

A Hands-on Introduction to Literate Programming

Tim Dennis (1), Harrison Dekker (2), and Juliane Schneider (3)

(1) Data Archive - UCLA Library, Orcid 0000-0001-6632-3812, (2) University of Rhode Island, (3) Harvard Catalyst Clinical and Translational Science Center, Orcid 0000-0002-7664-3331

A fundamental challenge for open science is how best to create and share documents containing computational results. Traditional methods involve maintaining the code, generated tables and figures, and text as separate files and manually assembling them into a finished document. As projects grow in complexity, this approach can lead to procedures which are error prone and hard to replicate.

Fortunately, new tools are emerging to address this problem. The workshop will introduce a solution that's gaining popularity in the R community utilizing the freely available RStudio development environment and other open source components. In the workshop we'll demonstrate how to create a "compilable" document containing all the text elements (including bibliography), as well as the code required to create embedded graphs and tables. We'll demonstrate how the process facilitates making revisions when, for example, a reviewer has suggested a revision or when there has been a change in the underlying data. We'll also demonstrate the convenience of integrating version control into the workflow using RStudio's built-in support for git.