Artifact Evaluations for Stronger Research Results

Dirk Beyer
dirk.beyer@sosy-lab.org

Stefan Winter
sw@stefan-winter.net

Tutorial Materials

- https://www.stefan-winter.net/ae-materials.html
- Linked from FSE 2025 program

Outline and Objectives

- Introduction and historical perspective
- ACM's artifact evaluation (AE) policy and terminology
- AE processes
- Common problems with research artifacts
- Timeline for AE chairs
- Recommendations for authors
- Recommendations for reviewers

Introduction and Historical Perspective

The Reproducibility Crisis in Science

- 2016: $>70\,\%$ of 1576 scientists unable to reproduce peers' results (https://doi.org/10.1038/533452a)
- In the following years: Numerous confirming reports (see https://en.wikipedia.org/wiki/Replication_crisis)
- Many follow-up studies, mainly focused on medical and social sciences
- How about computer science?

The Reproducibility Crisis in Science

- 2016: $>70\,\%$ of 1576 scientists unable to reproduce peers' results (https://doi.org/10.1038/533452a)
- In the following years: Numerous confirming reports (see https://en.wikipedia.org/wiki/Replication_crisis)
- Many follow-up studies, mainly focused on medical and social sciences
- How about computer science?
- NASEM report 2019: Root cause for non-reproducibility often lies in *artifact* deficiencies (https://doi.org/10.17226/25303)

The Reproducibility Crisis in Science

- 2016: $>70\,\%$ of 1576 scientists unable to reproduce peers' results (https://doi.org/10.1038/533452a)
- In the following years: Numerous confirming reports (see https://en.wikipedia.org/wiki/Replication_crisis)
- Many follow-up studies, mainly focused on medical and social sciences
- How about computer science?
- NASEM report 2019: Root cause for non-reproducibility often lies in *artifact* deficiencies (https://doi.org/10.17226/25303)
- Artifact evaluations in software engineering since 2011

Artifacts in Software Engineering (SE) and Programming Languages (PL) Research

Artifact: "a digital object that was either created by the authors to be used as part of the study or generated by the experiment itself"
https://www.acm.org/publications/policies/artifact-review-and-badging-current

Examples:

- Software tools
- Scripts to run experiments
- Data (raw or processed/aggregated)
- Documentation
- Mathematical proof (manual or automated)
- Audio and video materials

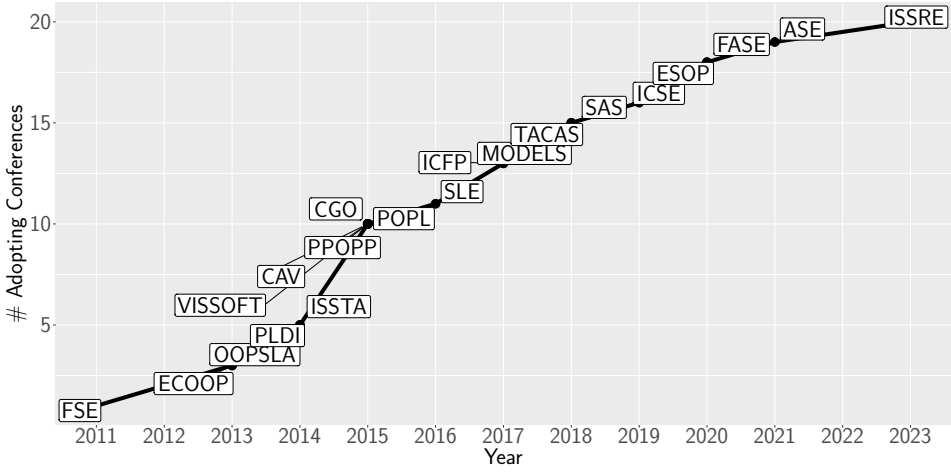
Systematic Assessment of Artifacts in SE/PL: Artifact Evaluations

- Pioneered at FSE 2011 & OOPSLA 2013
- Initial criteria (https://artifact-eval.org/about.html)
  - consistent with the paper
  - as complete as possible
  - well documented
  - easy to reuse, facilitating further research

Systematic Assessment of Artifacts in SE/PL: Artifact Evaluations

- Pioneered at FSE 2011 & OOPSLA 2013
- Initial criteria (https://artifact-eval.org/about.html)
  - consistent with the paper
  - as complete as possible
  - well documented
  - easy to reuse, facilitating further research

