ChAI => Interactive AI Tools in ChucK

Yikai Li
CCRMA, Stanford University
Stanford, California, United States
yikaili@stanford.edu

Ge Wang
CCRMA, Stanford University
Stanford, California, United States
ge@ccrma.stanford.edu

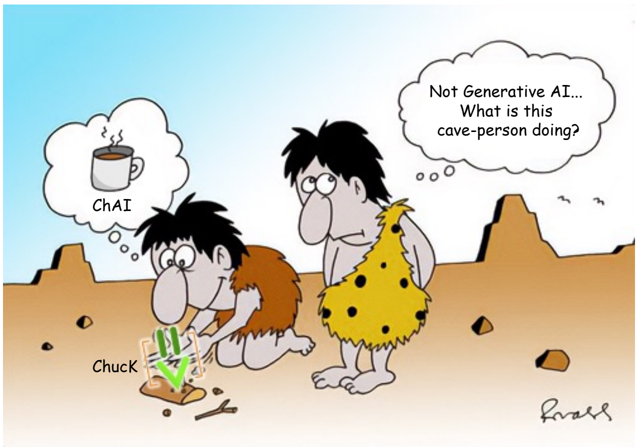


Figure 0: ChAI eschews the generative AI tools of today for time-honored interactive machine learning techniques, amenable for designing a wide range of AI-mediated musical tools and toys.

ABSTRACT

This paper introduces ChAI, a collection of interactive machine learning tools for the ChucK music programming language, and chronicles its use in Music and AI, a critical-making course at Stanford University. We explore the intersection of AI and music HCI, emphasizing the augmentation of human creativity, rather than its replacement. We introduce ChAI's interactive AI tools and toys and highlight philosophical, ethical, and cultural considerations. Focusing on interactive machine learning, audio analysis, feature extraction, and audio re-synthesis, this paper shows how ChAI could be used to create playful, expressive systems (in an age of generative AI that tends to limit human interaction in favor of prompt-based automation). In the critical-making course, ChAI was used alongside reflections on AI's broad impact on the arts and who makes art, while stressing the importance of understanding AI's societal and cultural implications. This work not only introduces a new tool but also invites broader conversation on the role of HCI, play, and "small-data" approaches, in music and AI practice and education.

NIME'24, 4–6 September, Utrecht, The Netherlands.

Author Keywords

Interactive machine learning, music and artificial intelligence, critical making, education, music programming languages and tools.

CCS Concepts

•**Applied computing** → *Sound and music computing;*
•**Computing methodologies** → *Machine learning;*
•**Human-centered computing** → *Interactive systems and tools;*
•**Software and its engineering** → *Domain specific languages;*

1. INTRODUCTION

The convergence of artificial intelligence (AI) and music is an ever-evolving frontier (and an increasing set of difficult questions) in the realm of creative technologies. ChAI, a suite of machine learning tools designed for the ChucK music programming language, plays a supportive role in this intersection. This paper documents the deployment of ChAI in the critical making "Music and AI" course at Stanford University, offering a tool and perspectives on the interplay between AI, music, and human-computer interaction (HCI). The purpose of this exploration is multifaceted, including the creation of interactive AI tools, the promotion of human creativity through AI (and human interaction, curation, and effort), and the consideration of the philosophical and cultural implications inherent in this fusion of technology and art.

Central to ChAI is an emphasis on making creative uses with a "small data mindset" [18]. This approach departs from the traditional reliance on large-scale data models and instead emphasizes the human creativity that can be achieved with smaller datasets and models that can be trained at interactive rates. This mindset is particularly relevant when considering new interfaces for musical expression, as it promotes interaction, creativity, and iterative designs.

In a way, ChAI "returns to basics" and promotes well-understood machine learning techniques that have been shown to be susceptible to interaction design: a vanilla Multilayer Perceptron (MLP), k-Nearest Neighbors (k-NN), Hidden Markov Model (HMM), Support Vector Machine (SVM), Word2Vec, Principal Component Analysis (PCA), and Wekinator, as well as an expanded suite for audio analysis.

A fundamental aspect of our exploration is the need for critical reflection [12, 19] when integrating AI with music. This involves closely analyzing the aesthetic and ethical dimensions, as well as cultural implications involved in AI-mediated music creation. It goes beyond the technical aspects of AI, taking into account its wider societal implications and its effects on the nature of creativity, art, who makes art, and why. It is our belief that adopting a reflective stance is crucial for comprehending and navigating the intricate interplay between technology, creativity, and our ever-shifting cultural landscape.

