University of Colorado-Boulder

Considerations Regarding the Proposed Open Code Policy

Summary:

This paper will focus the following aspects of how the proposed policy may affect the University of Colorado Boulder (University) and the broader researcher community: 1) use of open source code in the current NASA award framework and potential impacts of a new open code policy; 2) compliance and implementation issues; and 3) potential for modification of the current data rights language along with supplemental NASA guidance to promote easier open source use and distribution. The comments and conclusions contained herein are informed by work related to the contracting, compliance, IP and legal facets of University projects.

The following comments are those of the author and do not represent an official opinion or statement of the University or its employees.

Author:

Wiggs, Cody; Research Counsel, Office of University Counsel, University of Colorado

The following is based on the provision of legal and contracting support to multiple University departments in receipt of NASA funding. I believe my repeated involvement in these projects has been prompted by increased researcher acceptance of open source quality and a call for strict university compliance in all operational aspects. While I am obligated to make decisions with solely the interest of the University in mind, I have gotten to work extensively with researchers and have become familiar with their concerns and struggles related to open source utilization under federal awards.

In evaluating possible costs and benefits of the open code policy I will be working from the baseline assumption that anticipated policy would be applied broadly and uniformly to all institutions seeking NASA funding.

1. Use of open source code in the current NASA award framework and potential impacts of a new open code policy

Currently, NASA awards received by University contain mandatory terms or conditions dictating the ownership and distribution of data and computer software created through the project. Said terms can often have the unintended effect of delaying research. While the intention of the terms is often to ensure that the value and knowledge from the funded endeavor are made available to the public, the resultant terms often lack the flexibility to accommodate the use of open source tools or contributions by commercial partners. This is seen most often when the required terms lack a comprehensive framework for projects that wish to utilize a broad range of resources or/and preserve developments made to date.

Researchers are increasingly willing to release resources developed or modified with NASA funds under open source terms. However, current regulations can prevent such a release, or at minimum, create uncertainty among researchers about the ability to release code. For example, recently, a University primary investigator contacted me to ask if he could release some software to the public that he had created under a NASA project. This software was not a deliverable for the project, but the efforts to create the deliverable had yielded pieces of code during the

development of web and data portals that the researcher felt would be useful to the research community at large. As such, the release of the code was fully anticipated by the researcher. Unfortunately, upon further review of the relevant contract, it was found to contain 1852.227-14. The problematic language that we continue to encounter with projects involving the use of open source software or the release under an open source license is as follows; "Alt 4 - *(4)(i)* The Contractor agrees not to assert claim to copyright, publish or release to others any computer software first produced in the performance of this contract unless the Contracting Officer authorizes through a contract modification." The definition of computer software being, "*(i) Computer programs that comprise a series of instructions, rules, routines, or statements, regardless of the media in which recorded, that allow or cause a computer to perform a specific operation or series of operations; and (ii) Recorded information comprising source code listings, design details, algorithms, processes, flow charts, formulas, and related material that would enable the computer program to be produced, created, or compiled."* Having such a broad definition of computer software coupled with only one avenue for compliant release creates issues for the sharing of computer software through either open source release or detailed publication.

I would posit that before turning to a completely open code it may be best to reform the current regulations to accommodate for sharing and publication. Many pieces of the computer software captured under this provision aren't deliverables that NASA will have in its possession to publish. A lack of clarity on what is actually meant to be captured by the DFAR supplement makes it difficult for researchers to operate once a contract is in place. Should the best or most effective software for the project happen to be covered by an open source license that requires release or copyright, the practical use of such software is precluded under NASA contracts containing 1852.227-14[1]. Getting permission to publish or assert copyright to any new code requires the modification of a contract and the research timeline may not support delays for contract negotiations. Since many NASA projects are long term and complicated, it also is highly likely that an engineer along the line will use a piece of open source code which is commonly interpreted to be "unlimited free use" code by many. A lack of flexibility within the provision puts the funded institution in a difficult position. The need to come back mid project and negotiate the inclusion and publication of large pieces of mission critical software is, at best, a large impediment in a time sensitive mission plan. Returning to the example found in the previous paragraph, should a useful byproduct of the funded work be realized there is little incentive to negotiate a modification.

To provide the highest quality work, researchers will attempt to incorporate the highest quality tools. Each tool has different costs associated with the degree of initial quality. As a general rule, when moving up the spectrum from unrestricted open source products to restricted commercial products, the research tool becomes more reliable and capable. A tool at each portion of the spectrum can theoretically accomplish the same end goal but to get to that functionality a cost has to be absorbed somewhere between project initiation and completion or else it's likely that the value of the results will make up the cost differential. While the initial cost of a truly open source tool is low, development costs of said tool may be high, or the cost might be taken out of the potential contribution of the project due to suboptimal resources.

If a NASA funded researcher creates something with high utility from scratch to stay in compliance and within budget, they will want to control the utilization of said resource. Not

[1] 48 CFR 1852.227-14 Rights in Data – General (APR 2015)

being allowed such control will discourage the repeated creation of unique and unhindered resources. Should a broad open code policy be implemented, it should be malleable enough to allow for the creation of the appropriate incentive for the researcher. Due to both the scarceness of funds and increased competition, universities and commercial interests are unwilling to surrender a competitive advantage by contributing valuable resources to a NASA project which could result in the unintended infection of the resource.

Software derives value from its utility compared to competitors, and it maintains this value through a controlled distribution of the source code. Institutional personnel are wary of the potential loss of a resource's value when used in a federally funded project. If the resource owner agrees to the use of such resources in conjunction with research or services funded by the federal government, such use is only allowed at a high cost to the award recipient wishing to use a high value resource.

