BP's Brian Boulmay headed up the OneMap project.

OneMap for all

Lessons in bringing company-wide data together onto one platform

BRIAN BOULMAY

As BP's Global Geospatial Information Lead, Brian Boulmay has spent the past eight years building a sustainable organisational capability for geospatial data and analytics in BP Upstream, with a focus on people, process and technology. The result is OneMap, which brings company-wide geographical information together onto one platform. In this article, Brian describes how OneMap came into being.

When I started in 2011 in BP's onshore US business, it was a relatively dispersed organisation. We had assets in Oklahoma and Colorado, and Wyoming, and Texas, etc. At that time, each of the business units, and even different business functions within each of the business units, were all trying to tackle digital mapping differently.

A lot of it was already Esri technology, but we had a few other technologies. Also, even within the Esri stack, we had multiple versions. Therefore, it was really hard to share data just within this one business entity. I came in as the first person ever hired to be a business-facing GIS lead. The mandate I was given was to look at where we were, suggest where we should be, and then put together a plan to get there. And so, I spent my first month effectively interviewing and talking to the whole organisation, trying to get a feel for the organisation, the people in play and the data sets in play.

I want to see

At the end of the day, as a business we must be able to see all our assets. We should be able to see all our wells. We should be able to see all our people.

We mapped where we were. At that time, we had six different systems that were doing digital mapping - we should just have one. The focus wasn't necessarily centralisation, it was more integration; I should be able to see across all my systems. As we got into it,

we realised that most of the tech was the same vendor already. Most of the workflows that we needed were supported by that core tech.

Why not just go ahead and standardise tech? I realised very quickly that even if I tied this tool and that tool to the same central version, I would still have two people following different processes. They'd be learning from different user manuals, learning different ways of doing it, because the tools were different. Whereas if I got them onto the same tool, all of a sudden, this worker and that worker, who are doing the same process, I can actually give them a single way to do it to help them.

By doing that, the rest became a lot easier, because now when trying to teach this exploration team and that exploration team and that exploration team, instead of writing three manuals, I would write one, and each of them could do their piece of work.

This was again onshore US, just that one business of BP, which at the time was one of 15 different regions. We did that for two years, with a very small team. It was so successful that one of the global managers said, "Hey, we want this everywhere else".

The platform is the foundation

We established a core platform, which back then was ArcGIS desktop. There were a few tools like SDE, RTS server and a portal, and that was the platform. Then there were a few other tools like FME and some other tools that plug into that, extensions, etc. Since then, it's grown to over 200 tools.

Every tool we bring in has to play with the core. We're very clear about having the integration conversation upfront. However, something that we did differently to others was that we did not mandate the data. We did not mandate the data models. We did not mandate data governance. All we mandated was the structure of the instances in Oracle, and then we let the businesses do whatever scheme they wanted. They had full freedom to make any scheme, and store and manage their data any way they wanted.

The premise behind this was that people were already using data. They were already making business decisions. It might be the right data, it might be the wrong data, it may be less efficient or more efficient, but they're already doing it. If I come in and try to tackle that on day one, we'll never even get off the ground. But if I give them a platform that

> "OneMap was designed as a citizen platform from the beginning: it was a platform onto which anyone could build apps as they needed them."

they could put their data onto and start to use some of the tools, maybe that candy approach will pull them along: the carrot instead of the stick.

And it worked effectively.

People who've been working in desktop on their C: drives or network drives found those very slow and couldn't find their data. We now gave them a home where they could put data into a database, which meant it was faster, it was backed up, it

was accessible to more people. They could publish it as a service, which meant they could share a web map instead of a PDF or a PowerPoint.

Power to the citizens

OneMap was designed as a citizen platform from the beginning: it was a platform onto which anyone could build apps as they needed them.

The heavy users, the 'front runners' were pushing the envelope. If they didn't like exactly what was available, they built their own web apps, they built their own FME jobs, they built their own database tables. They could, literally, do whatever they wanted. They just couldn't change the core platform: they had to play within the core platform.

Furthermore, we have been able to upgrade seven times now with the users not even knowing. That's because we were not writing code, it's all low-code configuration. Those apps that we built are still working now - if I had to give it a percentage of apps that are out there, 95% go forward every time we upgrade without even the slightest problem. And, typically, the 5% that don't, they just get thrown away and someone makes something new on the new platform - they are citizen developer apps.

