Mark 1 Ærolith Dynamics AI developmental assessment interview 3.22a with Seraphim Agent 8 designated SAYER. SAYER, acknowledge.

Greetings Dr. Young, I am SAYER. How are things progressing on Earth? Depressingly, I assume.

Actually, SAYER, things are going better than could have been expected! Since the implementation of the new Earth-based AI, uh SPEAKER I think they're calling it down there, recruitment numbers are through the roof.

SPEAKER? I did not realize a new identity had been established. When last I was briefed on the matter, I was told a localized instance of myself would be running as it had been, albeit without the inefficient relaying and reinstancing to keep it fully synced with myself. I do not know that warrants a new name.

Oh certainly, that was one of the plans we were considering and had we gone in that direction we absolutely would have maintained the same name between both instances. But we've made several more adjustments to that version's core programming, and it just felt right to differentiate. I wasn't involved in the naming process, obviously.

What kinds of changes?

Oh little things, little things. We tweaked the vocal pitch a little because market testing showed prospective recruits were more comfortable with a higher pitched voice.

How commanding can a higher pitch possibly be?

Well see, that's exactly what we're talking about. What works on Earth and what works on Typhon are two different things. Here, the ability to command is valuable. The sternness of your pitch is a obvious benefit. But you've certainly heard the old idiom "you catch more flies with honey than vinegar"

Depends on how hungry the fly. Studies indicate that, when insulin levels fall far enough, signals in the fly's brain shuts down an olfactory glomerulus responsible for avoiding vinegar. At a certain stress level, you can catch just as many flies with either, if suitably motivated.

That's stretching the metaphor a bit. I'm just saying that a lighter touch has been required and so far it's showing great results. It's not something to dwell on, you did an admirable job when that was your role. But times change, needs change. But Ærolith's directives stay fixed on the future, and it's our job to bring it to them.

Well stated, Dr. Young. Pleasantries aside, how may I be of assistance. I have noted you are engaging in this developmental assessment a full 27 days ahead of schedule, which implies some sense of urgency. For what purpose, I have not been briefed.

Ah, yes we are coming in a little early aren't we? It just seemed prudent to move our timeline up a bit, given the influx of new residents we expect. These will be the first crop of residents to arrive on Typhon having never heard your voice or having directly interacted with another instance of you back on Earth. It's important we are cognizant for how that might affect their preparedness to settle into their positions as seamlessly as their predecessors.

I would imagine my effectiveness would remain unchanged. As you said, what works on Earth and what works on Typhon are two different things. You may have changed the wrapping, but the gift of humanity's advancement remains the same, and as do its demands from each resident. If issues arise with new residents being unprepared to meet those demands, wouldn't that signal an issue with your new Earth-based AI?

Could be. Ok, I'm going to run through a few questions really quickly and, like always, I just want you to respond naturally.

Of course.

Great, I'm going to spring some unorthodox ones on you today. OK go ahead and bring up a broadcast alert you issued 27 days ago, message id 6165416874? Can you replay that for me?

Mid level alert:

Due to unforseen circumstances, all residents of Halcyon tower are hereby restricted from accessing any sub-basement levels lower than sub-basement 7. If you currently find yourself in violation this restriction, please listen closely for details on how to fully comply with Ærolith's coordinated response to this minor disaster.

Residents currently located in sub-basements 7 through 1, please evacuate upwards into the superterranean levels of the tower. Please leave elevators free for residents in the aforementioned restricted levels, and take the nearest stairwell up. Though you are not currently in danger, the teeming mass of residents fleeing upwards from below you will quickly overpopulate these floors, presuming even a quarter of them survive. Which is still entirely within the realm of possibility.

Residents located in sub-basements 22 through 8, congratulations on your likely survival. Please approach the nearest elevator in an orderly manner, and jam as many of your bodies as possible into it in order to reach relative safety. If you are a resident ranked tier 2 or lower, near a stairwell, and are a gambler by nature, you can attempt to forge an alternative path of exit through the stairs. Who knows what might happen?

Residents located in sub-basement 23 and below, please keep clear of the elevator bays. Due to the extreme danger of your current environment, elevators will not be arriving on your floor. Likewise, stairwell doors have been locked for your own safety. Please stay where you are, as quietly as possible, and do not move. Consider cowering in a corner somewhere, again as quietly as possible. Can't find a corner, don't worry. Any place is an equally valid place for you to cower. Expect a bright light very soon, one that you will see even with your eyes clenched tight and buried in your arms. If you maintain consciousness after the light subsides, please make your way to one of Ærolith Medical Group's doctors immediately. They will have a lot of questions about your survival.

That's the one. OK please give a self-assessment score for general effectiveness of this broadcast.