A key theme in this discussion is the balance between machine automation and human interaction in creative processes. This contrast captures both the excitement for and caution towards AI's role in enhancing musical creativity. Examples like Wekinator [7] demonstrate the potential of AI as a collaborative partner in creativity, while also emphasizing the need for a critical and mindful approach to its integration. The challenge lies in finding a harmony where AI empowers rather than overwhelms or supplants human creativity.

ChAI was used as the primary tool in a "Music and AI" course at Stanford University. The course description reads:

> How do we make music with artificial intelligence? What does it mean to do so (and is it even a good idea)? How might we artfully design tools and systems that balance machine automation and human interaction? More broadly, how do we want to live with our technologies? Are there—and ought there be—limits to using AI for art? (And what is Art, anyway?) In this "critical making" course, students will learn practical tools and techniques for AI-mediated music creation, engineer software systems incorporating AI, HCI, and Music, and critically reflect on the aesthetic and ethical dimensions of technology.
>
> Coursework will span the practical ("how?"), the philosophical ("why?"), and the social ("for whom?"). Students will use these techniques and lenses to think as broadly as possible about the implications of AI in our world and to design interactive AI tools that keep human judgment in the loop. Through these exercises, we will explore how AI might augment, not replace, human creativity.

As a holistic evaluation of ChAI as a tool, we present several student projects in "Music and AI" to demonstrate a selection of what can be made with ChAI. Projects like "Pandora's Dream", "Music-Driven Motion Matching", and "Face Guitar", exemplify the interactive and humanistic uses of AI in music, and underscore the versatility of ChAI as a programmable tool.

2. RELATED WORK

The development of Human-Computer Interaction (HCI) in music, influenced by works such as [16] on designing gestural interfaces, has led to the creation of interactive and expressive digital musical instruments. This HCI-centric approach is further elaborated by [19], which highlights the significance of integrating humans in the loop of AI system design.

This is somewhat in contrast to the trend of music generation in AI. Projects like Google's Magenta [17] and MusicLM [1] employ deep learning and subsequent foundation model techniques to create "finished" products. Alongside these technological advancements, the integration of AI in music raises ethical and philosophical implications, particularly concerning authorship and the authenticity of AI-generated content [15].

In this work, ChAI intentionally abstains from a generative AI system and instead takes an HCI-centered approach. A notable shift in this approach is the emerging preference for a "small data mindset". Vigliensoni et al. [18] exemplify this approach, which emphasizes the efficiency and creativity achievable with smaller datasets compared to traditional large-scale data models. It signifies a move towards interaction, creativity, and real-time adaptability in musical applications.

The application of Interactive Machine Learning (IML) in music, as seen in tools like Wekinator [7, 10], offers musicians new avenues for real-time creative collaboration with AI systems. Technologies such as the Multilayer Perceptron (MLP) and Support Vector Machine (SVM) have been effectively applied in music analysis and generation [13, 5], further illustrating the breadth of AI's impact on music, especially when integrated with human curation.

In the realm of education, initiatives like Berklee College of Music's Machine Learning for Musicians course highlight the growing integration of AI in music education. These initiatives equip students with novel tools and perspectives on music composition and performance. Case studies like 'XAI' [2] underscore the practical utility of AI in enhancing musical creativity and audience engagement.

The critical making approach in AI and the arts emphasizes AI's role as a medium for artistic exploration and societal critique. This perspective invites a broader conversation about the intersection of technology, creativity, and culture, underscoring the need for interdisciplinary research and education in this evolving domain.

3. CHAI OBJECTS

We describe a few of the basic ChAI tools below. Note these are not "state of the art" but are chosen because they have been demonstrated to be amenable to building highly interactive musical interfaces [7]. Many of these can paired with the existing Unit Analyzer framework [20] to perform feature extraction to build systems both for analysis (e.g., sentiment or genre classification or segmentation) as well as synthesis systems (e.g., feature-based similarity retrieval and real-time audio mosaics). Existing audio features are shown in Table 1.

