

[NASA Mini-Summit for Open Source Software and Science](#) [1]

Submitted by superadmin on Fri, 2012-06-29 13:24 Tuesday, July 17, 2012 - 08:30 to 10:00
Tuesday, July 17, 2012 - 10:30 to 12:00

Event: [Summer Meeting 2012](#) [2]

Session Type: [Workshop](#) [3]

Identifier: doi:10.7269/P3WD3XH9

Collaboration Area: [Geospatial](#) [4]

Abstract/Agenda:

This session will host a series of Open Source for NASA Science Data Systems talks, focusing on:

- * Deployments
- * Licensing, Attribution and Redistribution issues in Open Source Software
- * Sharing Software Between Centers: Success Stories
- * Short Lightning Talks
- * How NASA open source software can help the ESDIS project
- * Legal/IP issues in Open Source Software

We will solicit between 4-6 30 min talks from leaders in the NASA Earth Science Data Systems Community and from leaders in the broader open source community from which NASA leverages (e.g., from NetCDF/UCAR, HDF5-group, GDAL, PostGIS, etc.). In addition, we will have 2 focused discussion sections with those in attendance to digest and learn from the talks given. Community feedback will be an important aspect of the workshop, as will sharing "tribal knowledge" from those in the open source community.

Detailed Agenda:

Presenters

8:30 - Introduction (Chris Mattmann)

8:40 - NASA ESDIS DRAG Open Source Effort - Andrew MitchellNASA - 25 mins + 5 mins questions

9:10 - ESRI Geo Portal Open Source Platform - Christine White/ESRI - 25 mins + 5 mins questions

9:40 - Discussion Topic: Process versus Project - should agencies develop an open source policy or use successful projects as a model?

10:00 - Break #1

10:30 - HDF Group - open source data management tools - Mike Folk/HDF Group - 25 mins + 5mins questions

11:00 - NOAA Open Source Success stories (ESRI Geo Portal) - Yuanjie Li/NOAA - 25 mins + 5mins questions

11:30 - TBD Speaker - 25 mins + 5 mins questions

12:00 - Adjourn and Further discussions

We expect strong synergy between the Geospatial Cluster and this workshop, including participation and attendance.

Notes:

NASA Mini-Summit for Open Source

What does Open Source Mean for HDF?

Mike Folk

What is HDF:

- a data model
- open file format
- open source software

Two version:

HDF4 & HDF5

- Both are relevant as far as the talk goes

History:

- U of Illinois, NCSA 1988-2006
- Non profit since 2006
- Best examples of open source project have the same roots as we do

Mission

- managing large complex data set
- provide services for users
- sustainability over long term
 - long term access and usability of data
 - left university to ensure this

HDF Communities

- Academia, government, commercial
- Large and complex data

What we do:

- EOS and JPSS, support in an enterprise way, whatever they need
- General maintenance, QA and Support

Open technology development

- they are a single maintainer open source company

Intellectual Property

- transferred to HDF Group for royalty on commercial profits
- they have a BSD license (Berkeley System Distribution)

- people can use software however they want to

Benefits on OSS, as it related to BDF

- trying before adopting
- if it almost works, you can modify it to make it work
 - usually better to go to them for help
- development activities are public
- freedom to develop tools that make HDF more usable
- Long term access
 - Code open, if the company stops, the code is still available

Aspects of OSS we're less sure about, as they relate to HDF

1. unpaid contributors can do much core work
 - has not been their experience
 - not software that a lot of people are interested in working on
 - we have people that they pay, which for them has worked better so far
2. given enough eyeballs, all bugs are shallow
 - 1. has not really worked for them
3. frequent development cycles are good
 1. when you have people writing applications with a particular API
 1. no documentation
 2. critical API
 3. change API,
 2. this can be detrimental to communities
 1. need to be slow and work with user community
4. OSS is easy to use
 1. Need to qualify this
5. OSS is low cost
 1. There are a lot of expenses
6. It is easy to run a community-based OSS project
 1. Managing that, you need to really invest
 2. Not simple
7. OSS business models
 1. All completely different

