Geolocation data: The value of precise information for improved underwriting

Participants


Jean Sullivan, Vice President, Insurance Software Sales Americas, Pitney Bowes


Ralph Groce, Chief Information Officer, Everest Re


Megan Thomas, Chief Underwriting Officer, Reinsurance, AXIS Capital


Michael Reilly, Managing Director, Accenture


Matthew Bishop, Deputy Property Manager for the Americas, SCOR Business Solutions


Kristen Bessette, Chief Actuary and business owner of the North America Data Strategy, QBE


Kevin Livermore, Executive Vice President, Willis Re


Rob Sabio, Vice President, Head of Corporate Catastrophe Analysis, W.R. Berkley

"Close enough" isn't nearly good enough anymore in the complex process of risk selection and modelling. Newly developed tools designed to provide more trustable intel on a given property can make all the difference.

In March, *Reactions* Editor-in-Chief Shawn Moynihan led a roundtable discussion sponsored by Pitney Bowes on the topic of geolocation data, the company's advancements in this technology, its possible wider adoption as a trusted source, and its valuable applications for insurers and reinsurers alike. An edited transcript of that discussion follows.

Jean Sullivan: Most people sitting here might think, "Why is Pitney Bowes sitting at this table?" They probably think that we're the 100-year-old mailing company located in Stamford, Connecticut. We've been in the software and data business for 40 years. Over that period of time we have evolved into offering the most precise geocoding available in North America and internationally, and we continually build data sets to augment our software. We have more than 300 insurance customers, we have 42 of the top 50, and all of the top 25 carriers use our data and software. We're also used by a number of reinsurance companies.

So why are they selecting Pitney Bowes? What makes us unique? It's three key areas. One is the precision of our geocoding: we provide the most accurate geocoding globally. Two is what we call our pbKey. It's a persistent and unique identifier that is now allowing companies to be able to build the location MDM. Three, there are a number of companies that are in the data space but they can't operationalise it. We've got the capability to not only provide insights from the data but to operationalise it, so whether it's risk management, underwriting and/or cat modelling, we're helping with those efficiencies and processes with insights provided by the data.

Shawn Moynihan: We have a terrific mix of representatives from both primary insurers and reinsurers here today. I'd like to start by having someone from each side articulate how geolocation data is used in your own risk-assessment process so that we can understand better how both sides of the house benefit from its use. Kristen, maybe you and Ralph could get us started?

Kristen Bessette: Catastrophe modelling is a good place to start. We're always trying to better understand our exposure to natural catastrophes, whether it's hurricanes, earthquakes, flood, fires in California, tornadoes, etc. Where that building specifically sits can greatly alter your exposure to all those events. In the case of a flood, for example, just being off by a few feet can make a big difference in terms of whether the water is in your space or not in your space, and the difference between a company's mailing address and the physical footprint of its building can really make a difference in your loss outcomes. We want to make the

best decisions and to inform pricing, underwriting, risk management and loss control, and help our customers protect their property.

Ralph Groce: It's critical to have that data and be able to operationalise it, and to have a single source of the truth. In some cases, a company's actuaries could be working from one data set while the underwriters are working from a different data set. In the end we need accurate data that everyone has access to, that is constantly being refreshed and updated.

Shawn Moynihan: Michael, you mentioned to me that you felt it was time to rethink and reimagine property underwriting in general – that mixed exposures, imprecise data and an inability to see aggregate views of risk in real time have impaired underwriting to a great extent. Can you expound on that?

Michael Reilly: If you look at the problems that are occurring in property underwriting today, there is the question of the precision of the data itself plus the fact that we're now trying to layer in lots and lots of different data on top of that. A building like the one we're in right now, for instance, has different floors; it has different tenants. The traditional way that we [gathered information on] that was through traditional relational databases. We had an address or a latitude-longitude location, but that really doesn't explain the complexity of the different levels of information that's here or the different levels of information to consider. The reality is, in order to truly underwrite well you need to have details on the entire structure.

