevent or retard microbial or physical deterioration of *microbial material* (3.17)

[SOURCE: ISO 20387:2018, 3.34, modified — Replaced “biological” by “microbial”.]

3.24

recipient

user

person or entity to whom/which the *microbial material* (3.17) and/or associated data is distributed

[SOURCE: ISO 20387:2018, 3.44, modified — Replaced “biological” by “microbial”.]

3.25**safe deposit**

service of MRCs for long-term preservation of *microbial materials* (3.17) with distribution restrictions at the discretion of the *depositor* (3.7)

Note 1 to entry: All information concerning the *deposit* (3.6) and the nature of the microbial materials are treated in the strictest confidence. Access to this type of microbial material is permitted only by documented request of the depositor.

3.26**storage**

maintenance of *microbial material* (3.17) under specified conditions for future use

[SOURCE: ISO 20387:2018, 3.47, modified — Replaced “biological” by “microbial”.]

3.27**unique identifier**

code that is associated with a single entity within a given system

Note 1 to entry: Such an identifier can establish an unambiguous relationship between each microbial material and its associated data.

[SOURCE: ISO 20387:2018, 3.50]

Note 2 to entry: It can include accession number and other local identifiers (local IDs) in the web-based context.

3.28**URI****Uniform Resource Identifier**

compact sequence of characters that uniquely identify an abstract or physical resource

EXAMPLE http://gcm.wdcm.org/Strain_numberToInfoServlet?strain_number=LMG%2026135 is an example of an URI in the Global Catalogue of Microorganisms database hosted by WDCM.

Note 1 to entry: URI is a kind of *unique identifier* (3.17). It is used in the field of computer science.

[SOURCE: IETF RFC 3986^[14]]

4 Data publication in MRCs**4.1 General requirements and recommendations for data publication**

4.1.1 The MRC shall develop, make available and frequently update online catalogues to keep pace with the in-house database and display modifications of information in a timely manner.

4.1.2 The MRC shall disclose data associated with the microbial material in the catalogue in accordance with the agreed terms and conditions presented in the applicable documents (e.g. MTA, MAA) or other contracts, if any. It is presupposed that these documents are in accordance with applicable statutory and regulatory requirements.

4.1.3 The MRC shall not disclose sensitive and other restricted information regarding safe deposit and patent deposit in any form as specified by relevant contracts, if any.

4.1.4 When the depositor requests a delay in the publication of the associated data, the MRC shall enforce an agreed timeframe and conditions, and may release the data into the catalogue after it is expired.

4.1.5 The catalogues of MRCs shall contain indispensable service information including ordering, shipping information, and basic template documents such as forms and agreements to comply with

requirements associated with the handling and shipping of microbial materials, which should enable the efficient use and be user-friendly.

4.1.6 Public data of microbial materials documented in the catalogue of MRCs should be shared to public databases such as GenBank <https://www.ncbi.nlm.nih.gov/genbank/>, Global Biodiversity Information Facility (GBIF) <http://www.gbif.org/>, Catalogue of Global Genome Biodiversity Network (GGBN) http://www.ggbn.org/ggbn_portal/, WFCC Global Catalogue of Microorganisms (GCM) <http://gcm.wdcm.org/>, if appropriate.

4.1.7 The online catalogue of MRCs should facilitate search capacity across community databases to support extensive searches for multiple characteristics of microbial materials and improve diversity of information.

4.2 Data fields of catalogue

4.2.1 The MRC shall properly arrange an accession form for online submission to deposit a microbial material and associated data efficiently, and it should apply a machine-readable tabular format, if applicable.

4.2.2 The MRC should ensure the availability of an MDS, and the MDS should include all the following data fields given in [A.2.1](#): a) accession number, c) organism type, d) status, e) biosafety and biosecurity, f) other restrictions, and [A.2.2](#): c) history of deposit. The MRC should select all or part of the optional data fields according to specific circumstances as specified in [Annex A](#).

