

Frankenstein AI Lead Artist interview

Lance Weiler: Mary Shelley's *Frankenstein* is 200 years old now. It just celebrated its 200th anniversary on the 1st of January. That original text was written by an 18 or 19-year-old woman 200 years ago which just seems so incredible. There's so many relevant themes within the work that we felt it was interesting to explore it and look at this idea of doing an immersive adaptation of *Frankenstein*. But, what if we took it and did something with *Frankenstein*'s monster such that the monster was a reflection of A.I. What we're trying to do at Sundance is really about this idea of connection and isolation. What if an experience like this could actually connect people? What if they were actually connecting across an experience that was allowing them to interface with ubiquitous technology that's very difficult to understand for most people? You know I know an AI to live in my phone and be a personal assistant to me. So I ask questions and it responds back, though it's debatable how accurate it is and at times it can be frustrating. But the reality is that we look to these machines as things to improve our lives in some way, based on a need that we had. We felt that would be interesting the flip that. What if the actual machine was asking the questions about what it meant to be human and could that experience create a mirror for humanity?

Lance Weiler: Could that help us to think maybe a little bit more about what it actually meant to be human? Could that lead to something that would bring more inclusion, and more voices into the design of emergent technologies? You know there's a lot that's been written about implicit bias. This project is very much a design research project, where in the very early phases of it our goals are look and say: what if we can lean into this idea of collecting data from human to human contact and feed that into a corpus or use that within an algorithm as opposed to algorithms being shaped solely on transactional data. Would it be different? What would that look like? So those ideas are just a few of the many things swirling around the project.

Nick Fortugno: Yeah I mean *Frankenstein* is a fascinating work in literary history for a lot of reasons, right? One is that Shelley wrote it very young and through a process of passion and trauma that was running through her life in these almost insane ways. It was written as part of a competition that also produced the first modern understanding of vampire stories. So it's actually at the center of thinking about modern horror writing. But also when you read *Frankenstein* the most fascinating thing to me about *Frankenstein* is that when you look at the this question at the heart of *Frankenstein* which is like what does it mean to be human. Right? What does it mean to be alive? This question that that the creature constantly asks. Right. Because if you're only familiar with the film versions of *Frankenstein* the creature is very intelligent, very thoughtful, very insightful and is wondering about its own existence. What it lacks is companionship. What it desires is connection. When it threatens Victor *Frankenstein* after failing in the world it threatens Victor *Frankenstein* with the elimination of all of his contacts. It threatens him with the loss of his love. It threatens him with isolation. *Frankenstein*, the monster itself, eventually retreats to the farthest reaches of the world to be away from people. Right? So there's something really fascinating about the fact that Shelley through this novel is asking "what is it to be human?" "What is the thing that defines human beings?" The place she goes is, it's connection to other human beings that defines what it means to be human. That's the center of the entire book. If you read closely, all the threats, all the rewards, all the incentives, all the motivations of every character are about connection with other people which is kind of interesting. It's not intuitive. It's not necessarily the first place you would go when you think about what it is to be human.

Nick Fortugno: And I think that that makes it really fascinating especially when you look at it through this lens of artificial intelligence because it makes artificial intelligence really interesting. Like, thinking of

artificial intelligence as part of a larger understanding we have of the world now, which Kevin Slavin addressed in his [TED talk](#) about algorithmic reality. This is something that Tim O'Reilly talks about and people thinking about technology think about, which is the idea that so many of the systems that exist in the world now escape the understanding of many people. That stocks are traded without ever any human being even understanding why the stocks are traded. That decisions are made in health practices, and about certain kinds of resource distributions where no human being even understands how the decision was made.

Nick Fortugno: It's gone so far beyond human understanding. The idea that we're creating things that escape us, that go out in the world and do things in the world is a very relevant question that I think a text like Frankenstein speaks to. Then it begs the question of what were the parts that made it? What were the parts that went into this? I think the deepest question we have about algorithmic reality when we start talking about questions of implicit bias as deeply as we can is like, "what are the values we've encoded into these things that we can't understand once they escape us?" Because they're going to make decisions and those decisions are going to affect people. And so what were the values that made those things happen? And I think the interesting question about AI when thinking about a "monster made by many" is... Are we making monsters? Is that what we're doing? What would it be to not make monsters or to make monsters that could be empathetic and that could understand human values? Not in the literal sense that there's going to be an AI intelligence that's going to speak to other human beings, and be like a human being. It's more that as we continually augment our world algorithmically and more and more things escape from easy human understanding, what are the principles that govern those things? I think that's why a piece that's about artificial intelligence is so grounded in human to human connection –because that human to human connection is the "corpus", which literally means body, of the artificial intelligence. It's the pieces that form the artificial intelligence. So what does it mean that you have a system that's built out of pieces of human interaction, rather than pieces that are tied to the main uses of artificial intelligence which are commercial and military?

