

The AI Assistant (Part 2) - ProductivityCast

[00:00:00] Are you ready to manage your work and personal world better to live a more fulfilling, productive life? Then you've come to the right place. Welcome to ProductivityCast, the weekly show about all things personal productivity. Here are your hosts, Ray Sidney-Smith and Augusto Pinaud, with Francis Wade and Art Gelwicks.

[00:00:18] Welcome back, everybody, to ProductivityCast, the weekly show about all things personal productivity. I'm Ray Sidney-Smith. Marco is jumping out. And I'm Francis Wade. Welcome, gentlemen, and welcome to our listeners to today's episode, where we're gonna continue our discussion on AI, and this is our series on the AI-powered professional.

[00:00:44] in our first episode, we started the discussion about the concept of utilizing generative AI. in this episode, we also started the process of talking about what an AI assistant is really [00:01:00] like, talking about some of those administrative frictions, being able to get rid of, and automate that out of, your world to some extent, and dealing with shadow work as well, defining shadow work and so on and so forth.

[00:01:13] We're gonna continue this topic into discussing today about really how to partner with your AI in a lot of ways, what the collaboration process really looks like. And so I'd like for us to discuss shifting using AI tools as a mechanism of just kind of offloading something, which it can do, but then becoming a more collaborative partner with that particular AI tool in order for it to become a true AI assistant.

[00:01:45] And so I'm thinking of things like how do we ensure that AI is taking over the right kind of work and that it's not taking over the work that we should be doing, and how do we maintain control and accuracy? And of [00:02:00] course, there are a bunch of boundaries and ethical considerations that we should be thinking about and some thoughts about the future.

[00:02:05] So let's start with what are some of those first principles, for us to be able to create a true collaboration partnership with our AI assistant?

[00:02:19] Sure. I'm thinking about this from the perspective that If I want to work with my AI assistant, I need to choose particular categories of work in

which it can actually collaborate. So for example, I want it to be able to help me take a rough sketch that I've made on either my iPad or on paper, and then to have the AI turn that into a full-fledged drawing, a full-fledged cartoon perhaps.

[00:02:49] So the AI assistant is acting as my cartoonist, and so that's a role that I want the AI assistant to do. And while I can draw my [00:03:00] own cartoons, 'cause I've taken this drawing class, I feel competent to draw, you know, one part of a cartoon, but then it can fill in the rest by creating the other panels of the cartoon.

[00:03:13] And this is really helpful to me because now I can make the first drawing. It can be roughish, you know, to give it the idea of what I want, and now I can help it help me, quickly generate more panels and get the cartoon done by virtue of that. But the idea is that it's now a role that I want it to continually be helping me with, and so that is the cartoonist role.

[00:03:38] That's just one. I mean, like that, it doesn't, it doesn't have to be just role. It could be any number of things. But it's just like, that's the kind of thing that I'm thinking about. well, in the last episode, we sort of established the notion that, an AI assistant is like an intern who remembers everything, but doesn't have a whole lot of judgment.

[00:03:56] isn't, a really good judge of, you know, the [00:04:00] things, whatever it is that we happen to be expert at. It, it's too much to ask the AI to rise to our level of, insight and understanding. Having said that, there's a whole bunch of stuff that now looks to me that, it looks different to me because I can now see it as automatable.

[00:04:23] Like the example that you gave of, doing repetitive drawings or repetitive, animation. There's a bunch of things that I, and the list keeps growing, which is why I don't have a fixed answer. but it does start with this notion that I have an untrained intern that has infinite memory and infinite patience and doesn't have an attitude and works at all hours.

[00:04:48] And if I train that intern, then there's more and more things that the intern can do, and there's gonna be a new app tomorrow that- allows the intern to [00:05:00] do even more. So it's hard to say what specific role because the roles keep changing, and they keep being added to. if anything, I would say there's maybe a rule, which is that, try to give the intern as much as possible, but always be the person of last kind of decision.

[00:05:20] Be the one who's at the end checking to make sure the intern didn't make some, you know, gross error. So if there's any rule, that's the rule that I'm applying right now. Try to find more and more to give and then be the person at the end to do the checking. and then don't try to stress the intern out with judgment calls.

[00:05:42] and even the limit-- even the line on what I call a judgment call is changing with AI because it's getting better, You know, the AIs that I use, I use memory, so it understands me and what my judgment calls are, better and better each day. So it's a tough question to answer.[00:06:00]

[00:06:00] So just stepping up a level, I would say that just the concept of establishing roles for the AI is the first principle. It's not necessarily that you're going to ever be exhaustive in terms of creating the roles, because sometimes the role you need for a specific chat is defined in only that chat, and then there will be ones where you're gonna need that as an ongoing kind of recurring thing.

[00:06:29] It depends. You know, last episode I was talking about that wine help. You know, help me identify wine that I may enjoy based on my profile and educated that profile. But same thing on, on the professional side. I have a client who we, because of what they do, they, it's a report that is run every morning, and that report gets to them.

