Jon Jay: A random chance approach would have gotten about 20%, would have anticipated about 20% of fires, and there's anticipated 71%.

Ed Comeau: You're listening to CRR Radio from the Vision 20/20 Project. Welcome to CRR Radio from the Vision 20/20 Project, my name's Ed Comeau. Today we have two fascinating conversations, Jon Jay, author of an article, Can Algorithms Predict House Fires? Who also did analysis to help the city of Boston predict the most likely places for house fires to happen is first up. Then after Jon, I talk with Nathan Armentrout who won a Hackathon competition in Louisville, Kentucky to invent a device that will detect fires in vacant properties and send an alarm to the Louisville fire department. Both of these conversations are about using today's technology and innovative ways to help make communities safer, so let's get into it. Jon, could you take a minute, introduce yourself to our listeners?

Jon Jay: Yeah, you bet. Thanks Ed. My name is Jon Jay, I'm a doctoral student at the Harvard School of Public Health.

Ed Comeau: I came across Jon when I read an article that he had written, Can Algorithms Predict House Fires? I listened to him talk at a recent meeting in Boston. Jon, fascinating article, can you kind of tell us a little bit about it?

Jon Jay: Yeah, thanks so much. The article, it starts with the ghost ship fire in Oakland, California last year, and traces through the idea that that was a property that was largely missing from city data sets. Maybe even more importantly, even if it hadn't been missing, the city didn't have an inspectional setup that would have necessarily gotten inspectors there to correct the building code violations that contributed to so many people dying in the fire. I looked into the ways that cities and fire departments in particular could be using data about properties better to predict risk and take care of it before fires happen. I found that there had been some very good work in this area. There have been projects in New Orleans, what they did there was they wanted to prioritize neighborhoods for a smoke alarm distribution program.

They did a very clever analysis using US Census data to figure out which neighborhoods were most likely to have the most houses that needed smoke alarms, and so that's where they targeted their program. Atlanta, Georgia had done one that was even closer to my interests because what they did is they went down to the property level for commercial building inspections. Actually, what they had originally done, they had been asked to go through the city and just find all of the commercial properties that they ought to be inspecting that they weren't. The problem was that they found so many properties that the fire department's next request was, "Can you figure out a way for us to prioritize these because there's no way we can hit all of them given our resources in a given year?"

That's where they came in with some more algorithmic data science approaches to prioritizing by risk. What they found is that risk varies so much at the property level that just by prioritizing you can do vastly better in your inspections than you would just by random chance. In their model, a random chance approach would have gotten about 20%, would have anticipated about 20% of fires, and theirs anticipated 71%.

Ed Comeau: You mean that by applying an algorithm to it, randomly they would've hit one out of five properties that might have had a fire just using random, versus using an algorithm they almost hit three quarters of the properties that might have had a fire, is that what you're saying?

Jon Jay: Making these comparisons it depends a lot on exactly what question you're asking and how you would be applying it operationally. Basically in this case it was, if you're accepting a 20% false positive rate, you've set your threshold so that you know you're going to be wrong some of the time. If you compared it to random chance, yeah, you'd do about three times better at that threshold. That's for a city-wide effort. I mean I think part of what's so powerful about that result is that it's for inspections of commercial properties, that's something that happens anyway. Using this approach it looks like, yeah, you more than triple the effectiveness of a program that you were already planning to do.

Ed Comeau: What kind of data do you use to reach that kind of decision? What sort of data sets are out there that tied into all that?

Jon Jay: I think that's part of what makes this so powerful. That's what we were looking for in Baton Rouge. A colleague and I worked on a similar project using not just commercial properties, but housing stock in Baton Rouge, Louisiana. What we wanted to show in particular in that project was that you could make predictions using data that every city's got. Any fire department can call the city assessor's office or the GIS manager's office and get these spreadsheets for every parcel in town, what's the size of the lot? And often you can get many building characteristics like the number of floors or the number of bedrooms. These are sort of standard urban planning data sets that might include the assessed value of the property and a bunch more variables that I think ... If you were coming from the urban planning world this is sort of just the routine property data that you've got for every property. That's the type of data they were using in Atlanta and the type of data that we used in Baton Rouge.

Ed Comeau: How were you able to tie that to fire safety? I mean, sure this data already exists, but what makes that connection now to fire safety?