In rethinking how this is done, we think there are three different things that need to change. The first thing that needs to change is we need to change how we capture and organise data. We need to move beyond relational databases to more modern and sophisticated data solutions that can handle location data's unique structures. Second, we need to layer analytics and machine learning on top of that data to improve data quality and deal with the differences in completeness and quality of many of today's data sources. For example, we know different counties collect data at different times and different pieces; in one county, I get information on the frame of the building and I trust that because that county updates it every six months. I can go to another county and that county has it from 10 years ago and has probably never updated it, and so I should not trust that source to the same degree. The last thing we need is better ways to present the information to property underwriters in the time, format, and visualisations they need to make use of it. You need to provide intelligence to this data to know when to trust it, so that I can interpret it in my underwriting process.

By rethinking the core of how we organise the data, by providing machine learning to improve the quality of the data, and by then building our analytics to be integrated to the underwriting flow and then designing that underwriting flow such that we are advising the underwriters at the point of need and the point of feed, if we make those changes then you'll start to see some pretty significant improvements that you can actually bring to property underwriting.

> "Our goal is to have people think of us as a locations intelligence software data provider – especially with our pbKey, our unique and persistent identifier"
>
> *Jean Sullivan, Pitney Bowes*

CONTINUED ON PAGE 28

Kristen Bessette: With better information, companies can look to write profitable business in areas or industries where they typically haven't written before because it was considered high-risk. But now those carriers can be more selective, based on what Mike just talked about. To be able to have that type of insight really changes the paradigm.

Michael Reilly: Brokers are becoming more demanding; they want you to quote at a higher volume than you've ever had to do before. The underwriting's getting more complex, and yet the market requirements are starting to come to quote more frequently, quote more often. So you have to find ways to do that efficiently. Missed exposures are still often the biggest loss; 40% of the time, losses are directly attributed to something that was overlooked in the underwriting process.

Kristen Bessette: From a resource perspective there's also a claims application: when a tornado happens you're trying to figure out what your exposure is, where to deploy your resources. Anything that can help you manage that more effectively has a direct impact on your business and your customers.

Jean Sullivan: Speaking of claims, the other trend that we're seeing is around aerial imagery and how that impacts claims. What they're finding is that just providing the imagery isn't enough. What we're hearing from our clients is it's really critical for them to understand the parcel boundaries and the property attributes, so that when they look at the imagery they can know precisely where a loss occurs. Because the policyholder experience is so important to carriers today, claims departments make those decisions quickly, write the cheque, and then also reduce the number of adjusters that they have to deploy, especially when you have wildfires and other natural disasters in areas that are difficult to access.

What we've also seen, which is really interesting, is a consortium of eight to 10 vendors that has been established to provide imagery to carriers. We're working with the consortium as it relates to our pbKey, our persistent unique identifier, and trying to be the link between what they're doing and then what the carriers need from

> "In some cases, a company's actuaries could be working from one data set while the underwriters are working from a different data set. In the end we need accurate data that everyone has access to, that is constantly being refreshed and updated"
>
> *Ralph Groce, Everest Re*

data sets such as property attributes and parcel boundaries. One of our customers said, "We absolutely cannot get it wrong that we thought it was this house that got burned versus that house. There is no room for error on the claims side." There is a huge benefit on the claims side in the new technology that's available today.

The other trend that we're seeing is that carriers want to be able to have better insight so when they renegotiate with their reinsurance carriers they can explain what their methodology is, that they truly do understand where the risks are within their book of business. That puts them in a better position when they're renegotiating with their reinsurance carriers or customers.

Ralph Groce: Regarding the imagery, I think the before-and-after also helps, before the claims and post-claim, because we also need to verify that the tornado did in fact rip the entire roof off, that the roof was not damaged beforehand and a fraudulent claim may have been put in, for example. I think it also helps as well with the wildfires in that obviously if that roof is gone the next day then I know that that person or the business has lost the building, and the claims cheque can be sent out immediately.

Megan Thomas: From a reinsurance perspective, we're highly dependent on the information that we're provided by insurance carriers. To us, having greater visibility and access to the primary or the policyholder information is key when you're in the treaty space. We have to put at the top of the pack those which we know have really good quality data, which have very good geolocation information and we know what their process is by the way they do their underwriting. Another area that we're interested in is also the topography as well as the location; that's another lens of data. That's probably a little bit tougher to get in some cases, but very important on the cat side.

Kevin Livermore: That's an interesting point, in terms of the elevation and the topography data. As time's gone on, we've seen the tools that people are using and the software applications, the cat models and so forth, have become increasingly sophisticated over time – and as an industry we have to make sure that the location level data is keeping up with the sophistication of the tools that you're putting them in.

