# Quality Guidelines for Research Artifacts in Model-Driven Engineering

CDN (Diego) Damasceno [a] and Daniel Strüber [a,b]

d.damasceno@cs.ru.nl, d.struber@cs.ru.nl

Radboud University Nijmegen [a] and Chalmers | University of Gothenburg [b]
ACM/IEEE 24th International Conference on Model Driven Engineering Languages and Systems

SCAN ME

MODELS 2021 Fukuoka
博多

Tweet to @damascenodiego

Radboud University

Adapted from *J. G. Cham, Prof. Smith's rules for advising grad students (& postdocs), 1st edition. Los Angeles, CA: Piled Higher and Deeper Pub., LLC, 2018.*

Research Artifact: *any **digital object** that is either **created by the authors** of a study or **generated by experiments***
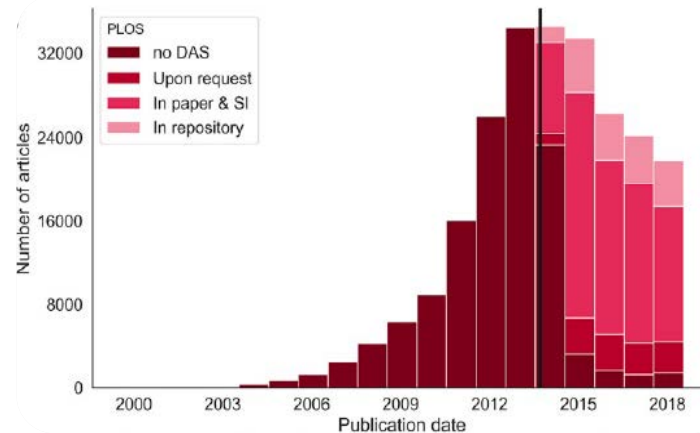
*Association for Computing Machinery. Artifact Review and Badging Version 1.1 - Aug. 24, 2020*
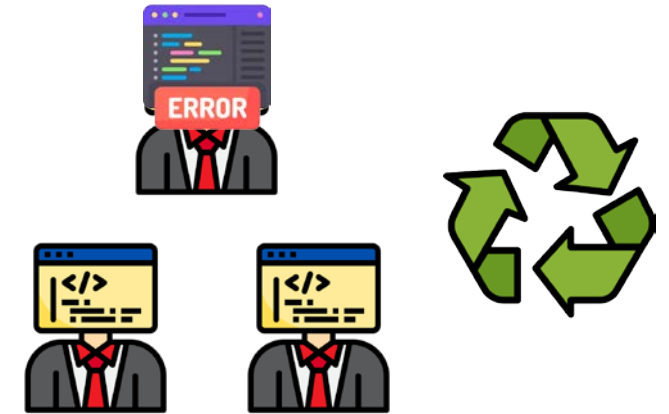
# Research Artifacts



**"validation of claims and result"**

*B. Hermann, et. al, 'Community expectations for research artifacts and evaluation processes', in ESEC/FSE 2020*
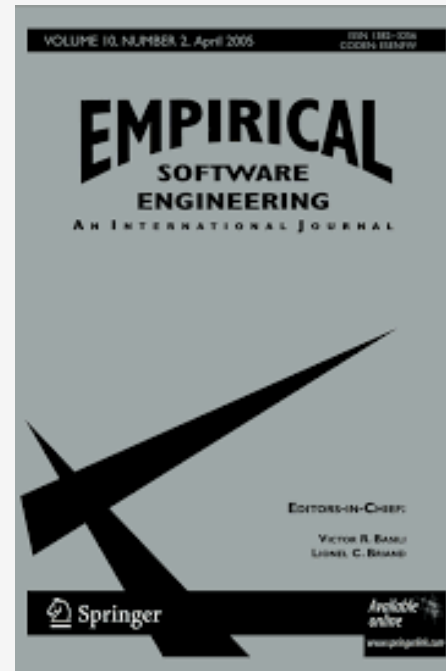


