EPISODE 859

[INTRODUCTION]

[00:00:00] JM: A grocery store contains fruit, vegetables, meat, bread and other items that can expire. In order to keep these items in stock, the store must be aware of how much food has been sold and what has gone bad.

When a food item is low in stock, the store needs to order more of that food from a central distribution system. Managing food inventory is not simple. Some kinds of meat might expire faster than others. Avocados do not become ripe at the same rate as apples. In order to keep the shelves stocked, there are manual workflows for checking the inventory and ordering a new inventory.

Afresh is a company that builds software for grocery stores. Afresh works with grocery chains that have a central distribution center. These grocery stores already have some software. At the back of the store, inventory management systems maintain records of the items that the store has on the shelves.

At the front of the store, checkout systems detect what has been sold to help update the inventory. When the inventory is running low, the store can order more inventory from the central distribution center so that trucks can deliver more inventory. Afresh improves the operational intelligence of the stores by detecting spoilage among items that are prone to expiration, such as fruit.

Volodymyr Kuleshov is the CTO and cofounder of Afresh and he joins the show to discuss the technical challenges of a grocery store and the software that Afresh is building to make groceries more intelligent.

Find Collabs is the company I'm building. It's a place to find collaborators and build projects. If you're interested in finding other people to work on a project with or if you want to find a cofounder for a new company, you can check out findcollabs.com. We're also having a hackathon with $2,500 in prizes, and some of the people who have created cool projects in the past have come on the Find Collabs Podcast. So if you're interested in learning more about the Find Collabs community, you can check out that podcast, and I would love to see you on Find Collabs.

With that, let's get on to today's show.

[SPONSOR MESSAGE]

[00:02:28] JM: As a software engineer, chances are you've crossed paths with MongoDB at some point. Whether you're building an app for millions of users or just figuring out a side business. As the most popular non-relational database, MongoDB is intuitive and incredibly easy for development teams to use.

Now with MongoDB Atlas, you can take advantage of MongoDB's flexible document data model as a fully automated cloud service. MongoDB Atlas handles all of the costly database operations and administration tasks that you'd rather not spend time on, like security, and high availability, and data recovery, and monitoring, and elastic scaling.

Try MongoDB Atlas today for free by going to mongodb.com/se to learn more. Go to mongodb.com/se and you can learn more about MongoDB Atlas as well as support Software Engineering Daily by checking out the new MongoDB Atlas serverless solution for MongoDB. That's mongodb.com/se.

Thank you to MongoDB for being a sponsor.

[INTERVIEW]

[00:03:50] JM: Volod Kuleshov, welcome to Software Engineering Daily.

[00:03:52] VK: Thank you for having me.

[00:03:54] JM: Yeah, it's great to have you. You work on Afresh, and we will get into what Afresh is. But first we should talk about grocery stores. I like to go to grocery stores. I like to look at fruits and vegetables and pick up stuff that is seasonal, but I've never worked in a grocery store. I have no idea how a grocery store operates, what the different employees do. Take me inside a grocery store. How does it operate?

[00:04:20] VK: So, in a typical grocery store, you have multiple departments. You would have produce, meat, deli, etc. Typically, each department has its own manager, and this manager does most of the work that goes into running this department. That includes some obvious things like helping the customers, stocking the shelves, unloading items that come in from the warehouse.

What really surprised me when I started working on this problem is that they also manage all the important operational decisions around their department, such as, "Which should I be ordering?" "Where I should I be placing it in the store?" "Do I need to run a discount? If I run a discount, what's the markdown in price?" "Do I need to add extra stickers to show that it's on add or it's in rebate?"

So these are decisions that really, really, really affect the bottom line of the store, which are made by the same person that kind of helps you and that stocks the shelves and they're doing it using very manual methods that are mostly based on their intuition. So they receive these ordering sheets, which are faxed to them from the corporate office, and then they just use this pen and paper process to make all these key operating decisions in the supermarket. So that's kind of what this observation is, part of what led me and my cofounders to start this company, Afresh, to try to bring some data-driven methods into their decision making.

[00:05:48] JM: What are the areas of waste within a grocery store?

[00:05:52] VK: Essentially, since a human does all these orders, again, without really any kind of data-driven support, their natural tendency is that, "I never want to run out of my items. My manager will really not like this. So I better order enough to have in stock."