4.2.3 The MRC should maintain a high quality and cross-referenced online catalogue to link the data to publicly available controlled vocabularies specified in [Annex A](#) and the authorized databases in accordance with [4.1.6](#) and [4.1.7](#).

4.2.3.1 The MRC should link each name of microbial material with publicly available nomenclature databases to facilitate data retrieval and accessing. The MRC shall indicate the changes of nomenclature in the printed catalogues.

4.2.3.2 The MRC should link each sequence (e.g. genomic, metagenome, cDNA, RNA and protein sequences) to publicly available databases.

4.2.3.3 The MRC should link bibliographic information including digitized literature and patent to publicly available databases and document the DOI numbers, if applicable.

5 Data management and data quality control

5.1 General

The data management of MRC shall be in accordance with ISO20387:2018, 7.10.

5.2 Requirements for in-house data management

5.2.1 The MRC shall establish, implement and maintain a data management system to ensure that the essential information is available in the database, e.g. the storage period and location of microbial materials, frequency of data update and backup, and review and inspection actions for the different categories of data.

5.2.2 The MRC shall establish and maintain a security system and implement a procedure for the management of data and related documentation to prevent access by unauthorized users, ensuring

information security, bio-security, protection of Intellectual Property Rights (IPRs) and depositor/user information.

5.2.3 The MRC shall establish and implement a strategic plan for the sustainability of the in-house data, especially critical details of operation in the preservation process by regular data quality audit, in order to fulfil traceability of microbial material and associated data, as well as potential risks.

5.2.4 The MRC shall establish, document and implement adequate data quality control (QC) procedures to reduce the risk of data inaccuracy and unexpected loss, in accordance with the QC specified in ISO20387:2018, 7.8.3., 7.8.2.8, and 7.2.8.9

5.3 Data management for the patent deposit of microbial material and associated data

5.3.1 Patent deposit related information shall be available exclusively to the depositor, and shall by no means be included in catalogues or any publicly available lists without a prior documented consent of the depositor.

5.3.2 The MRC may partly disclose patent related information including patent name, patent number and accession number in accordance with [4.1.2](#) to facilitate accessing of the patent deposited materials.

5.3.3 The MRC should document the MDS as specified in [4.2.2](#) and [A.10](#) c). It should document the essential information for viability testing and storage in the database.

5.3.4 The MRC shall establish, implement, and document a procedure for the handling of the patent deposit of microbial material and associated data. It is presupposed that procedures are developed in accordance with applicable statutory requirements.

NOTE For the parties to the Budapest Treaty, the required documents include the accession form (BP/1), official notifications to the depositor including the mandatory international form (BP/4), official receipt viability statements (BP/9) and notification of release to a third party (BP/14).

5.3.5 The MRCs should document major changes of patent deposit including the name of the authorized representing person, transfer of rights, fate of the culture after the prescribed duration of storage, as well as the dates when the microbial material was received, returned, accepted, and the date when it was validated, stored, renewed, withdrawn and destroyed in the database.

5.4 Data management requirements for the transfer and traceability of microbial material and associated data

5.4.1 The MRC shall ensure persistent identification of a microbial material and traceability of associated data in accordance with ISO 20387:2018, 7.5.

5.4.2 The MRC within jurisdiction of a CBD and NP Party ¹⁾ shall document the transfer related information of microbial material and associated data and link it to the related documents (e.g. PIC, IRCC, MTA or MAT) in the database, and introduce the administrative, technical and research output data in the catalogue.

a) Administrative data:

- 1) authorization of sampling and/or utilization (e.g. PIC);
- 2) scope of use;

1) Nagoya Protocol (NP) parties are states that are the parties to the Convention on Biological Diversity (CBD) signed Nagoya Protocol. The details of NP parties can be found in <https://www.cbd.int/abs/nagoya-protocol/signatories/>.

