Tegus Transcript

Vannevar Labs - Director, Defense & National Security at Dun & Bradstreet Corporation

Interview conducted on September 23, 2022

Director, Defense & National Security at Dun & Bradstreet Corporation. Expert a government employee for 11 years and an intelligence analyst for 8 years. Expert has worked with multiple clients to get an understanding of requirements and pain points of why they chose specific open-source intelligence platforms with AI/ML capabilities in place of others.

Director, Defense & National Security at Dun & Bradstreet Corporation (The). The expert is responsible for the partnership with D&B to help the Defense & National Security communities unleash the power of D&B analytical tools and underlying unique data sources to defeat national security threats.

Prior, the expert was the Senior Principal Success Manager at Mulesoft, leaving September 2021. The expert was responsible for building application networks: seamless frameworks of applications, data sources, and devices connected by APIs, whether on-premises or in the cloud.

Prior, the expert was the Director - DoD Programs at Babel Street, leaving May 2019. The expert was responsible for leading the Department of Defense for 3.5 years and owned the entire portfolio for the user base, expansion, and pre + post sales.

Q: Can you speak briefly about your ability to help our client better understand if and how the ML industry is servicing critical defense/intel end markets?
A: Right now ML is essentially helping to harness the immense amount of data that is used throughout intelligence analysis and analytics. The defense/intel end markers are looking for effectiveness and efficiency to get through so much data in a targeted manner.

Q: Can you speak briefly about your ability to help our client better understand the evolution of ML in critical defense, the solutions currently available, and any potential emerging technology to be on the lookout for?
A: Sure, this is one of the things I've been doing for the last 8 years, I've tracked about 30-35 platforms that have straight data sources to different machine learning and AI capabilities based on known requirements from these government companies, so I can speak very well about solutions currently available and potential technologies.

Q: Can you speak briefly about your ability to help our client better understand the capabilities and positioning of critical national defense ML solutions, such as In-Q-Tel, Vannevar Labs, Anduril Industries, etc? Have you ever evaluated and/or used any of the aforementioned providers?
A: I've used Palantir, I2, ShadowDragon, DataMiner, Battlestreet, and many others. I can speak about lots about capabilities and positioning of critical national defense ML solutions, and their differentiators, market share, etc.

Tegus Client

Hi, thank you for taking this call today. So the quick hit is I'm trying to understand more about the landscape of software and technology usage buying and in the DoD and to the extent possible the Intel communities, but I know high side is a little more difficult to tackle and then kind of companies that you've seen be successful in the space and how you would kind of evaluate that competitive landscape.

So we’re mostly interested in kind of the analytics and the open source Intel side. So kind of everything from collection, aggregation, modeling and then sort of visualization UI analytics output. So maybe as a starting point, I think before we jump in, I always think it's helpful to hear a couple of minutes on your background

from you and then we can jump into sort of the topics at hand.

Director, Defense & National Security at Dun & Bradstreet Corporation

So I got my Bachelors and Masters in International Affairs and Security, Georgia Tech, worked for the National Archives as a Research Specialist while I was in school once I got out, served as a DoD all-source Intel analyst and Targeter for about eight years, with the majority of my focus on AfPak region, Iraq augmented.

U.S. SOCOM spent 2.5 years deployed off and on throughout Afghanistan. Left the government after 11 years and went to work for a Software-as-a-Service company, I worked for Babel Street for about 3.5 years and served as their DoD and IC portfolio leads, and I was everything from demos and training presales to essentially account management, post sales and the expansion for those.

Quarter life, professional crisis left the start-up, went to a little bit of consulting on OSINT and targeting type stuff for a few months went to Salesforce MuleSoft and worked on the API management and integration side of data.

Ran their post-sale stuff for the DoD, Intel community and aerospace and defense. And have found my way back into more of the data and platform round now. Over the last six years, I've kind of done some consulting as it relates to open source intelligence platforms. The competitive landscape, the changes that have occurred from say, 2010 through now and kind of what different federal government organizations requirements, pain points are kind of what they've been looking for and how big have changed over that period of time.