Artifact Evaluation – Adoption in SE/PL



Evolution of Artifact Badges

Initial criteria
(https://artifact-eval.org/about.html)

- consistent with the paper
- as complete as possible
- well documented
- easy to reuse, facilitating further research

Evolution of Artifact Badges

Initial criteria
(https://artifact-eval.org/about.html)

- consistent with the paper
- as complete as possible
- well documented
- easy to reuse, facilitating further research
- Problem: *Prototypes* maybe not easy to reuse, but useful for reproducibility

Evolution of Artifact Badges

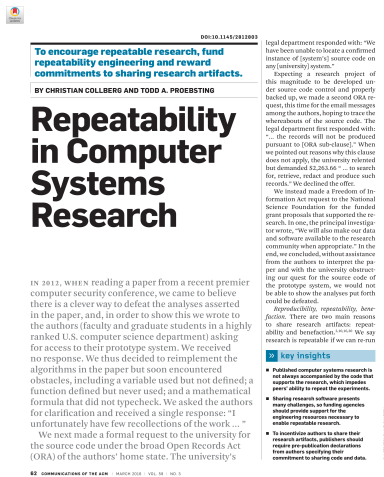
Orthogonal aspects:

- Functionality
- Reusability

Evolution of Artifact Badges

Orthogonal aspects:

- Functionality
- Reusability
- Availability
  - Intellectual property, licensing
  - Security
  - Privacy



contributed articles

DOI:10.1145/2812803

To encourage repeatable research, fund repeatability engineering and reward commitments to sharing research artifacts.

BY CHRISTIAN COLLBERG AND TODD A. PROEBSTING

Repeatability in Computer Systems Research

IN 2012, WHEN reading a paper from a recent premier computer security conference, we came to believe there is a clever way to defeat the analyses asserted in the paper, and, in order to show this we wrote to the authors (faculty and graduate students in a highly ranked U.S. computer science department) asking for access to their prototype system. We received no response. We thus decided to reimplement the algorithms in the paper but soon encountered obstacles, including a variable used but not defined; a function defined but never used; and a mathematical formula that did not typecheck. We asked the authors for clarification and received a single response: "I unfortunately have few recollections of the work ... "

We next made a formal request to the university for the source code under the broad Open Records Act (ORA) of the authors' home state. The university's legal department responded with: "We have been unable to locate a confirmed instance of [system's] source code on any [university] system."

Expecting a research project of this magnitude to be developed under source code control and properly backed up, we made a second ORA request, this time for the email messages among the authors, hoping to trace the whereabouts of the source code. The legal department first responded with: "... the records will not be produced pursuant to [ORA sub-clause]." When we pointed out reasons why this clause does not apply, the university relented but demanded $2,263.66 "... to search for, retrieve, redact and produce such records." We declined the offer.

We instead made a Freedom of Information Act request to the National Science Foundation for the funded grant proposals that supported the research. In one, the principal investigator wrote, "We will also make our data and software available to the research community when appropriate." In the end, we concluded, without assistance from the authors to interpret the paper and with the university obstructing our quest for the source code of the prototype system, we would not be able to show the analyses put forth could be defeated.

Reproducibility, repeatability, benefaction. There are two main reasons to share research artifacts: reproducibility and benefaction.[10,16,18] We say research is repeatable if we can re-run

key insights

- Published computer systems research is not always accompanied by the code that supports the research, which impedes peers' ability to repeat the experiments.
- Sharing research software presents many challenges, so funding agencies should provide support for the engineering resources necessary to enable repeatable research.
- To incentivize authors to share their research artifacts, publishers should require pre-publication declarations from authors specifying their commitment to sharing code and data.

© COPYRIGHT HELD BY AUTHORS/OWNERS

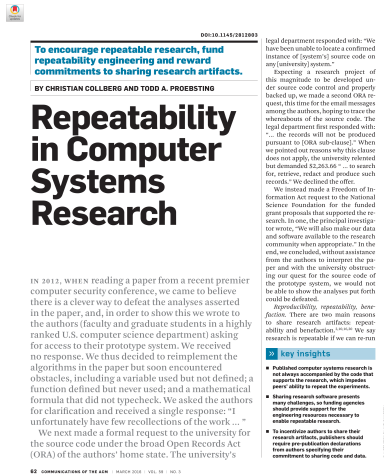
https://doi.org/10.1145/2812803

Evolution of Artifact Badges

Orthogonal aspects:

- Functionality
- Reusability
- Availability
  - Intellectual property, licensing
  - Security
  - Privacy

→ Separation of concerns



contributed articles

DOI:10.1145/2812803

To encourage repeatable research, fund repeatability engineering and reward commitments to sharing research artifacts.

