Not All Those Who Share Are Lost: Analyzing 25 Years of Cybersecurity Artifact Sharing Practices Through Automated Discovery

Daan Vansteenhuyse[1], Arthur Bols[1], Lieven Desmet[1], Victor Le Pochat[2], Jo Van Bulck[1], Marton Bognar[1]

[1]DistriNet, KU Leuven
[2]Independent

Abstract

Cybersecurity publications are often accompanied by artifacts (code and data) that support the paper's findings. In recent years, all top-tier cybersecurity conferences have adopted policies regarding artifacts and introduced artifact evaluation. However, large-scale trends in artifact availability and the impact of conference policies remain largely unexplored, and meta-science studies (often focusing on reproducibility) typically cover a small set of papers due to the extensive manual effort required.

In this paper, we enable larger-scale analyses by introducing *ArtiFinder*, an automated tool that accurately identifies artifact URLs in paper PDFs. Using ArtiFinder, we present the largest quantitative analysis to date of artifact availability and related effects in close to 9,000 papers published in the four leading cybersecurity conferences since 2000. Our longitudinal analysis reveals a steady increase in artifact sharing over time, with substantial variation across cybersecurity subfields. We find that recent policy changes are not consistently accompanied by increases in artifact sharing, but they coincide with clear shifts toward stable hosting services. We also replicate prior analyses on our dataset, examining trends in citation counts and repository popularity. Our findings highlight notable trends in artifact sharing and the guiding role of conference organizers, while our open-source tool and dataset enable future large-scale studies.

1 Introduction

Academic papers, especially in computer science and cybersecurity, often produce code or datasets, referred to as *artifacts*. These artifacts can range from scripts conducting experiments and validating the most important conclusions to standalone tools intended for broad use. Artifacts play an essential role in enabling further research and the replication of studies. Despite the strong tradition of open-source contributions in computer science, such artifacts are not consistently shared alongside academic publications. This lack of general availability raises concerns about reproducibility and impedes scientific progress.

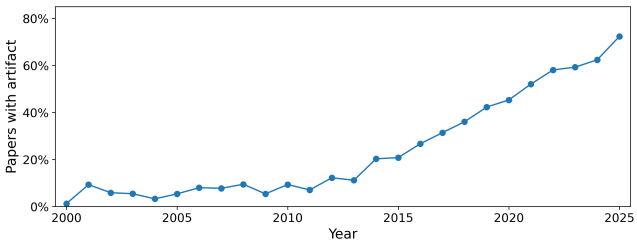

Figure 1: Artifact presence has been increasing since 2000, highlighting the importance of a longitudinal analysis.

Meta-science. Meta-science research [32, 36], the study of scientific research practices such as (the lack of) reproducibility [5] and open science, has seen increased interest across different fields. Already a decade ago, reproducibility issues in computer science research were highlighted by Collberg et al. [12], showing that only half of the artifacts associated with publications could be built. Since then, numerous meta-research studies have identified shortcomings in current research practices in cybersecurity [14, 25, 34, 35, 40, 47] and across computer science [13, 16, 26, 48, 49]. Unfortunately, reproducibility studies remain necessarily limited in scope due to the extensive manual effort required; previous work into artifact reproducibility reported spending four [39] or even eight person-years [40] of effort. Moreover, prior meta-science research on artifacts has primarily examined *reproducibility*, leaving longitudinal trends in artifact *availability* underexplored, covering at most 750 [40] or 2,000 papers [39].

Conference policies. To improve artifact availability and quality, conferences have started introducing explicit policies regarding artifacts. The most notable is the introduction of artifact evaluation (AE), where, similarly to the paper peer-review process, artifacts are evaluated by an artifact evalu-

ation committee (AEC), without it affecting the paper's acceptance. In computer science, AE was introduced in 2011 at ESEC/FSE [20]. Over the years, more venues instated AE, with WiSec [50] and ACSAC [2] introducing this practice in the cybersecurity community in 2017. Among the top-tier (CORE A*-ranked [29]) cybersecurity venues examined in this work, USENIX Security was the first to adopt AE, starting in 2020. With IEEE S&P running AE for its 2026 edition, all A* conferences now evaluate artifacts, albeit with varying policies. While AE imposes substantial demands on both authors and AEC members, its impact and benefits to the community remain understudied.

Existing datasets. In response to the growing number of papers with artifacts, community-driven infrastructure for collecting and even running artifacts has emerged. Services such as secartifacts [45], SEARCCH [6], or the now-discontinued paperswithcode [15] serve as a curated list of artifacts for various conferences. Importantly, these platforms are limited to papers that participated in AE (uploaded by the AEC chairs) or chose to upload their artifacts manually.

Current datasets, hence, pose several limitations for identifying papers with artifacts and analyzing broader trends in artifact sharing. First, not all papers that provide artifacts participate in AE, making such artifacts more difficult to discover. Second, existing aggregation websites remain incomplete. For example, despite ACM CCS having implemented AE since 2023, no publicly available list of its AE outcomes is available. Finally, since AE is relatively recent and current datasets focus primarily on papers that have undergone formal evaluation, they often lack older papers with artifacts. These limitations underscore that artifact sharing, despite being a longstanding practice, is difficult to study comprehensively.

This work. In this paper, we build ArtiFinder, a scalable artifact discovery tool, enabling large-scale analysis of artifact sharing. ArtiFinder can automatically discover artifact URLs from published PDF files by analyzing and scoring extracted hyperlinks based on a set of predefined heuristics. When comparing with various existing datasets, ArtiFinder identifies artifacts with an accuracy of 89.6–98.8%. Notably, ArtiFinder functions without the use of LLMs to avoid hallucinations and increase reliability and reproducibility, with the added benefit of limiting the costs, both economic and ecological.

Using ArtiFinder, we investigate 25 years of artifact sharing in top-tier cybersecurity conferences, focusing on large-scale artifact availability trends across close to 9,000 papers. First, we construct a dataset of all papers with artifacts from A* cybersecurity conferences from the past 25 years (IEEE S&P, ACM CCS, USENIX Security and NDSS). We find that, encouragingly, the share of papers with an artifact has grown from 1.2% in 2000 to 72.3% in 2025, as shown in Figure 1. Second, to understand the impact of conference policies and AE on sharing practices, we summarize relevant requirements from the calls for papers (CfPs) and calls for artifacts (CfAs) for cybersecurity conferences and overlay big changes with our dataset. Here, we find that the requirements are currently very heterogeneous across conferences and editions, while lower-ranked conferences barely implement policies at all. Finally, we characterize our dataset and replicate prior studies on our data, investigating the correlation of artifact availability and paper citations [8, 10, 27, 48, 49], AE processes and the number of papers with artifacts [40], the share of hosting providers [16], and participating in AE and repository popularity [16], with most findings showing no statistically significant impact of AE participation.

Terminology and scope. In this work, we define an artifact as a standalone digital object, consistent with the definition used by ACM [4] and most CfAs. Many aspects related to sharing and availability discussed in this paper are typically not relevant to self-contained papers that embed code or data, e.g., in the appendix. Our study does not investigate reproducibility; we focus on artifact *presence* and *availability*. Artifact presence determines whether the published paper contains a link to the artifact, while availability determines whether the linked resource is accessible at the time of analysis.

Contributions. To summarize, our main contributions are:

- We design and implement ArtiFinder, a fully automated tool capable of identifying linked artifacts in research papers with high accuracy.
- Using ArtiFinder, we construct a dataset of 3,922 research artifacts from A*-ranked security conferences since 2000, showing an increase of artifact presence from 1.2% to 72.3%.
- We summarize AE policies across conference editions and study the correlation between artifact sharing, AE participation, and other metrics, replicating prior studies.
- We provide recommendations to authors and conference organizers to increase the visibility of artifacts.

Open science. We open-source ArtiFinder at `https://github.com/DistriNet/ArtiFinder` and our dataset at `https://github.com/DistriNet/ArtiFinder-Data`. For full details on our open science contributions, please refer to the Open Science Appendix.

2 Artifact evaluation policies

In this section, we first give a brief background on AE, then perform a systematic overview of the policies used at A*-ranked cybersecurity conferences.

2.1 Artifact evaluation

To promote reproducible and reusable research, conferences introduced *artifact evaluation (AE)*, a process similar to the peer review of papers [16, 26]. AE is performed by an *artifact evaluation committee (AEC)*, which is separate from the program committee (PC) responsible for the peer review of the papers. AE is typically only conducted on accepted papers, with the outcome of the evaluation having no impact on the acceptance status of the paper.

The outcome of the evaluation is represented by means of *badges* that authors apply for, awarded by the AEC based on the properties and quality of the artifact. Although the criteria and definitions for these badges can vary between conferences [34, 39, 41], conferences typically award *Available*, *Functional*, and *Reproducible*, with ACM conferences also using *Reusable* and *Replicated* badges [4]. Generally, the Available badge is awarded if the artifact is stored on a stable hosting service, often required to be backed by a DOI. Artifacts with the Functional badge are judged to be well-documented, containing code that can be executed, with the Reusable badge awarded in case of exceptional quality. The Reproducible badge is awarded when the main claims of the paper can be independently verified using the provided artifact, whereas the Replicated badge represents validating the results without using the original artifact.

2.2 Policies at A* conferences

To understand what impact conference policies have on artifact sharing practices, we first conducted a systematic overview of these policies at A*-ranked security conferences[1]. We manually examined CfPs and CfAs (when present) for the past 25 years at USENIX Security (SEC) [46], ACM CCS [1], NDSS [31], and IEEE S&P [30]. Policies for 2026 editions are also included, as the websites were already available at the time of writing. Our methodology is based upon publicly available websites, it is, therefore, possible that we missed details or changes in policies that were only communicated to the authors directly. Table 1 depicts an overview of the most important policies, which are discussed in more detail next. We also contributed this data to CyCoAnalysis [7], a dataset collecting conference policies.