Tegus Client

Awesome. That's a perfect intro, and I think actually helps me structure the call a little bit. So I think maybe we might spend the first part, a little bit about your experience as an end user and kind of what you found as sort of key use cases for technology. Or these specific uses of technology in your day-to-day key pain points that you found over the years. Clearly, the tech landscape has evolved a little bit, but I imagine there's a lot of overlap. And so let's start on the end user side and then we'll move over to kind of more of the buyer side.

Director, Defense & National Security at Dun & Bradstreet Corporation

Sure, so while I was in the government, right, the first few years, that was essentially just an analyst working on a team, right? And so initial all source training stuff and biggest gap at that point in time, say, 2010, 2011 was there had been no real or true integration of open source intelligence data, right?

Like everything that was being used was mostly things that have been collected on the high side, right? And so the situational awareness aspect around OSINT was a gigantic intelligence gap, known requirement, known need.

And so I was one of the analysts on my team and then as I moved into more of a leadership role was actually betting the types of platforms in terms of open source intelligence collections so social media, open-source news type information, the beginning of what we now refer to as big data and platforms that help you do the collection, but then also the analysis of that data that you were bringing in.

And so it is the combination of that on one side, but then also looking at essentially network building type platforms in the intelligence sense, right? So i2, Palantir, Semantic AI, DataWalk, DataWalk came in later, right? But those are the main platforms that everyone throughout DoD, Intel Community, federal law enforcement uses on that side.

And so initially, everything was centered around data, right, as an intelligence analyst, the more data, the better, right? Like I want it all, give it all. So when you're talking about social media, and you're talking about Twitter and VK and regionalized type platform, you're talking about potentially millions and billions of pieces of information.

And so the analytic capacity on the back end of these platforms that are allowing you to do focused

collections were an on national like you're trying to find a needle in a stack of needles. And so you need to be able to sort through that information in order to get to what's truly important to you.

Tegus Client

Yes. So you're talking about sort of 2010 context and sort of the adaptation of that. How far would you say that has come in recent years? And I'll sort of break that down in a couple of ways, right? If you think about sort of quality of UI and ease of use. If you think about sort of model capability, integration of ML, et cetera.

And then probably the third part is potentially triage of that data collection at the outset, whether it's better labeling or aggregation of better sources or things like that? And maybe if you can talk about all three of those and where they sit in relation to 2010?

Director, Defense & National Security at Dun & Bradstreet Corporation

Yes. Sure. And it's interesting when you look back over the last, what is that, 12 or so years at kind of how everything has more changed and a lot of it, especially in this space, is solely based on the requirements of a few organizations because everything else tends to fall in line behind those, right?

And so on the DoD side, SOCOM is the driver. Everybody knows that. They drive what the requirements look like and everybody else tend to fall in line. And so initially, it was oh my God, we haven't had this type of situational awareness data, right? If you're on the DoD side, well, if we were rolling into some area in Afghanistan.

The only situational awareness we might have had was either humans that we had paid sources to go out and collect or any other information that convoys and conventional soldiers had been able to get when going through that area, right, indirect fire types of incidents. And so it was such a large intelligence gap.

Initially, it was just get me anything that can get me as much data to give our guys enough situational awareness to know what they're walking into. And so some of the initial platforms in that time frame, right, so 2012-ish to, let's call it, 2016, right? You're looking at the Babel Streets of the world. You're looking at the OpenIO, Sprawl.

What was the other big one. Crimson Hexagon was one, where they had access to the larger platforms worth of data, right? So initially, the Facebooks and the Instagrams of the world and Twitter. And then privacy starts to leak in here around that time frame and only has progressively become something that is going to truly mold and shape what this looks like moving forward.

And I'm sure we'll talk about that later. So it was access to the data from these large-scale platforms that were being used globally, right? And so the initial push was centered around the data and honestly less around the analytics and the visualization. But there are some key points that hadn't really been thought about before, I think, right?