Now we're seeing flood models in the United States where the underlying granularity is down to 30m in highly hazardous areas, that could be well within the parcel of an individual property – and so the idea that you can get by with an address which is geocoded at some point on the street is really being challenged by the increasing sophistication of some of these models that have really granular topography, and that's just on flood.

We're just coming out of the significant season of wildfires that we saw last year and there's a concept known as the Wildland Urban Interface, where the manmade structures meet the brushland, and at Willis Re we've developed a model that will help clients evaluate the distance to that. We're seeing homes that are within that brushland that can have a very significant loss potential and a very different outcome than homes that are nearby but just outside of that brushland area – so having sophisticated location data is important just to keep up with the tools that are now coming into the marketplace.

Michael Reilly: Agreed, and this goes back to why we're starting to think we need to change how we store the location data to begin with. First of all, those static relational databases, when I want to start to go add in that type of data, it doesn't work, it's not set up for it, it's not structured for it. I can run the analytics from time to time, but it's hard to just keep adding a row to that table because it doesn't work. I need a more nodal structure where I can actually add that information in and do more advanced machine learning and other types of analytics on it and begin to get different views.

The picture that we all want is a comprehensive picture of the location, and a comprehensive picture of the location encompasses environmental exposures, it encompasses political exposures, it encompasses traffic exposures – all those other things come into play. We need to start to bring those pictures together so that we achieve true insights.

Ralph Groce: A lot of the work that I'm involved in doing is bringing in all those disparate data sets, data elements and serving them up in a way that underwriters and other interested parties can utilise that information and make informed decisions. Again, getting a single source of the truth. If your teams aren't using the same tools you can wind up with a different set of outcomes and writing a different set of business in a different set of risk parameters that don't necessarily jibe.

Michael Reilly: The other thing that the industry is struggling with is not all data sets are equal; not all data sets have the same fill rates, not all data sets have the same quality.

Jean Sullivan: Quality is so much more important than quantity. There's so much data out there, and what you do with that is a big question.

Matthew Bishop: On the point of data quality, when we're talking about property underwriting of large commercial risk there's a real focus not just on the location itself but also the locations' interdependencies and potential supply chain impacts. The industry learned some hard lessons in 2011 with the Thailand floods and Japan earthquake, where some of the biggest claims were not mapped to damage at an insured's premise. Where there is contingent time element exposure, efforts should be made to also map supply chain locations as well.

I also believe the industry can do much better in terms of reporting proper insurance to value. The high loss activity seen in 2017 and 2018 has highlighted the fact that reported values are not always consistent with the actual loss figures.

Michael Reilly: This is where we're seeing a real evolution how to support the underwriter. We started with quality scores or predictive models that would help with pricing. We're now trying to find more sophisticated ways to inform and help the underwriter to make those decisions.

Jean Sullivan: What we're trying to do differently is to peel back the onion and actually give you more access to the insight of the data that we're providing you. A real-life use case would be wildfire. Traditionally people get scores: is it high risk, or not? What we do with our data is to allow you to get access to our insight – not just the score – to learn more about, what does this truly mean? Are there embers? Is it the vegetation?

> "From a reinsurance perspective, we're highly dependent on the information that we're provided by insurance carriers. To us, having greater visibility and access to the primary or the policyholder information is key when you're in the treaty space"
>
> *Megan Thomas, AXIS Capital*

Kristen Bessette: It's the difference between data and information; you really want to give people information in the right context, not always just data where they have to work hard to figure out how it's relevant.

Shawn Moynihan: One of the topics I wanted to address is the concept of sharing data sets, the feasibility of truly reliable information being utilised among different organisations. Obviously, different organisations have different standards and different opinions on which data sets to trust, as well as which ones are most relevant and valuable to them. Is that a pipe dream, or is that even a possibility? Rob, you and I talked about the Accord data set and some of the movement around that, maybe we could start there?