Submission. Since paper reviews and AE are usually independent processes, most conferences only conduct AE for accepted papers. Increasingly, however, conferences encourage or require submitting artifacts together with the initial paper submission. At CCS since 2022, authors are expected to provide artifacts to the reviewers at submission time if the paper's contributions rely on them or argue why this is not feasible. For its 2026 edition, USENIX Security and CCS require all submissions to make their artifacts available at submission time and explain their utility in a dedicated "Open Science Appendix".

[1] We consider conferences whose main focus is cryptography out of scope.

Table 1: Overview of artifact submission and evaluation policies at A*-ranked cybersecurity conferences.

Conference	Submission	Evaluation	Hosting	Appendix	Packaging	Infrastructure	Secartifacts	Awards
SEC'20	○	◐	○	○	◐	○	○	○
SEC'21	○	◐	○	○	◐	○	○	○
SEC'22	○	◐	◐	●	◐	○	○	●
SEC'23	○	◐	◐	●	◐	○	●	●
SEC'24	○	◐	◐	●	◐	○	●	●
SEC'25	○	●	●	●	◐	○	●	●
SEC'26	●	●	●	●	◐	○	●	●
CCS'22	◐	○	○	○	○	○	○	○
CCS'23	◐	◐	○	●	○	○	○	○
CCS'24	◐	◐	●	◐	●	○	○	○
CCS'25	◐	◐	●	◐	●	○	○	○
CCS'26	●	◐	●	●	●	○	○	○
NDSS'24	○	◐	●	◐	●	○	◐	●
NDSS'25	○	◐	●	●	●	○	◐	●
NDSS'26	○	◐	●	●	●	○	●	●
S&P'26	○	◐	●	○	●	●	○	●

The table shows whether artifacts should be provided at **submission** time; there is an artifact **evaluation** process for accepted papers; there are guidelines on the **hosting** provider of artifacts; evaluated papers should include an artifact **appendix**; there are guidelines on the **packaging** (format) of artifacts; the artifact needs to run on a public research **infrastructure**; the artifact evaluation process is described on **secartifacts** instead of the conference website; the conference **awards** outstanding artifacts. Half bullets (◐) indicate optional / partial policies.

Evaluation and hosting. In 2025, USENIX Security became the first (and so far only) conference to make participation in AE mandatory for the Available badge. Moreover, they introduced strict hosting requirements for this badge, mandating the use of a stable service with permanent links (see Section 4.3). It is specifically noted that GitHub, GitLab, or personal websites are no longer sufficient to be granted the badge, whereas a specific GitHub tag or commit was accepted at previous editions. CCS and NDSS were earlier to adopt hosting requirements, making a DOI link to the artifact mandatory already the year prior.

Appendix. Due to page limits, authors often do not have sufficient space to include detailed information about their artifact. Hence, artifact appendices were introduced and made mandatory for papers participating in AE starting with

USENIX Security in 2020, later adopted by NDSS. At CCS, creating such an appendix is optional. The artifact appendix is meant to contain all necessary instructions for running the artifact and reproducing the major research claims. This not only helps the AEC during the evaluation, but also other researchers who wish to independently verify results or reuse certain parts of the artifact in their own research.

Packaging. Although all conferences have guidance on the structuring and packing of artifacts, requirements vary greatly. S&P'26 has the strictest packaging requirements, with 10 required files or folders for code artifacts. This makes the evaluation process easier for AEC, but requires more effort from the authors. USENIX Security only enforces packaging requirements if the artifact is evaluated for the functional or reproducible badge. Throughout all AE editions, NDSS has always used the same packaging instructions, with the only requirement being the presence of a README with clear instructions on how to use the artifact. CCS has a similar policy, with the additional requirement of a LICENSE file.

Infrastructure. To facilitate the work for the AEC, who might not always have the same infrastructure as the authors, S&P'26 for the first time mandates that artifacts run on public infrastructure such as SPHERE [38], Chameleon [33], or CloudLab [11]. Exceptions can be made for artifacts requiring special infrastructure or a GUI.

Secartifacts. The secartifacts website [45] is a community-led project collecting AE results from various conferences. USENIX Security and NDSS started using secartifacts to host instructions and requirements for AE, making it a central hub for everything related to cybersecurity artifacts. In practice, however, the instructions for authors and the outcomes of the evaluations are often scattered on both the conference website and secartifacts.

Awards. To encourage authors to submit artifacts, conferences introduced *distinguished artifact awards* for exceptional artifacts. USENIX Security started this among A* conferences, with NDSS and S&P adopting it when starting AE.

Takeaway. As an overarching trend, policies change every year as new AEC chairs experiment with different approaches, similar to PC chairs experimenting with paper submission policies [7]. Given that AE is relatively recent, policies may converge in the future. With time, conferences set increasing requirements on stable hosting, packaging, and infrastructure for artifacts to support the long-term viability of both the AE process and artifact usage. Moreover, secartifacts has emerged as a central resource for AE.

Table 2: Overview of artifact submission and evaluation policies at A-ranked cybersecurity conferences using the same criteria as Table 1.

Conference	Submission	Evaluation	Hosting	Appendix	Packaging	Infrastructure	Secartifacts	Awards
ACSAC'17	○	◐	○	○	○	○	○	●
ACSAC'18	○	◐	○	○	○	○	○	●
ACSAC'19	○	◐	○	○	◐	○	○	●
ACSAC'20	○	◐	○	○	◐	○	○	●
ACSAC'21	◐	◐	○	○	◐	○	○	●
ACSAC'22	◐	◐	○	○	◐	○	○	●
ACSAC'23	◐	◐	○	○	◐	○	○	●
ACSAC'24	◐	◐	○	○	◐	○	○	●
ACSAC'25	◐	◐	○	○	●	●	○	●
ACSAC'26	◐	◐	○	○	●	●	○	●
CSF'20	◐	○	○	○	○	○	○	○
CSF'21	◐	○	○	○	○	○	○	○
CSF'22	◐	○	○	○	○	○	○	○
CSF'23	◐	○	○	○	○	○	○	○
CSF'24	◐	○	○	○	○	○	○	○
CSF'25	◐	○	○	○	○	○	○	○
CSF'26	◐	○	○	○	○	○	○	○
ASIACCS'26	◐	○	○	○	○	○	○	○

2.3 Comparison with A-ranked venues

While our dataset and analysis focuses on artifacts and policies at A*-ranked conferences, A-ranked conferences might also have an impact on the evolution of these policies, as evidenced by ACSAC being one of the pioneers for AE in cybersecurity. For this reason, we also analyzed the policies at A-ranked cybersecurity conferences to understand if they had an impact on the AE processes at A* venues. The results are summarized in Table 2, with the most important aspects highlighted below.

Overall, we found that only a few A-ranked conferences have strong policies on artifacts or an AE in place. ESORICS and RAID do not have any policies on artifacts, while CSF mentions since 2020 that "supplementary material such as proof scripts can be uploaded" along with the paper submission. Since 2022, EuroS&P has encouraged (made mandatory for a single edition in 2025) submitted papers to include a statement on open science and how artifacts will be shared, with the threat of retracting papers that fail to satisfy these promises. In its most recent CfP, ASIACCS follows the example of CCS in requiring authors whose contributions rely on artifacts to submit these at the time of paper submission. ACSAC stands out for having the longest-running AE process among all studied conferences, being the first to introduce awards and strong requirements on packaging and running

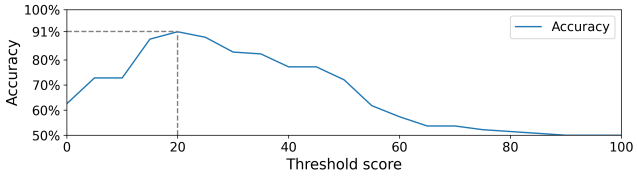

Figure 2: Since some papers do not have an artifact, the highest scoring link is only marked as an artifact if its score is above a threshold. This score threshold affects the accuracy. The final threshold is the score yielding the highest accuracy for our sample.

on public infrastructure. For most other policies, however, it does not enforce the same requirements as A* conferences. Somewhat uniquely, ACSAC explicitly involved artifact submission in the paper evaluation process in 2021, stating that "good artifacts will contribute positively to the paper evaluation", and even ran an early AE round in 2022, the results of which "may be factored into the decision for those papers that advance to round 2 of the reviewing process". However, this approach has not continued in later editions or been adopted by other venues.

3 The ArtiFinder tool

To address the limitations of existing artifact datasets identified earlier, we developed an automated tool, ArtiFinder, to automatically discover links to artifacts in a paper.

3.1 ArtiFinder design

ArtiFinder is based on the intuition that authors provide a link to their hosted artifact in the paper in a way that makes it clear that the link points to the artifact. Without a clear link, the artifact would be difficult to discover for the readers, reducing its value for reproducibility and future work. ArtiFinder is therefore composed of two main components: a link *scraper* and a *ranker*. When given a PDF file as input, the link scraper component extracts all hyperlinks from the PDF, leveraging Poppler [22]. This step also considers links in the text that are not clickable. Once all hyperlinks are identified, the ranker component scores and ranks the hyperlinks in order to find the most likely (link pointing to the) artifact. The scores are determined by heuristics, which are discussed in more detail in Section 3.2 and Appendix B. To determine the threshold score for a link to be considered the artifact, we analyzed the accuracy for 100 randomly chosen papers and found 20 to be the optimal score (cf. Figure 2). If multiple different links reach the threshold score, all are marked as artifact links.