And if you're searching in English across all of these different applications, you're only going to get, the Internet is essentially 50% in English right? But if you're only searching in English, you're only going to get 50% of the data, and if you're talking about the DoD and the intel community then you're going to miss out on a large portion of what it is that you truly want because they're not doing the majority of their stuff in English-speaking countries, right?

Though it is a popular language worldwide. And so the ability to be able to create searches in English and then essentially bring back post in the native language but then be able to translate and not necessarily a foreign speaker translation, but essentially an 80% translation that's going to give you the gist of what you need.

It was very, very important. And so that was something that platform started had either brought in initially because most of them were built by former Intel guys who saw the need and the requirement. But then it's, okay now that we have access to the data that we need right now, not only grows regionally and everything else moving forward.

But okay, now we're dealing with the language piece. Now we're dealing with millions and potentially billions

of records. How do we effectively search through the had information to get to what we need in an effective and an efficient manner. Because that's what these platforms truly do.

They allow what somebody could do themselves and would take them years to honestly be able to do it in hours and/or minutes to get to 100 documents or post to look at as opposed to one million, so that was the initial push. But once you get to about the 2018 time frame, there was kind of a divergence with all of the different open source intelligence platforms that are in the arena, right?

So you had some like Babel that were honestly jack of all trades and master of none. So they were trying to do all of these things. But then you had other platforms that picked one thing and were the best at that one thing. Niche type platforms like M3S is one. I think it falls under Raytheon now. But what M3S did is they had the best capability at being able to take audio in any language, translate it and put it into text.

And so depending on the requirement for a given organization, they may need that a lot more than they need some of the other types of stuff. And so it was a division of trying to be the one platform to rule them all which is what everybody wants because then everybody is buying your product versus being the best at one or two things, and that could be an analytic capability, that could be visualization of your analytics and results, or that could be an alerting capability.

I want something that is going to allow me to know as soon as something happens. Or is it something that's truly going to provide situational awareness and be able to provide me with a risk assessment of a given area that somebody might be going into or that we may have to pay specific attention to.

And so they've diverged from the focus on just data into a host of those things, and that was really driven by requirements from SOCOM, and that divergence is really data and visualization but the machine learning and analytic capability that allowed for quickness to an answer like efficiency and effectiveness, I don't need a 100% answer, I need a 70% answer right now.

Tegus Client

That's interesting. Because you sort of do hear and I actually have a little bit of history on this stuff. So I recall back in the '16, '17 time frame, just being like, okay, give me everything and not that kind of bleed towards no, I need sort of a recommendation engine style UI that allows me to get most of what I know very, very quickly.

That's interesting. So I wonder a couple of things on what you said. You mentioned Babel as a jack of all, master of none. It was actually platform potential. Is that how people want to buy? Or generally do they prefer single point solution? Or is it kind of all take whatever is best how do they think about that?

Director, Defense & National Security at Dun & Bradstreet Corporation

So that's a question I feel like has a lot of daggers in it, right? Because if you take out the data collection piece, what is the platform that when you talk about that comes to mind? Because there's one that comes to mind for me, and it has a very bad taste of everybody's mouth.

You look at Palantir, right, which has an amazing product. I mean, before that, everybody was using i2, and that was it. But they ended up dealing with a lot of issues as it came down to the data that was put into Palantir and whose data that was. And this is another step in the direction of data and privacy type issues, not just an open source collection, but like as you look at it for the U.S. national security arena.

And so they wanted to be that database for your unclassified and classified information that you could move up to the high side and bring everything together, the ability to search through anything you put in there and build out these networks. It also has the machine learning capability to help you with the types of search and the types of filtering, but it ended up leaving such a bad taste in people's mouth.

However, they have no choice but to use essentially one of two platforms. There are some other smaller ones that have teeny tiny pieces of the pie, the Semantic AI, the ShadowDragon, the DataWalks, but you're either an i2 family or you are a Palantir family. And which one is it going to be? And you look at things like U.S. SOCOM, which again, has always been the front-runner and trend setter as it relates to open source intelligence collection.

