

Esmeralda Conde Ruiz – Empfindsame Ohren und unempfindliche Infrastrukturen. Teil II

Ein Gespräch mit [Michael Klipphahn-Karge](#) | Bereich: [Interviews](#)

Das Interview wurde im [Original in englischer Sprache](#) geführt und bereits in w/k veröffentlicht.

*Übersicht: Von April bis Oktober 2022 war die international bekannte Komponistin, Dirigentin und Künstlerin Esmeralda Conde Ruiz (*1980/Spanien, lebt in Großbritannien) zu Gast im Schaufler Lab@TU Dresden. Während ihres Aufenthalts arbeitete sie mit Wissenschaftler*innen der TU Dresden (TUD) sowie mit Expert*innen anderer Institutionen zusammen und erforschte Ideen im Zusammenhang mit dem Leitthema des Labs Künstliche Intelligenz als Faktor und Folge gesellschaftlichen und kulturellen Wandels. Unser Redakteur Michael Klipphahn-Karge, derzeit selbst Stipendiat des Schaufler Kolleg @TU Dresden und Kunstwissenschaftler, hat Conde Ruiz zu ihrer Residenz in Dresden befragt.*

Kann KI beim Komponieren nicht nur – wie oft üblich – als Werkzeug, sondern vielmehr als Instrument eingesetzt werden? Wie kann der Mensch in diesem Fall auf die KI Einfluss nehmen, und wie wirkt der Einsatz von KI- Systemen zurück?

Was ich an Ihrer Frage interessant finde, ist der aufgefächerte Unterschied zwischen einem Werkzeug und einem Instrument. In meiner künstlerischen Praxis spiele ich z.B. Klavier, aber ich betrachte mich nicht als Pianistin. Daher ist das Klavier für mich eher ein Werkzeug. Denn ich verstehe etwa dessen Stimme, den Ton, wenn Sie so wollen, tatsächlich als ein beziehungsweise als mein Instrument. Ich habe Jahre damit verbracht, diese Fähigkeit zu perfektionieren, Stimmen und Stimmlagen zu verstehen, zu erforschen, zu analysieren. Computer und andere digitale Möglichkeiten nutze ich ebenfalls eher als Arbeitsgeräte – etwa, um Effekte zu erzeugen. Das Wichtigste dabei ist, dass die Klangerzeugung in meiner Hand liegt, nicht von der Technologie gesteuert wird, sie ist und bleibt für mich ein Werkzeug, um meinen Möglichkeitsraum zu erweitern.

Wenn KI ein Instrument wäre, würde ich mir gern die Zeit nehmen, um zu lernen, wie man es spielt, so wie man das Spielen jedes anderen Musikinstruments erlernen kann. Auch um es für mich wieder auf den Status eines Arbeitsgeräts reduzieren zu können. Dafür braucht man Knowhow, Anleitung und mehr Zeit – vor allem, um die erlernten Regeln zu brechen, die für die Beherrschung eines Instruments erst einmal nötig sind. Juniorprofessor Matthew McGinity (Immersive Medien, Institut für Software- und Multimediatechnik, TUD) sagte während unserer Zusammenarbeit im Rahmen der Residency etwas, das mich sehr berührt hat:

„Es kann derzeit nichts Sinnvolles mit Tiefgang geschaffen werden, da sich die Branche so schnell entwickelt, dass man mit den Aktualisierungen nicht Schritt halten kann. Es ist, als würde man jeden Tag ein neues Instrument spielen und man muss alles wieder neu lernen, anstatt es in seiner Zugänglichkeit als Werkzeug zu beherrschen und zu erforschen.“

Wenn das Instrument ständig aktualisiert wird, kann ich nicht wirklich etwas Sinnvolles schaffen, außer durch seine Nutzung aufzuzeigen, wie es funktioniert. Das ist das große Dilemma, was ich derzeit in den meisten KI-basierten Kunstwerken sehe. Natürlich könnte jemand, der aus dem Bereich algorithmischer

Musikproduktion und Live-Coding kommt, ganz anders auf diese Frage reagieren, und ich bin mir sicher, dass jemand anderes KI als Instrument dem Technologischen gegenüber disruptiv einsetzen könnte – doch für mich hat das, auch in Anbetracht der Zeit, einfach nicht funktioniert.