**"citation advantage"**

*G. Colavizza, et. al, 'The citation advantage of linking publications to research data', PLoS ONE, 2020*
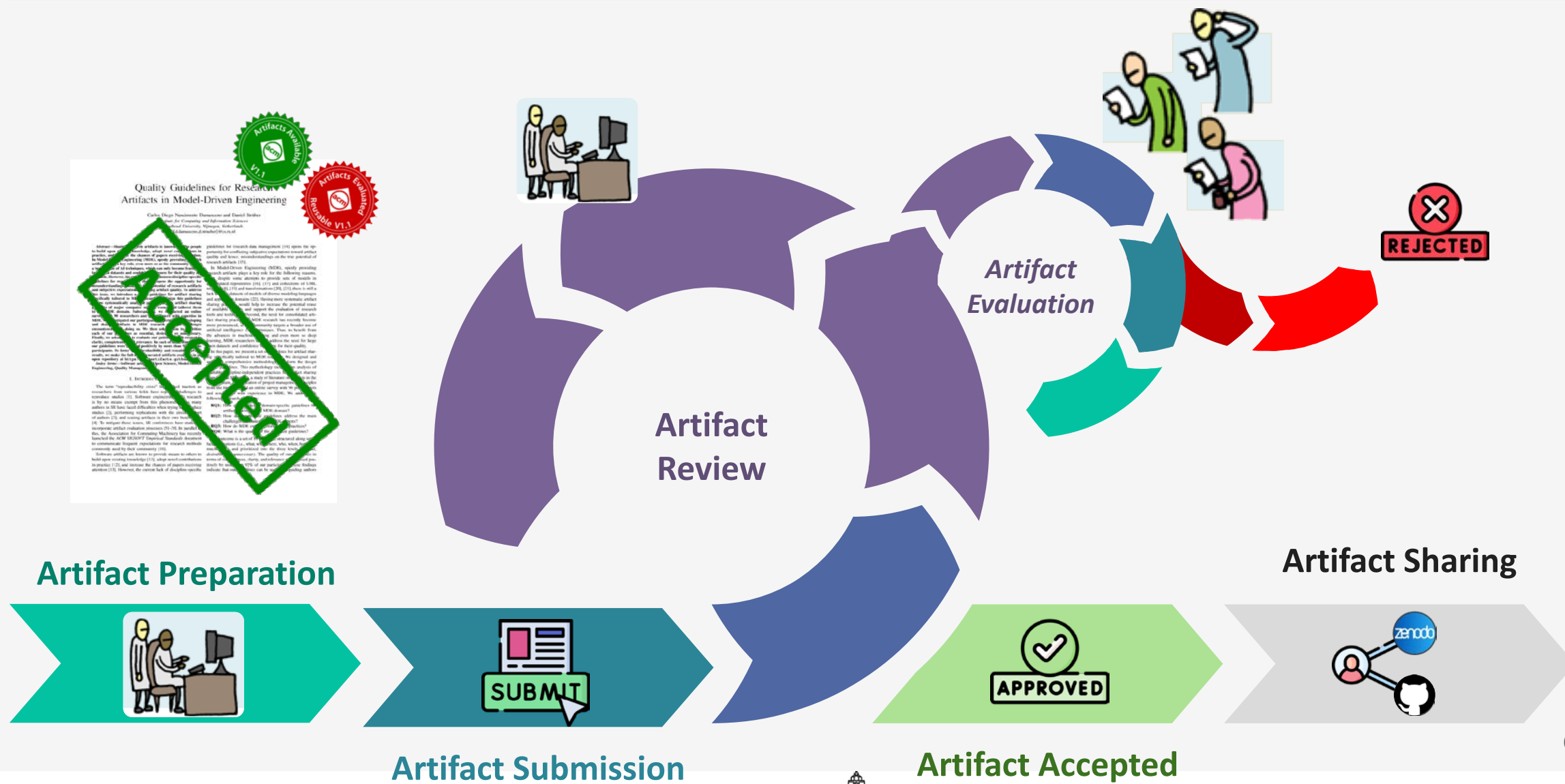


**"foster replicability and reusability"**

*R. Heumüller, et. al, 'Publish or perish, but do not forget your software artifacts', Empir Software Eng, 2020*

Radboud University

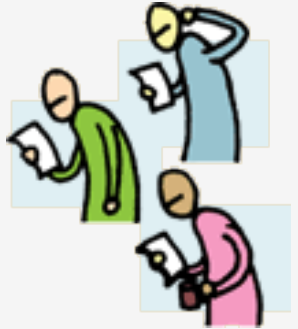# Artifact Evaluation (AE) in Software Engineering

Radboud University

# Artifact Evaluation (AE) in Software Engineering



**Artifact Preparation**

**Artifact Submission**

**Artifact Review**

**Artifact Evaluation**

**Artifact Accepted**

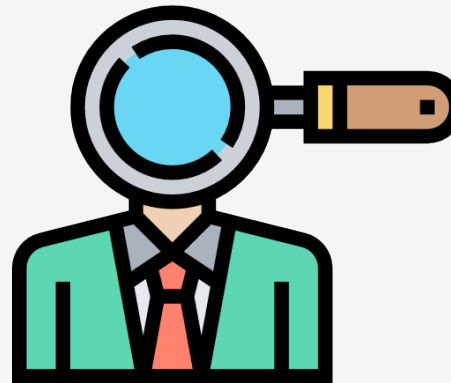**Artifact Sharing**

Radboud University

# Problem Statement



**Lack of consensus** *on well-defined* **quality thresholds**



**Subjective** *Quality* **Criteria**



**Limited experience** *with artifacts*

Hermann, B.; et al. **Community expectations for research artifacts and evaluation processes**. Proceedings of the 28th ACM ESEC/FSE. 2020. Available at: <https://doi.org/10.1145/3368089.3409767>.

