
Biotechnology — Data publication — Preliminary considerations and concepts

*Biotechnologie — Publication de données — Considérations et
concepts préliminaires*

iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

[ISO/TR 3985:2021](https://standards.iteh.ai/catalog/standards/iso/e52dd07f-b9c7-44b3-844e-d38d46c1c82b/iso-tr-3985-2021)

<https://standards.iteh.ai/catalog/standards/iso/e52dd07f-b9c7-44b3-844e-d38d46c1c82b/iso-tr-3985-2021>



iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

[ISO/TR 3985:2021](https://standards.iteh.ai/catalog/standards/iso/e52dd07f-b9c7-44b3-844e-d38d46c1c82b/iso-tr-3985-2021)

<https://standards.iteh.ai/catalog/standards/iso/e52dd07f-b9c7-44b3-844e-d38d46c1c82b/iso-tr-3985-2021>



COPYRIGHT PROTECTED DOCUMENT

© ISO 2021

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	v
Introduction	vi
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Abbreviated terms	4
5 Principles	4
5.1 General	4
5.2 Current technologies, approaches and their flaws	5
5.3 Standards and best practices to facilitate data sharing and reuse	6
5.3.1 Maximizing value to the payer	6
5.3.2 Data findability	6
5.3.3 Data machine and human interpretability	6
5.3.4 Using accepted controlled vocabularies and naming conventions	6
5.3.5 Biological annotation technology domain independence	6
5.3.6 Data locatability using multiple queries	7
5.4 Additional desirable attributes	7
5.4.1 Data linkage to a published and openly accessible document describing the experimental system	7
5.4.2 Data format linkage to a published and openly accessible document describing the format	7
5.4.3 Existing information technology	7
5.4.4 Development of tools and best practices for creating web friendly and search engine crawlable data documents	7
5.5 Essential considerations	7
5.5.1 Common annotation across multiple data sources	7
5.5.2 Keyword template	8
5.5.3 Embedding ontological descriptions	9
5.5.4 Pseudo-documents	9
6 Major challenges	10
6.1 General	10
6.2 Domain	10
6.3 Regionalization	10
6.4 Proprietary data	10
6.5 Large number of existing bio-ontologies, controlled vocabularies and terminologies	10
6.6 Large number of existing data repositories and corresponding domain specific data formats	11
6.7 Large number of funding agencies (e.g. national, educational, philanthropic, commercial)	11
7 Examples of existing national and regional standards or requirements for data sharing or publication	11
7.1 General	11
7.2 USA	11
7.3 Canada	11
7.4 European Union	11
7.5 Germany	12
7.6 China	12
7.7 United Kingdom	12
7.8 India	12
7.9 Japan	12
8 Existing legal requirements for data protection	12

8.1	USA	12
8.2	European Union	13
9	Timing of data publication	13
10	Costs of data publication	13
11	Archival data	13
12	Validation and verification of compliance	13
13	Affected stakeholder categories	13
Annex A (informative)	Searchability of scientific content on the web	14
Annex B (informative)	Example enhanced annotation of text documents	16
Bibliography		17

iTeh Standards
(<https://standards.itih.ai>)
Document Preview

[ISO/TR 3985:2021](https://standards.itih.ai/catalog/standards/iso/e52dd07f-b9c7-44b3-844e-d38d46c1c82b/iso-tr-3985-2021)

<https://standards.itih.ai/catalog/standards/iso/e52dd07f-b9c7-44b3-844e-d38d46c1c82b/iso-tr-3985-2021>