

Learning Health Systems
Computable Knowledge Publications
Instructions for Authors

Learning Health Systems welcomes submissions for computable knowledge publications that make *models* representing biomedical knowledge available as peer-reviewed software. Such a publication consists of two components: (i) One or more computable biomedical knowledge artifacts (CBKs), which are software artifacts, and (ii) a written human-readable manuscript describing the CBKs and their underlying models. The manuscripts must conform to the [Journal's general author guidelines](#) as well as the criteria described here.

CBKs within the Journal's scope are computable representations of models that result from an analytical, machine learning, and/or deliberative process. Examples include regression models, algorithms, computable phenotypes and guidelines, workflows, and policies; decision trees; calculators; Bayesian networks; and models derived from machine learning such as neural networks. *This is not an exhaustive list.* Models must relate to individual and/or population health. They can be in clinical, educational, research, or public health domains. A submission can include multiple CBKs only if they are logically interrelated in which case the instructions provided here for a single CBK apply to all.

The submission and review processes are explained below:

Short Proposal

Prospective authors should initially submit a short proposal describing their submission.

To propose a submission, send an email to Kathleen Young, the Journal's Editorial Assistant, at kayoung@med.umich.edu. The email should include:

1. the corresponding author's affiliation and full contact information, the proposed title of the publication
2. a brief description of the artifact(s),
3. citations to any pertinent previous publications,
4. a specification of the language and language version in which the artifact(s) are written.

The proposal will trigger a rapid determination by the Editorial Office of whether the proposed submission, and particularly its CBK, falls within the scope of the Journal and can be suitably tested as part of the review process as described below. Authors of prospective publications considered to be within scope will then be invited to prepare a complete submission for subsequent peer review.

Written Manuscript

The written manuscript should be between 2,000 and 4,000 words in length, and it can take many different forms, depending on the extent to which publications about the CBK and its development already appear in the literature. The following outline should be considered as guidance rather than prescription.

- a) An introduction describing the clinical, educational, research, or public health problem the model is designed to address. The introduction should also include a literature-based argument for the importance of the problem and may also site other models that have been developed to address this or similar problems.
- b) A methods section describing how the model was developed. The level of detail should align with previously published material about the CBK.
- c) A deployment section describing the experience of the CBK's use to date.
- d) A section describing the CBK's technical architecture.
- e) A discussion section as appropriate. This section might discuss limitations of the model, generalizability to settings other than where the model has been tested, and other topics. This section could also address user experience and provide tips on how to obtain maximum value from the deployed model.
- f) A technical implementation and test plan. This section will provide the specific information necessary to instantiating the model in a computing environment and ascertaining that it is functioning in accordance with specifications. Links to source code, additional required software, and any additional guides or documentation may be used freely and included in the references. Please make sure that supporting materials are well maintained and have permanent links.
- g) References, which must include all publications directly related to the CBK.

CBKs

The submitted CBKs may operate as “pure functions” that accept specified input data and return computed results. They may also take the form of statistical algorithms, stateful process models, queries, ETL pipelines, analytical models for machine learning, etc. Whatever the appropriate form of the model(s) represented, they should use commonly available software platforms, languages, and tools (for example, Python, JavaScript, R, or OWL). The editors express a preference for open-source, widely available technology. CBKs created via proprietary application packages, or that require such proprietary software to run them, are less likely to be accepted.

Published CBKs will not include the data used to train and/or validate the model encoded in the CBKs. Please make sure that sample data and common

configuration files are available when required to understand the deployment and use of the computable model. Authors, at their option, can furnish links to training and validation datasets as part of the CBKs' metadata.

Authors should provide: a) a machine executable or interpretable file (or files) containing the source for a CBK along with human-readable information required to run the model expressed in the CBK; b) a coded file containing test data that can be used to test the source code along with human-readable details describing the input data required by the model and the expected outputs the model should generate; c) metadata in human-readable form; d) a license specification in human-readable form; and e) any additional information the authors consider to be helpful to reviewers and users.

