



IEEE Standard for Bioinformatics Analyses Generated by High-Throughput Sequencing (HTS) to Facilitate Communication

DOI:

[10.1109/IEEESTD.2020.9094416](https://doi.org/10.1109/IEEESTD.2020.9094416)

Document Version

Proof

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

IEEE P2791 BioCompute Working Group (BCOWG), Mazumder, R. (Ed.), & Simonyan, V. (Ed.) (2020). *IEEE Standard for Bioinformatics Analyses Generated by High-Throughput Sequencing (HTS) to Facilitate Communication: IEEE Std 2791-2020*. IEEE. <https://doi.org/10.1109/IEEESTD.2020.9094416>

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact openresearch@manchester.ac.uk providing relevant details, so we can investigate your claim.



1 **P2791™/D3D4**
2 **Draft Standard for Bioinformatics**
3 **Analyses Generated by High-**
4 **Throughput Sequencing (HTS) to**
5 **Facilitate Communication**

6 Sponsor
7
8 **Standards Committee**
9 of the
10 **IEEE Engineering in Medicine and Biology Society**
11

12
13 Approved <Date Approved>

14
15 **IEEE-SA Standards Board**
16

17 Copyright © 2018 by The Institute of Electrical and Electronics Engineers, Inc.
18 Three Park Avenue
19 New York, New York 10016-5997, USA

20 All rights reserved.

21 This document is an unapproved draft of a proposed IEEE Standard. As such, this document is subject to
22 change. USE AT YOUR OWN RISK! IEEE copyright statements SHALL NOT BE REMOVED from draft
23 or approved IEEE standards, or modified in any way. Because this is an unapproved draft, this document
24 must not be utilized for any conformance/compliance purposes. Permission is hereby granted for officers
25 from each IEEE Standards Working Group or Committee to reproduce the draft document developed by
26 that Working Group for purposes of international standardization consideration. IEEE Standards
27 Department must be informed of the submission for consideration prior to any reproduction for
28 international standardization consideration (stds.ipr@ieee.org). Prior to adoption of this document, in
29 whole or in part, by another standards development organization, permission must first be obtained from
30 the IEEE Standards Department (stds.ipr@ieee.org). When requesting permission, IEEE Standards
31 Department will require a copy of the standard development organization's document highlighting the use
32 of IEEE content. Other entities seeking permission to reproduce this document, in whole or in part, must
33 also obtain permission from the IEEE Standards Department.

34 IEEE Standards Department
35 445 Hoes Lane
36 Piscataway, NJ 08854, USA

37

Abstract: ~~This standard establishes accurate and secure~~A major goal of this standard is to improve communication of bioinformatics protocols and data in order to facilitate bioinformatics workflow related exchange and communication between regulatory agencies, pharmaceutical companies, bioinformatics platform providers and researchers. ~~Accurate Detailed~~ communication helps ensure responsibility, reproducibility, verify bioinformatics protocol, track provenance information and promote interoperability. In addition, this standard also defines the assurance program for evaluating and certifying products against those requirements.

Keywords: genomics, next generation sequencing, high throughput sequencing, massively parallel sequencing, NGS, HTS, MPS, workflow, pipeline, bioinformatics, analysis, regulatory

OPEN SOURCE NOTICE; ~~This IEEE project incorporates open source software (download¹: Web view²) under a BSD 3-Clause license either within the draft or as a separate file. Any person contributing material to IEEE open source software during standards development, SA ballot, or Public Review is required to provide the appropriate license to IEEE (IEEE Contributor License Agreement or CLA). Please note that any Contributions to IEEE open source software that is submitted without first providing the appropriate CLA to IEEE will not be eligible for inclusion either in the draft standard or the open source software, and may not be considered by the comment resolution group. The appropriate CLA form (<https://opensource.ieee.org/community/cla/bsd3>) must be submitted to oscontrib@ieee.org, if applicable. Note that Federal or Crown employees must submit the CLA that includes Appendix A.~~
OPEN SOURCE NOTICE; This IEEE project incorporates open source software [File download: <https://gitlab.com/IEEE-SA/2791/ieee-2791-schema/-/archive/v1.3.0-alpha/ieee-schema-v1.3.0-alpha.zip>; Web link: <https://gitlab.com/IEEE-SA/ieee-2791/ieee-schema/tree/v1.3.0-alpha>] under a BSD 3-Clause license either within the draft or as a separate file. Any person contributing material to IEEE open source software during standards development, SA ballot, or Public Review is required to provide the appropriate license to IEEE (IEEE Contributor License Agreement or CLA). Please note that any Contributions to IEEE open source software that is submitted without first providing the appropriate CLA to IEEE will not be eligible for inclusion either in the draft standard or the open source software, and may not be considered by the comment resolution group. The appropriate CLA form (<https://app.box.com/s/2nlzxsru3jewdy9ello4rfkokwt9k841>) must be submitted to os-fb@ieee.org, if applicable. Note that Federal or Crown employees must submit the CLA that includes Appendix A.

Formatted: Font: Not Bold
Formatted: Font: Not Bold
Formatted: Font: Not Bold
Formatted: Font: Not Bold

Formatted: Font: (Default) Arial, Bold, Font color: Auto
Formatted: Font: (Default) Arial, Font color: Auto
Formatted: Font: (Default) Arial, Font color: Auto, Highlight
Formatted: Font: (Default) Arial, Font color: Auto
Formatted: Font: (Default) Arial, Font color: Auto, Highlight
Formatted: Font: (Default) Arial, Font color: Auto
Formatted: Font: (Default) Arial, Font color: Auto, Highlight
Formatted: Font: (Default) Arial, Font color: Auto
Formatted: Font: (Default) Arial, Font color: Auto, Highlight
Formatted: Font: (Default) Arial, Font color: Auto
Formatted: Font: (Default) Arial, Font color: Auto, Highlight
Formatted: Font: (Default) Arial, Font color: Auto
Formatted: Font: (Default) Arial, Font color: Auto
Formatted: Font: (Default) Arial, Font color: Auto

¹<https://gitlab.com/IEEE-SA/2791/ieee-2791-schema/-/archive/v1.3.0-alpha-rc2/ieee-2791-schema-v1.3.0-alpha-rc2.zip>
²<https://gitlab.com/IEEE-SA/2791/ieee-2791-schema/tree/v1.3.0-alpha-rc2>
The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

Formatted: Font: 8 pt
Formatted: Font: 8 pt
Formatted: Font: 8 pt
Formatted: Font: 8 pt
Formatted: Font: 8 pt

Copyright © 2018 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved. Published <Date Published>. Printed in the United States of America.