Radboud University

# Research Objective

> *How can one define **domain-specific guidelines** for **artifact sharing** in **MDE research**? (RQ1)*
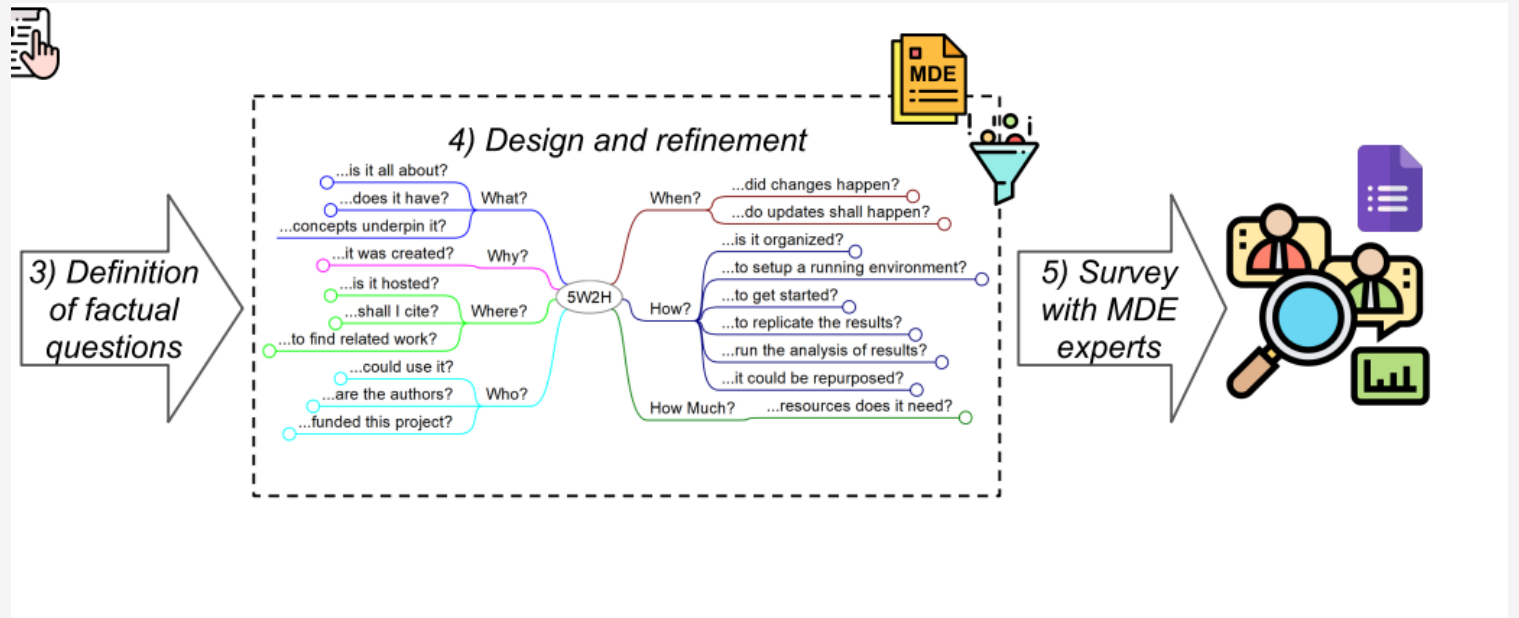>
> *Evaluate:*
> - ***Challenges** encountered by **MDE experts**? (RQ2)*
> - ***High-priority** practices? (RQ3)*
> - ***Quality** of the proposed MDE-specific guidelines? (RQ4)*

Radboud University

# Methodology

# Methodology

# 1) Identification of practices for artifact sharing

## Artifact Submission

An artifact submission consists of

- an abstract that summarizes the artifact and its relation to the paper,
- [for research and case-study papers] a .pdf file of the accepted paper (uploaded v comments into account (for tool papers, the submitted .pdf file will be used),
- a link to a .zip file (available for download) containing
  - a directory with the artifact itself,
  - a text file LICENSE that contains the license for the artifact (it is required th: mentioned above),
  - a text file README that contains detailed, step-by-step instructions on how
- SHA-256 hash of the zip file.

