On a Cross-functional Data Science Team

Jordan Walker

2015-02-17

1 Abstract

Diversity and communication are key characteristics of effective interdisciplinary teams [1]. CIDA teams are generally aligned according to disciplinary and skill set backgrounds (developers, system administrators, and project leads), despite operating in an interdisciplinary field, in which software engineers, scientists, project managers, and system administrators work together to produce technology solutions. Historically, communication barriers between CIDA teams have prevented beneficial overlap of information and ideas. In the interest of maintaining and enhancing a sustainable and effective CIDA, a formal mechanism to increase communication channels between CIDA teams may be necessary.

I propose adding a member of the developer team to the newly formed data science team in order to increase communication between these groups. This position will act to prevent this group developing as a "silo", which could negatively impact CIDA overall. Such a position will benefit the data science team by directing computer science expertise to the goals of the team. Finally, alternative paths for computer scientist career advancement may be exposed by this change; demonstrating these alternative paths will help for recruitment and retention of computer scientists in the future.

2 Rationale

Assembling effective teams to tackle complex challenges requires mixing diverse backgrounds and a blend of specialists and generalists [1]. The goal of team formation within CIDA should not be to collect a group of people with the same background, but rather to collect a group of people with mutual respect and complementary skills, that are willing to converge on a shared approach to solving the problems that are presented to CIDA. The existing data science team is utilizing some of CIDA's best-practices combined with novel scientific research to establish a new approach to the challenges CIDA is attempting to counter. With this approach, the team has an opportunity to distinguish itself within USGS, but will best accomplish its goals with a cross-functional team.

One of the directives of the data science team is to publish research that highlights the work of CIDA, the OWI, and the USGS. Much of this work will be done by utilizing APIs produced within and outside of the USGS. A developer with firsthand experience with APIs can contribute to data science research by most effectively leveraging these technologies, troubleshooting issues, and by bringing to the table knowledge of APIs that might otherwise be unknown to the data science group. The data science team, and CIDA as a whole, will also benefit by the input of a developer on architectural decisions made by the data science team, and by integration with existing CIDA work.

The goal of "cross-functional" teams is prevalent within CIDA, but defining that goal is somewhat difficult. One definition that seems to be making its rounds is that everyone on a team is able to take on any task that is presented to the team. While this is a goal that has its merits, a more widely accepted definition of a cross-functional team is a "group of people with different functional expertise working toward a common goal" [3]. By being a cross-functional team, the data science team will integrate better with CIDA

and be more able to accomplish its goals.

A growing issue within the development group is recruitment. Qualified persons are difficult to find, and harder to hire. One point that might aid in recruitment would be for different career paths that one could take. The RGE and EDGE [2] programs allow for advancement based on accomplishments in research and development. Currently this is an untested approach for computer scientists, especially for CIDA. Having this option available would help especially if looking for candidates with a more academic background (e.g. Computer Science or Electrical Engineering graduate students). Currently this path would most likely be set through efforts taken on by the data science team.

3 Duties

As a member of the data science team, the developer role would include:

- Act as a liaison between the data science team and the other developers.
- Keep track of services and APIs that might be of use to data science team and advocate for any additions that might be helpful.
- Act as a developer on projects that require the skill-set of a traditional developer but fit within the mission of the data science team.
- Keep up with trends in development of methods to work with big data, complex models, and scalable processing.
- Provide architectural reviews for software developed within the data science team.
- Perform research related to work being done within CIDA, OWI, and the Water mission area.
- Contribute to an RGE/EDGE portfolio and be evaluated for a position.

4 Transition

Acknowledging that this is a bit of a change, it would be best to plan the transition to avoid adversely affecting ongoing projects. The developer switching should start out at around 25% on the data science team, joining in on meetings while still working on projects already committed to. Then as planning and funding allows the allocation can increase to an eventual 100%.

Additionally, the increased capacity on the data science team will allowfor projects that fit with the goals of the team to be added and tackled as a joint effort between functional groups. The flexibility this will offer should ensure a self-sufficient data science team into the future.

5 Conclusion

The establishment of a data science team was a smart move for CIDA and the OWI. Furthering its effectiveness through tighter integration with the rest of CIDA will benefit both the team and CIDA. Working within a different work structure allows for multiple approaches to solving the science problems presented to the USGS. Introducing multiple functions into the separate approaches allows for fully functional teams which can complete tasks independent of others, while preserving the channels of communication that make CIDA a unique team and place to work.

References

[1] Kendra S Cheruvelil, Patricia A Soranno, Kathleen C Weathers, Paul C Hanson, Simon J Goring, Christopher T Filstrup, and Emily K Read. Creating and maintaining high-performing collaborative research teams: the importance of diversity and interpersonal skills. *Frontiers in Ecology and the Environment*, 12(1):31–38, 2014.

[2] US Geological Survey. Usgs research and development guidance. Accessed 17-Feb-2015.

URL: http://www.usgs.gov/humancapital/hr/rgeg.html.

[3] Wikipeida. Cross-functional team, 2015. Accessed 17-Feb-2015. URL: http://en.wikipedia.org/wiki/Cross-functional_team.