Editorial
On Writing Reproducible and Interactive Papers

Abstract—**While most publishers of scientific journals provide online archives of papers, in addition to or as a replacement for printed issues, the versions available online remain static documents. They fail to take advantage of the richness of interactivity that the digital world can offer. The time has come to offer authors, who submit their work for publication in the IEEE JOURNAL OF OCEANIC ENGINEERING, emerging technology options that will allow them to provide a richer experience to readers and to share data and code used in producing their results. Adoption of such technologies will also help in promoting data and code transparency, which is important for addressing growing concerns on reproducibility of research results. As an exercise in developing an interactive publication, this editorial is written as a Jupyter notebook that automatically generates the print version. The notebook can be modified and the results regenerated, if the reader so desires. We walk through the development of this notebook, for readers who wish to adopt this approach in their own writing.**

I. INTRODUCTION

While many aspects of academic publishing have evolved over the past three centuries, the essential nature of a paper as an archived immutable static document has not changed significantly. Traditionally, papers were published in printed journals. Now, most publishers provide online digital versions of papers, in addition to or as a replacement for printed papers. But the online digital version is still a static document, just like the printed counterpart. The digital version can be delivered much more quickly and at a lower cost. It can be searched. And it may provide clickable links. But it largely fails to take advantage of the richness of interactivity that the digital world has to offer.

With advancement in technology and a shift in attitude of scientists and researchers, change is inevitable, and even desirable. This editorial reviews some of the technology options that are now available for the IEEE JOURNAL OF OCEANIC ENGINEERING authors who wish to provide a richer experience to their readers. In Section II, I discuss some of the advantages of embracing such technologies, for the authors, reviewers, readers, and for the academic community as a whole.

As an exercise in understanding and demonstrating what it takes to develop interactive papers, I chose to write this editorial as an interactive *notebook*. The *print* version that you see is automatically generated from a Jupyter notebook that is available to the readers via the “Code & Datasets” tab on IEEE Xplore digital library. When the notebook is opened, the reader can see the code that was used to generate the figure in this editorial, and is able to change code or parameters, if desired. In Section III, we walk through the development of this notebook, for readers who wish to adopt this approach in their own writing.

Digital Object Identifier 10.1109/JOE.2018.2848058

II. SHARING CODE AND DATA SETS

The ability to replicate or reproduce research results is one of the key tenets of the scientific method. However, there are rising concerns in the scientific community on the reproducibility of results [1]. Data transparency and analytic methods (code) transparency can help in addressing a large part of this problem through an open research culture [2]. By sharing code and data, authors can ensure that readers are able to fully appreciate their work and reproduce it. More importantly, empowering the readers with code that implements the algorithms or methods, as described in the paper, encourages them to use and enhance the algorithms or methods in their own work. This is the spirit of academic progress, and the payback for open sharing can be seen in terms of increased citations from researchers who build on top of the work, in addition to the satisfaction one derives from seeing ones work being used and appreciated by others.

The idea of sharing data or code is not new. Several authors in our community have, in the past, hosted code or data on their own websites or in public repositories [3]–[5]. In some communities (e.g., robotics, image processing, and machine learning), this approach has become widely adopted and extremely successful. It has helped the community to grow and research advance in a way that would not have been possible without sharing. Some authors are concerned about losing competitive advantage if they share their code or data. However, experience from communities where sharing is common has shown that authors benefit immensely from synergies with other researchers and are able to advance the field more rapidly. Authors who share artifacts that others use and build on become *de facto* thought leaders and well-known researchers in their fields.

One of the shortcomings of hosting on public repositories or websites is that the hosting site is separate from the archive that stores the paper and may have a shorter lifetime or the uniform resource locators (URLs) may change over time. Even if the lifetime of the hosting site and URL can be ensured, the code often has dependencies (e.g., operating system version, external libraries, etc.) that might become unavailable or obsolete over time. In an effort to enable archiving of shared code and data, the IEEE embarked on a partnership with Code Ocean (https://codeocean.com), a cloud-based computational platform that enables users to upload, publish, and run code, without having to purchase or install additional software or applications on their computer [6]. This dramatically reduces the barrier to entry for readers wanting to reproduce results without having to contact the authors, or struggling to acquire appropriate hardware, tools, and dependencies. Once the code is published, it is linked to the associated article in the IEEE Xplore, allowing readers of the article to access it. A digital object identifier is assigned to the published code and the code can be directly cited. The authors maintain the copyright to the code and the published version remains unchanged even when readers change parameters or modify a copy for their own testing.

There are several additional advantages to sharing code and data. Algorithms presented in papers often have tunable parameters with specific values chosen by the authors to produce the results published in the paper. Similarly, analysis of experimental data usually depends on several tunable parameters. Reviewers and readers are empowered with the ability to change those values and study the results, once the code and data are shared with them. This promotes transparency and also has the potential to shorten the peer-review time by allowing interested reviewers to more deeply understand the intricacies of the algorithms and data processing techniques, without having to iterate through multiple review cycles with the authors. Sharing of code and data also promotes fair and objective benchmarking of algorithms and techniques. Such benchmarking exercises are critical in engineering research and yet often not performed because they are difficult to undertake. Finally, several funding agencies around the world now require research artifacts to be made publicly available. This requirement can be fulfilled with code and dataset sharing, as part of a peer-reviewed archived publication of the research results.

III. Research Notebooks As Papers

Sharing of code or datasets need not be onerous. Authors can set up a free Code Ocean (https://codeocean.com) account, create a new compute capsule, and upload their code and data in it. The code may be written in any language (or combination of languages) that the authors prefer (e.g., MATLAB, C, Python, etc.). If necessary, they install dependencies, such as external libraries, in the compute capsule and ensure that their code compiles, runs, and produces the desired results. They then publish the compute capsule and link it with the relevant paper on the IEEE Xplore. The shared code may produce a subset of the results in the papers.