BY CHRISTIAN COLLBERG AND TODD A. PROEBSTING

Repeatability in Computer Systems Research

IN 2012, WHEN reading a paper from a recent premier computer security conference, we came to believe there is a clever way to defeat the analyses asserted in the paper, and, in order to show this we wrote to the authors (faculty and graduate students in a highly ranked U.S. computer science department) asking for access to their prototype system. We received no response. We thus decided to reimplement the algorithms in the paper but soon encountered obstacles, including a variable used but not defined; a function defined but never used; and a mathematical formula that did not typecheck. We asked the authors for clarification and received a single response: "I unfortunately have few recollections of the work ... "

We next made a formal request to the university for the source code under the broad Open Records Act (ORA) of the authors' home state. The university's legal department responded with: "We have been unable to locate a confirmed instance of [system's] source code on any [university] system."

Expecting a research project of this magnitude to be developed under source code control and properly backed up, we made a second ORA request, this time for the email messages among the authors, hoping to trace the whereabouts of the source code. The legal department first responded with: "... the records will not be produced pursuant to [ORA sub-clause]." When we pointed out reasons why this clause does not apply, the university relented but demanded $2,263.66 "... to search for, retrieve, redact and produce such records." We declined the offer.

We instead made a Freedom of Information Act request to the National Science Foundation for the funded grant proposals that supported the research. In one, the principal investigator wrote, "We will also make our data and software available to the research community when appropriate." In the end, we concluded, without assistance from the authors to interpret the paper and with the university obstructing our quest for the source code of the prototype system, we would not be able to show the analyses put forth could be defeated.

Reproducibility, repeatability, benefaction. There are two main reasons to share research artifacts: reproducibility and benefaction.[10,16,18] We say research is repeatable if we can re-run

key insights

- Published computer systems research is not always accompanied by the code that supports the research, which impedes peers' ability to repeat the experiments.
- Sharing research software presents many challenges, so funding agencies should provide support for the engineering resources necessary to enable repeatable research.
- To incentivize authors to share their research artifacts, publishers should require pre-publication declarations from authors specifying their commitment to sharing code and data.

© COPYRIGHT HELD BY AUTHOR/OWNER

62 COMMUNICATIONS OF THE ACM | MARCH 2018 | VOL. 58 | NO. 3

https://doi.org/10.1145/2812803

The Current Artifact Badging System

Artifact Evaluation – Revised Badges

- ACM initiative 2017:

  https://www.acm.org/publications/task-force-on-data-software-and-reproducibility

- Minor revision 2020 for compliance with NISO RP-31-2021:

  https://doi.org/10.3789/niso-rp-31-2021

Artifact Evaluation – Revised Badges

- ACM initiative 2017:

  https://www.acm.org/publications/task-force-on-data-software-and-reproducibility

- Minor revision 2020 for compliance with NISO RP-31-2021:

  https://doi.org/10.3789/niso-rp-31-2021

Does the artifact work?

Artifact Evaluation – Revised Badges

- ACM initiative 2017:

  https://www.acm.org/publications/task-force-on-data-software-and-reproducibility

- Minor revision 2020 for compliance with NISO RP-31-2021:

  https://doi.org/10.3789/niso-rp-31-2021

Does the artifact work?

Is the artifact permanently available?

Artifact Evaluation – Revised Badges

- ACM initiative 2017:

  https://www.acm.org/publications/task-force-on-data-software-and-reproducibility

- Minor revision 2020 for compliance with NISO RP-31-2021:

  https://doi.org/10.3789/niso-rp-31-2021

Does the artifact work?

Is the artifact permanently available?

Can the results be confirmed?

