Ensuring Center for Army Analysis Compliance with the Federal Source Code Policy

Working Paper

Tom Spoon, Rick Hanson and Craig Flewelling

3 March 2017

Abstract

The Center for Army Analysis (CAA) should re-align its policy for the creation and releasability of Open Source Software (OSS), with the current federal policy mandated by Executive Order and detailed by the Federal Source Code Policy (FSCP). This re-alignment will take CAA (and other agencies) from the 1950s "communication channels of inability", and the mistrust which follows from that, to the modern era of facilitating communication, sharing technology, and collaboration. As the Office of the President of the United States notes, this will accrue manifold benefits across the federal government and, when warranted, with the public. Time and again, engagement across both communities — such as in the development and maintenance of Open Source Software and the collaborative work enabled by internet discussion boards and (pre-web) newsgroups — successfully brought vast, previously untapped technical resources to bear to solve complicated problems.

1 Background

Like many government agencies, the Center for Army Analysis (CAA) is both a consumer and producer of software. From decades-old legacy models and simulations that support Army and Department of Defense analyses, to the utilities, scripts, and tools of today that enable rapid research, analysis, and improved productivity, the federal employees at CAA have created a bevy of source code.

CAA's default position regarding sharing these creative works, particularly in the realm of modeling and simulation source, is one of closed-source and isolation. Objections to releasing source code range from security and classification risks, to loss of a competitive analytic edge, to lacking indemnity from potential consumers of CAA software, to the simple absence of a mandate to share. These concerns are more or less codified in the prevailing Army Modeling and Simulation Office (AMSO) releasability policy in AR 5-11 (2014), which fundamentally leaves the decision to share source code to CAA. Until recently, CAA's default position has been supported by legacy policy that defers to the preference of CAA leadership rather than mandating sharing and openness.

This becomes problematic when considering the overall benefits accrued to the government and the American public — that have been identified by the Executive Branch and formalized into policy — for *openness and sharing*, the exact opposite operating procedures of the legacy one of *closedness and isolation* maintained at some lower echelons. Also problematic is when agencies leverage the technology of open source (implying, mutually shared), meanwhile ironically, and to our eventual detriment, refusing to share the advancements they made upon them (when national security concerns are not involved).

2 Problem

Unfortunately CAA's position, and the policy in AR 5-11, is not compliant with current policy, namely the Federal Source Code Policy as directed by Executive Order and subsequent Office of Management and Budget, Executive Office of the President (OMB) policy. Initiated by President Obama, and continued by President Trump, the default position of the federal government towards source code is openness. This position emerged in 2013 under Executive Order No. 13642, "Making Open and Machine Readable the New Default for Government Information" (EO 13642). Via EO 13642, the Obama administration mandated sharing government data — including source code — between federal agencies, and embraced open-sourcing federal software to spur cost-reduction, reduce redundancy, and place tax-payer funded works in the public domain.

EO13642 authorized and directed the OMB to define and implement the federal policy to execute the order, which OMB detailed under the Digital Government Initiative (a.k.a. "The People's Code") in August of 2016. The ultimate result emerged in OMB M-16-21, "Federal Source Code Policy Memorandum," when the detailed Federal Source Code Policy (FSCP) became the governing policy for all federal agencies. Since the policy's inception, marked by the 2016 rollout of the central coordinating repository of government code `https://code.gov`, CAA has not been compliant with the mandates of the FSCP and the federal government's modernized policy on sharing and Open Source Software(OSS).

3 Paradigm Shift in Software Development: Embracing Open Source

DoD has been an early-adopter and general proponent of the utility of Open Source Software. The DoD Chief Information Officer (DoD-CIO/G6) paved the way for broad adoption and use of superior Open Source tools on DoD networks with its 2009 policy, "DoD-CIO/G6 Clarification Guidance Regarding Open Source Software (OSS)". In conjunction with the memorandum, DoD/CIO-G6 provided an official website addressing the benefits, risks, legalities, and practical questions surrounding the use of Open Source Software on DoD networks (`http://dodcio.defense.gov/Open-Source-Software-FAQ/`).

