International Standard

ISO 23494-1

First edition
2026-06

Biotechnology — Provenance information model for biological material and data —

Part 1: Design concepts and general requirements

Biotechnologie — Modèle d'information sur la provenance des matériels et données biologiques —

Partie 1: Concepts de conception et exigences générales

Reference number
ISO 23494-1:2026(en)

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

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 document 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).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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*.

This first edition of ISO 23494-1 cancels and replaces ISO/TS 23494-1:2023, which has been technically revised.

A list of all parts in the ISO 23494 series can be found on the ISO website.

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

Research in life sciences has undergone significant changes during recent years, evolving away from individual projects confined to small research groups to transnational consortia covering a wide range of techniques and expertise. The exchange of research data and biological materials has become essential for the research in life sciences and biotechnology, and consequently interoperability and quality measures of data have become imperative.

At the same time, several reports addressing the quality of research papers in life sciences uncovered an alarming number of ill-founded claims. The reasons for the deficiencies are diverse, with insufficient quality and documentation of the biological material used being the major issue.

Hence, there is urgent need for standardized and comprehensive documentation of the entire process, from collection, processing, storage and analysis of the biological material to data generation, data analysis, and ultimately, interpretation and use, by humans and machine, for example, in artificial intelligence (AI) applications. The resulting provenance information can serve as a quality indicator and can provide evidence of the reliability of the data, thus enabling transparency and comparability of research results.

The purpose of this document is the standardization of provenance information management for the biotechnology domain in a way that allows for meaningful data integration. To achieve this, provenance information needs to be prepared in such a way that interoperability between prevailing tools for data generation, processing and analysis is ensured. While well-established approaches for provenance information are available in information technology (e.g., OPM[1] or World Wide Web Consortium (W3C) PROV[2]), their implementation for the biotechnology domain and related fields in particular is still a pending issue (as became evident in the Electronic Health Records Systems for Clinical Research (EHR4CR) and Translational Research and Patient Safety in Europe (TRANSFoRm) projects [3]).

With the increasing adoption of data-intensive technologies, such as next-generation sequencing (NGS), high-throughput mass spectrometry as used for proteomics or metabolomics, or high-throughput microscopy in digital pathology, and their impact on data collection strategies, consistent and comprehensive documentation of data provenance has become a necessity.

In fact, experimental designs in life sciences have moved from individual experiments with a limited amount of data towards pipelines generating a vast volume of raw digital data using massively parallel acquisition systems demanding complex data processing workflows to extract biologically relevant information. This trend is particularly evident in NGS, where the actual data acquisition device at the wet lab-digital interface (i.e., the sequencer) is completely oblivious to the details of the experiment being performed, with all the specialization pushed to the protocols used by the sample preparation procedure and to the software pipelines processing the data. Software pipelines are continuously changing due to the evolution of analytical algorithms and reference data sets, which is having a significant effect on result concordance.

In addition, particular issues, relevant to scientific domains utilizing biological material and data obtained from humans, needs to be considered. These include aspects of data privacy, ethics or management of identities. Notably, issues such as withdrawal of an informed consent or communication of incidental findings require the implementation of appropriate mechanisms.

The major objectives for collecting and storing provenance information are summarized as follows:

— retrospective evaluation of experimental results and data analysis with respect to the influence of standard operating procedures (SOPs), processing methods, and process parameters;

— quality monitoring of biological materials and data entered in a computational workflow or analysis pipeline (e.g., against reference ranges and tolerances);

— automation of quality control procedures (e.g., comparisons of processing pipelines);

— profiling of sample and data analysis to identify bottlenecks;

— assessment of fitness for purpose of biological materials and data for the intended use.

To achieve these objectives, a digitally processable description of provenance information is required.

This overarching document will be complemented by additional horizontal standards and by appropriate vertical standards for specific fields (e.g., collection of biological material, data generation, processing of biological material and data). The basic requirements contained in this document do not impose any limitations to future, domain-specific standards based on this document.

The standardization of provenance information requires the conceptualization and essential specifications for the generation, management, provisioning and maintenance as described in this document. Not covered in this document are additional fundamental components such as a generic model for provenance information and extensions common to all kinds of provenance information, ensuring security, privacy and non-repudiation. For particular domains in biotechnology and life sciences in general, detailed specifications building on a common provenance model are required, covering provenance information describing:

— the life cycle of biological materials, including acquisition, processing, transport and storage;

— the data generation by analytical methods;

— the data processing and analysis in computational workflows.

This document provides definitions for relevant terms used and specifies fundamental requirements for provenance information generation, management and provisioning.

Biotechnology — Provenance information model for biological material and data —

Part 1: Design concepts and general requirements

1 Scope

This document specifies a general concept for provenance information of biological material and data, organizational roles, and requirements for provenance information management. The provenance information covers any information relevant to the traceability, quality and fitness for purpose of the biological material and data generated throughout the life cycle of the biological material from collection to analysis, including data originating from analytical procedures applied to the biological material and further processing of the data.

This document is applicable to organizations, authorities and industries that are:

a) acquiring, collecting, processing, testing, analysing, storing, or distributing biological material in biotechnology and biomedicine (e.g., biobanks, laboratories, biomedical research as well as biotechnological development or production);

b) generating, collecting, analysing, processing, or storing data on and related to biological material (e.g., biobanks, laboratories, developers, manufacturers, or other institutions and commercial organizations in biotechnology or biomedicine);

c) generating, collecting, analysing, processing, or storing data or digital objects in biotechnology and biomedicine (e.g., in vitro/in vivo/in silico diagnostics developers and manufacturers, or other institutions and commercial organizations in the domain);

d) manufacturing devices or software for the aforementioned tasks or providing facilities for these tasks.

This document is also applicable to providers of services related to provenance information management (e.g., provenance information generation, storage, provision, or validation).

This document can be used by customers, regulatory authorities, organizations and schemes using peer-assessment, accreditation bodies, and others for confirming or recognizing the competence of the aforementioned parties.

This document does not apply to biological material and data used for medical diagnosis, treatment and therapy.

NOTE 1 This document can be applied by organizations performing laboratory or research activities as well as other activities in biotechnology and biomedicine.

NOTE 2 International, national, or regional regulations, standards, or requirements can apply to specific topics covered in this document, e.g., for organizations handling human materials procured and used for diagnostic and treatment purposes.

2 Normative references

There are no normative references in this document.