Victoria Stodden
Yale Law School

Why Copyleft isn't Right for Scientific Code

An increasing proportion of our scientific intellectual output is being recorded in the form of computer software. For scientific results that involve computational or data analysis, often the complete methodology is embedded in the code. This has deep ramifications for the practice of the scientific method as code is not typically shared when findings are published making it practically impossible to replicate the results. For this reason a movement is underway to include data and code along with the publication of scientific discoveries (Gentleman and Lang 2004, Donoho 2009), opening a host of intellectual property issues including those of open licensing (Stodden 2009). In this paper I argue the normative structure in the scientific community, as opposed to that in the open source community, precludes the use of the copyleft or sharelike provision common to many popular open source software licenses. There three major reasons: the scientific ethos precludes directing an independent scientist's creative contribution; copyleft licenses make demands on downstream code, namely that they use the upstream license on the entire library of new code implying that two pieces of code under two different copyleft licenses cannot be mixed; and scientific knowledge is considered a public good, and as such members of society must be free to build on it including scientific collaboration with industrial partners.