Jon Jay: The really pivotal task there is to join those property data to past fires. What we had from Baton Rouge was four or five years of recent fires. For each of those one fires you match it with the property in your property data set. Once you've done that, now you can go back and what the algorithm's really doing is to look at your past fires and pull out the patterns that emerge across. Is this happening at bigger properties? Is it happening at older properties? Is it happening at

commercial properties versus residential properties? What the algorithm learns in that process, which we call training, that's what you deploy with your algorithm.

Ed Comeau: What were some of the big factors? I think you mentioned building size, assessed value, those are some of the big factors that play into making that prediction?

Jon Jay: Size and value were really powerful in Baton Rouge as well as latitude and longitude. My understanding is this is pretty well known in the fire literature is that within a city you get a big gradient across neighborhoods. I think latitude and longitude end up being strong predictors particularly for the type of model that we used, which is called a random force. The other strong predictors in the Baton rouge model included value and lot size. It was interesting because what we saw was increased risk at the high end of those two variables, so that you were talking about large properties that had high valuations, and in general those were bigger condo complexes, or college dormitories, or those kinds of places where there are many people and that increases risk. Then we also saw increased risk at the bottom of both of those variables where small properties with low valuations were the worst conditioned properties, and those also had elevated risk.

Ed Comeau: How about demographics, do those come into play at all?

Jon Jay: Not directly. On the projects that I've worked on we haven't had household level demographics. I think those are in the literature, at the census track level those are powerful predictors and I would expect that they would have correlations at the property level too. I think we do pick up a lot of that important variation within the property level data so that within a certain city you've already pulled in the neighborhood data by saying, "It's got these latitude and longitude coordinates, and then on top of that we know this is a three family house in that neighborhood." That might pull out a lot of the most important risk factor type information that you're also capturing, that you would otherwise capture with demographic data.

Ed Comeau: Now, you did this down in Baton Rouge, is this something that is unique to Baton Rouge, or would it be something that would be applicable in any other community across the country?

Jon Jay: Yeah, I think it's absolutely applicable to any other city. The city of Boston recently held a data competition basically to replicate this sort of model for the city of Boston. The people who did predictive models for the city ... For operations where the fire department is basically going by random chance, if you've got your inspectional list and it's just a random order, or the same order every year, something like that, the algorithms, we work hard to sort of get them very good at the very top end, but they're all just so much better than doing it randomly. These data, like I said, are available for many cities. One thing

that I think will be particularly interesting as it gets replicated in more and more places is that you can experiment with different predictors.

For example, in that Boston data competition I looked specifically at the city's 311 data set. Boston has a very robust 311 system where citizens call the city and they make requests or complaints. Boston generally receives many complaints about housing conditions or building conditions where neighbors have said, "The house next door is crumbling, or boarded up, or there's some other problems there," or residents have called and said, "My landlord's not taking care of this property, I want the city to come do something about it." One fifth of houses in Boston had a complaint like that during the period of around 2012 to 2016.

When you merge that in with all of this property level data and with fire data, I found that there was a really strong correlation between those complaints and fire incidents. Even controlling for things like the size of the property and factors like whether the property was owner occupied. Every city I think can and should do it a little differently, just because they're trying to answer slightly different questions and might be using it for slightly different tasks. But it's definitely the kind of data driven approach that I would foresee the majority of fire departments using within the next five or 10 years.

Ed Comeau: Now you said the data sets are pretty routinely available, or any large city would have it. To do this kind of an analysis, does it take complex software? Does it take a team of rocket scientists to do it? Or is this something that's within the reach of the ordinary city or the ordinary fire department?

Jon Jay: Everything that I've done has been with free software and does require a little bit of data science expertise. I think at this stage, as far as I know, there's not an off-the-shelf application that you could buy that would do this with the data that you've got and firehouse software or whatever. But I think it requires what I would consider a pretty basic data science skillset and I think it's an area where I would foresee people coming out with software dedicated to this task. I think an ambitious fire department, even if you didn't want to do the fanciest statistical modeling, this kind of data driven approach I think is the kind of in-house expertise that fire departments either have already or should be building just to be able to look at what kinds of properties are causing the most fires at which times? It doesn't need to be this very complex, precise statistical approach.

I think in many ways it might even be better from an institutional standpoint to start with building a culture of data of just understanding just so your rank and file firefighters could tell you off the top of their head, "What's the proportion of fires that we have in residential properties versus in commercial properties?" And start thinking along those lines and building up expertise in that way.

Ed Comeau: Is anybody building this as an open source sort of thing on GitHub or anything like that?

