HIFIS HELMHOLTZ FEDERATED IT SERVICES

HELMHOLTZ Open Science

HIRSE

HiDA HELMHOLTZ Information & Data Science Academy

Helmholtz Software Award

Helmholtz Software Award Evaluation Criteria

Introduction

Scientific software plays a crucial role in research, enabling scientists to conduct experiments, analyze data, and develop new methodologies. The quality and impact of such software vary widely, making it essential to evaluate them systematically.

There are numerous software projects in the Helmholtz Association, many of which are outstanding research achievements. The Helmholtz Research Software Award is intended to promote the development of professional, high-quality and sustainable research software and to recognize the commitment to software as the basis of modern data science.

This document outlines the relevant key criteria[1] for assessing scientific software that has been submitted for the Helmholtz Research Software Award; focusing on

A: Scientific Impact,

B: Use and Adoption within the Community,

C: Software Engineering Level,

D: Implementation of FAIR4RS and Open Science Principles, and

E: Young Software with High Growth Rate.

Each criterion is rated on a scale from fair (1 point) to outstanding (5 points).

[1] State 14.07.2025. The criteria were developed by the Helmholtz Task Group Research Software and the Helmholtz Open Science Office.

Criteria

A: Scientific Impact

Definition

This criterion evaluates the scientific impact of the software. Research software is regarded as high-impact if it demonstrably contributes substantial value to scientific research. A software has scientific impact if it e. g. enhances the analytic capabilities of research, or makes possible efficiency gains of research projects, or makes possible tackling new research questions. This is demonstrated by citations in in high quality publications or awards.

Indicators

Narratives explaining impact, quality and number of citations demonstrating impact, awards.

B: Use and Adoption within the Community

Definition

This measures the extent to which the software is used and adopted by the scientific target community. A widely adopted software benefits many users and integrates into standard workflows.

Indicators

community engagement expressed by number and distribution of researchers and research groups, references in research papers, download statistics, GitHub stars.

C: Software Engineering Level

Definition

High-quality engineering ensures maintainability, reproducibility, and reliability, long-term sustainability and ease of use.

Indicators

Development process, quality of documentation, test automation, continuous integration, version control, and adherence to best practices.

D: Implementation of FAIR4RS and Open Science Principles

Definition

The FAIR4RS principles ensure that scientific software and its reference data can be efficiently used, cited, and maintained. The FAIR4RS principles are part of the broader Open Science concept, in the current context this refers e. g. to open-source licensing.

Indicators

Quality of (citation) metadata, continuity of release citations, re-use potential, sustainability, adherence to open standards.

E: Young Software with High Growth Rate

Definition

The purpose of this category is acknowledgement of high potential software. High potential is to be understood as what is to be expected in the future of this software. Young but rapidly growing software may not yet have an extensive user base but shows strong momentum. The criterion can equally apply to establish software if the potential is connected to a major new development.

Indicators

Respective information in motivation letter from developers, high growth rate, Git statistics (commit frequency, contributor growth), user feedback, community interest.