Rob Sabio: Sure. A few years ago in a previous life I would attend meetings of an Accord data group, and they were trying to come up with a universal data set, if you will, one that would cover insurance and reinsurance and everything in between. Obviously, we all have different meanings; limits to me mean one thing and to you may mean another, and so they were trying to build a robust data set – a universal data set that would be able to satisfy all parties that were willing to contribute their data, and all parties involved would have a universal understanding of the data. Unfortunately, it never panned out, but I thought it was a great idea. Obviously, there was a lot involved with data security, confidentiality and what have you, but I think there is definitely a need to have a common understanding among us all – whether it be the insurance side, the reinsurance side or the brokerage side.

Kevin Livermore: Having an individual identifier for every structure shouldn't be a pipe dream. I think it's very much needed for risk management purposes, but you can also use it proactively from an underwriting perspective as well for risks that would fit into your appetite that you would

CONTINUED ON PAGE 30

like to pursue. It's an opportunity for growth.

Kristen Bessette: I think it's definitely possible to get better than we are today in terms of having that consistency. I don't think you'll have one data set that solves all our questions, so it's really defining where do you want to be consistent and where it helps us the most. Where are the biggest pain points? How do we have that common data set to allow us to exchange information freely with our partners? And then that can grow over time and solve some additional pain points as we go forward.

Ralph Groce: That's a good point you bring up Kristen, because I think that's what kind of brought that project to a halt, because you always had someone saying, "Well, what about this? What about that?" So yes, you do have to have a common data set.

Kristen Bessette: What are we trying to solve, right?

Ralph Groce: Yes, exactly, and then have a common goal. We may not be able to get all the 20 fields you want, but what if, say, if we get 15 of them, for example?

Shawn Moynihan: Obviously it's different for every organisation but the methodology behind the data sets that you use at all of your different companies is a factor. Where do you begin, in selecting which data sets you utilise and which vendors to use? What's most important to you?

Matthew Bishop: At SCOR we use a proprietary tool to map all location data, and the focus has been to overlay as much additional information on top of those locations, such as wind, EQ, flood and brushfire exposure. We can also see current events overlaid on that data. We're also working toward capturing more granular location specific date, incorporating risk protection and other details from engineering reports.

I think for us it's a constant challenge to try to add new relevant location data to the underwriting process. We're very excited by this journey and our clients really appreciate when we can inform them, in real time, when a CAT or man-made event is threatening their business. In that way we can bring real value to our clients by helping them with preparation and mitigation.

Megan Thomas: At AXIS we'll use a mix of vendor and internal models. We obviously rely a lot on our insurers to provide very detailed exposure information and we're using the information they give us into our approach that we want to take. It depends by peril, as well; some perils have a stronger vendor model than others.

Matthew Bishop: At SCOR there's a real focus for us on making sure that we're receiving the best data we can from the insured in terms of proper insurance to value. We also want to have a better understanding of contingent business interruption exposure and interdependency amongst the different locations, so as events unfold we can better understand the actual impact to our client, our clients' customers or our clients' suppliers.

Shawn Moynihan: Kevin, you and I discussed this, and I know you have some great examples of how better geolocation data can make all the difference when it comes to perils like wildfire and flood. Can you elaborate on that, why developing an elite data set is so important?

> "We started with quality scores or predictive models that would help with pricing. We're now trying to find more sophisticated ways to inform and help the underwriter to make those decisions"
>
> *Michael Reilly, Accenture*

Kevin Livermore: The insured's distance to the brushland is absolutely critical in determining its exposure to wildfire hazard, and I think the idea that your distance to fire station is going to solve all your problems has been somewhat refuted in the last couple of years. The California fires were uncontrollable for days; the distance-to-a-fire-station as a metric has been seen to not be very predictive of loss. Focusing more on the hazard and then the intersection of your location level data to the hazard has been the key for us.

In addition to the perimeter of the building, people are also starting to consider whether they have any additional factors surrounding the building that make it more susceptible. You've got the physical building, which may sit at different portions of a parcel, so it could be right next to a brushland or right next to a source of flooding, be it a river or a lake or what have you, and then you may have other features that either enhance or make the risk much worse. For example, you may have a house that's pretty far away from the brushland, but then you have a wooden deck that juts out all the way into the garden and actually becomes something that propagates fire. For example, vegetation that hasn't been cut back or growth in the setbacks around the house.

Shawn Moynihan: Ralph, I'd like your take on what differentiates a good data set from others. How do you delineate between those two? I'm curious about the decision-making around which data sets do you use that you feel are most "true."