And so with enough research there are potentially, we'll call it a sandbox where an organization puts together and goes out and does the research based on their known requirements for an organization like SOCOM and put together a whole bunch of tools and says, hey, here's seven tools. They have some similarities. They have some differences.

Some are better at one thing; some are better at another. Let us show you each of those tools and then pick one or two or all seven, and purchase access to those essentially as a sandbox or a test bed to figure out what's truly needed for your given organization. The piece on top of that comes to pricing of the products, budgeting for the organization and the red tape involved with government procurement.

Is the money coming at an enterprise level from a federal law enforcement agency? Are we going to buy 1,000 accounts for all intelligence analysts across pick your organization? Or is this a steeled office or a specific intelligence unit that has their own budget to buy so? Because then you have to rack and stack.

Well, this platform has everything, but it also costs 7x as much as this other one, and I can get a 50% or 70% answer from the cheaper one. So that one that does them all might be awesome, but they've priced themselves out of my ability to be able to purchase it. And also, hey, that looks great. Yes, of course, I want a BMW, but what I really truly need right now is the Toyota Corolla. I need to work course.

Tegus Client

Yes, super helpful. And I want to come back to the buyer point and buyer across this point in a second. One sort of, I guess, macro question, you mentioned being involved in AfPak for a long time. I assume subsequently, what you've been doing is maybe a little bit more peer competition. How would you say that, that aggregate shift is sort of changing any of these topics in terms of how we think about software needs, software buying, Intel collection and the answer may be not that much, I suspect not, but that's the very per answer as well.

Director, Defense & National Security at Dun & Bradstreet Corporation

Actually bigger than you think and especially in the last, let's call it, six months to one year. Since 9/11, we as an intelligence community, I mean even since the fall of the Soviet Union, our focus has shifted away from state actors. They're never completely gone, but like they haven't been the front runner or the primary. And so since 9/11, the intelligence community's focus has been around, nonstate actors, terrorist organizations.

But in the last year to two years since we've pulled out of Iraq and pulled out of Afghanistan completely, that focus has shifted, not completely away from nonstate actors and terrorist organizations, but back to state actors that haven't been a focus as much for the entirety of the community.

So Russia and the Ukraine, China, right? And so it's not that the requirements of the needs for these types of platforms are going away or wasting or waiting, but they're just shifting to something else. And so it's almost like data sources come back into play where they kind of fell off a little bit because now there is an even truer need for a very specific regionally focused social media platform data access.

The ability to effectively and efficiently translate that stuff. That's definitely not in English. The way in which people talk and the way in which people search for things are completely different. If somebody is going to target the President of the United States, they're not going to say, I'm going to kill Joe Biden or they're not going to say, I want to slap former President Trump in the face.

They're going to say, hey, I'd really love to punch the 45th Cheeto right in the nose, right? And so the use of an ontology also kind of feeds into that machine learning AI search capability to ensure that you're not missing post or headlines or anything that could be derogatory in nature and be what you're looking for.

So typing in a search saying, hey, I'm looking for post related to President Trump and this and that. Well, it's not just going to search for straight text President Donald J. Truck. It's going to search for all of the ridiculous things that people have called him in frustration, anger or whatever to ensure that you're not missing the type of results back for you to be able to sit through. I think I took myself off track there, but I think that, that's a very important.

Tegus Client

No, but I appreciate that. And to your point, you've mentioned a few times on the call the idea for not needing to get to 100% accuracy, but sort of thinking 70%, 80%. I imagine that in this context versus sort of a AfPak counterterror whatever, where you have very sort of contextual real-time maybe even kinetic context, accuracy of information probably is a little more necessary, very, very specifically, where they missed your point. It's sort of a gathering process that is more sort of a long-term longitudinal trying to understand where the world is going and less about what's happening tomorrow essentially.