*TACAS. TACAS 2019 - ETAPS 2019. Available at <https://conf.researchr.org/track/etaps-2019/tacas-2019-papers#Artifact-Evaluation>.*

Radboud University

# 1) Identification of practices for artifact sharing

We analyzed eight guidelines sets for artifact sharing:

1) The ACM Artifact Review and Badging

2) The EMSE Open Science Initiative

3) The Journal of Open Science Software (JOSS)

4) The Journal of Open Research Software (JORS)

5) The Guidelines by Wilson et al. (2017)

6) The NASA Open Source Software Projects

7) The TACAS artifact evaluation guideline

8) The CAV artifact evaluation guideline

**284** general-purpose research practices

*R. Heumüller, et. al, 'Publish or perish, but do not forget your software artifacts', Empirical Software Engineering, 2020*

Radboud University

# 2) Categorization of practices and 3) Definition of factual questions

- Five Ws and two Hs (5W2H) as **content tags**

  - **Gaining insights** on **types of questions** could be addressed

  - Present **practices as answers** to factual questions

- Mind mapping for **data representation**

  - Provide **directions** to the definition of **factual questions**

  - **Kick off** the creation of **domain-specific guidelines**
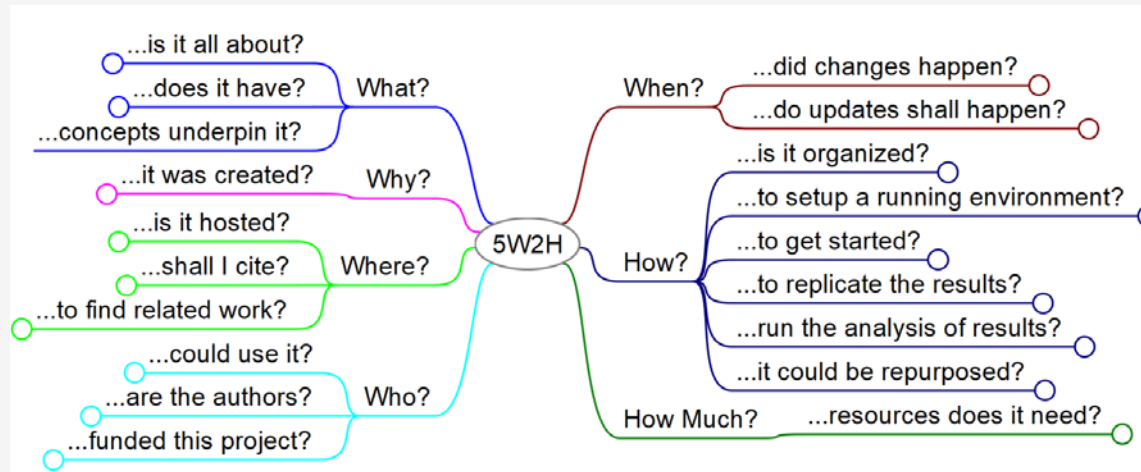


*2) Classification of best practices*

| Perspective | % |
|---|---|
| What | 31.44 |
| Where | 17.00 |
| Why | 3.97 |
| Who | 7.93 |
| When | 2.83 |
| How | 34.84 |
| How Much | 1.98 |

Radboud University

# 4) Design and refinement of the MDE-specific guidelines

SPECIAL SECTION PAPER

- ...........................................es)

## A taxonomy of tool-related issues affecting the adoption of model-driven engineering

Jon Whittle[1] · John Hutchinson[1] · Mark Rouncefield[1] · Håkan Burden[2] · Rogardt Heldal[2]

*4) Design and refinement*

...is it all about?
...does it have? — What?
...concepts underpin it?
...it was created? — Why?
...is it hosted?
...shall I cite? — Where?
...to find related work?
...could use it?

When? — ...did changes happen?
...do updates shall happen?
...is it organized?
...to setup a running environment?
How? — ...to get started?
...to replicate the results?
...run the analysis of results?

5W2H

- • E.g., Model semantics, syntax, and
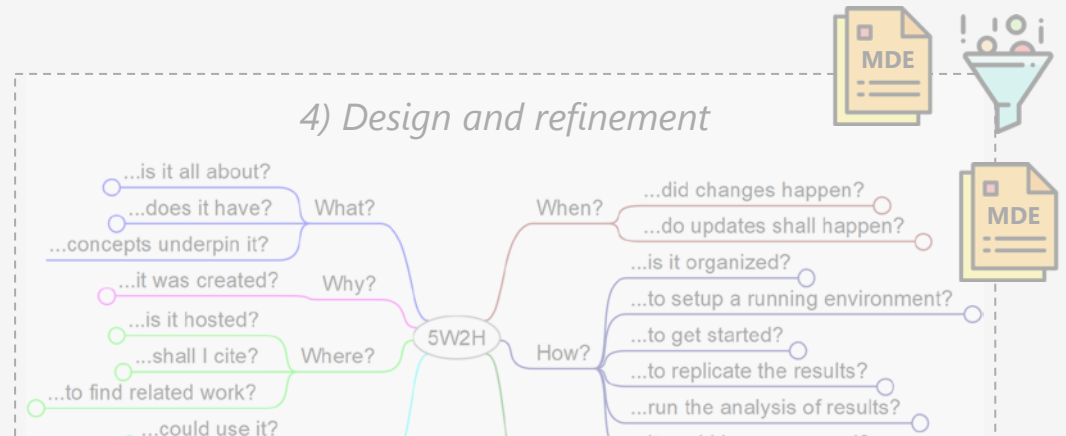
2017 IEEE/ACM Joint 5th International Workshop on Software Engineering for Systems-of-Systems and 11th Workshop on Distributed Software Development, Software Ecosystems and Systems-of-Systems (JSOS)

