ISO/IEC 23092-3:2025-05 (E)

Information technology - Genomic information representation - Part 3: Metadata and application programming interfaces (APIs)

Contents

Page

Foreword v

Introduction vi

1 Scope 1

2 Normative references 1

3 Termsanddefinitions 2

4 Abbreviated terms 2

5 Conventions 2
5.1 Character encoding 2
5.2 Bit Ordering 2
5.3 Syntax functions and data types 3
5.4 Graphic notations 3

6 Information metadata 4
6.1 General 4
6.2 Dataset group metadata 4
6.3 Reference metadata 5
6.4 Dataset metadata 5
6.5 Annotation table metadata 7
6.5.1 General 7
6.5.2 Anotation table general metadata 7
6.5.3 Anotation table analytics metadata 10
6.5.4 Anotation table linkages metadata 13
6.5.5 Anotation table history metadata 14
6.6 Metadata protection 15
6.7 Mechanism for extensions of the metadata set 16
6.7.1 General 16
6.7.2 Example for dataset metadata extensions 16
6.7.3 Example for obfuscating labels 16
6.7.4 Example for obfuscating sequences 17
6.8 Metadata profiles 17
6.8.1 General 17
6.8.2 Example of metadata profile -- Run 17
6.8.3 Example of metadata profile -- Genomic data commons 18

7 Metrics metadata 18
7.1 General 18
7.2 Syntax 18
7.3 Semantics 19

8 Clinical data linkage metadata 20
8.1 General 20
8.2 CDL Metadata protection 22

9 Protection metadata 22
9.1 General 22

9.2 Encryption of gen_info elements and blocks ... 23
9.2.1 General ... 23
9.2.2 EncryptionParameters carried in dataset group protection ... 23
9.2.3 EncryptionParameters carried in dataset protection ... 24
9.2.4 EncryptionParameters carried in annotation table protection ... 28
9.2.5 Key retrieval ... 33
9.2.6 Decryption ... 34
9.3 Privacy rules for the use of the genomic information ... 36
9.3.1 General ... 36
9.3.2 Example of use of privacy rules ... 37

9.4 Digital signature of gen_info elements and blocks ... 38
9.4.1 General ... 38
9.4.2 Signatures carried in dataset group protection ... 38
9.4.3 Signatures carried in dataset protection ... 38
9.4.4 Signatures carried in annotation table protection ... 40
9.4.5 Signatures carried in descriptor stream protection ... 42

10 Access unit information ... 42
10.1 General ... 42
10.2 genAuxRecord ... 43
10.3 genAux ... 44
10.4 genTag ... 44

11 Decoding process for metadata ... 45
11.1 General ... 45
11.2 Initialization of parameters ... 47
11.2.1 General ... 47
11.2.2 Properties ... 47
11.2.3 Parameters ... 48
11.2.4 Constants ... 48
11.2.5 Process ... 49
11.3 Macros ... 51
11.4 Decoding process ... 53

12 Application programming interfaces (APIs) ... 60
12.1 General ... 60
12.2 Structure of the API ... 60
12.3 Detailed specification of the API ... 61
12.3.1 Data types ... 61
12.3.2 Return codes ... 61
12.3.3 Metadata fields ... 62
12.3.4 Output structures ... 62
12.3.5 Filters ... 71
12.3.6 Genomic information ... 79
12.3.7 Metadata ... 84
12.3.8 Protection ... 87
12.3.9 Reference ... 91
12.3.10 Statistics ... 91

Annex A (normative) XML schemas corresponding to metadata information and protection elements ... 96

Annex B (informative) Example use cases of annotation table linkages metadata ... 98

Annex C (informative) XML schemas and XML-based data ... 100

Annex D (informative) Example of key transport ... 109

Annex E (informative) SAM interoperability ... 113

Bibliography ... 120