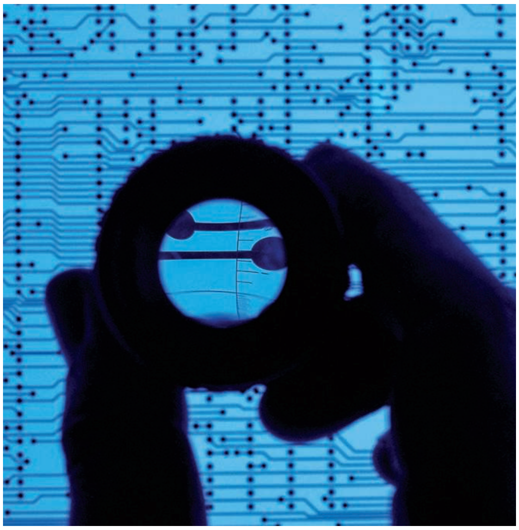
EDITORIAL

Solving reproducibility

The reproducibility problem in science is a familiar issue, not only within the scientific community, but with the general public as well. Recent developments in social psychology (such as fraudulent research by D. Stapel) and cell biology (the Amgen Inc. and Bayer AG reports on how rarely they could reproduce published results) have become widely known. Nearly every field is affected, from clinical trials and neuroimaging, to economics and computer science. Obvious solutions include more research on statistical and behavioral fixes for irreproducibility, activism for policy changes, and demanding more preregistration and data sharing from grantees. Two Perspectives in this issue (pp. 1420 and 1422) describe how journals and academic institutions can foster a culture of reproducibility. Transparency is central to improving reproducibility, but it is expensive and time-consuming. What can be done to alleviate those obstacles?



"Transparency is central to improving reproducibility..."

Most scientists aspire to greater transparency, but if being transparent taps into scarce grant money and requires extra work, it is unlikely that scientists will be able to live up to their own cherished values. Thus, one of the most effective ways to promote high-quality science is to create free open-source tools that give scientists easier and cheaper ways to incorporate transparency into their daily workflow: from open lab notebooks, to software that tracks every version of a data set, to dynamic document generation. Moreover, scientists who use open-source software are not locked into proprietary software platforms with unclear monetization plans. If philanthropy or government funds new tools that the open-source community can iterate and improve on, the per-dollar return on investment can far exceed the costs.

Infrastructural tools are now available, or in development, that should help to catalyze a change in scientific transparency. One example is the Open Science Framework (OSF), a free and open-source software platform for managing scientific workflow (supported by the Laura and John Arnold Foundation in partnership with the Center for Open Science). Among its many features, this platform can enable scientists to easily track the history of all versions of every document or data set and the exact contributions made by each team member. All project materials can be given persistent identifiers, and the tracking of provenance allows any subsequent research project to give proper credit to the original. Projects using this platform include the Shared Access Research Ecosystem project of the Association of Research Libraries and its partners. This project endeavors to connect scholarly metadata and allow the identification of various elements of a research project, such as grant proposals, journal articles, and data repository information.

Open-source platform innovations are growing. Other examples include the iPython project (supported largely by the Alfred P. Sloan Foundation), which offers a web-based computing notebook for users to create documents such as code, computational results, and narrative explanations. Although originally developed around the Python language, the project has expanded to cover other languages (such as R) under the banner of Project Jupyter. The Galaxy Project (funded by the U.S. National Science Foundation and others) provides a web-based platform for "data-intensive" biomedical research; and for managing bioinformatics and phylogenetic research on plants, there is the iPlant Collaborative.

The scientific community cannot depend entirely on volunteers or the private market to develop free platforms that address specialized scientific needs and encourage greater reproducibility. Ultimately, the infrastructure supporting science is a public good, just like the knowledge it produces. By supporting such infrastructure, the public and philanthropic sectors can make it painless for researchers to live up to their own values of openness and rigor.

– **Stuart Buck**

Stuart Buck is vice president of research integrity at the Laura and John Arnold Foundation, Houston, TX. E-mail: sbuck@arnoldfoundation.org

10.1126/science.aac8041

Science

Solving reproducibility

Stuart Buck

Science **348** (6242), 1403.
DOI: 10.1126/science.aac8041