IEEE is a registered trademark in the U.S. Patent & Trademark Office, owned by The Institute of Electrical and Electronics Engineers, Incorporated.

PDF: ISBN 978-0-XXXX-XXXX-X STDXXXXX
Print: ISBN 978-0-XXXX-XXXX-X STDPDXXXXX

IEEE prohibits discrimination, harassment, and bullying.
For more information, visit <http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html>.
No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher.

1 **Important Notices and Disclaimers Concerning IEEE Standards Documents**

2 IEEE documents are made available for use subject to important notices and legal disclaimers. These
3 notices and disclaimers, or a reference to this page, appear in all standards and may be found under the
4 heading "Important Notices and Disclaimers Concerning IEEE Standards Documents." They can also be
5 obtained on request from IEEE or viewed at <http://standards.ieee.org/ipr/disclaimers.html>.

6 **Notice and Disclaimer of Liability Concerning the Use of IEEE Standards** 7 **Documents**

8 IEEE Standards documents (standards, recommended practices, and guides), both full-use and trial-use, are
9 developed within IEEE Societies and the Standards Coordinating Committees of the IEEE Standards
10 Association ("IEEE-SA") Standards Board. IEEE ("the Institute") develops its standards through a
11 consensus development process, approved by the American National Standards Institute ("ANSI"), which
12 brings together volunteers representing varied viewpoints and interests to achieve the final product. IEEE
13 Standards are documents developed through scientific, academic, and industry-based technical working
14 groups. Volunteers in IEEE working groups are not necessarily members of the Institute and participate
15 without compensation from IEEE. While IEEE administers the process and establishes rules to promote
16 fairness in the consensus development process, IEEE does not independently evaluate, test, or verify the
17 accuracy of any of the information or the soundness of any judgments contained in its standards.

18 IEEE Standards do not guarantee or ensure safety, security, health, or environmental protection, or ensure
19 against interference with or from other devices or networks. Implementers and users of IEEE Standards
20 documents are responsible for determining and complying with all appropriate safety, security,
21 environmental, health, and interference protection practices and all applicable laws and regulations.

22 IEEE does not warrant or represent the accuracy or content of the material contained in its standards, and
23 expressly disclaims all warranties (express, implied and statutory) not included in this or any other
24 document relating to the standard, including, but not limited to, the warranties of: merchantability; fitness
25 for a particular purpose; non-infringement; and quality, accuracy, effectiveness, currency, or completeness
26 of material. In addition, IEEE disclaims any and all conditions relating to: results; and workmanlike effort.
27 IEEE standards documents are supplied "AS IS" and "WITH ALL FAULTS."

28 Use of an IEEE standard is wholly voluntary. The existence of an IEEE standard does not imply that there
29 are no other ways to produce, test, measure, purchase, market, or provide other goods and services related
30 to the scope of the IEEE standard. Furthermore, the viewpoint expressed at the time a standard is approved
31 and issued is subject to change brought about through developments in the state of the art and comments
32 received from users of the standard.

33 In publishing and making its standards available, IEEE is not suggesting or rendering professional or other
34 services for, or on behalf of, any person or entity nor is IEEE undertaking to perform any duty owed by any
35 other person or entity to another. Any person utilizing any IEEE Standards document, should rely upon his
36 or her own independent judgment in the exercise of reasonable care in any given circumstances or, as
37 appropriate, seek the advice of a competent professional in determining the appropriateness of a given
38 IEEE standard.

39 IN NO EVENT SHALL IEEE BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL,
40 EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO:
41 PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS;
42 OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY,
43 WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR
44 OTHERWISE) ARISING IN ANY WAY OUT OF THE PUBLICATION, USE OF, OR RELIANCE
45 UPON ANY STANDARD, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND
46 REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

1 **Translations**

2 The IEEE consensus development process involves the review of documents in English only. In the event
3 that an IEEE standard is translated, only the English version published by IEEE should be considered the
4 approved IEEE standard.

5 **Official statements**

6 A statement, written or oral, that is not processed in accordance with the IEEE-SA Standards Board
7 Operations Manual shall not be considered or inferred to be the official position of IEEE or any of its
8 committees and shall not be considered to be, or be relied upon as, a formal position of IEEE. At lectures,
9 symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall
10 make it clear that his or her views should be considered the personal views of that individual rather than the
11 formal position of IEEE.

12 **Comments on standards**

13 Comments for revision of IEEE Standards documents are welcome from any interested party, regardless of
14 membership affiliation with IEEE. However, IEEE does not provide consulting information or advice
15 pertaining to IEEE Standards documents. Suggestions for changes in documents should be in the form of a
16 proposed change of text, together with appropriate supporting comments. Since IEEE standards represent a
17 consensus of concerned interests, it is important that any responses to comments and questions also receive
18 the concurrence of a balance of interests. For this reason, IEEE and the members of its societies and
19 Standards Coordinating Committees are not able to provide an instant response to comments or questions
20 except in those cases where the matter has previously been addressed. For the same reason, IEEE does not
21 respond to interpretation requests. Any person who would like to participate in revisions to an IEEE
22 standard is welcome to join the relevant IEEE working group.

23 Comments on standards should be submitted to the following address:

24 Secretary, IEEE-SA Standards Board
25 445 Hoes Lane
26 Piscataway, NJ 08854 USA

27 **Laws and regulations**

28 Users of IEEE Standards documents should consult all applicable laws and regulations. Compliance with
29 the provisions of any IEEE Standards document does not imply compliance to any applicable regulatory
30 requirements. Implementers of the standard are responsible for observing or referring to the applicable
31 regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not
32 in compliance with applicable laws, and these documents may not be construed as doing so.

33 **Copyrights**

34 IEEE draft and approved standards are copyrighted by IEEE under U.S. and international copyright laws.
35 They are made available by IEEE and are adopted for a wide variety of both public and private uses. These
36 include both use, by reference, in laws and regulations, and use in private self-regulation, standardization,
37 and the promotion of engineering practices and methods. By making these documents available for use and
38 adoption by public authorities and private users, IEEE does not waive any rights in copyright to the
39 documents.

40

41

1 Photocopies

2 Subject to payment of the appropriate fee, IEEE will grant users a limited, non-exclusive license to
3 photocopy portions of any individual standard for company or organizational internal use or individual,
4 non-commercial use only. To arrange for payment of licensing fees, please contact Copyright Clearance
5 Center, Customer Service, 222 Rosewood Drive, Danvers, MA 01923 USA; +1 978 750 8400. Permission
6 to photocopy portions of any individual standard for educational classroom use can also be obtained
7 through the Copyright Clearance Center.

