



City Research Online

City St George's, University of London

Citation: Einbond, A. (2027). Beyond Automatic Genre Classification: Creative music information retrieval. In: Born, G. & Brackett, D. (Eds.), Music and Genre: New Directions. . Duke University Press. ISBN 9781478035183

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/38504/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Beyond Automatic Genre Classification: Creative music information retrieval

Aaron Einbond

Abstract

Since its beginnings in the late 1990s, the field of Music Information Retrieval (MIR) has led to technological and cultural impacts that set the stage for the more recent burgeoning of Artificial Intelligence (AI) technologies for music. Together, MIR and AI have reached ever-wider networks of listeners, users, developers, and marketers through applications like recommendation systems automatic genre classification. In contrast, how can musicians use similar technologies creatively while subverting and expanding their possibilities? This chapter introduces a group of composers and sound artists working in the period around 2008-2018 to exploit new resources for artistic creation offered by MIR. Unlike commercial MIR, their work resists the reification of genre categories by playing with, participating in, and forging novel micro-genres. Yet as MIR has permeated creation, it has demanded reassessments of definitions of creativity itself, offering perspectives that could also apply to more recent artistic creation with AI.

Keywords: Music information retrieval; machine learning; audio descriptors; composition; field recording

Before the recent rise of interest in Artificial Intelligence (AI) technologies for music, a more subtle, yet no less transformative, shift occurred in the first decades of the 21st century in the burgeoning field of music information retrieval (MIR). MIR, at that time, offered dual opportunities for music and genre: either the unprecedented mechanization of automatic genre classification, or alternatively, powerful means for creators to differentiate and personalize new practices. The former approach has been commercialized increasingly to reify genre, one of the themes of this volume, while the latter blurred existing genre boundaries to propose individual, difficult-to-classify artworks. At the same time, these works pointed to the new possibilities of an emergent creative genre of their own, predating yet significantly foreshadowing subsequent currents of AI-influenced artwork. In this way, the work discussed in this chapter exemplifies a tension in the theory of genre itself between the recognition of unique artistic statements versus the claim that no work exists independent of genre (Frow 2005, 28; Altman 1999, 12).

Beginning in the late 1990s, MIR explored the uses of audio features to organize information about digital sound (Fingerhut 2004). As one of the key catalysts of Artificial Intelligence (AI) technologies, MIR gained a place in scholarly and artistic research as well as commercial applications as it increasingly powered music software, websites, and apps. MIR and AI have changed the way music is classified, marketed, and experienced by nearly all digital listeners. This chapter suggests alternatives to commercial MIR by introducing a growing number of musicians—composers, improvisers, performers, sound artists, and programmers—who took advantage of the early decades of MIR, around 2008-18, to exploit it as a creative resource. In contrast to proprietary commercial software, these musicians took advantage of free and open-source tools, often developed by the artists themselves, permeating the boundaries between computer programming and other forms of music making. Still other artists, while not using MIR software directly, responded to the culture of music informatics creatively and played with its conventions and artefacts in their work. Whichever their technical method, these musicians revealed MIR to be a rich source of musical nourishment.

Like these other musicians, as a composer I became interested in how MIR could be exploited to artistic ends. In contrast to developers of MIR technologies in commercial applications, I have been more interested in situations where automatic classification fails: where musical materials do not fall into predictable groupings or where the audio content itself does not submit to easy analysis. This can lead to situations where the MIR technologies that power automatic genre classification instead can point to unexpected associations, fusions, and hierarchies. Such a creative *détournement* of music technology has occurred many times before and since MIR, and it also points to an important co-dependence of artistic exploration and genre. Artists find fertile territory at genre borders, and by repurposing the MIR technologies of genre classification, they could harness a pointed tool to critique existing genre relationships and cultivate generic ambiguity, multivalence, and inclusivity. Some of these musicians have described their practices as contributing to an “expanded field” (Ciciliani 2017) or “new discipline” (Walshe 2016) at the boundaries between composition, improvisation, performance, installation, theater, and software development. As such, they not only could defy received genre categories, but also participate in formulating new ones. These practices support a view of genre in constant flux, influenced by each individual artistic contribution.

Searchability

While MIR has taken many forms, they share the use of audio features and metadata to organize sound. An audio feature, or descriptor, is any characteristic of audio such as pitch, loudness, brilliance, or higher-level metadata attributed to these. Audio can also be tagged and searched by textual metadata, subjectively assigned by listeners. This was the approach taken by the early audio streaming service Pandora, whose website states “our team of trained musicologists has been listening to music across all genres and decades [...] creating a unique knowledge base identifying hundreds of musical details for each song.”¹ This is musicology on a grand scale indeed! Once this data has been assigned, including descriptors for genre and many other features, it can be used to organize, search, retrieve and recommend audio.

Another approach is to use machine listening, often in the form of proprietary software, to extract audio features automatically. Combined with ever more widespread AI techniques, specifically machine learning (ML), audio descriptors introduce the possibility of automatic genre classification. This is the approach that was pioneered by The EchoNest, founded in 2005 and acquired in 2014 by Spotify, allowing its algorithms to reach an ever-wider network of listeners, users, developers, and marketers. Before 2014, EchoNest was praised as “possibly the hottest music data company around right now” (TechCrunch 2012) who “pull the playlist picking strings” (Engadget 2011). Their acquisition by Spotify was celebrated as “a landmark moment for the MIR community” by members of the International Society for Music Information Retrieval (ISMIR).² Yet while Spotify has optimistically advertised “the right music [...] for every moment” to “soundtrack your life,”³ this populist message may belie increasing systematization and homogenization that reduces individual choice and creativity. As a case in point, before its acquisition, the EchoNest API (Application Program Interface) allowed developers to access its immense database and algorithms for public use. To cite a favorite example, MoreCowbell.dj, now defunct, was a web tool that accessed the EchoNest API to take a track uploaded by the user and automatically add samples of cowbell along with the voice of the actor Christopher Walken in reference to a renowned Saturday Night Live sketch (Mims 2010). Humor aside, this and other examples allowed users to subvert the powerful MIR algorithms developed by EchoNest for

¹ “Pandora Music Genome Project.” Accessed March 31st, 2026. <https://www.pandora.com/about/mgp>.

² Steven Toja, personal email sent to the [ISMIR-community] listserv, March 6th, 2014.

³ “Spotify.” Accessed March 31st, 2026. <https://www.spotify.com/uk/about-us/contact/>.

musical ends that the company could never have envisioned. Since Spotify acquired EchoNest, it replaced the latter's API with its own, limiting access to musical metadata. The MIR algorithms that power Spotify remain a "black box" of which the user has little knowledge or control. Since then, those who wished to use MIR creatively needed to look to open-access software instead.

Viewed historically, MIR represented a shift in sound potentially even more significant than the digital revolution that began much earlier: the revolution of audio searchability (Einbond 2013, 63-65). Kenneth Goldsmith, an experimental poet with his own distinctive digital aesthetics discussed further below, has analyzed how the advent of digital text in the mid-20th century brought with it the novel concepts of copying and pasting, more natural to the computer than to the typewriter (Goldsmith 2011, 6). Along with those came the string-searching algorithm (Boyer and Moore 1977), at the heart of computer programming languages and later internet search engines. This allowed writers efficiently to search digital text to find a desired word or phrase, a capability incalculably amplified by online tools like search engines. So much of our relationship to digital text relies on string searching that it is difficult to imagine the computer, Internet, or mobile device without it.