Table 3: Outcomes of ArtiFinder for identifying artifacts in different sets of papers.

	secartifacts (n=879)	**Olszewski** (n=742)	**Manual** (n=84)
Correct presence	87.0%	39.1%	6.0%
Exact link	58.6%	31.8%	6.0%
Alternative	28.4%	7.3%	0.0%
Correct absence	8.5%	54.2%	92.9%
No link	0.0%	45.8%	92.9%
Missing link	8.5%	8.4%	0.0%
Wrong presence	2.2%	4.4%	1.2%
No link	0.0%	4.4%	1.2%
Incorrect link	2.2%	0.0%	0.0%
Wrong absence	2.3%	2.3%	0.0%
Final accuracy	95.6%	93.3%	98.8%

3.2 Heuristics

ArtiFinder defines heuristics that are used for scoring the likelihood of links pointing to the artifact. These heuristics were refined during the development process and are often based on common intuitions about artifacts, such as the artifact page mentioning the paper title or authors. A full overview of the heuristics and associated scores can be found in Appendix B.

To speed up the ranking process, the heuristics are split into four phases that can be executed concurrently:

RawPhase To account for parsing errors, we check if a candidate link is a valid URL. If the hyperlink cannot be correctly parsed (e.g., due to illegal characters or an invalid TLD), it is excluded from further analysis.

UrlPhase By matching the URL host with a predefined list, an extra score is awarded to URLs leading to well-known hosting services (e.g., `github.com`, `zenodo.org`, and `doi.org`). Additionally, a DNS request is sent to check the availability of the domain.

LocationPhase Based on the location of a link in the paper (e.g., in a footnote or the introduction), an extra score is awarded. Keywords such as "source" or "artifact" in the paragraph containing the link also increase the score.

RequestPhase The final phase sends a request to each URL and checks if the authors or the title of the paper are present on the page. Optionally, for unreachable links, an archived version can be used (cf. Appendix C).

3.3 Accuracy

To evaluate the accuracy of ArtiFinder, we compare the artifacts found by ArtiFinder with a baseline of the two most

prominent existing artifact datasets in cybersecurity, as well as a manual sample. First, we compare the classification results on the subset of papers published at A* conferences that underwent AE and are listed on secartifacts [45]. Second, we use the manually identified artifacts from ML papers published at A* conferences from Olszewski et al. [40]. Third, since both datasets contain relatively recent papers (2013–2025), we randomly sampled 5% of the A* papers from each prior year (2000–2012), mostly consisting of papers with no associated artifact.

The results of the evaluation are shown in Table 3. For each paper, there are several possible outcomes depending on what ArtiFinder found and what is reported in the baseline dataset:

Correct presence ArtiFinder successfully identified a correct artifact link that was present in the paper. This link could be exactly the same as the one in the ground truth dataset, but in many instances, we found that the paper included an alternative link. For example, the secartifacts dataset might contain a DOI-backed link which was submitted for AE as required by the conference, while the paper links to a project website or a GitHub repository.

Correct absence ArtiFinder correctly concluded that the paper does not contain a link to an artifact. The *No link* outcome means that the baseline dataset also reports no artifact, while the *Missing link* outcome means that while a link was present in the ground truth, the paper does not contain this link.

Wrong presence is an edge case where ArtiFinder, incorrectly, identifies an artifact. Either when the ground truth does not have an artifact (*No link*), or when an *Incorrect link* was classified as artifact.

Wrong absence This outcome is often the result of the linked artifact containing little information that connects it to the paper (e.g., paper title, authors, venue) in such a way that ArtiFinder identifies no artifact.

Overall, we find a high accuracy for all datasets, ranging from 93.3– 98.8%. We also found a high share of papers that either report an alternative artifact URL compared to the one submitted to AE or omit the URL completely. This is further discussed in Section 7.1.

3.4 Ablation study

In the following, we describe a short ablation study on heuristics. The study quantifies the impact of two heuristics (`LocationInPaper` and `GitHubRepo`) on the classification performed on the subset of our dataset that has ground-truth data on secartifacts. The baseline accuracy on this dataset is 95.6% (cf. Section 3.3).

LocationInPaper. Since we expect artifact links to be mentioned in the body of the paper, this heuristic subtracts points for links in the references or appendix section, while giving bonus points for a link included as a footnote or in open science paragraphs. 81.6% of identified artifacts were granted points due to this rule. Removing this rule yields an accuracy of 88.8%, showing that it is not a dominant heuristic in the ranking process.

GitHubRepo. As a popular developer platform, GitHub links in a paper often lead to an artifact: in our sample, 51.7% of artifacts are a GitHub repository (cf. Section 4.3). Removing this rule yields an accuracy of 83.6%, indicating that GitHub-hosted artifacts are also detected by other heuristics.

3.5 Resource usage

While parsing the papers is a one-time effort, making it less resource-critical, the runtime of ArtiFinder needs to be reasonable to support the extraction of artifacts from large sets of papers. We conduct timing experiments on a random sample of 100 papers from our dataset, where 4 papers are parsed and ranked in parallel. The runtime depends on three factors: the length of the paper, the number of found links, and the reachability of these links. For our random sample, each PDF page takes on average 0.14 seconds to parse, resulting in an average total parsing time of 2.4 seconds per paper. Ranking the links is more expensive as an HTTP request is sent for each parsable link with an unpredictable response time. We limit execution time per link by employing a timeout of 3 seconds for each request, giving a final time of 10.4 seconds per paper, for an average of 46 links per paper. Appendix C expands on the performance impact of including an Internet Archive lookup for unreachable links.

4 Trends in cybersecurity literature

Thanks to the capabilities of ArtiFinder, it becomes possible to collect artifacts on a large-scale paper corpus with minimal manual effort. In this section, we perform the largest study to date on artifacts in cybersecurity, analyzing all publications at A* cybersecurity conferences from 2000 until 2025, totaling 9,054 papers. We first leverage DBLP [44] to retrieve metadata for all papers published at these conference editions, supplemented with Scopus [19] to extract citation counts and affiliations for all papers, and secartifacts [45] to retrieve badges of papers that participated in AE (except for the missing CCS editions)[2]. Afterwards, we feed the papers in PDF format to ArtiFinder, which identifies the artifacts. Due to parsing errors or missing papers, ArtiFinder was unable to process 102 papers, giving us a final corpus of 8,952 papers. From this corpus, ArtiFinder identified 3,922 linked artifacts.

[2]All data was collected in January 2026.

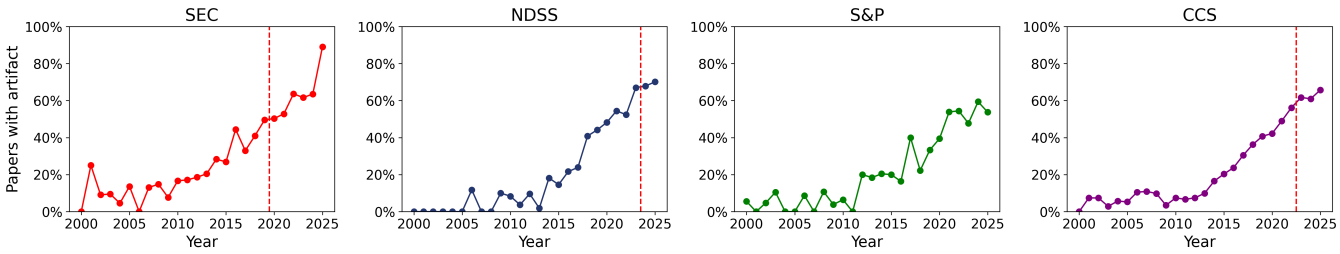

Figure 3: For each conference, the number of papers with an identified artifact present increases over time. The introduction of AE is indicated with a dotted vertical line for each conference.

Using this collected dataset, we re-run experiments from prior studies and answer several research questions. We first analyze *how often* artifacts are present, and how this relates to changing AE policies. We then look at *where* artifacts are published, and how this affects their long-term availability. Afterwards, we move to *who* publishes artifacts across subfields and institutions within cybersecurity research. Finally, we study *why* authors may want to make artifacts available, as artifact sharing might cause increased paper impact. Given the scope of our work and the automated nature of its artifact discovery and analysis, we leave the evaluation of *which* artifacts are shared and *how well* they are designed to future work on reproducibility.

4.1 Artifact presence

We start by analyzing summary trends on artifact presence over time. Although artifact evaluation for A* conferences only started with USENIX Security in 2020, Figure 3 shows that throughout the past 25 years, researchers have been releasing artifacts alongside their papers at all conferences. Clearly, recent years have seen a positive evolution of releasing research artifacts, also before the introduction of formal AE processes. One change clearly attributable to a single event is the surge in the share of papers with an artifact to over 80% at USENIX Security in 2025, which corresponds to the enforcement of making artifacts available. Between conferences, USENIX Security and NDSS have higher artifact rates than S&P and CCS, perhaps owing in part to the former two conferences embracing open access publishing early on, attracting researchers who also appreciate open science.

In prior work, Olszewski et al. [40] performed permutation testing to determine a correlation between the introduction of AE and artifact availability but found no statistically significant results. In contrast, when performing the same test on our dataset, we find an observed difference of 37.8% between artifact availability before and after AE with $p < 0.01$, showing a significant increase. However, this difference in results might be explained by our longer pre-AE window; future work could employ more advanced statistical techniques to study the precise impact of AE.

4.2 Participation in artifact evaluation

Although the overall share of artifacts has been steadily rising, the number of awarded badges does not completely follow this trend, as shown in Figure 4. This data is based on AE results from the secartifacts website, which at the time of writing only has results for NDSS and USENIX Security. While S&P understandably has no results due to the lack of AE so far, results for CCS are prominently missing, despite hosting AE since 2023. The two available NDSS cycles show relatively similar acceptance rates, while USENIX Security has seen a small decrease before a surge for the 2025 edition, due to its changed policy of mandatory AE for the Available badge. That policy change, however, only had a small impact on the number of awarded Functional and Reproduced badges. This shows that authors were only minimally influenced to also pursue the other badges.