8 Updating of IEEE Standards documents

9 Users of IEEE Standards documents should be aware that these documents may be superseded at any time
10 by the issuance of new editions or may be amended from time to time through the issuance of amendments,
11 corrigenda, or errata. A current IEEE document at any point in time consists of the current edition of the
12 document together with any amendments, corrigenda, or errata then in effect.

13 Every IEEE standard is subjected to review at least every ten years. When a document is more than ten
14 years old and has not undergone a revision process, it is reasonable to conclude that its contents, although
15 still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to
16 determine that they have the latest edition of any IEEE standard.

17 In order to determine whether a given document is the current edition and whether it has been amended
18 through the issuance of amendments, corrigenda, or errata, visit IEEE Xplore at <http://ieeexplore.ieee.org/>
19 or contact IEEE at the address listed previously. For more information about the IEEE-SA or IEEE's
20 standards development process, visit the IEEE-SA Website at <http://standards.ieee.org>.

21 Errata

22 Errata, if any, for all IEEE standards can be accessed on the IEEE-SA Website at the following URL:
23 <http://standards.ieee.org/findstds/errata/index.html>. Users are encouraged to check this URL for errata
24 periodically.

25 Patents

26 Attention is called to the possibility that implementation of this standard may require use of subject matter
27 covered by patent rights. By publication of this standard, no position is taken by the IEEE with respect to
28 the existence or validity of any patent rights in connection therewith. If a patent holder or patent applicant
29 has filed a statement of assurance via an Accepted Letter of Assurance, then the statement is listed on the
30 IEEE-SA Website at <http://standards.ieee.org/about/sasb/patcom/patents.html>. Letters of Assurance may
31 indicate whether the Submitter is willing or unwilling to grant licenses under patent rights without
32 compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of
33 any unfair discrimination to applicants desiring to obtain such licenses.

34 Essential Patent Claims may exist for which a Letter of Assurance has not been received. The IEEE is not
35 responsible for identifying Essential Patent Claims for which a license may be required, for conducting
36 inquiries into the legal validity or scope of Patents Claims, or determining whether any licensing terms or
37 conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing
38 agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that
39 determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely
40 their own responsibility. Further information may be obtained from the IEEE Standards Association.

1 Participants

2 At the time this draft standard was completed, the P2791 Working Group had the following membership:

3 **Raja Mazumder, Chair**
4 **Vahan Simonyan, Vice Chair**

5
6 Ogan Abaan 18 Paul Duncan 30 Rahi Navelkar
7 Jonas Almeida 19 Josep Gelpi 31 Asa Oudes
8 Gil Alterovitz 20 Carole Goble 32 Janisha Patel
9 Payal Banerjee 21 Jeremy Goecks 33 John Penn
10 Amanda Bell 22 Jonathan Jacobs 34 Megan Pottersbusch
11 Surajit Bhattacharya 23 Robel Kahsay 35 Jonathan Pryke
12 Lee Black 24 Jonathon Keeney 36 Stian Soiland-Reyes
13 Ben Busby 25 Charles Hadley King 37 Dan Taylor
14 Kristy Cloyd-Warwick 26 Jonathan LoTempio 38 Jason Travis
15 Ryan Connor 27 Xeandong Meng 39 Paul Walsh
16 Michael Crusoe 28 David Michaels 40 Jianchao Yao
17 Dennis Dean 29 Hiroki Morizono

41

42 The following members of the individual/entity balloting committee voted on this standard. Balloters may
43 have voted for approval, disapproval, or abstention.

44 *[To be supplied by IEEE]*

45 Balloter1 48 Balloter4 51 Balloter7
46 Balloter2 49 Balloter5 52 Balloter8
47 Balloter3 50 Balloter6 53 Balloter9

54

55 When the IEEE-SA Standards Board approved this standard on <Date Approved>, it had the following
56 membership:

57 *[To be supplied by IEEE]*

58 **<Name>, Chair**
59 **<Name>, Vice Chair**
60 **<Name>, Past Chair**
61 **Konstantinos Karachalios, Secretary**

62 SBMember1 65 SBMember4 68 SBMember7
63 SBMember2 66 SBMember5 69 SBMember8
64 SBMember3 67 SBMember6 70 SBMember9

71 *Member Emeritus

72

1 Introduction

2 This introduction is not part of P2791/~~D44~~, Draft Standard for Bioinformatics Analyses Generated by High-
3 Throughput Sequencing (HTS) to Facilitate Communication.

4 ~~The BioComputeP2791 specification enables the description of bioinformatic genome analysis workflows~~
5 ~~in a standardized way~~~~standardizes bioinformatics workflows in the genomic analysis space.~~
6 ~~BioComputeP2791~~ addresses the tremendous variability and uncertainty in communicating bioinformatics
7 workflows and data related to analysis as a result of high throughput sequencing (HTS). The need to
8 resolve issues in communication was felt particularly strongly between the United States Food and Drug
9 Administration (FDA) and the entities that submit any work to the FDA for regulatory analysis that
10 includes an HTS component^{1,2} (<https://doi.org/10.5731/pdajpst.2016.006734>, ~~and~~ [PMCS5510742](https://doi.org/10.5731/pdajpst.2016.006734))
11 ~~and~~ [PMCS5510742](https://doi.org/10.5731/pdajpst.2016.006734)). A plan ~~to~~ ~~for~~ ~~what would become~~ ~~BioComputeP2791~~ and initial goals of the project were
12 drafted in a collaboration between the George Washington University and the FDA in 2014. The project
13 has grown since then to include publications, workshops, applied use cases, and a large community of
14 participants and collaborators. ~~The P2791 Objects created according to this standard are standard-is~~
15 ~~intended~~ 1) to be both human and machine readable, 2) to be applied to genomic analysis workflows, and
16 3) to be able to capture ~~all~~ details related to a workflow in such a way as to facilitate efficient communication
17 and improve reproducibility and interoperability. ~~Every~~ ~~e~~ ~~Efforts~~ ~~is~~ ~~were~~ made to accommodate ~~any~~ ~~as~~
18 ~~many~~ tools, platforms or scripts ~~as possible~~, and to be adaptable to future developments in this field under a
19 unified set of descriptions to standardize and streamline the representations of such complex bioinformatics
20 processes.

