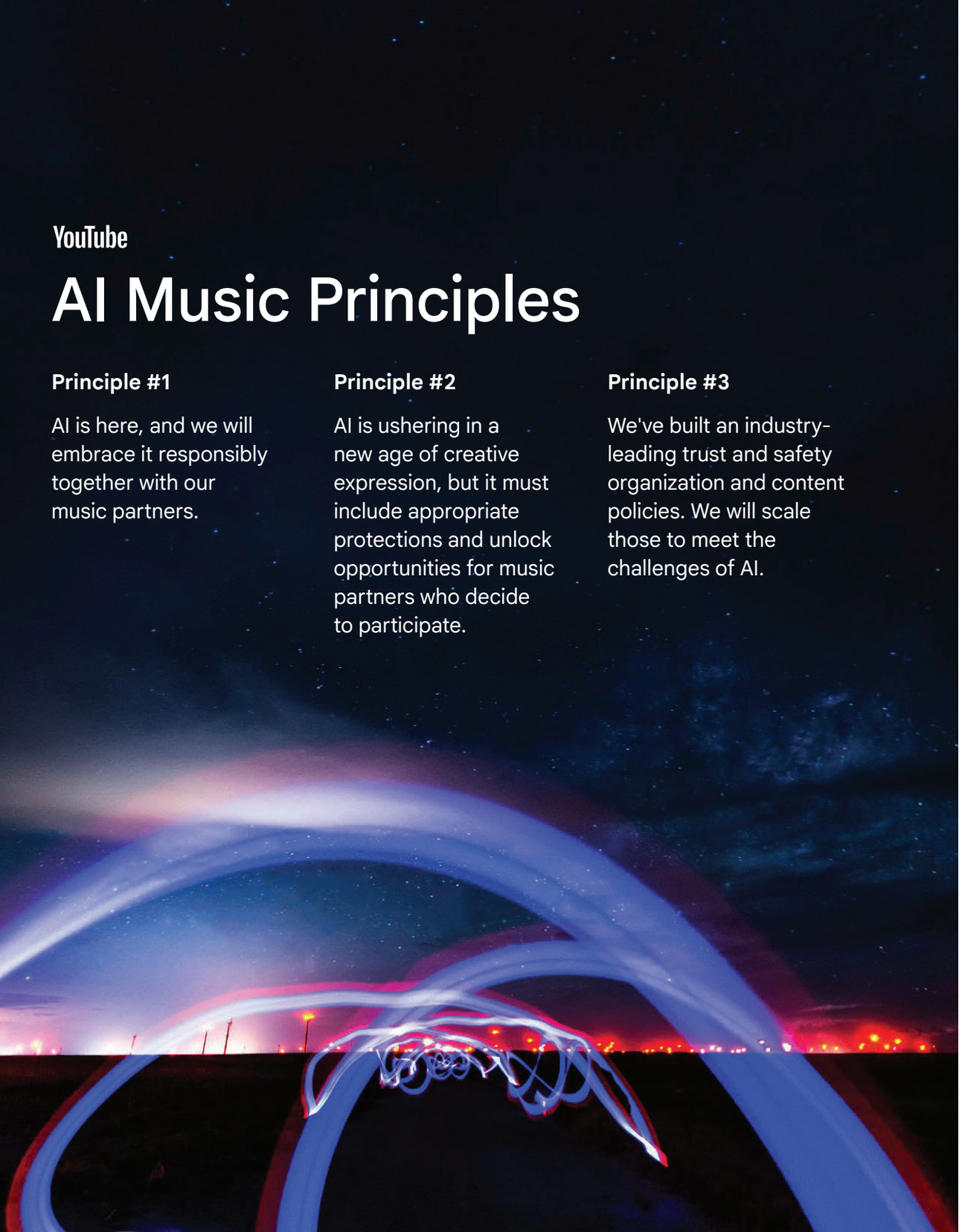

Insights from the intersection
of music and technology
SOUNDS OF
TOMORROW
Technology & Society
Google

SOUNDS OF TOMORROW

A collaborative research report with input from
artists, songwriters, and producers,
on the possibilities of generative AI and music

Google

Table of Contents



Music, AI, and the Human Spirit: Exploring a New Creative Era

Artists are always looking for new ways to get inspired, new ways to create great music and share their work, new ways for their songs to connect with people. Technology is and has always been a close collaborator in this.

For almost two decades, connecting artists, songwriters and producers with users worldwide has been core to our work. Making YouTube the best place for every artist and music fan means constantly strengthening those connections, cultivating deep relationships with our partners (one of our greatest superpowers) and creating best-in-class tools and experiences for them to tap into.

The arrival of generative AI has opened up new possibilities to explore and challenges to overcome. Over the past year, we've been hard at work building a foundation for this new era, one rooted in bold and responsible innovation. And together with our partners and the creatives themselves, we've begun exploring the potential of AI to expand the world of music.

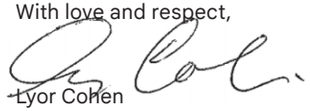
What stands true is that artists, songwriters and producers remain at the heart of music creation. What exhilarates me is where they lead us creatively as the aperture of sound widens.

Last summer we unveiled our Music AI principles and AI Music Incubator. Since then, we've brought dozens of music creators across the world, genres and backgrounds into the studio, testing early AI ideas, and getting their feedback on what works for their creative process. We want to build what's best for them and our incredible community of users. So far, it's been an invaluable learning experience.

This report dives into insights gathered from these sessions, covering learnings about generative music technology, design considerations for creative AI tools, and explores the potential risks and opportunities upon us. Ultimately, our goal is for our collaborative approach of building, testing, and learning to become a blueprint for how music generation tools are developed and deployed more broadly across the music and technology ecosystems.

I want the music community to have every tool they need to share their genius with the world. Our incubator and early experiments are key to building that foundation, figuring out what AI can mean for artists, songwriters, producers, and the whole industry. We don't have all the answers yet, but together, we're looking forward and making moves to uncover what's next.

