

Oliver Porth<sup>1</sup>

<sup>1</sup>Affiliation not available

October 24, 2022

The paper *Reproducibility of the First Image of a Black Hole in the Galaxy M87 from the Event Horizon Telescope (EHT) Collaboration* by Patel et al. describes the steps required to reproduce the M87 black-hole image from data artifacts released by the EHT collaboration. The paper also gives an overview of the applied software stacks and provides highly valuable containerized environments that enable easy execution of the three imaging pipelines that reproduce the ones from the EHT. The authors succeed in generating images and compare their products (images and statistics) against the published results, generally finding consistent results. The paper also critically reviews the documentation and reproducibility efforts made by the EHT collaboration itself.

I find the paper particularly interesting for two reasons:

1. It significantly lowers the entry barrier to EHT-type work and enables the public to redo the data processing steps behind the iconic black-hole image. Thus it can lead to increased participation in fundamental science and raise public interest in STEM research.
2. The work itself is a case study of how to disseminate data and software products in practice and can serve as an example to future efforts.

I recommend publications after incorporating the following suggestions, all of which should be straightforward to handle.

### Major comments

It should also be stated what this paper is not. It does not aim at an independent analysis of the EHT data since it only rebuilds the EHT pipelines. As far as I am aware, parameter choices are adopted from the EHT papers. It's important to point out this distinction in the paper. There have been several more or less independent analyses of the EHT data products by other groups (Arras et al., 2022; Carilli & Thyagarajan, 2022; Lockhart & Gralla, 2022; Miyoshi et al, 2022) and in particular the Miyoshi work has found different results (prompting communications such as <https://arxiv.org/pdf/2207.13279.pdf>). It could be easy to mis-quote the paper at hand for an independent analysis of the EHT results (as done by the EHT collaboration on their website

(<https://eventhorizontelescope.org/blog/imaging-reanalyses-eh- data>), which should be avoided.

I recommend adding a paragraph to the introduction where the aforementioned independent analyses are cited and a distinction between reproduction and independent analysis is made explicit.

I would like to hear the authors' perspective on how the extensive documentation needed to fully reproduce data products should be disseminated in an academic publishing process. Surely the research papers themselves are not the right place. Journals also often request DOIs for supplementary data. Is this compatible with the containerized approach? Further, in connection to funding agencies, FAIR principles

are often requested (<https://www.go-fair.org/fair-principles/>) in research data management (RDM) plans. An additional paragraph (e.g. in the conclusions) connecting this work with the RDM practices imposed by funding agencies and journals will be helpful to evaluate the practicability of the approach followed in this paper.

The validation of the data leading to Figure 2 seems insufficient as it merely shows the integrity of the data timestamps, not the data itself. The authors should comment on why this is sufficient, as the data itself (visibility phases and amplitudes) could still be tampered with. Perhaps the authors could clarify the scope of the validation.

It is not clear by reading the manuscript how much interaction with EHT members was needed to reproduce the results. Could the authors please point out where key information was not available in the publications/data release and input was needed?

A comment on reproducing the reproduced results: I have run all four Docker containers successfully but the eht-imaging one missed essential dependencies which I installed by hand (ehtim and matplotlib)... The authors should check their container (or find the one fixed by me here...)

## Minor comments

In section “Reproducing the EHT Images”, second paragraph, the authors state “... we only report the values with 0% systematic uncertainty”. Please clarify to which systematic uncertainties you refer. It isn’t clear if there might be a fundamental problem reproducing systematic uncertainties. Do they need further information that is not available in the data products? A brief clarification would help.

The DIFFMAP image statistics (Table 2) differs wildly from the original paper. The corresponding statement in the paper,

We also find a larger difference between the original and reproduced values for the DIFFMAP pipeline: this is consistent with the discussion of the different time averaging used in DIFFMAP

does not clarify. Please expand on your explanation, how does time averaging come into play in DIFFMAP?