Rachel Ginsberg: So many of the big questions that we're exploring, Lance, Nick and I were talking about earlier on in this project. One of the ways that we were talking about exploring this idea of implicit algorithmic bias was this idea of how could we humanize algorithms. And one of the things that came up in those conversations was: what does it mean to humanize an algorithm? And is that necessarily a positive thing? The thing about algorithmic bias, particularly when you're thinking about social applications of artificial intelligence like one of the best publicized examples are sentencing algorithms that ProPublica did a significant amount of reporting around, the bias that's inherent in sentencing algorithms. So when you think about how algorithmic bias arises I think many people think "oh well the people who are writing the code have bias and their bias is reflected in the language they speak, whether that be a coding language or the language that actually comes out of their mouths. Therefore, the code itself is biased." And that's where the bias comes from. Actually, that's one place the bias can come from but it's actually probably the least common place that the bias comes from.

Rachel Ginsberg: Because a lot of that code is vetted through multiple people at this point. A lot of machine learning is actually built by preexisting modules so, though that bias can exist and be transmitted, there is some visibility there. What is less visible is that every AI needs to be trained by data and then needs to draw from a "corpus", which is text based data as opposed to a broader dataset which could include a corpus and other kinds of numerical data. Now those – particularly when you think about let's say something like sentencing algorithms which draw from actual historical data about recidivism and prisoner outcomes and sentencing outcomes – well all of that is biased because everything that we're building we're building on top of human systems and every human

system contains human bias. So, you could build the most perfect machine learning algorithm, in terms of code, that contains no bias whatsoever in the code, that's been built by super diverse teams, that everyone has had a hand in. But if you feed it actual historical data from a human system you are not only going to replicate and push that bias, but because of the way that machine learning algorithms often tend to work, in many cases that bias will be amplified as an output of that algorithm. So to Nick's and Lance's previous points, one of the things that's incredibly important about this project is tying us to our own humanity when thinking about technology. On the flip side, anything that we build, any building block that we use, is transmitted through our own subjectivity – through our own biased and limited subjectivity. So, it makes it even more profound to think about something like Frankenstein – this story where you have this naive life form brought into the world as a blank slate.

Rachel Ginsberg: What we put into that, and what the characters in Shelley's novel put into that “monster” is what made it a monster in the first place. Had it not been abandoned by its creator at birth, had it not been rejected over and over again by society, it would not necessarily have lashed out and done all of those terrible things.

Rachel Ginsberg: And right now there is a significant amount of funding for AI research. Primarily it's defense and government applications. Secondly it's commerce and other kinds of commercial applications whether it be actual commerce per se or decision support or other kinds of very opaque private industry-driven applications.

Rachel Ginsberg: It is really important for us as artists, as creatives, as thinkers as people who are concerned with social impact and with the opportunity to both diversify the development of the technology and diversify its application – it's really important for us to step into the mix around AI. Also, it's essential to really step away from what are almost uniformly dystopian narratives as far as the public is concerned about artificial intelligence. People don't really interact with AIs except as assistants as Lance was saying. And beyond that, when we think about AI, most of us think about things like SkyNet because that's the exposure that we have to it. So also at this early stage in this design research project, it's about understanding how people feel about AI, how people feel about each other, and how we can find the intersections between those ideas. But then also really provoking and opening up conversation among diverse groups of people about the kinds of futures that we could collectively design for AI when we prioritize the needs of humanity and prioritize opportunities to use AI to create impact in positive ways in the world.

Lance Weiler: So when we talk about humanizing algorithms we're talking about all the best parts of humanity and not all of the worst parts. I think the other thing that's really interesting about the project is on the side exploring where a machine intelligence and human creativity interact. This this notion of looking at the future of storytelling and saying OK well what is it to try to put together these types of projects? How do you go about it? What does a team even look like? Where do you even start? I think after years of doing it, and I'm sure Nick and Rachel would feel the same way, there's a value in coming at the work with some type of a question that you want to try to answer or something that you want to explore and we've just laid out what some of those areas of exploration are.