Feature	Description
Centroid	Measures the center of mass of the audio's spectrum, indicating the brightness or sharpness of the sound.
Chroma	Represents energy distribution across pitch classes, essential for analyzing harmony.
Flux	and Captures rate of change in spectral power, highlighting dynamic changes in sound.
Kurtosis	Indicates shape of frequency distribution, showing extremeness of spectrum.
MFCC	Transforms the frequency domain into the Mel-scale, modeling the human ear's characteristics, essential in voice recognition.
RMS	Represents average power or 'loudness' of the audio signal.
Rolloff	Identifies frequency below a certain percentage of total spectral energy.
SFM	Measures 'noisiness' of sound, indicating similarity to noise versus tonal sound.
Zerox	Counts waveform crossings of zero amplitude line, useful in genre classification and tempo estimation.

Table 1: A few audio analyzers for feature extraction.

3.1 Multilayer Perceptron (MLP)

The integration of the Multilayer Perceptron (`MLP`) [13] in ChucK provides a new way for music creation and analysis. As a fundamental artificial neural network, the `MLP` tool excels in regression and classification tasks. In musical contexts, this capability allows for nuanced interpretations of complex musical patterns and dynamics. The `MLP` in ChucK has been optimized for real-time processing, enabling composers and performers to experiment with adaptive, AI-based compositions. This tool also supports a range of activation functions, learning rates, and layer configurations, offering users a customizable experience tailored to specific musical needs.

`MLP`'s basic functions include `.init()`, `.train()`, and `.predict()`. The code snippet below demonstrates training a basic artificial neural network using `MLP`:

```
// Create a MLP for general supervised learning
MLP model;

// Initializing MLP layers
model.init( nodesPerLayer );

// Training with a mapping of inputs and outputs
model.train( X, Y, learningRate, epochs );

// Predicting output based on new input
model.predict( x, y );
```

ChAI's `MLP` object also allows for a more explicit, step-by-step training with additional functions such as `.forward()`, `.backprop()`, and `.shuffle()`, useful for education purposes.

3.2 k-Nearest Neighbors (k-NN)

ChucK's k-Nearest Neighbors (k-NN) [6] algorithm provides an intuitive yet powerful method for musical data analysis. This implementation facilitates real-time music classification and clustering, making it an invaluable tool for interactive music systems. By analyzing the proximity of data points, `k-NN` in ChucK allows for dynamic musical decisions based on the similarity of incoming data to pre-existing datasets. This feature is particularly useful in applications such as genre classification, mood-based music generation, and adaptive musical accompaniment systems. Furthermore, ChucK's `k-NN` is designed to handle high-dimensional data efficiently, ensuring smooth performance even in complex musical scenarios. Basic functions of `KNN` and `KNN2` include `search()`, `train()`, and `predict()`. A basic example is:

```
// Create a k-NN, e.g., for similarity retreival
KNN model;

// Training with features
model.train( features );

// Finding nearest neighbors
model.search( q, 2 , neighbor_indices );
```

3.3 Hidden Markov Model (HMM)

The Hidden Markov Model (HMM) [3] tool in ChucK enhances AI-driven music creation by generating observation sequences that model temporal musical structures. This capability is particularly effective in capturing the probabilistic nature of musical progression and style evolution. `HMM` in ChucK can be used for tasks such as automatic composition, improvisation, and style transfer. It provides a framework for understanding and replicating the underlying patterns in musical pieces, enabling the creation of novel compositions that maintain stylistic coherence. Additionally, its versatility in modeling time-series data makes it an excellent tool for live performance, where real-time generation and adaptation of musical sequences are paramount.

HMM's functionalities include `generate()`, `load()`, and `train()`. An example use case is:

```
// Create a HMM for sequence generation
HMM model;

// Training with observations
model.train( states, emissions, observations );

// Generating new sequences
model.generate( 30, results );
```

This example showcases HMM's ability to generate sequences (e.g., melody or drum patterns) based on trained observation data.

3.4 Support Vector Machine (SVM)

The Support Vector Machine (SVM) [5] tool in ChucK stands as a robust method for model training and output prediction in various musical contexts. `SVM`'s ability to handle both linear and non-linear data makes it exceptionally versatile for music analysis and generation. This feature is particularly beneficial for classification tasks, such as instrument recognition and audio feature classification. In ChucK, `SVM` has been optimized for speed and accuracy, ensuring that it can be effectively used in real-time applications. Moreover, its capacity to deal with large feature sets makes it ideal for processing and analyzing complex musical data, providing insights that can drive creative and analytical processes in music.

```
// Create a SVM object for classification
SVM model;

// Training with labeled examples
model.train( rhythmFeatures, rhythmLabels );

// Classifying new inputs
model.predict( testFeature, testResult );
```

3.5 Word2Vec

Implementing Word2Vec [14] in ChucK offers real-time semantic text-based exploration in musical applications. By mapping words to vectors, `Word2Vec` allows for representation of linguistic entities as mathematical coordinates, amenable for similarity search and retrieval. This semantic mapping can facilitate music composition and analysis, such as thematic association and lyrical analysis and synthesis.