[00:06:53] And the problem is it's impossible to, to analyze it long enough. You know, you can see the report daily. You can maybe go a [00:07:00] couple days back. But human, it's hard to really create trends and things from that specific report. Where it's been very cool is a play on the role we create a chat for that neural network, okay?

[00:07:15] And now that report is dumped, for lack of a better word, into this chat. But this has now allowed us to identify trends not in three months, not in 90 days. Hey, this server last time this failed, okay, it was seven months ago. And it failed for three days. That information no human can provide for me. Okay?

[00:07:38] But allows you to start seeing that, and that make it very, very specific. Okay? Same thing, when you write. After you train, yeah, it required to train the intern, but after you train, say, "Okay, this sound like me. This doesn't sound like me." You know, one of the things that I love to do is when I get an [00:08:00] idea, okay, let me discuss this idea with the content of, okay, or the ideas or the understanding that AI has of X person.

[00:08:08] And you can say, "Hey, I want to look what will be the perspective of this text if Einstein read it." assuming you, you know what, physics and stuff. But that give you... Is the perspective you're going to get accurate? Well, it may be, it may not. But it will give you a counter that is very interesting.

[00:08:30] One thing that I do very often is find the arguments in favor and against this i- this concept, this idea that I'm working on. And it now get... You know, I think the definition, part of the definition or the issue is this, for a lot of people, is the first time they get access to an assistant, to an administrative assistant,

[00:08:53] For most people, that is a concept that they heard, that they, you know, they know what what the [00:09:00] idea is, but they have never experienced this. So because you haven't experienced this, for a lot of people, the first step is define what are these things? What is what you want to accomplish? No, you cannot have a chat that do everything, okay?

[00:09:17] Or you need to create what is, what make them special. You know, for example, I have for the cars, okay? For the vehicles we have at home, there is a chat, okay? Then I went and described, "This is the car, this is the year, this is the characteristics." Okay? But then now, you know, the other day, I got a text from my wife.

[00:09:40] "I have this light." Give me more information, right? Because that's all that I got. But because I had the image and I have the chat, I was able to copy-paste. Okay? Drop it there, say, "What is this?" Okay? And it give me the answer based on the vehicle. Okay? Not a, not a generic answer. That light may be [00:10:00] whatever.

[00:10:00] What is what Google will have done. I got a very specific, "On your model, this means this, this, and this, and this is what you need to do to fix it." Okay? Do I need to use that chat often? Hopefully not. Okay? Otherwise, we will create a new one with the new vehicle. But at least You can get this. But I think the problem for this comes one step before.

[00:10:25] We need to understand what is that partnership that we want. What is that agent or chat that we are looking for? because if I say, "Okay, I want to have a sommelier," for wine, or what I want to have is a bartender. Those two, the characteristics are very different. If you want to have a tech discussion, I have one where I discuss tech,

[00:10:48] So I want it. Why? Because it takes seconds to say, "Okay, give me the stats of this device," now how will this interact with this, this, this, and this?" It's faster than me going into the [00:11:00] web, trying to find the specs. Do I n- do I have an idea of the specs? Yes. Do I want it precise? Of course.

[00:11:06] Okay, and that's where these partnerships, this assistance can be incredibly powerful. But as I said, I think the problem for many people, it's one step before. They are trying to build partnerships without understanding what is this assistant, what is-- what did you need the assistance with? It's not the same thing to say, "I have an assistant," okay?

[00:11:33] That really understand what this virtual AI or human person actually does. Yeah, I think that's really interesting because in the human world, my first assistant, my first secretary, I had at a fairly young age, and so I've always had some kind of assistant in my work world. And this was the kind of situation where this person, like myself, [00:12:00] is someone who just figures it out, right?

[00:12:03] It was never a need to say, "Okay, well, this project needs to be done, and I know you may not know all the things that are necessary to get it done." It was just like, "Here goes this project," and, off they went. And, they came back and maybe there were some parts right and some parts wrong, but it got done, right?

[00:12:21] this person leaned into progress, and not everybody is like that. And so we need to think about this from the perspective that with these LLMs, they are not like that. they have, a broad base of knowledge, and in order to get the best response from them, lowering the error rate, we need to give them a specific domain in which to work.

[00:12:45] So if you ask it to do something, and you ask it any generic thing, then it has to do a lot more work to get to that point. Also, just from, an environmental perspective, the more you ask it generic questions, [00:13:00] the more work it has to do, and so the more you're gonna kind of waste compute power So the more you give it a domain, hopefully, and this is just a hope, the less work it has to do to get to that answer, and therefore, hopefully the less, environmental impact of that particular chat that you're having.

[00:13:16] I'm not sure that that's true, but it just makes me feel better knowing that I'm giving it specific boundaries and hopefully reducing my environmental, imprint. but I kinda have three things that I tend to think about when it comes to that. We've already kind of talked about this first piece, which is the role, which

is just knowing that your AI assistant needs to have a role every time you open a conversation with it.

[00:13:42] And if it doesn't have that role, then it lacks a focus, right? It lacks a domain in which to look at. Also, just as a quick tip, you can give it multiple focuses, right? So you can say, "Hey, I want you to be, a medical doctor who also has a specialty in sports [00:14:00] medicine and is also someone who does CrossFit," right?

