

# A New Kind of Article for Reproducible Research in Intelligent Robotics

By Fabio Bonsignorio

The reproducibility of experimental results is a key characteristic of the scientific method. Despite that, in robotics and artificial intelligence (AI)—maybe for good reason—replicating experiments in many cases has, so far, been limited or outright lacking. This fact hampers both research progress and results exploitation [2], [10] and becomes even more relevant when new editorial initiatives, such as [14], increasingly regard (intelligent) robotics as a science.

Reporting practices and formats are a key issue if we want to have reproduc-

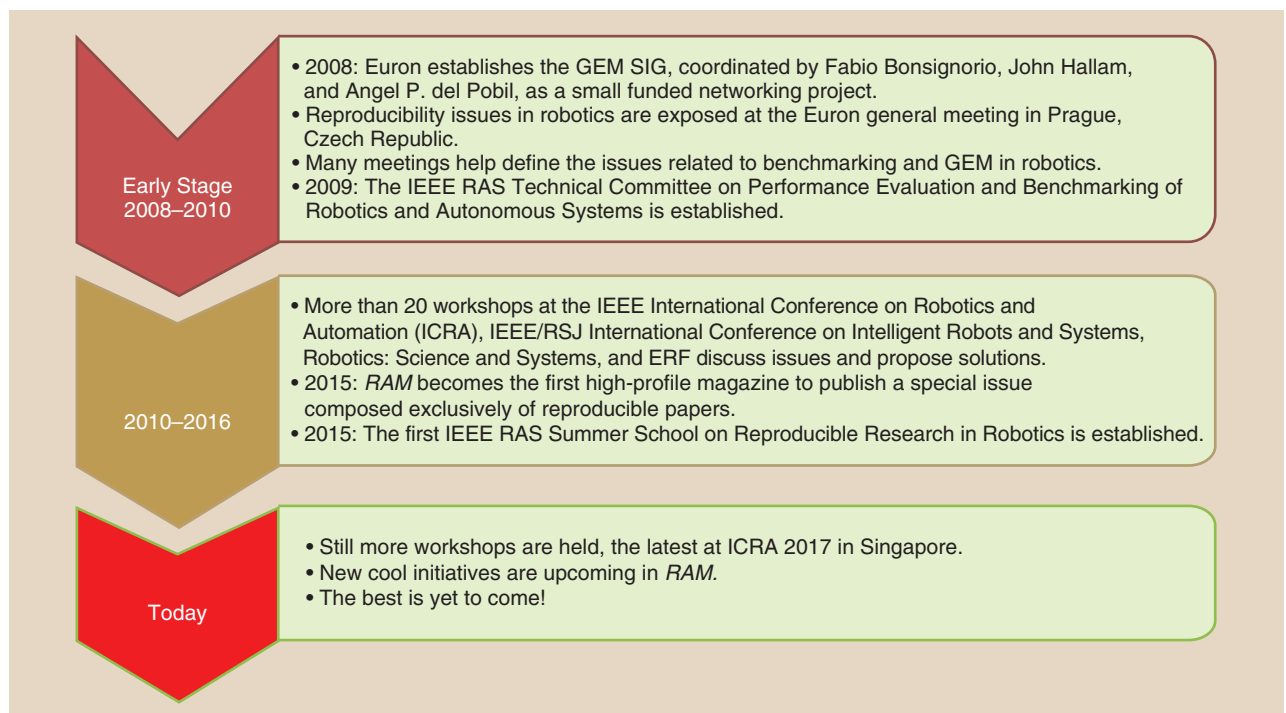
ible robotics and AI papers. After years of discussions in a long series of workshops [9] (Figure 1), the time is ripe for addressing this issue, and we are doing it! The first-ever special issue of a high-level, reputable robotics publication claiming the reproducibility of the published results was in this magazine in September 2015 [9] (Figure 2).

Reproducibility is now a priority for the IEEE, as shown by the fact that the organization recently decided to integrate the CodeOcean platform [15] in the websites of several magazines and journals. And we are going to do the same.