Jon Jay: I haven't seen it. I'm very happy to share all the code that I've used for these projects and I think it would definitely be ... I've seen that Vision 20/20 and other partners have been really thoughtful about the role that data is playing in firefighting. The National Fire Protection Association has also been doing a lot of work around data, and so I think it's definitely a good place for different partners to come together because so many of the tools are replicable across settings. Even if you don't want to replicate the exact thing that another city did, you could definitely be learning, and improvising, and innovating to suit your city's particular needs.

Ed Comeau: One of the things we were talking about was how this stuff is just so easily accessible that almost ... Well, I wouldn't say anybody can do it, but it doesn't take really esoteric tools to pull this stuff together does it?

Jon Jay: No, not at all. I think it's something that, right, even if someone in your fire department can't do it, someone in your city probably can. I think it's been a really interesting model to see the way that different cities have pulled in members of the general public to do this kind of work. I mentioned that Boston had run this open data competition recently, and they had people just come in who they do data science for a firm during the day, and they stay up late at night just to contribute to public safety in their cities, running very sophisticated models. I think even at a lower level of expertise there are data geeks out there who are very happy to contribute.

Ed Comeau: What got you into this? What turned you onto it in the first place?

Jon Jay: Yeah, so I worked last summer in the city of Lawrence, Massachusetts. I was there working with a group from the Harvard Kennedy School and we were working with the Community Development Department, basically the Urban Planning Department, and we were working with the Inspectional Services Department, and several other government agencies on housing related data. Basically sort of helping these agencies share the information that they had about different properties. Over the course of the summer there were several big fires, including one where I remember driving into Lawrence in the morning and smelling the smoke from this overnight fire, and just realizing that the data that I already had at my fingertips from this project might be useful.

By the end of the summer I was going in and looking through ... For each of these properties that had had a fire there was this really visible pattern that this wasn't the first problem that the city had had with most of those properties. Many of them had this long-standing pattern of there was something wrong with building code compliance or the police had been called to the property in the past. That was what really gave me the idea that we could be doing much more to predict risk at the individual property level.

Ed Comeau: Are there any other cities you know of doing this? You mentioned Baton Rouge, and New Orleans, and Atlanta. Are there others that are going down this road as well that you're aware of?

Jon Jay: I know Syracuse, New York had started working with ... There's a data science firm called Enigma that had worked with the team in New Orleans on their predictive model. I think they have picked this up. I know that American Red Cross was involved with the New Orleans project too, and so I wouldn't be surprised to hear that there were many more cities already starting to incorporate this, but those are the ones that I've come across.

Ed Comeau: We've been talking with Jon Jay who's a doctoral student at Harvard, really talking about the essence of his article here, Can Algorithms Predict House Fires? I'll certainly have a link to it in the show notes below, and his article in turn links to a lot of other great articles and sites on this stuff. Jon, I really appreciate you taking the time to chat with us today.

Jon Jay: Thanks so much Ed.

Ed Comeau: Some really great information from Jon Jay, and before we get into our conversation with Nathan Armentrout about how he is working with the Louisville Fire Department, I'd like to take a minute to tell you about some of the resources available to you from the Vision 20/20 Project. On our website you'll find information that can help you with your community risk reduction program including evaluation implementation guides, a fire safety materials generator, report on rural fire prevention with some recommendations on implementing fire prevention in these communities, videos, and much more. If you go to www.StrategicFire.org/CRR, you can check out the CRR portal for more information. Now, let's get back to our show.

Next on CRR Radio we're talking with Nathan Armentrout down in Louisville about what he's doing with the Louisville Fire Department. Nathan, could you please introduce yourself?

Jonathan A.: Yeah, absolutely. I'm founder of a new startup company called Eidolon and we are working with the city of Louisville to develop and implement a device that alerts the authorities when a fire in a vacant property starts. My background is in engineering where I worked to develop circuit boards and write software for IoT devices, so that's where my experience really helps with developing this device for the city of Louisville.

Ed Comeau: Nathan and I have already talked before the show here and I was really fascinated with the device that he's putting together. Maybe you can give a little bit of the backstory on it. How did it even come to be in the first place? How did this come onto your radar screen?

Jonathan A.: Well, the story's actually a great example of how a city can engage its tech community to solve a problem. Louisville Metro Government went to Level 1 Hacker Space, which is an open workshop where people who are passionate about making things come to hangout. The Louisville Metro Government said, "Hey, half of our fires involving two or more buildings involve a vacant property.

