The Lessons of Fin for This Generation of ‘AI Assistants’: The Real Problem is Managing Real World Liability as a Service vs. a Person

@lessin

In 2015 Andrew Kortina and I started ‘The Fin Exploration Company’ and set-out to build a human in the loop AAI (artificial artificial intelligence) assistant.

Our theory at the time, which we wrote in our ‘charter’ document was that lots of smart people were working on real AI, which would be coming, but wasn’t practical yet — but in order to figure out we would build an assistant service to figure out how to practically execute the ‘future’ of an on-demand assistant, that clearly everyone wanted, but outside of super rich people with full time actual assistants - no one had.

> “We believe that people will someday interact seamlessly with the universe of information and services through natural speech assisted by an “agent" that knows everything about us. The final interface is no interface. This is not a novel vision; it is one that has persisted through generations of science fiction. Many organizations are currently working on the deep problems that should eventually - given enough time and effort - deliver this future.” - Fin Charter 5/2015

Fast forward to today, and I get pinged weekly by folks looking for the ‘lessons’ of the Fin Exploration company, as they embark on their own journeys to build ‘AI assistants’ in various domains, with new fancy LLMs which were a lot of what we anticipated and wanted to be prepared for when we went after the problem years ago.

I do think that LLMs are cool technology that make a lot of things we did at Fin easier to execute. I do think there are niche domains where assistant services can work.

However overall, my advice to almost all of those that ask is that while I get the temptation to take a shot at assistant services - They will fail — for some non-obvious reasons.

Background Context on The ‘Fin Exploration Company’

When Kortina and I started Fin, there was a wave of messenger based ‘bot’ services popping up that were making wild promises about their ability to deliver great product experiences through a chat interface.

We got the marketing. Yes, we all want ‘Samantha’ from the Movie Her… we had all watched and read the same science fiction…

But we also knew way too much about the practical realities of these services in 2015 to at all believe the hype of the time. We also remembered the good old days of 'active buddy' on AIM - how many times this had been tried.

That said, we had time, investors willing to fund what we told them was likely a money losing 'exploration' … and we were very interested in the design challenge of the tech you would need to build to make a good assistant service experience — assuming you could 'plug' a flesh and blood 'Mechanical Turk' human in for the decision making and interactions that were no where near a reality in 2015.

Over several iterations we built a lot of very cool tech to chip-away and solve many of the problems of human in the loop assistance. A lot of this actually had some 'AI' to it (we called it machine learning back then. And it worked pretty well. Examples included:

1. Personal Identity & Preference Management and Recall
"get kara flowers" sounds like a simple enough request…"

but if you start actually trying to execute that there are many questions…. Who is 'kara', where does 'Kara' live… what preferences do you know about how much you want to spend on flowers for Kara in context. What colors do you like, what about her? What is she allergic to? How urgent is this? Is there a specific florist you like in the region? How would you likely trade your favorite florist, or flowers, vs. speed of deliver ability, reliability of service provider, etc.

If you are going to run a great assistant service the first thing you need to do is collect a lot of preferences…. And then figure out how to guess which ones apply to a given situation / apply them properly to requests as they come in. It isn't easy…. But we built a pretty great way to collect preferences in-context, store them, and then recall relevant information in the context of requests for 'agents' clicking the buttons to execute orders.

3. Task Queuing, Routing and Time Estimates
If you do tasks FIFO you are going to have a bad time. So, what order do you do what requests in? How long do you tell a person their task is likely to take to completion (and with humans how much human attention it will take). Given what agents are online and their relative experience with vendors or tasks when do you route back to them vs. new folks. How do you save and recall the 'state' of a task, etc.?

We built a pretty cool engine for taking tasks in, understanding what they were about, estimating how long they would take, and then trading around the order of execution on a given task (or sub-task based on 'triggered' work / a needed step in a process beyond the four walls of fin itself with an outside service provider.

The fun part about this — is that we did a lot of machine learned work here long before this generation of AI existed.

Unlike preference management which is still highly relevant, this part of what we did is somewhat less critical in today's world because rather than dealing with expensive human attention, and managing scarce human 'cores' in today's world there can be more machine scalability here… but i will say that many of the folks who come to me for advice still are using humans in the loop, at which point all of this becomes critical.