Ralph Groce: It's interesting, because from my perspective some advances in technology are obliterated in the notion of "good" versus "bad." We're moving away from traditional data storage, data bars, those kinds of things, and we're actually working on data lakes where we can bring all of these data elements together irrespective of traditional notions of "good" and "bad" and then laying machine learning on top of it and AI as well, so that we can help underwriters and actuaries begin to draw correlations between certain data elements that they themselves

may not have seen. And those correlations, as some people intimated over this conversation, aren't static themselves, so what mattered two or three years ago doesn't matter in the same way as it does today. So being able to update those correlations among certain data elements and certain data points and leveraging that technology are what's driving us toward better outcomes, better pricing and better quantitative assessments of risk.

Michael Reilly: I will say the metadata around the data is almost becoming as important as the data itself. When was it last updated? When did it actually last change? Do we have a quality view of it? Is this data reinforced by other data? Those sort of characteristics around it are equal or more important than the data itself because they're helping us know better when to use it.

Ralph Groce: I think that cat models can benefit from external data; however, we can't leverage the same old data sources and obtain an improved result. For example, capturing flood data is a crucial matter to solve, but with FEMA, some of their flood maps are out of date, or the newer data sets don't take into consideration the future impact of climate change. As an industry, we may have to look at non-traditional data sources and modelling techniques to improve the cat models and underwriting outcomes.

Jean Sullivan: The trend we've seen is that cat modellers are actually sending us carriers that have an interest in providing more accurate geocoding into the models. There is interest from the modellers, and they know they need to improve the accuracy of their geocoding. I think that both they and carriers realise that you can't just rely on that historical data that has been traditionally applied in the models.

Matthew Bishop: I think carriers will always be early in identifying the shortfalls in those models. A lot more can be learned though, for instance, when you look at the recent losses that have come out of perils like convective storm, brushfire and inland flooding. In the US, these have actually been more impactful than windstorm and EQ, so naturally there's a high demand for a better understanding of those risks.

> "Our industry is of vital importance to the world and there's an opportunity to do things better. If there's a chance to really think this through and take it to a new level, we can provide an extraordinary level of value for customers, and more generally for all our stakeholders"
>
> *Matthew Bishop, SCOR Business Solutions*

Jean Sullivan: I want to revisit one of the questions on standardisation. When you talk about the data lakes, we also see ourselves positioned uniquely from the perspective of location MDM. We're seeing trends around MDM – not the large MDM data warehouse, but how we build the location MDM. We have some companies for which we provide every address in the United States – it's about 190 million addresses – and they use this data from us and then overlay our risk perils so they can pre-score every address in the US.

Once they've done the pre-scoring, they can make decisions about the business that they want to target. What's going to be profitable? Where should they write? There are a lot of different ways in which companies leverage our data and our pbKey. Our goal is to become the standard as this data is exchanged among carriers, reinsurers and so forth.

Shawn Moynihan: As geolocation tools and data continue to evolve rapidly, what other new elements do you all want to see that will be useful to your risk selection?

Rob Sabio: I think what's amazing today is that there isn't even a consistency between Bing and Google Maps, for example. You take a structure such as the Empire State Building, which takes up a whole city block; one of them may geocode the building here, the other may place the building on the opposite side of the block. If you're running a terrorism accumulation, for example, that makes a big difference as to where exactly that point has been placed on the map. Today, there isn't even a consistency in the latitude-longitudes being assigned to a location. As technology progresses it'll be interesting to see how the Bings and the Googles are going to come to a consensus on where their mapping is actually placing these locations.

Ralph Groce: And are they motivated to do so?

Rob Sabio: There has to be some incentive for them to achieve consistency.

Matthew Bishop: I think certainly on the engineering data side, when a facility changes hands through acquisition there's still the potential to retain some of the history of that location which can help us with future underwriting. There's also the question of understanding the loss history: I think for us as carriers we're sometimes approached on a risk that may have suffered a past loss, perhaps a flood, but we're not made aware of that because too much time has passed. That information is important to underwriting and risk prevention, and we're getting smarter in how we retain information from past events impacting a location.

Megan Thomas: From a reinsurer perspective, a unique identifier of buildings would be very helpful. Towers of insurance are usually purchased by corporations that involve thousands of locations which you stack together, and in the event of a loss it can be quite time-consuming to correlate that quickly to understand where some of that significant exposure might be. Clean, easy access to information would allow a reinsurer to see with a high degree of visibility where all the locations are, that are being reinsured from each individual carrier.