In many ways, the policy and the website were merely addressing the reality of prevalent usage of OSS throughout DoD. Case in point: CAA initiated and maintains exceptions to Army policy for internal networks specifically to leverage the superior capability of both closed and Open Source Software that may not have a Certificate of Networthiness, or otherwise be "approved" software. Even legacy CAA models and simulations rely heavily on Open Source Software for development, scripting, and processing (both pre and post).

Still, the DoD/CIO-G6 policy focuses on access and use of Open Source Software, rather than on sharing existing federal source code or releasing source code into the public domain. Notably, the Army CIO/G6 has no additional policy addressing either the use or release of open source software. Until the emergence of the Federal Source Code Policy, DoD agencies had the ability to use Open Source Software for mission execution, without the mandate to share or otherwise release source code.

4 Benefits of Open Source and Sharing

Open Source Software fundamentally rests on the idea that the open sharing of information yields numerous second and third order benefits. In the case of software, making the source code available promotes the following principles.

- Publication

  - Particularly in the case of reproducible research, releasing the source code is a modern pre-requisite to validate any scientific or academic observations.
  - Similarly, making even amateur or experimental source code available serves as an entry point for establishing communication with other interested parties that may form a community.
- Collaboration
  - Freedom of information fosters communication, allowing interested parties to collaborate across various domains including research, development, issue tracking, patching source code, and more.
- Reuse
  - The "community" can freely build upon and extend existing works, ideally bypassing the need to constantly reinvent the wheel or duplicate effort.
  - Much like academia and industry, reusable source code allows the state of the art to advance more rapidly.

According to DoD (DoD/CIO-G6 Clarifying Guidance on Open Source Software, p. 4-5),

> *(i) The continuous and broad peer-review enabled by publicly available source code supports software reliability and security efforts through the identification and elimination of defects that might otherwise go unrecognized by a more limited core development team.*
>
> *(ii) The unrestricted ability to modify software source code enables the Department to respond more rapidly to changing situations, missions, and future threats.*
>
> *(iii) Reliance on a particular software developer or vendor due to proprietary restrictions may be reduced by the use of OSS, which can be operated and maintained by multiple vendors, thus reducing barriers to entry and exit.*
>
> *(iv) Open source licenses do not restrict who can use the software or the fields of endeavor in which the software can be used. Therefore, OSS provides a net-centric licensing model that enables rapid provisioning of both known and unanticipated users.*
>
> *(v) Since OSS typically does not have a per-seat licensing cost, it can provide a cost advantage in situations where many copies of the software may be required, and can mitigate risk of cost growth due to licensing in situations where the total number of users may not be known in advance.*
>
> *(vi) By sharing the responsibility for maintenance of OSS with other users, the Department can benefit by reducing the total cost of ownership for software, particularly compared with software for which the Department has sole responsibility for maintenance (e.g., GOTS).*
>
> *(vii) OSS is particularly suitable for rapid prototyping and experimentation, where the ability to "test drive" the software with minimal costs and administrative delays can be important.*

Open Source Software has had such undeniable success that even historically closed-source, proprietary companies like Microsoft have been making significant investments in opening their infrastructure and significantly contributing to the Open Source community, including helping historical "threats" like the Linux operating system.

5 OMB M-16-21, "Federal Source Code Policy Memorandum"

As a consequence of EO 13642, the OMB established the Federal Source Code Policy via memorandum M-16-21. The Federal Source Code Policy prescribes mandatory requirements for federal agency practices for sharing source code, open-sourcing software, and participating in the Open Source Software community.

5.1 Provisions

Major provisions of the FSCP include the following.

5.1.1 Government-wide Source Code Sharing

At a minimum, government agencies are mandated to make source code available to the federal government for interagency reuse.

5.1.2 Open Source Software Pilot Program

- Government agencies (including DoD) are automatically part of a 3-year Open Source Pilot Program. During the pilot program, policy mandates that 20% of newly created custom software is to be open source. This applies to software created by government employees, and software created in execution of a government contract.
- Open-sourcing software is highly encouraged, and even temporarily mandated for a portion of our new custom software projects for the duration of the pilot program. Open source software is intended to be broadly accessible, and developed with "open" practices, disseminated on — possibly 3rd party — platforms with established open source communities.