- 3) status of legislation and related regulations;
 - 4) MTAs.
- b) Technical data:
- 1) information about the in situ origin or the source of the microbiological material;
 - 2) information about the provider and recipient;
 - 3) information of collecting, including place, date of sampling, name of collector, and substrate;
 - 4) risk assessment of the microbial material including risk group/biohazard level, restrictions on distribution or safety precautions;
 - 5) taxonomical data and the reference;
 - 6) information necessary to ensure the long-term preservation of the microbial material including growth temperature, media composition, and specific growth conditions.
- c) Research output data:
- 1) information about modifications and unmodified derivatives of the material and associated data;
 - 2) scientific papers, patents, and any commercial results.

5.5 Unique identifier of microbial material and associated data

5.5.1 The MRC shall implement a unique identifier system in accordance with [5.5.1.2](#) to ensure efficient retrieval and access to the microbial material associated data.

NOTE The accession number of the legacy microbial material and associated data can be regarded as a unique identifier, only if the MRC can confirm its uniqueness among all the publicly available information in the context of life science and biotechnology. For example, an MRC can combine a unique acronym assigned when it registered in WDCM CCInfo database (<http://www.wfcc.info/ccinfo/>) with its local identifier to form a unique accession number.

5.5.1.1 The MRC should use and refer to pre-existing persistent identifiers (e.g. DOI) in priority.

NOTE Globally unique identifiers assigned by authoritative international organization (e.g. WDCM) and/or schemes can serve as persistent identifiers in the MRCs.

5.5.1.2 Unique identifiers shall be implemented, documented and maintained in accordance with ISO 20387:2018, 7.5.1 a).

5.5.2 The MRC shall issue a unique identifier in accordance with [5.5.1.2](#) to a microbial material at the time of deposit, and it shall take a long-term view to decide the scope and life cycle of data and the identifiers they have issued.

5.5.3 The MRC should document the procedure for assigning and maintaining the unique identifiers.

5.5.3.1 The change history of a unique identifier issued by the MRC shall be documented and/or retrievable.

5.5.3.2 The MRC shall not delete a unique identifier or reassign it to a different microbial material. The MRC should maintain a list of the obsoleted unique identifiers issued.

NOTE Obsoleted unique identifiers are outdated unique identifiers that will no longer be used.

5.5.4 The MRC, which also functions as an identifier practitioner (database provider), should implement a web-based identifier strategy to improve data sharing and interoperability of microbial material and associated data.

5.5.4.1 The MRC should issue persistent identifiers which can be transparently referenced and actionable (e.g. in a public index or search) anywhere, by anyone, and for any reason, and it should use the existing identifier platforms (e.g. Identifiers.org, W3ID) and services where appropriate.

5.5.4.2 The MRC can convert the local identifiers (e.g. accession number, isolate number) of a microbial material and associated data by adding prefix to persistent identifiers resolvable to a web address where the data or information about the entry can be accessed (e.g. URI).

5.5.4.3 The issued web-based unique identifiers, which can work in a web browser, should facilitate machine-readable mapping between identifiers including pre-existing ones.

EXAMPLE http://gcm.wdcm.org/Strain_numberToInfoServlet?strain_number=ICM%202499.

5.5.4.4 The MRC should present the cross-linked persistent identifiers in the online catalogue. The MRC should promote the citation of the issued unique identifiers by the users of microbial material and associated data in publications (e.g. scientific papers, patents) and other online platforms.

5.5.4.5 The MRC should document and cross-reference the unique identifiers involved in the microbial material and associated data transfer in the database to link the sample collected in situ to the derived microbial materials deposited in the MRC.

5.5.4.6 The MRC should implement a persistent identifier system, which can also facilitate to track and monitor the use of microbial material and associated data in both commercial and non-commercial activities, to meet the needs and expectations of all the parties to different agreements (depositors and recipients, government agencies, commercial concerns, etc.).

NOTE Three interconnected references are required to fulfil data concatenation, if the country is a party to the NP and has access legislation:

- a) unique identifier for the biological or ecological sample.
- b) unique identifier for each strain “extracted” from the sample.
- c) unique identifier for the PIC or for the IRCC related to the PIC.

5.6 Data quality control

5.6.1 General procedures and principles

5.6.1.1 The MRC shall produce the catalogues based on authenticated and validated information. It shall establish a documented procedure that involves all the major operations of the database including data updates, backup, cancellation, and recovery.