Artifact Evaluation – Alternative Badges

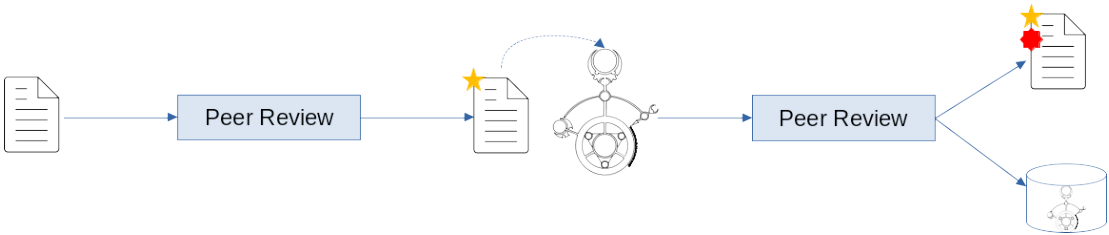
Alternative badges from other publishers have similar semantics:

Artifact Evaluation – Alternative Badges

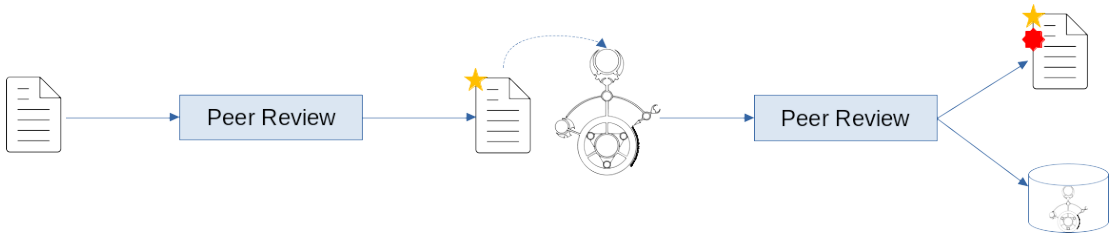
Alternative badges from other publishers have similar semantics:

ACM permits badge usage outside ACM if they comply with ACM definitions

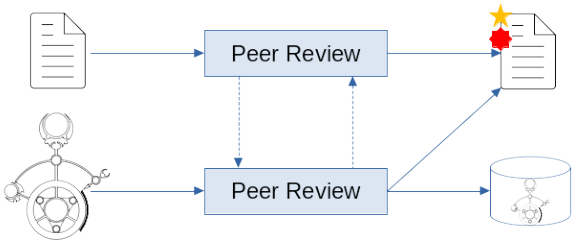
Artifact Evaluation – Processes



Artifact Evaluation – Processes



TACAS & CAV (tool papers), ECOOP (2022-2024)



ACM Artifact Policies and Terms

Different Levels of Research Reliability

Repeatability The measurement can be obtained with stated precision by the **same team** using the **same measurement procedure**, the **same measuring system**, under the **same operating conditions**, in the **same location** on multiple trials. For computational experiments, this means that a researcher can reliably repeat her own computation.

https://www.acm.org/publications/policies/artifact-review-and-badging-current

Different Levels of Research Reliability

Repeatability The measurement can be obtained with stated precision by the **same team** using the **same measurement procedure**, the **same measuring system**, under the **same operating conditions**, in the **same location** on multiple trials. For computational experiments, this means that a researcher can reliably repeat her own computation.

Reproducibility The measurement can be obtained with stated precision by a **different team** using the same measurement procedure, the same measuring system, under the same operating conditions, in the **same or a different location** on multiple trials. For computational experiments, this means that an independent group can obtain the same result using the author's own artifacts.

https://www.acm.org/publications/policies/artifact-review-and-badging-current

Different Levels of Research Reliability

Repeatability The measurement can be obtained with stated precision by the **same team** using the **same measurement procedure**, the **same measuring system**, under the **same operating conditions**, in the **same location** on multiple trials. For computational experiments, this means that a researcher can reliably repeat her own computation.

Reproducibility The measurement can be obtained with stated precision by a **different team** using the same measurement procedure, the same measuring system, under the same operating conditions, in the **same or a different location** on multiple trials. For computational experiments, this means that an independent group can obtain the same result using the author's own artifacts.

Replicability The measurement can be obtained with stated precision by a **different team**, a **different measuring system**, in a **different location** on multiple trials. For computational experiments, this means that an independent group can obtain the same result using artifacts which they develop completely independently.

https://www.acm.org/publications/policies/artifact-review-and-badging-current

Different Levels of Research “Reliability”

Repeatability Same team, same setup

Reproducibility Different team, same setup

Replicability Different team, different setup

https://www.acm.org/publications/policies/artifact-review-and-badging-current

Note: *Setup* includes the measured subject (e.g., software).