21 ~~BioCompute-P2791~~ is a standard and a ~~BioCompute-P2791~~ Object (~~BCO~~) is an instance of that standard.
22 High throughput sequencing (HTS), also referred to as next-generation sequencing (NGS) or massively
23 parallel sequencing (MPS), has increased the pace at which we generate, compute and share genomic data
24 in biomedical sciences. As a result, scientists, clinicians and regulators are now faced with a new data
25 paradigm that is less portable, more complex and most of all poorly standardized. The ~~BCO-P2791 Objects~~
26 ~~are written in uses a simple~~ JSON format to encode important information on the execution of
27 computational pipelines, or for the creation of knowledge bases. ~~BioComputeP2791~~ can be ~~considered to~~
28 ~~be~~ process oriented (for software pipelines) and/or product oriented (for knowledge bases). ~~So error domain~~
29 ~~can include information to do QA and/or QC.~~ The goal of using a ~~BCO-P2791 Object~~ is to streamline
30 communication of these otherwise difficult to elucidate details between stakeholders in academia, industry
31 and regulatory agencies.

32 ~~Encapsulating HTS data processing in a BCO will facilitate swift communications between the FDA and~~
33 ~~other stakeholders who seek regulatory review/approval hence reducing the burden and time to decision.~~

34 ~~The US Food and Drug Administration (FDA) and George Washington University (GW) have partnered to~~
35 ~~establish a framework for community based standards development and harmonization of HTS~~
36 ~~computations and data formats.~~ Standardized HTS data processing descriptions and data formats will
37 promote interoperability and simplify the verification of the bioinformatics protocols applied against data.
38 To do this, a schema has been developed to represent instances of computational analysis as a ~~BCOP2791~~
39 ~~Object~~. A ~~BCO-P2791 Object~~ includes:

- 40 — Information about parameters and versions of the executable programs in a pipeline
- 41 — Reference to input and output test data for verification of the pipeline
- 42 — A usability domain
- 43 — Keywords

¹ Alterovitz G et al. Enabling Precision Medicine via standard communication of NGS provenance, analysis, and results. *In press PLoS Biol.* 2018 Dec; 16(12):e3000099 DOI: <https://doi.org/10.1371/journal.pbio.3000099>.

² Simonyan V, Goecks J and Mazumder R. Biocompute Objects—A Step towards Evaluation and Validation of Biomedical Scientific Computations. *PDA J Pharm Sci Technol.* 2017 Mar-Apr;71(2):136-146

Formatted: Superscript

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

Formatted: Default Paragraph Font

1 — A list of agents involved along with other important metadata, such as their specific contribution

2 — ~~Knowledge of input data is intended to be captured according to existing efforts, such as~~
3 ~~Minimum Information Required about a Glycomics Experiment (MIRAGE)³, Minimum Information~~
4 ~~about a Proteomics Experiment (MIAPE)⁴, Standards for Reporting Enzymology Data (STREND A)⁵ and~~
5 ~~to be in accordance with Minimum Information Standards⁶ Knowledge of input data is intended to be~~
6 ~~captured according to existing efforts, including MIRAGE, MIAPE, and STREND A, and to be in~~
7 ~~accordance with Minimum Information Standards.~~ In addition to all the information captured in the
8 ~~BCOP2791 Object~~, the ~~BCOP2791 Object~~ itself ~~must-is intended to~~ be independent of the execution
9 environment, whether it is a local or a cloud-based infrastructure.
10

Formatted: Indent: Left: 0.14", No bullets or numbering

Formatted: Indent: Left: 0.14", Line spacing: single, No
bullets or numbering

11

³ Kolarich, Daniel; Rapp, Erdmann; Struwe, Weston B.; Haslam, Stuart M.; Zaia, Joseph; McBride, Ryan; Agravat, Sanjay; Campbell, Matthew P.; Kato, Masaki; Ranzinger, Rene; Kettner, Carsten; York, William S. (1 April 2013). "The Minimum Information Required for a Glycomics Experiment (MIRAGE) Project: Improving the Standards for Reporting Mass-spectrometry-based Glycoanalytic Data". *Molecular & Cellular Proteomics*. 12 (4): 991–995. doi:10.1074/mcp.O112.026492. ISSN 1535-9476. PMC 3617344. PMID 23378518

Formatted: Font: 8 pt

Formatted: Font: 8 pt

⁴ Taylor, C. F.; Paton, N. W.; Lilley, K. S.; Binz, P. A.; Julian Jr, R. K.; Jones, A. R.; Zhu, W.; Apweiler, R.; Aebersold, R.; Deutsch, E. W.; Dunn, M. J.; Heck, A. J. R.; Leitner, A.; Macht, M.; Mann, M.; Martens, L.; Neubert, T. A.; Patterson, S. D.; Ping, P.; Seymour, S. L.; Souda, P.; Tsugita, A.; Vandekerckhove, J.; Vondriska, T. M.; Whitelegge, J. P.; Wilkins, M. R.; Xenarios, I.; Yates Jr, J. R.; Hermjakob, H. (2007). "The minimum information about a proteomics experiment (MIAPE)". *Nature Biotechnology*. 25 (8): 887–893. doi:10.1038/nbt1329. PMID 17687369

Formatted: Font: 8 pt

Formatted: Font: 8 pt

⁵ Tipton, K.F., Armstrong, R.N., Bakker, B.M., Bairoch, A., Cornish-Bowden, A., Halling, P.J., Hofmevr, J.-H., Levh, T.S., Kettner, C., Raushel, F.M., Rohwer, J., Schomburg, D., Steinbeck, C. (2014) Standards for Reporting Enzyme Data: The STREND A Consortium: What it aims to do and why it should be helpful. *Perspect. Sci.* 1(1.6):131-137. DOI: 10.1016/j.pise.2014.02.012

Formatted: Font: 8 pt

Formatted: Font: 8 pt

⁶ Taylor, Chris F (2008). "Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project". *Nature Biotechnology*. 26 (8): 889–896. doi:10.1038/nbt.1411. PMC 2771753. PMID 18688244

Formatted: Font: 8 pt

Formatted: Font: 8 pt

1 **Contents**

2	1. Overview	1111
3	1.1 General	1111
4	1.2 Scope	2221
5	1.3 Purpose	2
6	2. Normative references	2
7	3. Definitions, acronyms, and abbreviations	3332
8	3.1 Acronyms and abbreviations	3332
9	4. BioCompute P2791 Standard	3333
10	4.1 General	3333
11	Annex A (informative) Bibliography	7775
12		
13		

Draft Standard for Bioinformatics Analyses Generated by High- Throughput Sequencing (HTS) to Facilitate Communication

1. Overview

1.1 General

The [BioComputeP2791](#) standard captures relevant information from a high throughput sequencing workflow as a [P2791 Object](#) in order to enable a user to understand and interpret the workflow efficiently and with high confidence. [BioComputeP2791](#) is a standard that [was initially created with the goal of improving efficiency in regulatory reviewis particularly well adapted to regulatory review](#). Pursuant to this, workflow steps and prerequisites to execute workflow steps are recorded in detail in [the a BioCompute P2791 standardObject](#). Information is recorded using key/value pairs in JavaScript Object Notation (JSON), adhering to the [P2791](#) JSON Schema.