4.3 Hosting platform

Prior studies of artifacts analyzed the share of hosting providers, finding the majority of cybersecurity artifacts on GitHub [40], and a balanced share between GitHub and Zenodo at EuroSys [16]. As shown in Section 2.2, the A* conferences are increasingly limiting which hosting providers are accepted for artifact evaluation. Dedicated *stable hosting services* cater specifically to academic research. They offer to host any code or data with the guarantees that the artifact is fully stable, i.e., that a given version cannot be changed (while allowing new versions), remains available in the long term, and can only be deleted after a manually evaluated request. These services also typically assign a DOI to the artifact itself. Examples of such services include Zenodo [21], FigShare [17] or Dryad [18]. Alternatively, authors may host artifacts on *code repository services* such as GitHub or GitLab. These repositories can be updated over time and have more advanced code browsing, discovery, and collaboration features. With Git, specific commits may be tagged to provide a stable snapshot. Finally, authors may host on *institutional or personal websites*, sometimes created specifically for the paper or artifact. This gives maximum flexibility in formatting and updates, and may therefore be more attractive for

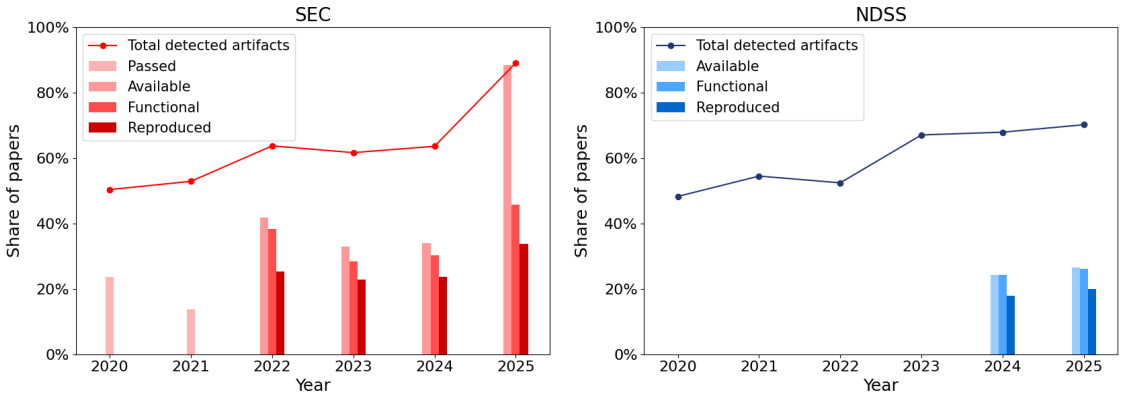

Figure 4: Share of papers awarded badges by AE for USENIX Security and NDSS per year.

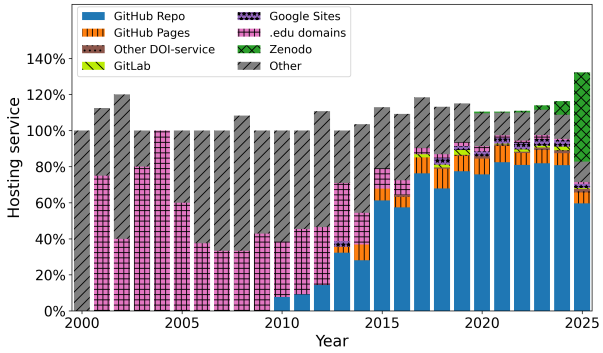

Figure 5: The share of services used for hosting artifacts in each year for all conferences. Note that the percentage in the graph can rise above 100%, as some papers link to multiple artifacts, detected by ArtiFinder when multiple links meet the score threshold (cf. Section 3.3).

continuously updated datasets or with authors who seek to publicize their work. However, these provide no guarantees for stability or long-term availability.

Figure 5 shows how the preferred hosting provider among cybersecurity authors has evolved over time. While artifacts were initially hosted on personal or university websites (e.g., using the `.edu` TLD), GitHub started gaining traction in 2010 (two years after its creation). Zenodo first appeared in 2020, coinciding with the introduction of AE, despite it being created already 7 years earlier. Notably, all but three artifacts stored with a DOI are from conference editions after 2020. Zenodo's adoption only became substantial after conferences began requiring stable hosting. In 2025, 49% of papers with an artifact included a Zenodo link, with GitHub still being the dominant service with 60% (totaling over 100% due to papers with multiple artifact links).

These trends are further explored in Figure 6 for individual conferences in the past 10 years. For the three conferences with an already established AE, the widespread adoption of stable hosting services only began after AE was introduced, and they are barely used at S&P. This figure also confirms that the surge of Zenodo in 2025 is mainly due to USENIX Security, where it is the most common hosting service, and where Zenodo is specifically noted as an example in the CfA. Notably, while both NDSS and CCS have required stable hosting services for artifact evaluation since 2024 (cf. Section 2.2), the share of Zenodo and other stable services is much lower than for USENIX Security in 2025. Conferences without formal AE tend to follow general trends, albeit at a slower pace. Although CCS has had an AE since 2023, the hosting services follow a similar pattern as S&P, which only starts evaluating artifacts for its 2026 edition. They are both following the same trend previously seen for USENIX Security artifacts, with a slow initial adoption of stable hosting services.

There is a surge of papers having multiple artifacts, visible as a percentage above 100% in Figures 5 and 6, being most apparent at USENIX Security and NDSS. This is primarily due to papers participating in AE that reference both the stable version submitted for evaluation and an actively maintained version, often on Zenodo and GitHub, respectively.

Dedicated domains. Although most of the hosting services used are free, some research projects acquire a custom domain name to gain more visibility or to group the artifacts of a longer line of research. We identify an artifact as being hosted on a dedicated domain if the hyperlink has no path and is a root domain. Figure 7 shows the number of these domains throughout the years, depicting an increase throughout the years and a decline in the last year, making up a total of 3.1% of all artifact links. In terms of TLD used, 42.1% of domains use `.org`, followed by 22.3% for `.com`, and 9.9% for `.net`.

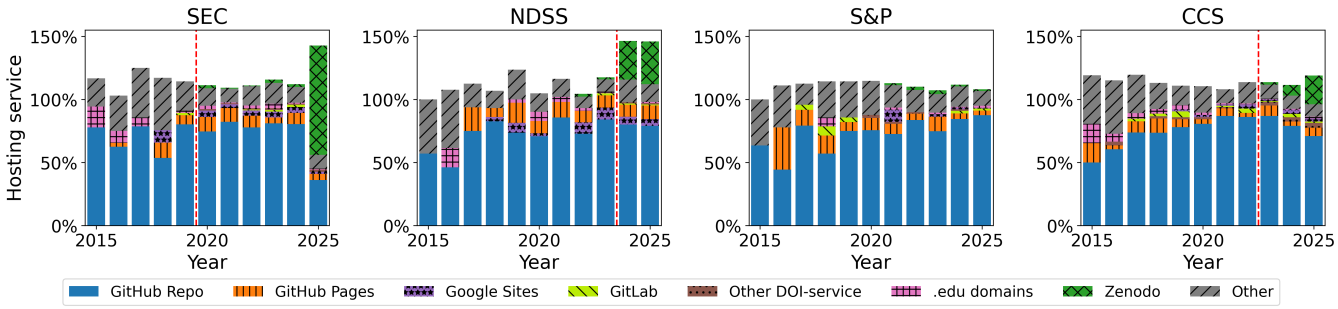

Figure 6: Domains used for hosting artifacts per conference. The introduction of AE is marked with a dotted line.

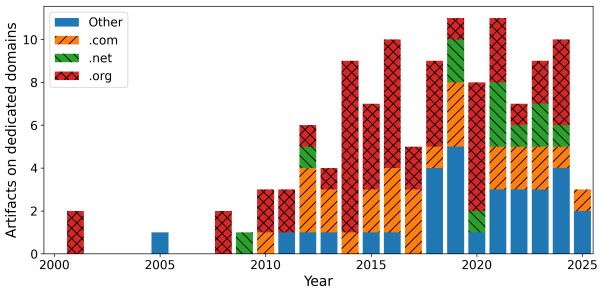

Figure 7: Number of artifacts hosted on dedicated domains.

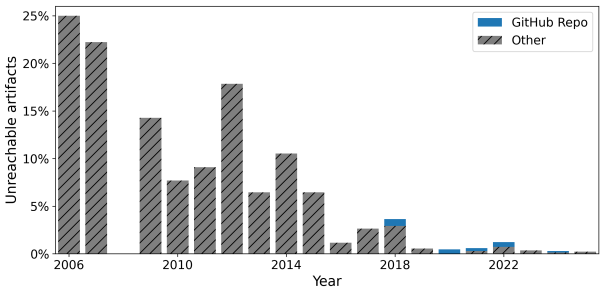

Figure 8: The availability of artifacts is highly dependent on their age, with mostly personal sites becoming unavailable.

4.4 Long-term artifact availability

Over time, artifacts become unreachable when websites are taken offline or domain names expire. Prior studies found up to 10% of artifact links to be unavailable [39], even some hosted on permanent storage services such as Zenodo [16]. During the ranking phase of ArtiFinder, the availability of links is checked, which enables identifying unavailable artifacts, especially when an Internet Archive snapshot is available (cf. Appendix C). Expectedly, the share of unavailable artifacts is steadily decreasing based on the time of publication, as shown in Figure 8. Importantly, all but five unreachable artifacts are hosted on personal or institutional websites, highlighting the value of stable hosting services. With AE pushing for such stable services, long-term artifact availability is expected to improve in the coming years.