Director, Defense & National Security at Dun & Bradstreet Corporation

Yes. It's such an interesting space, right? I mean you look at kind of where platforms have gotten to now. It's almost like they've actually gone back toward niche capabilities. You look at those. If I were to rack and stack them. You look at the five or 10 companies that have either been around the longest or the ones that are the most used or requested by DoD, IC, federal law enforcement.

And there are still a few that are trying to do all of the things. But the ones that have moved up faster in the last three years are ones that are doing specific things very well. And I really don't think that there is an affliction by national security organizations to say, look, I only want to buy one platform.

They have different requirements for different taskings or operations, and so they're going to need access to multiple tools. And that's where it makes it really easy to price yourself out of the market if you say, well, I'm the friggin BMW. In the Ferrari, you want me and nobody else. It's hubris, right? Like that's just not what the market looks like right now.

Tegus Client

So yes, I mean, to that point, let's maybe shift a little bit to the buyer side. and one thing I would love to hear about is you mentioned multiple times, SOCOM at the leading edge. I would love to hear sort of a couple of minutes on what that means contextually and why that has come to be? And yes, maybe we start there, a bunch of questions on this front, but that one was interesting.

Director, Defense & National Security at Dun & Bradstreet Corporation

Sure. I mean they've been the lead for a long time. I mean they have the money and the budgeting capability to honestly get money essentially because of Afghanistan and Iraq that they've had a lot of called OCO funding.

So funding for things to deal with the front that we were fighting on. And SOCOM is, you have Air Force, I know you've got Marines in there. I was deployed with somebody from every service as I was sitting out there under the umbrella of SOCOM.

And so those people have access to their own systems and tools from their own units. And so it really just was this melting pot of military intelligence being able to look at, deal with this current and newer problem of essentially asymmetric warfare.

And how to tackle that and also bring in everything that was the new hotness at the time as it related to social media. Because you can very easily track how these organizations, these non-state actors moved from making calls to each other to, that was no longer safe for them to them using e-mail and then that was no longer safe for them to them trying to hide in the white noise of large-scale social media, Facebook, Twitter.

That then was no longer safe for them, and they move to encrypted end-to-end communication like Telegram and WhatsApp. And so SOCOM budgeting capabilities, ability to put these types of tools that were that needed intel got into the hands of people at the front is what has led them to be that big brother to follow after in terms of targeting cycles, intelligence analysis and a full picture of what can be used, what can be done, what right looks like because they're trying to test what's out there.

Tegus Client

Totally. And I'll make an analogy that hopefully makes sense. Given sort of that hodge podge that you mentioned, given sort of multiple constituencies, all or many with owned budgets, even if it's more experimental. Is there sort of a bottoms-up motion that takes place in SOCOM? Or is typically the buying power hey, one unit tries something, it's kind of interesting and then the contract at a much higher aggregate level? How does that sort of genesis of sale typically happen with SOCOM specifically?

Director, Defense & National Security at Dun & Bradstreet Corporation

That's an awesome question. And it's very interesting SOCOM specific because it's usually a two-pronged approach. It's strategic from the top down, and then it's tactical from the bottom up. Because the typical people that are going in that are doing your business development, that are doing your sales, that are doing your account management.

These are all smaller companies that we're talking about. If you take Dataminr out of the picture and you look at GeoSpark, Babel Street, EchoSat, ShadowDragon, 35 others, they're smaller companies. And so the people that they've hired to fill those roles are subject matter experts because they've come from the community, they've come from federal law enforcement, that come from the military, that come from the IC, they know what the requirements are.

They know what the needs are, and they speak the language of the user because they were one. And so they reach out internally to their network and say, you've got to see this. You got to let me show this to you. And then any one of those companies typically goes and grabs somebody who formally had stars on their lapel as a big, huge rolodex to say, I need you to get me in front of the SOCOM J2, right? I need 30 minutes or an hour of his or her time so that I can present this to them. And then they'll push them down to other meetings. They'll say, well, yes, there's obviously a requirement and a need for that.