The text search suggested new relationships to audio, allowing users to interface with associated textual metadata like titles, artists, and genre labels. However, these actions on strings of text are fundamentally different from searching sound itself, which still remained elusive in the late 20th century. Early digital audio was dominated by analog metaphors, for example the tape splicing paradigm of Digital Audio Workstations (or DAWs), or the model of patch cables and effects units that inspired visual programs like *Max* and *PureData* (*Pd*). These metaphors of timeline and signal flow—both preserving a linear organization of time—did not yield to fundamentally different concepts for manipulating audio during the first decades of digitalization. Not until the early 21st century did increasing computational speed, memory storage, and technological innovation lead to effective audio searches, combining MIR with ML algorithms to give rise to automatic genre classification and recommendation.

Why and how would an experimental musician during the first decades of the 21st century be interested in MIR for non-commercial use? Foremost is the immense quantity of musical "big data" available. Musicians working with vast personal archives of audio databases, or collective archives of internet sound, may literally not have time to listen to all the samples they wish to sort or search. We have long passed the point, addressed by Jacques Attali (1977/1985, 126) among others, of having access to more music than we can physically listen to in our lifetimes. Machine listening offers an alternative. Despite the dystopian irony, MIR is a response to this dilemma, with all its positive and negative consequences. Just as Spotify can help listeners to stumble upon tracks, artists, or entire genres that they may not have known before, so can customized MIR software help artists to chance upon sonic resources that they may not have thought of unassisted. To do this, MIR has offered novel contexts to search for sounds both manually and automatically, for example by browsing a two- or three-dimensional plot of a space of sounds (see CatOracle screenshot in Fig 13.1 below) or by searching for a closest match to a specified audio parameter or parameters. Yet the excess of data available through MIR can lead to creative results that are far from automatic by offering musicians a rich supply of information to choose from subjectively and intuitively.

A related motivation is audio content itself: noise-based sounds may resist organization by other means, for example the neat classification by symbolic pitch and dynamics employed in Western notated music as well as many music technology applications, such as commercial orchestral sample databases. For at least the last half-century, musicians have sought models to understand complex sounds that strain against the hierarchies of Western harmonic theory through concepts

including *musique concrète* developed by Pierre Schaeffer (Chion 1983/2009, 37), *musique spectrale* associated with Gérard Grisey (2000), and *musique concrète instrumentale* coined by Helmut Lachenmann (Steenhuisen 2003). For the complex sounds of audio recordings, electronic instruments, non-Western instruments, and instruments performed and amplified in experimental ways, pitch and duration are often less informative than other audio descriptors that describe the timbre, or tone color of a sound. Timbre has long resisted the linear scales traditionally used for pitch or rhythm, instead suggesting a multi-dimensional organization that is not easily measured (Wessel 1979). However, MIR technologies have represented a major leap in the quantification of timbre through audio descriptors. While the possibilities are endless, some common timbral descriptors include spectral centroid, which correlates to the perception of brightness or brilliance related to the relative register of a sound spectrum. Another descriptor is periodicity, which is anticorrelated to noisiness and measured through the periodic repetition in a soundwave. MIR has allowed one to analyze, search, and compare these descriptors, cutting across traditional musical hierarchies to suggest ways to approach the wealth of sound around us in all its marvelous plurality.

Finally, for musicians whose exploitation of diverse skill sets, reference points, and sonic resources pushes them beyond received genre divisions, MIR has helped suggest idiosyncratic, unexpected, and highly personal taxonomies. Machine listening has been applied to sound that defies traditional classification to suggest novel families, communities, and systems, as I demonstrate in the following case studies. While these artists may be difficult to classify, they have been united in using MIR to permeate received genre boundaries and point to new constellations. Marko Ciciliani has referred to “Music in the Expanded Field,” alluding to Rosalind Krauss’s seminal article on sculpture four decades earlier: “Nowadays, many composers are working beyond the traditional boundaries of music, expanding into other media and practices” (Ciciliani 2017). These artists have been working across and between traditional fields of expertise in ways that would be difficult to imagine without technological mediation. By exploiting, subverting, reprogramming, and customizing the MIR tools used for automatic genre classification, they have inevitably muddled the received genre divisions for which these tools were frequently used. In doing so, they may define new emergent genres of their own. Jennifer Walshe, one of the artists featured in this chapter, coined the term “The New Discipline” in 2016, yet she cautions, “it isn’t a style, though pieces may share similar aesthetic concerns. Composers working in this way draw on dance, theatre, film, video, visual art, installation, literature, stand-up comedy” (Walshe 2016). While the data-driven topic of MIR might seem distant from these embodied practices, there are significant overlaps between the artists’ attitudes toward trans-disciplinarity, distributed authorship, intermediality, and collective creation. There is a further parallel in working methods that defuse existing genre boundaries to coagulate into new proto-genres. Even though these artists may have been working on the edges of existing genres, their works nonetheless already participated in new generic groupings the artists may not acknowledge or of which they may not even be aware. At the same time, writings by Ciciliani and Walshe helped to begin to gather awareness of generic commonalities among these practices as they continued to gather momentum over the past decade.

Style Melody

Genre has been an explicit preoccupation of composer Johannes Kreidler, who coins the term “Style Melody” (Kreidler 2011) as an answer to Schoenberg’s *Klangfarbenmelodie*. The latter describes Schoenberg’s proposal of a “melody” made up of a succession of different timbres, or tone colors (*Klangfarben*), where the changes of timbre may be just as salient as those of pitch (Schoenberg 1911/1978). In referencing Schoenberg, Kreidler suggests a parallel between timbral organization and genre classification, a recurring theme in this chapter’s examination of MIR. In

Kreidler's series of works *Stil I*, which he calls a "modular composition system," he leverages the auto-harmonization feature of 90s commercial software *Band in a Box* to change rapidly between the purported styles accompanying a repeated input melody. For example, in *Stil Ia*, the style changes approximately every five seconds, articulated by clangorous chords from the live performers.⁵

In this work, Kreidler deals in an unexpected sense with musical scales of genre: sequences of parameters that he navigates like the twelve-tone scale. It is a reference to, and parody of, serial technique, exploited for results that could hardly be a greater contrast to the aesthetic stereotypes of serialism. Of course, references to other aesthetic precedents can also be heard, such as postmodern stylistic plurality. However, scale in another sense applies here as well, in the technologically mediated speed with which Kreidler navigates a seemingly inexhaustible list of genre possibilities. We might ask whether each click through the style melody points to a different micro-genre, or whether the choice of the now vintage software *Band in a Box* represents a genre allusion in itself. In either case, Kreidler, both technologically and aesthetically, reaches beyond the sometimes-narrow genres of experimental concert music of the last century.