Empty repositories. While only 5 GitHub repositories are unavailable, this does not necessarily guarantee the completeness of the artifacts. Specifically, authors may create a repository as a placeholder, which is then never completed. We explore such repositories on GitHub by collecting artifact links pointing to repositories that are empty or contain a single file. After manual control, we find 68 GitHub repositories (2.6% of all GitHub artifacts) that do not contain an artifact.

4.5 Artifacts by paper topic

To explore whether certain subfields of cybersecurity are more likely to share artifacts, we apply topic modelling to our dataset. Concretely, we leverage the topic classification of Schloegel et al. [43] using BERTopic, which was used for studying CVE reporting behavior across cybersecurity topics. Table 4 shows the 16 most common topics together with the shares of papers with an artifact, with fuzzing and microarchitectural research having the highest percentage. However, this result could be influenced by the age of these papers; the mean publication year for these papers is 2021, while the mean publication year for all papers in our dataset is 2019. The topics with the fewest artifacts are encryption, mobile, and cryptography, which see an artifact sharing rate of less than 40%. There is no significant difference in the choice of hosting service between topics.

4.6 Authors and affiliations

Author count. Although creating artifacts requires more work from the authors, having multiple authors in a paper is not a strong indication for a higher chance of artifacts. We only find a limited point-biserial correlation ($\rho = 0.15$, $p = 2.2 * 10^{-48}$) between the presence of an artifact and the number of authors on a paper, visualized in Figure 9.

Table 4: Artifact sharing rate by paper topic for the 16 most common topics in our dataset.

Topic	# Papers	Artifact	Topic	# Papers	Artifact
fuzzing	178	73.6%	networking	101	47.5%
microarchitectures	119	70.6%	social issues	127	47.2%
vulnerabilities	347	57.6%	intrusion detection	111	46.0%
hardware	172	54.7%	cryptocurrencies	209	45.5%
machine learning	607	51.2%	computations	155	45.2%
privacy	842	49.6%	mobile	319	39.2%
malware	165	49.1%	encryption	113	38.9%
browser security	140	48.6%	cryptographic	1678	35.0%

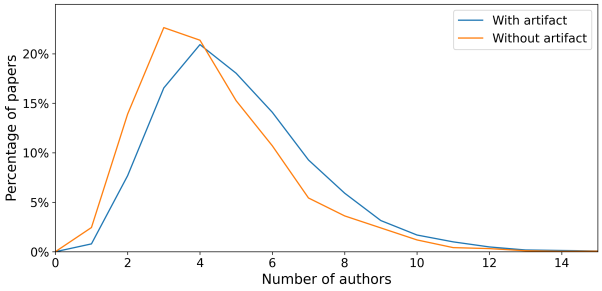

Figure 9: The number of authors is only weakly associated with the presence of an artifact.

Institutions. We also observe a large variance in how often different institutions publish artifacts for their papers. Table 5 shows the ten institutions with at least 50 papers that have the highest artifact sharing rate. The full list can be found in our artifact. Aside from universities or public institutions, private companies also regularly publish at cybersecurity conferences. Possibly due to strict intellectual property regulations, certain private companies are found at the bottom of the artifact share ranking, with artifact sharing rates as low as 8.0%. Note that these statistics might also be influenced by other factors, such as different institutions focusing on different subfields of cybersecurity (cf. Section 4.5).

4.7 Popularity metrics

Although the main driver for artifact release is to enable reproducibility and future work, it might also influence the visibility and impact of the paper [26]. Previous research has found conflicting results regarding the correlation between artifact availability and paper impact, mostly measured in citation counts [8, 10, 27, 48, 49]. We recreate such an experiment on our dataset by collecting citation counts from Scopus [19], shown in Figure 10.

Since more recent papers are likely to have fewer citations than older papers, we use an ordinary least squares (OLS) regression to quantify the influence of artifact release and publication year on citation count. OLS determines the influence of artifact presence on citations while normalizing for publication year. While the regression model suggests that papers

Table 5: Institutions that published over 50 papers at A* conferences with the highest share of artifacts per paper. Only the top 10 are shown.

Institution	Artifact share	Papers
CISPA - Helmholtz Center for Information Security	69.3%	274
Technische Universitat Graz	66.7%	69
Singapore Management University	66.1%	56
Virginia Polytechnic Institute and State University	65.2%	66
Nanyang Technological University	64.4%	73
Zhongguancun Laboratory	63.6%	77
Zhejiang University	63.0%	219
Huazhong University of Science and Technology	62.5%	56
Hong Kong University of Science and Technology	62.2%	90
The Hong Kong Polytechnic University	61.7%	81

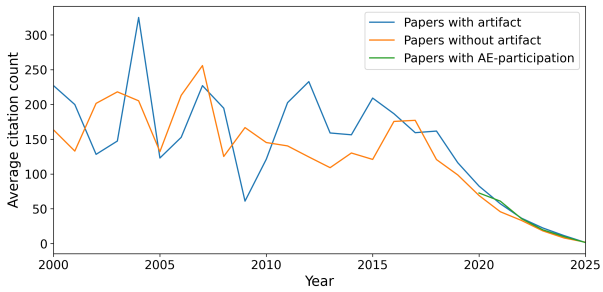

Figure 10: The average citation count for papers with and without an artifact per year.

with artifacts receive 4.9 more citations than those without (95% confidence interval of $[-5.5, 15.3]$), the p-value (0.358) and other diagnostics (cf. Appendix A) indicate statistically insignificant results. Applying the same method to assess the impact of AE participation on citation count also yields inconclusive results ($p = 0.662$), with a 95% confidence interval of $[-4.4, 7.0]$ more citations for papers participating in AE. It is, however, possible that other factors, such as authors or institutional reputation, influence these results.

Similar to prior work [16], we also compared popularity metrics of artifacts hosted on GitHub that participated in AE with detected artifacts that chose not to undergo AE. Figure 11 shows large fluctuations due to a few high-impact repositories, hindering any meaningful correlation calculation.

5 Case study: generalizability of ArtiFinder

While our artifact analysis focused on elucidating trends in A*-ranked cybersecurity venues, the automation in ArtiFinder enables easily extending the dataset with additional venues (subject to access to the papers). To show the generalizability of ArtiFinder beyond A* venues, we run it as a proof-of-concept on 617 papers published from 2017 until 2025 at ACSAC, an A-ranked cybersecurity conference with the longest-standing AE in cybersecurity (cf. Section 2.2). We perform

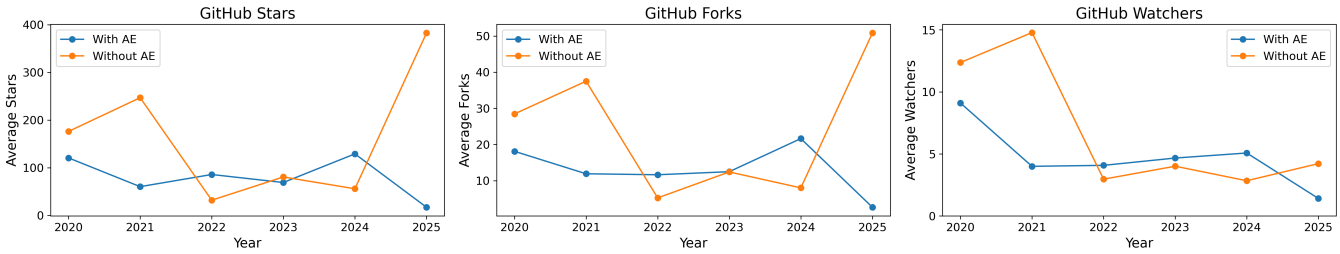

Figure 11: Average GitHub Stars, Forks, and Watchers based on AE participation.

Table 6: Outcomes of ArtiFinder for identifying artifacts for ACSAC papers.

	secartifacts (n=269)
Correct presence	67.7%
Exact link	51.7%
Alternative	16.0%
Correct absence	21.9%
No link	0.0%
Missing link	21.9%
Wrong presence	5.2%
No link	0.0%
Incorrect link	5.2%
Wrong absence	5.2%
Final accuracy	89.6%

this analysis without making any changes to ArtiFinder after our initial calibration (Section 3).

We perform our accuracy analysis with secartifacts as a baseline. The results are shown in Table 6. While the accuracy is slightly lower than for A* venues (89.6% instead of 93.3–98.8%), the most notable insight is that, despite the long history of AE, 21.9% of ACSAC papers with an artifact did not include a link to it in the text, a much higher rate than at A* venues (cf. Table 3).

In Figure 12, we characterize the share of hosting platforms at ACSAC. The majority of the artifacts are hosted on GitHub, with a small number on other domains, mostly project websites. ACSAC does not impose any requirements on hosting providers for AE, which means that this distribution reflects the preference of the authors, showing strong similarities with S&P from Figure 6, also imposing no restrictions.

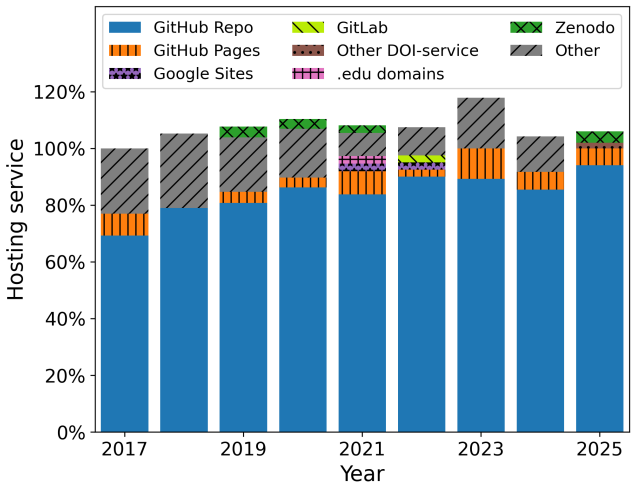

Figure 12: Hosting providers chosen by ACSAC authors, where GitHub is the most popular, with limited DOI-backed services.

