What the Label Misses: A Hybrid Multidimensional Taxonomy for Analyzing Heterogeneous Uses of GenAI in Popular Music

Andrés Mondaca Sepúlveda
Master's Programme in Music, Music Technology, UNAM
andresmondaca.sepu@gmail.com

Abstract

The proliferation of generative artificial intelligence (GenAI) in popular music has outpaced available analytical frameworks. Terms such as "AI Cover", "AI Artist", and "human-AI co-creation" group heterogeneous practices under ambiguous labels, obscuring substantive differences in works, authorship, and modes of human-machine integration. Moreover, existing critical and regulatory frameworks tend to rely on unidimensional criteria, lacking tools to systematically describe or compare contribution and integration across specific musical works. As a methodological proposal, this article introduces a multidimensional taxonomy for analyzing the use of GenAI in popular music singles. The taxonomy articulates three dimensions: (i) a use typology classifying GenAI's role at the vocal, musical, and lyrical levels, (ii) a contribution axis estimating GenAI's input across nine parameters of each single, and (iii) an integration axis measuring human-machine coupling at structural and procedural levels. A pilot application on key cases demonstrates its analytical capacity.

1 INTRODUCTION

While music creation using artificial intelligence systems dates back to the last century with works such as the *Illiac Suite* (Hiller & Isaacson, 1957) or David Cope's *Experiments in Musical Intelligence (EMI)* during the 1980s (Cope, 1989), it is only in the last decade that we have begun to see a significant impact of these systems, particularly generative AI (GenAI)[1], on popular music.[2]

Although this landscape is still recent, it has given rise to various approaches to the technology. At one extreme, we find pieces that generate specific musical material using symbolic systems, such as *Daddy's Car* (2016) by Carré and Pachet, created with Flow Composer (Pachet et al., 2021), or the album *Chain Tripping* (2019) by YACHT, composed using Google's Magenta system (Roberts, 2019). Others use GenAI solely for vocal transformation via timbre transfer, such as *Heart on My Sleeve* (2023) by Ghostwriter977 (Robinson, 2023) or *DEMO #5: nostalgIA* (2023) by the Chilean producer FlowGPT (McGowan, 2024). At the opposite end of the spectrum are songs generated almost entirely by generative systems for direct distribution, such as *How Was I Supposed to Know?* (2025) by Xania Monet (Anderson, 2025) or *Dust on the Wind* (2025) by The Velvet Sundown (Hiatt, 2025).

Despite this variety, critical frameworks operate with overly broad labels that blur the differences in both creative processes and artistic outcomes. The label "AI Music" itself is used by the press and the public to group heterogeneous projects without distinguishing the type of system or how the generated material was processed, which makes it "crucial to specify the particular form being discussed rather than resorting to generalizations" (Avdeeff, 2025).

New terms also operate ambiguously. Several authors use "AI Cover" for songs where GenAI replaces the vocal timbre of a pre-existing recording with that of a well-known singer (Ünal & Taylan, 2025; Galuszka, 2024; Baris, 2024); yet searches for the term on YouTube or Spotify mostly return reversions generated entirely with GenAI in a different genre[3], without necessarily involving

1 Computational techniques that learn patterns from training data to produce seemingly novel, meaningful material, in this case audio, symbolic, or textual (Feuerriegel et al., 2024), as distinct from rule-based or purely algorithmic composition systems.
2 Following Tagg (1982), we define popular music as music conceived for mass distribution to socioculturally heterogeneous groups, and stored and distributed in non-written form, which in today's digital context includes distribution on streaming platforms.
3 This inconsistency can be explained by the evolution of practices associated with the term: those based on vocal cloning emerged as a trend around 2023 (Feffer et al., 2023) and remain present, but style-transfer reversions, enabled by refinements in platforms such as Suno or Udio that support work with reference audio, became a trend around 2025 and displaced the former in current search results.

vocal cloning.[4]

The label "AI Artist" behaves similarly: the specialized press applies it to "virtual personalities" whose music is generated primarily by GenAI (Zellner, 2025), in line with Shin et al.'s (2025) "AI Singers", while Cross (2025) reserves it for human artists working with generative systems. There is "significant disagreement as to who or what deserves to be called an AI Artist" (Browne, 2022).

Terms such as "co-creation" or "collaboration" are proposed to distinguish human–GenAI composition from pieces that automate most of the process through text-to-music platforms (Morris et al., 2025; Pons et al., 2025; Avdeeff, 2025), yet what co-creating with a GenAI system means, and where its limits lie, remains unelaborated. Cross (2025) even argues that the system's autonomy does not preclude collaboration; on the contrary, it is a necessary condition for transcending its instrumental use and making "a distinctive contribution to the creative process."

