*Learning Health System*s
Pilot Project on Computable Knowledge Publications
Instructions for Authors

Learning Health Systems welcomes submissions of "Computable Knowledge" papers that make biomedical models and algorithms available as peer-reviewed software. Computable knowledge papers will consist of two components: one or more software artifacts to be called digital *Knowledge Objects* (KOs) and a written (human-readable) manuscript that describes the KOs. The written manuscript must conform to the journal's general author guidelines, as well as policies specific to computable knowledge papers described here. If a paper includes multiple KOs, they must be interrelated in a way that forms a logical collection.

Learning Health Systems is initiating this publication type via a pilot project through which many of the procedural details of the submission and review processes will be worked out interactively with the submitting authors. After a sufficient number of computable knowledge papers (N ~ 10) have been accepted for publication through the pilot stage, the Journal will adopt a more permanent set of policies and procedures.

1. Submitting a Paper During the Pilot Phase

During the pilot phase, prospective authors of computable knowledge papers should initially submit a short proposal describing their submission.

To propose a Computable Knowledge paper send an email to Kathleen Young, the Journal's Editorial Assistant, at kayoung@med.umich.edu. The email should include the corresponding author's affiliation and full contact information, the proposed title of the publication, type of publication (EP or IP: see below) proposed, a brief description of the KO(s), citations to any previous publications pertinent to the KO(s), a specification of the language and version in which the KO(s) are written; and, if available, a link to a record in the source code repository (e.g., GitHub) where the KOs currently reside.

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

2. Two Types of Computable Knowledge Publications

The Journal welcomes Computable Knowledge submissions of two types:

The first type is a full-length *Computable Knowledge Enhanced Publication* (CK-EP). It would describe a previously unpublished scientific effort resulting in a computable artifact, and provide one or more knowledge objects.

The second type is a shorter *Computable Knowledge Implementation Publication* (CK-IP). This brief would serve as a "bridge to implementation", describe technical deployment and implementation details, and share a computable artifact. Such a brief can complement other previously published (or soon to be published) articles. CK-IPs will enable sharing of peer-reviewed computable artifacts, not previously available in any systematic way, and their implementation in practice.

For both types, the written component of the paper will be available as a pdf from Journal's website, in the same manner that all other Journal articles are available. An embedded link in the article pdf will enable download and local execution of the computable Knowledge Object(s).

3. <u>Structure and Scope of Knowledge Objects (KOs)</u>

KOs encapsulate biomedical knowledge in computable form. KOs 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; computable 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.* Prospective authors should consult the Journal's Editorial Office for an opinion as to whether their model(s) are in scope. Models must relate to individual and/or population health. They can be in clinical, educational, research, or public health domains.

Conceptually, KOs have three components: 1) a "payload": the executable code that implements the model, 2) service specifications: a representation of the data required by the model and the result the model generates, 3) "metadata": a specified set of fields describing the model. The components of the KOs we expect authors to submit, during this pilot phase, are detailed in Section 4 of this document.

The submitted KOs should generally operate as "pure functions" that accept specified input data and return computed results. They should be coded in specified versions of commonly used languages such as, but not restricted to, Python, Javascript, R, or OWL. KOs created via proprietary application packages such as Excel or SAS, or that require such proprietary software to run them, are out of scope.

Published KOs will <u>not</u> include the data that were used to train and/or validate the model encoded in the KOs. Authors, at their option, can furnish links to training

and validation datasets as part of the KOs' metadata. Training and/or validation data must reside external to the Journal's digital KO repository maintained by the publisher, John Wiley and Sons.

4. <u>KO Submission Requirements</u>

When submitting the computable KO component of their paper, authors should provide: a) an executable file or files containing the KO source code along with (in human readable form) information required to run the model expressed in the KO; b) a coded file containing test data that can be used to test the source code along with (in human readable form) details describing the input data required by the model and outputs the model generates; 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. More details are provided below.

   a) <u>Source Code and Information To Run It</u>: Authors should submit their source code with enough additional information so that the Journal staff and, ultimately, those who access the published version of the computable knowledge paper can directly inspect the source code and execute it. 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 KO(s) to be used. When the article is published, this information will be included as part of the written manuscript's "technical implementation and test plan".

   b) <u>Input/Output Specifications</u>: 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 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 KO. These submitted data files must not contain Protected Health Information (PHI) of specific individuals.

   c) <u>Metadata</u>: During the pilot phase, we are asking authors to provide, in human-readable form, the structured metadata that they believe are required to characterize their KOs. This will help us converge on a standard metadata set for the Journal. Prior to publication, we may also ask authors to provide additional metadata, as this standard set evolves.

d) <u>License</u>: 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. For the pilot phase of the project, the Journal will not pre-specify eligible license types.

During the pilot phase, Editorial Office team members will work with authors to clarify the components of their submissions. Authors with questions are encouraged to contact the Editorial Office.

5. <u>Written Manuscript</u>

The written manuscript (maximum of 4000 words for CK-EPs or 2000 words for CK-IPs) should take the general format of a scientific report, consisting of:

a) An <u>introduction</u> 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 also cite other models that have been developed to address this or similar problems.

b) A <u>methods</u> section describing how the model was developed. If the model results from a quantitative analytical process, the method description should include as appropriate: a description of the training data and its source, the analytical methods used, and cross-validation methods employed. If the model results from a deliberative process, the method description should describe who engaged in the deliberative process, and how the process was conducted to reach a consensus product.

c) A <u>results</u> section that represents the model in human-readable form where possible; for example, a table of regression coefficients and other parameters that characterize the results of the analytical process. If these details are already published elsewhere, a reference (with URL) to the publication will suffice as long as the article referenced is accessible without fee. If these results are published behind a pay wall, the written paper submitted to this Journal must include these details with attribution.

d) A <u>discussion</u> section as appropriate. This section should 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.

e) A <u>metadata</u> section that includes, in the pilot phase, both the common metadata specified by the Journal and the author-provided additional metadata.

f) A <u>technical implementation and test plan</u>. 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.

6. Copyright and Intellectual Property

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 KO(s).

Prior to publication of a computable knowledge paper, authors will be asked to attest that: 1) any errors, inaccuracies, misrepresentations, and malfunctions in the coded KO(s) 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 KOs and the submitted test data do not contain any protected health information.

If the author is submitting to *Learning Health Systems*, without substantial modification, a CK-IP paper offering a computable version of a model to complement a paper that has been previously published in human readable form, *Learning Health System*s will not consider this to be dual publication. However, 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.

7. Review of Computable Knowledge Papers

Computable knowledge papers will be rigorously reviewed. The written paper will be evaluated by at least two external reviewers, following criteria closely paralleling review criteria for research papers.

The KOs will separately evaluated by the Editorial Office for technical conformance to the specifications described above. The payload(s) will be tested as part of the review process, using the input data provided by the author, and compared with the corresponding output results also provided. This process will result in a "test report" which will be considered 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 KOs are revised and successfully retested. User interfaces provided as part of the KO will not be reviewed unless the purpose of the KO is to create or support user interaction.