Ist es möglich, der KI eine eigene künstlerische Stimme zu geben und so einen individuellen Klang zu finden oder zu entwickeln? Welche Rolle spielt dabei der Mensch – einerseits als derjenige, der die KI bedient und programmiert hat, und andererseits als derjenige, dessen Stimme durch die KI ersetzt werden kann?

Es ist schwer zu sagen, wie eine sozusagen *eigene* künstlerische Stimme einer KI klingen könnte, die unabhängig von menschlicher Vorstellung und damit rein aus der Technologie erwachsen ist. Ich kenne nur Stimmen von Menschen, die menschliche Stimmen nachahmen oder synthetische Stimmen, die auf dem menschlichen Verständnis davon beruhen, wie eine *natürliche* Stimme zu klingen hat. Zählt man das Kopieren einer menschlichen Stimme dazu, KI eine eigene künstlerische Stimme zu geben, dann wäre *Holly+* etwa ein Schritt in diese Richtung. Diese Methode, synthetische Stimmen als Instrument zu verwenden, wurde von Holly Herndon entwickelt, einer in Berlin lebenden Klangkünstlerin und Sounddesignerin. Jeder kann mehrstimmige Audiodateien auf eine entsprechende Website hochladen, um eine individuelle und automatisiert erstellte Aufnahme der Tonspuren zu erhalten. Diese Datei enthält dann wiederum die eigene Audiodatei – nur mit Herndons Stimme gesungen oder mit ihrer Stimme neu gesprochen.

Die Rolle, die der Mensch dabei spielt, ist entscheidend: Da die synthetischen Stimmen immer menschlicher klingen – man denke nur an Sprachassistenzsysteme –, wird derzeit auch darüber diskutiert, ob die Stimme, die man hört, als *nicht-menschlich* oder künstlich bezeichnet werden sollte, damit überhaupt klar wird, dass dort kein *echter* Mensch spricht.

Andererseits werden etwa bei Kund*innengesprächen am Telefon Aufnahmen mitgeschnitten und in ein KI-Modell eingespeist, um die KI so an menschlichen Stimmen zu trainieren und dafür zu sorgen, dass ihre Stimme menschlicher klingt. Ergo: Ohne den Menschen gibt es folglich keine synthetische, keine künstliche Stimme.

Eine weitere faszinierende Entwicklung in diesem Zusammenhang ist derzeit der Klang von KI-Übersetzungssoftware. Die Modelle sind unglaublich, zeigen aber auch die Unterschiede in der Stimmqualität und -ausbildung, etwa je nach Region. In Googles KI-Übersetzer etwa haben bestimmte Länder wie Armenien gar keine Stimme, während Bosniens Stimme sehr roboterhaft und Kataloniens unglaublich menschlich und exakt klingt. Bestimmen also beispielsweise der Wohlstand und die politische Lage eines Landes die Qualität und Genauigkeit einer künstlichen Stimme? Das wäre eine Untersuchung wert!

Welche Rolle spielen in diesem Zusammenhang Autorschaft, künstlerische Autonomie und Originalität?