You talk to the people below me, but then you already have it coming up from the bottom where you've attempted to set trials in place to get your platform in the hands of some of these users. Now neither one of those are the two correct path for how you're actually supposed to present technology to an organization like there's legitimate form to fill out that says, I have this new technology would love to be able to write a white paper about it, present it to essentially a technology board and then have them kind of put feelers out to that organization's community.

But like that takes a long time. And when your job is selling things, well, then let's go three from to do all three and see what gets us there faster. But the Perl strings are held by the people at the top, but I can tell you the number of times that I've demoed a platform to a general officer. Or somebody from the senior executive service with other GS-9 and 11 analysts in the room, and I can wow the socks off of the person with the Perl string, but you know what they're going to do immediately is turn to the GS-9 and turn to the GS-11 and go, is this what you need?

So it's important to come from the tactical and be able to get in front of and know what the analysts want, need, desire, but also have something that you can show that impresses at least something upon that person that's actually going to be holding the Perl string especially in SOCOM, right?

Tegus Client

And like in rough numbers because I don't want to get into too much sensitivity, but like when we're talking about either sort of individual teams or units within SOCOM and then SOCOM more broadly, how would you think about sort of spending power or budget on things like a recommendation engine, discovery platform, analytics, things like that, these kind of software products, if you were to sort of put market math on it, are we talking a couple of hundred million dollars, are we talking about several billion? Are we talking huge? How do you think that?

Director, Defense & National Security at Dun & Bradstreet Corporation

That's a horrible question because especially when you pick your organization, right, N-Com, SOCOM, whatever. N-Com right for the A2 for the Army. It 100% depends on who's sitting in the seat. And so when the current SOCOM J2 came in they didn't have the same thoughts on OSINT in general and budget for OSINT platforms. And so something that the credit tester worked so hard to bring in and integrate and secure

money for is the first thing to get cut.

So it's insanely important to know the likes and dislikes of that senior intelligence person that's coming in. Are they geared more toward the three in the operations side of the house? Or are they the TrueBlue intelligence officers and analysts who want to ensure breadth and depth of data and situational awareness?

Or are they more toward the upside and want, hey, give me an answer now. I need a 50% answer in five minutes, not 100% answer in two days. But like as it's kind of wax and wane and to be honest, I'd say it's gone down some since we've pulled out Afghanistan and Iraq because a lot of the funding that was being used was funding specific to those engagements. Because it was easier to get that quickly to pull from essentially a pot of money that could be used for anything for that as opposed to having to put traditional line items in the budget 18 months out.

But if you look at some of the large-scale contracts that have been won, so the cattle contract I think the three big front runners for that were Dataminr, GeoSpark and Babel Street and I mean, you can find this if you search if you haven't already.

But like it ended up going to Dataminr, it was contested at least twice by one or both of the other organizations. And that's like a 250-plus million contract for five years. And it literally just put the Dataminr platform, and the name of their platform is what first response.

And so they are the front runner for first alert. So for an alerting capability, like your focus is U.S. Paycom. I need to know immediately when anything happens in the Paycom AOR because I'm an analyst that's focused on North Korea, right? If they fire something off, I need to know immediately and you're going to get that the most quickly, honestly through social media. But that is a big contract, right? I mean, $250 million over five years.

But I would guess that the majority of 30 companies on the list that I track and go back to like Dataminr is the 800-pound gorilla in the room. They took $330-plus million in funding in their last round. They're in family well backed. I think Twitter has a 5% stake in Dataminr, which I would only assume gives them excellent access to data from one of the primary sources. But the rest of the company, I can't imagine that there's another one that tops $40 million, $50 million total revenue.

And a lot of these don't have the commercial requirement like the commercial side of the house doesn't have the same types of requirements. So the majority of sales for any of those platforms is generally into the government unless it's an executive protection type use case and then even in Amazon, it's only going to buy a couple of accounts.

So the last after that while you're thinking the integration of data. There has been a drive, and it's happened in the last two years. All of these platforms, most of them have not been resellers of the data that they're bringing in.