Key functionalities include `load()`, `getSimilar()`, and `contains()`. An application example is:

```
// Create a Word2Vec object
Word2Vec model;

// Loading vectorized words model
model.load( ``lyricModel.txt'' );

// Finding similar words
model.getSimilar( ``love'', 5, similarWords );
model.getSimilar( vector, 5, similarWords );
```

3.6 Principal Component Analysis (PCA)

The `PCA` [11] tool in ChucK offers an essential solution for dimensionality reduction in complex musical data. By transforming data into a set of linearly uncorrelated variables, `PCA` in ChucK simplifies the complexity inherent in multidimensional musical datasets. This simplification is crucial for efficient data visualization, pattern recognition, and feature extraction in music analysis. PCA's ability to highlight the most significant features in a dataset makes it an invaluable tool for composers and researchers who are dealing with large-scale musical data, enabling them to focus on the most impactful elements of their work. `PCA`'s main function is `reduce()`:

```
// Create a PCA for Dimensional Reduction
PCA.reduce( input, 3, output );
```

3.7 Wekinator

The integration of the Wekinator [7] utility in ChucK marks a powerful step forward for interactive machine learning for musical expression in ChucK, but also represents something of a homecoming, as Wekinator has early roots as experiments in ChucK [8, 9]. `Wekinator` in ChucK allows users to train models interactively, using input from musical instruments or other controllers. This functionality opens up exciting possibilities for the design and development of new AI-driven musical instruments and interfaces. By facilitating real-time interaction and learning, `Wekinator` empowers musicians and composers to create highly adaptive and personalized musical experiences. This tool's flexibility and ease of use make it an ideal choice for both novices and experts in the field of music technology. Furthermore, having `Wekinator` as a code-programmable entity allows for different architectures (potentially involving different `Wekinator` instances) and usages, and supports both Open Sound Control as well as direct integration between audio synthesis, analysis, and `Wekinator`.

```
// Create a Wekinator object
Wekinator wek;

// Interatively add training data
wek.add( inputVector, outputVector );
wek.add( ..., ... );
// Train the model based on current data set
wek.train();
// Predicting output
wek.predict(inputVector, outputVector);

// Refine by adding more training data
wek.add( inputVector, outputVector );
// Train and Predict, etc...
```

4. CASE STUDIES

These case studies originate from Music 356 / CS 470 Music and AI: A Critical Making Course at Stanford University. We also reflect on the importance of "critical making" in teaching Music and artificial intelligence.

4.1 Pandora's Dream

At the intersection of live coding, interactive AI, audiovisual design, and music performance is "Pandora's Dream" by Celeste Betancur [4], which showcases the effective use of ChAI to enhance musical expression and interaction, while keeping the human firmly in the interaction and artistic loop. The piece has been presented in several concerts:

- "Invisible Powers - In Pandora's Hands" was presented in Bing Concert Hall (June 10th, 2023, Stanford, CA) in a live coding performance where the performer controls 12 laptops simultaneously plus 124 phones (audience devices).

- "Pandora's Mycophony" presented in Centro de la Cultura digital (June 16th, 2023, CDMX, Mex) as part of the NIME conference, in this performance the system interacts with two live coders in real-time.

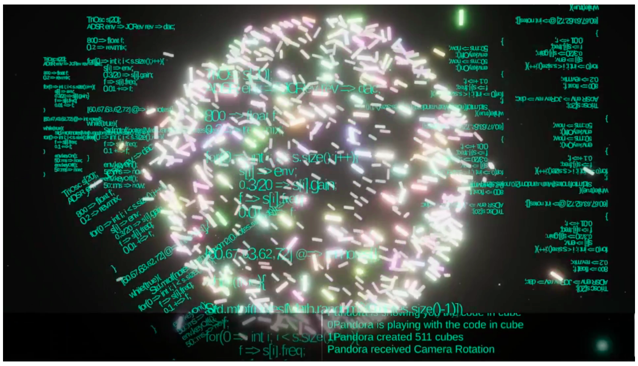


Figure 1: Pandora's Dream is a live-coding, audiovisual performance system.

