

Data Scientists, Data Management and Data Policy

Sylvia Spengler

National Science Foundation, 4201 Wilson Boulevard,
Arlington, Virginia, USA
sspengle@nsf.gov

US science agencies have or will soon have a requirement that externally funded projects have “data management plans.” Projects with a large budget or a tradition of data access and repositories do not see the impact as significant. However, the impact of the requirement can be particularly challenging for single investigators and small collaborations, especially in multidisciplinary research. These data represent what is known as Dark Data (Heidorn, 2008) in the long tail of science, where the data sets may be relatively small and the funding and expertise for handling also small. But just developing tools or putting computer scientists with the investigators is not sufficient.

The challenges exist at multiple levels: in the data themselves, where many different formats and systems are used; social and cultural issues of reward, recognition, rights and protections; in the persistent issues of investment and valuation. This panel will explore the challenges in this shotgun wedding, illustrate some success stories and explore ways that agencies can support the necessary changes.

Reference

Heidorn, B.P.: Shedding light on the dark data in the long tail of science. *Library Trends* 57, 280–289 (2008)