

Geointeresting Podcast Transcript

Episode 3: Chris Rasmussen

June 22, 2015

Welcome to Geointeresting, the official podcast of the National Geospatial-Intelligence Agency. This week on the show we are celebrating the 2015 GEOINT Symposium with conversations with some of the most innovative and influential people within NGA.

Today we have Chris Rasmussen, NGA's public software development lead. Chris has championed NGA's official presence on the co-sharing site GitHub, overseeing the launch of more than 20 repositories on the platform. He is also spearheading the agency's Pathfinder program, an off-site, unclassified lab, to answer key intelligence questions. It is safe to say Chris is an agency trailblazer when it comes to pushing NGA to operate openly, transparently and unconventionally.

Stay tuned for Geointeresting!

NGA: So Chris, what sparks your interest in trying to push some of the agency projects you are working on to the public?

CHRIS: Well, the whole push towards GitHub started within the domestic-support circles within the domestic support we provide in HADR situations — humanitarian disaster response — and in some of the domestic-support stuff. It made perfect sense if we were building high-quality stuff for our domestic partners, our HADR partners to use; that is a completely unclassified field. The first thing that went out was GeoCube, and it was a hit. It was well-done; it just made perfect sense to start with that. If we are talking about getting on the same sheet of music in a common operational picture, it just made total sense to put that source code on GitHub. So internally, it just lined up, and it's grown much beyond the domestic support branches and things within the agency to different types of software packages, much, much further beyond that, but that was the originating piece. It just made perfect sense to do it and not overdo it with for-official-use controls. It was the perfect case to do that.

NGA: So a couple weeks ago we had the opportunity to have GitHub governmental evangelist, Ben Balter, on the show to discuss NGA's open-source initiatives. Tell me a little bit about the process in getting the agency onto GitHub and why it is so important.

CHRIS: Sure. So it took a while. Like I said, this was a first, and so it took a while. I've talked about this before, but most of my time during that selling it throughout the organization was myth busting. People think that because it exists on the internet, all the horror stories associated with internet-based chatrooms and the Yahoo comment section are the same thing; they are not. This is a very highly professionalized engineering culture, and just because it works over the internet doesn't mean it is going to bring all the associated baggage with certain things and biases people will have about internet-based communications. This is a very professionalized group, so "random people" cannot just show up and put in things; it doesn't work that way. You have to be a developer. You make your suggestions. They go into a bucket. We discuss the changes. We look at them. Then it is merged in. So it is a very controlled environment. There are very highly scientific groups that work in crowd-tasking platforms and things like that —



NGA
NATIONAL GEOSPATIAL-INTELLIGENCE AGENCY



astrophysicists; those are internet-based. So the issue was people coming up with kind of peanut-gallery stories that had nothing to do with the highly professionalized open-source software-development and engineering culture, so that was number one. So we had to figure out a way; there was no rear-view mirror for this. We had to make up all the forms; all the points of contact. All of it had to be done from scratch. So we craft a way to document intellectual property. If a contractor works on something under contract, we have to document intellectual property with the contractor and the contractor's company; convey the expectations that we are now going to open source the list to the public; that was new. People use the term open source a lot, but when you go out on a platform of millions with a legally reviewed open-source license, that is very different than claiming it is open source and emailing it to your friends, point to point. This is the real legal deal. So that had to be crafted on the contract side. On the government side, government employees are not entitled to copyright protection under U.S. law. So anything a government employee develops has to be released in the public domain. That can blend with other licenses, and it gets a little technical and nerdy on the legal side there. But we had to craft the government intellectual property forms, the contracting forms, and then we had to figure out a way to review it and come out with a security process to check it out to the public, which was created as well. So there is a couple of pieces. Document the intellectual property, who made what, how many people involved. Then there is the security checkout process of what it does. Then there is a conversation with security; there is a determination that is made. And then I went in, and I'll make the repo, the repository; I'll name it. In GitHub everything is lowercase and with slashes; that is kind of the convention. And then I turn over the rights to the NGA account to the lead developer of that project, who has the most intimate knowledge of the code. And then they start uploading. I'll write the 'read me' file. I'll pick the license. And then they come in and start uploading the code. The kind of expectations we have with our developers, anyone that's working for NGA if they are government or working on behalf of NGA as a contractor; they have to respond to all requests within 24 hours. So that doesn't mean if it is a big build. If it is a lot of code coming in they can just say, "Thanks. Gotcha." That is much much better than a lot of these government black holes; that you put in a comment, and you have no one commenting back to you. It just doesn't inspire a sense of community. So they have to be fun with it and be responsive. So that is kind of the process.