5.6.1.2 The MRC should digitize physical documents and archive all the records regularly to keep consistent with the database records.

5.6.1.3 The MRC should retain the information relating to a microbial material indefinitely and identify the microbial materials that are no longer available as a living material.

5.6.2 Personnel responsibility and authority of data management

5.6.2.1 The MRC shall keep detailed records of the history of modifications. The MRC should designate properly qualified staff for data inputs, validations, corrections and deletions.

5.6.2.2 Restricted and confidential data regarding safe deposits or patent deposits shall only be accessible to authorized staff, and the related data management operations should be under the control of the security manager.

5.6.2.3 The MRC should introduce appropriate measures (protocols, tools and standards) in their own IT systems to ensure reasonable security of information, such as authentication by user ID and password, encryption, encryption of messages and restriction of IP addresses, or other evolving technologies.

5.6.3 Data content and format checking

5.6.3.1 The MRC shall maintain documented information about the quality control workflow and plan for at least five years to ensure the validity of the data, e.g. the audit date, next scheduled audit date, scope of audit, audit content, auditor, responsible person, responsibility and authority, method, comments, results and action, signature.

5.6.3.2 The MRC should establish, implement and maintain a procedure for spell checking during the data input process, in order to correct typographical errors and other data content errors by manual and/or automatic validation methods. It is recommended to establish a thesaurus/ontology to check the correctness of the input if applicable.

5.6.3.3 The MRC should also adopt procedures to detect errors in both new data and pre-existing data in current databases to improve their quality and consistency.

5.6.3.4 The MRC shall select the data format and data representation according to the specifications in [4.2](#) to ensure the quality and consistency of data sets. The date and time documentation should be formatted in accordance with ISO 8601-1, ISO 8601-2 and ISO 20387:2018, 7.1.3.

NOTE The date can be expressed as YYYY-MM-DD (e.g. 2018-04-25) and time can be expressed as hh:mm:ss (e.g. 04:26:55).

5.6.3.5 The MRC should ensure the accuracy of taxonomic/scientific names of microbial materials by linking the data field of name to publicly available taxonomic databases or rechecking against the recognized source of valid taxonomic names in accordance with the related international nomenclature codes [e.g. International code of nomenclature for algae, fungi and plants (ICN), International code of nomenclature of prokaryote (ICNP)].

5.6.4 Recommendations for improving data interoperability

5.6.4.1 The MRC should distribute the microbial material associated (meta)data in standardized formats in accordance with the most relevant standards for data exchange and integration to ensure robustness and timeliness of data sharing, as well as standardized protocols to provide data in human-readable and machine-readable formats to facilitate interoperating with various bioinformatics tools.

5.6.4.2 A list of all the names that have been used for the same material, such as vernacular names, synonyms, common names, misapplied names and misspellings, should be documented as metadata and easily accessible to the users.

5.6.4.3 For the purpose of data exchange and integration, the MRC should ensure the data field of accession number (see [Annex A](#)) is self-resolvable and able to function as a passport for a certain microbial material.

5.6.4.4 According to relevant principles (e.g. FAIR Guiding Principles^[18]), MRCs should use unified or compatible data formats such as controlled vocabularies if applicable for data management and exchange.

5.6.4.5 The MRC should select standardized data schema and protocols to maximize the chance of interoperability between databases.

5.6.4.6 The MRC should use a formal, accessible, shared, and broadly applicable language for (meta) data. English should be chosen as a preferred language of data (in addition to the local language if different). A standardized approach should be adopted for certain scientific symbols.

5.6.4.7 The MRC should ensure the (meta)data of microbial material are retrievable by unique identifiers using an appropriate communication protocol and ensure (meta)data include qualified references.

STANDARDSISO.COM : Click to view the full PDF of ISO 21710:2020

Annex A (informative)

Recommended data sets

A.1 Name and taxonomy of microbial material

a) **Name:**

Data Type: String or link

Description: The current scientific name of the species contains genus name and species epithet, or at least the taxonomic name indicated the lowest taxonomic category that can be identified at present, in accordance with the recommendations specified in 4.2.2, or it can include a link to the authorized taxonomic databases for reference. Exceptions can be justified for microbial materials deposited as species complexes.