6 Discussion

6.1 AE policy impact

Over the years, conferences adopted artifact policies to encourage authors to publish reproducible research. In our dataset, we observed that the artifact sharing rate has steadily increased over the years, even before the introduction of AE (cf. Section 4.1). We also observed that similar policies can have different outcomes depending on enforcement. Requiring authors to submit their artifacts for availability evaluation at USENIX Security 2025 has increased the share of papers linking to an artifact to over 80% (but with limited effect on the share applying for higher badge categories). CCS requires authors to upload an artifact on submission, but only if the paper's claims rely on it, and participation in AE is not mandatory. Seemingly, this less strict policy did not have the same effect on artifact presence in the published papers as at USENIX Security.

However, policy changes did influence *where* artifacts are hosted. Zenodo is increasingly used, likely in part due to USENIX and NDSS enforcing stable storage for the Avail-

ability badge. Although CCS has had a similar requirement since 2024, the rise of Zenodo or other DOI-backed services is not as noticeable there, likely meaning that authors often do not link to the version of their artifact that underwent AE. Overall, the positive evolution of using stable hosting services ensures a higher availability of artifacts, although code repositories such as GitHub also largely guarantee long-term availability. Importantly, we observed that policy changes at one conference did not seem to influence the choice of hosting providers of papers published at other conferences.

With availability on a stable service becoming mandatory, we see a rise in papers linking to multiple artifacts. This often happens when linking to both the artifact submitted for evaluation and a project website or an updatable repository. The former benefits reproducibility, as the exact code and data used for the paper is available long-term (Section 4.4), while the latter allows authors to continue developing their work, while also keeping it compatible with future software or hardware (also benefiting reproducibility). On the flip side, this also leads to fragmentation, and researchers might accidentally build on an outdated version of the artifact if they do not find the up-to-date variant. As a concrete example, the secartifacts dataset currently only contains the link used for artifact evaluation, which might be different from the maintained version.

6.2 Incentives for artifact sharing

User studies of AEC members and authors showed that the main incentive for publishing artifacts is transparency and reproducibility [26]. Prior research is conflicting on whether artifact sharing causes a positive influence on metrics such as citation counts [8, 10, 27, 48, 49], and our analysis also did not reveal statistically significant results. As more papers include artifacts, authors might experience peer pressure to also do so. The absence of an artifact risks being seen as less trustworthy due to the perceived lack of reproducibility. Yet, while this pressure boosts the number of available artifacts, it does not necessarily guarantee their quality. Consequently, AE for Functional or Reproducible badges may become increasingly important to ensure that published artifacts are genuinely useful and usable for other researchers.

Furthermore, the lack of incentives or time pressure might withhold an author from publishing an artifact [37]. Conferences address this by only organizing AE after paper acceptance notification, giving authors more time to finalize the artifacts. Moreover, the *Distinguished Artifact Awards* can also function as an incentive for some authors to release an artifact. Similar awards include the ACSAC Cybersecurity Artifacts Competition and Impact Award [3], organized since 2022 for all cybersecurity papers (regardless of publisher), which might encourage the release of artifacts, also beyond the ACSAC conference.

We found a large variance in the share of papers with artifacts across institutions. Some institutions might pose requirements or have higher expectations concerning open science, possibly as a result of funding requirements [28]. On the other hand, research performed in an industry context might prevent the open release of research artifacts, which cannot even be counteracted by conference policies.

6.3 Meta-research challenges

Over the course of this research project, we identified several shortcomings in the current research ecosystem that affect the ability to conduct meta-research work, particularly regarding artifacts and reproducibility.

Despite the efforts of the secartifacts website to collect all information relevant to AE, results of AE are not always accurately listed or are scattered across various sources. During the development of ArtiFinder, we identified missing artifact links, badges, and incorrect titles in the results as reported on secartifacts. While we contributed fixes for these issues, such issues may arise again in the future. Most notably, although CCS has conducted AE since 2023, its results are not reported on secartifacts. For the 2023 edition of CCS, the results of AE are completely unknown, as artifact appendices were not required, and the results are not published anywhere online.

To gather metadata for all papers, we leveraged DBLP and SCOPUS for citations. Unfortunately, these databases are not always complete and can have missing entries, both for published papers and the papers that cite them. To combat this, relevant conference websites could be scraped, but this is hindered by inconsistent layouts and restrictive terms of service. Additionally, other services that could be used to broaden our dataset, such as Crossref [42], are limited to papers with a DOI, which USENIX does not distribute.

Finally, we also observed that information contained in paper PDFs is also often incomplete. Some authors choose not to include a link to their artifact, even if it is published or even submitted for AE. Additionally, whether the obtained badges are displayed on the published paper itself also depends on the conference edition.

7 Recommendations

Based on our findings, we list some recommendations for both authors and conference organizers to improve artifact availability and discoverability, boosting reproducibility.

7.1 Artifact sharing

Thanks to ArtiFinder, we are able to identify artifacts in a paper. Yet, it should be possible to more easily find artifacts without the need for an automated tool. Artifact appendices can fulfill this role by centralizing all information about the artifact in the paper. Unfortunately, these appendices are currently only added for papers that passed AE and require sig-

nificant time investment from the authors. Even for artifacts that were—for various reasons—not submitted for evaluation, such an appendix could help to quickly identify the artifact. Moreover, (similar to statements on ethical concerns) short statements on open science should be encouraged or mandated, as is done now at e.g., USENIX Security and EuroS&P. This permits readers to find relevant code or datasets in a quickly identifiable location in the paper, without requiring a detailed appendix from the authors.

Similarly, unifying the awarded badges and always including them in the published paper would help with identifying papers that have (evaluated) artifacts. At the same time, we found numerous papers that were awarded badges but do not mention the artifact in the text. Although the badges indicate to a reader that the paper has an artifact, it cannot easily be found, and might not even be publicly available (especially if no Available badge was awarded). Having a tighter connection between the badges and the artifact (e.g., a clickable Available badge leading to the artifact) could clarify these cases and increase the discoverability of artifacts.

7.2 Unified requirements

Section 2.2 identified several differences in policies between conferences. Additionally, some conferences publish evaluation guidelines split up over different locations, such as secartifacts and their own website. With differing badge requirements, artifact appendices, and packaging instructions, we believe this process should be unified to streamline AE. By having a standardized way of sharing research artifacts in the community, it would become easier for both authors and AEC to manage expectations and converge towards more uniform and easily reusable artifacts. Finally, we observed that some papers do not follow the requirements for linking to a stable hosting service in the published version of the paper, which could be more strictly enforced if desired by conference organizers.

8 Related work

Our work identifies artifact sharing trends in top-tier cybersecurity papers over the past 25 years. By developing a tool capable of automatically identifying artifacts, we are able to significantly reduce the workload typically associated with meta-science studies. In this section, we compare related work to highlight its differences and similarities with our research.

Meta-science and reproducibility. The practice of conducting meta-science in computer science and cybersecurity research is well-established [9, 24, 36]. For example, Vandewalle found in 2019 that only 25% of published papers in the field of image processing have available code [48]. Similarly, in machine learning security, Olszewski et al. [40] discovered that less than half of papers published in Tier 1 security conferences have available code and identified common reproducibility problems. Similar to our results, they noted a significant increase in availability over the years, showing the continuous effort of the community towards more reproducible results. In a more recent study, focusing on applied security, Olszewski et al. [39] found that AE did not significantly improve the reproducibility of artifacts and showed the inconsistencies between evaluation processes due to different guidelines. Crowder et al. complement reproducibility research on code by analyzing dataset artifacts for security measurement papers [14]. By manually checking artifacts, they identified shortcomings in current data sharing practices and listed recommendations. In systems research, Collberg and Proebsting [13], van der Kouwe et al. [47], and D'Elia et al. [16] all provided recommendations to publishers and authors on how to best tackle the reproducibility issues, while also showing how artifact sharing improved over time. Compared to other research, we limit ourselves to the identification of artifacts without studying their effective reproducibility.

Artifacts. While previous research on artifacts often required substantial time investment, ArtiFinder significantly cuts down this cost. Closely related to our work, Winter et al. developed a tool to identify links in a PDF, still requiring manual labeling [49]. Although efforts have been made by the community to create an overview of available research artifacts, these are limited to papers participating in AE [45]—relying on the conference organizers for the data—or are manually curated datasets often limited to a small set of venues [6] or a subset of topics or types of research [39, 40, 43]. As we have shown, these datasets might also suffer from omissions or errors. To help address this issue, we are working with the maintainers of secartifacts to expand their dataset and create a central database of research artifacts in cybersecurity.

9 Limitations and future work

ArtiFinder identifies artifacts by parsing all hyperlinks in a paper, so an artifact can only be identified if it is linked in the paper. By comparing our dataset with results from other sources, we found that 8.4–21.9% of papers with an artifact did not include a link to it, preventing us from detecting and analyzing these artifacts.

To enhance the visibility of historical artifacts, we have contacted the maintainers of secartifacts and, specifically for non-AE papers, are developing an integration of our results into the platform, with a disclaimer and the option to manually verify our results. Currently, secartifacts lists 1,112 artifacts for top-tier papers; integrating our dataset would increase this by 275% to 4,165. We are also discussing how to integrate alternative (non-stable) links into this dataset.

While we believe that the accuracy of ArtiFinder (cf. Section 3.3) is sufficient to perform our analysis and make in-

formed, data-driven decisions about policies, it certainly introduces some errors in our dataset. We encourage future work to extend or adapt our open source implementation of ArtiFinder to further improve the detection accuracy and correct mistakes in our dataset. This could involve extending the heuristics, possibly even with a carefully considered LLM-based approach. We also encourage researchers in other fields to leverage our tool to identify artifacts outside of cybersecurity.