[Key/value pairs Information in P2791 Objects isare](#) organized by domains;

- The Provenance Domain - tracks metadata [about the P2791 Object](#)
- The Usability Domain - tracks what was done
- The Extension Domain - provide user-defined fields
- The Description Domain - captures a description of external resources, pipeline steps, and the relationships of I/O objects
- The Execution Domain - describes information needed for deployment, software configuration and running applications in a dependent environment
- The Parametric Domain - captures all parameters that customize a computational flow
- The Input and Output Domain - contains a list of global input and output files

- ~~The Error Domain - describes errors, including the limits of detectability, false positives, false negatives, statistics confidence of outcomes, and description of errors (i.e. empirical or algorithmic).~~

~~This document should be read in conjunction with the open source P2791 JSON Schema files (<https://w3id.org/2791/>) which are referred to from the text, for instance “*provenance domain.json*” refers to (https://w3id.org/2791/provenance_domain.json). Files in the schemas folder are kept separate for organization. References in the BCOP2791 Object schema (*\$ref*) to these files should be replaced with the proper domain from the appropriate file. For example, line 142~~4~~ of “*p2791object.json*” (“*\$ref*”: “*provenance domain.json*”) is a reference to the structure specified in the *provenance domain.json* file. The BCO-P2791 Object Schema builds on the JSON Schema by adding domains in a way that facilitates the communication of bioinformatics workflows. A description of the domain files follows. In addition, two examples have been generated by the community of users¹, as well as a tool to automate the creation of a file using the BCO schema standard².~~

1.2 Scope

This standard establishes ~~accurate-detailed~~ and ~~secure-structured~~ communication of bioinformatics protocols in order to facilitate bioinformatics workflow related exchange and communication between regulatory agencies, pharmaceutical companies, bioinformatics platform providers and researchers. ~~Accurate-Detailed~~ communication helps ensure responsibility, verify bioinformatics protocol, track provenance information and promote interoperability.

1.3 Purpose

The standards allows for the cross platform communications of complex computation from inception to manufacturing of medical products and ~~services~~. ~~Another goal of this standard is to improve efficiency and speed of communication services, resulting in decreased costs of drug discovery and review, and accelerated delivery of treatment to patients.~~

2. Normative references

The following referenced documents are indispensable for the application of this document (i.e., they must be understood and used, so each referenced document is cited in text and its relationship to this document is explained). For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments or corrigenda) applies.

- ¹<https://github.com/biocompute-objects/examples><https://www.w3id.org/2791/examples>
- ²Alterovitz G et al. 2019. “IEEE 2791 Schema: 1.4”. GitLab. December 2, 2019. <https://gitlab.com/IEEE-SA/2791/ieee-2791-schema/-/archive/master/ieee-2791-schema-master.zip> (supplement to <https://gitlab.com/IEEE-SA/2791/ieee-2791-schema>)
- ³Bray, T. “The JavaScript Object Notation (JSON) Data Interchange Format.” IETF Tools, INTERNET STANDARD, Dec. 2017, tools.ietf.org/html/rfc8259.
- ⁴“JSON Schema: A Media Type for Describing JSON Documents.” Edited by H Andrews and A Wright, IETF Tools, Internet Engineering Task Force, Mar. 2018, tools.ietf.org/html/draft-handrews-json-schema-01.

Formatted: No bullets or numbering

Formatted: Font: Italic

Formatted: Font: (Default) Courier New

Formatted: Font: (Default) Times New Roman, Italic

Formatted: Font: Italic

Formatted: Font: (Default) Courier New

Formatted: Font: Italic

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

Formatted: Font: 8 pt

1 [BioCompute Schema](#).⁶
2 [JSON Schema](#).⁷
3 [description_domain](#)⁸
4 [execution_domain](#)⁹
5 [io_domain](#)¹⁰
6 [parametric_domain](#)¹¹
7 [provenance_domain](#)¹²
8 [usability_domain](#)¹³
9 [Common Workflow Language](#)¹⁴

10 3. Definitions, acronyms, and abbreviations

11 For the purposes of this document, the following terms and definitions apply. The *IEEE Standards*
12 *Dictionary Online* should be consulted for terms not defined in this clause.¹⁵

13 3.1 Acronyms and abbreviations

14	BCO	BioCompute Object
15	JSON	JavaScript Object Notation
16	FHIR	Fast Healthcare Interoperability Resources
17	SCM	Source Control Management

Formatted: Highlight

Formatted: Highlight

18 4. BioCompute P2791 Standard

19 4.1 General

20 This document describes the P2791 standard for describing bioinformatic workflows. A **BCO-P2791**
21 **“Object”** is an instance of the P2791 standard, and is a text file written in JSON data structure that shall

⁸“JSON Schema Validation: A Vocabulary for Structural Validation of JSON.” Edited by A Wright and H Andrews, IETF Tools,
Internet Engineering Task Force, Mar. 2018, tools.ietf.org/html/draft-handrews-json-schema-validation-01.

Formatted: Font: 8 pt

⁶Fully-organized BioCompute Object-containing all domains in proper JSON Schema format

Formatted: Font: 8 pt

<https://www.w3id.org/biocompute/schemas/1.3.0/biocomputeobject.json><https://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/biocomputeobject.json>

⁷<https://tools.ietf.org/html/rfc7159> <https://json-schema.org/latest/json-schema-core.html>.

⁸https://www.w3id.org/2791/biocompute/schemas/schemas/1.3.0/description_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/description_domain.json

⁹https://www.w3id.org/biocompute/schemas/1.3.0/execution_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/execution_domain.json

¹⁰https://www.w3id.org/biocompute/schemas/1.3.0/io_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/io_domain.json

¹¹https://www.w3id.org/biocompute/schemas/1.3.0/parametric_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/parametric_domain.json

¹²https://www.w3id.org/biocompute/schemas/1.3.0/provenance_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/provenance_domain.json

¹³https://www.w3id.org/biocompute/schemas/1.3.0/usability_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/usability_domain.json

¹⁴<https://www.w3id.org/cwl/v1.0/>

Formatted: Font: 8 pt

¹⁵<http://dictionary.ieee.org>.

Formatted: Font: 8 pt

1 consist of all domains required by the **BioCompute-P2791 Schema**¹⁶
2 (<https://w3id.org/biocompute2791/schemas/1.3.0/biocomputeobject.json>). The P2791 Schema is the formal
3 definition of the standard against which instances of the standard can be validated. JavaScript Object
4 Notation (JSON) is a textual format used by both instances of Objects and the formal P2791 Schema, and
5 the JSON Schema is the language used to express the P2791 Schema.