Example 1: Rhodosporidium toruloides

Example 2: pFBAOT6

Example 3: Paenibacillus sp

1) **Subspecies name:**

Data Type: String or link

Description: Contains subspecies epithet and (if different from species epithet) author(s).

2) **Infrasubspecific names:**

Data Type: String or link

Description: Contains variety, designation, epithet, authors and reference. Does not apply to all strains.

3) **Variety name:**

Data Type: String or link

Description: Variety name, author(s).

4) **Former name:**

Data Type: String or link

Description: Name, author(s). Contains former epithet and (if different from species epithet) author(s).

5) **Special former name:**

Data Type: String or link

Description: Name, author(s). Contains special former epithet and author(s).

6) **Misapplied name:****Data Type:** String or link**Description:** Name, author(s). Names that have been wrongly applied to the current species, and may also be correctly applied to another species.b) **Naming history:****Data type:** Text or links**Description:** Taxonomic name, date, the reasons for modifications, and who did the modifications.

NOTE It can include subfields including synonym, misspelling, common name, basionym, misapplied name to indicate different name types of former names.

c) **Taxonomy level:****Data type:** Controlled vocabulary**Description:** All the taxonomic levels in which the microorganism is classified, including kingdom, phylum (or division), class, order, family, genus, species, subspecies and all the other intermediate levels.d) **Taxonomy ID:****Data Type:** String or links**Description:** An identifier for the set of taxon information.e) **Serotype:****Data Type:** String**Description:** Serotype name and author(s). Required for some medically important species.**A.2 Strain administration****A.2.1 General**a) **Accession Number:****Data Type:** Mandatory, string**Constraint:** Not null**String Description:** The unique acronym of an MRC followed by a delimiter and a unique number to identify a microbial material.**Example:** JCM 10021b) **Other Collection Numbers:****Data Type:** String**Description:** Accession number registered in other MRCs regarded as having the same biological origin, separated by semicolon. Equivalent numbers of the element in other MRCs under which the element is distributed.**Example:** ATCC 10657; ATCC 18533; CBS 349; IAM 12256; IFO 0880; JCM 10301; NBRC 0880; NRRL Y-1588

c) **Organism Type:**

Data Type: controlled vocabulary

Constraint: Not null

Description: Recommended values can be Fungi, Yeast, Bacteria, Archaea, Microalgae, Cyanobacteria, Protozoa, Plasmid, Phage, Virus, cDNA.

Example: Yeast

d) **Status:**

Data Type: controlled vocabulary

Constraint: Not null

Description: Nomenclatural status of the strain. Value “no”, or a text field for taxonomic type strain of the species following the type definitions based on the International Code of Botanical Nomenclature (ICBN) (Chapter II, Section 2, Article 9). Does not apply to all strains.

Example1: Type of *Candida albicans*

Example2: Ex-neotype

Example3: Ex-type

Example4: Ex-lectotype

Example5: Ex-holotype

Example6: Ex-isotype

e) **Biosafety and biosecurity:**

Data Type: String

Constraint: Not null

Description: Precautions for internal policy and restrictions for distribution, import and export of cultures, required by the pathogenicity class (WHO standard, WHO Laboratory biosafety manual Third edition Geneva 2004 and/or national standard) corresponding regulations.

Example 1: BSL2

Example 2: National standard for biosecurity 3

Example 3: In relation to plant quarantine.

Example 4: Transfer only in biohazard laminar flow cabinet

f) **Other Restrictions:**

Data Type: string

Constraint: Not null

Description: Precautions for internal policy and restrictions for distribution, import and export of cultures for a certain country.

Example 1: Not to be released to unauthorized recipients.

Example 2: There are NO known Nagoya protocol restrictions for this strain.