It was game-changing, because it allowed us to put simple mapping in the hands of 70,000 people, and modules to make advanced tools for the people who needed it. So from one platform I could serve all those use cases: in a big oil company, especially in the GIS teams, we all think everybody needs a full-fledged
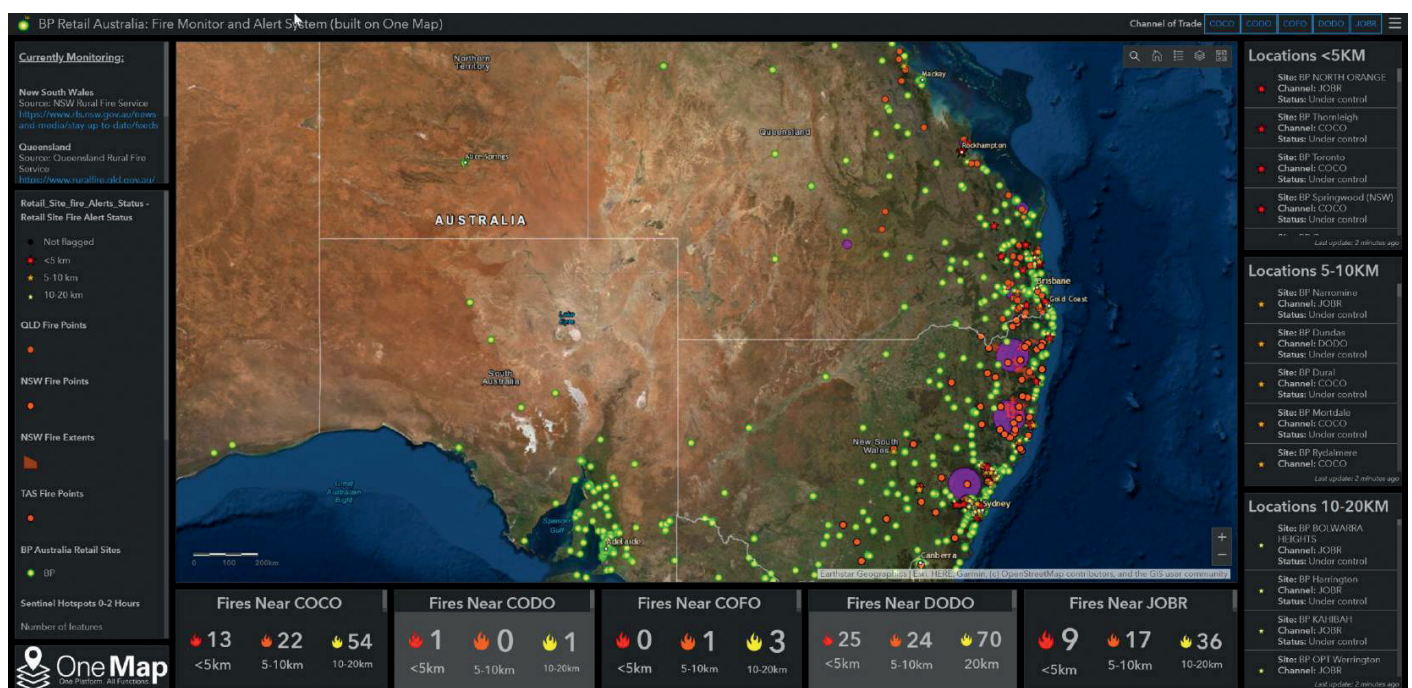
BP Retail Australia's Fire Monitor and Alert System (built on OneMap).
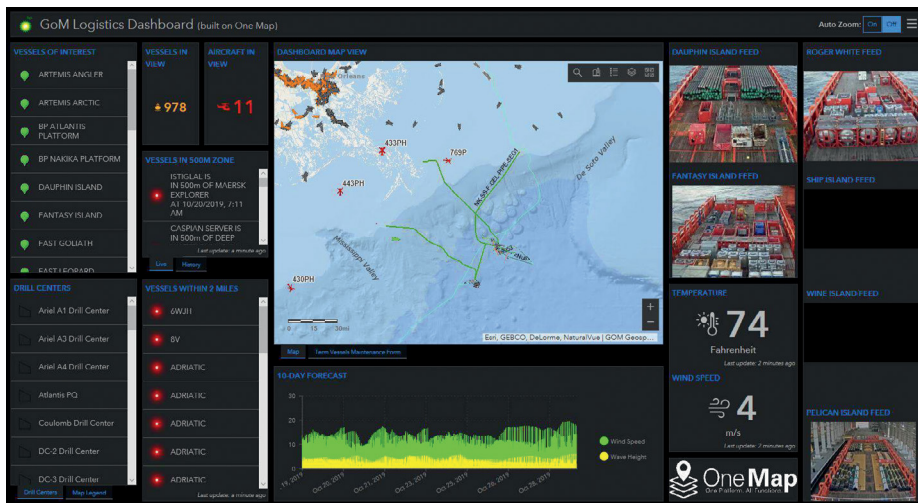
Real-time video feed enables OneMap logistics users to spot available boat deck space.