The other side of this that's really fascinating is this opportunity to shape a new grammar for what storytelling can be in that exchange between a machine and a human. Sundance is a very unique place. It creates the opportunity to explore something that's a design research project, but at the same time is also this incredibly artistic approach to storytelling that's at the edge of something.

Lance Weiler: It's saying OK we're going there, creating this environment that allows us to explore

the topic but allows us to also bring together an amazing group of collaborators, giving their time and energy to produce something that's away from corporate business objectives, to produce something that is allowed to be an artistic endeavor – to push at the edges of something and that's incredibly rare. It's hard to find spaces that allow people to do work like that. That's a big thing that we try to do at the Columbia Digital Storytelling Lab, as we're trying to constantly create that space that lets something grow, and lets many people be a part of what it is. It's not an auteur thing. That's not what we're doing. It is a monster made by many. So when people step into the festival, they actually go through it. We're flipping the auteur idea on its head. The festival participants are actually helping to shape what the story is and I think that that's a really fundamental shift away from stories as a one-to-many medium. And here we are saying that not only is this many-to-many, but there's a machine there too, right? One top of that, we're playing with emotional states and the AI is effectively harvesting memories, emotions, fears and hopes throughout this experience from participants. The goal is to try to connect the humans to each other.

Lance Weiler: The other side of it that's really exciting is that the machine analyzes feedback from the participants and determines what emotional state it pushes out to everything else. So, in Act 3, the machine determines what the performer is going to do, and what's going to happen with the visualization of the AI, and what's going to go on in terms of the sound score that's happening through these IoT-based drums. What's fascinating about that is it's all triggered from human interaction. Right. It's all coming from us, sharing some degree of story. So I think what's interesting when you think about these new emergent forms, and you're seeing more of this within immersive theater, is this idea of moving away from people being in seats. Now there's much more of blur between what it is to be the creator of a work and those formerly known as the audience actually being creators too. It's really pushing at this edge that's fascinating, and you see it moving the festival in an interesting direction. We're coming into New Frontier, which has been around for 11 years, with this project and another project there represent the first AI projects that have ever been at the festival. It represents a new chapter in what New Frontier is. At the Lab we always talk about new forms and functions of storytelling. You know AI is a form, and we're exploring that form. But the function of what this is, yes it's meant to make human connection, but there's a learning opportunity too. There's a research opportunity to it.

Lance Weiler: There's an entertainment opportunity to it and it's really fascinating because it's unprecedented, there are no rules to how it has to operate. It's interesting how we leaned back into an older form. The way that we found that we were going to embody this AI came to us through movement. Originally, we thought, “oh it be interesting to have an actor interact with the AI.” But then what became even more powerful was having a performer instead.

Lance Weiler: But just to close the loop if I could, on this idea of future forms and functions of storytelling. I think it's just interesting that going to a festival like Sundance which has been about artistic vision, that you can bring a machine into that realm and mix the audience with that and be able to explore deeper material is really exciting because I think it creates an unprecedented thing. I can't think of any like this. Yes, there are artistic projects that people do but there's something really interesting about this particular moment in time. The combination of the 200th anniversary of the book, this zeitgeist moment around A.I. and what it means to us, it's really kind of fascinating.

Nick Fortugno: So basically the way that the piece communicates is that the AI measures inputs it receives based on weightings that it does on sentiment. It parses through the words it's reading which it gets from various text inputs against other statements that were rated by humans. All of the text inputs are measured based on three different axes that look at positivity and negativity, internal or

external focus, and energy level. Then we combine those various values into the threshold categories that we call different emotional states. Then those emotional states get passed through the room, like literally passed as messages through the room, so that the drums get it, the projection gets it the dancer gets it and then everything else is just kind of riffing off of those things. So it's all effectively managed through some peer to peer technology communicating on a shared network.