7 A **valid BCO-Object shall be written in must conform to the P2791 JSON Schema** (see section 4.3) JSON
8 Schema¹⁷, and therefore invokes all of the requirements of the JSON Schema- (while a valid Object file
9 **must conform to the schema, the schema file is not technically required to create the Object file**). Later
10 versions of P2791 may be updated for conformance with future JSON Schema versions. The minimum
11 requirement to execute the standard is the fully organized **BCO-P2791 Object** containing all domains in
12 proper JSON Schema format. Pursuant to JSON schema, the **required-fields required for a valid Object** are
13 listed at the top of the **BCO2791Object.json file**.

15 The fully-organized BCO file is hosted in the schemas folder, along with related files. All the files in the
16 schemas folder repository are linked together (using JSON pointers as described by the JSON Schema),
17 being referenced by the 'biocompute2791object.json' file. For development purposes, these files are used
18 to track changes, but some are not required to adhere to the standard. Those required for a complete BCO
19 are the *biocomputeobject.json*, *description-domain.json*, *execution-domain.json*, *io-domain.json*,
20 *parametric-domain.json*, *provenance-domain.json*, and *usability-domain.json*. The *error-domain.json* is
21 an optional domain to further describe empirical and algorithmic sources and measures of error for a
22 bioinformatics workflow (<https://w3id.org/2791/schemas/1.3.0/error-domain.json>) is an optional domain
23 that further describes a bioinformatics workflow, and the *extension-domain.json* is an optional domain that
24 contains user-defined fields.

26 At its top-level, Objects have the following three required metadata fields. The top three lines of a BCO (:
27 "bcg-spec version", "bcg-object_id", and "checksumE-Tag") are metadata that describe
28 the BCO. These lines are external to all domains. Everything except for the *checksumE-Tag*,
29 *bcg-object_id*, and *bcg-spec version* should shall be included in the generation of an ETag (see
30 <https://tools.ietf.org/html/rfc7232#section-2.3>) - which can be "strong" or "weak" (see
31 <https://tools.ietf.org/html/rfc7232#section-2.1>)¹⁸. It is recommended that the ETag be deleted or updated if
32 the object file is changed (except in cases using weak ETags in which the entirety of the change comprises
33 a simple re-writing of the JSON) should be included in the generation of the hash.

35 *bcg-object_id* is a string that follows the JSON Schema format of namespace/ref. The checksum is
36 calculated on all following lines. should shall be a unique identifier. Users are free to number BCO Object
37 files in the manner of their choosing, however, in order to avoid naming conflicts, it is recommended that a
38 domain namespace be registered with a registration authority, such as the one at
39 <https://www.biocomputeobject.org/registry.html>, such as **BioCompute Registry**. For example,
40 <http://www.w3id.org/biocompute/1.3.0/schemas/ABC-biocomputeobject01.json> <http://www.w3id.org/example/gcom/exampleproject/1.3.0/schemas/ABC-object001.json>, where "ABC" is a registered domain, and
41 "*biocomputeobject001.json*" is an arbitrary identifier, chosen by the owner of that domain.

45 The remaining top level fields are domains that partition workflow into meaningful subunits. These are the
46 Description domain, Error domain, Execution domain, IO domain, Parametric domain, Provenance domain,
47 and Usability domain.

48 Files in the schemas folder are kept separate for organization. References in the BCO schema (*\$ref*) to
49 these files should be replaced with the proper domain from the appropriate file. For example, line 141
50 ("*\$ref*": "*provenance-domain.json*") is a reference to the structure specified in the

¹⁶ Fully organized BioCompute Object containing all domains in proper JSON Schema format

¹⁷ <https://www.w3id.org/biocompute/schemas/1.3.0/biocomputeobject.json> <https://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/biocomputeobject.json>

¹⁸ P2791 Standard for Bioinformatics Computations and Analyses Generated by High-Throughput Sequencing (HTS) to Facilitate Communication <https://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/biocomputeobject.json> JSON Schema <https://json-schema.org/latest/json-schema-core.html>

¹⁹ Internet Engineering Task Force (IETF), "Hypertext Transfer Protocol (HTTP/1.1): Conditional Requests," Internet Engineering Task Force, ISSN: 2070-1721, [Online]. Available: <https://tools.ietf.org/html/rfc7232> [Accessed: Nov. 4th, 2019].

Field Code Changed	... [1]
Formatted	... [2]
Formatted	... [3]
Formatted	... [4]
Formatted	... [5]
Formatted	... [6]
Formatted	... [7]
Formatted	... [8]
Formatted	... [9]
Formatted	... [10]
Field Code Changed	... [11]
Formatted	... [12]
Formatted	... [13]
Formatted	... [14]
Formatted	... [15]
Formatted	... [16]
Formatted	... [17]
Formatted	... [18]
Formatted	... [19]
Formatted	... [20]
Formatted	... [21]
Formatted	... [22]
Formatted	... [23]
Formatted	... [24]
Formatted	... [25]
Formatted	... [26]
Formatted	... [27]
Formatted	... [28]
Formatted	... [29]
Formatted	... [30]
Formatted	... [31]
Formatted	... [32]
Formatted	... [33]
Formatted	... [34]
Formatted	... [35]
Formatted	... [36]
Formatted	... [37]
Formatted	... [38]
Formatted	... [39]
Formatted	... [40]
Formatted	... [41]
Formatted	... [42]
Formatted	... [43]
Formatted	... [44]
Formatted	... [45]
Formatted	... [46]
Formatted	... [47]

1 ~~provenance_domain.json~~ file. The BCO Schema builds on the JSON Schema by adding domains in a way
2 ~~that facilitates the communication of bioinformatics workflows. A description of the domain files follows.~~
3 ~~In addition, two examples have been generated by the community of users¹⁹, as well as a tool to automate~~
4 ~~the creation of a file using the BCO schema standard²⁰.~~

Formatted: Font: Italic

6 The Description Domain of an **BioCompute**-Object²¹ contains a description of external resources, pipeline
7 steps, and the relationship of I/O objects (https://w3id.org/2791/description_domain.json).