## Revisiting Criteria for Description of MDE Artifacts

Fábio Paulo Basso[1,2], Cláudia Maria Lima Werner[1], Toacy Cavalcante Oliveira[1]
{fabiopbasso, werner, toacy}@cos.ufrj.br
[1]Systems Engineering and Computer Science Department, Federal University of Rio de Janeiro, RJ, Brazil
[2]Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul, Bento Gonçalves, RS, Brazil

- Analyzed two studies on **MDE tools/**

- **Final guideline set:** 19 factual questi

Radboud University

# 5) Survey with MDE Experts

- Time window: **April-May 2021**

- Participants invited via e-mail:

  - Ex-members of MoDELS AECs

  - Coauthors of papers (in the last 3 years):

    - MoDELS and SoSyM linked to an **artifact**

  - PlanetMDE list <planetmde@univ-grenoble-alpes.fr>
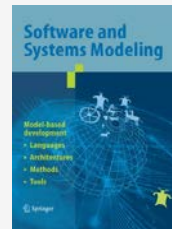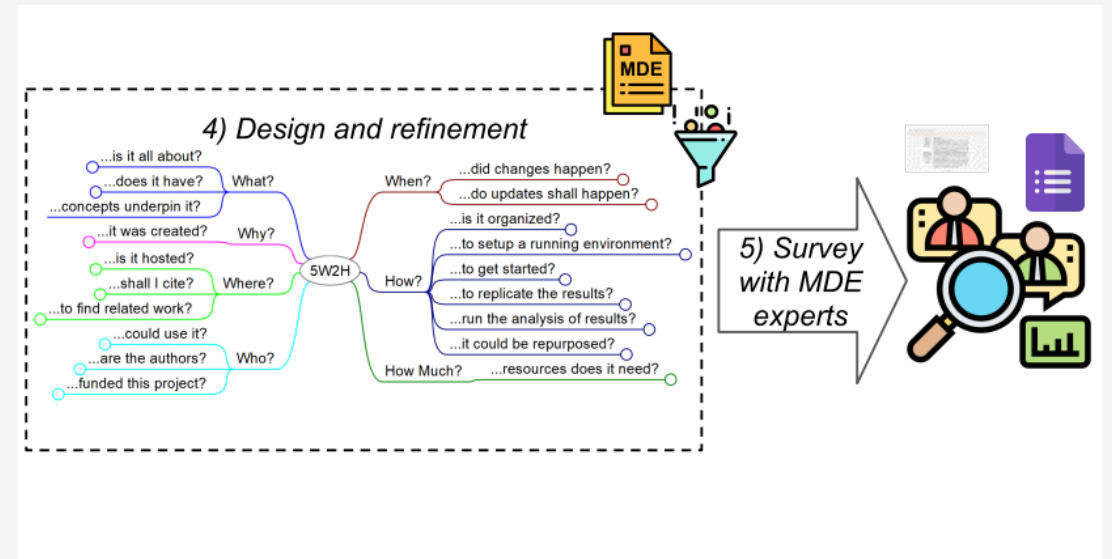
# Results

# Guidelines for MDE Research Artifacts Sharing (RQ1)

- Our guidelines included:

  - **84 best practices + 19 factual questions**

- **By products** of this study have been:

  - Versioned on **GitHub**

  - Archived on **Zenodo**

  - Shared on **arXiv**



https://mdeartifacts.github.io/

DOI 10.5281/zenodo.5109401

Radboud University

TABLE IV: Demographics - Primary role and gender

| Primary role | # Male | # Female |
|---|---|---|
| Industrial Practitioner | 7 | 0 |
| Industrial Researcher | 7 | 0 |
| Academic (Pre-Phd) | 10 | 4 |
| Academic (Post-Doc) | 18 | 5 |
| Academic (Professor) | 35 | 4 |

**90 participants**

TABLE V: Have you made contact/contacted someone for the purpose of artifact reuse?

| Made contact? | Been contacted? | # | % |
|---|---|---|---|
| No | Yes | 13 | 14.4 |
| Yes | No | 13 | 14.4 |
| No | No | 15 | 16.7 |
| Yes | Yes | 49 | 54.4 |

**Meaningful collective experience**

8. Which challenges have you encountered during the sharing and use of artifacts in MDE research projects?

Challenges are issues that make the sharing and use of artifacts difficult. If you have encountered multiple challenges, please start a new line for each.

