Journal Replication Policy Use Case
Summary Description

Reproducibility is a core principle of scientific research. And there is now widespread recognition that many of the findings published in peer-reviewed journals, especially those in the medical, health, behavioral and social sciences, do not describe methods and data in sufficient detail to support verification and replication. [1,2,3] Simultaneously, pressure has increased from various sources to make data more openly available – and principles of open data transparency are gaining more recognition. [4,5] Consequently, professional societies, editorial boards and individual editors are increasingly adopting more detailed and actionable data citation, sharing, and archiving policies. [6]

The characteristics of data used as evidence for articles submitted even to single journal varies widely. Such data includes well-known, previously published data, such as that from official statistics and community-managed databases (e.g. Genbank, Flybase); data obtained from third parties; data collected by the authors of the submission generally for their research; and derivative datasets prepared specifically for a publication which may integrate, correct, annotate and recode data from multiple sources.

Journals policies may allow this data to be disseminated in a variety of ways: through structured "data citation" to data already persistently and publicly available; direct attachment to an article as "supplemental materials"; or deposition prior to publication within a journal-specific "replication archive", a community or field-specific archive, or (rarely) within university institutional repository.

Journal editors and publishers are concerned with a range of disclosure threats: the main threats include identification of an individual subject; revealing sensitive locations (such as archaeological sites); and disclosure of proprietary information obtained from third parties.

Editors and publishers use a number of disclosure limitation mechanisms: the most common of approach is to require deposited data to be deidentified; or to be deposited in an external archive that vets access through data use agreements. In many cases, however, sensitive data are simply exempted from the policy.

The emerging challenges in this area are related to the variety of data and the limited resources available for vetting it. Journal editors and publishers are

particularly concerned with developing policies that are strong enough to strengthen replicability; can be implied without intense case-specific scrutiny; and recognize common disclosure threats while still permitting researchers who are able to provide the necessary protections access to the data.

References:

[1] B. D. McCullough, Kerry Anne McGeary and Teresa D. Harrison "Do Economics Journal Archives Promote Replicable Research?" Canadian Journal of Economics 41(4), 1406-1420, 2008.
[2] Vasilevsky et al. (2013) On the reproducibility of science: unique identification of research resources in the biomedical literature. PeerJ 1:e148 http://dx.doi.org/10.7717/peerj.148
[3] Ioannidis, John PA, et al. "Repeatability of published microarray gene expression analyses." Nature genetics 41.2 (2008): 149-155.
[4] Uhlir, (editor), For Attribution: Developing Data Attribution and Citation Practices and Standards. National Academy of Sciences. http://sites.nationalacademies.org/PGA/brdi/PGA_080121
[5] http://openeconomics.net/principles/
[6] http://jordproject.wordpress.com/about/

Examples

1. Data Sharing Systems for Open Access Journals

The OJS-DVN project aims to create an easy-to-use flexible workflow for data publication in open access journals through integrating the Public Knowledge Project's *Open Journal System* (OJS) and Harvard *Dataverse Network* (DVN) systems.

The project is extending both system to support enforcement of data deposition and citation policies, to facilitate peer-review workflows for submitted data, and to enable automatic deposition of replication data into designated replication repositories.

See:
http://projects.iq.harvard.edu/ojs-dvn/blog/welcome-ojs-dvn-project-blog-and-website

2. American Political Science Association Data Access and Research Transparency [DART] Policy Initiative

The American Political Science Association, which is the largest and highest profile professional association in the field, and publisher of the discipline's flagship journal, is in middle stages of developing new policies for data sharing and citation.

These include explicit discussion of sharing of proprietary and confidential data. The current recommendations include using deidentification, data use agreements, and consent language. Confidentiality and proprietary concerns may still be used to withhold data from all other researchers under these policies, subject to documentation of the rationale.

See:
http://www.apsanet.org/media/PDFs/QUANT%20DART%20Guidelines%20for%20Data%20Access%20and%20Research%20Transparency%20July%2028%202013%20Ver%207.pdf

http://www.apsanet.org/media/PDFs/QUAL%20DART%20Guidelines%20for%20Data%20Access%20and%20Research%20Transparency%20August%207%202013.pdf