

Artifact Checklist (V20201122).

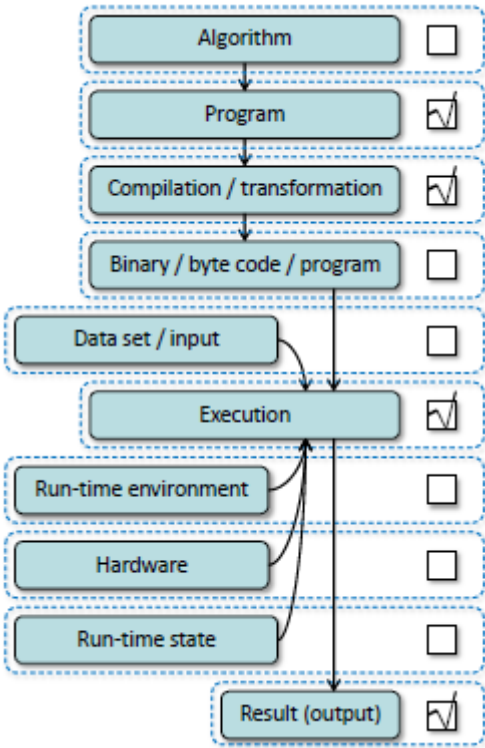
Here we provide a few informal suggestions to help you fill in your [Artifact Appendix with the Reproducibility Checklist](#) for submission while avoiding common pitfalls.

Abstract

Briefly and informally describe your artifact including minimal hardware and software requirements, how it supports your paper, how it can be validated, and what is the expected result. It will be used to select appropriate reviewers. It will also help readers understand what was evaluated and how.

Checklist

Together with the artifact abstract, this check-list will help us make sure that reviewers have appropriate competency and an access to the technology required to evaluate your artifact. It can also be used as meta information to find your artifacts in Digital Libraries.



Fill in whatever is applicable with some informal keywords and remove unrelated items (please consider questions below just as informal hints that reviewers are usually concerned about):

- **Algorithm:** Are you presenting a new algorithm?
- **Program:** Which benchmarks do you use ([PARSEC ARM real workloads](#), [NAS](#), [EEMBC](#), [SPLASH](#), [Rodinia](#), [LINPACK](#), [HPCG](#), [MiBench](#), [SPEC](#), [cTuning](#), etc)? Are they included or should they be downloaded? Which version? Are they public or private? If they are private, is there a public analog to evaluate your artifact? What is the approximate size?
- **Compilation:** Do you require a specific compiler? Public/private? Is it included? Which version?
- **Transformations:** Do you require a program transformation tool (source-to-source, binary-to-binary, compiler pass, etc)? Public/private? Is it included? Which version?
- **Binary:** Are binaries included? OS-specific? Which version?
- **Model:** Do you use specific models (ImageNet, AlexNet, MobileNets)? Are they included? If not, how to download and install? What is their approximate size?
- **Data set:** Do you use specific data sets? Are they included? If not, how to download and install? What is their approximate size?
- **Run-time environment:** Is your artifact OS-specific (Linux, Windows, MacOS, Android, etc) ? Which version? Which are the main software dependencies (JIT, libs, run-time adaptation frameworks, etc); Do you need root access?

- **Hardware:** Do you need specific hardware (supercomputer, architecture simulator, CPU, GPU, neural network accelerator, FPGA) or specific features (hardware counters to measure power consumption, SUDO access to CPU/GPU frequency, etc)? Are they publicly available?
- **Run-time state:** Is your artifact sensitive to run-time state (cold/hot cache, network/cache contentions, etc.)
- **Execution:** Any specific conditions should be met during experiments (sole user, process pinning, profiling, adaptation, etc)? How long will it approximately run?
- **Metrics:** Which metrics are reported (execution time, inference per second, Top1 accuracy, static and dynamic energy consumption, etc) - particularly important for multi-objective benchmarking, optimization and co-design (see [ACM ReQuEST ML/SW/HW co-design tournaments](#)).
- **Output:** What is your output (console, file, table, graph) and what is your result (exact output, numerical results, measured characteristics, etc)? Is expected result included?
- **Experiments:** How to prepare experiments and replicate/reproduce results (OS scripts, manual steps by user, [IPython/Jupyter notebook](#), [automated workflows](#), etc)? Do not forget to mention the maximum allowable variation of empirical results!
- **How much disk space required (approximately)?:** This can help evaluators and end-users to find appropriate resources.
- **How much time is needed to prepare workflow (approximately)?:** This can help evaluators and end-users to estimate resources needed to evaluate your artifact.
- **How much time is needed to complete experiments (approximately)?:** This can help evaluators and end-users to estimate resources needed to evaluate your artifact.
- **Publicly available?:** Will your artifact be publicly available? If yes, we may spend an extra effort to help you with the documentation.
- **Code licenses (if publicly available)?:** If you workflows and artifacts will be publicly available, please provide information about licenses. This will help the community to reuse your components.
- **Code licenses (if publicly available)?:** If you workflows and artifacts will be publicly available, please provide information about licenses. This will help the community to reuse your components.
- **Workflow frameworks used?** Did authors use any workflow framework which can automate and customize experiments?
- **Archived?:** Note that the author-created artifacts relevant to this paper will receive the ACM "artifact available" badge *only if* they have been placed on a publicly accessible archival repository such as [Zenodo](#), [FigShare](#) or [Dryad](#). A DOI will be then assigned to their artifacts and must be provided here! Personal web pages, Google Drive, GitHub, GitLab and BitBucket are not accepted for this badge. Authors can provide this link at the end of the evaluation.