CONTINUED ON PAGE 32

Kristen Bessette: The data and technology are moving very fast, and we have volumes and volumes of it. To me, the question is, how do you make use of it? Do we have the right people to actually take the flood of new technology and new data and actually apply it to really the core things that we do in a way that makes us better? To me, that's really the question we have to answer pretty quickly, because the data and the technology are outpacing the skills to operationalise it. At the end of the day, it has to tie back to a better customer experience, to premium, or to losses, the core pieces of our business.

Ralph Groce: Success will be measured on how you use the data, and there are going to be companies that isolate the most important data and use it very wisely and other companies that for one reason or another will get stuck in the weeds.

Michael Reilly: I think we're at an inflection point. Carriers that take a step back, re-think how to store data, organise data, assess data and then rethink how to start to apply data are the ones that are going to be able to take the leap forward. Those that are going to stick with the models that they have always used are going to start to struggle. Machine learning and AI are all ready to go. It's just about taking the time to step back, reimagine what this space looks like and then start anew – and then the sky's the limit.

Ralph Groce: If you've ever watched chess players, particularly when you're playing with the clock, there are a million moves they could make and 10 moves they should make relative to whatever strategy they're looking to employ. That's where we are now. There are a million things you can do with this data, but there are 10 things you should do with this data. It's about figuring out what those 10 relative things are, those 10 questions, those 10 problems, those 10 issues, those 10 opportunities.

Shawn Moynihan: If there's a consensus to be reached today, what would it be?

Kristen Bessette: I think we all clearly know how important the data is, how important it is to our future, and I think all of us have talked about a lot of different ways that that's manifested in our organisations. To me, that might be the takeaway, that you have people from all these different parts of the industry who came here and devoted time to talking about having more granular, really good data to help us make better decisions. I think we all know how important that is and want to invest in that, and there are a lot of different ways that manifests itself. This type of information is our future, it's coming at us very, very quickly, and all of us are thinking about how to leverage that and to capitalise on it.

> "With better information, companies can look to write profitable business in areas or industries where they typically they haven't written before because it was considered high-risk. But now those carriers can be more selective"
>
> *Kristen Bessette, QBE*

Megan Thomas: There's not a one-size-fits-all solution because we all have different appetites, we all have different approaches, we're all coming from different areas depending on what our books are made up of. But getting access to information that is relevant and accurate and of good quality will better empower your decision-making.

Michael Reilly: My hope would be that there's at least a consensus that staying the course of "business as usual" is not success. The technologies, the data, the information has gotten to a point where it is time for every business to check its pulse. Even in a fast chess game, every once in a while you pause, you take a little bit longer on a move to rethink your strategy. We're at that point where we need to take that time and rethink the next step a little bit because I think there are some fundamental decisions as to where you place your bets next that will have huge implications for the future.

Ralph Groce: I agree, and I think it's really important in doing so that we look forward. It's easy to get caught up in the most recent event, such as the wildfires in California, and not considering brushfires in New Jersey. As we develop our systems and we collect this data we need to be thinking for the next event, not the last event, so that when flood happens in Malaysia versus, say, Thailand, we're not caught completely off guard. I think we have a short attention span in terms of looking back on the latest loss event. We need to be thinking about how we're capturing data for future unknowns.

Matthew Bishop: These are really exciting times. To Mike's point, we should take this opportunity to rethink "business as usual." Our industry is of vital importance to the world and there's an opportunity to do things better. If there's a chance to really think this through and take it to a new level, we can provide an extraordinary level of value for customers, and more generally for all our stakeholders.

Jean Sullivan: The consensus I'm hoping for is from the perspective of change and rethinking, that our organisation is perceived in a new way. Our goal is not to have people think of us as a mailing company. We're hoping that our role expands, especially with our pbKey and what we're doing around that unique and persistent identifier, and we believe we can offer consistency amongst carriers and reinsurers that are using our data. We're leading the industry in terms of providing the most accurate geocodes globally, relevant location risk data, and providing consistency through the pbKey across the enterprise. Our unique offerings improve the underwriting, actuarial, risk management and claims processes for carriers, reinsurers and brokers.