5.1.3 code.gov

Per the policy, agencies will coordinate and publicize their software with the OMB-managed website `https://code.gov` . This website serves as an accessible repository of all known government source code, to facilitate discovery and ease reuse.

5.2 Participation

The FSCP encourages participation in the OSS community. Here is the relevant excerpt taken from the FSCP memorandum, pp. 8-10.

> *When agencies release custom-developed source code as OSS to the public, they should develop and release the code in a manner that (1) fosters communities around shared challenges, (2) improves the ability of the OSS community to provide feedback on, and make contributions to, the source code, and (3) encourages Federal employees and contractors to contribute back to the broader OSS community by making contributions to existing OSS projects. In furtherance of this strategy, agencies should comply with the following principles:*
>
> - *Leverage Existing Communities:*
>   - *Whenever possible, teams releasing custom-developed code to the public as OSS should appropriately engage and coordinate with existing communities relevant to the project. Government agencies should only develop their own communities when existing communities do not satisfy their needs.*
> - *Engage in Open Development:*
>   - *Software that is custom-developed for or by agencies should, to the extent possible and appropriate, be developed using open development practices. These practices provide an environment in which OSS can flourish and be repurposed. This principle, as well as the one below for releasing source*

code, include distributing a minimum viable product as OSS; engaging the public before official release; and drawing upon the public's knowledge to make improvements to the project.

- *Adopt a Regular Release Schedule:*
  - *In instances where software cannot be developed using open development practices, but is otherwise appropriate for release to the public, agencies should establish an incremental release schedule to make the source code and associated documentation available for public use.*
- *Engage with the Community:*
  - *Similar to the requirement in the Administration's Open Data Policy, agencies should create a process to engage in two-way communication with users and contributors to solicit help in prioritizing the release of source code and feedback on the agencies' engagement with the community.*
- *Consider Code Contributions:*
  - *One of the potential benefits of OSS lies within the communities that grow around OSS projects, whereby any party can contribute new code, modify existing code, or make other suggestions to improve the software throughout the software development lifecycle. Communities help monitor changes to code, track potential errors and flaws in code, and other related activities. These kinds of contributions should be anticipated and, where appropriate, considered for integration into custom-developed government software or associated materials.*
- *Documentation:*
  - *It is important to provide OSS users and contributors with adequate documentation of source code in an effort to facilitate use and adoption. Agencies must ensure that their repositories include enough information to allow reuse and participation by third parties. In participating in community-maintained repositories, agencies should follow community documentation standards.*
  - *At a minimum, OSS repositories maintained by agencies must include the following information:*
    - *Status of software (e.g., prototype, alpha, beta, release, etc.);*
    - *Intended purpose of software;*
    - *Expected engagement level (i.e., how frequently the community can expect agency activity);*
    - *License details; and*
    - *Any other relevant technical details on how to build, make, install, or use the software, including dependencies (if applicable).*

5.3 Exceptions

The FSCP acknowledges exceptions to policy where legislation, or national security precludes the release of source code. If source code is either classified, or classified as a "national security system" under 44 U.S. Code § 3542, the software is excepted.

5.4 Management and Oversight

DoD (and Army) CIO(s) are required to coordinate with the OMB CIO to define and execute an implementation plan for the OMB guidance. OMB provides quarterly processes that oversee the growth, maintenance, and overall progress of both the pilot program and compliance with the Federal Source Code Policy.

6 AR 5-11 "Management of Army Modeling and Simulation, 30 May 2014"

AR 5-11 is problematic for a number of reasons. First, there is no mention (specifically nothing precluding) distribution of models and simulations as Open Source Software. Further, the distribution processes defined by the AR, specifically for interagency — even internal Army distribution — seems to directly contradict both the Executive Order and the OMB implementation memorandum. Further, none of the regulations referenced in AR 5-11 address the possibility of Open Source Software — or even acknowledge it per se — nor do they address the legal and regulatory issues and mandates regarding OSS in the executive directives specified by the EO and the OMB memorandum. In general, AR 5-11 should be refreshed and/or rewritten to account for the FSCP and to clean up antiquated terminology.