Urheberschaft, künstlerische Autonomie und Originalität sind in der Kunst seit jeher umstritten. Es gibt endlose Gerichtsverfahren zum Einsatz von KI, die die Branche bereits gespalten haben. Ein Beispiel: *Stable Diffusion 1*, ein auf maschinellem Lernen basierendes Text-zu-Bild-Modell von *Stability AI*, generierte Bilder im Stil des Künstlers Greg Rutkowski, zum Entsetzen des Künstlers. Das Modell wurde mit seinen Bildern trainiert, ohne dass er davon wusste. Der Schaden, der dem Künstler und anderen dadurch entstanden ist, ist immens und kaum wieder gutzumachen, denn dieses System wurde sehr genau darauf trainiert, ein Werk im Stil Rutkowskis als urheberrechtsfreies Bild zu produzieren. Ein kürzlich vorgenommenes Update von *Stable Diffusion 2* erschwert das Kopieren von Künstler*innen, was

nun wiederum die Nutzer*innen verärgert. Die Motivation für die Implementierung des Updates war vermutlich weniger der Schutz der Künstler*innen, als vielmehr die Vermeidung rechtlicher Konsequenzen.

Das Problem wird noch komplizierter, da viele dieser Modelle auf Abonnements basieren. *Stable Diffusion* ist für die Open-Source-Gemeinschaft bestimmt und kann durch eigene Anpassungen und Codes zu individuellen Versionen weiterentwickelt werden. Das bedeutet auch, dass es schwieriger zu kontrollieren und zu verstehen ist, was mit dem Trainingsdatensatz passiert, was damit erstellt und was hinzugefügt wurde und wird. Denn anders als Menschen können diese Modelle weder vergessen noch verlernen.

Ein weiterer interessanter Fall ist der von Jason M. Allen, der dieses Jahr den ersten Platz im US-Kunstwettbewerb *Colorado State Fair* für sein Werk *Théâtre D'opéra Spatial* gewonnen hat. Sein Preisträgerbild in der Kategorie *Digitale Kunst* löste eine Debatte darüber aus, was es eigentlich bedeutet, Künstler*in zu sein. Denn die Jury verstand nicht, dass und wie das Bild mit *Midjourney* erstellt wurde – einem Programm für KI, das Bilder aus Textbeschreibungen erstellt.

Was das für tonbasierte Arbeiten bedeutet, ist noch unklar. Sound ist aufgrund der Länge der Stücke komplexer, und es ist viel schwieriger, Töne zu komprimieren als etwa die Pixel eines Bildes. Die derzeitigen KI-basierten Text-Ton-Modelle sind auch nicht annähernd so beeindruckend wie ihre visuellen Äquivalente – aber man kann erahnen, dass sie zeitnah wesentlich leistungsfähiger werden. Was kommt als nächstes?

Wie beurteilt man also ein Kunstwerk, das mit entsprechender KI-Software erstellt wurde? Verlangt man den Code, eine originäre Partitur oder die Trainingsdaten zu sehen, um Urheberrechtsfragen zu klären? Immerhin ist das Urheberrecht eines der wichtigsten Gesetze zum Schutz der Schöpfungen von Künstler*innen.

Welche Rolle spielt der Gedanke der Vernetzung von Mensch zu Mensch, das Agieren in Kollektiven und sozusagen der tatsächliche Kontakt – etwa in Chören –, wenn man diese Konzepte des gemeinsamen Musizierens an den digitalen Möglichkeiten misst?