Different Levels of Research "Reliability"

Repeatability Same team, same setup

Reproducibility Different team, same setup

Replicability Different team, different setup

https://www.acm.org/publications/policies/artifact-review-and-badging-current

Note: *Setup* includes the measured subject (e.g., software).

Important: Only reproducibility mandates artifact (*setup*) sharing.

Terminological Variations

Property	NISO[1] (ACM since 2020)	VIM[2] (ACM pre 2020)
Repeatability	Same team, same setup	Same team, same setup
Reproducibility	Different team, same setup	Different team, different setup
Replicability	Different team, different setup	Different team, same setup

[1] https://doi.org/10.3789/niso-rp-31-2021

[2] https://doi.org/10.59161/JCGM200-2012

ACM Badge Categories

Artifacts evaluated

Artifacts available

Results validated

These [badge categories] are considered independent and any one, two or all three can be applied to any given paper depending on review procedures developed by the journal or conference.

Artifacts Evaluated

Two levels are distinguished, only one of which should be applied in any instance

- Documented
- Consistent
- Complete
- Exercisable

- All properties of “Functional”
- Well documented and structured
- Meets community norms and standards

Artifacts Available

- Publicly accessible
- DOI or link + unique object identifier
- Long-term retention policy (≥ 10y)
  - Zenodo, Figshare, Dryad, ...
  - *Not* GitHub, institute website, ...
- "Formal evaluation" not strictly needed

Results Validated

- Subsequent study from other authors exists
  - uses some of the original work's artifacts
  - confirms results
  - deviations from exact results tolerable if conclusions do not change
- By definition not artifact-related

Recommendations for AE Organizers

Timeline

1. Process/submission-system decisions
2. Load/resources planning/decisions
3. AE committee assembly
4. AE timeline planning and call for artifacts (CfA)
5. Awards and process evaluation planning
6. Bidding, Assignment, Evaluation
7. Publication preparations

Timeline

1. Process/submission-system decisions
2. Load/resources planning/decisions
3. AE committee assembly
4. AE timeline planning and call for artifacts (CfA)
5. Awards and process evaluation planning
6. Bidding, Assignment, Evaluation
7. Publication preparations

Load and Committee Planning

- AE takes much time: ≤ 5 artifacts/person
- Recommendation: 3 artifacts/person & 3 reviews/artifact
  $\rightarrow$ Committee size = number expected submissions ($+$ x %)

Load and Committee Planning

- AE takes much time: ≤ 5 artifacts/person
- Recommendation: 3 artifacts/person & 3 reviews/artifact
  $\rightarrow$ Committee size = number expected submissions ($+$ x %)
- Whom to recruit: Mix of junior (30-40 %) and senior artifact reviewers
  https://doi.org/10.1145/3368089.3409767

Load and Committee Planning

- AE takes much time: ≤ 5 artifacts/person
- Recommendation: 3 artifacts/person & 3 reviews/artifact
  $\rightarrow$ Committee size = number expected submissions ($+$ x %)
- Whom to recruit: Mix of junior (30-40 %) and senior artifact reviewers
  https://doi.org/10.1145/3368089.3409767
- How to recruit: PC/open nominations, scan prior committees

The Call for Artifacts (CfA)

- Timeline: Submission $\rightarrow$ bidding $\rightarrow$ pre-assessment ("kicking-the-tires") $\rightarrow$ review $\rightarrow$ submission $\rightarrow$ discussion $\rightarrow$ author notification

The Call for Artifacts (CfA)

- Timeline: Submission → bidding → pre-assessment ("kicking-the-tires") → review → submission → discussion → author notification
- Pre-assessment and author response: Early! 3-10 days after assignment

The Call for Artifacts (CfA)

- Timeline: Submission → bidding → pre-assessment ("kicking-the-tires") → review → submission → discussion → author notification
- Pre-assessment and author response: Early! 3-10 days after assignment
- Discussion: Min. 1 week

The Call for Artifacts (CfA)

- Timeline: Submission → bidding → pre-assessment ("kicking-the-tires") → review → submission → discussion → author notification
- Pre-assessment and author response: Early! 3-10 days after assignment
- Discussion: Min. 1 week
- General CfA advice: Do not restrict nor extend ACM's badge definitions