7 Forcing Functions

There is, at a minimum, a federal mandate for sharing source code across the government, and a mandate to open-source 20% of our custom code during the course of the 3-year Open Source Pilot Program. Under the Federal Source Code Policy, CAA must share our source in an open, unimpeded manner with other government agencies so that there is government-wide reuse and cost saving.

Some agencies, such as NASA and the US Army Research Laboratories, are choosing to cut to the chase, and both open source and openly develop their code on GitHub. GitHub is the largest 3rd-party open source community that offers source code hosting services, and meets every prescription of the FSCP guidance for "Participation in the Open Source Community." GitHub repositories are then registered with `https://code.gov` to satisfy the discoverability and coordination requirements in the FSCP.

8 Toward Open Sourcing MARATHON 4

MARATHON 4 is written in an open source language (Clojure), managed with open source tools (Git), and has emphasized unclassified development from inception. MARATHON 4 is intentionally written and maintained in such a way as to facilitate sharing and discovery, particularly to enable flexible development among remote work locations and to enable sharing of code for research purposes, peer-review, external verification, and publication in professional forums like MORS, WinterSim, and INFORMS. MARATHON 4 is, for all intents and purposes, open-source ready and entirely compliant with the practices established by the FSCP. Consequently, MARATHON 4 is an obvious open source release candidate, preferably hosted on GitHub.

8.1 Practical Benefits of Open Source Via GitHub

- It allows for flexible team-based collaboration.
  - Developers can work remotely, from home, the office, at odd hours, etc. Using Git, we have a rich collaborative platform for managing the source code, enabling concurrent, asynchronous development that maximizes development team productivity without sacrificing version control. This complements existing technology like Defense Collaboration Services (DCS), allowing teams to communicate in real-time to resolve issues, learn about the software architecture, and even modify the source code.
- There is empirical evidence at CAA of its usefulness/value.

  - CAA has repeatedly maintained a developer shortfall; MARATHON 4 is a shining example of the scarcity of developer talent. The sole developer (Mr. Spoon) was allowed to continue working remotely because of his decision to maintain MARATHON 4 development in an unclassified format, thus enabling exactly the kind of remote/telework opportunity mentioned above. CAA has been able to avert the loss of critical infrastructure development precisely due to the flexibility enabled by distributed version control, unclassified development, and openness.
  - With the addition of new team members, leveraging GitHub as a synchronization point has already been incredibly useful for distance-based training, collaboration, source code revision, and real-time pair-programming.
  - This very document has been collaboratively built and revised on GitHub by CAA personnel.
- It is industry-standard version control.
  - GitHub provides a seamlessly integrated suite of tools that enhance the Git distributed version control system (DVCS) developer experience with
    * source code repository hosting;
    * web-based interface for examining source code history, diffs, branches, etc.; and
    * web-based issue tracking, team communication, and other collaboration features.

8.2 Compliance with the Federal Source Code Policy

- At a minimum, MARATHON 4 must be shared with other federal agencies.
- Hosting as an open source project, hosted on GitHub, satisfies the existing Federal Source Code Policy, in addition to the spirit of the executive order.
- MARATHON 4 could be used to fulfill the 20% mandate for open-sourcing custom software during the current FSCP pilot program period.

9 Possible Objections and Risks

- "Army Policy Prevents Us From Doing So"
  - The AMSO guidance in AR 5-11 contradicts (or in the best case, is ignorant of) the Federal Source Code Policy. The apparent reflexive response to "not share" with federal agencies, and international partners, is contrary to both the spirit and the policy codified by EO 13642 and OMB M-16-21.
- "We should protect Army / CAA interests by not sharing source code."
  - The numerous benefits delineated by the DoD/CIO-G6, as well as decades of empirical confirmation that "sharing is beneficial" from the software industry and academia support an alternate prospect: CAA would be protecting CAA / Army interests by taking advantage of the massive benefits of open source collaboration, and by complying with policy set forth by the Executive Office of the President of the United States.
- "We should run the model, they don't need the source. They can ask us for the analysis."
  - This service-minded aspect of Army M&S is detailed in AR 5-11 and is the predominant "business model" that CAA and many analytic agencies follow. Sharing code does not equate to sharing expertise. Indeed, the dominant open source software business model is to provide support and service in exchange for

remuneration. Many clients or sponsors simply lack the developer talent or inclination to modify the source code, and will still be interested in the services provided. The legacy service-based model can — and will — survive, with the added benefit of collaboration and possible community engagement.

