

JOB INTERVIEW ON THE PHONE WITH AN AUTOMATED SPEECH RECOGNITION SYSTEM

Setting

Amy has a phone interview with you all, about our workshop today.

Collectively, you are an AUTOMATED SPEECH RECOGNITION SYSTEM, C.

How to read

The speech you are programmed to say will be handed to you on a piece of paper, you just read one question and pass the script on to the next person.

A: Hi, it's Amy, I'm here for my interview

C: Thank you for calling, you have reached the Urgent Publishing Workshop Application Call Line.

A moment of silence please.

Welcome, interviewee number 231. You are currently applying to organise a workshop at the Urgent Publishing conference. Please inform us of the title of your workshop. *[hand the phone over to the person next to you]*

A: Say it Ain't So

C: Oops! Didn't catch that. Please repeat. *[hand the phone over to the person next to you]*

A: Saaaay it Aaaaain't Sooooo

C: Thank you. Next question. What is the concept of your workshop? *[hand the phone over to the person next to you]*

A: We were drawn to the Session focus on Ursula K. Le Guin's *Carrier Bag Theory of Fiction*, where, Le Guin tells us that stories come in bags. In computer text processing, there is actually a technique called BAG OF WORDS! It's one of the simplest techniques: you make a list of all the words that appear in a text and then count how many times they appear in a sentence. What's funny though, is the contrast between Ursula Le Guin's take on the subject and Zellig Harris' take, who coined the linguistics expression in 1954.

C *[INTERRUPTS AMY]*: Maximum time capacity reached for this question. Next question. Explain further the concept of bag-of-words. *[hand the phone over to the person next to you]*

A: Where Le Guin talks about bags-of-words, she uses it as a metaphor to move away from one-track hero narratives to multiple track stories. Placing emphasis on collection, multiplicity and holding onto words we like. In contrast, the linguist and statistician Harris says, "*language is not merely a bag of words but a tool with particular properties which have been fashioned in the course of its use*". For Harris, the bag of words is a tool of observation, of looking inwards into existing structures.

We have partnered this theme with voice, our voices being one of the most essential ways we tell stories and spread knowledge. We recognise developments in voice technologies as violently forcing out non-normative voices, and therefore non-normative stories and knowledge.

C: What benefits can be drawn from participating in your workshop? *[hand the phone over to the person next to you]*

A: Erm ... well ... we hope to introduce you to a wide range of source material that we have been gathering in our bag, around the subject of voice technologies, voices in public space, and who's voices can be heard in these domains. We have prepared scripted material to experiment with our voices in space, and practice speaking out loud together. These exercises are devised for us to think about how our voices are captured by speech recognition technologies, specifically speech to text software, but also how they might evade capture. We have planned practical activity to experiment with our voices being captured or evading capture. We are using open source technology that we want to share with you too.

C: Next question. Please expand on your choice of word - scripts. *[hand the phone over to the person next to you]*

A: When I use the word I mean a written document, that I have composed from different sources, which I intend to be read out loud by a group of people. They are a way to keep our source material together in a bag. A way to have them speak together.

C: And what about the scripts that are used by speech-to-text recognition algorithms? *[hand the phone over to the person next to you]*

A: You could say the same thing about them, just that they are not meant to be read by people, they are meant to be read by computers. Though people did write them.

C: We have processed your earlier answers. Please, could you add a quote from Ursula K. Le Guin to support your argument *[hand the phone over to the person next to you]*

A: Ooh erm well She proposes alternative readings with her bags, she says "*if one avoids the linear, progressive, Time's - (killing) – arrow mode of the Techno-Heroic and redefines technology and science as primarily cultural carrier bag rather than weapon of domination, one pleasant side effect is that science fiction can be seen as a far less rigid, narrow field ... less a mythological genre than a realistic one*"

C: Next question. How does this relate to Urgent Publishing? *[hand the phone over to the person next to you]*

A: Ooh that's a hard one. Well. We feel it is urgent to raise awareness to digital discrimination. With voice recognition technologies, some parts of the population are not recognised. This is problematic because it denies them access to infrastructure, information, knowledge, tools. On the other hand, these minoritarian groups are also being attacked, with power structures collecting, voice profiling and tracking their voice in order to exclude them from services, incarcerate them and remove their identity from the public realm.

C: Could you summarize? *[hand the phone over to the person next to you]*

A: Mmmm ... Your voice is like a finger print but easier to get?!

C: And, to return to the Publishing aspect? *[hand the phone over to the person next to you]*

A: Thinking about publishing as presenting bags of words, or a bag of voices. We want to work through polyvocal approaches to sharing stories that retain their multiplicity. Much of publishing is focused around singular figures: the author is singular when they receive credit; how can we present the references an author builds on, the conversations with peers they've had, the editors that have helped crystallize the writing that often fade into the background.

C *[INTERRUPTS AMY]:* Maximum time capacity reached for this question. Next question. *[hand the phone over to the person next to you]*

A: Oh but I haven't finished. I just want to say one more thing. When we talk about machine learning algorithms, datasets are another form of publishing that is more concealed ... one that bears serious consequences.

C: Can you give us a concise outline of the workshop activity?

A: Yes! So the first half of the workshop, before lunch, will involve exercises in speaking and listening - like now - and we will develop ways you can be heard and not heard - captured or evading capture. First we will get to know Pocket Sphinx, the open source speech recognition engine that we are working with.
Then we will get to know some of our source material by experimenting with some scripted dialogue. Before lunch we will choose elements of dialogue that we want to be heard, and that which we don't. We will then return our speaking voices to pocket sphinx, who will document and write down your voice. Pocket sphinx will also make your bag of words.

C *[INTERRUPTS AMY]:* Maximum time capacity reached for this question. Next question. *[hand the phone over to the person next to you]*

A: Oh and I will just explain the second half before we move on!
We'll come back from lunch to look at some spoken and performed source material.
Then we will have a read of our bags of words, and discuss how we want to take them back out of writing and into action.
We'll introduce exercises in score making and performing, and we will record responses to make a collectively published outcome.

C *[INTERRUPTS AMY]:* Maximum time capacity reached for interview. Next question. Do you have any final remarks? *[hand the phone over to the person next to you]*

A: Erm ... we hope that we can record your voices on my zoom recorder, to publish our work together. If anyone would not like to be recorded then please say, we will only be recording moments that you choose to present to the zoom and an audience.

C *[INTERRUPTS AMY]*: Thank you interviewee number 231. We have completed your interview. You may now hang up and await our response.
Goodbye.

A: bye?!