4. Recipes for executing tasks
Great, so how do you actually 'execute' a given task once it is time to do it?

We built an engine for scripting the steps / processes of various tasks that an assistant would need to complete (from scheduling all the way to booking flights, etc). This included collecting any of the missing critical preference information — the 'steps' to follow which we learned for getting a good outcome (when to check in with user or ask for feedback), what to do when you hit problems or delays, etc.

This next generation of assistants probably can auto-learn / do a first pass at these steps better than we could in our era .. but my bet is that for a long long time any assistant oriented service using LLMs is going to need a recipe engine to script and control how the steps in a task are completed.

5. Shared CRM of People and Services; How to 'Get Smarter' at Scale
On the internet every company and service screams at you that 'they are the best at X'… but really, who does good jet-ski rental in Key West? What actually is the best sushi in Palo Alto?

One of my favorite parts of Fin is that by having a central service doing tons and tons of tasks for people, you get to build a real index of the best services and people in the world at given tasks, and re-use those … this is way better / more interesting, and more powerful than the marketing you get online or online reviews. People you work for inform you of what they know in order to get things done, and with every interaction you get smarter for the whole 'network'.

Those who know me will recognize that this was my absolute favorite part / dream of getting fin 'right'. We built a pretty cool open-ended system for leaning, structuring, and recalling quickly and efficiently these types of answers far beyond some dump wiki :)

6. Ranking and Feedback Loops
So you do a task, but did you do a good job? And how do you do a better job next time

The QA of assistant service tasks is also pretty interesting problem we spent time working on (solving would be strong but we made a lot of progress in our time). You do get user feedback at the extremes of good and bad, and that is useful — and you can solicit star feedback and get some sense of density, but any system you build for open ended assistant work is going to need constant improvement and building tools and process for driving that QA and improvement loop was another big spot of work.

7. API integrations for efficiency

There are lots of semi-public databases out there to help you… of flights, of hotels, etc. as well as personal integration points into calendar and contacts

Obviously, for specific domains, there are APIs and data-sources that aren't publicly available to integrate with that helps you eek out a bit more efficiency and quality on various tasks. This wasn't a major bit of work for us, but we did do some of the obvious stuff, and while not technically challenging it is technical table stakes to do a lot of this to provide good services.

What Did & Didn't Work About Fin

Fin was quite loved by many - starting with Me and Kortina — and with mostly a group of a few thousand folks in the tech world that as founders or investors had more money than time.

We must have booked millions of dollars of hotels and flights over time (we never tracked), dinner reservations, scheduling, tons and tons of on-demand purchasing where someone would just say "buy me X" into their microphone and it would magically show up at their door. No interface needed, no bullshit, no questions even!

Even better than 'one request' interactions, were zero-work real world interactions. People in particular loved setting up 'recurring' tasks where they could fire once 'do X for me weekly' or 'send my mom flowers once a month' etc. where you went form a 'one click' interaction in the real world to a zero-click one.

It was a great experience in the many cases that it worked well, and for people who didn't mind that it was expensive. But ultimately we had trouble figuring out how to scale it with good margins and quality.

So, after a while, we chose to spin down the assistant service and go deep on a tool-chain for instrumenting and optimizing humans doing operations work (I think at peak 50,000+ customer service agents were on the platform as we helped big companies understand work process cross-app and tool to optimize quality and efficiency).

Looking at the world of today there are things we did / spots where LLMs do really help and would make a fin-like service better.

I think in particular todays LLMs would be better than even we were at preference collection and structure, and getting the full 'context' of a request being made.

That said, i think there are quite a few places where the challenges we figured out at Fin with running an 'assistant' service persist / are not at all solved by LLMs… and the all broadly boil down to this fact:

When you get between a customer and an open ended set of services with real world consequences that you don’t control, you are assuming a massive amount of liability for outcomes you can’t guarantee.

1. **"The 1% Blame Game -** The Assistant Front End Will Get Blamed for Any Downstream Service Mistake" — it doesn’t matter how good you get, when you start using a digital assistant to interface with the real world, there is going to be a non-zero error rate. The particular problem for assistant services is that (a) tiny error rates if you use a service constantly end up showing up constantly and (b) the more the service ‘chooses’ for you (which is valuable), the more liability it takes on for the choice. You blame whomever you told to do the task, not who they subcontract to - so the problem becomes that any assistant business ends up assuming a ton of liability for its choices, which either leads it to be super conservative and limited (not useful), or be constantly making mistakes.