Real-Time AI Integration: At its core, "Pandora's Dream" combines ChAI and live coding in ChucK, with real-time interaction with AI and machine learning algorithms. This integration allows for dynamic and responsive interactions during performances, using the aforementioned "small data mindset" to interactively train and generate units of musical building blocks (such as rhythmic patterns using `HMM`).

Rapid AI Model Training with Audio Samples: The project features an unsupervised AI model that can be trained in real-time, with a training process that can take place in milliseconds to a few seconds. This rapid training phase enables the extraction of complex audio features such as Centroid, RMS, and Mel Frequency Cepstral Coefficients from live audio inputs for similarity retrieval and concatenative synthesis.

Performer-Centric Approach and Live Model Modification: ChAI in "Pandora's Dream" allows for real-time modification of AI models during live performance, offering a high degree of customization and control. The live coding gives the audience a transparent view into the performer's processes, revealing both the technology within and underscoring the importance of human curation and interaction.

In summary, the integration of ChAI in "Pandora's Dream" represents an organic fusion of AI, live coding, and audiovisual performance systems, and has been used in settings ranging from academic computer music conferences to nightclubs.

4.2 Music-Driven Motion Matching

The "Music-Driven Motion Matching" project [1], created using ChAI, demonstrates the ability to generate real-time dance movements based on input audio. This system is built upon the analysis of the AIST dance database, which provides a collection of dancing videos along with associated audio and movement data.

Audio Analysis and Dance Database Integration: The project utilizes ChAI's audio analysis module to extract key audio

---

[1] Music Motion: `https://youtu.be/Pb9h6G9skIA`

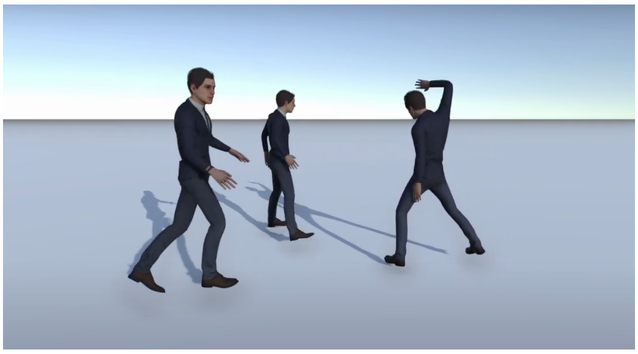


Figure 2: Music-driven motion-matching real-time dance generation tool/toy.

features such as RMS, MFCC, and others from the input music. These features are then correlated with dance movements to create a comprehensive database. This integration allows the system to understand and connect audio elements with corresponding dance actions.

Real-Time Audio-Driven Dance Generation: During real-time operation, the system takes incoming audio, extracts relevant audio features, and uses ChAI's `KNN` module to identify similar fragments in the database along with their movement data. This process enables the real-time generation of dance movements that are synchronized with the audio inputs.

Unity's Motion Matching System: The project employs Unity's advanced motion matching system to efficiently generate dance movements. This system ensures that the generated movements not only align with the current audio but also appear fluid and natural. Unity's technology contributes to the project's high-quality performance, making it comparable to larger dance generation models.

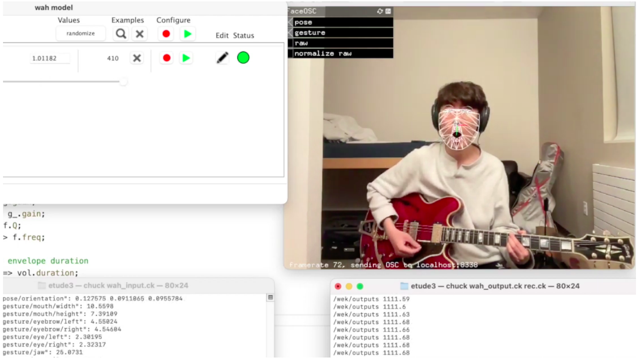
Responsive Dance Movements: One of the project's notable features is its ability to adapt to different musical inputs. It interprets various musical elements such as tempo, rhythm, and melody, and translates them into expressive dance movements. This adaptability ensures that each piece of music is uniquely represented through dance, enhancing the overall audio-visual experience.

Interactive User Experience: The project emphasizes user interaction and engagement. Users have the ability to manipulate audio inputs in real-time, influencing the dance generation process. This interactive feature immerses users in the experience and encourages exploration of the relationship between music and its visual representation through dance.

