

VOA Codebase

Paul Harrison - JBO

with thanks to mailing list contributors

Where do I download the VO Software?

- ✦ For most of the VO standards we do have high-quality implementations
- ✦ However, these tend to be controlled locally, at the national project or institutional level.
- ✦ it is hard to guarantee continued support in the long term
- ✦ it is not clear to third parties which is the "right" one to use.

Existing VO Codebases*

	language	Active	multi-org	VCS	contributors	content
<u>GAVO DaCHs</u>	Python	✓	✗	✓	≈	client/server
<u>OpenCADC</u>	Java	✓	✗	✓	3	library/server
<u>volute</u>	mixed	✓	✓	✓	30	documents/data models/ libraries
<u>VAO</u>	Java/Python/ C/.Net	✓	✗	✓	≈40 fewer active	client/server/ library
<u>astrogrid</u>	Java	✗	✗	✓	3 (activish) ≈20 (historic)	server/client/ library
<u>CDS</u>	Java	✓	✗	✗	?10	developer's corner - libraries

* with apologies to those that did not fit my criteria

Astropy

- an exemplar collaboration? Nice looking website!
 - can we learn from it?
- have published a paper (44 authors)
- uses GitHub
- 200+ contributors (on github 56 watchers, 239 Starred, 180 forks)
- built on top existing libraries (Numpy - Scipy)
- language specific (Python!)
- strongly curated - both software and web-site
- affiliated packages



Need to decide

- ✦ Hosting service
- ✦ Version Control system
- ✦ Software curation rules
- ✦ Coding standards

Hosting Provider

- ✦ will affect the optimal way to organise and manage the repository and how people might contribute.
- ✦ need to consider what facilities that hosting provider offers - e.g. bug management
- ✦ however, never allow being too dependent on particular provider

Version control

- ✧ Dichotomy between “distributed” and “centralized”
 - ✧ with Git and Subversion the “brand leaders” of each kind.
 - ✧ there are others... e.g. mercurial
 - ✧ “religious” issue!
- ✧ curation and contribution workflow will be strongly affected by the choice.

Codebase Curation

- ✦ Needs a small number of master curators that coordinate the whole project.
 - ✦ these individuals will need to donate considerable time - especially when working towards initial launch
- ✦ Code curators for each software “module”
- ✦ desirable that knowledgeable “outsiders” can easily contribute patches

Coding standards

- ✦ Aim to
 - ✦ produce consistent and (re-)usable APIs
 - ✦ have comprehensive user and developer documentation
 - ✦ control external dependencies
- ✦ Make it easy to do the right thing
 - ✦ have some short lists of “never do x” and “always do x” rules

Code Donation

- ✦ Idea will only fly if existing code is contributed to kick-start it.
- ✦ Not sufficient merely to include code. To be a cohesive project implies
 - ✦ reworking of existing code to conform to project standards.
 - ✦ possibly some loss of “ownership” for original author
 - ✦ however they should be the module “lead curator”
- ✦ Incentives - needs to be worth the contributors’ time in terms of being able to use other parts of the framework.
- ✦ What to do about existing competing code?

Nice to have

- ✦ generated documentation
 - ✦ way to republish to hosting web site, but independent of any particular publishing technology
- ✦ Agile process *cf* Standards Documentation
- ✦ incubator/affiliated modules
 - ✦ not formally in the repository, but working towards inclusion.
 - ✦ way to advertise what is being proposed

Libraries vs Services.

- ✦ Libraries and client code easier to integrate in this sort of project - services more messy because
 - ✦ services often run in some form of container
 - ✦ need to be configured
 - ✦ often have to talk to other components - e.g. databases
- ✦ Perhaps services would always be affiliated projects

Conclusions

- ✦ Needs to be “self organising”
 - ✦ but with some dedicated individuals
- ✦ Community needs to be in broad consensus
 - ✦ has political/personality implications
- ✦ Can the IVOA decide/agree on this?
 - ✦ better for a smaller like-minded group to initiate?

Personal opinion

- ✦ One repository per implementation language
- ✦ Astropy has momentum if implementation language is python - no point competing - just join in
- ✦ I would be willing to be a “lead curator” in a Java repository
 - ✦ I am anyway reviving parts of astrogrid for JBO publishing
- ✦ Volute should be endorsed as the IVOA document and data model repository
- ✦ Use GitHub! - mindshare again....