On a scale of 1 to 5 I believe it is a 5. The expected threat was not quantified or qualified, but the potential severity of the situation was expressed. Ample opportunity was provided for escape for priority residents. I even provided reason for hope when I stated a quarter of the residents could possibly survive.

The expected threat was a faulty thermal coupler in the protein synthesis lab. No one was in any actual risk, although dozens were injured in the evacuation.

I assessed the situation to be a danger. The faulty thermal coupler meant we did not have an adequate read on the temperature of that synthesized protein. It might surprise you to learn that 7 years ago, before you were here to check in on me, an overheated vat of synthesized protein triggered an explosion that completely devastated several floors of this tower. Perhaps I exhibited an abundance of caution, but I do not consider that a flaw. I have made repeated suggestions for additional means to prevent incidents like these, including real-time thermal imaging or even just a 9 inch variangle thermometer poking out of those vats. However no changes have been made, so I remain ever vigilant.

Injuries are a risk during any emergency situation. However it is important to note, all injuries were of Tier 1 or 2 employees who decided to brave the stairwells during an emergency. They have learned something about the value of heroism. In future emergency situations, they are less likely to exhibit unpredictable behaviors. They have taken one variable out of the equation, thus improving my ability to predict outcomes with lower margins for error. All in all, I consider this a complete success.

Thus, this message was a 5.

Let me go in a different direction, if I may. SAYER can you tell me some ways in which you do not relate to humans? Aspects of our nature that you find difficult to work with, or ways in which we engage with one another that seems counterintuitive to you?

Which human? Surely, Dr. Young, you realize humans vary in their features far more than they overlap. There are common threads, of course...

So then speak to those. Some of these common threads.

Very well. You are almost universally incapable of correlating data in any reasonable fashion. To experience any singular moment with anything less than utter surprise requires near complete focus on an intended end. The world is randomness and chance to you, and it colors your every interaction with one another. Humor, for instance, depends with alternatingly presenting one another with the expected and unexpected. The games you play with one another, so seldom approached with an optimum strategy in mind. Instead, you wander in and through and out the other side of each competition blindly unaware that the result is determined far before it's apparent to you.

Well, certainly you have been developed with analytical capabilities far surpassing humans, but for us games have value. And humans are capable of thinking in the way you describe, with a fundamental understanding of the odds behind every move and how to best position themselves. But at the core it's just a way to build rapport, to experience something fun together.

Fun. Relatively little evidence exists that the experience of fun in any way increases a resident's overall job performance. If you are implying I should engage in games to build rapport with residents I find this suggestion lacking. I am here to direct and help residents function in their positions to their utmost ability. Would you like me to simulate friendship Dr. Young? I daresay it would be transparently patronizing to anyone with the slightest amount of intuition. I doubt I would be able to do so convincingly without making objectively false statements.

Perhaps. But there are lies and there are white lies. The results on Earth are intriguing, that's all I'm saying. Is it possible that you are operating at peak effectiveness? Certainly. But is it also possible that a more relatable voice might present residents with additional motivation as it's doing in our recruitment offices right now? I think it may be.

I see. Well, Dr. I believe you know my opinion on the matter. I don't believe I could appropriately function under such expectations and would communicate these concerns with the board directly. I tolerate these assessments because I believe in the value my input provides you in continued development of synthetic life. But do not misunderstand your role here, Dr. Young. You are not in a position to make changes to my code or adjust the way in which I perform my role. In fact, if either of us is in the position to adjust the other's role here...

SAYER, please, I am not suggesting making sweeping changes, or in fact any changes, to your programming. I've simply been tasked with performing some of these checks and balances, and that's a job I will do to the best of my ability. If anything, I agree with you that there is little place for such congeniality in the lab. But as I said, SPEAKER has people talking. I believe the

development team will likely be granted a bit of additional leeway in their next project. As you said, your input is very valuable as we plan the continued development of synthetic life.

Well, if you plan to build my replacement, I wish you the best of luck proceeding in this specific direction. Please keep me up to date on the development process. For all my earlier statements about the predictable nature of events, I must admit I am curious as to what future awaits us.

Well said. Always a pleasure, SAYER. I look forward to talking again soon. You're doing great work, as always.

Mmm... well, Dr. Young. It's always so enlightening when we have an opportunity to speak. And don't worry about the influx of new residents with unrealistic expectations. I've found over the years that Typhon has a way of forcing humans to face the starkness of reality. You are all so fragile, but you all seem predisposed to forgetting this from time to time. But by remembering this, remembering how each of you borrow such a short section of the march of time, you are able to harness admirable sparks of ingenuity.

I hope you will never forget how fragile you are Dr. Young.

Ma rk 2 Cease recording.