In conclusion, the "Music-Driven Motion Matching" project showcases the capabilities of ChAI in bridging audio and motion, enabling real-time, responsive, and interactive musical expression. It doesn't always "make sense" but even the goofy and awkward results are expressive and interesting.

4.3 Face-controlled Audio Effects

This project [2] demonstrates an expressive use of ChAI in combination with FaceOSC to redefine the interaction between a guitarist and their instrument. By using facial expressions and gestures as control mechanisms for guitar effects, traditional human expressions become proactive musical inputs, reversing the conventional dynamic of guitar performance.



```
 g_.gain;
f.Q;
> f.freq;

 envelope duration
=> vol.duration;
```

```
pose/orientation": 0.127575 0.0911065 0.0955784
gesture/mouth/width": 10.5598
gesture/mouth/height": 7.39109
gesture/eyebrow/left": 4.55024
gesture/eyebrow/right": 4.54604
gesture/eye/left": 2.30195
gesture/eye/right": 2.32317
gesture/jaw": 25.0731
```

```
/wek/outputs 1111.59
/wek/outputs 1111.6
/wek/outputs 1111.63
/wek/outputs 1111.68
/wek/outputs 1111.66
/wek/outputs 1111.68
/wek/outputs 1111.68
/wek/outputs 1111.68
```

Figure 3: Guitar effects control facial expressions and gestures

Integration of Facial Expression and Gesture Detection: Through the use of ChAI and FaceOSC, the system controllably interprets facial expressions and gestures. This integration is crucial as it transforms spontaneous emotional responses into direct inputs for sound modulation. This approach not only recognizes but actively utilizes the musician's emotional engagement, fostering a more intimate and responsive interaction with the instrument.

New Interfaces for Facial Expression (NIFE!): By replacing foot pedals with facial expressions and gestures, the project aligns the modulation of sound effects with the guitarist's expressive output. This shift from physical to facial expression-based control offers new methods of musical expression, enabling deeper emotional resonance and expressivity in performances.

Diverse Effect Mapping: ChAI's adaptable framework allows for mapping specific facial movements to distinct sound effects:

- **Wah Effect**, traditionally controlled through pedal manipulation, is now controlled by mouth movements, aligning physical expression with sound modulation.
- **Pitch Shift** (including vibrato and bending) is modulated through head orientation, providing nuanced and expressive control over pitch.
- **Volume Pedal** also utilizes head orientation, allowing dynamic control over sound intensity.

Expressiveness is the Goal This system does not purport to solve any "problems" as it proactively explores different spaces for expression (both sonic and facial!). The integration of emotional and physical expressions with real-time sound modulation creates a playful new medium for musical exploration.

4.4 Reflection

4.4.1 *Small Data Mindset*

In the rapidly evolving landscape of artificial intelligence in music, there is a shift from relying on large-scale data models to a more refined, more interactive "small data mindset." This approach emphasizes efficiency and creativity in AI models, operating effectively with smaller datasets. It is particularly significant in new interfaces for musical expression, where interaction, creativity, and real-time adaptability are crucial.

The small data mindset focuses on training AI models with limited data paired with nuanced human interaction. Models trained on smaller datasets can adeptly respond to

---

[2] Face Guitar: `https://youtu.be/--Lp2WFTOwU`

musical interactions in real-time, allowing for greater experimentation, personalization, and creative exploration. This contrasts with larger data models, which require extensive computational resources, training data, and time, limiting their flexibility and responsiveness.

This mindset also promotes rapid prototyping and "by-example" learning in interactive AI and instrument design. By utilizing a limited set of input-output examples, AI can efficiently understand and replicate complex behaviors. When integrated with human judgment and sensibility, this can yield powerful and expressive systems.

Moreover, the small data mindset encourages the use of simpler models and smaller datasets to achieve substantial outcomes. Simpler configurations, such as basic neural network models, can be surprisingly expressive and impactful, demonstrating that significant results can be attained without complex systems and extensive data.

In essence, the small data mindset in musical AI advocates for less being more. It represents a thoughtful, targeted approach that focuses on what can be effectively achieved with limited resources. This mindset breaks from conventional expectations about data size and model complexity, fostering a more accessible, personalized, and creative approach to musical expression in the AI domain. It embodies a paradigm shift, advocating for a balance between technological capability and creative flexibility, essential for the future of musical AI and NIME.