Finally, to scope our work, we did not study the reproducibility of artifacts as in previous work [8, 39, 40]. We hope that our dataset enables future replication studies across a much larger and longitudinal corpus. Similarly, our statistical analyses mostly focused on re-running experiments from prior work; more advanced methods would likely be able to identify the impact of conference policies and artifact sharing more accurately.

10 Conclusion

In cybersecurity, authors often produce artifacts to support the findings of papers. We complement existing work in meta-science by developing ArtiFinder, a high-accuracy tool that automatically identifies artifacts in research papers. This significantly reduces the time needed to perform a large-scale analysis of artifact presence and availability. Using ArtiFinder, we create a dataset of 3,922 artifacts published in top-tier cybersecurity venues and conduct the largest evaluation to date of artifact sharing. We find an increasing share of papers with artifacts, from 1.2% of papers in 2000 to 72.3% in 2025. While this increase cannot solely be attributed to AE policies, they do highly influence the choice of hosting service, leading to increased artifact availability over time.

In our research, we also identify remaining challenges with artifact sharing and encourage the community to adopt a unified approach to ensure long-term artifact discoverability.

Acknowledgments

We would like to thank everyone who shares artifacts for their papers, as well as those who promote and support artifact sharing and AE, including those serving on AECs, for their contribution to open science. We are grateful to the anonymous reviewers for their helpful feedback on this paper and to Anjo Vahldiek-Oberwagner for productive discussions and advice on contributing to secartifacts. This research is partially funded by the Internal Funds KU Leuven and by the Cybersecurity Research Program Flanders. Victor Le Pochat is currently employed by the European Commission. The views and opinions expressed herein are personal and do not necessarily reflect those of the European Commission or other EU institutions.

Ethical Considerations

Our work can be placed in the wider body of meta-science, where we do not uncover new vulnerabilities and only analyze publicly accessible artifacts. We adhered to the rate limits imposed by API services and complied with best practices regarding web crawling.

For our research, we identify five stakeholders: authors of analyzed papers, conference organizers, artifact hosting providers, institutions, and the research community at large.

Authors. Since we only used published papers and publicly accessible information on the internet, we believe that individual authors are not negatively impacted. We refrained from highlighting issues with specific papers and only showed aggregated statistics. Upon the publication of our dataset and integration into secartifacts, we will provide a mechanism to correct mistakes, accompanied by a disclaimer that our data is the result of automated extraction.

Conference organizers & committee members. With our research, we hope to give conference organizers insights into trends related to artifacts, which can inform decisions on new policies. These policies would then (hopefully positively) also impact the evaluators and authors.

Artifact hosting providers. We discuss and compare different hosting providers and their characteristics, but we do not aim to make a value judgement. Our goal is to highlight advantages, which can then be adopted by different providers, better aligning with the needs of authors and conference organizers.

Institutions. In our research, we quantify the share of artifacts per institution, which could be interpreted as a positive or negative attribute. To limit negative associations, we only highlight the institutions that share the most proactively, and we explain that there are legitimate reasons (e.g., intellectual property restrictions) for not sharing artifacts.

Research community. The community will benefit from the publication of our results, as it allows the discovery of published artifacts and enables future meta-science research.

Open Science

The initial version of ArtiFinder and our collected dataset is archived at https://doi.org/10.5281/zenodo.20412201. Due to copyright restrictions, the paper PDFs cannot be distributed, but all the findings of the paper can be reproduced using this archive, which also includes scripts to collect, aggregate, and visualize data.

We submitted a pull request to secartifacts at https://github.com/secartifacts/secartifacts.githu

b.io/pull/147, adding our dataset to the website, and to CyCoAnalysis at https://github.com/CyCoAnalysis/dataset/pull/1, contributing the collection of conference policies regarding artifacts.

We are continuing the development of ArtiFinder at https://github.com/DistriNet/ArtiFinder, where we invite contributions, and maintain an active version of the dataset at https://github.com/DistriNet/ArtiFinder-Data, open to manual corrections and extensions.

References

[1] ACM SIGSAC. ACM CCS History, 2025. URL: https://www.sigsac.org/ccs/ccs-history.html.

[2] ACSAC. Call for Submissions - Security Conference, Security Training & Security Networking - ACSAC 2017, 2017. URL: https://www.acsac.org/2017/cfp/.

[3] ACSAC. Cybersecurity artifacts competition and impact award - ACSAC 2022, 2022. URL: https://www.acsac.org/2022/submissions/artifacts_competition/.

[4] Association for Computing Machinery. Artifact review and badging, 2020. URL: https://www.acm.org/publications/policies/artifact-review-and-badging-current.

[5] Monya Baker. 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604):452–454, May 2016. doi:10.1038/533452a.

[6] David Balenson, Terry Benzel, Eric Eide, David Emmerich, David Johnson, Jelena Mirkovic, and Laura Tinnel. Toward Findable, Accessible, Interoperable, and Reusable Cybersecurity Artifacts. In *Proceedings of the 15th Workshop on Cyber Security Experimentation and Test*, CSET '22, pages 65–70, August 2022. doi:10.1145/3546096.3546104.

[7] Marton Bognar, Lieven Desmet, and Frank Piessens. CyCoAnalysis: A dataset of cybersecurity conference metadata for meta-science and policy decisions. In *IEEE Security and Privacy Workshops (SPW)*, May 2026.

[8] Nicolas Bonneel, David Coeurjolly, Julie Digne, and Nicolas Mellado. Code replicability in computer graphics. *ACM Trans. Graph.*, 39(4), August 2020. doi:10.1145/3386569.3392413.

[9] Thomas E. Carroll, David Manz, Thomas Edgar, and Frank L. Greitzer. Realizing scientific methods for cyber security. In *Proceedings of the 2012 Workshop on Learning from Authoritative Security Experiment Results*, LASER '12, pages 19–24, July 2012. doi:10.1145/2379616.2379619.

[10] Bruce R. Childers and Panos K. Chrysanthis. Artifact Evaluation: Is It a Real Incentive? In *2017 IEEE 13th International Conference on e-Science (e-Science)*, pages 488–489, October 2017. doi:10.1109/eScience.2017.79.

[11] CloudLab Team. CloudLab, 2025. URL: https://www.cloudlab.us/index.php.

[12] Christian Collberg, Todd Proebsting, and Alex M Warren. Repeatability and benefaction in computer systems research, 2015. URL: https://web.archive.org/web/20250223094935/https://repeatability.cs.arizona.edu/v2/RepeatabilityTR.pdf.

[13] Christian Collberg and Todd A. Proebsting. Repeatability in computer systems research. *Commun. ACM*, 59(3):62–69, February 2016. doi:10.1145/2812803.

[14] Anna Crowder, Allison Lu, Kevin Childs, Carson Stillman, Patrick Traynor, and Kevin R.B. Butler. Data to Infinity and Beyond: Examining Data Sharing and Reuse Practices in the Computer Security Community. In *2025 IEEE Symposium on Security and Privacy (SP)*, pages 2678–2696, May 2025. doi:10.1109/SP61157.2025.00180.

[15] CSPaper. Paper with code redirects to GitHub – what it means for CS researchers, 2025. URL: https://forum.cspaper.org/topic/118/paper-with-code-redirects-to-github-what-it-means-for-cs-researchers.

[16] Daniele Cono D'Elia, Thaleia Dimitra Doudali, Cristiano Giuffrida, Miguel Matos, Mathias Payer, Solal Pirelli, Georgios Portokalidis, Valerio Schiavoni, Salvatore Signorello, and Anjo Vahldiek-Oberwagner. Lessons learned from five years of artifact evaluations at EuroSys. In *Proceedings of the 3rd ACM Conference on Reproducibility and Replicability*, pages 108–120, 2025. doi:10.1145/3736731.3746152.

[17] Digital Science. Figshare, 2025. URL: https://figshare.com/.

[18] Dryad. Dryad, 2025. URL: https://datadryad.org/.

[19] Elsevier. Scopus | Abstract and citation database | Elsevier, 2025. URL: https://www.elsevier.com/products/scopus.

[20] ESEC/FSE Organization. Call for Artifact Evaluation | ESEC/FSE 2011, 2011. URL: http://2011.esec-fse.org/cfp-artifact-evaluation.

[21] European Organization For Nuclear Research and OpenAIRE. Zenodo, 2013. URL: https://www.zenodo.org/, doi:10.25495/7GXK-RD71.

[22] freedesktop.org. Poppler, 2005. URL: https://poppler.freedesktop.org/.

[23] Florian Hantke, Stefano Calzavara, Moritz Wilhelm, Alvise Rabitti, and Ben Stock. You Call This Archaeology? Evaluating Web Archives for Reproducible Web Security Measurements. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, pages 3168–3182, November 2023. doi:10.1145/3576915.3616688.

[24] Cormac Herley and P.C. Van Oorschot. SoK: Science, Security and the Elusive Goal of Security as a Scientific Pursuit. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 99–120, May 2017. ISSN: 2375-1207. doi:10.1109/SP.2017.38.

[25] Ben Hermann. What Has Artifact Evaluation Ever Done for Us? *IEEE Security & Privacy*, 20(5):96–99, September 2022. doi:10.1109/MSEC.2022.3184234.

[26] Ben Hermann, Stefan Winter, and Janet Siegmund. Community expectations for research artifacts and evaluation processes. In *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pages 469–480, November 2020. doi:10.1145/3368089.3409767.

[27] Robert Heumüller, Sebastian Nielebock, Jacob Krüger, and Frank Ortmeier. Publish or perish, but do not forget your software artifacts. *Empir Software Eng*, 25(6):4585–4616, November 2020. doi:10.1007/s10664-020-09851-6.