Example 3: Commercial use and transfer to third parties is completely excluded for this strain.

g) **License:**

Data Type: Link to documents

Description: A legal document giving official permission to carry out certain activities with the microbial material and associated data. For NP or just CBD contracting parties, it includes PIC/MAT, MAA, and MTA documents.

h) **Country of origin:**

Data Type: Controlled vocabulary

Description: Country from which the microbial material and associated data was obtained. Country or sea names should be chosen from the INSDC country list (<http://insdc.org/country.html>), or the GAZ ontology (v 1.512) (<http://purl.bioontology.org/ontology/GAZ>).

NOTE It can be mandatory for MRC of NP parties.

i) **Access rights:**

Data Type: Controlled vocabulary

Description: Information about who can access the microbial material and associated data or an indication of its security status. Access Rights may include information regarding access or restrictions based on privacy, security, or other policies.

A.2.2 Deposita) **Depositor:**

Data Type: String

Description: A list (concatenated and separated) of names of people, groups, or organizations who deposit a microbial material in the MRC.

b) **Date of deposit:**

Data Type: Date

Description: The date the MRC receives a microbial material and its associated completed accession form, and uses the format [YYYY]-[MM]-[DD] if applicable.

c) **History of deposit**

Data Type: String

Constraint: Not null

Description: History of the transfer of the microbial material and associated data from isolator to depositor, including name of isolator, institute of isolator, the name of depositor and the institute of depositor.

Example 1: IFO 0880 < – NI (K. Okunuki)

Example 2: < LMG 1625 < – NCIB 8894 < – R. H. Parsons.

A.2.3 Preservationa) **Remarks on validation:**

Data Type: String

Description: The date, methods (e.g. microscopy, DNA barcoding and sequencing) and results of the accuracy of strain designation, viability and purity of the microbial material.

b) **Date of accession:**

Data Type: String

Description: The date the MRC accepted a microbial material for preservation, using the format [YYYY]-[MM]-[DD] if applicable.

c) **Conditions for storage:**

Data Type: String

Description: Storage conditions and special care needed including temperature (e.g. -80°C), light source, intensity, periodicity, and cryoprotectant.

Example: Requires DMSO for liquid nitrogen (10 % glycerol)

d) **Remarks on storage:**

Data Type: String

Description: Remarks on the general ability to survive, sporulate, etc. after accession.

Example: Does not endure storage at 4°C

A.2.4 Distribution

a) **Form of supply:**

Data Type: Controlled vocabulary

Description: Form in which the microbial material is available and will be sent to the recipient.

Example 1: Lyophilized

Example 2: Active culture

NOTE It can include data fields including information on the purchase form, e.g. quantity, unit price, total price, customer number, purchase order number, commercial register and number (if available), date, invoice address, shipping address, signature, and other necessary contact details.

A.3 Environment and history

A.3.1 Collection

a) **Geographic Origin:**

Data Type: Controlled vocabulary

Description: The geographic location, at least the original country where the microbial material is isolated from, the recommended format can be [Region], [City], [Country].

Example 1: Doi Inthanon National Park, Chiang Mai, Thailand

b) **Substrate:**

Data Type: String

Description: Substrate from which the microbial material was isolated (e.g. soil, water, blood, leaf, etc.). Name of host plants/animals can be a kind of substrate.

Example 1: Oil pollution

Example 2: Vegetative residues

c) **Habitat:**

Data Type: String

Description: Indicates the biotope where the species is found. A spatial region or named place. The field contains the location information of latitude, longitude, altitude and depth. The format and representation of altitude and depth should be specified. Physical factors such as; humidity, range of temperature, pH, and light intensity for the environment as well as biotic factors such as the availability of food and the presence or absence of predators.

Example 1: Beech forest

Example 2: Water body

1) **Latitude:**

Description: Enter latitude geographic coordinates as decimal fractions. Do not enter in degrees, minutes and seconds format.

2) **Longitude:**

Description: Enter longitude geographic coordinates as decimal fractions. Do not enter in degrees, minutes and seconds format.

3) **Altitude:**

Description: The altitude is the vertical distance between Earth's surface above sea level and the sampled position.