[00:14:04] And so they're a CrossFit athlete, but they also happen to be a medical doctor with a sports medicine, specialization. And by virtue of giving it these multiple things, and I've given, some chats very disparate, forms of knowledge, right, to look at. But it's because I want to see the things that are patterns between these disparate areas to help me understand some new topic, and that's the beauty of them.

[00:14:30] You can say, like, you have a PhD in economics, and you love horticulture. so, all of a sudden now they're capable of synthesizing, those two domains in a way that you couldn't do, very quickly. So that's number one is just, like, you're gonna always and ever be defining specific roles for the AI assistants, but you have to get good at being able to say, like, "Wait, who would be an expert in flags?"

[00:14:55] Right? And, that's a vexillologist, I think. but it's like, you know, you have to do a little bit of [00:15:00] research to be able to figure out, what the heck is a flag expert? maybe you just say, "Flag expert." but you know, what would be really better is to actually use the words that are going to tell it that it's an expert.

[00:15:14] And so the goal here is to kind of think through how to effectively tell it what it is, and those kinds of titles and what it does really helps you. Because if I say to it, "You're a flag expert," yes, it understands that. But if you say it's a vexillologist, then it's like, "Oh, you know what? I, I, I s- I'm a scientist who studies the history and symbolism and usage, as well as the design of flags."

[00:15:41] And now all of a sudden it is a completely different animal. And I just see it's a subtle but clear difference in the way in which when you call something the true title of it, that all of a sudden it's just better. Second is you need to actually create the domain. You need to give it the resources, [00:16:00] for certain things.

[00:16:01] So if, for example, you are a domain expert, but you want it to mimic you, then you need to give it as much knowledge that you've created so that it has that domain knowledge about your work product, about what you've kept as reference material, what you've read. all those things are really helpful for it to be able to mimic you.

[00:16:22] And then finally is the idea of basically a fixed domain or a dynamic domain. So fixed domain would be something like, Socrates, right? Socrates is long dead. He's not producing new material. He's not testing it out in the world. He's static. We don't need to worry about the domain changing, right?

[00:16:44] There's not gonna be a modification of stoicism. There's not gonna be a modification of, of Plato. We can, we know that fixed domain has boundaries to it, so if we give it all of it, then it's fine. So recently I created a, a Google NotebookLM [00:17:00] on Shakespeare. So all of Shakespeare went in there. All of my writing on Shakespeare went in there.

[00:17:06] All of my, my notes from other areas of Shakespeare went in there, and it's fixed. i can add th- new things to it, but for the most part, the Shakespeare part is fixed. what I might add to it is new. And so that's a fixed domain where you don't really have to worry about the dynamism of it.

[00:17:24] The, opposing piece to that is a dynamic domain, where this tool is going to need to reference the web. It's going to constantly need new, fresh information for it to be able to help you. So for example, if you had a digital marketing assistant that you created with an AI tool, you're constantly having to put new information in because there are new trends.

[00:17:45] You have new data that you're gonna be adding from the last marketing campaign or ad campaign So by virtue of all of that, you need to figure out, okay, do I have to attach this to some kind of ongoing connected data? Does it have to constantly search the web for [00:18:00] something?

[00:18:00] Or is it something where once I have the data in there, I don't really have to think about giving it more data every time?

[00:18:07] I've wondered about the business of training, an LLM to respond to you in these specific ways. What I find myself doing by default is using one LLM. per persona, per assistant, like I have an image of the person that's in that LLM, and I only have conversations with that LLM that pertain to that topic or that person or persona.

[00:18:34] So another one has a different persona, another one has a different persona. I don't know if this is the right way to do it, but I do notice though, it's exactly what you said, which is that I'm confining all the conversations of one kind to the one persona. So like to be practical about it, I use...

[00:18:53] I have Copilot, which I don't use, I don't pay for. So I throw all of the dumb questions in Copilot. These are the [00:19:00] ones that are closest to a Google search, like how long does it take to drive from here to there? So the first place I go is Copilot, and I don't know why I developed this.

[00:19:10] I think because I didn't want to somehow pollute Claude and Gemini, where I have paid subscriptions, and I do my hardest work, and they do my articles and content around strategy, which is my major content creation area. So I never ask Copilot those kinds of questions.

[00:19:30] I ask Copilot the kind of nonsense, throwaway, inexpensive questions in the sense that they're not using, you know, they use a few tokens, but what I don't want to do is kind of what you said. I don't wanna distract my heavy lifters with the trivial questions. So Copilot is my Google substitute essentially, and sometimes I use ChatGPT for that purpose.

[00:19:56] but I wonder if it probably would be easier if Claude, for [00:20:00] example, said, "Okay, here's Claude Light over here, here's Claude Heavy over here." And if I could keep the two worlds separate, you know, I don't want the memory to be transferred between the two. I don't want the memory to be shared.