With love and respect,



Lyor Cohen
Global Head of Music, YouTube/Google

The hands of jazz composer Maxwell Powers in 1948 playing the 'keys' overlayed with a more recent photo of hands using another sort of keyboard. (George Pickow/Getty)

Creating Together: The YouTube Music AI Incubator

It is critical to develop advanced technologies in collaboration with society. For this reason, starting in 2023 teams from YouTube, Google DeepMind, and Technology & Society have worked together on a program called the YouTube Music AI Incubator. The teams have spent hundreds of hours with artists, songwriters, producers and industry professionals across 17 countries.

Over the course of the past year, the Incubator sessions have reached a wide variety of music professionals at the top of their game. Participants included global superstar Anitta, internationally renowned and revered songwriter, producer and entrepreneur Björn Ulvaeus, innovative genre-defying artist d4vd, acclaimed musician, composer and producer Don Was, Colombian sensation Juanes, hitmaking producer Louis Bell, visionary composer Max Richter, influential songwriter and producer Rodney Jerkins, iconic singer-songwriter Rosanne Cash, 3x Grammy award-winning songwriter and producer, Ryan Tedder of OneRepublic, acclaimed rapper, multi platinum musician, entrepreneur and philanthropist Yo Gotti, and the estate of American musical icon Frank Sinatra, amongst others from around the world.

> Can AI advance music in a positive way? The verdict was: yes, it can, and we need to do it responsibly, together.

During the YouTube Music AI Incubator sessions, Googlers interviewed these professional musicians, artists, songwriters and producers about music and AI. The team shared in-progress technologies with participants and wondered together: Can AI advance music in a positive way? The verdict was: yes, it can, and it must be explored responsibly, together. By the time each research session ended, the groups were imagining new musical creations at the speed of sound.

Incubating innovation

In a typical incubator session, artists, songwriters, producers, and their teams are invited to explore experimental prototypes from Google. These sessions can take place in a studio, at someone's home, or a Google office — whatever works for the participants. The group sometimes spends hours co-creating: generating new music, extending existing tracks, blending sounds, and more. For example, one early prototype let a participant transform live vocals into a horn section. These research sessions are private so the participants can feel comfortable sharing their honest feedback.

Consistently, participants asked about authorship, ownership, future use cases, and how this technology and the music produced could be shared. Through these conversations, the YouTube and Google teams gained substantial insights about the participants' creative processes, experiences, and needs to understand rights protection and monetization possibilities, and together, they pushed into the unknown spaces of this new technology.

Impressions from sessions

At the beginning of the sessions, many participants were eager to learn about the capabilities of the models and wondered how the technology could positively impact their creative workflows. Most also expressed concerns, seeking assurances that human creativity would be necessary and preserved for the future. However, as they experimented with the prototypes and tapped into their ingenuity and expertise, many saw the potential to extend their creativity and save time.

Hearing how these participants imagined the future was magical. They thought about emerging artists and younger musicians, and their own roles in preserving and advocating for their craft. Different participants had different definitions of musical skill — to some, musicianship is being able to play a guitar; to others, it means working a sequencer — and differing opinions on technology in art. Yet even the most skeptical participants asked for more and new capabilities or safeguards that would help them in their work.

