Checklist for FAIR Sharing of Data in Autotuning Research

This document contains a checklist of recommended information to share. Information that gets automatically collected by using our scripts [1] or those that are present in our proposed JSON schemas [2] is marked by ⊡.

- General information
  - ☐ name of the dataset for easier future reference
  - ☐ DOI and link to repository
  - ☐ contact information to authors
  - ☐ how to cite
  - ☐ licence and usage restrictions
  - ☐ link to related papers
- Measurements
  - ⊡ kernel experiment time
  - ⊡ validation time
  - ⊡ compilation time
  - ⊡ overhead details (search method overhead, autotuner overhead, model overhead)
  - ⊡ timestamp
  - ⊡ if possible, additional measurements, such as power consumption, profiling counters or clock frequencies
- Tuning space
  - ⊡ names and types of tuning parameters
  - ⊡ values or ranges of tuning parameters
  - ⊡ conditions of tuning parameters
  - ☐ details about how different types of invalid data points are handled
  - ⊡ details about how the results are validated
  - ☐ description of the effects of tuning parameters
  - ☐ details about search space, i.e. raw dataset or run dataset
- Computational problem and its implementation
  - ☐ explanation for non-experts

[1] `https://github.com/odgaard/TuningSchema/blob/T4/metadata.py`

[2] `https://github.com/odgaard/TuningSchema/blob/T4/TuningSchema.json` and `https://github.com/odgaard/TuningSchema/blob/T4/resultsSchema.json`

  - ☐ common programming patterns it fits into
  - ☐ memory- or compute-bound
  - ☐ source code location and version
  - ⊡ programming language used
  - ⊡ grid and thread size
  - ⊡ kernel argument details
- Search method and models
  - ⊡ hyperparameters of the search method
  - ⊡ budget
  - ⊡ performance metric and optimization objective function
  - ☐ details about how models were created and trained
- Environment and execution
  - Input data
    - ☐ size and other relevant characteristics
    - ☐ whether it is included within the dataset
  - Hardware
    - ⊡ details about the device and the model
    - ⊡ chipsets and memory specifics
    - ☐ details about how power consumption is measured
    - ⊡ details provided by the recommended Supercomputing conference environment script
  - Software
    - ⊡ software specifics, OS and compilers
    - ⊡ details about compilation
    - ⊡ details about execution environment
  - Data processing
    - ☐ details about how data were acquired
    - ☐ details about how the autotuner was set and executed
    - ☐ details about data processing and filtering
    - ☐ if relevant, details about analysis and visualization
    - ☐ software and scripts used for dataset acquisition, processing, analysis and visualization