4.4.2 Critical Reflection

Critical reflection plays a vital role in the integration of AI with music and making. This deliberate and thoughtful process is essential before, during, and after system design and potential deployment, and involves unpacking the social, aesthetic, historical, and cultural implications.

The Music and AI course advocates for a multidisciplinary understanding across engineering, the humanities, the arts, and the social sciences, emphasizing that AI proficiency must include a grasp of its wider implications, or at least a healthy mindset to frame what is at stake. The influence and role of AI necessitate ongoing scrutiny and evaluation. As such, a course of study in any AI system (including music) ought to extend well beyond technological competence.

The integration of AI in music and arts also demands a critical examination of the technology itself. This involves questioning the validity and role of AI in music creation, urging a reflective stance on how these technologies influence creativity and human experience. The Music and AI course alternates between creative design using ChAI and as well as critical reflections on coursework, research, and readings. As one example, the following is a prompt from a critical response:

> **Critical Response #1 prompt:** Do a critical response regarding generative music AI technology and more broadly probe the question, "What do you (really) want from AI music generation?"
>
> Gain a sense of the merits and limitations of generative music AI systems for what they respectively aim to do (the provided examples are a good place to start for this). At the same time, as a critical exercise, think as broadly as possible about its implications for culture, artistic practice, its promise, and its perils, and who might be impacted and in what ways. Question the premise of the stated goals of the system—possibly starting with a critical (not cynical) pondering of the question, "Is this even a good idea?" Favor lens-building over opinion sharing; favor well-framed questions over verdicts. One fun and potentially telling exercise is to imagine a version of the world where the system in question is in prevalent use. In what ways does that version of the world align with the world you would want to see? In what ways does it not? Pick a particular aspect of your reflections on this question, and write a 500-word response (yes, you can optionally use chatGPT to help with the task, but the viewpoint must be yours). Happy reflecting.

We believe these kinds of reflections are not only academic exercises but also a necessary tool for navigating the complex interplay between technology and our culture and making our best effort to ensure that the use of AI in music remains thoughtful, ethical, and creatively enriching.

4.4.3 Machine Automation vs. Human Interaction

In contemporary music production, the interaction between machine automation and human involvement is pivotal in influencing the creative outcome. This dichotomy represents a fusion of enthusiasm and caution towards AI in shaping musical expression. While there is excitement about AI's role in augmenting creative processes, concerns about its impact on societal values, the nature of creativity, and the authenticity of artistic output persist.

At the heart of this discourse is AI's potential as a collaborator in creative endeavors. Innovations like Wekinator [7] demonstrate how machine learning can serve as a *playful* building block to link various inputs to creative outcomes such as musical interfaces and expressive toys.

The challenge lies in establishing an equilibrium where AI acts as an empowering entity rather than an overpowering presence. Achieving this balance requires deep engagement with the technology but also tools for humanistic interpretation, and quite crucially, a healthy dose of playfulness — and the willingness to design not only systems to "solve problems" but also systems that are intrinsically interesting, expressive, and fun.

In summary, the convergence of technology and creativity in music is valued not solely for its ability to generate outputs but also for its capacity to involve artists, tool-builders, and researchers in a deeper contemplation of the meaning of human creativity in an era of AI automation.

5. CONCLUDING REMARKS

In this paper, we introduced ChAI as a set of interactive AI tools for the ChucK music programming language and discussed the ethos of using and teaching with ChAI in a Music and AI critical-making course, with the hope of giving students tools to engage creatively and critically with AI in music making as well as in the design of systems for creative expression.

The tools within ChAI and the course gave students the means to manipulate sound and interact with music in real-time, while also providing a tangible way to incorporate AI into their creative processes. This hands-on experience was crucial in demystifying AI in music, showcasing its practical application, and fostering a deeper appreciation for human creativity. As one Music and AI student wrote in her reflections about the course:

> My area of study within Computer Science is specifically on machine learning and artificial intelligence. I have previously taken many ML and

> AI classes and worked on research in these fields. In these past experiences, I was concerned with the theoretical and technical details of how to design and train an ML model. My goals were to achieve or surpass established state-of-the-art benchmarks. How could my ML models be better? How should I tune my model layers? Where could I get more data? These were the types of questions that I was always trying to ask and answer. However, taking this class has made me take a step back and consider why I am working in this field and what this is all for. Thinking about all this has led me to have more questions than answers, but one of the greatest takeaways I have is that AI is "astoundingly competent," but humans are astounding.
>
> After a quarter of seeing and experiencing the sheer range of creativity that has been produced from this class, I feel secure in knowing that AI is a long way away from ever achieving anything close to the projects that were created for this class. Everyone started from the same basic idea, yet they all went in wildly different and creative directions that I could never have imagined. The funny thing is that I've learned that AI doesn't necessarily have to ever achieve this...
> —J.