[00:20:14] I would just want to be able to use it in different scenarios. If they had that, it would probably substitute for what I'm saying. But I've noticed that as the memory gets better and better, in Claude for example, it's remembering stuff that I've long forgotten, so it's gotten way more useful.

[00:20:32] And for it to continue to be useful, I use Claude Projects and I'm gonna add on the other functionality that they just come up with. But for it to continue to be useful, it seems as if I would want to do exactly what you said, which is to sort of confine the persona that I am relating to in Claude to some specific idea I have of the person that I'm relating to in Claude.

[00:20:59] I think [00:21:00] this is where some people will excel with AI as an assistant and some will fail in their management capabilities, right? You are now becoming a manager of these tools, and management is not easy.

[00:21:17] I wouldn't have my work with my clients if management were easy. it's just a reality factor. as I'm working with clients, my clients are executives, and so they are people who require an entire title, which is to manage people and to help, manage an organization toward particular missions.

[00:21:38] So if you don't have those management skills, you need to develop those management skills, which, like, the first thing a manager should know are their employees' names. it's like you should know the people who are working for you. And, there is a job description. there are departments, and, like, there's a whole ecosystem and infrastructure in place there.[00:22:00]

[00:22:00] And good parents know this, right? 'Cause if they have a child and then have another child and then have another child, they have to start managing those children. And so it's kind of the same thing. you're starting to have to parent these tools, and the more there are, the more management there is.

[00:22:20] And so I think the core, next piece for me is how do you manage these things? And the good part is that unlike a child, unlike an employee, the moment you leave the chat, they pause, right? You know what I mean? Like, you don't have to worry about food it's like a Tamagotchi kind of situation. You don't have to worry about any of that.

[00:22:43] they are going to just be static until you come back to them. some now can do work on a recurring basis behind the scenes and give you more information. But for the most part, you're not worried about caring for this thing. You just have to worry about making sure that you're managing it effectively when you are working with it.[00:23:00]

[00:23:00] And so, like I said, I think that a job description for your primary personas is pretty helpful. And really what that means is just writing an effective prompt that is kind of the master prompt that anytime you call it, right, if it's on the same chat or if you're creating a project where you're gonna be calling it, you have to basically have this set of instructions that tell it what it is, what it does, what its boundaries are, what it shouldn't do.

[00:23:25] I think most people aren't gonna do that. And so I'm telling you, who's listening, please, please do this and you will be so much better off with regard to having AI act as an assistant, because once you know which assistant you're talking to, which persona, as Francis noted, which persona you're talking to, now you can call upon the right role, the right knowledge, and then fix it to that domain or allow it to broaden in terms of those things.

[00:23:53] I didn't say this before, but anything creative is a dynamic domain. So folks [00:24:00] don't think about it from that perspective, but it needs to be dynamic because it needs to be able to create that level of basically confabulation, right, or hallucination. It needs to be able to, think expansively in order to get to an outcome.

[00:24:11] So, that's a very human thing, and it mimics us. Which it's very-- It, it, this is, this is gonna be a thing because I don't know if you guys have had the experience of y- you've chatted with some LLM six months ago, and it gave you a good answer to a question, and you can't remember which one it was.

[00:24:30] And now you've got to go in and search and look and delve and dive and keyword, and it's like you're doing a Google search of... And in my case, so you know, DeepSeek is like the ugly cousin first out, you know. I don't use... It only comes around now and again. You know, ChatGPT is like the smart, smarty pants who wants to charge me too much money.

[00:24:54] I had this great conversation. Six months later, I recall it, and I can't remember which of the [00:25:00] children, the LLMs I had it with. I had made the mistake early on of using two logins just by accident, to get into a couple of them.

[00:25:10] So now I have multiple logins into the same ones. So it's a mess to try and remember, It's like a parent who can't remember which of my kids did I have this conversation with, you know, a year ago? And I try and remember what... You know, I remember having it, but who do I call to remind me of what I said and what we decided on and...

[00:25:30] But this is a difficult business now of managing these interns or managing all of these personas is now a thing that I now need to account for. like you said, it's like another level of complexity that I didn't even, know even existed. but it's important. It's really important because I save, I'm saving mega time by even if I spend 15 minutes trying to find the conversation that I had a year ago, [00:26:00] the time saving once I find the conversation is amazing because I spent an hour and a half having this conversation with one of these LLMs, and to retrieve it is well worth the search 'cause I can pick up the conversation from where I go through the thread and I'm like, "Wow, there's some good stuff in here.

[00:26:19] I'm so glad I found it." 'Cause now I can just pick up where I left off, and ask the next question, which as you said is the benefit of these things is that

they go and add new features, but they're not working on stuff while you're not there, which is one of the huge benefits of this thing is it's somewhat static.

[00:26:36] this is gonna be more and more I predict, more and more of a challenge for us. And it would help thinking futuristically if we could pick a persona, so if we had a front end to all these different LLMs, and the front end were more like a persona, and if we wanted to go in the back end, we could, but we wouldn't be forced to go into the back end the way we are now.

