Members Forum

Research environments are becoming increasingly more complex in the 21st century, with a concomitant upsurge in the amount of data produced. Along with collecting and managing large, complex data sets, graduate students must be aware of the most common reasons for data loss and the strategies to prevent them.

The State of Scientific Data Management

In November 2012, Doucette and Fyfe (2013) conducted an online survey of more than 350 master’s and doctoral students in six subject areas (i.e., Geography, Psychology, Sociology, Chemistry, Physics, and Earth Sciences) from nine different research universities in Canada. This work assessed graduate students’ behavior, attitudes, and education related to managing research data. Some of the most remarkable outcomes from the survey were as follows:

- 14% had “re-collected data that had been previously collected because [they] could not find or open the file.”
- 17% had “lost a file and been unable to re-collect the data.”
- 40% were unsure, disagreed, or strongly disagreed with the statement, “I have provided enough documentation that a research peer or future grad student could use my data.”

The majority of students agreed or strongly agreed that:

(a) there is a value in reusing or repurposing their research data in the future (74%); (b) management of research data is important for a research group (83%); and (c) they were confident in their abilities to manage research data (90%). However, almost 38% of the students did not have written or verbal policies related to research data management (RDM) within their research group.

- The majority of respondents who had re-collected data or lost a file also agreed or strongly agreed with statements related to importance (81%) and abilities (77%) related to RDM.

Though this study did not include st

Data Management

by Martín Battaglia, Wade Thomason, and John Fike

Published online July 6, 2017