- "We'd expose ourselves to security vulnerability."
  - MARATHON 4 is not a national security system, as defined by 44 U.S. Code § 3542. MARATHON 4 is merely an instantiation of AR 525-29, a publicly available document detailing Army Force Generation. Started as a purely unclassified development effort, MARATHON 4 maintains that the source code for the simulation — including comments, notional test data, and related documentation — neither requires nor includes classified information. Rather, only the data upon which MARATHON 4 is applied, and the resulting analysis, if performed on a secure network with classified input, would be classified.
  - The security benefits of sharing and open sourcing are well-known, even within DoD and the Army. Per the (DoD-CIO/G6 OSS FAQ, "Q: Doesn't hiding source code automatically make software more secure?"):

    > *Even when the original source is necessary for in-depth analysis, making source code available to the public significantly aids defenders and not just attackers. Continuous and broad peer-review, enabled by publicly available source code, improves software reliability and security through the identification and elimination of defects that might otherwise go unrecognized by the core development team.*
    >
    > *Conversely, where source code is hidden from the public, attackers can attack the software anyway as described above. In addition, an attacker can often acquire the original source code from suppliers anyway (either because the supplier voluntarily provides it or via attacks against the supplier). In such cases where only the attacker has the source code, the attacker ends up with another advantage.*
    >
    > *Hiding source code does inhibit the ability of third parties to respond to vulnerabilities (because changing software is more difficult without the source code), but this is obviously not a security advantage. In general, "Security by Obscurity" is widely denigrated.*

  - The Office of the Secretary of Defense recently piloted a successful program, called "Hack the Pentagon", to harden Pentagon defenses by engaging the broader community of (third party) security experts to test DoD software systems and services at the Pentagon. The DoD News article "DoD Announces 'Hack the Pentagon' Follow-Up Initiative" outlines the results:

    > *. . . the pilot program . . . allowed more than 1,400 registered hackers to test the defenses of select open source DoD websites such as Defense.gov. Hackers who identified security gaps that qualified as valid vulnerabilities were then rewarded with a corresponding bounty price. As a result of this pilot, 138 unique and previously undisclosed vulnerabilities were identified by security researchers and remediated in near real-time by the Defense Media Activity.*

- "We have no obligation to release if no-one asks."
  - The Federal Source Code Policy mandates that our non-exempt software, like MARATHON 4, is — at a minimum — advertised via `https://code.gov` and accessible to other federal agencies for reuse.
  - The Federal Source Code Policy mandates that 20% of created or acquired custom software must be released as open source during the current pilot program, which lasts until 2019.

- "Contractors will just repackage it and sell it back to us."
  - If a contractor uses the source code to make something even marginally better, then under the EO and OMB guidance we (the federal government) actually should get their modifications back in full. Other agencies devoting resources to improve MARATHON 4, with CAA controlling the integration and merging of improvements, serves to extend the range of support for MARATHON 4 development, further helping the chronic developer capability gap at CAA.

- "CAA will be legally liable for support and any problems users encounter if we open source."
  - The default open source posture provided by the FSCP, and DoD/CIO-G6, precludes this possibility. Where applicable, open source licenses vetted by DCS/CIO-G6, specifically indemnify the original author of the code and provision no warranty for fitness of use or guarantee of support.

10 Desired End State

Ideally, CAA will join the ranks of other Federal agencies and embrace the general modernization of government technology, specifically the realization of the benefits of Open Source Software development and sharing source code. The tendency to reflexively lock down source code without assessing the benefits — let alone the current mandate — to share our knowledge across the government and the public domain, serves to ensure isolated, resource-constrained development devoid of the known value of external collaboration. In pursuance of modernizing the Army technology space and in accordance with the Federal Source Code Policy, CAA and AMSO should lead this effort from the front.

In an ideal world, the benefits of sharing source code and allowing for interoperability with other agencies (even individuals like college students, researchers, or industry professionals) can pay dividends in improving the source, aiding in verification, and generally building a community of interest.