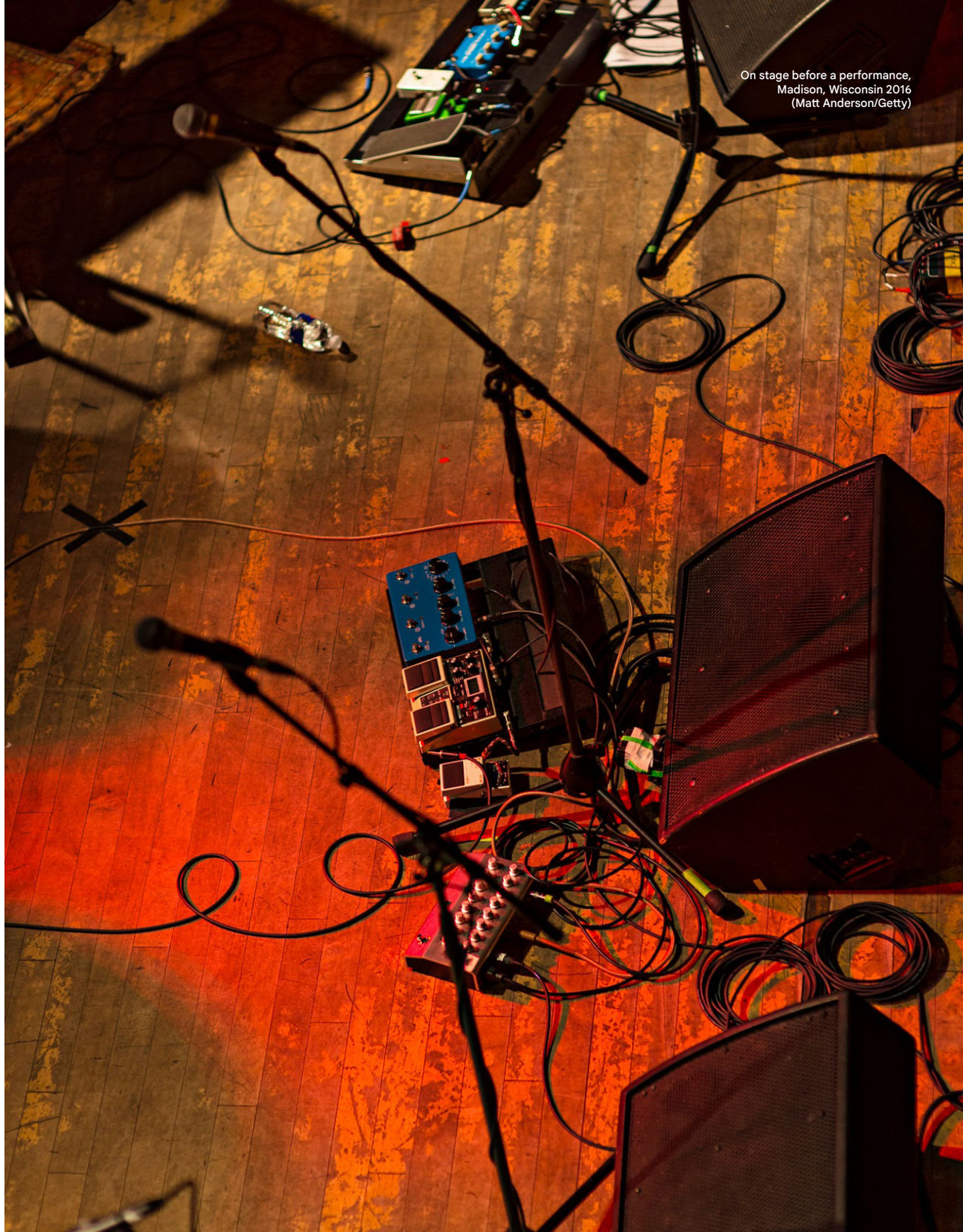
Listening to learn

Building AI responsibly requires groups from all disciplines, including design and research science, to align AI capabilities to benefit artists and communities. In response to the incubator sessions, teams from Google DeepMind, YouTube, and Technology & Society worked together to integrate feedback from participants directly into ongoing research, design, engineering work. The Music AI Sandbox has evolved significantly as a result of the insights from, and conversations with, artists met through the incubator program.

> We sometimes spend hours co-creating: generating new music, extending existing tracks, blending sounds, and more.

Beyond creating collaborative research experiences to advance the technology, these sessions provided an open forum for candor, offering the opportunity for participants of the Incubator to voice critical questions and concerns. The open space for dialogue about what the future can hold tended to inspire: many attendees expressed cautious excitement for the future of the tech, and enthusiasm for how they might use it to advance their craft. These insights and critical questions provide the backbone of this report, and motivation to continue the program, and others like it at Google.

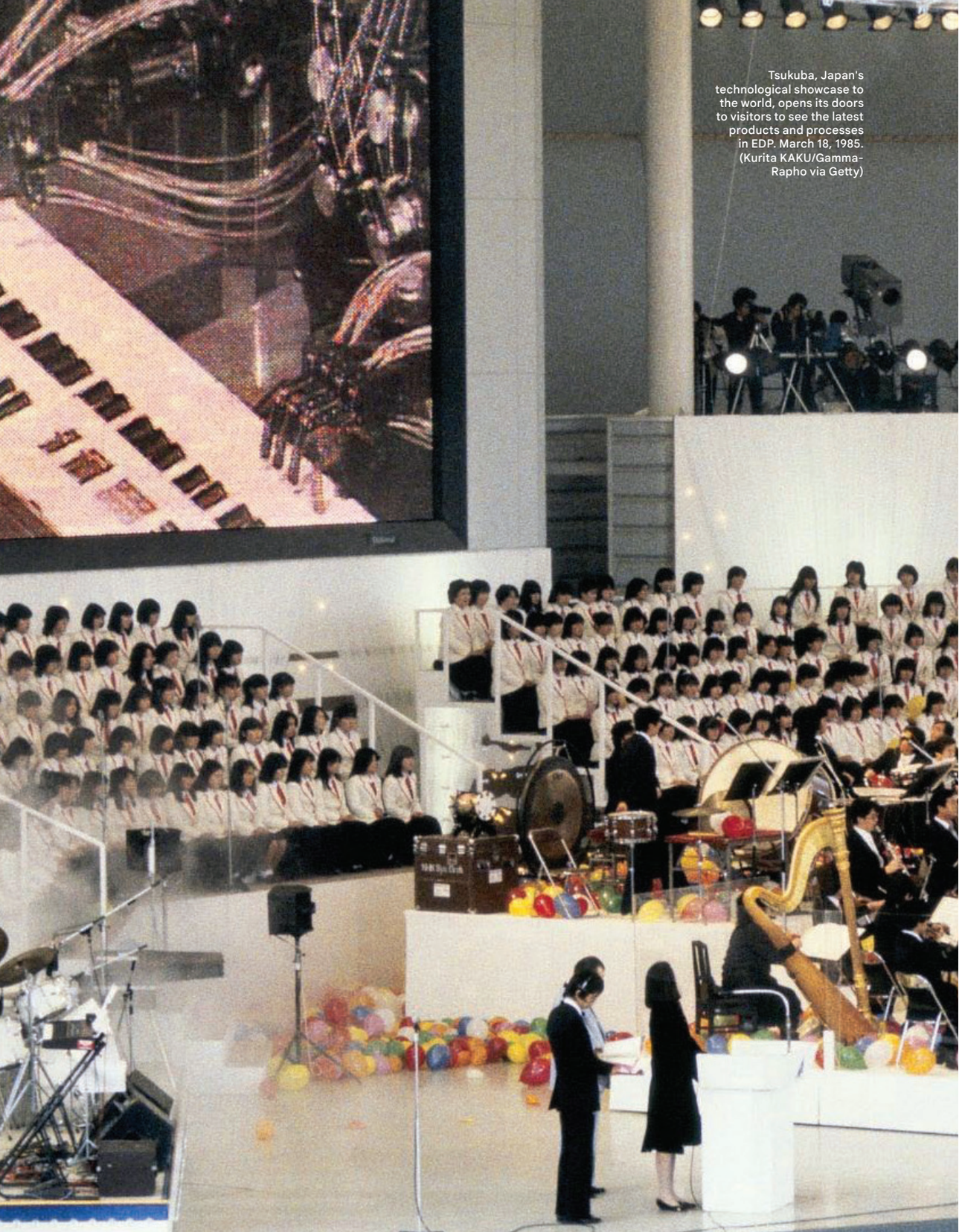
In sum, the program has reinforced the value of co-creation as a methodology for shaping research investments, but more importantly, it is an invaluable method for evolving culture at the intersection of technology and society: one of collaborative learning, trust, and respect.



On stage before a performance, Madison, Wisconsin 2016 (Matt Anderson/Getty)

Imagine if generated music could meet and match the mood

The party host scans the friendly faces, packed but happy in the hot, narrow kitchen. Reaching behind him, he dims the lights, and starts to usher people out, pressing palms into sweaty backs. "To the living room," he says over the sound of upbeat folk music. He twists a dial on the generative stereo and the beat begins to pick up. It's a new take on the same song but faster, and with a house beat. People bounce on their heels and ice clinks in drinks, both music and party goers phasing into a new direction towards a song they love in a style they've never heard. The low light catches on eyes, cheeks, lips, and teeth. As the group shifts towards the couch and chairs, the music shifts too, settling into lower tempo and volume — another refresh of the signature style of the artists of the evening, but still recognizable, and familiar. Low rumbles of talk picks up in the lulls, and the party flows on.



