

STM Annual US Conference 2017

The Future Decade of the Researcher

Trends in Peer Review: **Data, Software, and Reproducibility in Publication**



Bernard Rous
ACM Publications Director Emeritus
April 26, 2017
Washington, D.C.



Association for
Computing Machinery

Advancing Computing as a Science & Profession

OUTLINE

- Reproducibility in computer science – context
- ACM Reproducibility Task Force
- Lessons learned
 - Review process
- CS Goals for Artifact Reviewing
- Terminology and Badging
- Pilot Integrations
- Best Practices Summary

Context

- Experimental research
- Tremendous variability
 - Volume and types of data
 - Instrumentation
 - Algorithms
 - Computational resources
- Different traditions re reproducibility
 - Biomedical and Pharmaceutical research
 - Computer science research

ACM Task Force

- ACM Task Force on Reproducibility
 - Working towards common solution
 - Integration with publication
- Reviewer and Reader
 - Similar needs wrt reproducibility
- The Impediments
 - Recreating experimental environments

Lessons from the Field

- Early Days
- Do not mandate artifacts
- Do not tie to article acceptance
- No Single or Double Blind reviewing
- Provide Motivation to develop new habits
 - Credit for Reviewers
 - Branding For authors - Levels of “Reproducibility”

Review of Artifacts

Known Reviewers

- Practical efficiency
- Decoupled from article acceptance
- Different reviewers employed

Goals for Artifact Reviewing

- Develop new habits of documentation and specification
 - Move towards structured metadata descriptions
- Artifacts as primary research objects
 - Not just *supplements* to article
 - Independent DOI for citation and linking
- Enable re-use for further development
 - Encourage liberal user license



ACM Badging: Artifact Evaluation



- Artifacts Evaluated
 - *Functional* - The artifacts associated with the research are found to be documented, consistent, complete, exercisable, and include appropriate evidence of verification and validation.
 - *Reusable* - The artifacts associated with the paper are of a quality that significantly exceeds minimal functionality. That is, they are very carefully documented and well-structured to the extent that reuse and repurposing is facilitated. In particular, norms and standards of the research community for artifacts of this type are strictly adhered to.



ACM Badging: Validation of Results



- Results Validated - This badge is applied to papers in which the main results of the paper have been successfully obtained by a person or team other than the author. Two levels are distinguished:
 - Results Replicated - The main results of the paper have been obtained in a subsequent study by a person or team other than the authors, using, in part, artifacts provided by the author.
 - Results Reproduced - The main results of the paper have been independently obtained in a subsequent study by a person or team other than the authors, without the use of author-supplied artifacts.





Association for
Computing Machinery

Advancing Computing as a

Integration Project

Funded by Sloan Foundation



- Need for integrations
- Three examples
- Video as Independent Artifact
 - DOI = <http://dx.doi.org/10.1145/3076216>



Best Practices Summary

- *Clarify basic definitions, evaluation criteria, and branding for:* replicability, repeatability, reproducibility, re-usability, availability
- *Motivate and incentivize:* authors, reviewers, program committees, editorial boards
- *Adopt/invent standard metadata descriptions:* for software, data, methodologies
- *Enable:* artifact evaluation processes in automated submission workflows
- *Encourage:* sharing of artifacts
- *Define:* acceptable storage and packaging formats
- *Support and integrate:* internal and external data and software depositories
- *Identify, cite, and link:* artifacts as first-class publication objects
- *Curate and preserve:* artifacts for future re-use
- *Develop legal framework:* for artifact owners, users, publishers



Feel Free to Visit (And participate in Survey)

- <http://dl.acm.org/reproducibility.cfm>



BLANK SEPARATOR



Association for
Computing Machinery

Advancing Computing as a Science & Profession

ADDITIONAL MATERIAL

- Following slides provide further lessons learned and examples published in the ACM Digital Library

TOMS

Companion Publication and Supplemental Artifacts

<http://dl.acm.org/citation.cfm?id=2786970&picked=prox&CFID=609384431>

- Editorial description of “Replicated Computational Results Initiative”
- Article Citation Page gets Editorial Note and
 - Link to Reviewer Report
 - Link to Process Description
- PDF gets logo
 - Linked to description
- Reviewer of software gets publication with
 - Links to Author’s article and Editorial Description
- Supplemental Files contain artifact



Further Lessons from the Field

- Artifacts must be first class objects: identifiable, citable, and linkable
 - Standard metadata descriptions, DOI assignment
 - Stand-alone and/or components of article(s)
- Artifact Review and Badging independent of artifact publication
 - Proprietary interest
 - Reader trust



More Lessons and Examples

- Provide Access to artifacts
 - Publisher archived and served
<http://dx.doi.org/10.1145/2699878>
 - External repository links
<http://dx.doi.org/10.1145/2688500.2688501>
- Develop a Legal Framework
 - For serving artifacts
 - Ownership, user rights, publisher liability