Access to the source code does not imply knowledge of how to build, execute, modify, or extend the model; nor does access engender an innate desire to do so. The established model-as-service approach still works under the open source paradigm.

MARATHON 4 can directly benefit from open development and hosting on GitHub by taking advantage of the open source paradigm under the auspices of the governing Federal Source Code Policy.

11 Recommendations

1. CAA should comply with the Federal Source Code Policy.
   - CAA should provision the sharing of source code with federal agencies, and advertise repositories on code.gov and / or code.mil as appropriate.
   - CAA should comply with the provisions of the pilot program from M-16-21, that 20% of newly-created (or acquired) custom software must be released as open source.
2. AR 5-11 "Management of Army Modeling and Simulation, 30 May 2014" should be made consistent with the Federal Source Code Policy.
   - AR 5-11 does not account for the paradigm shift toward Open Source Software development and hence is now incongruous with the governing orders regarding the Federal Source Code Policy.
3. MARATHON 4 should be hosted on GitHub to enable collaborative team development.
   - CAA should actively leverage modern technology (GitHub and DVCS) to address the long-standing MARATHON developer capability gap.
   - Until CAA refines its position on compliance with the Federal Source Code Policy, MARATHON should be maintained as a private repository on GitHub to enable rapid development and verification in the near-term.

  - Private GitHub repositories are available, but require additional $25/month funding to support 5 developers for an organization.
  - Should CAA decide to openly develop MARATHON, GitHub hosting is free for public repositories.

4. CAA should follow the example set by US Army Research Laboratories and release MARATHON 4 into the open as public domain software.
   - The U.S. Government has no copyright or intellectual property claim to MARATHON 4 or any taxpayer-funded creative work.
   - MARATHON does not meet the exceptions provided by the Federal Source Code Policy, namely the legal, classification, or national security system (44 USC 3542) exceptions.
   - Open development can only serve to strengthen the quality of MARATHON 4 by easing collaboration and community engagement, while fulfilling the Federal Source Code Policy mandate for the Open Source Software Pilot Program, and establishing CAA as a leader in modern federal Open Source Software development.

12 References

- Code.gov: `https://code.gov`
- Code.mil: `https://code.mil`
  - This is run by Defense Digital Service of the US Digital Service.
  - Defense Digital Service: `https://www.dds.mil`
  - US Digital Service: `https://www.usds.gov`
- AR 5-11 (Management of Army Modeling and Simulation, 30 May 2014)
  `http://www.apd.army.mil/epubs/DR_pubs/DR_a/pdf/web/r5_11.pdf`
- AR 25-1 (Army Information Technology, 25 June 2013)
  `http://www.apd.army.mil/epubs/DR_pubs/DR_a/pdf/web/r25_1.pdf`
- DoD-CIO/G6 (Clarification Guidance Regarding Open Source Software (OSS), 16 October 2009)
  `http://dodcio.defense.gov/Portals/0/Documents/FOSS/2009OSS.pdf`
- DoD-CIO/G6 (Open Source Software FAQ)
  `http://dodcio.defense.gov/Open-Source-Software-FAQ/`
- Federal Source Code Policy Memorandum
  `https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2016/m_16_21.pdf`
  `https://sourcecode.cio.gov`
  `https://sourcecode.cio.gov/Exceptions`
- Executive Order 13642, May 9, 2013 "Making Open and Machine Readable the New Default for Government Information"
  `https://www.gpo.gov/fdsys/pkg/DCPD-201300318/pdf/DCPD-201300318.pdf`
  - Initial guidance for the Federal Source Code Policy
- Obama's Digital Government Initiative
  `https://obamawhitehouse.archives.gov/sites/default/files/omb/egov/digital-government/digital-government.html`
  `https://obamawhitehouse.archives.gov/blog/2016/08/08/peoples-code`
- Definition of "national security system" from "44 U.S. Code § 3542 - Definitions"
  `https://www.law.cornell.edu/uscode/text/44/3542`
  - This term shows up in the Federal Source Code Policy Memorandum (esp. in section "Exceptions")
- Hack the Pentagon Results, "DoD Announces Hack the Pentagon Follow-up Initiative"
  `https://www.defense.gov/News/Article/Article/981160/dod-announces-hack-the-pentagon-follow-up-initiative`