So, that will naturally lead them to over-ordering, and the problem with perishable food is that, of course, it goes bad. So, stores are incurring massive amounts of food waste because of these manual ordering processes, which then in turn really, really hurts their profit margins, which are really thin, which are just normally really thin.

Basically, so you are asking about food waste more generally. It is a big problem both at the supermarket level and at the level of the entire food supply chain. So, if for example, you look at the entire U.S. food supply, the statistics are that around 50% of all the food that gets produced ends up being discarded before it gets to the consumer. So this is a huge financial problem, a huge environmental problem. Of all of these food waste, about 40% occurs at the supermarket level and downstream. So that's kind of the, I guess, part of what we're trying to do with our company is we're trying to make these manual processes more efficient by making them more data-driven and then at the same time it will significantly reduce the food waste that these supermarkets and then also the downstream consumers who buy this product, it will reduce the waste that they're incurring, which will really help the bottom line for customers and also it will be good for the environment.

[00:07:36] JM: Grocery stores do have some technology. There's a cash register. There's a conveyor belt at the checkout. Some places have a self-checkout. There're also security camera systems. There's probably some other software, back office related software. But generally speaking, grocery stores represent a kind of domain that software is starting to affect more and more in a way that's really positive, because we have all these domains throughout our lives where you just walk into these domains, like an airport. I think of an airport, for example. Airport is actually maybe not a great example, because they're a little bit sophisticated.

But you walk into an airport and you're just like there's so much manual work going on or work that is like sub – Like where people are using computers but they're using computers in a sub-optimal, not really modernized way, and that's to be expected, because mobile computing is fairly new. Cloud computing is fairly new. Usable machine learning interfaces are really new. So there's so much room to modernize. So much room to save resources, to save waste in a really positive-sum way.

But before we get to that ideal, the stuff that you're working on, what's the status quo? What kinds of technology is installed within a modern grocery store?

[00:09:04] VK: Yeah. I think you named a lot of the main ones, kind of these devices that you described are where we get most of our data. So, obviously, there's the cash register. These some fairly old equipment, various scanned ons and handheld devices to scan boxes as they arrive as they move through the store. So these would typically have barcodes and they – Basically by scanning them, there's a record that's preserved of the box. So leaving the warehouse, arriving to the store. Then as they sold, these items are also being tracked and as they go through the register.

You mentioned security cameras. That's actually a really exciting opportunity. So, our system is currently not using them, but there is lots of exciting work in this area to also be able to use, or in the area – There's exciting work that are just being done in grocery stores around trying to use that data a little bit better, that we have improved computer vision.

But I think these processes are surprisingly manual and very old. This has traditionally been a very tech-adverse industry, and only now are they waking up to the need to adapt technology faster. So, there're a few reasons. Traditionally, a lot of the technology that they're trying to sell them now that's centered around fresh food was not as necessary, because fresh food was seen as just the way to get a person to come to the store and then they would sell them these expensive packaged goods where they would make most of their profits.

Now, a lot of the not perishables are being sold online. People are buying them through Amazon and they get delivered to their door. So now, the importance of fresh food is becoming greater than ever. At the same time, they're facing a lot of competitive pressure from companies at Amazon that are now deploying their own automated grocery stores, and this is really making them realize the importance of adapting these technologies that we're describing.

[00:11:12] JM: Explain what Afresh does.

[00:11:14] VK: So, Afresh tries to – We're building technology that automates the supply chain decisions that are being made at the supermarket level. So, effectively, we deliver a product to this department manager that I just described, and this system from his perspective, it's a tablet, which runs their software. This tablet replaces the pen and paper workflow that we were just discussing.

So, right now, they get faxed these order sheets. Now, all that information comes via the tablet. You can think of this tablet as being control center for their grocery store. So they can use it to do their orders, and if they want to set markdowns, they can also do it through the app. Basically, we envisioned this as being the control center through which they can set all the operations through the store.

Now, the key part of this system where we're going to deliver a lot of value is on the backend side where we have a machine learning engine that we have developed, which looks at their historical data and automates all these decisions that they're making.