2. **“Depth of Error Impact in the Real World** — When F-ups happen, you as the front end are on the hook for big tickets” — the most useful stuff assistants can do are in the real world with real money consequences. Book this business meeting at a great restaurant, get me a car, flight, hotel, etc. And when those things go wrong, they go terribly wrong with real world time and money consequences. The surface area of cost and consequence is dramatically dramatically bigger than it is when your Doordash food is cold. The problem is a financial one. Every time you take on a high ticket item you are basically taking on the liability of a big refund or huge fight if you don’t deliver … and if you are just getting paid as an ‘assistant’ vs. something like a 10% or 20% commission on everything, you just don’t have the cash to give reasonable compensation to your customers and keep your model working.

3. **"I could have done it better / faster myself"** — Anytime a service isn’t perfect, the customer / client thinks ‘I could have done it better or faster myself’ … even when that is 100% untrue. This was a constant issue with Fin both on quality and efficiency. To a hilarious extent. People always think they can do things 2x faster than they really can, and are dramatically better at finding answers or the right service than they are… but again, when someone gets in the middle of that transaction in most cases the liability assumption even when things don’t go terribly wrong creates constant stress.

Conclusion: An AI Assistant Ecosystem is a Beautiful Dream … but a Long Way Off And It Is Hard To See Successful Businesses Being Started Given Low Switching Costs

The dream is beautiful.

An ‘AI assistant’ stores who I am / my preferences and context perfectly and acts on my behalf. It is faster than me at using Kayak and finding all competing bids on flights. It knows how to pick a hotel room and when I would prefer an airbnb. My mom always has fresh flowers as though i had time to call the florist every week and send ‘happy Monday’ note.

It feels at first blush tantalizingly close — Storing preferences sure. Going and using websites for me, is infinitely doable…. And we know it is awesome because some people have PAs and love them!

Beyond theory, it feels like if it starts happening it is just the start to a way better world. Let’s drop the stupid websites of people marketing services to each other mediated through google and instagram! If you are a chef, just tell the platform you are a chef and where you are and get a few folks to vouch for you… and you have a business.

If you are a person the AI can know from real experiences of real people what the ‘good’ services, and since it works for the customers not for the businesses, the business model is flipped and it can find you way way better stuff than you can from marketing pages online!

At scale it even fixes capitalism, because the service ‘truth’ engine of who delivered and who didn’t — with a tight coupling to the customer and real feedback vs. stars changes everything for a wide wide swath of industries.

That said, I am really not holding my breath.

Practically these services are going to be fraught in terms of managing and dealing with the risk of acting on behalf of people.

Service providers will mess up. The front end ‘assistant’ that picked the provider will be blamed. If the bouncy house guy get a flat tire, then the customer will think rightly or wrongly ‘I would have known not to trust that guy’

People will lose money, cars wont show up, the food will be cold, etc. and just as DoorDash eats the cost when a delivery goes bad, the default will be the ‘assistant’ eating costs — which get very expensive very fast when you go beyond dinner.

People will be unhappy with choices even when the service was delivered as planned. If the rafting tour wasn’t fun enough, the assistant will take the blame and the customer / user will think ‘I would have known this was going to be boring if I had done the task myself!’

So in this respect, the problem isn’t really with assistants, it is with most people.

And the worst part of these ‘assistants as AI businesses’ … is that unlike real world assistants there is going to be zero loyalty.

Real assistants f-up all the time. They also pick services that then under-deliver…. Even the best ones. But the reality is that if you work with an assistant they have so much context in their heads that it is very hard and expensive to switch. You stick with them even with the error rates because the switching cost is too high.

But digital assistant companies? Well if you assume (as I do) that your preferences will be 100% portable thanks to things like GDPR / download your history… .then you can boot up a new digital assistant service whenever you want. The data is not defensible.

And in this respect, when it comes to digital assistants there will be literally zero loyalty. You will fire and change them constantly, which is a very very hard way to build a business.

Will there be any assistant verticals that work? Sure… I am not sure they will be high margin, but there are places that will work (heck if I had time i even have a few ideas on 'small' ones i would want and would work well!) — But the beautiful dream from 'HER' where everyone has a personal assistant and no one bothers to ever click on websites anymore, that is miles and miles away.