Tsukuba, Japan's technological showcase to the world, opens its doors to visitors to see the latest products and processes in EDP. March 18, 1985. (Kurita KAKU/Gamma-Rapho via Getty)

Discovering New Musical Possibilities

For nearly a decade, Google has been collaborating with musical creatives of all kinds — artists, songwriters, and producers. Collaboration programs like Magenta, Artists + Machine Intelligence, and Google Art & Culture, have shown how partnering across disciplines helps creatives explore new opportunities and invent new forms of expression. The YouTube Music AI Incubator program, a project from Google DeepMind, YouTube, and Technology & Society, has been another fruitful engagement, inspiring new technology and research directions.

> The Music AI Sandbox is an experimental prototype designed to supercharge creativity within a musician's existing workflow.

Composing with prompts in Music AI Sandbox

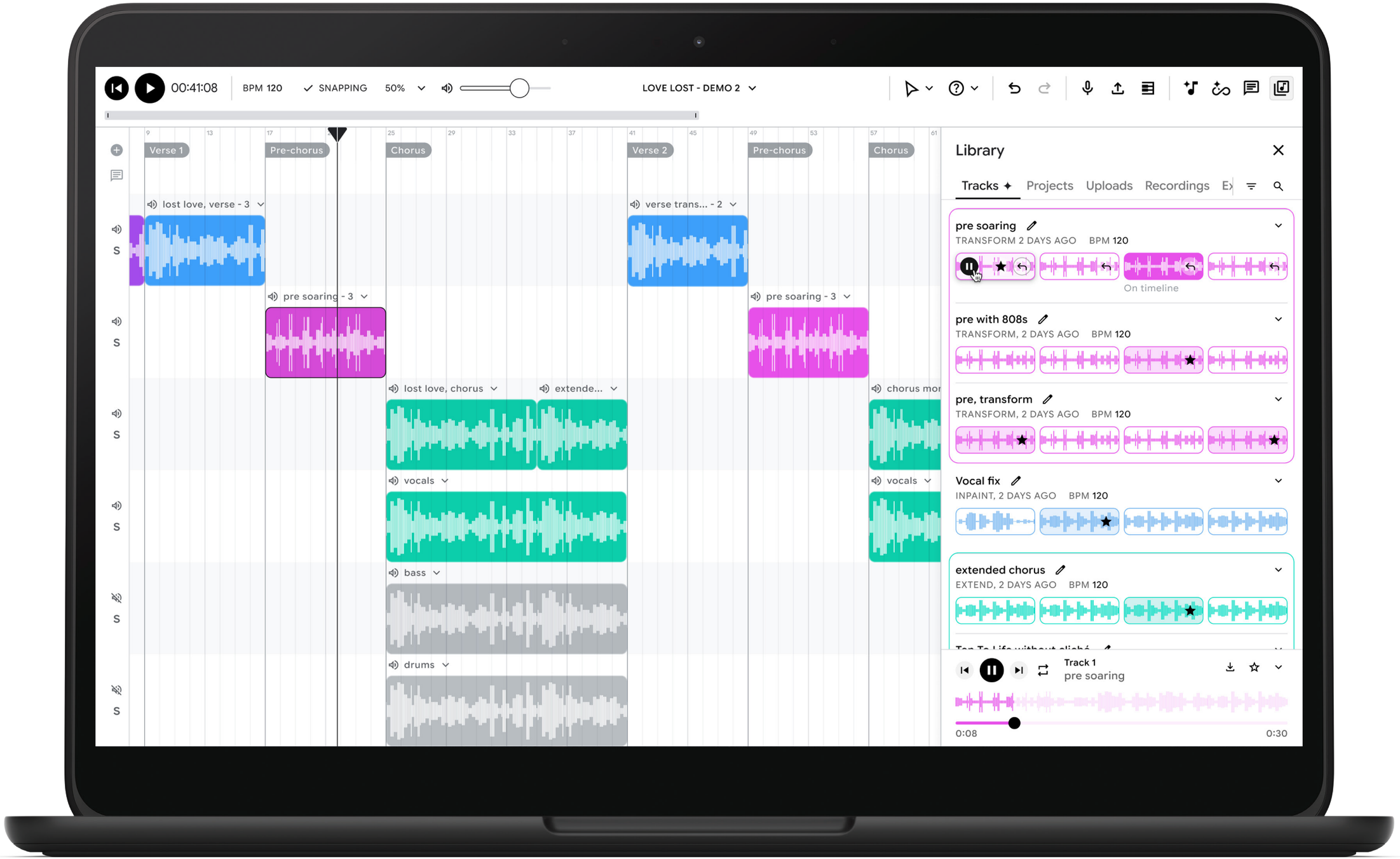
Starting in 2023, Google has been developing a suite of experimental tools for music professionals, called the Music AI Sandbox. These prototypes are built on state-of-the-art AI technology from Google DeepMind that excels at generating high-quality music with instrumentals and vocals, transforming sound or continuing a piece of music in new directions, and giving users more nuanced control of the output's style and performance.

The Music AI Sandbox is designed to supercharge creativity within a musician's existing workflow. In research sessions that inspired the development of the Sandbox, as part of the YouTube Music AI Incubator, professional artists, musicians and producers pushed the boundaries of generative music technologies. By using these tools to create, compose, and explore sounds, participants sparked new creative ideas and music clips they imagined might be useful as part of a composition. Many participants mentioned that tools like the ones they tried could speed up the process for creating new music.

Tech plus taste, from artists & musicians

When participants in the YouTube Music AI Incubator sessions used the prototypes, there came a moment when the artist, songwriter, or producer started to see the capabilities of the models, and get a handle on how they might use them in their own workflows.

The Music AI Sandbox is designed to supercharge creativity within a musician's existing workflow.