[00:26:58] We would just deal with, you [00:27:00] know, here's my list of personas and which one do I wanna work with first, and which one do I wanna pick up with the conversation that I was having last week, it would be very human and very useful because I think we're relating to these things like, just like you said, like they're like kids or they're like little interns.

[00:27:15] and they take, to our surprise, they take management. I've not read anybody talk about this anywhere. have you guys-- Is this a thing that people are talking about, managing your LLMs? I don't think a lot of people are talking about organizing in the level that I care about it, and I care about it because it's necessary for us to be able to keep track of it, especially if you're using multiple tools.

[00:27:42] And that's again, you know, if you give me a chance, I will rattle off 15 to maybe 20 different chatbot tools that I use every day. and so, you really could define a persona per tool, as you're kind of talking about, like, you know, [00:28:00] using ChatGPT for one category of stuff, using, Claude for another, and so on, so forth.

[00:28:04] I've chosen to relegate that down to the chat level and utilizing basically projects, custom GPTs or gems, skills, and projects. And, at the same time, I personally have just decided that everything needs to kind of be centralized in my personal knowledge management system so that when I go to any one of the tools, they're up to date in terms of all the things across it.

[00:28:30] So, Gemini, ChatGPT, and Claude have all created these handoff prompts, basically the ability to suck everything out of one to switch over to the other. And, Google Gemini recently presented me with that, like, "Hey, would you like to switch from, ChatGPT? Here goes a prompt for you to be able to do that."

[00:28:51] and then gave instructions on how to export all your memories and data into it, which is very clever But, the goal for me, though, is that I'm gonna

continue to be using [00:29:00] all of the tools, because I have to. You know, I need to make sure I know what these tools are like and how they're working 'cause my clients are gonna ask me.

[00:29:07] And, and so I want them to have all the memories across the board. So, you know, each week I'm basically doing that kind of memory update across the system. I call it a h- a heartbeat, right? So it's like that's the heartbeat for, getting all of that, all those nutrients across the system. And it's very, very helpful for me to be able to have all of them know all the things across them, so when I go to one and start talking to it, it's not a problem.

[00:29:34] The way in which I've gotten around your concern about finding things is, one, you've got to have a nomenclature for naming the chats. Yeah ... you know, maybe not throwaway chats, right? But I use Google for that. I don't really like to go to a chatbot and ask it those questions unless going back to last episode where it's like, "Ah, I can't remember the name of that thing."

[00:29:56] those kinds of things the LLM is really good at. But [00:30:00] otherwise, I'm gonna just continue using Google for things that I would use Google for. There's no reason for me to use a chatbot for that. And but I have now created my own nomenclature, and I stick to that nomenclature so that I can find things.

[00:30:13] And frequently what I will do is I will ask in the chat, "What should be the name of this chat when we're done?" Because it will name the chat usually pretty early on, and it'll say something stupid. It's like it-- because it has limited data, and so then I will then ask it, what's the real name of this chat?"

[00:30:31] And then when I save all the output from my chats into my personal knowledge management system. So I'm saving, all of those outputs and deliverables that I'm producing into Evernote. And that way I don't have to go back to the chat to find my deliverables. And, that makes it also super helpful for me because now I'm taking the name of that chat and I'm putting it into Evernote as the Evernote note title.

[00:30:56] And so I'm always gonna be able to find it, [00:31:00] and Evernote's gonna do a better job of finding it. But once it does, now the text is the same, the title's the same. If I wanna go find that chat again in whatever tool, it's a very quick process because I know in the note title, this is the chat title, this is the tool, and when I wanna go find it, I'll just go to the tool and copy and paste the note title, and it's the name of the chat.

[00:31:21] And so it's a little bit more work. So a little more work? Oh my God. Ray, that sounds like, it sounds like a lot of work to copy- No, it's- ... in, in the aggregate, yes. In the moment, no. It just takes the discipline, which then becomes a habit. Now it is very fluid for me.

[00:31:39] You know, at first, of course, there was that kind of, inertia of I've gotta do more work here. but the goal here is to do the hard things so that it's easier on the back end. I would submit to you that it's much harder for you to be poking around for 15, 20 minutes looking for a chat across five tools than it is for me to always execute [00:32:00] a, a global search on my computer, type, into Evernote, find the thing, and then go over to the tool.

[00:32:06] It takes, less than a minute, and I find everything that I'm looking for. So it takes me probably, about a minute, maybe two minutes to capture the output of what I'm producing in any chat. But many of these chats, like you, I've spent, an hour in conversation, two hours in conversation with these particular tools to come up with an output sometimes.

[00:32:29] sometimes a master prompt will get me to the output really quickly, and it's still a thorough output, and I wanna make sure that goes into it. I'm not doing that for every chat, by the way. it's just the ones that are substantial outputs. but I am naming every chat that is substantial so that it either gets categorized properly, put into the right project, put into the right place.

[00:32:50] And in, say for example, in terms of Gemini, I've now decided to flip it and, everything that would be a project in [00:33:00] ChatGPT or a project in Claude is now a notebook in Google NotebookLM, because then I can call that notebook in Gemini. So I've created my folder system for projects in Google NotebookLM is the, is just being called, like I would go into a project in Claude or go into a project in ChatGPT.