In the meantime, we are in the middle of what has been dubbed a *reproducibility crisis* hitting well-established scientific fields ranging from medicine to psychol-

ogy [3]–[5], [13]. For example, a recent study [11] discovered that only about a third of psychology papers are reproducible. The situation is better in cancer research [12] but is still not optimal. However, the situation in robotics and AI is different. While, in other disciplines, a shared methodology for performing experiments has been in place for a long time and the problems might come from organizational, societal, and sometimes ethical causes, in robotics, the problems are of a methodological and even epistemological nature [9, pp. 32–35]. In the September 2015 *IEEE Robotics and Automation Magazine (RAM)* special issue [9], we gave authors a large degree of latitude in terms of how to define reproducibility and good reporting

Digital Object Identifier 10.1109/MRA.2017.2722918  
Date of publication: 13 September 2017



**Figure 1.** The robotics community has been aware of replicability issues for years. GEM SIG: Good Experimental Methodology Special Interest Group; ERF: European Robotics Forum.

formats. We obtained a large number of high-quality submissions, ranging in topic from manipulation to wearable systems, marine robotics to bipedal locomotion. The success of the special issue showed that the community is ready for this. We are now defining some rules and have designed a full publishing process to move from an initial pioneering effort to a more organized process.

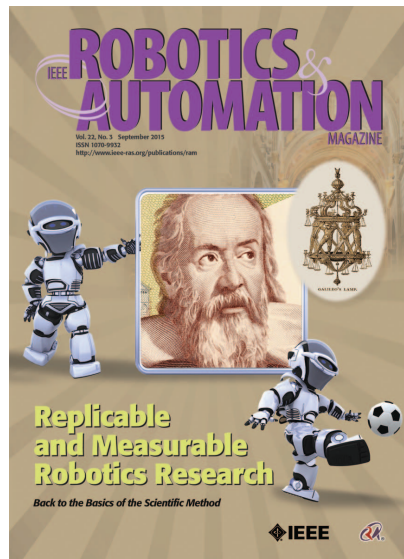
### The RAM R-Article

We have decided to introduce a new kind of article in the magazine, an article specifically designed to be reproducible, and we are calling it the *R-article*.

The authors will have to share whatever they believe is necessary and sufficient to reproduce their results and confirm that in the cover letter. Of course, since experimental outcomes are produced in statistical terms, in most cases it would be almost impossible to achieve exactly the same results, i.e., the same inputs and the same results. Moreover, the description part of the report will at least have to comply with the criteria outlined, e.g., in [1], and summarized in Figure 3.

This new kind of article should include, at least,

- a description, i.e., an almost ordinary journal article with text, figures, and multimedia, complying at a minimum with the good experimental methodology (GEM) guidelines
- data sets stored in CodeOcean, and all the necessary information tests that have been run; only in well-justified special cases will we accept different sharing platforms; videos of experiments should not be considered as data sets but as multimedia



**Figure 2.** The September 2015 RAM special issue cover.

attachments if they are not actually part of the data elaborated by some algorithm in the robot software

- complete code identifiers and preferably downloadable code (executable files may be enough) on CodeOcean—the algorithms as they have been coded and all necessary library and middleware information; also in this case, only in well-justified special cases will we accept different sharing platforms
- a hardware (HW) description or HW identifier (we need to know the physical details of the robots to be able to repeat robotics experiments); by HW description, we mean the full design (at the level of detail necessary for replication) of the robot or robots used in the experiment.

It is important that the code and HW are well identified to be able to replicate experiments. This might be

one of the main bottlenecks for replication of robotics experiments.

### The R-Article Life Cycle

As we aim to foster the spread of a reproducibility culture in our community, the introduction of the R-article is not enough. We will provide a venue for publishing reproduced experiments (this could be a nice opportunity for Ph.D. students, but not only for them). It will be possible to publish a short article about the results replication of an R-article—what we might call an *r-article*. Such articles will be peer reviewed like any other RAM article and will undergo a quick data and code consistency check. Similarly, the authors of the original R-article will be able to submit, again, in the form of a short peer-reviewed article, a reply to the authors of the *r-article*, again, with a data and code consistency check. We will call this a *Reply article*.

The outcome will be a two-stage, high-quality review process. The first stage will be little more than the already very rigorous RAM review process. The second, more accurate stage will be the reproduction of the results by the community. We think the process will improve over time as we adapt and expand the data and code-sharing platform, integrate other tools, and refine the procedure. As the process is implemented, we will learn from experience—and, of course, comments and suggestions are welcome. Feel free to contact me at [fabio.bonsignorio@santannapisa.it](mailto:fabio.bonsignorio@santannapisa.it) or the editor-in-chief at [ieerameic@gmail.com](mailto:ieerameic@gmail.com) if you need any further clarification.

### Call for R-Articles

From this issue's publication date, we will begin to accept R-articles for RAM. We will start to publish them as soon as the first accepted R-article comes out of the review pipeline. We are proud to say that we are the first top-tier robotics publishing venue to do this, and we sincerely hope that many others will follow. It will be a learning process for sure, but we begin on a solid foundation of wide discussion and strong efforts in the community in the past years. Please help us to make history! Let's start!

