What does it take to do reproducible computational science? What stands in our way?

Bill Rider, Sandia National Laboratories

In conducting and documenting computational science research there are a number of distinct steps that are well defined and generally agreed upon. Among these steps would be defining the basic concept (innovation/algorithm/question/application/result), literature review, derivation (and proofs where appropriate), implementation (i.e., coding in a programming/macro language C, C++, Fortran, Python, Matlab, Mathematica, etc.), debugging/testing, calculations, quality assurance work/V&V, writing and ultimate submission with associated peer review. To conduct "reproducible research" requires additional steps that would necessarily allow scrutiny of previously "private" steps in the research such as the details of a derivation, computer implementation, and the quality assurance process. Some of these steps have crept into publishing practice, as is the case with V&V. This additional scrutiny would then invite additional attention to the packaging and automation of steps that might have been heretofore much less formal. This would likely have the immediate impact of improving the manner in which this work in conducted. Moreover, the availability of these details would likely accelerate follow-on work and provide the basis for faster prototyping of extensions.

All of these impacts are generally positive, but must be countered with increased regulation of information for a variety of reasons including security-related concerns; export control/ITAR laws, intellectual property laws, proprietary information and the editorial policy of publications. Each of these provides a barrier of one sort or another for producing reproducible research that outstrips any of the technical challenges. These barriers are distributed and rather unevenly across organizations engaged in computational science research resulting in the specter of creating a culture of the "have's" and the "have not's" in reproducible computational science research.