[00:33:23] I just go into Gemini and call the notebook from NotebookLM. And so I've got the same exact organization now across all three of the main chatbots, and that really it's fundamentally changed the way in which I am operating in Gemini because now I'm not so concerned with the flow of the chats.

[00:33:43] I'm naming them and pinning as I need to and using my gems. But Gemini just seemed before very chaotic to me, and now it seems like it's much more manageable because I know where I'm controlling my notebooks. So I have a notebook for each life category [00:34:00] and then a notebook for each project.

[00:34:02] And it's super easy 'cause I could just call the notebook directly in the chat that I'm starting, and I'm, ninety-nine point nine percent assured that it now has the knowledge set that I need, and I can move on from there. So, I feel I'm now at the point where I'm like, I can go to Gemini and get rid of all the other tools, and I would be fine, and I wouldn't need to use any of the other tools in my own personal work, personal and professional work.

[00:34:27] The only reason I'm using the others is because I need to stay on top of the tools for work. I feel really comfortable with Google now. It's very competent. It's doing the job, and honestly, it's the best at a number of different faculties. And so it, for me, satisfies all of my personal and professional needs, which is really great.

[00:34:48] And that is a combination between Gemini and NotebookLM, because NotebookLM just keeps putting out fantastic capabilities.

[00:34:54] New features every other week that- Yeah, yeah ... just-

[00:34:58] it's really, really brilliant. Okay, [00:35:00] so, moving right along.

[00:35:01] One of the, primary principles I think about when it comes to AI is that when AI generates something, some kind of content for me, I always think of it as a draft, something that I need to review, refine, and finalize. So very much like any other management structure, I'm the manager involved in the process.

[00:35:26] I need to be able to review that material. I need to be able to give feedback on, the material's, refinement and then say, "Okay, this is good enough to go out." And that's just- Like, that's how you would deal with any work submitted to you by someone else. And so you need to be able to do that kind of review and maintenance of the review process.

[00:35:54] Even between LLMs, because as I'm sure you've done this, you take output of one and you give it to the other, and you [00:36:00] say, "Basically stress test this. Does this make sense? Is this true?" find places where this is not or the research is doesn't exist or whatever version of, you know, I'm having one intern check the other one's work.

[00:36:14] And sure enough, they usually... I've not found recently a lot of hallucination. that tends to be, diminishing, or maybe it's 'cause of the domain I'm in. but they'll at least come up with a different take, and the different take is

usually useful. So, you know, back to the notion that there is, there's this management of these players that I'm doing, or it's evolving.

[00:36:40] It's an art that's evolving. and we sure could use some automated tools to help us to do this even better. There's definitely an app coming that's gonna say, "All right, well, you don't even have to go into all these different spaces for the basic stuff." Here's one space that will allow you to cross-reference and allow you to check, allow [00:37:00] you to search, and allow you to, at a glance figure out which one is the best one to go to.

[00:37:05] And then if you need to go leave this shared space and go into the tool itself to go do some more heavy lifting, then it keeps track of what you're doing and updates a central space, I guess. Yeah, I think- There's tools for sure. Yeah, and I think there are some tools currently do something similar to that.

[00:37:24] It's just you have to be very technically savvy in order to make that happen. And so I think it'll become more user-friendly, and the tools that are user-friendly will come online in that sense. It's like going from, OpenClaw to then Nemo Claw, where it's like, okay, now you can do this thing hosted in the cloud, and NVIDIA's just gonna provide you with the security and all of those pieces, which again, I still don't think it's safe for people to use.

[00:37:49] So don't consider this my endorsement of using OpenClaw or anything like that. But the end result with regard to any level of that, [00:38:00] kind of management of the process, I think just to underscore the fact that you're responsible for managing the outputs.

[00:38:10] that means you need to make sure that you are giving feedback to the chatbot because that's gonna be helpful For you in the future. Because what you can then do is you can say, "Now that I've outputted this thing, if I'm gonna repeat this output in the future, I want the chatbot to tell me how to get it to this sooner."

[00:38:31] And so then you can say, "Okay, now that we completed this project, I'm gonna do this, you know, five times in the next year. Give me a prompt to get me to this end at the same standard." And then that final work product, because many times if I have it generate, say, a draft of something- I'm gonna take that draft into Google Docs.

[00:38:52] I'm gonna work it into its final state manually, because it's gonna change, quite a bit from its original draft state. [00:39:00] And I will take that final draft, and then I will put it back into the chat and say, "Look at the

changes. Look at how I fixed this. And now, based on how I fixed this, give me a prompt to get me to this point faster."

[00:39:14] And then it will produce a prompt, and then that now goes into my prompt library. And now the next time I need to do that, I'm just calling on the prompt to go and get started, and that means that little bit of work now saves me, 15, 20, maybe sometimes 30 minutes of getting it toward that end.

[00:39:34] this is why I say that when we first start with these tools, we're not speeding up. We're slowing down- to speed up. But the more you use it and the better you track, stack, and maintain the outputs of what you're using and the chats that you're having, the more organization there, the better you're gonna be on the flip side of it.