**66 participants reported challenges**

Radboud University

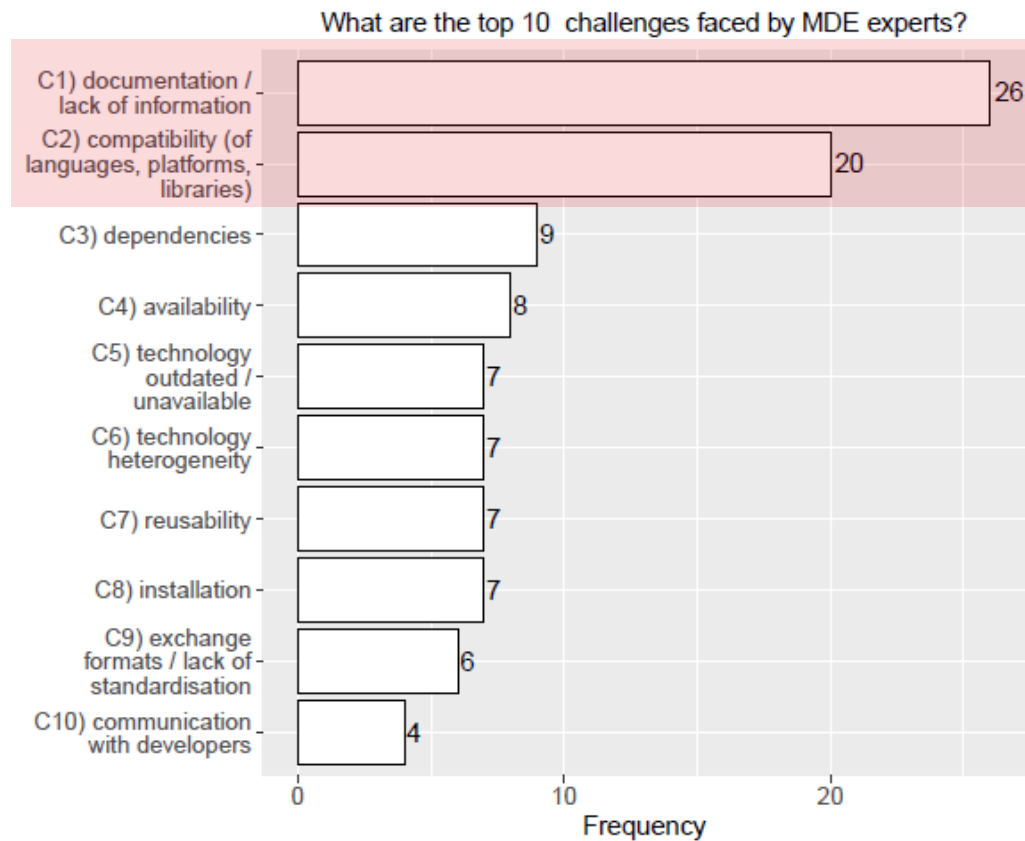# How do our guidelines address challenges encountered by MDE experts? (RQ2)

**What are the top 10 challenges faced by MDE experts?**

| Challenge | Frequency |
|---|---|
| C1) documentation / lack of information | 26 |
| C2) compatibility (of languages, platforms, libraries) | 20 |
| C3) dependencies | 9 |
| C4) availability | 8 |
| C5) technology outdated / unavailable | 7 |
| C6) technology heterogeneity | 7 |
| C7) reusability | 7 |
| C8) installation | 7 |
| C9) exchange formats / lack of standardisation | 6 |
| C10) communication with developers | 4 |

Fig. 4: Top 10 challenges faced in MDE artifact sharing

**TABLE VI: Traceability matrix for the 5W2H perspectives and Top 10 challenges encountered by MDE experts**

| 5W2H | Question | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| What | 1.1) What is it all about? | ✓ | | | | | | | | | |
| | 1.2) What does it have? | ✓ | | ✓ | ✓ | | | ✓ | | | |
| | 1.3) What underpins the artifact? | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | |
| Why | 2.1) Why it was created? | ✓ | | | | | | | | | |
| Where | 3.1) Where is it hosted? | ✓ | | | ✓ | | | | | | |
| | 3.2) Where shall I cite? | ✓ | | | | | | | | | |
| | 3.3) Where to find related work? | ✓ | | | | | | | | | |
| Who | 4.1) Who could use it? | ✓ | | | | | | ✓ | | ✓ | |
| | 4.2) Who are the authors? | ✓ | | | | | | | | | ✓ |
| | 4.3) Who funded this project? | ✓ | | | | | | | | | |
| When | 5.1) When did changes happen? | ✓ | | | ✓ | | | ✓ | | | |
| | 5.2) When do future changes shall happen? | ✓ | | | ✓ | | | ✓ | | | ✓ |
| How | 6.1) How is it organized? | ✓ | | ✓ | | | | ✓ | ✓ | | |
| | 6.2) How to setup a running environment? | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| | 6.3) How to get started? | ✓ | | | ✓ | | | ✓ | | | |
| | 6.4) How to replicate the experiment? | ✓ | | | ✓ | | | ✓ | | | |
| | 6.5) How to run the analysis of results? | ✓ | | | ✓ | | | ✓ | | | |
| | 6.6) How could it be repurposed? | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | |
| How many | 7.1) How many resources does it need? | ✓ | | ✓ | | | | ✓ | ✓ | | |