In a more extreme example from the first decade of this century, Kreidler's electroacoustic work *product placements* (2008) samples 70,200 sources in 33-seconds. Although the MIR techniques may be rudimentary—Kreidler did not reveal what his specific search-and-retrieval criteria may have been nor did they seem to matter—the algorithms themselves were less important than the scale and speed of their application. Further, the documentation and discourse surrounding the work spilled beyond the genre of electroacoustic music: Kreidler has stated "the work is a network"⁶ and further explains in a documentary video, "in this piece, the musical composition, the essay, the sculpture, the performance, and the entire discussion surrounding are materials: one could say it's a multimedia theater work."⁷ The video shows him using a patch with software *Pd* to retrieve the samples, and probably also to fill out automatically the paperwork for GEMA, the German performance rights organization. This process resembles the style changes in *Stil I*, but now at millisecond speed. Listening to the work itself, one has the sense that its thousands of samples span many genres, making it an ironically condensed example of genre fusion. However, while this collage may have antecedents in postmodern methodologies, the millisecond scale of the samples mostly prevents them from being recognized as genre referents, despite a few tongue-in-cheek exceptions. Instead, the resulting work sounds more like a contribution to the genres of noise or glitch music.⁸ Yet the extramusical dimensions push the work's reception beyond these genres onto the borders of many more, notably conceptual and performance art. As Sol LeWitt defines, "in conceptual art the idea or concept is the most important aspect of the work" (1967). Indeed, Kreidler's infamous composition is most often shared through a description of its concept, or the video of the extramusical performance, rather than through the medium of the audio work itself.

Comparing Kreidler's use of basic MIR technology to commercial MIR approaches described above, his two works follow the model of accessing textual meta-data to parse large-scale audio databases: the genre melody is produced by serializing a sequence of subjectively assigned genre "descriptors" produced by the makers of the *Band in a box* software, like a peripatetic version of

⁵ Johannes Kreidler. "Stil 1a (Martin Jaggi & Leopold Hurt)." Accessed March 31st, 2026. https://youtu.be/sblIn_Zgu5o.

⁶ "Das Werk ist ein Netzwerk" (Groetz 2009, 7)

⁷ English subtitles accompanying *YouTube* video "Johannes Kreidler GEMA-Aktion product placements Doku". Accessed March 31st, 2026. <https://youtu.be/TRbsM2MXP20>

⁸ Johannes Kreidler. "product placements (music piece)." Accessed March 31st, 2026. <https://youtu.be/nYqnaiQpe1c>

Pandora's music recommendation system. *product placements* relies on textual metadata associated with the 70,200 samples that compose the work in order for them to be retrieved and concatenated by *Pd*. It does not, as far as its documentation indicates, take advantage of machine listening to process these samples. Still, Kreidler demonstrates how much computational, and compositional, power is enabled by searching musical metadata using even the simplest MIR techniques.

Material Network

During a similar time period to Kreidler, Maximillian Marcoll approached musical metadata, as well as audio transcription, in an imaginative way in his *Compounds* series (2008-2014). Like Kreidler, he describes a "material network" that he mediates through a process of "decentralized composing."⁹ This term could be read both in reference to his own algorithmic compositional processes of this period, and to the multiple actors that produced his recorded sonic material, including found and homemade samples and field recordings. Marcoll also built and performed with his own instruments, further decentralizing the locus of authorship, as in *Compound No.5: Construction Adjustment* for "Black Box" and electronics.¹⁰

Marcoll represents his material networks graphically as trees that could be interpreted as a highly personal genre classification. Marcoll writes:

The network structure portrays connections between materials which have inherited something from one another, and other connections, which essentially result from phenomenological similarities. All objects of the network belong to "identities," rough classifications of acoustic material, mostly determined by their origin, which build the titles of the pieces (Marcoll 2009, 3).

After grouping his audio into these "identities," Marcoll applies subjective tags to each, symbolized visually by lines connecting the nodes in the graph. Formally the data structure represented by Marcoll's network is a card catalogue (*Zettelkasten*) as derived from the work of sociologist and systems theorist Niklas Luhmann (Marcoll 2009, 4).¹¹

As with Kreidler, Marcoll utilized textual metadata to perform MIR in an original way, building a homegrown search-and-recommendation system. He also went deeper into MIR techniques to analyze musical content; but far from mechanical, he sees this as an additional site of invention. To transcribe his collected sonic materials for instrumental performance, Marcoll used his own software *quince* for a machine-assisted process of listening and notating. As with many MIR-influenced composers, software development has been an integral part of Marcoll's process. But he emphasizes that the output of the software is not automatic: it forces him to "work 'with the ears' and arrive at each decision manually and consciously even."¹² Even when using his purpose-built software, Marcoll refuses to accept automatic results and subverts them creatively through his compositional subjectivity.

⁹ "Materialnetzwerk," "dezentrales Komponieren" (Marcoll 2012, 40).

¹⁰ Maximillian Marcoll. "Compound No.5: CONSTRUCTION ADJUSTMENT." Accessed March 31st, 2026. https://youtu.be/E_mDeWJjGaM.

¹¹ See images in Marcoll 2012. Accessed March 31st, 2026. https://marcoll.de/text/articles/Marcoll_Compounds_2012.pdf

¹² "mit den Ohren' zu arbeiten und jede Entscheidung manuell und bewusst selbst zu treffen" (Marcoll 2012, 42).

Everything is Important

The emergent genre “The New Discipline” proposed by composer and performer Jennifer Walshe grouped together the work she and her peers have been doing over the past decades to put the body and voice of the performer at center stage, literally and figuratively (2016). In doing so she emulated contemporaneous research in embodiment and voice studies, bringing those topics into the heart of creative dialogue. Perhaps ironically, Walshe has since then also been profoundly influenced by machine bodies and voices growing out of recent research in AI and MIR. As MIR increasingly infiltrated our listening culture since the beginning of this century, its aesthetic fingerprints have become an expressive resource for subjective artistic expression. Walshe has referenced the surface textures, pops, and glitches of MIR, creating works that look and sound as if they could have been generated algorithmically whether in fact they were or not. In *The Total Mountain* (2014) for voice and video, texts assembled from Twitter feeds are sung by Walshe’s costumed alter-egos, performing a visual counterpart to the shuffle of identities and voices experienced when browsing social media. Walshe has curated these texts, costumes, and vocal pastiche styles with care while suggesting the volatility of a random Twitter feed, as the social media platform was called at the time. In the program text that accompanies the work, Walshe tellingly quotes Goldsmith: “We don’t read: we skim, parse, bookmark, copy, paste, and forward. We become information hoarders and amateur archivists who frantically collect, store, and move artifacts that we’ll never interact with.” To create the work, Walshe trolled the internet herself, as an MIR algorithm might, to produce the kind of fragmentary, chaotic juxtapositions that were becoming an increasing part of our daily lives at the time.¹³ Similarly, in *Everything is Important* for voice, string quartet, and video, pop-up windows proliferate on the video screen as if they have been opened by a malicious computer virus.¹⁴

Walshe’s early interest in MIR turned out to be prophetic as MIR became the input to increasingly prevalent AI and ML technologies. Since then, Walshe has taken ML as a critical point of inspiration, writing about her unique aesthetic viewpoint (Walshe 2023) and incorporating it in her creative work. In *There was a visitor*,¹⁵ the video that accompanies Walshe’s vocal and electronic performance refers to ML algorithms that presumably were used to generate some of the vocal and electronic sounds. The algorithms may be real or imagined, but perhaps it does not matter: just as on the Internet we are not aware whether we are interacting with a bot or a human, neither are we in Walshe’s music. In *IS IT COOL TO TRY HARD NOW?*, Walshe subjectively imitates the output of an actual AI algorithm from the Google DeepMind project’s WaveNet.¹⁶ Through her live vocal performance, she reembodies the disembodied AI agent as it attempts to reproduce the clicks and pops of human speech. As she explains:

I’ve learned to reproduce these examples with my voice and have used my versions in pieces such as *IS IT COOL TO TRY HARD NOW?* [...] The process of listening to the output of a neural network and trying to embody those results; the experience of trying to take sounds produced by code and feel my way into their logic – for me, that is part of the texture of being a person, of being a composer in 2018. (Walshe 2018)

In each case, Walshe creates work in response to the sonic texture of ML and MIR, as well as the visual and discursive language with which it is received, discussed, and propagated: its “meta-data.” Beyond harnessing these technologies only as tools, Walshe’s MIR poetics embrace the full

¹³ Jennifer Walshe. “The Total Mountain.” Accessed March 31st, 2026. <http://milker.org/thetotalmountain>.