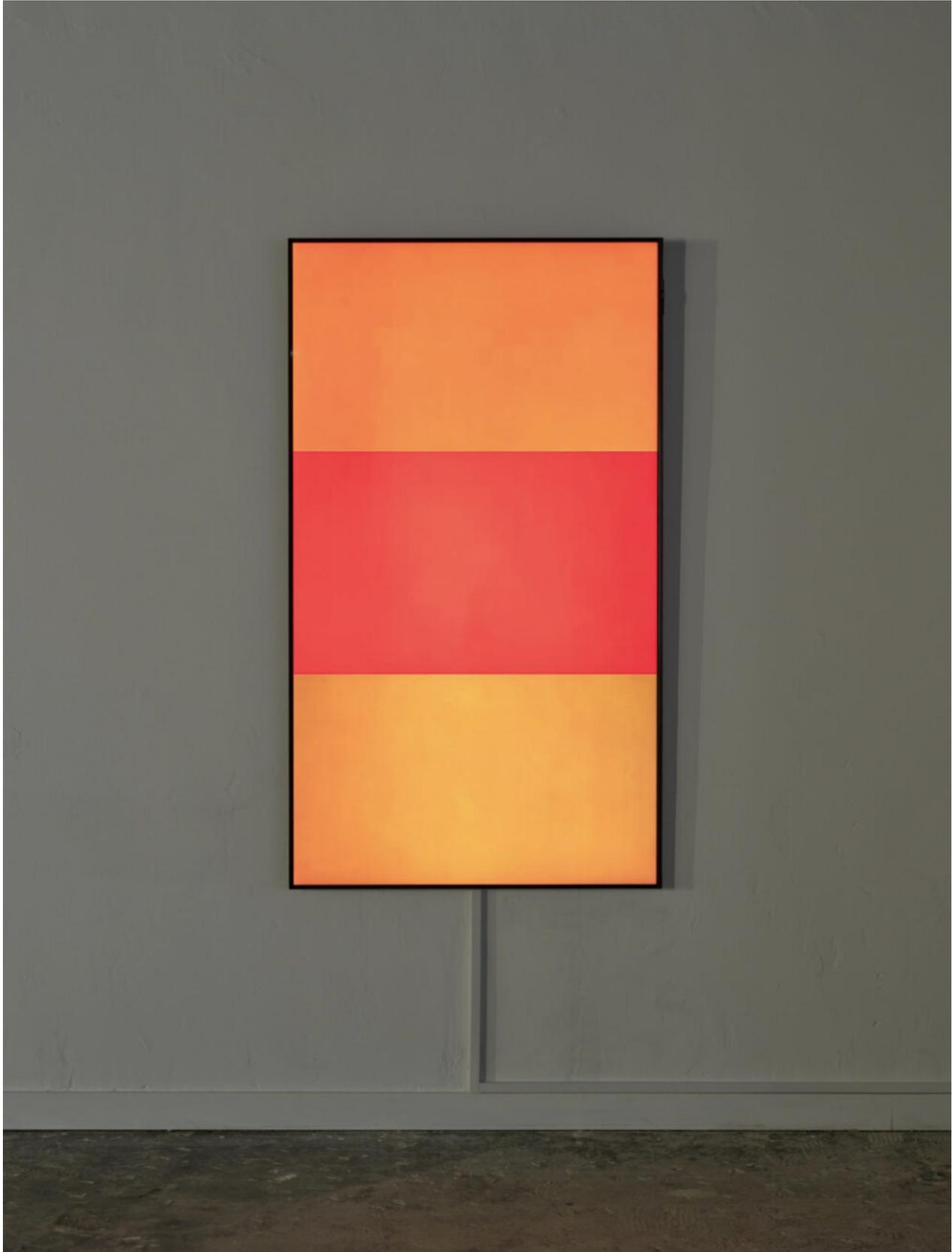
Die Kunstkritikerin Lucy Lippard sagt in *Sweeping Exchanges: The Contribution of Feminism to the Art of the 1970s*:

„Wir halten es für selbstverständlich, dass Kunst zu machen nicht nur bedeutet, sich selbst auszudrücken, sondern eine viel umfassendere und wichtigere Aufgabe ist: sich als Mitglied einer größeren Einheit oder Gemeinschaft auszudrücken.“

Ich stimme ihrem Statement in jeder Hinsicht zu. Ich denke, das ist es, was mich immer wieder zur Arbeit mit Gruppen und Chören zurückführt. Ohne sie kann ich meine Arbeiten nicht erschaffen, und durch sie bekommen meine Arbeiten überhaupt erst einen tieferen Sinn.