So, effectively, the system repopulates and sets the knob of this control panel to what we think this should be doing, and then they only need to review our recommendations and either accept or maybe slightly overwrite some of our recommendations if they don't agree with them. So you can think of this process as – The ultimate result is this sort of automated supermarket where once they accept our recommendations, the store replenishes itself and sets prices and does any price markdowns that are necessary and effectively in this self-driving, autopilot mode with respect to all the key operating decisions that a person currently has to make.

[00:13:08] JM: I can imagine workers within a store walking around with tablet computers. Maybe it's an iPad or an Android device or something, and that software would let them make decisions. What kinds of decisions would the human be making? Would this be like the humans fulltime job? Do they just become the kind of driver of the grocery store where they're like walking around, they're looking at stuff and then they're kind of getting some signals from other things and there's like kind of a human computer interaction thing?

I guess, I want to better understand what the human computer symbiosis is. What the computer is offering and what the human is making decisions on.

[00:13:55] VK: Yeah, definitely. Yeah, I think that's a really great question, because this symbiosis is really crucial to what we're doing. So, first of all, the amount of time that the person currently spends on this ordering and on kind of reviewing their inventory and so on, this will significantly diminish. So we envisioned a world where these department managers now have more time to spend with helping customers and making sure that the shelves look good and doing all the important things that a human will do better than the machine.

Now, in the time that's left and the time that they spent actually using the tablet, they would do several things. One would be to review what the system is recommending and make sure that it makes sense to them. Then, also, by reviewing them, they get to input some of their knowledge into the system.

Part of the reason why these technologies that I'm just describing, so these supply chain automation technologies, have not been deployed in the fresh food vertical is because there are a lot of elements of what's in the real-world and is crucial that are not being captured in these typical data streams. So, for example, a cash register and what gets delivered. There're still many elements that are not captured in their database.

So, by giving them this tablet, on one hand, we can push our recommendations to them and then also we can pull certain aspects of what's happening in the store. So, this gives our algorithm data to make better decisions that could not be made otherwise if we were just doing a centralized system that only integrates with their main warehouse.

So, this is actually a really interesting design problem, I have to tell you, that we're still iterating on many kinds of designs that will collect various sorts of data. But you can think of them providing rough, high-level information about what's in the store or some elements about the quality of an item that would come in. Of course, these are all these queries that we would ask the – These are also data-driven.

For example, if we see that given their historical data, we now have a large arrow bar on what's in the store. Then we can them, "Well, okay, can you tell me how many cases of this particular item you have in the backroom, because now the data is looking – Basically, my model is now putting high arrow bar in that, and you could query that question in order to improve my accuracy." So it's this twofold process that's really interesting. This kind of bidirectional process that's interesting.

[SPONSOR MESSAGE]

[00:16:41] JM: When I'm building a new product, G2i is the company that I call on to help me find a developer who can build the first version of my product. G2i is a hiring platform run by engineers that matches you with React, React Native, GraphQL and mobile engineers who you can trust. Whether you are a new company building your first product, like me, or an established company that wants additional engineering help, G2i has the talent that you need to accomplish your goals.

Go to softwareengineeringdaily.com/g2i to learn more about what G2i has to offer. We've also done several shows with the people who run G2i, Gabe Greenberg, and the rest of his team. These are engineers who know about the React ecosystem, about the mobile ecosystem, about GraphQL, React Native. They know their stuff and they run a great organization.

In my personal experience, G2i has linked me up with experienced engineers that can fit my budget, and the G2i staff are friendly and easy to work with. They know how product development works. They can help you find the perfect engineer for your stack, and you can go to softwareengineeringdaily.com/g2i to learn more about G2i.

Thank you to G2i for being a great supporter of Software Engineering Daily both as listeners and also as people who have contributed code that have helped me out in my projects. So if you want to get some additional help for your engineering projects, go to softwareengineeringdaily.com/g2i.

[INTERVIEW CONTINUED]

[00:18:32] JM: The advantage of this domain is that inevitability is on your side. Eventually, these supermarkets are all going to need to refurbish themselves as software integrated shipping areas. So that's great. The problem is you are kind of, by definition, early, because you're building the technology that they're going to adapt. So what makes it tricky is that there's so many different areas that you want to solve, and I imagine if you aren't super familiar with the grocery domain, these things can start to look so interconnected that it's really hard to find a specific problem to solve within the grocery domain.