We want to know when they catch fire so they don't spread. Can you invent something?" In partnership with Level 1, the city hosted a Hackathon and we ended up having 27 makers from the community compete with all kinds of different solutions. Everything from hacking smoke detectors to our solution called Casper. Ultimately, our team ended up winning that competition and we've been working with the city for about a year now to develop it from prototype to production ready of this device.

Ed Comeau: Now, before we get into Casper itself, I'd like to talk a little bit about the Hackathon. I mean is this a common sort of thing that the government might go to a place like this? How do these function? How do they work?

Jonathan A.: Yeah, so this is kind of a novel approach to solving the problem. Louisville had a problem, they couldn't find a solution on the market, and so they went to the maker space. Maker spaces are kind of a new thing, they've been around about five to 10 years mostly. There's usually one in every major city and community, sometimes multiple. These folks who hangout here, myself included, just love building things, it's just what we do. After we come home from our jobs we can't think of anything else to do so we just keep building new things, and work with technologies, and experiment, and try all sorts of different ideas. Sometimes they turn out to be products, sometimes they turn out to be fire-breathing ponies. It could be anything from fun to helpful. This was a really unique approach the city took where they challenged the community to come up with the solution and the community responded.

Ed Comeau: What were some of the solutions other people came up with? Again, before we delve into yours, I'm just kind of curious what some other people came up with?

Jonathan A.: The second place team was working on a way to hack into a smoke detector directly and find some way to embed a wireless device inside of it so they could detect when it was going off and then send a message to Louisville Metro Government that way. There was a lot of different permutations of that, different combinations of that idea. Some were just trying to outright build the smoke detectors themselves with some connectivity. They took some smoke sensors, and IR sensors, and different things like that, and then just tried to build it completely from scratch and kind of work with the constraints of the Hackathon. With vacant properties, the challenge is that they don't have electricity so the goal was to come up with a device that could work in those constraints.

Ed Comeau: Well, let's talk about Casper. Why don't you describe what you put together there.

Jonathan A.: Yeah. Casper is an acronym for Completely Autonomous Solar Powered Event Responder, which is kind of descriptive. CASPER, what is does is it listens for smoke detector sirens inside vacant homes and then alerts the authorities and neighbors of the imminent danger usually through text message or with a phone call. The reason we went with this approach, the completely autonomous part,

is that we didn't want to try to reinvent the wheel. Smoke detectors are pretty common and pretty reliable, so we just wanted to make a device that would listen to them and then let the government know when a problem is occurring. Now, the unique thing about CASPER versus other DIY home security products is that it's solar powered and it doesn't rely on someone paying the electric bill, which is perfect for vacant properties. Then obviously event responder is just it responds to events. Right now it uses a microphone to sense the smoke detector sirens, but we can certainly start to listen for other events in the future that the city's requested.

Ed Comeau: Physically, what does this thing look like?

Jonathan A.: It looks kind of like the size of your typical home router, your wifi router at home, except for on one side it's a solar panel and the rest of the box just holds a number of different components. It's got the microphone in there, we have a specialized circuit that listens just for smoke detector sirens only. Then we have a small computer inside of it that processes that waveform and then if it says, "Hey, it sounds like a smoke detector," it sends a message to the cloud using a cell modem. You don't have to have wifi at the house or rely on large wifi networks or anything like that, it works independently.

Ed Comeau: So since it's solar powered, I assume it needs to be near a window and does that create a problem in terms of listening to a smoke alarm that might be going off somewhere in the house like down in the basement or something like that?

Jonathan A.: We get really good coverage with just one CASPER per house across multiple floors and stories of a building. CASPER is typically installed inside of the vacant property in front of the best window that we can find, typically a south facing window if we can. Once CASPER is installed, it's microphone has enough sensitivity to pick up on smoke detector sirens throughout the house. Smoke detector sirens are very loud, and with our sensitive microphone we can pick them up through multiple stories, whether or not doors are closed or not, different things like that. We have also installed CASPER outside in some pilot homes that we tried this device on. It's able to also pick it up through exterior walls if needed. There's a lot of different options with installing CASPER to meet lots of different architectures and housing styles.

Ed Comeau: You're saying it's sending out the signal over the cell service. Where does it send the signal to? Can you specify who gets it, or even how many people get a signal?