They're either collecting it through developer APIs like with the Twitter developer API or they're going out and scraping the data. Because they're not a reseller of the data, they don't have the ability when an organization comes to them and says, hey, that's great.

We love the analytics that you have in all this, but we want to be able to ingest all of the data that you have access to, to be able to bring it into something we're building ourselves or something where we're pumping all of our data into to be able to look at everything in the same place be that on the low side or be that on the high side.

I mean the end all BI would be able to take all of this open source data, air gap it up to the high side and combine it with all of that classified collection and information you have to build out sift through and build out networks with fully rounded information.

And the problem with the majority of those platforms is that they haven't had the ability to do that, they can't. Because if they do sell that data, which is what these organizations want now, and there's a very clear push toward wanting to have access to platforms where they'll open the back end and allow them to use the data.

And so not having that is making it so that organization’s usage numbers will decrease in the number of accounts that they're procuring will decrease. But at the same time, that is one hell of a scary thing. Where

does that data go. Okay, awesome. We'll build the integration, open up the API, pump that data in, well then do they really need you anymore if they have all your data?

Tegus Client

Yes. No, that's helpful. I want to actually go back to something you said earlier that we didn't touch on, but I think interplays interestingly with this were just the question of privacy, which I know has become more and more impactful.

As you think about that macro shift from counter terror to sort of great power competition and some of the sort of foreign action that you might be working with here. You talked about a platform like Twitter, which clearly is going to be impactful here.

But if you look at something more like Weibo or some of the Russian social networks or things like that. How do you think that calculus changes in terms of data aggregation? There's obviously an offensive element to that, that's a little bit complex. But that would strike me as being potentially and slightly different call it, problem or context where data access might be less of an issue? I don't know.

Director, Defense & National Security at Dun & Bradstreet Corporation

Well, so I'm going to take the last part first, and then I'll go back. So talking about the Weibo's or the VK of the world, right, or even smaller, more strategically regionally focused data sources, right? To be honest, there are a lot of data aggregators out there.

And for those I mean, I've seen firsthand, or heard stories where smaller, 30 40-person start-up in this type of platform category, it’s a lot of the data that they try and go out and get it built on requirements of an organization.

They come to them and say, hey, I don't have this. You give me that, I'm going to need to buy more accounts, right? And so when you talk about some of those Chinese and some of those Russian-based social media platforms, like the horror stories are CEO and Chief Strategy Officer, actually meet with people from the company, and they request such ridiculous things like, hey, I'm going to need Mr. CEO, your Social Security number, your copy of your passport, right?

And I wouldn't give that type of information up, not that pick one or two countries don't have the majority of that information already, but in terms of that, going through other essentially third parties and data aggregators I mean they've made deals to get either to get access to the data or have found a way to go out and get it themselves.

In my opinion leaving plausible deniability on the company because they're not the ones collecting it. They're just buying it from a third party with access to what a client needs. Going back to the privacy piece, it is something that's come up repeatedly, right? Pre pandemic, oh my god, I always forget the Senators name. It starts with a W.

Tegus Client

Wyden?

Director, Defense & National Security at Dun & Bradstreet Corporation

Yes. Thank you. It is Wyden. Pre pandemic, huge push on talking about the use, the collection, the use of this type of data. We, as Americans want our privacy. We have the right be the first amendment to say what we want to say, but we want to be able to put that out there to as many people as humanly possible, but then also don't want to have that used against us in any manner.

And so the legislation that is currently bounding the use of open source intelligent social media data, pre-date to the Internet. It was written in the early 80s and so this has been coming for a very long time. And it only gets further enforced by the use of information like AdTech data. But you've seen in the last couple of

years, maybe the last 1.5 years. So the pandemic slowed that down. It's slowed everything down.