4) **Depth:**

Description: Depth is defined as the vertical distance below local surface, e.g. for sediment or soil samples depth is measured from sediment or soil surface, respectively. Depth can be reported as an interval for subsurface samples.

d) **Collected by:**

Data Type: String

Description: Name of the collector.

e) **Date of Collection:**

Data Type: Date

Description: The date on which the microbial material and associated data was sampled.

Example 1: 1988-09-21

Example 2: Before 1989

A.3.2 Isolation

a) **Isolated From:**

Data Type: International environmental ontologies list

Description: Information about the substrate or host from which the strain/ microbial material (e.g. virus, plasmid, cDNA) is isolated from, and the environment or isolation source.

Example 1: Soil

Example 2: Rat faeces

Example 3: Yeast cake for Ethanol fermentation from rice

Example 4: Fermented cheese

b) **Isolated by:**

Data Type: String

Description: Name of isolator, original number of isolate, date. Information separated by commas (or parentheses).

Example: I.S. Nogood, I.S.N. 1548, 16 Oct 1982.

c) **Method of isolation:**

Data Type: String

Description: Method used to isolate the microbial material. The isolation medium (abbreviated) may be given in parentheses.

d) **Date of isolation**

Description: The date in which the microbial material was isolated. The format [YYYY]-[MM]-[DD] shall be used, if applicable.

A.3.3 Identification

a) **Identified by:**

Data Type: String

Description: A list (concatenated and separated) of names of people, groups, or organizations who assigned the taxon to the subject. If a second microbiologist reached a different identification, also the date (year) is to be entered between parentheses. After the delimiter "<":

- 1) The natural/cultural material on which the identification was based;
- 2) The method used, e.g. rDNA (16S, 18S, 5.8S, and/or ITS) sequencing, DNA hybridization, crossing, characterization methods including flagellar staining, GC content of DNA, protein analysis by MALDI-TOF, etc.

b) **Date of identification:**

Data Type: Date

Description: The date on which the microbial material was identified; using the format [YYYY]-[MM]-[DD], if applicable.

A.4 Biological interactions

b) **Symbiosis:**

Data Type: String

Description: Only for mutualistic or commensalistic symbioses. Type of symbiosis (lichen, ectosymbionts or endomycorrhiza), scientific name of symbiont or of group of symbionts, can be separated by "with".

Example 1: Endomycorrhiza with Neottia

Example 2: Ectomycorrhiza with *Abies alba*, *Picea abies*

Example 3: Lichenized with *Coccomyxa*

c) **Mycoparasitism:**

Data Type: String

Description: Qualification of the parasite or symptoms, and host (indicated by scientific name), separated by “”:

Example: Destructive mycoparasite: *Rhizoctonia solani*

d) **Pathogenicity:**

Data Type: String

Description: Any specific disease that the strain can cause. Name and/or symptoms of disease, susceptible group of organisms, separated by “:”. If possible, use scientific names of organisms. The names are organized through the thesaurus file of substrata.

Example: Leaf spot: *Hordeum*; sheath rot: *Triticum*

e) **Allergenicity:**

Data Type: String

Description: Type of allergy and susceptible group of organisms.

Example: Skin irritation: rabbit

f) **Toxicity to other organisms:**

Data Type: String

Description: Symptoms, susceptible group of organisms, organs and degree of toxicity.

Example: Skin lesions: rabbit < at 0,01 % (HT-2 toxin).

g) **Antagonistic activities against other organisms:**

Data Type: String

Description: Action and name of susceptible organism, and responsible mechanism. Nematophagous activities are also entered here. Information can be separated by “:” or “<”.

Example 1: Growth inhibition: *Heterobasidion annosum*

Example 2: Capturing nematodes < adhesive network

A.5 Sexuality

a) **Sexual behaviour:**

Data Type: String

Description: Indicates the type of sexual behaviour.

b) **Sexual state:**

Data Type: String

Description: Indicates the actual sexual condition (e.g. mating type) of the strain, if applicable.