If there was some specific problem that you could find, for example, like let's say just apples. Let's say you could just optimize when a store should order more apples. If you could just simply integrate with the cash register system. Let's say you just integrated with the cash register system and the inventory management system. Let's assume those things are accurate. You could simply know how many apples have been ordered. How many apples have been purchased, and based on that information, and perhaps some other inputs about how fast apples spoil or maybe like recent science about how apple spoilage has gotten better or something. You could probably integrate those things together and simply say, "Hey, grocery store, we're going to provide you with a model that allows you to more intelligently order your apples. Will you let us trial this?" They let you trial it, and then you reduce their apple spoilage. You make them some money, and then they, like no brainer, are willing to buy software from you.

I just bring that up as an example, because I want to know how you actually enter into this market, because selling software to a grocery store, I imagine, is really, really, really hard and you have to prove to them some specific thing that you can do for them that is advantageous. So how do you enter this market?

[00:20:42] VK: Yeah, that's a great question. Yeah. So, I guess the first quick answer is that we found that there is so much need for this tool and there is basically the amount of money that they lose on this food that goes to waste. It's really huge, and they're willing to work with a relatively young company, like ours, and they're willing to do pilots with us, because they just see the immense potential that a technology like this can have.

Yeah. Basically, that's kind of the short answer. Now, maybe the slightly longer answer is that I can tell you a little bit more about how we structure this process and how we got into this market in the first place. The story – So our initial partner, we started working with them, me and my two cofounders, Matt and Nathan, when we were still students at Stanford.

So I was finishing a BSc program and they were finishing their MBA. We got data from a supermarket chain that was interested to work with us while we were still students, and this was structured as a research project where we would look at their data and try to understand their sources of food waste, which is interesting in itself, because food waste is a societal and environmental problem that is interesting to many people in academia and nonprofits. To them, this partnership was interesting, because maybe we would find a better way, some good way to reduce their food waste, which I think we did.

Initially, we started this project as students and this was a research project. Then we found a way to – So we basically looked at their dataset, looked at their waste. Tried to understand why and when it happens. Then we came up with a system that would automate some of these decisions, basically focusing on ordering. Then we came up with a framework for evaluating and trying to predict how much value this would deliver using simulation on their historical data.

So, then we gave them a presentation of these results and they were very excited and we saw that there was a lot of demand among other supermarket chains. So, that's basically how we decided to form this company.

Then, currently, the way that we sell this to software companies is very similar. We have several steps. Initially, we get historical data dump from them and we use this historical data dump to build a model and to, again, run a simulation on their historical data to compare their historical performance with a performance they would have had if they had used our algorithm.

We then deliver a report with kind of charts and metrics and we show them what our algorithm would have done and how this would have impacted waste. We explain to them how our model works and how our simulation works and what are all the assumptions. We try to be conservative. Then we let them see this report. We show them what's our expected ROI. Then based on this, they determine whether it makes sense for them to proceed. Then the next step for us is to deploy the system in a small number of stores, usually between one to five. Then once that goes well, we move on to larger deployment.

[00:24:07] JM: So you get a historical data dump from a prospective grocery store customer.

[00:24:14] VK: Yes.

[00:24:15] JM: What are the datasets and the machines that they are giving you in that data dump?

[00:24:25] VK: Basically, they give us pretty much everything that they record internally in their database, or data warehouse, and this includes mainly things like their historical sales, their historical orders, their historical prices, their shipments between stores. Sometimes they track their waste, although it's not always accurate. So we can get a feed of that. Historical adds, markdowns, etc.

So it's a lot of time series data and it's everything that – So, these are things that are usually recorded either at the cash register or at various places in their supply chain when they move boxes and when they scan them with a gun.

[00:25:03] JM: So, are most grocery stores this well teched out, or are we talking about some specific set? Because this is like somewhere in between the technological sophistication of Whole Foods, which is very sophisticated, versus your random mom and pop grocery store or like a farmers market. How many grocery stores are this technologically sophisticated, where they're tracking all these data?

[00:25:29] VK: We work with supermarkets that are self-distributed. Meaning that they own their warehouse. This is typically – When you look at chains that start to have 25, 30 stores or more, then it makes sense for them to have their own warehouse and then there're a lot of standard tools and processes that are used when managing a warehouse, which then record all of these data. Basically, these are kind of midsized to large chains that we work with. So we don't work with a small mom and pop supermarkets or small mom and pop corner grocery stores. In the U.S., about half of the stores, half of the grocery stores in U.S. would be part of a chain which would have this kind of technological sophistication that I described.