Die digitalen Möglichkeiten sind daher aufregend für mich, weil es darum geht, kollektive und neue, integrativere Wege der kollaborativen Werkgenese zu eröffnen. Kürzlich haben z.B. Kinder auf TikTok ein vollständig von Gamer*innen generiertes, dezentrales Spiel entwickelt, das sie gemeinsam erschaffen haben, weiterentwickeln und natürlich auch spielen – weltweit. Was für eine unglaubliche kollektive Leistung, die zeigt, dass Technologie ein verbindendes Werkzeug sein kann und Coding beispielsweise das Potenzial zur universellen Sprache mitbringt. Mein Kunstwerk *Cabin Fever* ist etwa vollständig digital entstanden, ein audiovisuelles Kunstwerk, das auf Zoom mit professionellen und

Laien-Sänger*innen aus der ganzen Welt geschaffen wurde.



Installation *Cabin Fever* beim TULCA Festival, Galway (2022). Foto: Ros Kavanagh.

Viele der involvierten Personen haben online Bande geknüpft und rasch versucht, sich auch im

wirklichen Leben zu treffen. Das Hauptaugenmerk aller Interpret*innen lag generell darauf, in Kontakt zu bleiben und sich verbunden zu fühlen, auch wenn die Technologie letzten Endes ihre Stimmen oft unhörbar machte. Das bestätigte mich in meinen großen Hoffnungen, dass sich bald Wege finden werden, das globale gemeinsame Musizieren auch technologisch sichtbar zu machen und Partizipation zu erleichtern.

In der Forschungs- und Entwicklungsphase von *Cabin Fever* sprachen wir mit vielen Chören auf der ganzen Welt, die keinen zuverlässigen Internetzugang hatten oder denen es aus anderen Gründen schwerfiel, an diesem Projekt teilzunehmen. Die Begeisterung, sich zu verbinden und gemeinsam aufzutreten, war weltweit zu spüren.

Johanna Elisabeth Moeller und Katrin Etzrodt vom TUD-Institut für Kommunikationswissenschaft ermöglichten es, einen Auszug aus *Cabin Fever* bei ihrem *Exploring Socio-Technical Research Workshop* im Sommer 2022 zu zeigen, einem transdisziplinären und methodenübergreifenden Projekt. Die Arbeit mit Forscher*innen zu teilen, war für sie und auch für mich sehr inspirierend und löste viele Gespräche über digitale Gemeinschaftsmusik aus und auch darüber, wie damit zusammenhängende Möglichkeiten genutzt werden können, um wiederum neue Gemeinschaften zu ermöglichen.

Welche Rolle spielt für Sie die Idee eines Publikums im Allgemeinen, etwa in Bezug auf die experimentelle Beteiligung anderer? Und wie wichtig ist der damit verbundene performative Aspekt?

In meiner Arbeit verwische ich gern die Grenze zwischen Publikum und Darsteller*innen. Für mich sind sie alle Teil des Stücks. Ohne das Publikum, die Beobachter*innen, Personen, die das Geschehen reflektieren, bleibt es nur eine Probe. Auch bringt das Publikum buchstäblich Körper in den Raum: Ohren, die hören, hängen wie Verstand und Emotion mit körperlichen Erfahrungen zusammen – das Wahrnehmen von Vibrationen, die Reaktionen und Klänge, die nicht zum Stück gehören, wie die Geräusche des Publikums, aber auch der Körper als Resonanzraum und Widerstand, gar das ganze Raumerlebnis an sich, der Ton und die Präsenz des Ortes. Dahingehend ist meine Arbeit audiovisuell, alle Teile sind gleichwertig, und meine Arbeit existiert nicht ohne ein Publikum, das zuhört.

Meine Klanginstallationen und Performances werden folglich immer mit Blick auf ein Publikum erdacht, das sich frei im Raum bewegt. Räume klingen anders, wenn Körper in ihnen sind, die Resonanz verändert sich. Wir machen immer Tests, bevor ein Publikum den Raum betritt. In gewisser Weise ist das Publikum also schon bei den ersten Skizzen dabei: materiell und klanglich.