C1=Documentation  C2=Compatibility  C3=Dependencies  C4=Availability  C5=Tech. outdated/unavailable
C6=Tech. Heterogeneity  C7=Reusability  C8=Installation  C9=Exchange formats  C10=Communication

Radboud University

# How do MDE experts prioritize the practices? (RQ3)

## TABLE VII: Practices for MDE artifact sharing: 23 top-priority practices (out of 84 in total)

| 5W2H | Question | Practice | Priority | | |
|---|---|---|---|---|---|
| What | 1.1) What is it all about? | Indicate the context of its development (e.g., domain, problem, project) | 64.4% | 33.3% | 2.2% |
| | | Report its name | 71.9% | 24.7% | 3.4% |
| | | Indicate its main functionalities supported (e.g., modeling language, model analysis) | 81.1% | 18.9% | 0.0% |
| | 1.2) What does it have? | Include everything required for replications (i.e., complete) | 78.9% | 21.1% | 0.0% |
| | 1.3) What underpins the artifact? | Indicate modeling languages used to develop it (e.g., UML, SysML, BPMN) | 55.1% | 40.4% | 4.5% |
| | | Indicate libraries/frameworks used and their respective versions (e.g., Eclipse release) | 62.2% | 33.3% | 4.4% |
| Why | 2.1) Why it was created? | Indicate its objective/goal (e.g., replicability, reusability) | 51.1% | 44.4% | 4.4% |
| Where | 3.1) Where is it hosted? | Repository is open and public (e.g., GitHub, Zenodo, Figshare) | 77.8% | 17.8% | 4.4% |
| | 3.3) Where to find related work? | Give credit to data obtained from other sources (e.g., author, repository) | 58.9% | 41.1% | 0.0% |
| Who | 4.1) Who could use it? | Deposited under an explicit open license (e.g., reported in a LICENSE file) | 53.9% | 41.6% | 4.5% |
| | 4.2) Who are the authors? | Indicate the names of its authors | 73.3% | 24.4% | 2.2% |
| | 4.2) Who are the authors? | Indicate the authors's contact details (e.g., email, ResearchGate, website) | 56.7% | 36.7% | 6.7% |
| When | 5.1) When did changes happen? | Tracked using version control (e.g., GitHub, GitLab, BitBucket) | 61.1% | 36.7% | 2.2% |
| How | 6.1) How is it organized? | Files and folders shall have self-explaining names matching content | 51.7% | 47.2% | 1.1% |
| | 6.2) How to setup a running environment? | The artifact shall provide a step-by-step tutorial build the source code | 60.7% | 37.1% | 2.2% |
| | | The artifact shall provide instructions for downloading | 77.5% | 18.0% | 4.5% |
| | | The artifact shall provide instructions to install it | 86.7% | 12.2% | 1.1% |
| | 6.3) How to get started? | The artifact shall include instructions for running it on minimal test data | 62.9% | 33.7% | 3.4% |
| | | The artifact shall include step-by-step instructions for running it (e.g., README) | 86.7% | 13.3% | 0.0% |
| | 6.4) How to replicate the experiment? | Provide manual/automated instructions for the complete/partial replications | 57.3% | 41.6% | 1.1% |
| | | The artifact shall include the complete set of test models considered | 58.4% | 39.3% | 2.2% |
| | 6.5) How to run the analysis of results? | Provide a clear description of measurements and metrics used in the paper | 51.7% | 41.6% | 6.7% |
| How Many | 7.1) How many resources does it need? | Indicate the system/environment settings where it was successfully evaluated | 53.3% | 42.2% | 4.4% |

