Opal Electronics is running a challenge with the goal of finding talented engineers who are interested in solving hard problems.

The challenge

Create a convincing AI voice clone (of yourself?) that's capable of passing as you on a phone call—to both the receiver and a third party observer. Ideally, it'd be exceedingly difficult to tell when it is your AI clone talking and when *you* have taken over.

There will be three tracks:

1. Phone tree and IVR bypass: Agents must be capable of bypassing most typical implementations of IVR systems and achieving provided objectives. The idea is to focus on the experience that both the recipient and an observer of the phone call would hear. An example: When you simply provide an AI agent with DTMF tone playing capabilities, you will observe it preamble and postamble, ie: it will *respond* to the automated prompts since it is unable to differentiate between a real human and the automated system. How might you solve this problem?
2. Resistance to prompt injection and pressure: Language models learn to be 'helpful assistants' during post-training, which makes it challenging to coax them to roleplay as a real human (you) during phone calls. How would you approach this problem, giving them *durable* personalities that are resistant to prompt injection and prevent them from folding like a deck of cards the moment they receive pushback from a human?
3. Latency and response quality: Humans have a very interesting adaptation to deal with drifting latency and lag on a phone call. We automatically incorporate a *backoff* when we detect ourselves cutting someone off and understand how to pace ourselves when for instance, dictating phone numbers/spelling. This adds up to a more realistic phone call clone. How would you incorporate these mechanics into a VAD system or into audio response pipeline to make the clone seem more realistic.

Valid submissions are able to solve at *minimum* one of these problems and great solutions strike a good balance between these tracks.

We care not purely about task completion, but *sprezzatura,* and the grace and apparent ease with which your agents navigate their tasks. Do they know not to talk back to a robot IVR recording? Can the person tell when you take over control? Do they know how to *faff* and deflect when they face pushback?

Your constraints

- Custom voice: Must use a custom voice agent and not a default voice model. Building a phone call agent with a generic voice is *significantly* simpler, since they already have good emotional range, expressivity and better latency. We're interested in people who

want to tackle the hard challenge of making an agent that can pretend to be you, so that it can do all your boring calls that you don't want to handle yourself
- Not overfit: How flexible is your solution to a variety of scenarios? Can you use your voice clone to file a complaint? Can it convince someone at a popular restaurant to give you a table? Can it handle the sticky task of breaking up with your girlfriend? Can it navigate a tedious IVR menu without sounding like an idiot?

We want to have people take a crack at the hardest parts of the puzzle, so we'll provide client libraries and a boilerplate code that's already wired up to clone voices, place calls with your voice.
Your task will be to take this clone call to the next level, and build a replica of yourself so good that it could fool your mother.

Why phone calls?
Phone calls are an interesting medium. We think that there's many interesting applications of a convincing clone of yourself and are interested in exploring them:
- Making bookings and reservations: It's often easier to convince a human to sneak you in even when something is technically 'impossible'
- Convincing your elected official to listen to you: People find it harder to say no to a real human. Can we use this ability to campaign for things that we want?
- Tired of repeating the same instructions over and over: Screen your phone calls, give delivery drivers instructions that you've said a million times before without breaking a sweat when you get a call.
- And many other fun ideas.