NGA: So after all that work, what has been the response you have seen from developers and coders about NGA sharing code publicly?

CHRIS: So when Balter was here, he talked about GeoCue, and that was the first to go. It was decent software. A lot of times there is scripts and really exocentric stuff an organization makes. I mean, most software is designed for very specific purposes; it is to turn the toaster on at 8 a.m. in the lobby. I mean, that is what a lot of software purpose is built for. But GeoCue was just really well-done. It was a hit that went out. Looking back and revisiting history a little bit, had we put out some of these other ones first, I don't think we would have had as big of a splash. There was some novelty in the sense that it was an intelligence agency doing this, so there was that. And then the first thing that went out was really good. So I think that the community was built from a combination of being a first mover and actually offering something that was really, really solid. That really pushed it. It surprised a lot of people that number one, we were doing it, and number two, that something that high quality was outputted.

NGA: Do you have plans to use any of the updates or improvements coming out of this?

CHRIS: Yes. So that is part of the process. So there have been substantial improvements to all of the projects. Now they tend to be configuration fixes and some minor-medium things. Yes, people want the “free feature” that comes in, but a lot of the times that is going to be built by your in-house talent that you are paying people to do. The trick is, how many full-time people do you devote to work in this transparent space? And then, do you want to be able to catch those spare cycles of people’s random contributions or time that they spent just wanting to help? So the issue is you are always going to have full-time people or people that you pay to do this, but if they work in a transparent space, you are going to be able to catch those much, much smaller spare cycle edits. So the things that have come through have been really good. Have they been these monstrous configuration; these big-feature-type things that have come in? No, they haven’t, but still the question is, do you want that fix; yes or no? It’s not whether it was big enough; it’s, do you want anything at all “for free?” And the answer is yes, of course you would.

NGA: So, switching to the other project hat that you’ve been wearing recently — can you break down the GEOINT Pathfinder project for us and what exactly is it going to deliver for NGA?

CHRIS: Sure. So the GEOINT Pathfinder was a recommendation that was given to Mr. Cardillo, and he signed off on it in March. Now, the logic is business analogies are imperfect in government, but they are instructive. The idea behind the Pathfinder was that sometimes companies — whether through mergers or acquisitions, or they come up with something that is totally different than the mainline business of the organization — they spin off an independent unit to let it grow. So headquarters doesn’t map its entities to it, and everyone just kind of projects what they’re used to doing on to it, but then it gets choked out. It is a real common strategy within the business world. If you look at the commodization of the commercialization of GEOINT in particular, most of the value this agency needs to drive will come from the unclassified side of the house. That is so fundamentally different than how we operate now that we needed to spin out a separate entity and let it breathe. So the recommendation was to have a unit that will only use unclassified sources and methods to answer these questions and output unclassified outputs. And the reason classified information was denied is we have done several of these experiments before. We need to up our open-source intelligence game. We start off strong for about two weeks; then we default right back to legacy behavior. So that was simply denied to inspire a sense of hustle and creativity. And it is going to be a separate kind of entity; it is going to be partially off-site. There will be some folks here, people teleworking and people in other labs throughout. But the idea was to disconnect a little bit to drive out new value. So what the Pathfinder will do, in addition to answering the questions; it is going to be a lot of pathfinding that these are the hiring practices that we need to pursue further. These are the policy rewrites that have to take place in order to be a much more powerful open-source-resource organization. These are the security-clearance requirements. Some of the things we are going to find out with Pathfinder; the question is, are security clearances required for all personnel? The answer might be no, because if you are doing high-quality unclassified research, you might not need to spend the money to run a clearance or basically, to securitize a facility to store classified information. It is going to impact almost every arm of the organization. What needs to take place? What needs to change in the enterprise? What needs to change in the security? What needs to change in contracting? What needs to change with how we interact with our partners; how we interact with the public? All of those things are going to change. And so the Pathfinder is going to help find what those are. And this is so complex that the only way to do this is through active experimentation. You can out business school this thing to death and outplan the plan the plan and study the study the study to come up with a recommendation. You have to do it to find out