Lessons

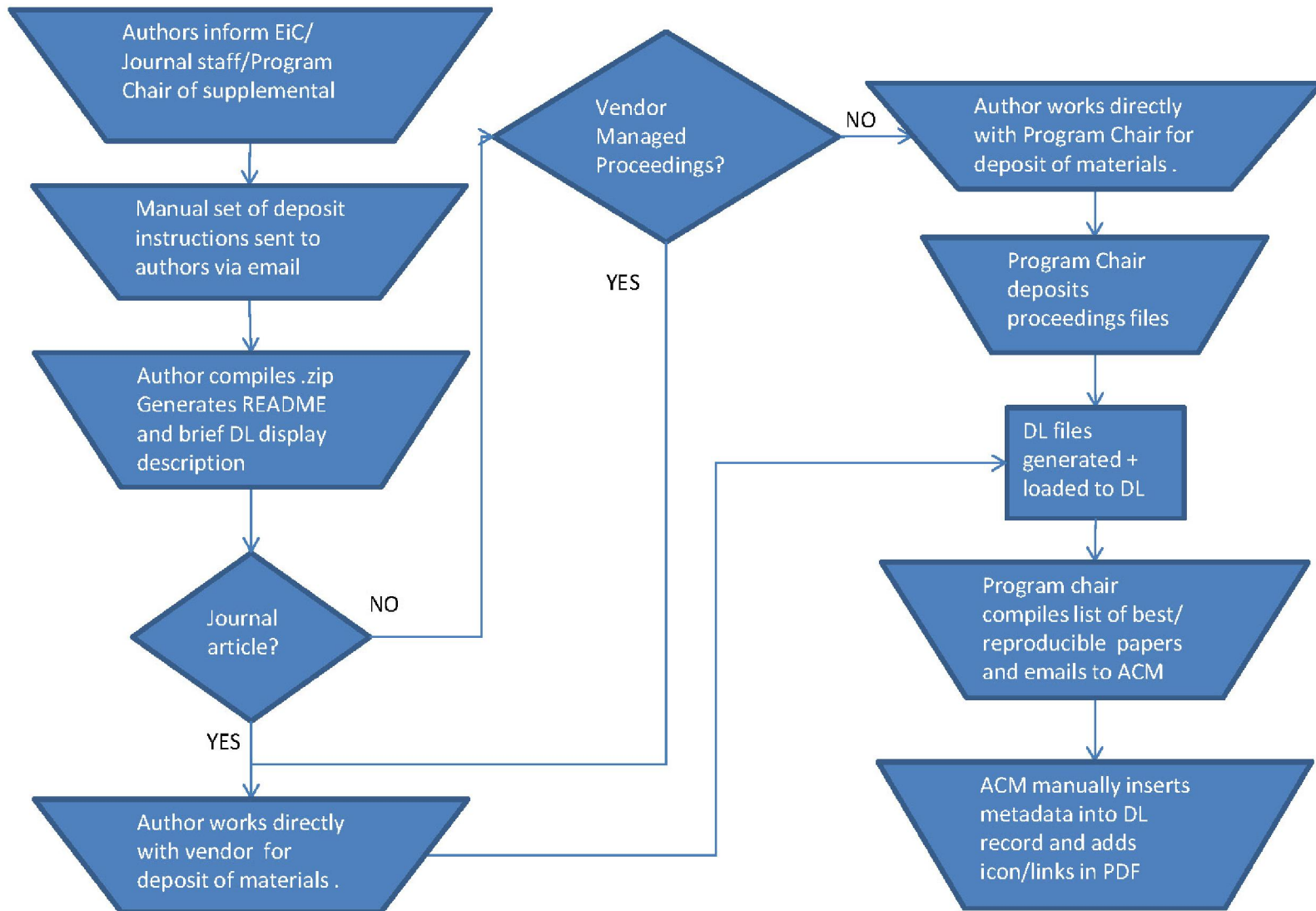
- Author support tools and services
 - Building “wrappers”, encapsulation, lightweight virtual machines
- Integration
 - [External data repositories and software curation platforms](#)



Current Status

- Source materials amassed & organized by Task Force
- Individual journals and conferences deploying review processes and branding
 - Disjoint from article peer review
 - Disjoint from publication
- Manual (post-publication) curation in ACM Digital Library
 - Editorial Notes
 - Links
 - Local branding
 - Supplemental files

Manual Work Flow



Examples SIGMOD



<http://dx.doi.org/10.1145/2723372.2737793>

- The following have a NOTE on the Citation Page about Reproducibility
- The Note contains a link to the Process used to obtain the badge.
- The PDF has the logo. In the case of the first DOI in the list, that logo has an active link to the explanation.