The Call for Artifacts (CfA)

- Timeline: Submission → bidding → pre-assessment ("kicking-the-tires") → review → submission → discussion → author notification
- Pre-assessment and author response: Early! 3-10 days after assignment
- Discussion: Min. 1 week
- General CfA advice: Do not restrict nor extend ACM's badge definitions
- Provide clear guidance for packaging and documentation

Awards

- Academic evaluations and reward based on citation counts
- Evaluated artifacts not linked with higher citation counts
  $\rightarrow$ Alternative reward mechanisms needed!
  https://doi.org/10.1145/3540250.3549172
- Distinguished artifact/reviewer awards as intermediate remedy

Publication Preparations

- Send badge information for submissions to publication chairs
  (also send the submission numbers/IDs for the *papers*)
- Describe AE process, outcome, committee in proceedings preface

Publication Preparations

- Send badge information for submissions to publication chairs
  (also send the submission numbers/IDs for the *papers*)
- Describe AE process, outcome, committee in proceedings preface
- If possible:

Publication Preparations

- Send badge information for submissions to publication chairs
  (also send the submission numbers/IDs for the *papers*)
- Describe AE process, outcome, committee in proceedings preface
- If possible:
  - Check camera-ready papers for artifact links
    Ideally insist on data availability statement at the end of the paper

Publication Preparations

- Send badge information for submissions to publication chairs
  (also send the submission numbers/IDs for the *papers*)
- Describe AE process, outcome, committee in proceedings preface
- If possible:
  - Check camera-ready papers for artifact links
    Ideally insist on data availability statement at the end of the paper
  - Check artifacts behind such links to correspond to evaluated version

Recommendations for Artifact Authors

Artifact Badges

Unless there are legal or ethical restrictions, go for the Available badge.

Artifact Badges

Unless there are legal or ethical restrictions, go for the Available badge.

AE Evaluated	Available Badge Status	Papers with Artifact Link	Is Accessible
AE	Av. Badge	676	675 (99.9%)
	No Av. Badge	473	431 (91.1%)
NonAE	Av. Badge	67	65 (97.0%)
	No Av. Badge	1148	1032 (89.9%)

https://doi.org/10.1145/3540250.3549172

Artifact Badges

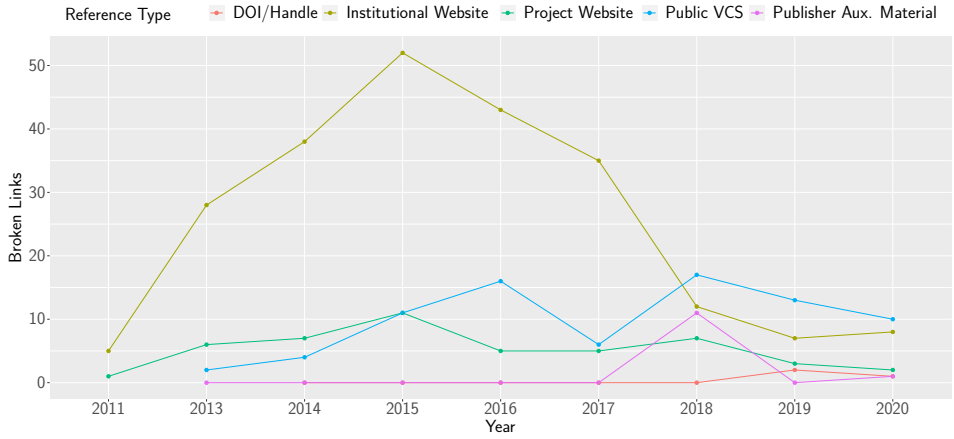
Unless there are legal or ethical restrictions, go for the Available badge.

AE Evaluated	Available Badge Status	Papers with Artifact Link	Is Accessible
AE	Av. Badge	676	675 (99.9%)
	No Av. Badge	473	431 (91.1%)
NonAE	Av. Badge	67	65 (97.0%)
	No Av. Badge	1148	1032 (89.9%)

https://doi.org/10.1145/3540250.3549172

Artifact Hosting

If you go for Available, use a DOI-issuing platform for submission.