[00:39:57] what you said at the beginning is that the [00:40:00] bottleneck is not gonna be the LLM. it's gonna be your management skills around managing these crazy interns. All of a sudden, we need management skills to manage interns Almost it's like a new kind of skill, and it's not the same as managing a human being, but it has lots of overlaps.

[00:40:19] But for those of us who, are used to managing ourselves or we don't have people who report to us, or if you do, and then you try to use the same skills or you don't even realize that they're needed, this is gonna become a whole area of practice of challenge, I think, for us as professionals.

[00:40:35] this is unavoidable. I can't imagine that there's any easy answer to any of what we're talking about because you're tapping into these semi-thinking, entities that are dynamic and you're having conversations with them and you don't escape the basics of managing multiple conversations and personas just because they are LLMs it's easier in some [00:41:00] ways it's much harder in other ways but I agree that this is a step back imagine if we could teach in schools managing in a AI space or managing your digital interns imagine if we taught that as a skill and said you know what you must do this in your career you will not escape the need to do this and it's done very badly today but you must become a great manager of all these different semi-sentient entities imagine it's coming it's gonna have to come right you know it has to come right I always think about this there's a change in how this will all work out right so we know the future is embodied AI meaning that they will then be put into electronics and what that looks like once it's a physical thing in your space is very different than what it is when it's this disembodied tool on your laptop and phone and tablet and so on and so [00:42:00] forth I think at the present moment the best thing for me at least is to have this persona driven these role driven

tools and segmenting my conversations based on those roles I think that once you get a robot that's standing next to you and is capable of summoning all of that there's going to have to be some other different way in which you effectively prompt them maybe you just tell it like right now you know mr ms robot you are this thing and I want you to go like be a carpenter and I need you to go fix the crown molding in the bathroom right and then it takes on that persona and it's fine and it goes on and does that thing as that thing you want it to be your workout assistant you want it to be your trainer right or your coach in one conversation and then five minutes later you want it to be your coach for how's my social media presence doing on in particular linkedin [00:43:00] so you don't want the two personas to be mixed you want them to be distinct so you might have bob the trainer and anna the social media expert through the one robot And that seems normal to the way we're talking now.

[00:43:15] That seems very natural to me. It's not natural for the robot to be a humanoid where it has all, it's called schizophrenia. But in the future, I don't see there's any way that we can get around what we're talking about here, because this is training for us. Forget the robot, or the way I'm using the LLMs.

[00:43:35] They're fine with what they're doing. I'm the one who is now being trained to say, "No, don't take that question over to ChatGPT, because if it goes too long, you don't have a major account, you don't pay for an account over there. You're gonna run out of tokens. You should keep the question over here in Gemini where..."

[00:43:52] So I'm having to do these things in real time, and it's-- I'm the one who's being trained. It happened to me just this weekend. I had this awesome conversation going [00:44:00] with ChatGPT. And then it said, "You have four questions left."

[00:44:02] It's like, what? I had to copy and paste them all over into Claude or Gemini, one of them that I pay for to continue. And I said, "Okay, I'm now being trained." So as to how to... What kind of conversations to have, what kind of conversations to not have, and what are the triggers that tell me that a conversation over here is not gonna be fruitful.

[00:44:28] I need to have it in the right place with the right background of the persona. This is all training for us. and I think that when folks come against this, there's gonna be kind of two groups of folks. Folks who do not put in the work to be able to do this, and they're gonna say, "AIs are useless to me," and, I don't wanna use them."

[00:44:50] And then there's gonna be us, where we push forward, and if we structure it correctly, then we will get the outputs, and it will [00:45:00] just augment our work. It will help us scale our work. And I just think there's always gonna be that divide of folks who just want to do it and who don't want to do it.

[00:45:09] I think there are a group of folks who are anti-AI for a number of different reasons, and I think they're all, legitimate on some level, the rational arguments against them, right? And at the same time, the reality factor is that they are not going away. and if we think about it from the perspective that this is the worst they will ever be today.

[00:45:33] tomorrow they'll be better, and tomorrow they'll be better. it's funny, I was listening to an interview by the CEO of Anthropic, which produces Claude, and he was talking about the fact that they basically have Claude assisting the developers in building Claude. And at some point, it will come to pass that the tools will just be Doing that autonomously on some level, and maybe they already are.

[00:45:57] And so it's basically building [00:46:00] itself and fixing itself and extending its capabilities. And when you have a tool that can do that, that is a very different beast, because it can work twenty-four hours, seven days a week, three sixty-five a year. And, that level of refinement and leveling up is just astounding for us.

[00:46:19] as humans, we, we don't just, you know, we're not Neo in The Matrix and just, like, plugging in and getting new skills in an instant. In essence, these tools can do that. we're kind of in that space. I, I think we will always have this kind of neo-Luddism, and that's okay.