[28] Horizon Europe. Horizon europe programme guide, 2025. URL: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf.

[29] ICORE. ICORE conference rankings, 2026. URL: https://portal.core.edu.au/.

[30] IEEE. IEEE Symposium on Security and Privacy, 2025. URL: https://sp2026.ieee-security.org/past.html.

[31] Internet Society. Previous NDSS Symposia, 2025. URL: https://www.ndss-symposium.org/previous-ndss-symposia/.

[32] John P. A. Ioannidis, Daniele Fanelli, Debbie Drake Dunne, and Steven N. Goodman. Meta-research: Evaluation and improvement of research methods and practices. *PLOS Biology*, 13(10):1–7, 10 2015. doi:10.1371/journal.pbio.1002264.

[33] Kate Keahey, Jason Anderson, Zhuo Zhen, Pierre Riteau, Paul Ruth, Dan Stanzione, Mert Cevik, Jacob Colleran, Haryadi S. Gunawi, Cody Hammock, Joe Mambretti, Alexander Barnes, François Halbach, Alex Rocha, and Joe Stubbs. Lessons learned from the Chameleon testbed. In *Proceedings of the 2020 USENIX Annual Technical Conference (USENIX ATC '20)*. July 2020. URL: https://www.usenix.org/system/files/atc20-keahey.pdf.

[34] Jan H. Klemmer, Juliane Schmüser, Fabian Fischer, Jacques Suray, Jan-Ulrich Holtgrave, Simon Lenau, Byron M. Lowens, Florian Schaub, and Sascha Fahl. How transparent is usable privacy and security research? a Meta-Study on current research transparency practices. In *34th USENIX Security Symposium (USENIX Security 25)*, pages 5967–5986, August 2025. URL: https://www.usenix.org/system/files/usenixsecurity25-klemmer.pdf.

[35] Victor Le Pochat. Reflecting on research practices. *Communications of the ACM*, 67(5):37–39, May 2024. doi:10.1145/3651965.

[36] Victor Le Pochat and Wouter Joosen. Analyzing Cyber Security Research Practices through a Meta-Research Framework. In *2023 Cyber Security Experimentation and Test Workshop*, pages 64–74, August 2023. doi:10.1145/3607505.3607523.

[37] Tom Longstaff, David Balenson, and Mark Matties. Barriers to science in security. In *Proceedings of the 26th Annual Computer Security Applications Conference*, ACSAC '10, page 127–129, 2010. doi:10.1145/1920261.1920281.

[38] NSF. SPHERE Testbed, 2025. URL: https://launch.sphere-testbed.net.

[39] Daniel Olszewski, Allison Lu, Anna Crowder, Nathaniel Bennett, Seth Layton, Sri Hrushikesh Varma Bhupathiraju, Tyler Tucker, Siddhant Kalgutkar, Hunter Ver Helst, Carson Stillman, Kevin R. B. Butler, Sara Rampazzi, and Patrick Traynor. Reproducibility in Applied Security Conferences: An 11-Year Review on Artifacts and Evaluation Committees. In *Proceedings of the 3rd ACM Conference on Reproducibility and Replicability*, pages 96–107, July 2025. doi:10.1145/3736731.3746151.

[40] Daniel Olszewski, Allison Lu, Carson Stillman, Kevin Warren, Cole Kitroser, Alejandro Pascual, Divyajyoti Ukirde, Kevin Butler, and Patrick Traynor. "Get in Researchers; We're Measuring Reproducibility": A Reproducibility Study of Machine Learning Papers in Tier 1 Security Conferences. In *Proceedings of the 2023*

ACM SIGSAC Conference on Computer and Communications Security, pages 3433–3459, November 2023. doi:10.1145/3576915.3623130.

[41] Daniel Olszewski, Tyler Tucker, Kevin R. B. Butler, and Patrick Traynor. SoK: Towards a unified approach to applied replicability for computer security. In *34th USENIX Security Symposium (USENIX Security 25)*, pages 469–488, August 2025. URL: https://www.usenix.org/system/files/usenixsecurity25-olszewski.pdf.

[42] PILA. Crossref, 2025. URL: https://www.crossref.org/.

[43] Moritz Schloegel, Daniel Klischies, Simon Koch, David Klein, Lukas Gerlach, Malte Wessels, Leon Trampert, Martin Johns, Mathy Vanhoef, Michael Schwarz, Thorsten Holz, and Jo Van Bulck. Confusing value with enumeration: Studying the use of CVEs in academia. In *34th USENIX Security Symposium (USENIX Security 25)*, pages 2887–2906, August 2025. URL: https://www.usenix.org/system/files/usenixsecurity25-schloegel.pdf.

[44] Schloss Dagstuhl. dblp: computer science bibliography, 2025. URL: https://dblp.org/.

[45] secartifacts. Security Research Artifacts, 2025. URL: https://secartifacts.github.io/.

[46] USENIX. USENIX Security Symposia, 2025. URL: https://www.usenix.org/conferences/byname/108.

[47] Erik van der Kouwe, Gernot Heiser, Dennis Andriesse, Herbert Bos, and Cristiano Giuffrida. SoK: Benchmarking Flaws in Systems Security. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 310–325, June 2019. doi:10.1109/EuroSP.2019.00031.

[48] Patrick Vandewalle. Code availability for image processing papers: a status update. In *WIC IEEE SP Symposium on Information Theory and signal Processing in the Benelux*, 2019. URL: https://lirias.kuleuven.be/2815281.

[49] 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. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ESEC/FSE 2022, pages 145–156, November 2022. doi:10.1145/3540250.3549172.

[50] WiSec. Call for papers, posters, demos - WiSec 2017, 2017. URL: https://web.archive.org/web/20170322014332/http://wisec2017.ccs.neu.edu/call-for-papers-posters-demos.html.

A OLS Output

Table 7 shows the full results of the OLS regression test from Section 4.7.

Table 7: Full OLS regression results of citations and artifact presence.

Variable	Coef.	Std. Err.	t-stat	P> |t|	[0.025	0.975]
const	256.0147	7.570	33.822	0.000	241.177	270.853
artifact	4.8926	5.322	0.919	0.358	-5.540	15.325
year	-9.6009	0.425	-22.579	0.000	-10.434	-8.767
Fit Statistics			**Diagnostics**			
R-squared	0.069		Omnibus	15855.745		
Adj. R-squared	0.068		Prob(Omnibus)	0.000		
No. Obs.	8,216		Skew	14.971		
F-statistic	302.0		Kurtosis	369.618		
Prob (F-stat)	2.71e-127		Durbin-Watson	1.904		
Log-Likelihood	-55,808		Jarque-Bera (JB)	46319455.585		
AIC	1.116e+05		Prob(JB)	0.00		
BIC	1.116e+05		Cond. No.	62.8		

B ArtiFinder heuristics

Table 8 lists the heuristics used in ArtiFinder.

C Internet Archive

Whenever a link is unreachable, ArtiFinder can fetch its archived content from the Internet Archive. By using this method, we aim to assess more links and reach a higher accuracy. Otherwise, ArtiFinder would have to skip the final phase of its ranking process, which is used to check if, e.g., the title of the paper is present in the content of the link.

However, this approach comes with two limitations. First, the Internet Archive is, as a service, unstable and has regular downtime. According to our own experience and services such as downdetector.com[3], daily outages and failed requests are common. This can lead to a different ranking of links depending on whether content can be fetched from the archive. Inclusion of this service, however, will always positively impact the results, providing a more accurate view of unreachable links.

Second, the Internet Archive lookup service is quite slow and is severely rate-limited, exchanging accuracy of ArtiFinder for speed [23]. While ArtiFinder normally has a timeout of 5 seconds for fetching a link, this is increased to 60 seconds for Internet Archive requests to account for its slow service. Additionally, rate-limits are in place with a

[3] https://downdetector.com/status/internetarchive/

Table 8: Overview of heuristics employed by ArtiFinder to determine the artifact and their scores.

	Description	Score
Parsable	Checks if the hyperlink is parsable as a URL	-100
JoinedSentence	Poppler often merges the beginnings and ends of sentences, which is identified as a link	-100
NotDomain	Manual blocklist of domains that do not host an artifact, mainly to exclude venue websites	-100
GithubRepo	Checks if the link leads to a GitHub Repo	+10
GithubStable	Extra points if the GitHub link is a stable link, i.e., specific commit or tag	+2
DoiZenodo	A DOI link that leads to a Zenodo archive	+10
ZenodoArchive	Checks if the link is a Zenodo archive	+10
TitleInUrl	Does the title of the paper appear (partially) in the hyperlink?	+10
LocationInPaper	If the link is in the references, we do not award points; links in a paragraph or footnote are awarded bonus points	+10, +15
LinkParagraphContext	The presence of certain keywords (available, artifact) influences the score	+15
FailedRequest	The hyperlink is resolved; for failed requests, we default to Internet Archive. If the archive does not have the page, the rest of this phase is cancelled	-20
TitleInContent	The full title of the paper can be present in the retrieved HTML	+20
AuthorInContent	The authors of the paper can be present in the retrieved HTML, points are based on the number of authors present	+0–10
PartialTitleInContent	The title of the paper can be present in the retrieved HTML, points are based on the amount of title present	+0–20
Citation	Parses the HTML to check for the presence of a BibTeX citation that matches the paper	+50
Created	Specifically for GitHub and Zenodo, the date the repository was created can be at most 3 years in the past.	-20
FinalNotDomain	Employs the same checks as NotDomain, but on the final URL in case the original hyperlink redirected	-100

maximum of 15 requests per minute[4]. These limitations, on average, increase our execution time to 317 seconds per paper when using the Internet Archive.

[4] https://archive.org/details/toomanyrequests_20191110