¹⁴ See photo in “Jennifer Walshe: Work List.” Accessed March 31st, 2026. <http://milker.org/jenniferwalsheworklist/>.

¹⁵ As performed at Café Oto, London, February 6th, 2018.

¹⁶ DeepMind. “WaveNet: A generative model for raw audio.” Accessed March 31st, 2026. <https://deepmind.com/blog/article/wavenet-generative-model-raw-audio>.

circle of human and technological co-creation: how we read, write, critique, hear, listen, imitate, and respond to the technologies ever more present in our daily lives. These creative contributions distributed between human and non-human agents were one of her central claims in formulating The New Discipline: through collaboration with other artists, researchers, machines and influences outside of the past boundaries of experimental music, Walshe pioneered a nascent cross-disciplinary genre bringing together elements of many fields.

Corpus-Based Synthesis

One of the hallmarks of MIR and musical big data has been access to a large database or *corpus* of samples as source material. But how have MIR tools permitted artists to navigate these corpora of unprecedented size to creative ends? One technique is corpus-based synthesis (CBS), or the closely related technique concatenative synthesis, in which the sample corpus is the source of sound material to be spliced and recombined at the millisecond level. This is akin to the splicing technique Kreidler used to create *product placements*, except instead of his conceptual application, in CBS the choice and sonic content of the concatenated samples is of central creative importance. CBS has been implemented by Diemo Schwarz and colleagues at IRCAM (Institut de Recherche et Coordination Acoustique/Musique) through the *CataRT* (Schwarz 2007), currently distributed for software *Max* and *MaxforLive* in the packages *CataRT-MuBu*¹⁷ and *SKataRT*,¹⁸ respectively, and as a standalone application.¹⁹ In combination with the freely distributed package *MuBu for Max* (Schnell *et al.* 2009), these packages include libraries of external patches that *Max* users can download to extend the program's functionality.²⁰ Like the electroacoustic technique of granular synthesis, CBS relies on short grains of recorded audio concatenated into a new output. But with searchability possible through MIR techniques, the selection process presents new means of organization. CBS goes beyond the traditional stochastic approach of earlier techniques like granular synthesis where each grain is chosen randomly from an input sound and then modulated according to finely tuned parameters. In CBS, instead, grains are chosen by the computer based on the closest match to one or more audio descriptors that can be specified in multiple ways. These "target" descriptor values can be input directly or chosen graphically by navigating a multidimensional interface using a mouse, drawing tablet, or other controller. They can also be taken from a target grain of audio from a recorded or live source, in which case the CBS algorithm searches for the closest matching grain in another sound. The descriptors specified can relate to any audio parameters, including more traditional scales of pitch or duration. However, the most interesting, and unexpected results occur when the descriptors are drawn from timbral features, like brightness and noisiness as defined above. Whatever the selection method, CBS facilitates large-scale navigation of vast audio databases. In comparison to the laborious process by which an electroacoustic musician scans through a library of recordings or a hard drive of sound files for the perfect sound, the CBS algorithm can instantaneously search a corpus of many gigabytes to identify a sound with desired characteristics. While the search algorithms that power CBS are not as sophisticated as the neural networks referenced by Walshe, they represent a basic yet powerful form of ML.

CBS arose in the early 2000s and traces its heritage not only to granular synthesis but also Plunderphonics and microsound (Schwarz 2006) among other influences where microscopic (or rather microphonic) snippets of sound are juxtaposed potentially to bring elements of many genres together. A connection can be made to the website *Every Noise at Once* where visitors can

¹⁷ "CataRT-MuBu." Accessed March 31st, 2026. <https://cycling74.com/packages/catart-mubu>.

¹⁸ "SKataRT." Accessed March 31st, 2026. <https://forum.ircam.fr/projects/detail/skatart/>.

¹⁹ "CataRT Standalone." Accessed March 31st, 2026. <https://forum.ircam.fr/projects/detail/catart-standalone/>.

²⁰ "MuBu." Accessed March 31st, 2026. <https://cycling74.com/packages/mubu-for-max>.

rapidly audition and recombine audio samples of thousands of genres.²¹ Perhaps these explorations into genre recombination constitute an emergent genre of their own; however, CBS also presents opportunities for musicians to question, experiment with, and redefine genre. For example, a user could fill a corpus with samples drawn from multiple contrasting sources or genres and ask how they relate in terms of audio descriptors and where are points of contact and overlap. Or the user could ask the computer to “recommend” the samples from one collection that are the closest match to those from another. This could be heard as a home-grown music recommendation system, opening up the “black box” of algorithms kept hidden by proprietary platforms like Spotify. In this way, CBS can be heard as a kind of highly subjective genre-classification system, in which the creative MIR artist can define groups of samples with related characteristics or compare and contrast one group with another. These idiosyncratic genres may have little to do with the cultural context of genre discussed elsewhere in this volume, but they represent a productive subversion of the MIR tools of automatic genre classification.

More recently CBS has seen further implementations and extensions inspired by *CataRT*, including the project *FluCoMa* (Fluid Corpus Manipulation)²² for software platforms *Max*, *Pd*, and *SuperCollider*, seeking to bring “breakthroughs of signal decomposition DSP and machine learning to the toolset of techno-fluent computer composers, creative coders and digital artists.” Some of the earliest artistic explorations of the technological results by Lauren Hayes, Olivier Pasquet, Leafcutter John, and Rodrigo Constanzo revealed the potential for the software in electroacoustic composition and live improvisation contexts. Performer, composer, and programmer Constanzo uses *FluCoMa* and other tools to gather elements of his personal performance environment into comprehensive packages that he distributes as free software including *Confetti*, *Cut Glove*, *C-C-combine*, and *The Party Van*.²³ Constanzo describes *C-C-Combine* as “play anything with anything” in a video demo that shows him improvising on percussion accompanied by a corpus of electric guitar samples.²⁴ Similarly in “Kaizo Snare” he uses *FluCoMa* tools connected to a contact microphone scratched against his prepared snare drum to trigger bell-like metal percussion samples.²⁵ We could call this live MIR: the computer uses the real-time analysis of audio features from the percussion performance, captured through high-quality contact microphones, to search and retrieve samples from a database of pre-recorded samples that follow the improvisation as it unfolds. The descriptors chosen as criteria for similarity searches can have a decisive influence on the sonic output, and Constanzo reports that in “Kaizo Snare” and other recent works he analyzed loudness (perceptual amplitude), spectral centroid (a measure of brightness) and spectral flatness (a measure of how much the sound spectrum contains sharp frequency peaks). This means that connections between the live performance and pre-recorded corpus will favor these qualities rather than other descriptors, for example pitch.²⁶ The chosen samples are output through loudspeakers subtly positioned behind Constanzo’s percussion setup to fuse and confuse the live and pre-recorded sounds, complementing Constanzo’s improvisation with a battery of percussion samples that expand his drum set beyond its physical limits. Constanzo’s system using *FluCoMa* acts as a millisecond-speed musical recommendation system, responding to his live improvisation with samples that best match his performance, like a hyperactive version of Spotify.