Each participant brought with them instincts honed over years of feeling, understanding, and precisely channeling emotions through music. As in any well-crafted communication, structure is key: there is a hard-earned craft to arranging musical elements into a song that hits right in the heart. Expert musicians hear and control how a song is structured: subtle shifts in phrasing, carefully arranged timbres, precisely timed silences. The Music AI Sandbox surfaces important musical patterns, such as swing, tonality, and vocal melody, so artists can manipulate them to achieve their intended emotional impact.

Beyond producing what artists, songwriters, and producers expect, the tools can also invite new ways to compose and arrange that aren't in line with traditional musical taste or training. This divergent, generated output — though 'weird' and 'wrong' according to participants — sometimes showed a new creative path they may not have found on their own. The generated outputs helped participants move beyond their rule book, to pursue experimentation and surprising new works knowing they could learn more quickly or fail faster.

For professional musicians, this element of randomness often made unexpected connections. New input styles uncovered maps to new musical worlds: crossing genre boundaries and time traveling, blending instruments from other eras, geographies and cultures, leading to fresh collaborations between musicianship of past and present, making the sounds of tomorrow.

This divergent, inhuman output — though 'weird' and 'wrong' according to participants — sometimes showed a new creative path they may not have found on their own.

Respectful & responsible development

Today, society is seeing advances in generative AI that technologists only dreamed of a decade ago. In the research sessions, the musical fluency of the music AI models often surprised participants. The technologies also gave many participants pause, and cued important questions about how best to respect and protect the contributions of artists, songwriters, and producers.

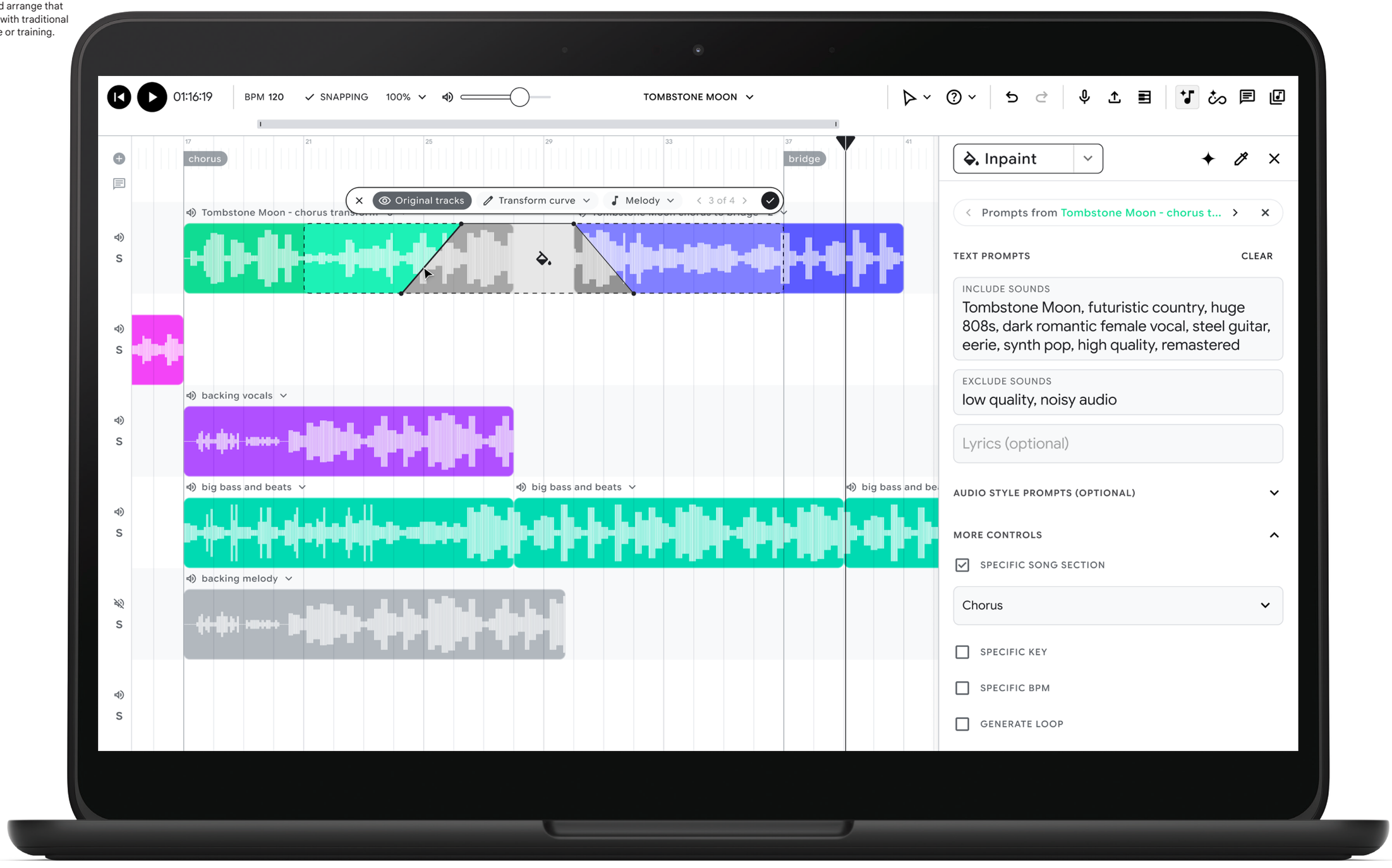
Identifying and managing generated content is a deep area of research and technical investment at Google, and we've made significant advancements over the years to support the interests of copyright holders, the creative community, and music fans.