In particular, one goal of the publication of computable knowledge is to move beyond disseminating knowledge for *inspection* and to provide it in a form more nearly ready for *use*. Collections of CBKs submitted to the Journal should be relatively easy to set up, run, and validate for reviewers and readers. Because of the wide range of types of computable knowledge considered for publication we cannot be too specific on requirements for deployment and testing, but any submissions should be readily usable by other researchers with similar interests and experience and should fully demonstrate the claims made for the model with a combination of test scripts and data, preferable self-validating.

More details are provided below.

- a. Source Code and Information to Run It: Authors should submit their source code with enough additional information to inspect and execute it so that the Journal staff and, ultimately, those who access the publication can directly inspect the source code and execute it. In the case of submission of compiled or packaged artifacts, make sure that links to the source code are available. This additional information may include a list of software dependencies and how to resolve them, a set of instructions for compiling or interpreting the source code to execute it, and other information necessary to enable the artifact(s) to be used. In the case of compiled or packaged source files, make sure that links to the source code are available. This additional information will be published as a part of the written manuscript's "technical implementation and test plan".
- b. Input/Output Specifications: Authors should submit descriptions of the input data required by the model, in human-readable form, with as much detail as possible: including variable names, acceptable value sets, units of measurement, and any standards (and versions of these standards) used for data representation. Authors will submit a data file with examples of the *inputs* to their model they submit along with, for each input record, the corresponding value(s) of the output(s) that should be produced by the model. Ideally, the examples of input data will include sets that are

representative of the range of data that users of the model will be employing, including “edge cases” in addition to records that would be more focused on illustrating the basic functioning of the artifact. In cases where the inputs and outputs are not typical transactional, record-based data, for example, local data used to produce a trained model configuration, additional means or verification should be supplied. These submitted data files must not contain Protected Health Information (PHI) of specific individuals.

- c. Metadata: We are asking authors to provide, in human-readable form, the structured metadata that they believe are required to characterize their artifact(s). This will help us converge on a standard metadata set for the Journal. Please be careful to separate domain-metadata about the model from technical and descriptive metadata about this computable implementation of the model (e.g., the language the model is written in versus the biological system modeled). Prior to publication, we may also ask authors to provide additional metadata according to the emerging metadata standards for CBKs.
- d. License: Software submitted to the Journal must be attached to a specific license that allows testers and readers to access and run the submitted source code using the test data submitted. The Journal will not pre-specify eligible license types.

Copyright and Intellectual Property Considerations

Because *Learning Health Systems* is an open-access journal, authors retain copyright of their work and retain full intellectual property rights thereto. This policy will apply to the published written article and its related CBK.

Prior to publication, authors will be asked to attest that: 1) any errors, inaccuracies, misrepresentations, and malfunctions in the CBK are the sole responsibility of the authors; 2) the software does not contain components that constitute embedded intellectual property, use of which without permission would violate the license type specified by the authors; and 3) the CBK and the submitted test data do not contain any protected health information.

If the author is submitting to *Learning Health Systems* a paper offering a computable version of a model to complement a paper that has been previously published in only human-readable form, *Learning Health Systems* will not consider this to be dual publication. In that case, the written manuscript should focus on design, technical deployment, and implementation details and be minimally redundant with previously published work. All specific verbiage used in previous publications must be represented as quotations. All human-readable components of the publication must contain citations to all relevant prior publications.

Review Process

CBK submissions will be rigorously reviewed. The written manuscript will be evaluated by at least two external reviewers, following criteria closely paralleling review criteria for research papers.

Separately, the CBK will be evaluated by the Editorial Office for technical conformance to the specifications described above. The CBKs will be tested as part of the review process, using the input data provided by the author, and compared with the corresponding output results. This process will result in a “test report” considered a part of the review returned to the authors. If the Editorial Office cannot reproduce the expected results, this will be reflected in the test report and the submission will not be acceptable until the CBK is revised and successfully retested. User interfaces provided as part of the CBK will not be reviewed unless the purpose of the CBK is to create or support user interaction.