Comments/Questions:

1. If they became a community model maybe they'd have more community involvement
 1. Such as documentation or other areas that are not in regards to coding
2. Contradiction: Only HDF group understands the complexity, so if the HDF group went away then no one would understand the complexity
 1. Maintaining a knowledge core of developers and supporters will help
 2. If the core is not there at least you have the source code

Ken Casey: NOAA Open Source

- NOAA National Oceanographic Data Center
- Real world demonstration of OSS
- NODC
 - Data services
 - Have done usability tests (will present on Thursday)

- NODC
 - Space data to numerical models, heterogeneous, a lot in HDF, and such
- Scientific Stewardship
 - Acquire
 - Archive: understandability over long term
 - Access
 - Add Value

NODC Archive Paradigm

- human and machine interfaces
 - Google, data.gov, GOS, OAS, Geoportal Server Web App
- Machine to Machine
 - CSW, Geoportal, REST API, Open Search
- Metadata: FGDC, ISO, Ad hoc, minimal standardized metadata
- Geoportal Highlights (additional capabilities)
 - Multiple service links
 - Temporal search
 - Browse by keywords
 - Data visualization
 - Cart functions
 - Working with ESRI developers
- Other Improvements
 - Ocean Locator
 - Spatial search
- Discover by Browse
- Thumbnail graphics
- Users not just scientists, --schoolchildren to scientists

Machine to Machine

- Open Search
- ARCMAP
- Geoportal server -REST URL
 - Allows users/community to do something with it

Geoportal Server Knowledge Sharing

- working to make sure partners are aware
- knowledge and community

Open Source

- Force multipliers exist
- Choose open source and standards first
- Pay external experts to develop features
- ESRI
- OPeNDAP, Inc

Comments/Questions:

1. What does connect with Fed-wide platform?
 1. Still in progress
 2. Will working out connections

3. Connections should be easier than in past
2. What got from talk, start with what's out there from OSS and if it doesn't work develop what you need. There's some good to starting internally, for example some of the long-term archive parts of NASA. Present a lot of value, doesn't believe that open source should be the starting part.
 1. What Ken thinks is that some of the most general items, such as searching a catalog, open source makes sense
 2. One of the most custom items they have broken down from a monolithic item to a more loosely coupled item so that they can apply some open source items to the older systems

Discussion

Concerns:

1. Licensing
2. Redistribution
3. Open source "help desk" syndrome versus community
 1. Need to nurture the environment actually want to create this environment or learn
 2. Meritocracy: have various structures, established wait-times, and established social interactions
 3. Cross fertilization
4. Fear of messing up code:
 1. community models
 2. start small, people you know and trust
5. trust is the heart of role, unless there's trust people won't come to them for data or give data
 1. hard to quantify, but in favor of Open source

Session Leads:

Name: [Chris Mattmann](#) [5]

Organization(s): [Jet Propulsion Lab](#) [6]

Notes takers:

Name: [Angela P. Murillo](#) [7]

Email: amurillo@email.unc.edu [8]

Creative Common License: Creative Commons Attribution 3.0 License

Teaser: This session will host a series of #OpenSource for NASA Science Data Systems talks w/strong synergy between the Geospatial Cluster

Keywords: [Open Source](#) [9]

Source URL: <https://commons.esipfed.org/node/437>

Links

- [1] <https://commons.esipfed.org/node/437>
- [2] <https://commons.esipfed.org/event/summer-meeting-2012>
- [3] <https://commons.esipfed.org/session-type/workshop>
- [4] <https://commons.esipfed.org/collaboration-area/geospatial>
- [5] <https://commons.esipfed.org/node/330>
- [6] <https://commons.esipfed.org/taxonomy/term/197>
- [7] <https://commons.esipfed.org/node/633>
- [8] <mailto:amurillo@email.unc.edu>
- [9] <https://commons.esipfed.org/taxonomy/term/303>