²¹ Every Noise at Once. Accessed March 31st, 2026. <https://everynoise.com>.

²² “Fluid Corpus Manipulation.” Accessed March 31st, 2026. <https://www.flucoma.org>.

²³ Rodrigo Constanzo. “Software.” Accessed March 31st, 2026.

<http://www.rodrigoconstanzo.com/software/>.

²⁴ Rodrigo Constanzo. “C-C-Combine.” Accessed March 31st, 2026. <https://youtu.be/K5AgH1leBUU>.

²⁵ Rodrigo Constanzo. “Kaizo Snare.” Accessed March 31st, 2026. <https://youtu.be/MTWklm1oXWQ>.

²⁶ Rodrigo Constanzo. “Traces of Fluid Decomposition: Rodrigo Constanzo, 'For When The Bell Tolls'”. Accessed Oct. 28th, 2021. <https://youtu.be/spYWGa0Okzs>.

While Constanzo harnesses a highly personal, customized system, his approach also offers potentials for wider use. A similar technique was commercialized by software startup Mogeess, which marketed a contact microphone and set of apps that could be connected to a smartphone or tablet. Akin to Constanzo, the Mogeess promotional video advertised “transforming any object around you into a unique musical instrument.”²⁷ Originally derived from work by Schwarz and colleagues at IRCAM, Mogeess’s microphone connects to a solid surface and analyzes its acoustic vibrations using a piezoelectric element. Unlike other microphones that respond to vibrations through air, this means the microphone can pick up vibrations that may not be audible to the naked ear, allowing the Mogeess software to substitute its own synthesized sounds in their place. Like Costanzo’s live improvisation, the software then retrieves correlating samples from a pre-recorded database or synthesizes new sounds using physical modelling algorithms.

Viewed through the lens of genre, Mogeess did not classify music, nor even sound in the usual sense, but near-silent acoustic vibrations from the everyday objects around us. These new categories are then used to trigger a synthesized performance. In this way Mogeess added some creativity and choice back into commercial MIR: the user decided to what object to attach the microphone, how to “play” the object, and how to map the resulting audio signal to audio synthesis. But it was not as open as FluCoMa or Costanzo’s custom-programmed packages, which users could modify and adapt in their own way.

Audio Mosaic

The view of MIR as an experimental resource for subjective genre classification and recommendation can be heard as a backdrop for my own compositional work. However, rather than an improvisational context, I craft this approach into notated scores for instruments and electronics. This porous approach to genre has also led me, parallel to Walshe and other artists she describes in “The New Discipline,” to move beyond the traditional genre of electroacoustic concert music toward a practice that crosses other musical and disciplinary genres. It has also pushed me beyond traditional artistic disciplines entirely, to embrace software development as part of my practice, like other composers introduced in this chapter. After I was initially inspired by Schwarz’s *CataRT*, in collaboration with Schwarz and colleagues, I was the first composer to create a work with a live interactive audio mosaic (Einbond 2009),³¹ and since then other artists like Constanzo have developed and personalized the technique further. I later developed *CatOracle*, distributed freely as part of the *CataRT-MuBu* package, to extend *CataRT*’s capabilities for audio mosaicking, corpus-based transcription, and computer improvisation (Einbond 2016). *CatOracle* is the principal software tool I used to produce my compositions *Central Park* and *The kind of problem a city is*, discussed below.

I work with CBS to take advantage of audio recorded in full detail with all its variety and richness. This is a decisive benefit over many other synthesis techniques: any recorded sound can be a source, making it particularly well adapted to the excess of audio recordings around us. This includes noisy sources, both acoustic and electronic, where traditional musical symbolic means of organization, such as pitch and duration, may be inadequate. Using a wide range of timbre descriptors, CBS mobilizes the unique MIR capability of automatically searching and retrieving audio content through machine listening. When CBS is applied to imitate a longer target sound by matching it to shorter samples from the corpus, then the result can be termed an *audio mosaic*, just as a visual mosaic represents a larger image with smaller tiles. A key goal is musical fusion: if samples are chosen based on similar descriptor values, and therefore similar timbres, they may

²⁷ “Mogeess brings you an entire world of creative possibilities.” Accessed March 31st, 2026. <https://youtu.be/30xWm0fyqsc>.

³¹ *Beside Oneself* for viola and electronics, premiered at IRCAM, Paris, February 8th, 2008.

sound to the listener as if they emerge from the same source. I borrow here the concept of perceptual fusion from the pioneering work of Albert S. Bregman (1990, 245), and it can be heard as a noisy analogy to the fusion of partials in a harmonic series: while sounds with similar harmonic characteristics may sound like they belong together in the musical context of pitch, I suggest that sounds with a noisy spectrum can sound fused when their non-pitch descriptors have similar values. Even when the “target” sound of an audio mosaic is unheard, it can still exert its faint presence over the shorter samples that are fused together in its image, analogous to a harmonic fundamental in a tonal context.

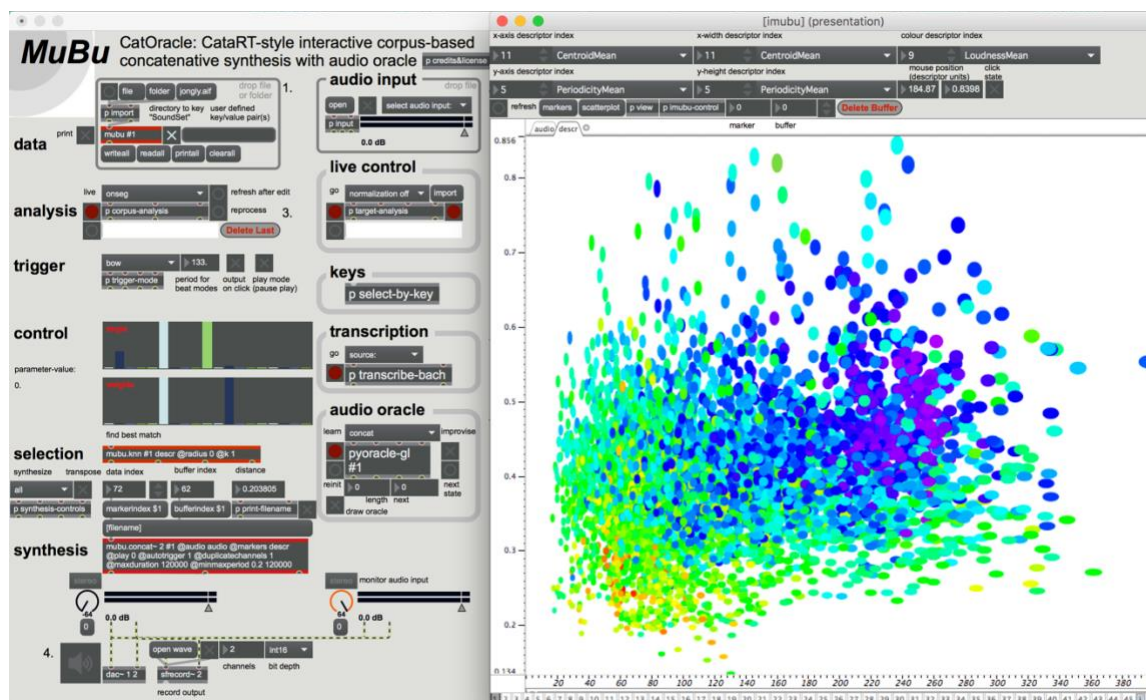
I have described my audio mosaic technique as “*musique instrumentale concrète*” (Einbond 2016) in reference to its background in Pierre Schaeffer’s *musique concrète* and Helmut Lachenmann’s *musique concrète instrumentale*. Short instrumental samples are mapped to a mosaic of a longer field recording by a comparison of their timbral descriptors, and then re-transcribed into a detailed instrumental score to be reinterpreted live. The result is an acoustic texture where noisy instrumental performance techniques invoke the concrete sounds of the field recordings that they accompany. By using MIR techniques to cross the two influential genres of *musique concrète* and *musique concrète instrumentale*, my work does not rest comfortably in either one but perhaps proposes a nascent genre of its own.