Jonathan A.: Yeah, so it comes up through our company's cloud service. In Louisville Metro's case, we work with them to set up lists of phone numbers to call when an event happens. Right now a person from the Office of Performance Improvement from Louisville Metro Government gets the phone call. A select number of firefighters can get a phone call as well. Then of course myself and the rest of the team at Eidolon also get an alert when an event happens. While we've been piloting this technology, we've been working hard to make sure that there's no

false positives and annoying features like that. We can change that access list based on whatever the city wants to allow. We're looking in Louisville to implement a website where neighbors of vacant properties can choose to get alerts as well. Often times vacant property fires happen late at night when everyone's asleep, and there's not always a good warning when you hear a fire going on. This is definitely a way for the city to open up a community engagement path and make sure everyone's protected.

Ed Comeau: Tell me a little bit about the pilot test, how's that been going?

Jonathan A.: We concluded the pilot at the end of April. We tested 10 CASPER devices across 15 different homes in Louisville. All different types of homes, multiple unit homes, homes with awnings, homes that are really close to other homes. It's just endless possibilities there. We were really trying to put it through its paces and make sure that CASPER works really well. Put it next to homes with train tracks running by, next to factories, in all sorts of different situations. Louisville was really pleased with the results of this pilot, it proved really three things about it. One, when we went to go test the devices monthly they all reported and heard the smoke detector sirens correctly, so that was always a good gut check test each month.

It had zero false positives over the course of those six months, so nothing where we went out and said, "Hey, there was nothing going on, or a school bus, or an ambulance, or something weird." Then the best news of all is that it actually caught two times when a smoke detector was going off inside a vacant house. Not something we set up as a test, it was a real life event and it was a great example of what was happening. It was a great example of how CASPER is supposed to work for cities. The fun thing about it, while their building wasn't on fire, the smoke detectors were going off and we were able to document it and show that the device was working even outside of our testing.

Ed Comeau: Can you monitor the health of the device through your cell service?

Jonathan A.: Yeah, so we can monitor the battery voltage and keep close tabs on how the devices are working. Another kind of fun story about the pilot we did is at one of the houses we noticed that the CASPER device was dying really quickly, and so we went out there and noticed someone had taped a couple phone book pages in front of the device to kill it. That's kind of one of those interesting things where working with the city we have to communicate to neighbors and let them know what the device is doing so they feel safe and really understand what it's doing.

Ed Comeau: What are the next steps?

Jonathan A.: Well, we're working with Louisville to expand the program from 10 devices to hundreds of devices so we can start targeting the hotspot neighborhoods where vacant properties are a problem. We're looking at kind of retooling CASPER and

making it extensible. We've also talked to a number of other cities who have these problems as well, and we said, "What would your dream device look like?" We've taken that feedback and started adding in the ability to maybe detect if someone kicked in the doors using an extra sensor, or maybe a motion sensor outside to see if there's any type of drug activity. We're really looking to find different ways of how CASPER can not only help just fire departments and cities, but also the vacant abandoned properties, offices, and police departments as well.

Really CASPER can become a platform that helps cities manage vacant properties and mitigate their dangers effectively. I never knew that this was a huge problem for cities, and it's helpful to take your engineering skills and apply it to something that's not only just got business impact but definitely a social impact as well. In fact, we've gone out and interviewed a number of neighbors of these houses, and to hear their stories of near misses and things like that. You talk to these people, they're really awesome folks and they deserve a lot of help.

Ed Comeau: I have to say, I was so intrigued when I first heard about this and that's really why I reached out to you. I figured our listeners would be just as intrigued too. I think it's just an incredible device and I can see a lot of potential for it. Really, kudos to you for coming up with it.

Jonathan A.: Well, thank you. It's been a really great experience and we're excited to start spreading CASPER to multiple cities here soon.

Ed Comeau: Well, today we've been talking with Nathan Armentrout down in Louisville about the new device he has come up with through a Hackathon called CASPER, and we'll have more information in the show notes along with a link to his company there where you can find out more information. Thanks for joining us today Nathan.

Jonathan A.: Yeah, thank you Ed.

Ed Comeau: Well, we certainly covered a lot of ground in today's podcast. You'll find links to much of the information that both Jon and Nathan mentioned in the show notes and on the CRR Radio Podcast page. While you're there be sure to check out our other podcasts including a talk with National Fire Academy Superintendent Tanya Hoover, an overview on how New Orleans is using data to focus its smoke alarm installation program, CRR in Rogers, Arkansas with IFC President Tom Jenkins and much more. CRR Radio is produced by me, Ed Comeau, and is a production of the Vision 20/20 Project. For more information please visit us as www.StrategicFire.org, and thanks for listening.