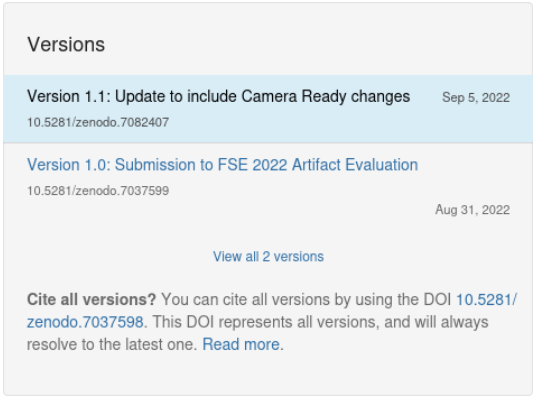
https://doi.org/10.1145/3540250.3549172

Artifact Hosting

If you use a DOI-issuing platform for submission, don't rely on *concept DOIs*.

Artifact Hosting

If you use a DOI-issuing platform for submission, don't rely on *concept DOIs*.



Artifact Packaging

Include a Docker container or VM, if possible.

- Reduces overhead for reviewers

Artifact Packaging

Include a Docker container or VM, if possible.

- Reduces overhead for reviewers
- Reduces likelihood of missed dependency documentation

Artifact Packaging

Include a Docker container or VM, if possible.

- Reduces overhead for reviewers
- Reduces likelihood of missed dependency documentation
- Reduces likelihood of result deviations

Artifact Packaging

Include a Docker container or VM, if possible.

- Reduces overhead for reviewers
- Reduces likelihood of missed dependency documentation
- Reduces likelihood of result deviations
- Ideally also ship Docker/Vagrant files for transparency

Artifact Documentation

Follow the CfA's documentation requirements (if none, use FSE 2018) and include a license (as file)

Search Term	Matched Artifacts		Avg. Word Count	
	AE	NonAE	AE	NonAE
No match	13	12	–	–
^read.*me	84	86	**1 389**	645
^install	**6**	1	324	593
^doc/	1	8	2 431	13 901
^copyright	0	1	0	268
^license	**50**	46	850	1 220

https://doi.org/10.1145/3540250.3549172

What to Include in an Artifact README?

- Abstract: summarize contents, purpose, and required computing resources
- Contents: list the important files/directories and table of contents
- TL;DR: simple instructions to run the tools and perform small experiments
- System requirements and installation guide
- Instructions to execute the tools: command line, input/output, configurations
- Instructions to perform the experiments: demo and full evaluation
- Instructions to process and understand the experimental results
- Others: known issues and reference logs for listed commands

README — Takeaways

- Write artifact README like a short paper

README — Takeaways

- Write artifact README like a short paper
- Do not assume the users/reviewers have prior knowledge about your tool

README — Takeaways

- Write artifact README like a short paper
- Do not assume the users/reviewers have prior knowledge about your tool
- Selection criteria for demo runs (smoke tests):
  - Can finish within a reasonable amount of time (1-2 hr)

README — Takeaways

- Write artifact README like a short paper
- Do not assume the users/reviewers have prior knowledge about your tool
- Selection criteria for demo runs (smoke tests):
  - Can finish within a reasonable amount of time (1-2 hr)
  - No errors/exceptions/timeouts (otherwise make sure to explain them)

README — Takeaways

- Write artifact README like a short paper
- Do not assume the users/reviewers have prior knowledge about your tool
- Selection criteria for demo runs (smoke tests):
  - Can finish within a reasonable amount of time (1-2 hr)
  - No errors/exceptions/timeouts (otherwise make sure to explain them)
  - Ideally, the selected tasks can partially validate the claims in the paper

README — Takeaways

- Write artifact README like a short paper
- Do not assume the users/reviewers have prior knowledge about your tool
- Selection criteria for demo runs (smoke tests):
  - Can finish within a reasonable amount of time (1-2 hr)
  - No errors/exceptions/timeouts (otherwise make sure to explain them)
  - Ideally, the selected tasks can partially validate the claims in the paper
- Check **consistency** between the paper and artifact!

Recommendations for Artifact Reviewers

Artifact Hosting

- Many hosting platforms undermine double-blind reviews by tracking IP addresses
  - URL redirects
  - Personal/institute websites
  - ...