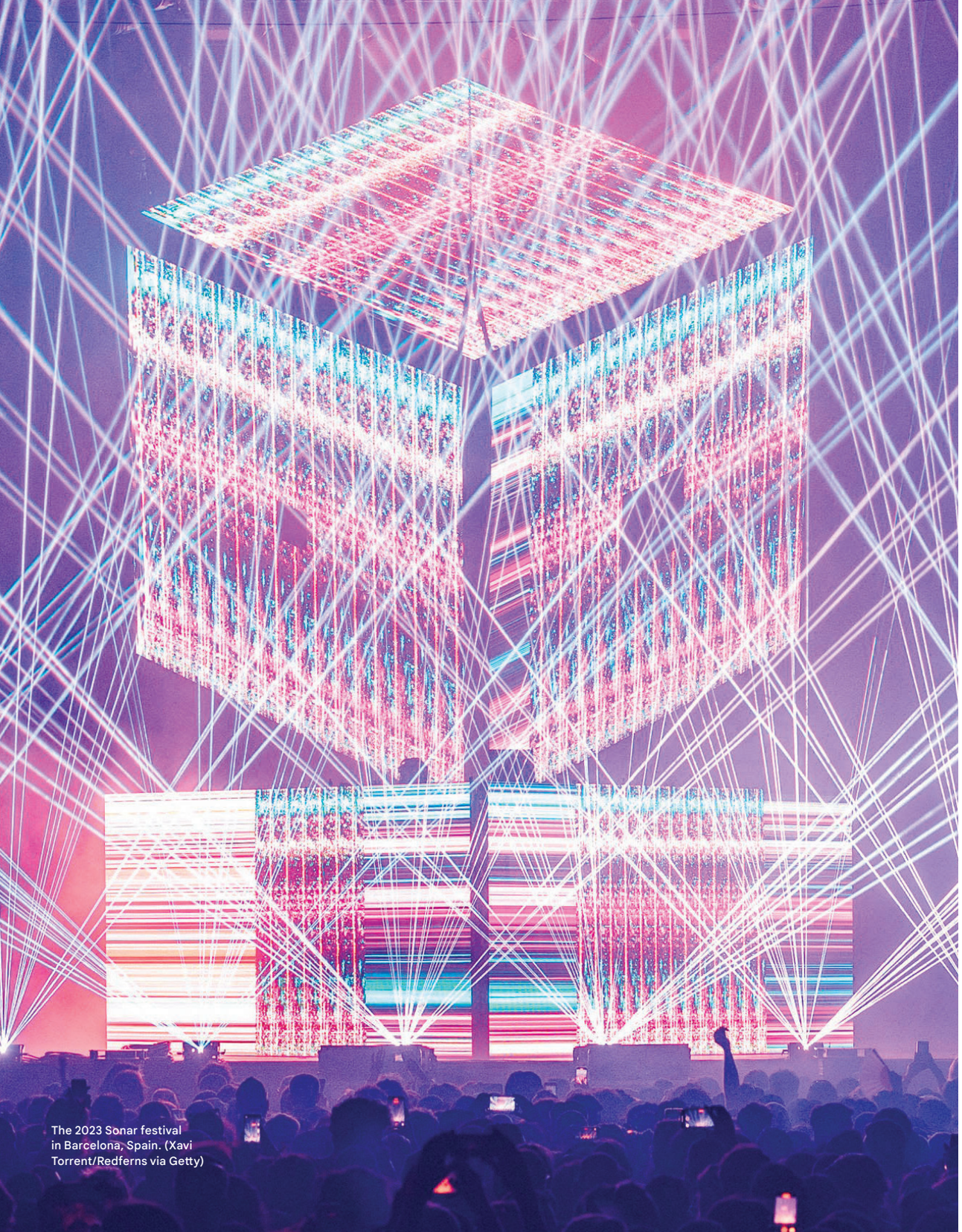
Outside of the Music AI Sandbox, YouTube, and GoogleDeepMind have already launched a few key technologies for managing generated media across Google platforms. For example, the SynthID toolkit from Google DeepMind watermarks AI-generated content. These tools embed digital watermarks directly into AI-generated images, audio, text or video. Meanwhile, Content ID, YouTube's best-in-class rights management technology, ensures rights holders get paid for use of their content and has generated billions of dollars for the music industry over the years.

Google is building for the future, and committed to delivering tools that support and empower creativity for all. This is made possible by putting the artist first: supporting those who create the music we all live by.

Hegeler & Ehlers, Pianino with automatic self-playing mechanism, 1850-1859 (Sepia Times / Universal Images Group via Getty)

Beyond producing what artists, songwriters, and producers expect, the tools can also invite new ways to compose and arrange that aren't in line with traditional musical taste or training.





The 2023 Sonar festival in Barcelona, Spain. (Xavi Torrent/Redferns via Getty)

Key Learnings: 8 Insights for Music AI Makers

Participants in the YouTube Music AI Incubator program asked and answered a broad range of questions about AI, music and the tools in development in the Music AI Sandbox. Their thoughtful responses, including new ideas and critical issues, are summarized below.

Hearing the music reduces skepticism

Many participants started out doubting that they could use AI to produce music with real depth, but their opinion changed when the music started to play. Initial amazement was followed by emphasis that we need to continue to build with and for the creative community.

Music models can expand imagination

Participants valued experimenting with AI to explore new and unconventional sounds. One participant talked about how musical creativity can be blocked by knowledge of conventional music production, such as being stuck in what's musically correct. In contrast, the unpredictable nature of music AI sets it apart from other software used in music production, with this unpredictability seen not as a flaw, but as a strength.

Multimodal inputs unlock musical expression

Steering music with natural language initially felt limiting, even though novel prompts and word combinations resulted in useful outputs. Many participants were looking to hum, tap, swipe, merge, and compose without words. They also wanted affordances to extract elements like beats or vocals, match lyrics to melody, and pull sheet music.

AI models make it easy to explore complex sounds

Artists found that sounds that would be difficult to create, like crowds chanting or choirs singing, were useful to generate as placeholders to get a feel for how a piece might work. This helped them sketch out new ideas to see where to invest in new musical partnerships.

Real-time interaction makes generating fun

With some of the experiments, it's possible to change the generative music in real-time, and shift tempo, mood, or genre by adding to the existing prompts. This let participants 'jam' with generative music, creating sections of music live which they then wanted to use as part of a larger composition. Still, there's more work to do to show musical structures in the tools so artists and songwriters can control them better.

Different disciplines require different designs

Artists, songwriters, and producers each have unique needs that need to be addressed by the design of the AI experimental tools. While producers loved being able to craft the sound of a perfect drum loop, songwriters were more interested in the relationships between lyrics, vocals and rhythm. Participants saw potential in the experiments as an inspiration tool, a song starter, a song finisher, or a pitch tool. However, every participant desired controls that kept them in the driver's seat so they could leverage their specific skills.

