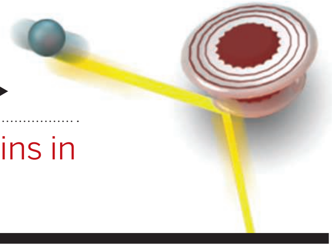
Turning the tap of electron flow p. 1026 ▸

From syngas to olefins in one step p. 1030



PERSPECTIVES

ILLUSTRATION: D. BONAZZI/@SALZMANART

RESEARCH INTEGRITY

Liberating field science samples and data

Promote reproducibility by moving beyond "available upon request"

By **Marcia McNutt,[1]* Kerstin Lehnert,[2] Brooks Hanson,[3] Brian A. Nosek,[4] Aaron M. Ellison,[5] John Leslie King[6]**

Transparency and reproducibility enhance the integrity of research results for scientific and public uses and empower novel research applications. Access to data, samples, methods, and reagents used to conduct research and analysis, as well as to the code used to analyze and process data and samples, is a fundamental requirement for transparency and reproducibility. The field sciences (e.g., geology, ecology, and archaeology), where each study is temporally (and often spatially) unique, provide exemplars for the importance of preserving data and samples for further analysis. Yet field sciences, if they even address such access, commonly do so by simply noting "data and samples available upon request." They lag behind some laboratory sciences in making data and samples available to the broader research community. It is time for this to change. We discuss cultural, financial, and technical barriers to change and ways in which funders, publishers, scientific societies, and others are responding.

POLICY

Repeating a study from start to finish using new samples and equivalent procedures under identical conditions is the ideal. This may be practical in laboratory sciences but is rarely possible in field sciences. Objects of study might be ephemeral (Superstorm Sandy), exceptionally rare (*Dreadnoughtus*), or forever changing (succession in a forest or how climate affects a prairie ecosystem). Nevertheless, transparency and reproducibility have substantial value for field sciences. Independent analysis of original data can uncover statistical or coding errors, data selection bias, or problems with observations that are "too good to be true." Original analyses may be augmented with new techniques to test novel questions. Data and samples can be combined across studies for more precise or generalizable tests [e.g., the Paleobiological Database (*1*) or PetDB database (*2*)].

Such efforts must recognize that motivations for promoting transparency and reproducibility vary by stakeholder. Researchers want to produce knowledge in new directions and to get credit for their contributions. Funders want to see greater value from their investment. Journals want to facilitate reproducible science. Repositories want to support their communities and streamline data flow.

FUNDING, PUBLISHING, AND CULTURE CHANGE. Transparency and reproducibility in scientific research require investment. Quality control is costly, and even large projects can have difficulty curating data to make them useful for others. For example, many data sets in ecology and evolution publications are incomplete or inaccessible (*3*). Funders that support domain-specific data and sample repositories do so because they advance science by facilitating preservation and reuse, as well as support data professionals who help with collection, management, and curation.

U.S. and UK funding agencies require publicly supported researchers to provide data management plans and use open repositories for data and samples. Investigators are often allowed reasonable amounts of time between collection and deposition (e.g., up to 2 years) for quality control or publication priority, but making data and samples available sooner can advance science more rapidly [e.g., (*4*)]. The U.S. National Science Foundation's PASSCAL (Portable Array Seismic Studies of the Continental Lithosphere) and OOI (Ocean Observatory Initiative) projects are examples of programs that are acclimating communities to more progressive data-sharing policies, demonstrating research gains, and undermining arguments against open data.

Initiatives such as OpenAIRE and the Center for Open Science are advancing principles and guidelines for transparent research and publication. More than 500 journals have indicated intent to review the TOP (Transparency and Openness Promotion) (*5*) publication guidelines for potential implementation. However, journals that require that data and samples be deposited upon publication may suffer if journal staff lack resources to verify deposition, evaluate metadata, or check code accessibility of code or other materials.

"All should credit data creators and accelerate recognition of the value of data in the...system."

Funding agencies and journals can guide expectations and set requirements, but top-down mandates alone are unlikely to foster needed cultural changes in scientific communities. Research culture prioritizes publications, innovation, and insight, which puts data stewardship and reuse far down the list. Data professionals in large projects too often are invisible team members. To change this, bottom-up approaches are needed to earn community buy-in and effect culture change to recognize researchers who develop and curate original data (*6*): Principal investigators can lead by example; universities can offer education in data and samples management; and scholarly societies can recognize excellence in data and samples stewardship in their awards, selection of fellows, and leadership. Development of data journals with citable output can provide data experts acknowledged leadership roles.