2. Positive or negative impacts of all future NASA funded science code being required to be OS

Research institutions, in addition to the absorbing the currently unaccountable cost discussed above, will be bear a large upfront cost in training employees to comply with a new open code policy. Resources are stretched thin in all projects, however ensuring compliance with a new open code policy will place an obligation on institutional employers to train employees that can't be recouped through project funding. If the open code policy is applied broadly, each competitor for funding will need to make a representation of the effect that compliance will be ensured. The only way to make such a representation accurately will be through detailed training and compliance plans. Implementing oversight for training and compliance stands to impose a high cost as any new policy or regulation creates uncertainty and requires thorough analysis. Larger institutions may be able to pay this cost and will be able to have assurances that compliance will be maintained. Smaller institutions that aren't repeat participants in this area will most likely not be able to comply with a broad open code policy. As such, there is another cost/benefit analysis that needs to be made by such institutions. According to recent data, regulatory compliance (not including costs associated with hospital operation) accounts for up to 11 percent of the operating expenses for an institute of higher education. In addition to the added expense, between 4 and 15 percent of a faculty members' time is spent focusing on compliance issues. These costs generally increase for non-tier 1 research institutions.[2]

In general, open source software and associated licenses are not a familiar issue for general counsel, and some organizations seeking NASA funding may not have the resources or the knowledge to ensure compliance with an open code policy. A lack of understanding or man power may lead to unknowing acceptance of all attached terms without a realization of the associated requirements. This situation results in a detriment to a more familiar or risk averse participant in competing for funding. Accepting terms without the necessary compliance evaluation leads to a participant moving quickly and unintentionally making unintentional misrepresentations. Institutions who are more familiar with the funding process and implications of agreeing to certain award terms will take longer to evaluate the ability to comply and will be at an inherent disadvantage in the award process. Speed and ease of negotiations can often influence the distribution of funds. Even for sophisticated institutions ensuring compliance with

[2] The Cost of Federal Regulatory Compliance in Higher Education: A Multi-Institutional Study (Boston Consulting Group), https://news.vanderbilt.edu/files/Regulatory-Compliance-Report-Final.pdf

policy related to unfamiliar and difficult concepts such as open source software will lead to increased delay and difficult negotiations. Should funding opportunities be missed due to attempts to ensure compliance with the open code, researchers may choose research strategies and resources with a strong preference to guarantee compliance upfront. More cautious institutions may be motivated to seek exceptions, rework the research plan or simply proceed without institutional oversight. Using these strategies and resources may not produce the most cutting edge research or results, but it will allow for consistent funding. Alternatively, institutions unintentionally accepting terms with which they are not able to comply could be prevented from implementing proposed research or resources in favor of compliance with NASA policies or other contractual duties. Such a situation would prevent maximization of the potential research. Alternatively, the situation may lead to the unintentional violation of policies or agreements. A realization of either hypothetical institution's scenario creates a lower level of results and which outweighs the benefit of access to open code. While being able to recreate and test results is of monumental importance to furthering scientific innovation, budget and funding realities have reduced opportunities to conduct such work. However, it is difficult if not impossible to qualify the value certain research aspects against each other. Is it more beneficial to the collective to have readily accessible and testable deliverables, or is more value extracted from complex research lacking some transparency? This is a question that may need evaluation in order to effectively implement an open code policy.

While some of the indirect costs associated with an open code policy may be high depending on the form and application of the policy, there is undeniable utility in making a broad swath of information automatically available to the public. Such a release is the preferred strategy for most researchers anyway. However, a mandatory release of all code developed under a NASA award creates a great deal of uncertainty when trying to collaborate and find necessary research tools. A better approach may be to refine the structure of 1852.227-7014 to contain a default position that all code first developed will be released after a certain trigger. An effective trigger would be temporally based in relation to delivery or creation of computer software. Such a timeline would provide a guideline for the evaluation of computer software. Researchers working on federally funded projects that produce software have much fewer opportunities to secure funding through alternative routes like university technology transfer.[3] They need to be able to further their research through an individualized manner. Either they have something that is commercially valuable or they are able to secure a share of royalties for their lab while still giving NASA unlimited rights, or they are able to release their work under open source conditions to garner funding and notoriety. Researchers know their fields and technology best and are in the best position to create the most value. Both choices result in the code reaching the public. The increased value of having the code refined and made available to the public has the potential to overcome the initial cost to access the code. However, the initial cost often includes dedicated support and updates of the original code. Whereas, a free release through an unfamiliar mechanism might lead only to open code with no support or continued developement.

Even though the open code policy currently only at a request for comment stage researchers have had strong reactions to the potential outcomes of the policy. Most researchers

[3] The Economic Contribution of University/Nonprofit Inventions in the United States: 1996-2013. Prepared for the Biotechnology Industry Organization by Lori Pressman, David Roessner, Jennifer Bond, Sumiye Okubo and Mark Planting, March 2015

already release the majority of their developments under an open source license though GitHub. They are familiar with that the expectations and work that comes with providing their code to the community. However, concerns were expressed about unknown requirements of an open code policy. It is yet unclear if any sort of support by the contributing research will need to be maintained. In the event that a piece of software is receiving modification from community members, there may be a necessity to coordinate associated software change processes, including issue tracking, code branching/modification, testing and continuous integration and release. Researchers would also like clarification on how uniformly an open code policy would be implemented. There is concern that tools that have been developed or used in conjunction with a non-federal project may need to be used in future work. It would be difficult for a researcher to either disentangle certain pieces of the software that don’t have the potential for compliance with the open code, or forego the use of a valuable development when a large portion of the software is compatible with such a policy.