Description

How to access

Describe, how reviewers can access your artifact:

- Clone repository from GitHub, GitLab, BitBucket or any similar service
- Download package from a public website
- Download package from a private website (you will need to send information how to access your artifact to AE chairs)
- Access artifact via private machine with pre-installed software (only when access to rare hardware is required or proprietary software is used - you will need to send information and credentials to access your machine to the AE chairs)

Please describe approximate disk space required after unpacking your artifact (to avoid surprises when artifact requires 20GB of free space). We do not have a strict limit but strongly suggest to limit the space to several GB and avoid including unnecessary software packages to your VM images.

Hardware dependencies

Describe any specific hardware and specific features required to evaluate your artifact (vendor, CPU/GPU/FPGA, number of processors/cores, interconnect, memory, hardware counters, etc).

Software dependencies

Describe any specific OS and software packages required to evaluate your artifact. This is particularly important if you share your source code and it must be compiled or if you rely on some proprietary software that you can not include to your package. In such case, we strongly suggest you to describe how to obtain and to install all third-party software, data sets and models.

Note that we are trying to obtain AE licenses for some commonly used proprietary tools and benchmarks - you will be informed in case of positive outcome.

Data sets

If third-party data sets are not included in your packages (for example, they are very large or proprietary), please provide details about how to download and install them. *In case of proprietary data sets, we suggest you provide reviewers a public alternative subset for evaluation.*

Models

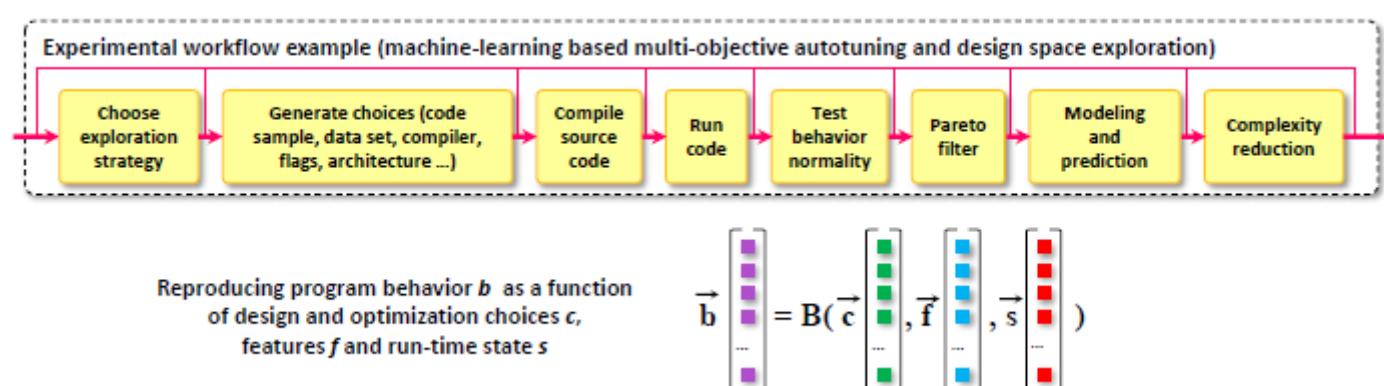
If third-party models are not included in your packages (for example, they are very large or proprietary), please provide details about how to download and install them.

Installation

Describe the setup procedures for your artifact targeting novice users (even if you use VM image or access to remote machine) .

Experiment workflow

Describe the experimental workflow and how it is implemented, invoked and customized (if needed), i.e. some OS scripts, [IPython/Jupyter notebook](#), [portable CK workflow](#), etc. See [past reproduced papers](#) and an example of the experimental workflow for multi-objective and machine-learning based autotuning:



Evaluation and expected result

Describe all the steps necessary to reproduce the key results from your paper. Describe the expected result and the maximum allowable variation of empirical results (particularly important for performance numbers and speed-ups). See the [SIGARCH empirical checklist](#), the [NeurIPS reproducibility checklist](#) and the [AE FAQ](#).

Experiment customization

It is currently optional since it is not always trivial. If possible, describe how to customize your workflow, i.e. if it is possible to use different data sets, benchmarks, real applications, predictive models, software environment (compilers, libraries, run-time systems), hardware, etc. Also, describe if it is possible to parameterize your workflow (whatever is applicable such as changing number of threads, applying different optimizations, CPU/GPU frequency, autotuning scenario, model topology, etc).

Notes

You can add informal notes to draw the attention of reviewers about specific requirements to evaluate your artifact.

After introducing the unified artifact appendix and checklist in 2015, we continue updating it based on public [discussions](#), [past Artifact Evaluation experience](#), [the feedback from researchers](#), [PL AE](#), and the [ACM reviewing and badging policy](#). Our goal is to come up with a common methodology, [unified artifact appendix with the reproducibility checklist](#) and an [open reproducibility platform](#) for artifact sharing, validation and reuse. If you have questions or suggestions, do not hesitate to get in touch with the AE community using the [dedicated AE google group](#).

This document was prepared by [Grigori Fursin](#) with contributions from [Bruce Childers](#), [Michael Heroux](#), [Michela Tauffer](#) and other colleagues.

Maintained by [Grigori Fursin](#) and the [cTuning foundation](#) (2015-cur.)