REPOSITORIES, CITATIONS, AND CURATION. Although general-purpose repositories are useful for data that do not fit easily within a specific domain, the value of preserved data is best maximized by discipline-specific repositories. They can provide services informed by community priorities, which can promote cultural change.

Repositories can improve quality control and compliance, but progress will be faster if repositories move beyond curation and provide value-added services, such as certifying to journals completion of deposition requirements. Because deposition alone does not guarantee that data or samples are discoverable or usable, machine-readable, quality-controlled, public metadata would help researchers find, understand, and use the resources. Automated services can promote cost-effective and data-conscious research cultures, while providing incentives to data collectors by demonstrating impact through metrics of views, downloads, and data and sample uses. The Open Science Framework (http://osf.io/) connects repositories so authors can find domain-specific repositories and provides links to general-purpose repositories for unusual classes of data (*7*).

Repositories can expand assignments of digital object identifiers (DOIs) to data sets to aid citation. International initiatives, including DataCite (www.datacite.org) and the IGSN e.V. (www.igsn.org, International Geo Samples Numbers), develop and promote common conventions for unambiguous identification and citation of data and samples to enhance usability, to guard against loss, and to provide credit to creators. Publishers can adopt resolutions such as the Joint Declaration of Data Citation Principles (*8*) that specify data citation in the references. Journals can publish descriptions of data sets and methods. All should credit data creators and accelerate recognition of the value of data in the scientific reward system.

Despite many efforts, there remains widespread disagreement regarding data and sample availability and metadata, as well as uneven sample deposition across the field sciences. Journals might use DOIs for data or accession numbers for curated samples (e.g., IGSNs), but these are not used routinely. Journals that ensure precise descriptions of variables, data, or sample provenance and that provide details concerning collecting permits and requirements or restrictions for reuse can help. However, such information must come from those closest to the data and samples—the researchers themselves. Scientists and technical staff must know how to create such information and be supported by funding agencies to do so. Repositories could provide user-friendly software and training to field teams to overcome resistance to sharing data and samples.

Scientific societies can work with stakeholders to set guidelines for provenance of metadata; access; and security or ethical restrictions (e.g., for protecting human subjects or rare and endangered samples); and cooperate with all journals, not just those they own or sponsor. Community partnerships [e.g., COPDESS (Coalition for Publishing Data in the Earth and Space Sci-

[1]Editor-in-Chief, Science, Washington, DC 20005 USA. [2]Columbia University, Palisades, NY 10964 USA. [3]American Geophysical Union, Washington, DC 20009 USA. [4]University of Virginia and Center for Open Science, Charlottesville, VA 22903 USA. [5]Harvard University, Petersham, MA 01366 USA. [6]University of Michigan, Ann Arbor, MI 48109 USA. *E-mail: mmcnutt@aaas.org

ences) (*9*)] can provide discipline-specific metadata and effective quality control with secure badging for journal verification. The Research Data Alliance (RDA) is developing approaches and infrastructure for the publishing community.

Finally, not every sample can be saved. Museums and other special-purpose repositories (e.g., ice-core labs) face resource and space limitations. Curators must decide what to keep. Samples supporting peer-reviewed publications should have priority. Digitized samples and collection information or other metadata will facilitate remote examination. Digital catalogs can provide persistent access to metadata on samples used in publications. These should include information on access linked to publications via resolvable unique identifiers such as the IGSN. The System for Earth Sample Registration (SESAR), iDigBio, and Cyverse provide examples of metadata profiles.

By working together, stakeholders can create a virtuous cycle of increasing data and sample accessibility. The days when scientists held on to samples and data hoping to squeeze out one more publication are ending. Sharing can be more productive than hoarding when researchers get credit for use of their data or samples. The citation advantage for papers with open data (*10*) suggests that stakeholders help themselves by promoting transparency and reproducibility. ■

REFERENCES AND NOTES

1. Paleobiological Database, https://paleobiodb.org/#/.
2. D. B. Pet, http://www.earthchem.org/petdb.
3. D. G. Roche, L. E. B. Kruuk, R. Lanfear, S. A. Binning, *PLOS Biol.* **13**, e1002295 (2015).
4. D. Lindenmayer, G. E. Likens, *Ecol. Soc. Am. Bull.* **94**, 338 (2013).
5. B. A. Nosek, G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. A. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. L. Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, T. Yarkoni, Promoting an open research culture. Science 348, 1422–1425 (2015).
6. J. A. Mills *et al.*, *Trends Ecol. Evol.* **30**, 581 (2015).
7. S. Buck, *Science* **348**, 1403 (2015).
8. Data Citation Synthesis Group, Joint Declaration of Data Citation Principles, M. Martone, Ed. (FORCE11, San Diego CA; 2014); www.force11.org/group/joint-declaration-data-citation-principles-final.
9. B. Hanson, K. Lehnert, J. Cutcher-Gershenfeld, *Eos* **96**, 10.1029/2015EO022207 (2015).
10. H. A. Piwowar, T. J. Vision, *PeerJ.* **1**, e175 (2013).