But you're starting to see it come back up now, especially with the legislation that's gone through the house and going to the Senate on privacy of this type of data. But companies like Apple and Google have already foreshadowed what that's going to look like. When you talk about AdTech data, essentially somebody embeds an SDK in on App. And so do you read all the terms and conditions of use on a free app that you download from the App Store? Actual question. Do you? I don't.

I'd ask that question to literally hundreds of people, and I've only had two people tell me, yes. One was my former roommate from college, weird kid, but I guess I'm not surprised. And the other was like a former SES from NSA, right?

Nobody else reads it. It's long, it's in legal lease, but like deep down in there and the free app says, we're going to sell your data. And people are like, oh, well, it's anonymized data. Sure. There's a 32 digit hexadecimal, anonymized whatever for the device that it came from, but give that to an intelligence analyst, and I can tell you who it is if you give me like 30 minutes.

And so the direction of that's going to go is being foreshadowed by Google and Apple especially with updates that they've made to app updates and downloading apps now. I mean I have Apple everything, right? And so I download a new app. And what's the first thing it says as soon as you've installed it says, do you want to allow this app to track your location, right?

No. Well, if you're doing in a free weather app, well, that would seem dumb because then it's never going to be able to tell you what the weather is, where you are, right? You'll have to go in search, and everybody wants an easy button, right? But if it's the Home Depot app or Candy Crush, whatever it is that the kids are playing today, right? Like why would that need to have my location data.

And so with them, essentially, and you saw Zuckerberg go crazy on Apple when they instituted that because they get so much money from that type of data, and those types of marketing campaigns because that's what it truly was.

It was marketing data that when people on the intelligence side, found out what it could do, it would essentially make it easier than trying to obtain actual phone records or having to get a warrant for something because, well, this is all anonymized and how could you ever, right?

So it will go that way. I mean the privacy issue; I truly believe will go that way. And that will make it very difficult for a lot of these types of platforms. Especially for ones where it allows you to dig into a username or something very tactical and specific versus alerting need to a natural disaster that has happened, like a Tsunami off the coast of the Philippines or an earthquake in Peru or lightning strikes across the Gobi, right?

Or something where it's looking at aggregated social media information to provide risk assessment. Those are not going to have the types of privacy issues that's something that is allowing you to dig into all of the posts from a given user for the last four years.

Tegus Client

Yes, makes sense. So I guess my follow-up question would be besides Dataminr, which you mentioned, and obviously, like Palantir, i2, we talked about a bit. But besides Dataminr anybody else that you think is particularly noteworthy or generating a ton of buzz momentum, anything like that? Or are you finding generally that like, hey there are folks that are very good at one acute thing, and they will continue to be good at that one acute thing, but probably not going to scale that much.

Director, Defense & National Security at Dun & Bradstreet Corporation

I mean it goes to the niche aspect, right? So like I mean Dataminr always the 800-pound gorilla. They have a gigantic contract. GeoSpark, which is more centered around like risk assessment, and risk analysis of country level, region and cities. I mean they have a government-wide IDIQ which just makes selling it very easy. ShadowDragon is one that a smaller company, maybe 30 people, but it's been around for, I don't know, five, six years, I think.

But in terms of building out networks and access to data not bring your own data, they've made a really big push up the lift in the last couple of years. Let's see what else is on there. There are some that have definitely gone down.

Fivecast and Skopenow are two that have moved up were either nonexistent four years ago but have made their way up to the top. Fivecast is actually an Australian-based company, which has been making pushes into like federal law enforcement and the DoD. Recorded Future has been around for quite a while. They're always hovering around the mid-tier. PrimerAI, I would put them at the top of like the machine learning AI, analytics capability, but they're also more of a bring around data type of tool.

I would say the only other one I would say is Cobwebs, which has an AdTech feature, but they come with their own issues because I believe they're actually Israeli based and so that make selling into the DoD and IC a little more difficult.

Tegus Client

Cool. Super helpful. Well, listen, this has been fantastic. I really, really appreciate super thorough, and I love that we were able to cover both, as I said, the end user side and the buyer side. So thank you so much for your time. Have a good day.