Central Park

In common with other composers cited above, MIR methods encourage distributed creativity where many voices join to contribute to the finished work. My process often begins as a collaboration with performers to record samples that will become the basis for the composition: in the case of *Central Park*, the performers of the ensemble loadbang with the unique instrumentation of baritone, bass clarinet, trumpet, and trombone.³² I have described this process as “radically personalized” because the resulting composition is fine-tuned to the unique timbres of the performers’ voices, bodies, and instruments (Einbond 2013, 68). With loadbang, I began our initial sampling sessions with a text that itself recalls the idea of “information retrieval” in an analog form: *Capital: New York Capital of the 20th Century* by, or rather compiled by, Kenneth Goldsmith in 2015. By collecting quotations from hundreds of sources, Goldsmith “rewrites” Walter Benjamin’s *Arcades Project*, exchanging Paris for New York and the 19th for the 20th century. Echoing Benjamin, Goldsmith divides his volume into 52 chapters clustered around generic titles, a radical taxonomy of urban life reminiscent of online tags, hyperlinks, or musical metadata. This is no surprise; as Goldsmith observes, “in many ways, the way we read *The Arcades Project* points toward the way we have learned to use the Web” (Goldsmith 2011, 155).

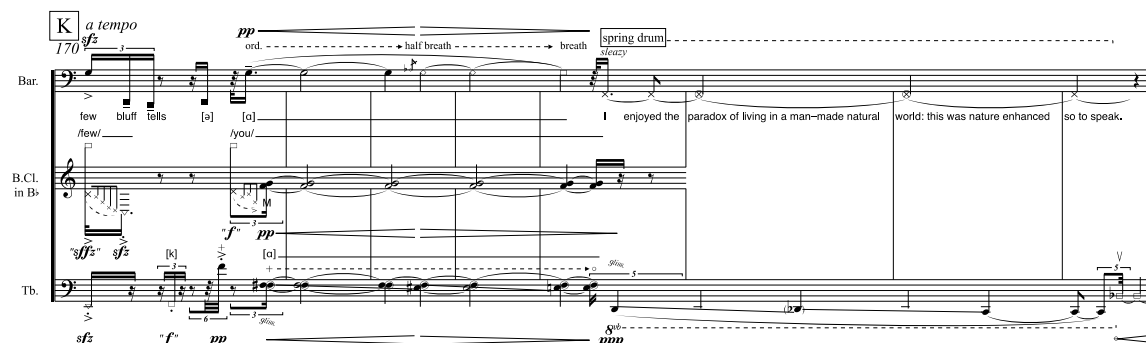
For my work with loadbang, I asked each of the four members to perform texts from the chapter “Central Park”: spoken, sung, whispered, inhaled, growled, and fried, further transformed by brass mutes and preparations including megaphone and spring drum for the voice, aluminum foil on the bell of the bass clarinet, and water poured into the trumpet. To give a sense of the scale of the resulting corpus, a screenshot shows the samples by one of the performers, baritone Jeff Gavett, loaded in *CatOracle*, representing a total of 43 minutes of audio in 145 sound files divided into 11,079 grains (Fig 13.1). Each dot represents one grain, plotted according to timbral descriptors of spectral centroid (correlating to the sound’s brightness, periodicity (anticorrelating to noisiness), and shaded according to loudness.

³² loadbang. Accessed March 31st, 2026. <https://www.loadbang.com>. Fittingly, the ensemble’s name is a playful reference to *Max* software, which I and other composers have used to create work for them.



<Fig 13.1 here>

Fig 13.1 *CatOracle* screenshot showing a corpus of vocal samples performed by Jeff Gavett of loadbang.



<Fig 13.2 here>

Fig 13.2 Excerpt from *Central Park* showing one musical phrase re-transcribed for baritone, bass clarinet, and trombone.

A parallel inspiration for the work is Robert Queneau's *Exercices de style* (1947), itself an encyclopedic investigation of genre in which a page-long description of an everyday Parisian scene is re-told 99 times using different literary devices. Analogous to Queneau's work, *Central Park* proceeds from a single two-minute field recording of the New York landmark, re-heard repeatedly through a multilayered process of reading, sampling, and transcription. The distinctive colors of the sampled texts are used to transcribe and re-transcribe the field recording drawing on each performer's unique sonic palette and personality. Both the original field recording and the performers' samples are analyzed to obtain a list of timbral descriptors, and the ML search algorithm matches the most similar sample grains to assemble them into an audio mosaic. Finally, the mosaic is transcribed into a detailed instrumental score, through my subjective interpretation

of the results, which in turn allows the performers to reinterpret it live, in effect performing CBS acoustically.

This is fundamentally different from searching textual metadata: here the search matches grains of audio content itself, as in the algorithms used for automatic music classification and recommendation by Spotify. The resulting transcriptions bear a shadowy imprint of the original recording, tracing faint hints of its color, texture, and internal timing. Applied successively, the effect is emphasized further: the field recording yields a version for solo baritone, then trombone, and in turn bass clarinet, each one a degree further from the original, as shown in the score excerpt (Fig 13.2). The transcriptions in the final composition are deployed in a formal patchwork, progressively moving closer to the original field recording and revealing larger segments of it, until concluding with the full recording accompanied by a close transcription for solo trumpet. Thus, the form itself is guided by the idea of audio search and retrieval, or alternatively, the metaphor of a dim memory rehearsed as it becomes progressively clearer. At the same time, the work's four interlocking solos are like the paths of four characters crossing the park from different times and directions: a palimpsest of perspectives from the ensemble and myself in our shared city.³³

While the free association of texts, instrumental sounds, and field recordings facilitated by MIR techniques touches on genres ranging from experimental poetry to concert music to *musique concrète*, the work also alludes to the historical genre of musical phonography, as suggested by the title's veiled allusion to another work of musical naturalism, *Central Park in the Dark* by Charles Ives. As Ives describes, "This piece purports to be a picture-in-sounds of the sounds of nature and of happenings [...] when sitting on a bench in Central Park on a hot summer night." (1906/1973, 31). Like Ives, I hope the juxtaposition of multiple layers of musical associations will transport listeners to their own memories of urban parks, or to Central Park itself, real or imagined.