Monetization, attribution, and control are critical concerns

Virtually all of the incubator participants felt monetization, attribution, and control for generated music were important topics to address. For example some asked, who owns music created with music AI tools? How do songwriters participate? Many pondered answers, suggesting scenarios on how these issues could be tackled. To be clear, for YouTube, these are not an 'if' but a 'when' and 'how'. These topics are at the heart of YouTube's second AI Music Principle and front and center in their generative AI work.

Education and accessibility are keys to empowerment

Musicians imagined the experiments could be useful for beginners or emerging producers with limited technical knowledge or resources. They imagined a collaborative partner or co-writer to spark ideas, help complete their works by generating bars, or improve their sound with tonal tweaks. One participant imagined how a younger version of themselves could have gotten further along creatively, faster, with the prototypes they tried.

A series of cymatic patterns which are often created by imaging sound through liquids

Imagine if you could play AI with your movements

The dancer steps and turns, filling the mirrored studio with music matched to their movements. Sneaker squeaks catch the beat. The dancer's arms play the air like a harp, and a pounding baseline thunders. Turning to the sound table, they shake their head, and the sound of wind chimes flutters through the room as sweat drops fly. Laughing, the sound engineer turns the generation off. "Not quite right," she agrees. She types in a few adjustments: shimmer, foghorn, ragtime, MRP. The dancer stretches a foot out and flexes their toe on the sensor mat, testing the changes by varying the pressure and speed. "Let's go again." The dancer nods, and, and with a thumbs up from the engineer, leaps into the sound.

Crowd surfing during a music festival on June 21, 1993 in New York City.
(Bill Tompkins/Getty)

Safeguards for Sharing: 3 Key Questions for Getting it Right

During the research sessions, participants in the YouTube Music AI Incubator program explored music AI tools and prototypes in development. The Google team collected their feedback and questions to understand how best to support the needs of artists, songwriters, and producers — and their fans — both with the tools and on Google platforms. Participants described how AI could influence the way they think about music technology, their craft, and what they bring to the table as creatives. They also asked questions about future use cases; about where generated music could be shared, how, and with whom.

How might artists and fans make music together?

Song snippets can inspire remixes or new ways to sample, and participants imagined new ways to build off of their past work. However, there are situations where artists don't want songs generated in their voice, such as if the output is silly, obscene, or misrepresents their viewpoints. Also, artists simply might not want someone else using generated synthetic outputs of their voice, regardless of how it is used. To prevent this, YouTube is testing new ways to allow creators to add fresh and unique music to their content, balanced with strong guardrails. For example, by developing tools that allow partners to detect and manage AI-generated music content that mimics an artist's singing voice.

How might listeners identify generated content?

Participants wondered if generated tracks or clips could take attention from an artist's original album, and wanted a clear distinction for AI-generated content. Since the beginning of the incubator program, YouTube has launched prominent AI labeling – watermarks,

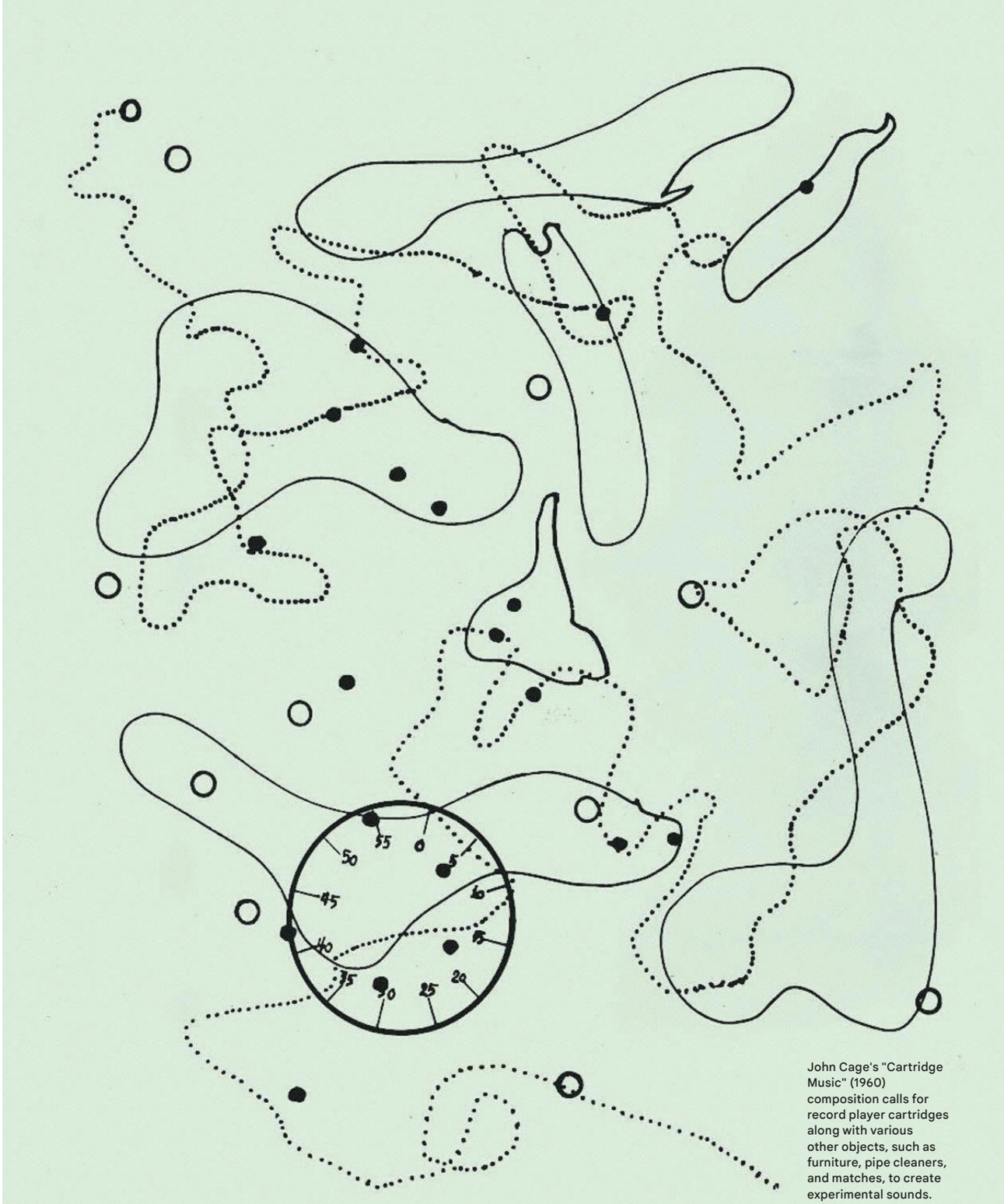
badges, and hashtags – across the platform. This year, YouTube announced a new tool in Creator Studio requiring creators to disclose to viewers when realistic content – content a viewer could easily mistake for a real person, place, scene, or event – is made with altered or synthetic media, including generative AI. To support this, YouTube established policies to encourage users to tag AI-generated or manipulated content as such when they upload it. To enforce this, if a creator consistently fails to identify generated content as such, they'll be subject to content removal, suspension from the YouTube Partner Program, or other penalties.