Association for
Computing Machinery

Advancing Computing as a Science & Profession

PPoPP

- The Citation Page for the first DOI has no note but see the Source Materials tab for
 - a link to the AEC explanatory page and
 - a link to github for the artifact
- PDF has the PPoPP reproducibility logo
- <http://dx.doi.org/10.1145/2688500.2688501>

JEA

The ACM Digital Library as Preservation Repository

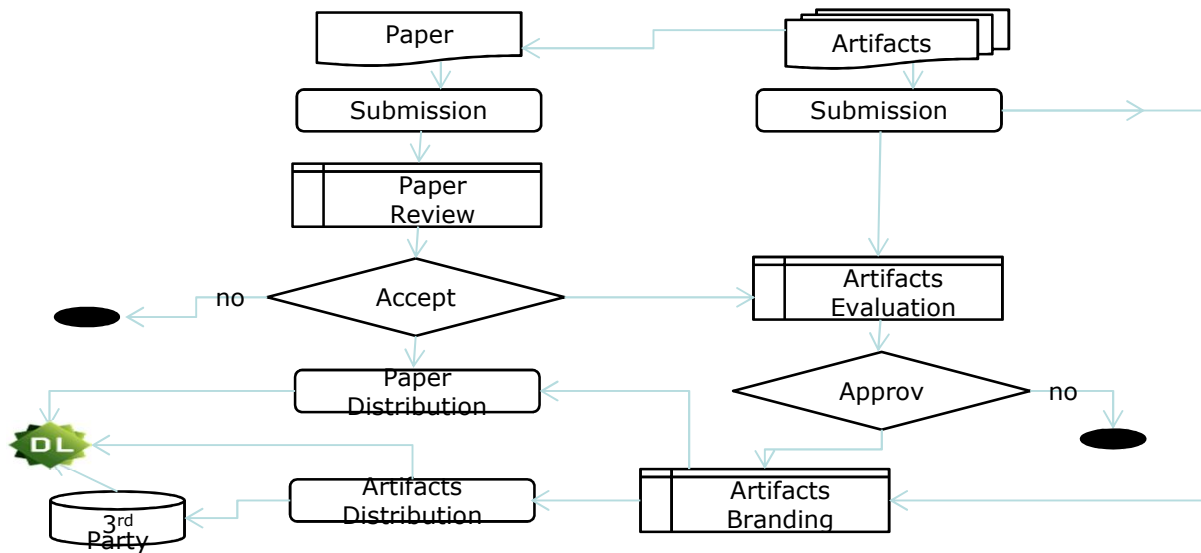
<http://dx.doi.org/10.1145/2699878>

-
- Citation Page has no Editorial Note
- But Source Materials include supplemental files
 - Extensive Readme file “A Guide to using the associated software”
 - The Software
 - Not refereed
 - Ownership, user rights, and disclaimer

TOMACS

- The Citation pages for the two DOIs have no Notes as of yet.
 - They are not linked to each other.
 - But see the PDF for the first DOI – it has the TOMACS reproducible logo (with no link yet to process used)
 - There are no supplemental files.
-
- <http://dx.doi.org/10.1145/2883608>
 - <http://dx.doi.org/10.1145/2893479>





Other Artifact Examples from DL:

CACM, “presentation”

<http://dl.acm.org/citation.cfm?id=1467267>

CHI, “Preview videos”

<http://dl.acm.org/citation.cfm?id=2732509>

CFP, Audio

<http://dl.acm.org/citation.cfm?doid=564566>

SIGGRAPH, Multimedia

<http://dl.acm.org/citation.cfm?doid=945317>



Current Limitation

- All cases are manually curated
 - In various stages of completion
 - Without uniformity of treatment
- No standard definitions, branding, or artifact descriptions
- The current method does not scale.



Cases we expect to support:

All Artifacts and/or associated Papers are “Branded”. Distribution from DL or linked.

Artifacts submitted with papers:

- Evaluation and approval of Artifacts required for publication of paper.
 - Artifacts are available | not available for distribution.
- Evaluation of Artifacts independent of publication of paper.
 - Artifacts are available | not available for distribution.
- No Evaluation of Artifacts performed.
 - Artifacts are available for distribution.

Artifacts only:

- Evaluation and approval required for distribution.
- No Evaluation performed.

Requirements for Scaling

- *Agree basic definitions, evaluation criteria, and branding for:* replicable, repeatable, reproducible, re-usable, verifiable (and availability)
- *Motivate and incentivize:* authors, reviewers, program committees, editorial boards
- *Enable:* artifact evaluation processes in automated submission workflows
- *Provide:* easy-to-use rerun environments
- *Adopt/invent standard metadata descriptions:* for software and for data, standalone or as component of article
- *Identify, cite, and link:* artifacts as first-class publication objects
- *Define:* acceptable storage and packaging formats
- *Encourage/require:* sharing of artifacts
- *Specify legal framework:* for serving and using data and software artifacts
- *Support and integrate:* internal and external data and software depositories
- *Curate and preserve:* artifacts for future re-use



Ideal Workflow