9 The Error Domain contains information related to the bounds of detection (such as the minimum sequence
10 depth and minimum sequence coverage), and statistical analyses of the pipeline (such as the false negative
11 and false positive rates). ~~Fields in the Error Domain are open-ended and not restricted nor defined by the~~
12 ~~current P2791 standard. It is recommended that the keys directly under **empirical_error** and~~
13 ~~**algorithmic_error** use a full URI. Resolving the URI should give a JSON Schema or textual~~
14 ~~definition of the field. Other keys are not allowed in **error_domain**. Fields in the Error Domain can be~~
15 ~~determined algorithmically (by repeatedly invoking the pipeline with the same data) or empirically (by~~
16 ~~invoking the pipeline with different data, often synthetically generated data).~~

Formatted: Font: (Default) Courier New

Formatted: Font: (Default) Courier New

Formatted: Font: (Default) Courier New

18 The Execution Domain of an **BioCompute**-Object²² ~~describes details of~~ contains information needed for
19 deployment, software configuration, and running applications in a dependent environment
20 (https://w3id.org/2791/execution_domain.json). This may include scripts, drivers, environment variables,
21 and other software prerequisites.

23 The IO Domain of an **BioCompute**-Object²³ is a list of global input and output files that may exist on local
24 machine or on another machine (https://w3id.org/2791/io_domain.json). It does not include references to
25 intermediate files.

27 The Parametric Domain of a **BioCompute**P2791 Object²⁴ includes any parameters used in a workflow
28 (https://w3id.org/2791/parametric_domain.json). ~~This is intended for use when parameters are changed~~
29 ~~This is typically used only in the context of parameters changed from default settings for ease of~~
30 ~~understanding.~~

Field Code Changed

32 The Provenance Domain contains metadata related to the **BCO**Object
33 (https://w3id.org/2791/provenance_domain.json) and is not used for computation^{25,26}. It is used to track the
34 flow of data from original source to final computation, and includes contributors, reviewers, and
35 versioning. ~~In the event that a P2791 Object retrospectively references an existing Object (such as an~~
36 ~~example Object), the **derived_from** field within the Provenance Domain shall reference the specific~~
37 ~~Object by **object_id** field. In the event that the Object is an example Object or is created de novo~~
38 ~~without reference to existing work, this field is not included. In the event that the Object is an example or~~
39 ~~template P2791 Object, best practice is to state this in the Usability Domain, along with relevant details~~
40 ~~(such as completeness of data, whether data is real or artificial, etc.).~~

Formatted: Font: (Default) Courier New

Formatted: Font: (Default) Courier New

¹⁹ Example repository

²⁰ BCO-Editor

²¹ https://www.w3id.org/biocompute/schemas/1.3.0/description_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/description_domain.json

²² https://www.w3id.org/biocompute/schemas/1.3.0/execution_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/execution_domain.json

²³ https://www.w3id.org/biocompute/schemas/1.3.0/io_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/io_domain.json

²⁴ https://www.w3id.org/biocompute/schemas/1.3.0/parametric_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/parametric_domain.json

²⁵ https://www.w3id.org/biocompute/schemas/1.3.0/provenance_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/provenance_domain.json

²⁶ <https://github.com/biocompute-objects/BCO-Specification/blob/master/provenance-domain.md>

1 The Usability Domain of an **BioCompute** Object²⁷ is a plain language description of what was done in the
2 workflow (https://w3id.org/2791/usability_domain.json). This ~~is not used for computation, and~~ should align
3 with the actual steps described elsewhere in the **BCO** Object. The Usability Domain conveys the purpose of
4 the **BCO** Object, ~~and which also improves searchability findability of the BCO Object. It is recommended~~
5 ~~that a novel use of the BCO could result in the creation of a new entry with a new Usability Domain.~~

6
7 The Extension Domain allows a user to define additional fields and is optional. ~~A separate folder called~~
8 ~~extension_domain exists within the schemas folder. Two Extension Domain example files exist in the~~
9 ~~extension_domain folder that describe how a BCO can include a reference to FHIR (Extension Domain~~
10 ~~example: FHIR²⁸) and/or to SCM (Extension Domain example: SCM²⁹). The Extension Domain is for the~~
11 ~~inclusion of any additional structured information. A valid JSON schema for each extension used in this~~
12 ~~domain is expected to be specified. The schema should be name spaced, and it is recommended that~~
13 ~~resolving the namespaced URI will provide the extension's JSON Schema. The URL should be provided in~~
14 ~~the required "extension_schema" field. If execution portability is desired, then the included script~~
15 ~~should be in the Common Workflow Language v1.0³⁰ (<https://w3id.org/cwl/v1.0/>) or later format. In order~~
16 ~~to avoid potential naming conflicts, it is recommended that users register their domain with~~

17
18 Additional helpful resources have been created, including a Community User Guide for Best Practices
19 [B1]. This document describes ways in which the schema has been used and is known to be effective, using
20 these to derive best practices. In addition, a repository of examples [B3] exists, which includes the use of
21 optional Error Domain. A BCO Editor [B4] tool has also been generated. The BCO Editor is an example
22 implementation of the schema, and can be used to create and edit BCO's. Finally, a script [B5] to validate
23 that documents have been created according to the BCO schema is also available for use. This python tool
24 will check a document to ensure that it has been created according to the current BCO Schema.
25

Formatted: Font: (Default) Courier New

²⁷ https://www.w3id.org/biocompute/schemas/1.3.0/usability_domain.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/usability_domain.json

²⁸ https://www.w3id.org/biocompute/schemas/1.3.0/extension_domain/fhir.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/extension_domain/fhir_extension.json

²⁹ https://www.w3id.org/biocompute/schemas/1.3.0/extension_domain/scm.jsonhttps://github.com/biocompute-objects/BCO-Specification/blob/master/schemas/extension_domain/scm_extension.json

³⁰ <https://w3id.org/cwl/v1.0/>

Formatted: Font: 8 pt

Formatted: Font: 8 pt

1 **Annex A**

2 (informative)

3 **Bibliography**

4 Bibliographical references are resources that provide additional or helpful material but do not need to be
5 understood or used to implement this standard. Reference to these resources is made for informational use
6 only.

7 [B1] Community User Guide for Best Practices.
8 <https://w3id.org/biocompute/1.3.1>[https://github.com/biocompute-](https://github.com/biocompute-objects/BCO_Specification/blob/master/user_guide.md)
9 [objects/BCO_Specification/blob/master/user_guide.md](https://w3id.org/biocompute/spec/1.3.0/user_guide.md)[https://w3id.org/biocompute/spec/1.3.0/user_guide-](https://w3id.org/biocompute/spec/1.3.0/user_guide.md)
10 [md](https://w3id.org/biocompute/spec/1.3.0/user_guide.md)

Formatted: Highlight

11 [B2] JSON Schema: A Media Type for Describing JSON Documents. [https://json-schema.org/latest/json-](https://json-schema.org/latest/json-schema-core.html)
12 [schema-core.html](https://tools.ietf.org/html/draft-handrews-json-schema-01) <https://tools.ietf.org/html/draft-handrews-json-schema-01>[https://json-](https://json-schema.org/latest/json-schema-core.html)
13 [schema.org/latest/json-schema-core.html](https://json-schema.org/latest/json-schema-core.html)