How might platforms protect the integrity of artists and musicians?

In interviews, artists were frank. One participant shared that her biggest fear was that artist's voices would be stolen and used without their permission for disparagement of others. Google has been researching mitigations for fears like this, investing in both counterabuse technologies and teams of highly trained moderators to keep platforms safe. The same technology that Google research groups use for information integrity can be leveraged to solve problems of attribution and monetization for music. Teams at Google DeepMind, YouTube, and Technology & Society are already imagining and researching how to support artists by tying generated songs back to the original work.

> However, there are situations where artists don't want songs generated in their voice, such as if the output is silly, obscene, or misrepresents their viewpoints.

Collaborating so deeply with artists, songwriters, and producers has helped turn these questions and concerns into teams working to build the actual capabilities. Teams at Google have and will continue to action as much feedback as possible into their work — from new research to new features. These questions have also deepened YouTube's commitment to the YouTube AI Music Principles, which will continue to guide work on this technology in the future.



John Cage's "Cartridge Music" (1960) composition calls for record player cartridges along with various other objects, such as furniture, pipe cleaners, and matches, to create experimental sounds. (Photo 12/Getty)

Imagine if you could use AI for collaboration, sharing, and learning

Angling her head to hear better, she squeezes the strings against the love-aged fretboard of the storage-scuffed ukulele. She sits up, knuckles curved, fingertips white, pinkies out — tongue too. The phone facing her beeps: placement looks good. She strums the open strings, and listens as the app generates feedback and reminds her to strum with her thumb, hand facing down. The kid adjusts, and strums again, and the app generates harmony. She listens as the app describes a new chord and finger placement, building off the previous. Bolder now, the ukulele's sweet soprano rockets around the little room, sending an eavesdropping crow squawking past the window. The kid plays the chord series again, and the phone burbles a melody alongside, showing sheet music as they go. Save this track? "Save as 'Grandpa's Song'."

Developed at the Google Arts & Culture Lab, Living Archive, a tool for choreography powered by machine learning, generates original movement inspired by Studio Wayne McGregor's 25-year archive, creating a live dialogue between dancers and his body of work. (Wayne McGregor's Sulphur 16, 1999, Photo: Ravi Deepres)

Music, AI, and the Human Spirit: Exploring a New Creative Era

Brilliant minds from Pythagoras to Bach have observed how close in spirit music and math are. Once you learn their intricate patterns and structures, you can say something fundamental about both the world around us, and through some greater alchemy, something fundamental about the experience of being human.

AI in the field of music is the ultimate marriage of music and math. Until recently, the complex and densely layered nature of music – its melodies, rhythms, and vocals – have been difficult for AI systems to understand. That's changing.

For years, Google has worked with musicians exploring new ways to create compelling art and music with machine learning. Dating back to our Magenta project in 2016, we released an array of AI tools over the ensuing half decade. In November 2023, we announced Lyria, our most advanced AI music generation model to-date, along with a suite of music AI tools, and an experiment in collaboration with YouTube called Dream Track, which allowed a select group of creators to generate entirely new songs in the style of a few participating artists, through a simple text prompt.

There's a long history of technology and artistry pushing each-other ever-forward. In that tradition, we want artists and creatives to have at their disposal powerful new tools for bringing their ideas to life and sharing them with the world.

Until recently, the complex and densely layered nature of music – its melodies, rhythms, and vocals – have been difficult for AI systems to understand. That's changing.

Guided by YouTube's principles for partnering with the music industry on AI technology, we're working alongside artists, songwriters and producers to build the future of music, today. When we talk to these artists, we often hear their particular enthusiasm for music education and the need to make sure the next generation of artists have opportunities and tools they need. AI could help anyone learn to be more musically expressive. It could be a tutor, listening to your work and suggesting areas for improvement, a collaborator, or a source of inspiration.

Of course, building this future requires an equally robust and parallel focus on safety, responsibility and transparency. Building on previous work, we'll continue to research and innovate in technologies that enable responsible deployment like watermarking. And we'll continue engaging artists, songwriters, the music industry, and wider creative community to set the standard for the responsible development and deployment.

Demis Hassabis, CEO
Google DeepMind

Contributors

Arathi Sethumadhavan
Blake Cunningham
Chris Reardon
Courtney Marr
Demis Hassabis
Denise Su
Doug Eck
Grace Nicklin
Hema Manickavasagam
James Manyika
Jenna Rubenstein
Jessi Liang
Jody Medich
John Guilyard
Karen Chappell Arellano
Lauren Resnick
Lyor Cohen
Mariana De Felice
Matt Miller
Matthew McClernon
Michelle Chung
Mira Lane
Priya Jhakra
Roxanne Pinto
Suzana Carlos
Tom Hume
Vivien Lewit

Google Technology & Society DeepMind YouTube

The Grand Opera, situated in Harbin, a UNESCO-listed "City of Music"where China's first ever orchestra was established. (Hufton+Crow/View Pictures/Universal Images Group via Getty)