To take this a step further, I wondered if it would be possible to share the code, data, and text for the paper in a way that allows a reader to reproduce the paper in its entirety. With this editorial, I have done exactly that. This editorial is written using a Jupyter notebook, which is shared on Code Ocean.[1] The notebook includes the text of this editorial, all equations, the Python code to generate the figure, and all the citations. Simply running a shell script in the compute capsule will generate the pdf version of this editorial. But an interested reader may edit the notebook and change some of the parameters or code and regenerate the results.

To illustrate how code, equations, and figures are included in the notebook, I include a simple technical section (see Section III-A) in this editorial and encourage readers to explore the Jupyter notebook and try modifying it themselves.

[1]The Code Ocean compute capsule Digital Object Identifier is https://doi.org/10.24433/CO.0a11ea60-3b65-4550-9bb4-94c6a2ea5fbb.

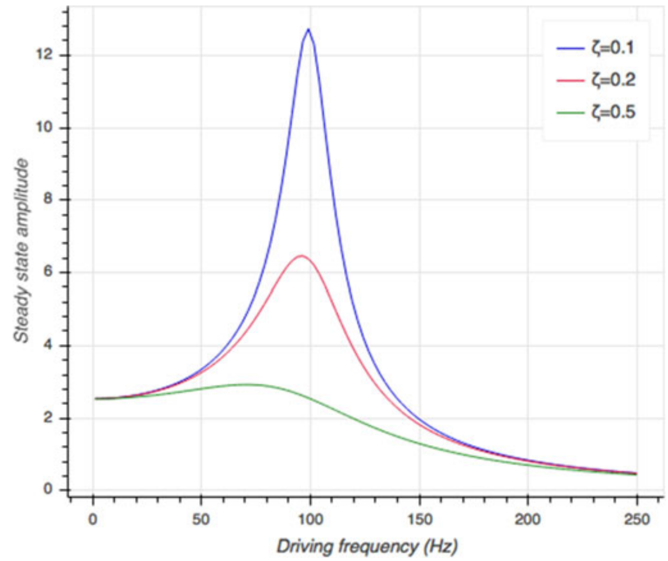

Fig. 1. Variation of steady-state oscillation amplitude with driving frequency for an oscillator with undamped frequency of 100 Hz, a driving amplitude $A = 10^6$, and various damping constants.

A. Illustrative Example

For the purposes of illustration, consider a damped harmonic oscillator driven by a sinusoidal signal

$$\frac{d^2x}{dt^2} + 2\zeta\omega_0\frac{dx}{dt} + \omega_0^2 x = A\sin(\omega t) \tag{1}$$

where $x(t)$ is the quantity of interest, ζ is a damping constant, ω_0 is an undamped resonant frequency, and $A\sin(\omega t)$ is a driving sinusoidal signal. The steady-state solution of (1) is given as follows:

$$\begin{aligned} x(t) &= \frac{A}{Z_m\omega}\sin(\omega t + \phi) \\ Z_m &= \sqrt{(2\omega_0\zeta)^2 + \frac{(\omega^2-\omega_0^2)^2}{\omega^2}} \\ \phi &= \arctan\left(\frac{2\omega\omega_0\zeta}{\omega^2-\omega_0^2}\right). \end{aligned} \tag{2}$$

Fig. 1 shows the steady-state amplitude of a 100-Hz oscillator as a function of driving frequency. We see that the oscillator exhibits a strong resonance for small values of ζ and the resonance frequency is slightly lower than the driving frequency of 100 Hz.

Readers who wish to see plots for a different value of the undamped resonant frequency or the driving amplitude can change it in the online compute capsule easily.

IV. Conclusion

Sharing code and data has many potential benefits for authors, readers, and the ocean engineering research community as a whole. The IEEE has partnered with Code Ocean to make code sharing easy for authors who wish to do so. The code may be written in any language that the authors prefer, and they may generate all or part of the results in the paper. Modern tools, such as Jupyter notebooks, enable authors to combine the manuscript of a paper and the code necessary to generate the results in the paper as a single document. Sharing these notebooks is a powerful way to publish a *dynamic* paper that allows an interested reader to explore the results in the paper in a much deeper way than previously possible.

While technologies and tools needed to publish dynamic papers with shared code and data are now available, they are still far from maturity. Early adopters of the technology are welcome to submit papers to the journal using these tools. I also welcome inputs and suggestions from readers, authors, editors, and reviewers on how to improve the journal's manuscript management process to serve such papers better.

MANDAR CHITRE
Editor-in-Chief

REFERENCES

[1] M. Baker, "Reproducibility crisis?," *Nature*, vol. 533, pp. 452–454, May 26, 2016.
[2] B. A. Nosek *et al.*, "Promoting an open research culture," *Science*, vol. 348, no. 6242, pp. 1422–1425, 2015.
[3] M. B. Porter and Y.-C. Liu, "Finite-element ray tracing," *Theor. Comput. Acoust.*, vol. 2, pp. 947–956, 1994.
[4] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 701–717, Oct. 2013.
[5] P. van Walree, R. Otnes, and T. Jenserud, "Watermark: A realistic benchmark for underwater acoustic modems," in *Proc. IEEE 3rd Underwater Commun. Netw. Conf.*, Aug. 2016, pp. 1–4.
[6] D. Goldstein, Introducing code ocean in IEEE Xplore, *Institute*, Jun. 2017.