where the complexities are and make the recommendations. So it is a learn-by-doing approach as well.

NGA: So you are on an elevator with the president and DNI Clapper, and you have a chance to pitch them the Pathfinder program. How do you sell them on the value of operating in an unclassified environment?

CHRIS: So for real quick, the real elevator speech was the commoditization rates of GEOINT are so high that the majority of value will come from unclassified sources. We have to figure out how to drive niche, new content and different thinking with mostly unclassified information. That is not to say classified information is going away. It is just going to be more focused, and it won't be the default. It needs to focus on its strong suit, not defaulted to answer every question. So I would say that we need to position ourselves to be a high-power, open-source, analytic entity, and we are not structured right now to do that, and that is what the Pathfinder is supposed to try to attempt to do.

NGA: In a perfect world, what do you hope to achieve through the Pathfinder process? Will you be producing reports? Will it be new code? What do the results look like?

CHRIS: Both. So there would be three parts. One would be the content, of course; the analytic writing outputs that will take place. We are not necessarily writing reports with logos on them. We are building up a living knowledge base. If someone wants to snapshot off a paper from the living content, that can be done, but the end state is not to produce more electronic paper. There are other ways to tell stories than pushing out reports. There are just other ways to do it. And the reason that GitHub works so well is it is on a distributive version control, meaning people can see all the changes that are happening within the source code, basically, through a Wiki. You can compare and see what the changes are. We are going to adopt the same philosophy with the content. But rather than source code, it is going to be written content — links to things, pictures of screen shots to things; look at this dataset. It needs to be a living knowledge base. That will be the analytic output piece. The second part will be the code that is developed there. There is a lot of tech here. Within the Pathfinder, we are going to keep a four-to-one analysis-to-developer ratio. So for every four end users, there will be one embedded developer to have that feedback loop; to have those conversations; to make those applications; to query the data; to be able to drive those outputs. So it is going to have a very tech-focused element. We can open source all of that stuff as well that we make, if we like, on GitHub. The last part will be the lessons learned. And I think this will be the key. It is documenting. What did we do? Why was it different? What needs to change back at headquarters? It is a combination of the analytic output with the IT and the lessons learned.

NGA: Operating out in the open and transparency is the priority of ODNI and of NGA Director Cardillo. What would you say to concerns that by doing so, the intelligence community might be giving away too much of the secret sauce?

CHRIS: Well, there is secret sauce. However, if you look at the phrase sources and methods, it's a common phrase you hear that you have to protect sources and methods. Well, take out half of that equation: sources. They are not all unclassified, right? You really have to focus on the methodology. And you have to know what is happening in industry and academia and in the private sector to know what is unique and what is stock, right? So every methodology, just