[00:26:20] JM: Yeah, that actually makes a lot of sense, because I think about like Safeway or H-E-B. Those are the grocery chains that I'm most familiar with. They're pretty well-run operations. I'm thinking about it now. Earlier I kind of categorized grocery stores as a place where technology has not invaded that much, but it does feel like Safeway and H-E-B. If I think back 10 years ago, they are much more well-run today than they were in the past. So I can imagine them actually having a good amount of data.

[00:26:55] VK: Again, this data depends on the supermarket chain. Definitely, we've seen a lot of variance in how much data they store, and it has never gone back to like 20 years ago. It's always on the order of like between 1 and 10 years of data that they have. So they definitely have not been recording this data for a long time, but in the last decade or so, they have gradually been adapting more and more technology and they have been understanding gradually the importance of data. Now it has become easier and cheaper than ever to store this data. So, the last several years is when this data has become truly available and available for companies like us to use.

Also something else that I want to add is, I think a lot of these – So, a store like Safeway, definitely has a lot of technology, but something that we noticed, and that's one of the key ideas or key observations that are important to our business is that a lot of these technology has been developed for non-perishables. Because, traditionally, most of the profits of these stores have been from non-perishables. These technologies for non-perishables, they don't work for perishables, because the data may be noisier and then the decision is more difficult and needs to be more accurate and there are some things that we don't observe in the data and we need to collect our kind of tool basically via this app that I was describing to you. Yeah, even though they do have a lot of technology, there is a big lack of technology for fresh perishable food, and that's a really big problem that we're trying to solve.

[00:28:27] JM: Yeah, the data dump problem brings up an interesting and generalizable question I have for you. So, when they give you this big data dump, I imagine there is a lot of data cleaning that you have to do. They probably have some multiple different systems internally. There's probably some kind of data matching problems, like kind of figuring out how to join one dataset to another.

But once you have that solved and you do your proof of concept with them. The advantage is you can say, "Look, that took us a longtime, but now we understand your systems really well. Here is like five different ways you could make your systems more efficient. We want to help you do that." You could become a kind of systems integrator for these grocery stores as well. So it's a very interesting proof of concept, like initial engagement with the grocery store company. I think this maps to what a lot of different SaaS companies are doing right now, like infrastructure SaaS companies. Whether they're selling to banks or selling to insurance companies or selling to oil companies. They kind of have to do this systems integrator kind of like consultative approach, but that consultative approach gives them a really, really good in to developing a close relationship and figuring out, first of all, what do you need from us today. Second of all, what are the SaaS tools that we should be building that will be desirable in five years?

[00:30:04] VK: Yeah, that makes sense, and I agree with the part of what data cleaning does. So there's a saying in machine learning that 80% of the work is about understanding, kind of getting the right set of data, and then 20% of it is actually modeling part.

So, in our case, basically what we found is when we start with a supermarket, often not all the data is good, but we can really separate well the good parts from the bad parts. So, off the top of – When we get initially this dataset, we can often cover about 50% of the store just without doing a lot of complicated data work. Then we can really quickly, by investigating the remaining 50%, we can really quickly kind of go and identify the good parts of that data and do some cleaning only on the data where we need to do this work.

Again, there are a lot of very standard ways of doing this that are common across different supermarket chains. Effectively, when the data is bad, is often bad in very similar ways. We can easily reuse our existing cleaning modules to new customers, which makes cleaning the remaining fraction of the data that is good relatively easy and easier with each new customer that we're facing. We're actually putting a lot of thought into how to architect our software that we use for setting up our ideal pipelines such that it's maximally easy to [inaudible 00:31:38] process and quickly apply our existing clean rules that we know to the data of our customers.

[00:31:45] JM: Describe a typical real-world problem with a grocery store that you can solve with machine learning.

