Introduction

Researchers in all fields use software tools for purposes such as collecting data, analyzing results, creating graphs and charts, and looking up information. However, the vast number of options available leads to inconsistencies in areas such as file types, storage, and output. There is currently no unified way to handle information from a large research project.

We interviewed researchers who are currently working on biomedical-related projects about their use of software in their research. This information went into Professor Stafford's proposal that was sent to the National Institute of Health. Professor Stafford hopes to get funding to create a software infrastructure aimed at helping biomedical researchers reach their goals more quickly and efficiently.

Example Interview Questions for Researchers

- What kind of software do you use in your research?
- How do you obtain this software?
- What kind of software do you use in your research?
- Do you integrate the software or run programs separately?
- What are your chief quality concerns with software you are using?
- What problems do you frequently come across in your software use?
- How do you evaluate software use to be sure your goals will be met?
- What are your chief quality concerns with software you currently use?
- How do you obtain this software?
- Do you integrate the software or run programs separately?
- What are your chief quality concerns with software you currently use?
- How do you evaluate software use to be sure your goals will be met?
- What problems do you frequently come across in your software use?

Researcher Interview: Dr. Peter Castaldi

Dr. Castaldi is currently involved in three active research projects.

1. Locating the Genes that Cause Emphysema
   - Emphysema is a chronic obstructive pulmonary disease (COPD) that is often caused by exposure to toxic chemicals or tobacco smoke. Dr. Castaldi is working on locating the genes that cause this disease. Dr. Castaldi uses a software called Haploview, a menu-driven software program to input data points and find where on the genome certain variations are located.

2. Tracking Medication Skipping Habits of the Elderly
   - Dr. Castaldi’s second project involves tracking the medication skipping habits of people over 75 who have chronic respiratory disease. This involves a database of responses from a survey. The SAS and STAT software programs are used for statistical analysis of the results.

3. Building a Predictive Model for Alpha-1 Anti-Trypsin Deficiency
   - Alpha-1 Anti-Trypsin Deficiency is a genetic disorder caused by defective production of alpha1-antitrypsin, which causes symptoms such as shortness of breath and wheezing. Dr. Castaldi is working on building a model that will predict which people will develop the disease. This project also utilizes the SAS and STAT software.

Typical Problems in Research Software

Time
- Researchers use many different software programs in their projects, picking up new skills as they go. They sometimes rely on outside sources, such as a colleague with programming experience or outside help.

Compatibility
- Researchers receive data from many sources, and accepting all kinds of data types, from Excel spreadsheets to databases, graphs, and text documents, is critical to putting all the pieces of a research project together.

Output
- A research software program should output data in ways that are easily adapted to presentation of results: graphs and charts should be neat and readable, and data from projects like surveys should be in a presentable format.

Efficiency and Consistency
- A program must be able to calculate results relatively quickly and with utmost efficiency, and also obtain consistent results from the same data set.

Simplicity
- Researchers use many different software programs in their projects, picking up new software utilities as they need them. An all-encompassing software utility for researchers might instead use a library of software plugins that are added as needed.

Conclusion

Biomedical researchers currently employ a variety of programs and software-use strategies to reach their goals. There is no unified infrastructure that can be used to integrate different programs’ results or combine data into a single, useful output. Programming must be done by the researchers themselves or by hiring outside help, and various programs are generally run independent of each other. This can result in slow work where software is involved, due to such issues as data incompatibility. Biomedical research projects are multifaceted, and any software utility created to handle all aspects of research projects must contain certain elements that researchers require. Speed, ease of use, compatibility, and efficiency are all such essential factors that the end product must have in order for it to be a reliable research tool.