- 1) Is it an experimental paper?
- 2) Are the system assumptions/hypotheses clear?
- 3) Are the evaluation criteria spelled out explicitly?
- 4) What is being measured and how?
- 5) Do the methods and measurements match the criteria?
- 6) Is there enough information to reproduce the work?
- 7) Do the results obtained give a fair and realistic picture of the system being studied?
- 8) Are the drawn conclusions precise and valid?

**Figure 3.** The Euron GEM guidelines in a nutshell [1]. The guidelines were meant as an adaptation of the basics of the scientific method to robotics and AI.

## References

- [1] F. Bonsignorio, J. Hallam, A. P. del Pobil, Eds. (2008). GEM guidelines: Euron GEM Sig report. [Online]. Available: <http://www.heronrobots.com/EuronGEMSig/downloads/GemSigGuidelinesBeta.pdf>
- [2] F. Bonsignorio, A. P. del Pobil, and E. Messina, "Fostering progress in performance evaluation and benchmarking of robotic and automation systems," *IEEE Robot. Autom. Mag.*, vol. 21, no. 1, pp. 22–25, 2015.
- [3] Challenges in irreproducible research. *Nature* (special). [Online]. Available: <http://www.nature.com/nature/focus/reproducibility/>
- [4] (2013, Oct. 21). How science goes wrong: Scientific research has changed the world. Now it needs to change itself. *The Economist*. [Online]. Available: <http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong>
- [5] (2013). Trouble at the lab: Scientists like to think of science as self-correcting. To an alarming degree, it is not. *The Economist*. [Online]. Available: <https://www.economist.com/news/briefing/21588057-scientists-think-science-self-correcting-alarming-degree-it-not-trouble>
- [6] J. Claerbout and M. Karrenbach, "Electronic documents give reproducible research a new meaning," in *Proc. 62nd Annu. Int. Meeting Soc. Exploration Geophysicists, Expanded Abstracts*, vol. 92, 1992, pp. 601–604.
- [7] F. Bonsignorio, J. Hallam, A. P. del Pobil, "Defining the requisites of a replicable robotics experiment," in *Robotics: Science and Systems 2009 Workshop on Good Experimental Methodologies in Robotics*. [Online]. Available: <http://www.heronrobots.com/EuronGEMSig/downloads/Seattle/RSS09GEMPaperBonsiLast.pdf>
- [8] F. Amigoni, M. Reggiani, and V. Schiaffonati, "An insightful comparison between experiments in mobile robotics and in science," *Autonomous Robots*, vol. 27, no. 4, pp. 313–325, 2009.
- [9] *Special Issue on Replicable and Measurable Robotics Research: Back to The Basics of the Scientific Method*, F. Bonsignorio and A. P. del Pobil, Eds. *IEEE Robot. Autom. Mag.*, vol. 22, no. 3, 2015.
- [10] E. Guglielmelli, "Research reproducibility and performance evaluation for dependable robots," *IEEE Robot. Autom. Mag.* vol. 22, no. 3, p. 4, 2015.
- [11] Open Science Collaboration, "Estimating the reproducibility of psychological science," *Science*, vol. 349, no. 6251, pp. 943–952, 2015.
- [12] B. A. Nosek and T. M. Errington, "Reproducibility in cancer biology: Making sense of replications," *eLife*, vol. 6, p. e23383, Jan. 2017.
- [13] J. Wilson. (2017, June 14). Promoting reproducibility by emphasizing reporting: PLoS ONE's approach. *PLoS Blogs*. [Online]. Available: <http://blogs.plos.org/everyone/2017/06/14/promoting-reproducibility/>
- [14] G.-Z. Yang, J. Bellingham, H. Choset, P. Dario, P. Fischer, T. Fukuda, N. Jacobstein, B. Nelson, M. Veloso, and J. Berg, "Science for robotics and robotics for science," *Science Robotics*, vol. 1, no. 1, p. eaal2099, 2016.
- [15] CodeOcean. Discover and run scientific code. [Online]. Available: <http://www.codeocean.com>



# IEEE connects you to a universe of information!

As the world's largest professional association dedicated to advancing technological innovation and excellence for the benefit of humanity, the IEEE and its Members inspire a global community through its highly cited publications, conferences, technology standards, and professional and educational activities.

Visit [www.ieee.org](http://www.ieee.org).

Publications / IEEE Xplore® / Standards / Membership / Conferences / Education



IEEE