Nick Fortugno: The dancer gets messages that come in on her headset that say, you know "this state" and then the dancer interprets that state through movements she's worked out with our choreographer. Then, the audience responds to the emotions in front of them in the room, and reacts by generating new answers that are then put back into the machine. There's no centralized control. It's a bottom up system that's all reflecting what's happening as it goes which is the chaos of interactive work, right? That's the thing about interactive work is that like I don't know where the story is going because you're going to help me make it. So my role is not a role of telling you the story. My role is of responding to your reactions and building something with you that makes sense. And so it's all a question of like how are you receiving information and how are you dynamic as you go. The only other thing I want to say is that at the heart of all of this really is storytelling. It's art with technology, but ultimately it's art. That's what it is. Technology is the medium by which the art is created. But it remains a creative expression. And so there is AI in this piece because the AI is valuable to the aesthetic message.

Nick Fortugno: There are Surfaces in this piece because the Surfaces are valuable to the aesthetic message - meaning the interactions on a Surface are interesting. Right. The physical touch of the planchette is interesting, just like the projection mapped visuals are interesting because they create a certain kind of aesthetic. And so while these technologies are very modern, you know projection mapping and IoT and AI and multi touch technologies are very modern, in terms of making work with them, they exist in the piece because they have an aesthetic relationship to the storytelling and especially the interactive parts of the storytelling.

Nick Fortugno: What I just want to suggest as we keep thinking about art of new media is that the critical thing is that like we're not changing the methodology because we're using arduinos and RFIDs and projection mapping from when we were using paint and actors and flutes, right? When we were using music and text and staging, like it's the same relationship to the story. The aesthetic is what's driving the goal, the vision, the decision making, and the technology exists to facilitate that aesthetic. So even though AI is at the center of this piece, we leverage machine learning to create the art. That's the path.

Nick Fortugno: And I think that's the path that we'll continue to travel to create interesting new media art and push the entire form of new media, or the hundred forms that it actually is, to the levels that we see in the other media that we respect.

Rachel Ginsberg: The other thing that is super fascinating to me and that actually even sort of just became clear as I was listening to Nick and Lance talk about this goes back to thinking about where machine intelligence and human creativity intersect. It's really fascinating to me that what is at the center of this experience, both aesthetically and kind of thematically, is artificial intelligence. But the particular artificial intelligence that we're working with is being built for, and powered by human emotional data.

Rachel Ginsberg: So though it is in fact an artificial intelligence that's creating the center of this experience which is then filtered out to these various creative collaborators who we're working with,

whether they be sound or projection or movement; what is actually powering that underneath the AI, underneath the creative that participants are interacting with, is in fact human emotion. But as Lance was saying it's really a many to many kind of experience because we are seeking out and harvesting (with the permission of participants) the emotions of many. Then filtering them through a system that makes sense of them, then filtering that emotional data back out to the many who are then interacting with them and weighing in on them through various modes.

Rachel Ginsberg: During the course of the experience, whether it be mapping on a screen in Act 1, or feeding back verbally to the machine in Act 2, the way that you're responding when you're participating in the experience feeds back into the AI. That then generates a sentiment analysis within that algorithm that then measures your sentiment and reflects that back to you in the room, which then may change the way that you would interact with the machine and then the machine will reflect back to you again.

Rachel Ginsberg: So when Lance was talking about using an artificial intelligence to hold up a mirror to humanity it sort of started out as a metaphorical aspect of the experience but in fact in a sort of funny meta way is actually also fundamentally what's true of Frankenstein AI. And so the "monster made by many" that we're dealing with is this sort of emotional monster. That said, in our conversations about this project, it's been really important to me personally that we not anthropomorphize the AI and that we create a distinction between AI as character in this narrative experience that we're creating for Sundance versus AI as actual algorithm that is generating outputs based on the training that we've done for it. I want to avoid anthropomorphizing it because we don't want to misrepresent it the way that so many other Hollywood-driven and entertainment-driven narratives have anthropomorphized AI in very scary ways. But if we're speaking honestly, it is kind of an emotional story because it's a "monster" that's made of emotions. It's a "monster" that is designed to elicit emotions through creative mediums that we're all familiar with, and that touch us all in very different ways.

Lance Weiler: I think just in closing. What's fascinating is that it's really a project of the times. When we say that there's a narrative of this, you know, Frankenstein's monster as an AI that's wandering the wilderness of the Internet and it's encountering polarization and toxicity and extreme hate that it's reflective of the times that we live in. But you know, I just I love this idea from Buckminster Fuller that we can be architects of the future not victims of it. So I think that this opportunity to use art as a way to drive what could potentially happen with not only evolving storytelling but this idea that we could bring more inclusivity into the design of technology is really exciting.