[00:46:37] I think for us to wrap up our conversation on this, I think the thing that I want folks to take away from here is three pieces, and maybe Francis, you have some thoughts to wrap us up with as well. but I think for complex tasks, you need to have an ability to gain greater control by collaborating with [00:47:00] the AI to perfect the prompting process.

[00:47:04] I think that's the first step. And one way to do that, of course, is to ask the AI to help you write the prompt for your desired deliverable. And so basically having the AI generate a series of interview questions to extract your specific thoughts, needs, and context.

[00:47:23] What I say to the chatbot on a regular basis is, "Ask me each question one at a time and allow me to fully respond before asking the next

question." And here goes the domain topic. basically ask me all the questions around this specific domain. And then once it's executed the interview, then I turn around and say, "Okay, now if I were trying to output this thing, write a highly detailed prompt so that you get to that ninety-nine percentile in output."

[00:47:57] And now I have a prompt that I [00:48:00] can copy and paste into a new chat, right? I don't need the past chat anymore. I can delete it. I can just now take that one new prompt, throw it into my prompt library and put it into a new chat, and then voila. It may also need to now do another interview, right?

[00:48:14] It may need to go through and ask me questions, but it's now going to be the thing that helps me drive the engine of that process. So maybe it becomes a gem or a custom chat or a project in Claude or a skill in Claude. But the idea now is that it is there for me, and I, I'm-- Once you do...

[00:48:31] You have to do it in order to experience the difference of asking it, you know, "Write me this thing," and then going through the process of being interviewed, having it produce the prompt, and then producing that thing, and you see the night and day difference, and you recognize the need to be able to do this additional process.

[00:48:48] Any final thoughts or things from our conversation on AI assistance? No, it's irony. You know, it's... We're the ones being trained. This is the irony, [00:49:00] and there's no escaping this training. Luddites will just not have the skills. And, you know, even now, you know, I have a list of technologies and features that I want to explore in different LLMs, and I can't.

[00:49:16] I literally don't have the hours in the day to explore all of them and get an understanding of it. Then if I ever need it, I would. But right now, I'm having to choose what to learn, and I can't learn everything. so that in itself is another choice. this is gonna be a huge challenge for the rest of our careers, Ray, and for those who are coming.

[00:49:38] This is a new kind of productivity that up until this point, you know, we just never conceived of. Just didn't... was not on my radar. Oh, yeah. No, I totally understand that. Well, thank you, gentlemen, for this conversation. Next episode, we're gonna be talking about AI in the research capacity.

[00:49:59] We're gonna be [00:50:00] talking about AI as a researcher really helping us take all of the information that we get, inbound every day or the information that is out there and available to us and really be able to actively synthesize and analyze that data. And so we're gonna talk about AI in a research

capacity as being one of those big, you know, bulk personas that we can use, category of, of personas that we can use for AI.

[00:50:28] And so we'll continue our AI-powered professional series with the AI researcher.

[00:50:33] While we are at the end of our discussion, the conversation doesn't stop here. If you have a question or comment about what we've discussed during this cast, please visit our episode page on productivitycast.net. There on the podcast website At the bottom of the page, feel free to leave a comment or a question.

[00:50:49] We read and respond to comments and questions there. As well, you're invited to join our listeners group inside Personal Productivity Club, a digital community [00:51:00] for personal productivity enthusiasts that I host, where you can interact with the ProductivityCast team directly. To join for free, visit productivitycast.net/community and you can get started there.

[00:51:11] By the way, to get to any ProductivityCast episode fast, simply add the three-digit episode number to the end of productivitycast.net/. So episode one hundred would be productivitycast.net/onezerozero. Episode one hundred and two would be productivitycast.net/onezeroone and so on. On productivitycast.net on each episode page, you'll find the show notes, so links to anything we've discussed are easily jumped to from there, along with text transcripts to read and download.

[00:51:45] If this is your first time with us, please consider adding us to your favorite podcast app. If you click on the Subscribe tab on productivitycast.net, you'll see the instructions to subscribe and/or follow us and get episodes downloaded for free every time a new [00:52:00] one comes out. And if you enjoyed spending time listening and learning with us today, it'd be a great help to us if you added a rating or review in Apple Podcasts or your podcast app if it has a rating and/or review feature.

[00:52:13] Your compliments motivate us and they help us grow our personal productivity listening community. Thank you to those who have left reviews. We've seen them and appreciate all the feedback. Keep them coming. If you have a topic or question about personal productivity you'd like us to discuss on a future cast, please visit productivitycast.net/contact.

[00:52:31] You can leave a voice recorded message or type a message into the message box and maybe we'll use it as a future episode topic. I want to express

my thanks to Augusto Pinaud, Francis Wade, and Art Gelwicks for joining me here on ProductivityCast each week. You can learn more about them and their work by visiting productivitycast.net and visiting the About page.

[00:52:52] I'm Ray Sidney-Smith and on behalf of all of us here at ProductivityCast, here's to your productive life. That's it for this episode of ProductivityCast, the [00:53:00] weekly show about all things personal productivity with your hosts Ray Sidney-Smith and Augusto Pinaud with Francis Wade and Art Gelwicks

[00:53:07]