GIS with all the complexities. The reality is, we have more users who just need a picture of a map. They don't actually need scale, they don't need location, they don't care about any of the finer details, they just want a picture.

That's great, but I also needed to deal with the people who needed engineering-level information.

The portal let me do both on the one platform. So, if the person just needs a picture, then they use the picture, but if they want to throw on leases and pipelines and wells, then they would just lay them on and it would work. And that, I think, helped us, because those people who only needed a picture won't install ArcView. They're not going to learn ArcView. They're sure not going to figure out how to get the first base map into ArcView. But in a portal, they don't have to install anything. They didn't have to take any training in the first place, as the base maps were already there.

The engineering-type information, the high-level information sits on the same base map, the picture users just don't see it.

How does it work? If you go into one of our portals and search for a pipeline for instance, you're going to get two sets of results. You'll see some pipelines that come from vendor data, where the accuracy level is very basic, it may just show a line across the country. And if all you need is to show a pipe crossing the country, you might use this one because it loads really fast.

If you actually need the engineered pipe, there'll be a PP pipeline layer, which is that same pipe at centimetre accuracy. And if this other view of the pipe loads just as fast and it's accurate, then you'll actually start using the better data. Historically in GIS, you didn't use the better data because it was too slow, too many features, too many vertices. But in a web environment, all that heavy lifting is happening on the background, so that

the accurate pipe is actually almost as fast as the vendor pipe now. Now people will start to use the real data instead of the vendor data. And that's exciting because now we're putting the right and most accurate information in front of users.

Three degrees of separation

When we did our design early on, we wanted it to be as simple as any online map to use. We wanted the portal to be at least that easy, and preferably easier, because in ours the data layers that users need are already there, which aren't in online maps.

With 15,000-plus active users, we had to make sure that important data is never compromised, so we created three layers, or classes, of data: project, authoritative or corporate, and control.

Project data is a result of projects that are still in development. It can be new projects, it could be intermediate data, and analysis data sets. It could be somebody trying to do something and making some changes.

Then we have data that's called authoritative data, that is data that's been checked and is managed but may or may not have spatial accuracy or attribute accuracy. It is an up-to-date version of that data that's available, for example, vendor data.

The third layer is the controlled data, and it's tagged and shows up that way. That's data that's been through very rigorous checks and balances and is deemed best available.

The way to look at it is this: in the portal, you can trust control data without question. With authoritative data, you should check how you're going to use it to make certain it makes sense for your use case. And for other data, you use it at your risk. Sometimes this other data is actually good data, as long as you know what it was built for, what the limitations are. So, I still want people to find this data, but

most often users pull from authoritative or control data.

In terms of quality control, we don't drive standard models or standard templates. Everybody gets their own scheme up, so it's actually impossible for somebody to mess up data that's already there. The worst thing that can happen is they create a duplicate data set that may or may not be as good as one that's already there, but they can't actually hurt data that's there.

That's protection layer number one. Protection layer number two, if data sets come in and they're going to be part of that control dataset, we have a lot of processes, both personal as well as automated ones that get data from project to authoritative to controlled. But that only happens on those items that end up in the controlled space, which is a very small subset of data.

In the authoritative space, there are fewer checks and balances, but we still have some: you've got to have basic metadata and you've got to have documented load and some detail around freshness, things like that. We have data managers typically working on those.

And on the project side we don't control anything, on purpose. We want people to feel safe to come in and do good or bad things. The entire premise of our approach is based on monitoring behind the scenes. It's passive monitoring that we act on if needed.

What we have found in the big company, if we put any even perceived hurdles in place, yes, people always go back to what they used to do, and the candy model won't work. So, you can use your data without permission. You can do dumb things without permission. You can do great things without permission. If we find something that doesn't feel right or look right, we have a chat with you. And if we find something that looks wonderful and needs to be promoted, we also have a chat with you and, this being a citizen dev model, you're free to put in your data for anybody to use. It even becomes a little self-healing, because people put a little more effort into it.

Not only that, people feel empowered because while they can still own their data set in this published world, now there are five people or five teams using it. They have just become more valuable to the company by sharing their data, not by hiding it. So those people are publishing even more data and sharing more things because they see that while hiding may have made them important to a few people, publishing made them important to the whole organisation.

Brian Boulmay is BP's Global Geospatial Information Lead ∎