Während des Aufenthalts in Dresden schwebte mir zunächst ein Stück ohne Menschen vor, aber ich begriff rasch, dass an KI-Systemen unvorstellbar viele Menschen beteiligt sind. Daher änderten sich auch meine visuellen Vorstellungen. Ich sah plötzlich Menschen in einem quadratischen, lichtdurchfluteten Raum stehen und zuhören, sah eine Bewegung von A nach B: eine Bewegung, interpretierbar als Systemarchitektur eines KI-Modells.

Für w/k ist es wichtig, das hier stattfindende Zusammenspiel von Kunst und Wissenschaft möglichst genau zu bestimmen. Beschreiben Sie bitte exemplarische Verbindungen zwischen Kunst und Wissenschaft, die für Sie in den einzelnen Komponenten Ihrer Projekte relevant wurden.

Die Zeit mit den Forscher*innen und die für mich sehr wertvollen Gespräche und neuen Impulse haben die Richtung meiner Arbeitsansätze und Ideen stark geprägt. Nachfolgend möchte ich einige Beispiele aufführen:

Miriam Akkermann, Juniorprofessorin für empirische Musikwissenschaft an der TUD, verbrachte einige

Monate damit, sich mit mir über meine ursprüngliche KI-Chor-Idee auszutauschen. Ihr gefiel besonders der Ansatz, den menschlichen Klang nicht zu kopieren, und sie stellte Überlegungen an, wie man ein solches Modell theoretisch aufbauen könnte, was man messen könnte, sprach mit mir über Ideen und inspirierte mich zu neuen Überlegungen.

Sebastian Merchel, Wissenschaftlicher Mitarbeiter des Lehrstuhls für Akustik und Haptik (Institut für Akustik und Sprachkommunikation, TUD), führte mit mir Tests mit Geräuschbeispielen verschiedener Maschinen durch, die ähnlich klangen wie die Geräusche von Serverfarmen, die ich bei meinem ersten Besuch vor Ort gehört hatte. Wir untersuchten gemeinsam, wie sich diese Geräusche in Bezug auf Frequenz, physikalische Vibration und Hörerlebnis unterscheiden. Wir probierten unterschiedliche Bodenschwingungen für verschiedene Klänge aus. Da unsere Methoden zur Nutzung eines Testpublikums sehr ähnlich sind, hoffen wir, auch in Zukunft zusammenzuarbeiten. Ziel ist es, ein Kunstwerk zu schaffen, das Vibrationen einbezieht, und Sebastian Merchel in die Lage versetzt, Experimente und weitere Forschungen über das Kunstwerk und die entstehenden Rezeptionsperspektiven durchzuführen.

Orit Halpern, Professorin und Inhaberin des Lehrstuhls für Digitale Kulturen und gesellschaftlichen Wandel (Institut für Germanistik und Medienkulturen, TUD), hat sich viel Zeit genommen und mit mir ihr fundiertes Wissen über digitale Infrastrukturen geteilt. Weitere Forscher*innen aus Halperns Team außerhalb der TUD, wie Nanna Thylstrup, Nelly Yaa Pinkrah und Michelle Pfeifer, haben sich dem Gespräch angeschlossen und uns an ihren neuesten Erkenntnissen über die Dekolonialisierung von KI und Spracherkennungsalgorithmen teilhaben lassen, die wir in Form von öffentlichen Veranstaltungen diskutieren konnten.