ACKNOWLEDGMENTS

These recommendations resulted from a workshop supported by the Laura and John Arnold Foundation. Attendees who helped shape these recommendations include P. G. Brewer, B. Clement, C. Duke, D. Erwin, L. Ferguson, R. Freckleton, A. Friedlander, T. Gardner, B. Grocholski, L. Hale, S. Hageman, J. Kratz, W. J. Kress, M. Morovati, M. Parsons, K. Ram, M. Ramamurthy, H. J. Smith, A. Sugden, M. Uhen, J. VanDecar, T. Vision, N. Wigginton, and R. Wakimoto.

10.1126/science.aad7048

PHYSICS

Electrons go with the flow in exotic material systems

Electronic hydrodynamic flow—making electrons flow like a fluid—has been observed

By **Jan Zaanen**

Turn a switch and the light goes on. The layman's perception is that this is like opening a tap so that the water starts running. But this analogy is misleading. The flow of water is governed by the theory of hydrodynamics, whereby the behavior of the fluid does not require knowledge of the motions of individual molecules. Electrical currents in solids, however, are formed from electrons. In metals, these do not collide with each other, but they do scatter from lattice imperfections. The resulting "Knudsen flow" of electrons is reminiscent of the avalanche of balls cascading through a dense forest of pins, as in a Pachinko machine. On pages 1058, 1055, and 1061 of this issue, evidence is presented that electrons can actually yield to the laws of hydrodynamics (*1*–*3*). What is additionally surprising is that these observations are in agreement with mathematical techniques borrowed from string theory (*4*). These techniques have been applied to describe strongly interacting forms of quantum matter, predicting that they should exhibit hydrodynamic flows (*5*).

The experiments have been made possible by progress in new materials and nanofabrication techniques. Two of the papers report on complementary aspects of the electron hydrodynamics in graphene (*1*, *2*). The third paper deals with an oxide material that exhibits highly surprising transport properties. By confining the electrical currents to nanoscale pipes, hydrodynamic flow is demonstrated (*3*).

The flow of substances is governed by simple conservation laws: Matter, energy, and electrical charge are naturally conserved, while in a perfectly homogeneous space the velocity of an aggregate of matter is not changing either; that is, momentum is also conserved. A classical fluid, such as water, looks like a dense traffic of colliding water molecules exchanging momentum at a very high rate. However, their combined momentum does not change unless the space they are moving in is made inhomogeneous by, for example, putting the water in a pipe such that the overall momentum relaxes and the kinetic energy turns into heat. Electrons in solids, however, move in a background of static ions, breaking this translational invariance, and imperfections occur even in the most perfect periodic crystals. It is now a matter of numbers. Could it be the case that an individual electron can lose its momentum because of scattering from the ionic disorder before it meets another electron (Knudsen flow) (see the figure, left panel), or will the electron fluid equilibrate first through many electron-electron collisions without noticing the imperfections (hydrodynamic regime) (right panel)?

To better understand the situation, we must invoke quantum physics. On the microscopic scale, electrons in solids are strongly interacting, but quantum many-body systems submit to the principle of

> ***"...hydrodynamic flows are much richer than the diffusive currents that have been the traditional mainstay of solid-state electronics."***

renormalization, in which the electrons' behavior is dependent on the scale at which the system is observed. In conventional metals, the renormalized electrons increasingly ignore each other as the energy decreases. On the macroscopic scale, the electrons behave like the individual balls of the Pachinko machine. However, it might well happen that the effects of the interactions increase as the energy decreases (giving rise to a complex quantum soup), and until now we did not have the mathematical tools to describe transport in the resulting highly collective quantum state. Recently, it has been shown that the mathematical machinery developed by string theorists can

Instituut-Lorentz for Theoretical Physics, Leiden University, Leiden, Netherlands. E-mail: jan@lorentz.leidenuniv.nl