[00:31:52] VK: As I mentioned, one kind of the really, really, really important problem is around their ordering, and that's what we've been focusing on the most. So, how much of a particular good should I be ordering on any given day? So how many tomatoes or how many avocadoes or how many pounds of salmon should I be ordering into the store every day that I'm making an order? This is a huge problem again, because the only information that they have right now is maybe some numbers that tell them how much they sold on the same day a year ago or how much they sold last week. Then they have to go look at these numbers. Then using their kind of intuition and what they remember from their recent sales, they have to guess, "Okay, how many more this item do I need to order." Again, they tend to typically order, which causes waste, which is really bad for their bottom line, and using data-driven methods, is much more efficient.

[00:32:53] JM: Can you talk through in any more detail like how that would apply to a specify to a product? Let's say I wanted to better figure out how many avocadoes to order. I want my automated system to tell me that rather than having a human walk up to the avocado area that two hour intervals throughout the day and try to gauge how fast the avocadoes are being consumed. Walk me through like what would be the development and the deployment of that kind of model and how it would be consumed.

[00:33:24] VK: Yeah. So we've built several models. I think the way that we've been breaking up this problem is – So there's this concept in kind of machine learning called model waste planning, which means you perform decisions by first learning a model of the world and then you solve an optimization problem where you make the optimal decisions under this kind of model of the world you have incurred from data.

So, the modeling problem is more of a supervised learning problem. So, we basically train a realistic model to predict how much they will be selling given various input parameters. So the key source of uncertainty about the future is how much are they going to be selling a particular given good. That's a function of many parameters such as the price, the day of the year and the weather, any kind of holidays, events. Whether it's an add. There's a lot of parameters, some of which we can control, some of which we cannot control. So we need to learn based on their historical data the mapping of all these variables.

Now, the output of this model is a probability distribution. Yeah, it's a probability distribution over a given horizon. Then this basically defines then the expected utility of the store given the decisions that they were making. Then our goal is to find the best sequence of decisions that optimize the expected utility or the expected objective function of the store under this probability distribution.

So that's kind of the high-level algorithmic idea. Yeah, I guess I can tell you a little bit more about how we implement this or how we deploy this. Maybe you can actually tell me what do you think would be the most interesting.

[SPONSOR MESSAGE]

[00:35:09] JM: Deploying to the cloud should be simple. You shouldn't feel locked-in and your cloud provider should offer you customer support 24 hours a day, seven days a week, because might be up in the middle of the night trying to figure out why your application is having errors, and your cloud provider's support team should be there to help you.

Linode is a simple, efficient cloud provider with excellent customer support, and today you can get $20 in free credit by going to linode.com/sedaily and signing up with code SEDaily 2019. Linode has been offering hosting for 16 years, and the roots of the company are in its name. Linode gives you Linux nodes at an affordable price with security, high-availability and customer service.

You can get $20 in free credit by going to linode.com/sedaily, signing up with code SEDaily 2019, and get your application deployed to Linode. Linode makes it easy to deploy and scale those applications with high uptime. You've got features like backups and node balancers to give you additional tooling when you need it. Of course, you can get free credits of $20 by going to linode.com/sedaily and entering code SEDaily 2019.

Thanks for supporting Software Engineering Daily, and thanks to Linode.

[INTERVIEW CONTINUED]

[00:36:38] JM: Why don't we actually take a step back? Let's talk about the software stack of Afresh more broadly and then we'll revisit the machine learning side of things.

[00:36:47] VK: Okay, cool.

[00:36:48] JM: So, like you've got a bunch of different problems you have to solve. You have technology choices to make. So, like you said, you've got this initial point of proof of concept that you need to establish with the grocery store. You get a big data dump. You need to have some way to load that data dump into your data infrastructure. You data infrastructure is sitting on some kind of cloud provider. So you have to choose whatever cloud provider you use.

You're using some data engineering tools. You're using some kind of machine learning framework, I imagine, or you've rolled your own. Then once you build this thing, you build a model. You're able to output the results of that and then you present it to them. Then once you are integrating them with your full application stack, I don't know how much of this you've built yet. But you have like a mobile application that's sitting on a tablet somewhere. You've got a backend server that's sitting somewhere. You've got to have a place where you host your models. You've got lots of choices in cloud infrastructure and SaaS services.

I know we can't go through your entire stack, but can you outline the contours of it and maybe just point out any interesting technological decisions, any software choices, vendor choices, that were surprising to you that you ended up making.