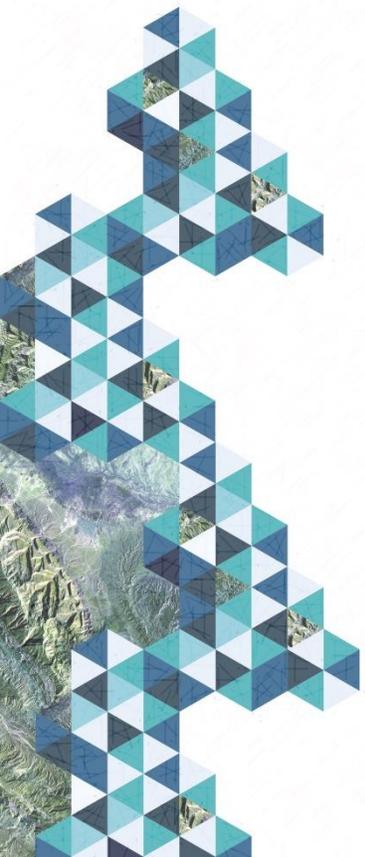
because it is an intelligence agency, does not inherently make it unique or classified. So we have to really focus the ask on, is this methodology truly unique? And in order to do that, you have to understand, what are the standard practices within the data science communities — what other companies are doing with this — to find out if this is just a common methodology of just the modern knowledge worker. So some of these conversations we are having now; you see them in the STEM debates, all the way down to elementary schools. Even on the college side, some of the best schools in the country are having the best conversation. What skills does a confident knowledge worker have moving forward in the 21st century when there is more data [and] things are becoming more statistically driven. And how do you collaborate with people you are not necessarily sitting next to? And how do you interact within these platforms? That is a huge, huge, huge ask, and I view we are intelligence officers; however, [also] knowledge workers. [There is] so much overlap between what a hedge-fund analysts does and what an intelligence analyst does, particularly on the strategic questions. A lot of these strategic questions are equally as boring that hedge fund folks are looking after; that we want to go after. So there are tactical things that have a sexiness to it. However, a lot of the sociopolitical questions and things like that; there are many, many different groups that are asking the exact same questions. And just because we ask them doesn't necessarily make it inherently spooky. So Mr. Cardillo's referred to this in the past, this self-imposed layer of caution. We just need to not overdo it with the spookiness when it is not required because it just puts self-imposed restraints on something that doesn't need to be there.

NGA: So, I've heard that NGA might have some involvement with Google's Summer of Code initiative. Can you share the details of that project?

CHRIS: Sure. So that was the GeoWave; I think maybe the third or fourth offering that we open sourced on the GitHub. And, basically, what it does is it adds a spatial and temporal layer over a humanlow, which is a big data system that was open sourced by the NSA. That project has kind of been the Eclipse Foundation. It's a big open-source software entity out there. There is a couple of them: there is the Apache Foundation, the Eclipse and others. Their whole reason for existence is to promote open-source software projects. So the Eclipse Foundation was kind of looking at GeoWave; they're thinking of starting up a subgroup called Location Tech, focused just on geospatial technologies. A lot of the big data systems in particular are designed to do text-based analytics. Geospatial is kind of an afterthought. GeoWave is a good project that is kind of bridging that. So GeoWave is the big data ecosystem with heavy emphasis on geospatial. That is kind of an emerging market. So the Eclipse Foundation is building out this working group to focus solely on that. They looked out the NGA's GitHub stuff and said, "Hey, this is pretty good. This is the bridge that we are looking for." They have a brand, and they have been in the open-source space for a while. NGA is in the early process of potentially joining the Eclipse Foundation. So by us investigating to join the Eclipse Foundation, we were able to kind of piggyback on their brand, and they won a spot on the Google Summer of Code, a global competition. This is a big deal. So the GeoWave is going to be in the Google Summer of Code kind of through the Eclipse brand, and we are doing that now. So this is a big deal. As when Ben Belter was here — he was involved with the Google Summer of Code. It is a huge global competition, and it is going to be really cool. So we are benefiting from kind of a test run with the Eclipse Foundation. But, regardless, to use a baseball analogy, an old one, in fact, from Bull Durham, "This is the show. This is a big deal." Which is really, really cool. This is a global competition, and to be a part of that; we've only been at this for a year with the entire open-

source thing in GitHub. So we are moving pretty fast, and that's pretty cool to be in that kind of circle and brand already.

NGA: I'd like to thank Chris for joining us on the show today. For more information on the GEOINT Pathfinder project, you can visit our website at www.nga.mil. We have a special GEOINT Symposium site with news and information, including a schedule of where and when Chris is speaking.



NGA.mil for more information