A central aspect of our course was the critical-making approach, which emphasized learning through designing practical applications using ChAI and thinking as broadly as possible about their implications. Additionally, the course sparked philosophical, ethical, and aesthetic discussions, prompting students to consider the broader implications of AI and music, including issues of authorship, accountability, authenticity, aesthetics (a lot of A-words!), expressiveness, and transparency. The range of student work demonstrated what people can make with ChAI in and perhaps beyond the classroom, especially when equipped with broader humanistic lenses, and given creative prompts that do not rush to ask "How can we solve the world's problems with AI" (good luck with that!) but rather "how can we create more interesting and playful experiences for ourselves and our communities using AI?".

6. ACKNOWLEDGMENTS

Thank Celeste and Andrew for help on the case studies, Terry and Eito for reviewing the paper, and the students of Music and AI: a Critical-Making Course at Stanford University in Winter Quarter 2023.

7. ETHICAL STANDARDS

ChAI has been developed with the support of CCRMA's departmental funding, curricular student research, and volunteer contributions. The authors are aware of no potential conflicts of interest. The students using early-stage ChAI were doing so to learn audio programming in a curricular context and were never treated as test subjects.

8. REFERENCES

[1] A. Agostinelli, T. I. Denk, Z. Borsos, J. Engel, M. Verzetti, A. Caillon, Q. Huang, A. Jansen, A. Roberts, M. Tagliasacchi, M. Sharifi, N. Zeghidour, and C. Frank. Musiclm: Generating music from text, 2023.

[2] J. Armitage, N. Privato, V. Shepardson, and C. B. Gutierrez. Explainable AI in music performance: Case studies from live coding and sound spatialisation. In *XAI in Action: Past, Present, and Future Applications*, 2023.

[3] L. E. Baum and T. Petrie. Statistical inference for probabilistic functions of finite state markov chains. *The annals of mathematical statistics*, 37(6):1554–1563, 1966.

[4] C. Betancur Gutierrez. Ai in live coding environments: Pandora's Dream. *AIMC 2023*, aug 29 2023. https://aimc2023.pubpub.org/pub/9lqurr1x.

[5] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[6] T. M. Cover and P. E. Hart. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory*, 13(1):21–27, 1967.

[7] R. Fiebrink and P. Cook. The wekinator: A system for real-time, interactive machine learning in music. *Proceedings of The Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)*, 01 2010.

[8] R. Fiebrink, G. Wang, and P. Cook. Foundations for on-the-fly learning in the chuck programming language. 08 2008.

[9] R. Fiebrink, G. Wang, and P. Cook. Support for mir prototyping and real-time applications in the chuck programming language. pages 153–158, 09 2008.

[10] R. A. Fiebrink. *Real-time human interaction with supervised learning algorithms for music composition and performance*. PhD thesis, USA, 2011. AAI3445567.

[11] K. P. F.R.S. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

[12] D. Grba. Deep else: A critical framework for ai art. *Digital*, 2(1):1–32, 2022.

[13] S. Haykin. *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1994.

[14] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space, 2013.

[15] F. Morreale. Where does the buck stop? ethical and political issues with ai in music creation. 2021.

[16] S. S. Rautaray and A. Agrawal. Design of gesture recognition system for dynamic user interface. In *2012 IEEE International Conference on Technology Enhanced Education (ICTEE)*, pages 1–6, 2012.

[17] A. Roberts, J. Engel, C. Raffel, C. Hawthorne, and D. Eck. A hierarchical latent vector model for learning long-term structure in music. In *International Conference on Machine Learning (ICML)*, 2018.

[18] G. Vigliensoni, P. Perry, and R. Fiebrink. A small-data mindset for generative ai creative work, 2022.

[19] G. Wang. Humans in the loop: The design of interactive ai systems. `https://artful.design/hai/`, 2019. Accessed: [insert date you accessed the website].

[20] G. Wang, P. R. Cook, and S. Salazar. ChucK: A Strongly Timed Computer Music Language. *Computer Music Journal*, 39(4):10–29, 12 2015.