[00:38:07] VK: So, I guess there's maybe three or maybe four high-level pieces. Obviously, one kind of – Or like three or four classes of – Or like three or four high-level sections of our system. Obviously, there's the frontend part. So, we have a mobile app that we kind of deploy in the stores, which is connected to an API backend that powers this app. So there're some interesting design choices that need to be made. Some of these are a little bit – I think some of these maybe a little bit proprietary, because we think that kind of the user experience is something that's really, really kind of interesting and differentiating. That thing, I will probably cover at a higher level than the other parts.

Then another part which I think sits in the middle of our system, it's our data backend. So, we receive a lot of data from the customers, and an interesting technical challenge is how do we actually store this data effectively such that it's easy to access by a machine learning system and they have API backend and such that it's – Yeah. We can safely and easily access this data. So there's a storage backend in the middle.

Once that's out, there's an ETL system that feeds data, that daily dumps data into this data backend. So, there's actually an interesting question on how we design this ETL that connects from the customer and feeds data into the system. Then, finally, there's a compute backend which sits closely with the data backend, where we want to launch our machine learning jobs and where we want to – So, doing both training and entrance and then, yeah, save the results of all these runs to the data backend. Yeah, if I were to put it in four parts, I would say frontend, data, ETL, compute. I can go into these if you want.

[00:40:12] JM: Yeah. Why don't we delve into the data infrastructure side of things? We've done all these shows about data infrastructure recently talking about different streaming platforms, different machine learning frameworks. There's a whole lot of, I guess, development going on in the data infrastructure space right now. I would say there's much – I mean, there's a lot of churn in the kind of Kubernetes infrastructure space too. But I think it's nothing compared to the data infrastructure space, where you have so many different options, very few best practices that have been developed. Give me your take on data infrastructure.

[00:40:47] VK: Yeah. I think that we've been approaching this problems steps, and I think, initially, we just did the easiest and fastest way to get the first system out there and it was not optimal. We've been gradually iterating on this initial growth type and adding features that we thought are the most crucial, and I think we're still in this process. So I can tell you maybe a little bit about the past and about the future.

Basically, the very first thing we started with was big data warehouse, where we store all these big data dumps that we get from our customers. So, we're currently running on Azure, and so we use their service for that, and we experimented on a few others. We've used a few others at various points, but because of a combination of kind of technical and business constraints, we went with that option. Then have been gradually adding additional components, which is like additional data storage components and kind of creating [inaudible 00:41:47] layer or the machine learning system to be able to access this data effectively.

Then I think some of the future vision that we have is around using technologies such as Spark. So we're actually currently architecting our system to use Spark both for handling the ETL and the data warehouse parts of it. I can tell you some of the business requirements, or – Yes, some of the –

[00:42:16] JM: Yeah, please.

[00:42:16] VK: Technical requirements that are driving this decision. So, as I mentioned, something that's really crucial for us is to be able to quickly integrate with our customers and quickly get good data from them. This needs to be, basically, one important way in which we can do this quickly is by reusing components from our earlier – Basically, by reusing components across different customers. By using systems such as Spark, which has a lot of really nice APIs around, say – We're mostly working with Spark DataFrame API. That allows us to structure our code and architect it in the form of modular and reusable components that we can easily deploy for other customers in the future. Basically, reusability is one big thing that we think we're getting form Spark. Another one is –

[00:43:12] JM: Sorry. That's model reusability? So like if you build a model for avocadoes, you could reuse it from one store to another.

[00:43:18] VK: No, not quite. These would be things that – For example, we sometimes see that an item is being sold under to ID numbers, for example.

[00:43:28] JM: Ah! Oh! The PLU numbers.

[00:43:30] VK: Yes. So [inaudible 00:43:31] internal identifiers. So that's a common problem. But now we're at the stage where we can really quickly fix such errors by using standard cleaning rules where, well, okay, these two data feeds perfectly match with each other if we – And then with this other data feed. Then we can combine them in a certain way. Then we can just kind of add that module that is written in a general way that's not specific to any customer.

[00:43:56] JM: So the double PLU number question, it raises something in my mind that's kind of an adjacent topic to this whole software thing. I feel constant anxiety whenever I go into a grocery store, because I have no idea if I'm supposed to get the organic or the nonorganic, because I worry that the organic title has been entirely captured by kind of just the regulatory changes, like ways that people can get around the spirit of organic, and the markup is oftentimes 2x or 1.5x the price.

