Findings from the DIL Interviews: Data Conversion and Interoperability

Skills in this competency may include:

- Is proficient in migrating data from one format to another.
- Understands the risks and potential loss or corruption of information caused by changing data formats.
- Understands the benefits of making data available in standard formats to facilitate downstream use.

Additional skills mentioned by interviewees:

- Understands the advantages of different formats of files.
- Ability to code.

Average Ranking of Importance (5=essential): Faculty=4.13, Students = 4.24

Faculty responses:

In considering this competency, the theme of potential data loss in the conversion process recurred multiple times. Many faculty sensed that students were not considering the potential for loss or corruption when converting files. One faculty made a connection between understanding how data can be manipulated and ensuring the quality of the data. Another saw this as an important skill for students to develop not just for working in his lab but for getting jobs after graduation.

Most faculty reported that competencies with data conversion and interoperability were generally underdeveloped in students. Faculty reported that their students acquire their knowledge and skills in this competency in a variety of ways: classes, peers and experience. One faculty stated that his students need to have more experience with how conversion can affect their data. Another mentioned that students must be aware of what data may be lost during data migration and what open standards for file formats are.

Student responses:

Nearly all of the students reported converting data as a part of their work in the lab, although most did not mention conversion as a distinct stage of the data lifecycle. Conversions ranged from a simple cutting and pasting, to identifying the meaningful elements of the data and extracting them into a usable format. Students responded to questions of data conversion and interoperability by discussing conversion of raw data (i.e. Access files to text files, sensor data to Excel) as well as processed data (i.e. converting images created in gnuplot to GIF or JPEG, converting a figure to a table).

No concerns of data loss during the conversion process were brought up by students. A few students reported checking the data after converting it to ensure that data loss had not occurred.