Artifact Hosting

- Many hosting platforms undermine double-blind reviews by tracking IP addresses
  - URL redirects
  - Personal/institute websites
  - ...
- Solution: obfuscate IP address (proxies, tor)

Artifact Hosting

- Many hosting platforms undermine double-blind reviews by tracking IP addresses
  - URL redirects
  - Personal/institute websites
  - ...
- Solution: obfuscate IP address (proxies, tor)
- Better: Inform AEC chairs

Artifact Pre-Assessment ("Kicking-the-Tires")

- Can the artifact be downloaded?
- Are HW requirements (GPU, x86-46 vs. ARM silicon) met?
- Are input data or external software dependencies included or (if not) accessible?

The sooner you realize, the sooner the AEC chairs can react.

Serve the Community

- Almost every artifact is better than no artifact.

Serve the Community

- Almost every artifact is better than no artifact.
- Help authors improve artifact quality.

Serve the Community

- Almost every artifact is better than no artifact.
- Help authors improve artifact quality.
- Reject broken artifacts that cannot or will not be improved.

Checklists

- Goal: Fair and unbiased assessments

Checklists

- Goal: Fair and unbiased assessments
- If none provided, create one based on CfA and structure your review accordingly

Checklists

- Goal: Fair and unbiased assessments
- If none provided, create one based on CfA and structure your review accordingly
- For a template, see our extended abstract.

Further Reading & Hands-On Experience

Reading Suggestions

- Christian Collberg, Todd A. Proebsting: "Repeatability in Computer Systems Research" (https://doi.org/10.1145/2812803)
- Robert Heumüller, Sebastian Nielebock, Jacob Krüger, Frank Ortmeier: "Publish or perish, but do not forget your software artifacts" (https://doi.org/10.1007/s10664-020-09851-6)
- Ben Hermann, Stefan Winter, Janet Siegmund: "Community expectations for research artifacts and evaluation processes" (https://doi.org/10.1145/3368089.3409767)
- Christopher S. Timperley, Lauren Herckis, Claire Le Goues, and Michael Hilton: "Understanding and improving artifact sharing in software engineering research" (https://doi.org/10.1007/s10664-021-09973-5)
- Stefan Winter, Christopher S. Timperley, Ben Hermann, Jürgen Cito, Jonathan Bell, Michael Hilton, and Dirk Beyer: "A retrospective study of one decade of artifact evaluations" (https://doi.org/10.1145/3540250.3549172)

Preparations for the Hands-On Part

- Select an artifact from our list:
  https://www.stefan-winter.net/ae-materials.html
- Or from FSE 2025: https://dl.acm.org/toc/pacmse/2025/2/FSE

Preparations for the Hands-On Part

- Select an artifact from our list:
  https://www.stefan-winter.net/ae-materials.html
- Or from FSE 2025: https://dl.acm.org/toc/pacmse/2025/2/FSE
- See if you can find a link to the artifact in the paper

Preparations for the Hands-On Part

- Select an artifact from our list:
  https://www.stefan-winter.net/ae-materials.html
- Or from FSE 2025: https://dl.acm.org/toc/pacmse/2025/2/FSE
- See if you can find a link to the artifact in the paper
- If not: Look at our link list on the website (unless FSE 2025)

Preparations for the Hands-On Part

- Select an artifact from our list:
  https://www.stefan-winter.net/ae-materials.html
- Or from FSE 2025: https://dl.acm.org/toc/pacmse/2025/2/FSE
- See if you can find a link to the artifact in the paper
- If not: Look at our link list on the website (unless FSE 2025)
- Check if it is small enough to download over the coffee break

Preparations for the Hands-On Part

- Select an artifact from our list:
  https://www.stefan-winter.net/ae-materials.html
- Or from FSE 2025: https://dl.acm.org/toc/pacmse/2025/2/FSE
- See if you can find a link to the artifact in the paper
- If not: Look at our link list on the website (unless FSE 2025)
- Check if it is small enough to download over the coffee break
- If not: Copy files from one of our USB drives (unless FSE 2025)

Preparations for the Hands-On Part

- Select an artifact from our list:
  https://www.stefan-winter.net/ae-materials.html
- Or from FSE 2025: https://dl.acm.org/toc/pacmse/2025/2/FSE
- See if you can find a link to the artifact in the paper
- If not: Look at our link list on the website (unless FSE 2025)
- Check if it is small enough to download over the coffee break
- If not: Copy files from one of our USB drives (unless FSE 2025)
- Have fun!