The kind of problem a city is

The kind of problem a city is, for two percussionists, two keyboard players, and electronics, blurs genre boundaries between artistic media more explicitly, touching on sound installation, radio documentary, and site-specific theatre, among others. It takes radical personalization (Einbond 2013, 68) one step further by highlighting close collaboration with individual performers not only to sample the musical materials for a piece, but also to record interviews with the musician about their experiences. These personal histories are woven into the work as a portrait of my long-term collaborators Yarn/Wire, focusing upon our shared history together and relationship to our hometown of New York. The interview texts are complemented by excerpts read by the performers from Jane Jacobs's *The Death and Life of Great American Cities*, a book that not only defines an approach to urban planning still influential today, but itself blends genres between rigorous scholarship, personal anecdote, and poetic language:

A city cannot be a work of art. [...] Although art and life are interwoven, they are not the same things. Confusion between them is, in part, why efforts at city design are so disappointing. [...] Artists, whatever their medium, make selections from the abundant materials of life, and organize these selections into works that are under the control of the artist. (Jacobs 1961/1993, 485)

While a city may not be art, Jacobs's writing takes on artistic resonances that I amplify in my composition. As she describes, I make a selection from both the materials of her book and the

³³ loadbang, "Central Park by Aaron Einbond". Accessed March 31st, 2026.
<https://youtu.be/CM64PvHCxdc>.

sonic and human fabric of New York (Fig 13.3). Or, as musicologist Tim Rutherford Johnson writes, “Einbond constructs a map of New York made up of individual locations, chosen by the lives and personalities of the four musicians.” (2019, 10)

During the same sessions in which we recorded interviews and spoken texts, I asked the performers to bring instruments that I could sample to reproduce the minute coloristic and rhythmic details of the urban scenes and texts: mbira, kalimba, and melodica as one might hear from street performers, an array of small percussion to evoke the sounds of an outdoor grill (Fig 13.4), and piano and spring drum prepared with knitting needles, aluminum foil, and coins that lend themselves to imitate the sound of the M train, the elevated subway that passes loudly outside the kitchen window of one of the performers. Transcriptions were made first by visiting locations mentioned in the interviews and Jacobs’s book to make field recordings capturing the settings and experiences described there (Fig 13.3). As in *Central Park*, I compared these field recordings to the database of the performers’ samples using *CatOracle* and edited the transcriptions by hand into a detailed instrumental score that is reinterpreted live by the performers alongside the field recordings.



<Fig 13.3 here>

Fig 13.3 Recording Brooklyn Bridge Park, New York, with Russell Greenberg of Yarn/Wire.



<Fig 13.4 here>

Fig 13.4 Russell Greenberg plays small percussion (shekere, caxixi, tambourine, and mbira) in a transcription of Brooklyn Bridge Park (Photograph: Bruno Mathez).

Although the transcription process is computer-assisted, it is far from mechanical. One motivation for using MIR approaches is quantity of data: I recorded hours of samples with Yarn/Wire that, together with descriptor analyses, produced gigabytes of data. MIR provides a wealth of information—an overabundance of options—from which I must select during the compositional process. However, rather than purely automatic, MIR offers an expressive palette just as broad as any other compositional technique, if not more. In the daily work of composing, I may transcribe a single field recording into an instrumental mosaic over and over before choosing the one version as the basis for further development, or multiple multiple versions I composite to incorporate desired characteristics of each. Thanks to MIR, each transcription step is fast, but the entire process of iteration, evaluation, and subjective editing of the results is slow. The numerous variables involved in transcription are controlled subjectively: there is no single best way to render the sounds of a subway for prepared piano and spring drum, just as no two individuals hear these sounds the same way, or no two experiences of urban life are the same. It is the process of selection, which Jacobs describes above, that provides the key to creativity.

In performance, the field recordings, transcriptions, interviews, and texts are complemented by staging and lighting to situate the performers and audience in a theatrical dramaturgy that turns the space into an imagined map of New York. The many voices from which the work is sourced spill out of the confines of the genre of concert music toward the borders of performed installation, radio play, sonic documentary, personalized map, or pocket theater to ask the unanswerable question of what sounds like home.³⁴

³⁴ “*The kind of problem a city is*” released on *Cities*. Yarn/Wire. multi.modal MM03. LP. Accessed March 31st, 2026. <https://youtube.com/playlist?list=PLd-yAHCtHwIO5e-C1c7ZucRWPLOO4Flcu>

Moving Information

These are a few diverse examples, among many more, of ways MIR has been used beyond automatic genre classification to challenge received categories and, in the process, posit new ones. The use of MIR by the artists discussed above around the period 2008-2018 has exploded in the following years as MIR data become the inputs to ML models of ever-growing size. As a result, now more than ever, these technologies are demanding a much-needed reassessment of concepts of originality, authorship, and expression. They challenge the musical work as a unique site of authenticity, instead suggesting it to be a distributed practice involving many creative voices in dialogue. Processes of collage and intertextuality that began with modernism and postmodernism are accelerated to unprecedented scales and speeds. Now, in the age of AI and MIR, artistic invention can be acknowledged as less to do with producing information and more to do with retrieving, selecting, organizing, and reproducing it.

This is not new to artists, who have always used appropriation and borrowing as a creative resource, nor to many scholars who have observed the multiplying artistic implications of “reproduction” over the last century, from Walter Benjamin onward (1968, 217). However, some critics have been slow to catch up, and criteria for originality still require re-evaluation. As early as 2010, literary theorist Marjorie Perloff writes of “moving information” to highlight the increasing emphasis not on creation of content, but on “mastery of existing information and its dissemination” (Perloff 2010 cited in Goldsmith 2011, 1), even more relevant today than when it was written. Goldsmith and Perloff question tired tropes of originality, the former instead proposing “that today’s writer resembles more a programmer than a tortured genius, brilliantly conceptualizing, constructing, executing, and maintaining a writing machine” (Goldsmith 2011, 1-2). These observations seem prophetic considering the “writing machine” of AI such as large language models (LLMs) that would become popular over a decade later.

Critiquing literary reception of his time, Goldsmith notes that practices of reproduction have long been accepted in other fields, for example among visual artists from Marcel Duchamp to Andy Warhol (7). These artists and the reception of their work have revealed that, rather than compromising artistic expression, the act of reproduction is inherently creative: “The secret: the suppression of self-expression is impossible. Even when we do something as seemingly ‘uncreative’ as retyping a few pages, we express ourselves in a variety of ways” (Goldsmith 2011, 9). This point of view, from over a decade ago, still offers a promising way forward that could benefit more recent reflections on creativity with AI. Or, turning to the musicians in this chapter working with MIR, by retrieving musical information that already exists from one source or another, and repurposing it in novel contexts, they use reproduction to create profoundly personal musical statements.

Which brings us back to genre: by subverting the technologies used increasingly to reify genre through automatic classification, creative MIR has been offering potent means to make work that is difficult to classify through crisscrossing and problematizing genre boundaries. This permeability is not new: as Benjamin writes “all great works of literature found a genre or dissolve one” (1968, 201). Yet it takes on particular resonance in this century: writer David Shields elaborates, “I’m interested in the generic edge” (2010, 65); “[t]he books that most interest me sit on a frontier between genres. On one level, they confront the real world directly; on another level, they mediate and shape the world” (2010, 69). Similarly, the musicians in this chapter defy easy classification to participate in many genres and forge new ones at the boundaries between composition, improvisation, performance, theater, documentary, and software development: whether Ciciliani’s expanded field, Walshe’s New Discipline, or an incipient genre of creative MIR. To repurpose a question by media theorist Alexander Galloway, “besides producing a class of objects already designated as literature did they, or can they, produce other

things, which may or may not be literature, and which may be even more interesting than literature?" (Galloway 2014, 3). To answer this question in relation to music of the early 21st century, creative MIR is one place to look.