Eine weitere fantastische Gelegenheit war die Einladung von Matthew McGinity, den ich bereits erwähnt habe, an seinem zweiwöchigen transdisziplinären IXLAB-Workshop für Informatikstudierende mitzuarbeiten. Gemeinsam mit seinem Team ergründeten wir die Erzeugung von räumlichen Klangerlebnissen, indem wir Echtzeitsysteme konstruierten, um Klang und Musik in Strukturen und Umgebungen zu verwandeln, die in einer virtuellen Umgebung erkundet werden konnten. Seine Studierenden untersuchen den Einsatz generativer und emergenter Systeme und wie diese Systeme in Verbindung mit Musikinstrumenten, Klängen und der Stimme als Schnittstellen für die Gestaltung virtueller Welten genutzt werden können. Konzepte im Zusammenhang mit Emergenz, Komplexität und Echtzeit sowie mit sinnesübergreifenden Wahrnehmungsverfahren inklusive der Einbindung musikalischer, räumlicher und modalen Kognition sind Schlüsselpunkte dieser Analyse. Insbesondere forschten die Studierenden an Klangerfahrungen durch die Brille der *4E-Kognition* als einer Aktivität, die Personen sich durch Handlung erschließen können und die in den Raum eingebettet und erweitert ist. Wir verlegten den gesamten Workshop ins *COSMO* im *Kulturpalast* Dresden und nutzten das Foyer als Augmented-Reality-Raum. Da die Chöre der Philharmonie im selben Gebäude beheimatet sind, besuchten wir deren Proben und führten erste Gespräche, um Wege der Zusammenarbeit zu finden.



IXLAB-Workshop (2022). Foto: Schaufler Lab@TU Dresden.

Welche Vorteile hat es, als Künstlerin im Bereich der Wissenschaft zu arbeiten, und wie kann die Wissenschaft von Ihrer Arbeit profitieren? Was können Wissenschaftler*innen von einem dezidiert künstlerischen Umgang mit Forschungsmaterial, Technologien usw. lernen?

Wenn die Voraussetzungen stimmen, kann diese Zusammenarbeit für beide Seiten sehr inspirierend sein. Für eine echte Interferenz muss man die eigene Komfortzone verlassen, dann können neue Ideen gemeinsam auf fruchtbarere Weise erforscht werden. Zuhören können, Geduld haben, sich Zeit nehmen, etwas anderes tun, neugierig und offen sein, Raum für eine solche Begegnung zwischen Wissenschaft und Kunst schaffen – das sind die wesentlichen Punkte für ein produktives und wertschätzendes Miteinander.

Ich habe schon viele interdisziplinäre Projekte durchgeführt. Der Unterschied bei der Zusammenarbeit mit Forscher*innen unterschiedlicher akademischer Disziplinen besteht darin, dass man Gespräche auf einer trans- und interdisziplinären Ebene führt, die das eigene Denken auf eine neue Art und Weise herausfordern und inspirieren können: Erkenntnisse, Methoden, andere und neue Denkweisen, die man anders nie erreichen könnte, entstehen – sie durchdringen und bilden sich wechselseitig ab. Gemeinsam gelingt es, einen anderen, neuen Blick auf die Welt zu wagen.

Links:

▷ www.econderuiz.com

▷ Symposium *Zukunftsmusik. What the future might sound like:*
<https://www.youtube.com/watch?v=UXx3w9pM898>

Im Frühjahr 2023 wird Conde Ruiz eine Ausstellung im *HELLERAU - Europäisches Zentrum der Künste*, einem der wichtigsten Zentren für Gegenwartskunst, Tanz und Musik in Deutschland, präsentieren. Sie wird ein ortsspezifisches Kunstwerk präsentieren, das die Synthese jener Ideen darstellt, die sie während ihrer Residenz gesammelt hat: <https://www.hellerau.org/de/event/24-7/>

Beitragsbild über dem Text: Symposium *Zukunftsmusik. What the future might sound like*, TUD (2022).
Foto: André Wirsig/© Schaufler Lab@TU Dresden.

Tags

1. Artist in Residence
2. Esmeralda Conde Ruiz
3. Interviews
4. Kooperation
5. Künstliche Intelligenz (KI)
6. Michael Klippahn-Karge
7. wissenschaftsbezogene Kunst