Given that you have a unique view into the supply chain of groceries, do you have any perspective as to how contaminated or how captured the idea of organic produce has become?

[00:44:44] VK: I think whether organic food is superior to the nonorganic food, that has like a very, very long – That's a very long discussion. But in terms of their ability to label this data, I think it's actually pretty – I'm pretty confident that when – That the organic items are labeled properly, but you might have some issues and that if you go to the cashier and then they scan your organic or nonorganic apple, it might get scanned under the other code that I think you might see. But I will personally not be too worried that before the cashier, the label will be wrong, that I will be a bit less worried.

[00:45:24] JM: All right. Fair enough. A couple of more questions related to actual software. You've done research on accurate uncertainty and deep learning, and I think this is a very interesting term, accurate uncertainty. As I understand it, it means you are trying to find how certain you are. What your probability distribution is. Rather than just say, "Yeah, I don't know how certain we are. I don't know what the actual uncertainty is. I just know that I'm on unsure." You're talking about finding an actual measurable numerical version of explaining your uncertainty. Explain the term accurate uncertainty using an example from the grocery domain.

[00:46:10] VK: Yeah, sure. So, this is something – The notion of uncertainty and using probabilities is something that's really important and really essential to the machine learning part of our system. It's this idea that – Well, first of all, what is uncertainty and why do you need a probability distribution?

So let's say that I need to determine how many apples I'm going to buy. If today I need to order apples, how many apples should I bring into the store? So this really depends on the demand and how much am I going to sell.

Traditional forecasting method, like ARIMA or point average, things like this, you would get a single point estimate of this number. So this is not sufficient for our purposes, because we also need to transform this into a decision. What we actually need, it's a distribution or all the possible outcomes that will happen in the future. With some probability I will sell 10 cases, but some probability I will sell 9 cases, 8 cases and so on.

Then by assigning an accurate probability to each of these outcomes, I can then compute the expected value of any kind of decision that I will be making, and this allows me to actually make much better decisions, because I'm actually much more – I'm tracking of what's going to happen in the future in a much better way. That's one reason why the probabilities are important, and that reason has to do with the model kind of knowing what it knows and knowing what it doesn't know and it's really important when we're building a system that works with a human operator. Then the system can hold off from making a recommendation if it sees that it doesn't have enough data to make a good decision. So that's another very important implication.

Then you ask whether you need to have accurate uncertainties. Well, effectively, these probabilities, they should correspond to the true, real-world probability of what will happen once you make this decision. So, this is where – The way that we're thinking of this is that kind of the empirical and the predictive probabilities in the long-term should match.

So if I assigned something, if I assigned an event, a probability of 90% to an event. Well, then it'd better be true 90 times out of a hundred times when I made this prediction. So that's something that we're baking in into our system and it allows to make better decisions.

[00:48:25] JM: All right. Final question; how will robots change grocery stores?

[00:48:32] VK: I mean, there are many applications. I mean, there are some applications which are applications of robots. Now, even in kind of midsized supermarkets, there is a lot of exciting work that's been done by companies to try to use these robots to assist – To actually count inventory. That's another big area where they're being used. Then I think in bigger stores, there's of course a lot of automation in the warehouse and things like this that are happening.

I don't think that robots are going to need overtaking grocery stores, because, ultimately, the differentiating – What's going to really differentiate the supermarkets is going to be the quality of service and the human element and whether you will be – The friendliness and the helpfulness of the staff. So, I think that in the future, as we automate a lot of the things that machines are good at, humans will have more time to do the things that they're good at, and these are things like helping a customer and ensuring a really good experience. So, I think, basically, the world will be good because humans will have more time to do the things that they're good at, and the more tedious decisions, those will be left to machines and algorithms and robots and so on.

[00:49:43] JM: Volodymyr, thanks for coming on the show. It's been really fun talking to you. I think Afresh is a really and a futuristic company. So, best of luck.

[00:49:50] VK: Awesome, yeah. Thank you, Jeff. This has been super fun as well. Thank you for having me. Thanks.

[00:49:54] JM: Awesome. Thank you.

[SPONSOR MESSAGE]

[END]