Priority legend: ☐ Essential ☐ Desirable ☐ Unnecessary

# What is the quality of the proposed guidelines? (RQ4)
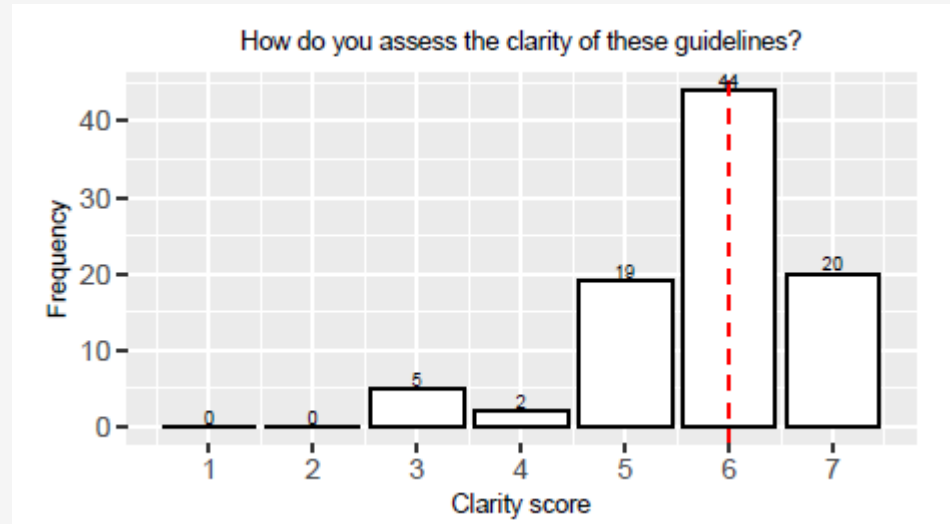
# What is the quality of the proposed guidelines? (**Improvements**)

"In terms of clarity, the questions use some abstract terms, in particular sharing. I was somehow confused by this term since sharing MDE artifacts may be associated with a research paper or not. Specially when the artifact is produced in an industrial context." [P58]

Radboud University

# Implications

# Implications for

**Artifact authors**

1. Toolkit for artifact creation, sharing and maintenance in MDE research

2. Drive authors to top-priority challenges in reusing MDE artifacts

**AEC members**

3. Complement other initiatives, e.g., ACM SIGSOFT Empirical Standards

4. Kick off the creation of venue-specific guidelines or FAQs

Radboud University

# Final Remarks

# Final Remarks

Radboud University
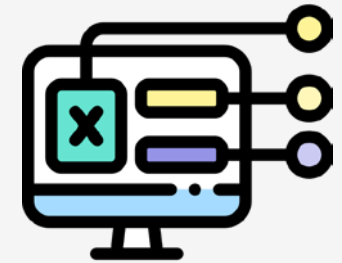
# Future Work



*Other artifact types and viewpoints*
*(e.g., non-coding artifacts, industrial research)*
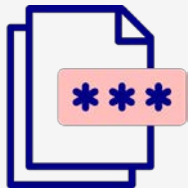


*Sub-group analysis*
*(e.g., gender, roles)*



*Other domains*
*(e.g., software product lines)*



*Retrospectives after artifact reviewing/usage*



As a \<user role\>

I want \<goal\>

so that \<benefit\>.

*User stories from*
*artifact stakeholders*

Radboud University

# Questions?

To access the guidelines…

SCAN ME

Carlos Diego N. Damasceno [a] and Daniel Strüber [a,b]

d.damasceno@cs.ru.nl, d.struber@cs.ru.nl

Radboud University [a] and Chalmers | University of Gothenburg [b]

ACM/IEEE 24th International Conference on Model Driven Engineering Languages and Systems

Tweet to @damascenodiego

Radboud University

# 5) Survey with MDE Experts

| Topic | Description |
|---|---|
| Demographics data | Questions about the participants (Q1) gender and their (Q2) current primary role |
| General experiences with artifacts | How would you rate your experience in (Q3) artifact development and sharing and (Q4) reusing artifacts in MDE research?; (Q5) Have you ever submitted an artifact for evaluation? Have you ever (Q6) contacted other researchers or (Q7) been contacted by other researchers asking for help on reusing their artifacts? |
| Challenges in artifact sharing | (Q8) Which challenges have you encountered during the sharing and use of artifacts in MDE research projects? |
| Evaluation of the Guidelines | We asked participants to rate the (Q9-34) relevance of each one of the 84 practices and, if needed, recommend additional guidelines. |
| Final evaluation | How do you assess the (Q35) clarity, (Q36) completeness, and (Q37) relevance of these guidelines? Open field for (Q38) additional remarks or (Q39) providing e-mail, if wanted to stay updated about our results. |

# 5) Survey with MDE Experts

**10.** **1.2) What does it have?** *

The artifact shall provide...

*Mark only one oval per row.*

|  | Essential | Desirable | Unnecessary | No answer |
|---|---|---|---|---|
| a description of its directory structure and content | ⬭ | ⬭ | ⬭ | ⬭ |
| everything required for replications (i.e., complete) | ⬭ | ⬭ | ⬭ | ⬭ |
| no more assets than necessary for replications (i.e., concise) | ⬭ | ⬭ | ⬭ | ⬭ |
| a preprint of its associated article | ⬭ | ⬭ | ⬭ | ⬭ |

Radboud University