14 [B3] Example Repository:
15 [https://www.w3id.org/biocompute](https://www.w3id.org/biocompute/examples)[/examples](https://www.w3id.org/biocompute/examples)[w3id.org/biocompute/examples/HCV1a.json](https://www.w3id.org/biocompute/examples/HCV1a.json)

Formatted: Highlight

Formatted: Font color: Auto

16 [B4] BCO Editor. https://w3id.org/biocompute/beo_editor/html/js[https://github.com/biocompute-](https://github.com/biocompute-objects/beo_editor)
17 [objects/beo_editor](https://github.com/biocompute-objects/beo_editor)

18 [B5] Validation Script:
19 <https://w3id.org/biocompute/1.3.1/validate.py><https://w3id.org/biocompute/validate.py>

Formatted: Highlight

20 [BioCompute Registry](https://github.com/biocompute-objects/BCO_Specification/blob/master/validate.py)https://github.com/biocompute-objects/BCO_Specification/blob/master/validate.py

Formatted: Highlight

21

Page 4: [1] Change	Unknown	
Field Code Changed		
Page 4: [2] Formatted	Keeney, Jonathon	9/16/2019 3:29:00 PM
Font: Italic		
Page 4: [3] Formatted	Keeney, Jonathon	9/16/2019 3:29:00 PM
Font: Italic		
Page 4: [4] Formatted	Jonathon Keeney	7/16/2019 11:40:00 AM
Font: Italic		
Page 4: [5] Formatted	Jonathon Keeney	7/16/2019 11:40:00 AM
Font: Italic		
Page 4: [6] Formatted	Jonathon Keeney	7/16/2019 11:40:00 AM
Font: Italic		
Page 4: [7] Formatted	Jonathon Keeney	7/16/2019 11:40:00 AM
Font: Italic		
Page 4: [8] Formatted	Jonathon Keeney	7/16/2019 11:40:00 AM
Font: Italic		
Page 4: [9] Formatted	Jonathon Keeney	7/16/2019 11:40:00 AM
Font: Italic		
Page 4: [10] Formatted	Jonathon Keeney	7/16/2019 11:40:00 AM
Font: Italic		
Page 4: [11] Change	Unknown	
Field Code Changed		
Page 4: [12] Formatted	Jonathon Keeney	7/16/2019 11:41:00 AM
Font: Italic		
Page 4: [13] Formatted	Jonathon Keeney	8/23/2019 2:15:00 PM
Font: (Default) Courier New, Highlight		
Page 4: [14] Formatted	Jonathon Keeney	7/16/2019 11:41:00 AM
Font: (Default) Courier New		
Page 4: [15] Formatted	Jonathon Keeney	8/23/2019 2:15:00 PM
Font: (Default) Courier New, Highlight		
Page 4: [16] Formatted	Jonathon Keeney	7/16/2019 11:41:00 AM
Font: (Default) Courier New		
Page 4: [17] Formatted	Jonathon Keeney	7/16/2019 11:41:00 AM
Font: (Default) Courier New		
Page 4: [18] Formatted	Jonathon Keeney	7/16/2019 11:42:00 AM
Font: (Default) Courier New		
Page 4: [19] Formatted	Jonathon Keeney	8/23/2019 2:15:00 PM
Font: (Default) Courier New, Highlight		

Page 4: [20] Formatted	Jonathon Keeney	7/16/2019 11:42:00 AM
Font: (Default) Courier New		
Page 4: [21] Formatted	Jonathon Keeney	8/23/2019 2:15:00 PM
Font: (Default) Courier New, Highlight		
Page 4: [22] Formatted	Jonathon Keeney	7/16/2019 11:42:00 AM
Font: (Default) Courier New		
Page 4: [23] Formatted	Jonathon Keeney	9/5/2019 12:54:00 PM
Default Paragraph Font		
Page 4: [24] Formatted	Jonathon Keeney	8/23/2019 2:15:00 PM
Font: (Default) Courier New, Highlight		
Page 4: [25] Formatted	Jonathon Keeney	7/16/2019 11:44:00 AM
Font: (Default) Courier New		
Page 4: [26] Formatted	Jonathon Keeney	7/16/2019 11:45:00 AM
Highlight		
Page 4: [27] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Font: Italic		
Page 4: [28] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Default Paragraph Font, Font: Italic		
Page 4: [29] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Default Paragraph Font, Font: Italic		
Page 4: [30] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Default Paragraph Font, Font: Italic		
Page 4: [31] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Default Paragraph Font, Font: Italic		
Page 4: [32] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Default Paragraph Font, Font: Italic		
Page 4: [33] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Font: Italic		
Page 4: [34] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Font: Italic		
Page 4: [35] Formatted	Jonathon Keeney	8/23/2019 2:12:00 PM
Font: Italic, Highlight		
Page 4: [36] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Font: Italic		
Page 4: [37] Formatted	Jonathon Keeney	8/16/2019 3:36:00 PM
Font: Italic		
Page 4: [38] Formatted	Jonathon Keeney	7/16/2019 11:48:00 AM
Code Char, Font: (Default) Courier New		

Page 4: [39] Formatted	Jonathon Keeney	7/16/2019 11:48:00 AM
-------------------------------	------------------------	------------------------------

Font: (Default) Courier New

Page 4: [40] Formatted	Keeney, Jonathon	9/8/2019 4:00:00 PM
-------------------------------	-------------------------	----------------------------

Font: 6 pt

Page 4: [41] Formatted	Keeney, Jonathon	9/8/2019 4:00:00 PM
-------------------------------	-------------------------	----------------------------

Font: 6 pt

Page 4: [42] Formatted	Keeney, Jonathon	9/8/2019 4:00:00 PM
-------------------------------	-------------------------	----------------------------

Font: 6 pt

Page 4: [43] Formatted	Keeney, Jonathon	9/8/2019 4:00:00 PM
-------------------------------	-------------------------	----------------------------

Font: 6 pt

Page 4: [44] Formatted	Keeney, Jonathon	9/8/2019 4:00:00 PM
-------------------------------	-------------------------	----------------------------

Font: 6 pt

Page 4: [45] Formatted	Keeney, Jonathon	9/8/2019 4:00:00 PM
-------------------------------	-------------------------	----------------------------

Font: 6 pt

Page 4: [46] Formatted	Keeney, Jonathon	11/26/2019 1:06:00 PM
-------------------------------	-------------------------	------------------------------

Font: 8 pt

Page 4: [47] Formatted	Keeney, Jonathon	11/26/2019 1:06:00 PM
-------------------------------	-------------------------	------------------------------

Font: 8 pt