References

- Altman, Rick. 1999. *Film/Genre*. London: British Film Institute.
- Attali, Jacques. 1977/1985. *Noise: The Political Economy of Music*. Trans. Brian Massumi. Minneapolis: University of Minnesota Press.
- Benjamin, Walter. 1968. *Illuminations*. Trans. Harry Zohn. New York: Schocken Books.
- Boyer, Robert S. and J. Strother Moore. 1977. "A Fast String Searching Algorithm." *Communications of the Association for Computing Machinery*. 20 (10): 762–772.
- Bregman, Albert S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: The MIT Press.
- Bucklet, Sean. "Nokia Music joins the auto-playlist club, brings Pandora-like song-finding to Lumia 800." *Engadget*. Accessed July 9th, 2021. <https://www.engadget.com/2011/11/15/nokia-music-joins-the-auto-playlist-club-brings-pandora-like-so/>.
- Chion, Michel. 1983/2009. *Guide to Sound Objects*. Translated by John Dack and Christine North. Paris: Éditions ^{SEP} Buchet/Chastel. Accessed October 18th, 2021. <http://ears.pierrecouprie.fr/IMG/pdf/Chion-guide/GuidePreface.pdf>
- Ciciliani, Marko. 2017. "Music in the Expanded Field: On Recent Approaches to Interdisciplinary Composition." In Rebhahn, Michael and Thomas Schäfer, eds. *Darmstädter Beiträge zur Neue Musik*. Mainz: Schott, 22-35.
- Crook, Jordan. "The Echo Nest CEO On What Big Data Means to The Music Industry." *Techcrunch*. Accessed July 9th, 2021. <https://techcrunch.com/2012/03/23/the-echo-nest-ceo-on-what-big-data-means-to-the-music-industry/>.
- Einbond, Aaron. 2016. "Musique Instrumentale Concrète: timbral transcription in *What the Blind See* and *Without Words*." *Open Music Composer's Book Vol. 3*. Paris: Editions Delatour / IRCAM, 155-171.
- Einbond, Aaron. 2013. "Subtractive Synthesis: noise and digital uncreativity" in Aaron Cassidy and Aaron Einbond, eds. *Noise In And As Music*. Huddersfield, UK: University of Huddersfield Press, 57-76.
- Einbond, Aaron, Diemo Schwarz and Jean Bresson. 2009. "Corpus-Based Transcription as an Approach to the Compositional Control of Timbre." *Proceedings of the International Computer Music Conference (ICMC)*, Montreal, 223-26.
- Einbond, Aaron, Diemo Schwarz, Riccardo Borghesi, and Norbert Schnell. 2016. "Introducing CatOracle: Corpus-based concatenative improvisation with the Audio Oracle algorithm." *Proceedings of the International Computer Music Conference (ICMC)*, Utrecht, 140-46.
- Fingerhut, Michael. 2004. "Music Information Retrieval, or how to search for (and maybe find) music and do away with incipits," *Proceedings of the IAML-IASA Congress*, Oslo.
- Frow, John. 2005. *Genre*. London and New York: Routledge.

Galloway, Alexander R., Eugene Thacker, and McKenzie Wark. 2014. *Excommunication: Three Enquiries in Media and Mediation*. Chicago and London: University of Chicago Press.

Goldsmith, Kenneth. 2011. *Uncreative Writing*. New York: Columbia University Press.

Grisey, Gérard. 2000. "Did you say spectral?" Translated by Joshua Fineberg. *Contemporary Music Review*, 19: 3, pp. 1-3.

Groetz, Thomas. 2009. "'Kunst muss verdächtig sein': Der Komponist und Aktionskünstler Johannes Kreidler," *MusikTexte* 123, 5–10.

Ives, Charles. 1906/1973. *Central Park in the Dark*. Ed. Jacques-Louis Monod. Hillsdale, NY: Boelke-Bomart.

Jacobs, Jane. 1961/1993. *The Death and Life of Great American Cities*. New York: Random House/The Modern Library.

Kreidler, Johannes. 2012. "Stilmelodie" in Claus-Steffen Mahnkopf, Frank Cox, and Wolfram Schurig eds. *Music Material Today (New Music and Aesthetics in the 21st Century*, vol. 8). Hofheim: Wolke Verlag.

LeWitt, Sol. 1967. 'Paragraphs on Conceptual Art', *Artforum* Vol.5, no. 10, pp. 79-83.

Marcoll, Maximillian. 2012. "Compounds — Dezentrales Komponieren." *Positionen* #93, 40-42.

Marcoll, Maximillian. 2009. "Compounds." Accessed July 9th, 2021.
http://www.marcoll.de/archives/texts/compounds_eng.

Mims, Christopher. 2010. "Auto-Mash-up Your Favorite Tracks." *MIT Technology Review*. Accessed Oct 13th, 2021. <https://www.technologyreview.com/2010/05/20/203274/auto-mash-up-your-favorite-tracks/>.

Perloff, Marjorie. 2010. *Unoriginal Genius: Poetry by Other Means in the New Century*. Chicago: University of Chicago Press.

Rutherford Johnson, Tim. 2019. "The World In Pieces: Aaron Einbond." *TEMPO* 73, pp. 7-12.

Schnell, Norbert Axel Röbel, Diemo Schwarz, Geoffroy Peeters, and Riccardo Borghesi. 2009. "MuBu & Friends—Assembling Tools for Content Based Real-Time Interactive Audio Processing in Max/MSP." *Proceedings of the International Computer Music Conference (ICMC)*, Montreal, pp. 423-26.

Schoenberg, Arnold (1911/1966/1978). *Theory of Harmony*. Translated by Roy E. Carter. Berkeley and Los Angeles: University of California Press.

Schwarz, Diemo. 2006. "Concatenative Sound Synthesis: The Early Years." *Journal of New Music Research*, vol. 35, no. 1, pp. 3-22.

Schwarz, Diemo. 2007. "Corpus-Based Concatenative Synthesis." *IEEE Signal Processing Magazine*, vol. 24, no. 2, pp. 92–104.

Steenhuisen, Paul. 2004. "Interview with Helmut Lachenmann—Toronto, 2003." *Contemporary Music Review*, 23:3, pp. 9-14.

Walshe, Jennifer. 2016. "The New Discipline." *Borealis Festival Catalogue*. Accessed July 9th, 2021. <http://www.borealisfestival.no/2016/the-new-discipline-4/>.

Walshe, Jennifer. 2018. "Ghosts of the Hidden Layer." *MusikTexte 159*, pp. 29–38. English version accessed July 9th, 2021. <https://texte.musiktexte.de/mt-159/207/ghosts-of-the-hidden-layer>.

Walshe, Jennifer. 2023. "13 Ways of Looking at AI, Art & Music." *Unsound Dispatch*. Accessed March 31st, 2026. <https://unsoundfestival.substack.com/p/unsound-dispatch-13-ways-of-looking>

Wessel, David. 1979. "Timbre space as a musical control structure," *Computer Music Journal*, vol. 3, no. 2, pp. 45-52.