Meanwhile, in the Americas, rights management organizations have responded to the rise of GenAI songs with binary or one-dimensional frameworks to describe, categorize, and regulate these works. In Mexico, for example, INDAUTOR stated that "works created exclusively by artificial intelligence systems are not registrable" (2025), without specifying what happens with works partially created using these systems. For its part, the SCD in Chile states that "it is not possible to register as a work results generated solely by AI tools without verifiable creative input from the registrant," but it provides a toggle on its registration platform that allows users to indicate whether AI was used partially in the music and/or lyrics and which tool was used for that purpose (SCD, 2025). Another case is the joint statement by ASCAP, SOCAN, and BMI, in which they state that they will "accept registration of musical compositions partially generated using artificial intelligence (AI) tools" (2025). Notably, to distinguish qualifying works, BMI proposes a four-level unidimensional categorization based on the use made of these systems (Broadcast Music Inc., n.d.).

The problem with these proposals is that the distinction between works created without, partially, or entirely with these systems is insufficient to account for the diversity of uses in creative work with them. On the one hand, solutions like INDAUTOR's tend to be reductionist by grouping very different proposals into just two categories. On the other hand, more detailed solutions such as BMI's, being one-dimensional, mix criteria of content, creative process, and post-production, generating noisy categories that hinder an orderly systematization for comparing the pieces.

Based on the above, this research proposes a theoretical framework that generates clear definitions for the terminology surrounding these practices and strengthens categorization through three independent dimensions that describe compositions from different angles: use (operational/discursive axis), contribution (output axis), and integration (process axis), using complementary methods—qualitative for the first axis and quantitative for the other two.

2 RELATED RESEARCH

While most research on music and GenAI adopts a technical approach, focused on the development, analysis, and categorization of generative systems (Briot et al., 2020; Fernández & Vico, 2013; Herremans et al., 2017), there are also studies focused on musical pieces as *output* and on the creative process.

In the first case, Huang et al. (2020), Morris et al. (2024, 2025), and Pons et al. (2025) analyze and compare musical pieces created using GenAI; the latter two also publish *datasets* that categorize them according to the systems used and how they were employed, along with other parameters described qualitatively. Although none report a detailed methodology for replicating their categories, these works form a key foundation for our proposal.

In the second category are the works of Vear (2021) and Vear et al. (2023), who, drawing on Small's concept of *musicking* (1998), focus on the process of musical creation with GenAI systems, with an emphasis on live musical performance. Their distinction between concurrent, collaborative, and co-creative interaction will be revisited in our integration axis.

3 THE TAXONOMY

3.1 Design

This proposal offers two key distinctions from previous research. First, it is multidimensional: it independently measures the mode of use, the amount of generative material in the output, and the creative process. Vear et al.'s (2023) framework comes closest to this approach, but its questions yield only qualitative descriptions. Hence the second distinction: our dimensions operate in a hybrid manner[5], qualitative (typology of use, level descriptions) and quantitative (contribution and integration axes), ensuring a clear and robust methodology[6] that facilitates its use in future research.

[4] Definition aligned with Ventayen (2026).

[5] Although Pons et al. (2025) and Morris et al. (2024, 2025) report a quantitative analysis, their approach corresponds to a "mixed methods" approach to qualitative content analysis (Mayring, 2014). In our case, quantitative methods operate as discrete classification rules, not as tools for analyzing results.

[6] In accordance with Nickerson et al. (2013), we understand the robustness of the taxonomy as the presence of "enough dimensions and characteristics to clearly differentiate the objects of interest." Likewise, we adopt their hybrid approach between the conceptual-deductive (such as the division of parameters on the contribution axis) and the empirical-inductive (such as the design of labels in the usage typology).

Second, while this research does not propose the creation of a new dataset, it does establish criteria for defining the corpus of pieces to be categorized: popular music singles created with some degree of GenAI involvement in musical composition, vocal identity, and/or lyrics[7], recorded in a fixed format and distributed via digital platforms. In this regard, we differ from Pons et al. (2025), who, in addition to singles, categorize works in various formats such as albums, installations, or performances. To balance this filtering, our criteria include casual creators, relevant for amplifying the impact of GenAI music through social media trends that have significantly increased the visibility of these systems.[8]

It is important to note that the model we propose is a flexible and extensible prototype (Nickerson et al., 2013) since, although it is designed to categorize and describe singles, it can serve as a modifiable basis for categorizing other musical formats or even other types of art created with GenAI. Its application, in any case, is not automatic: each item is coded by an analyst applying judgment to the available evidence; the taxonomy structures that judgment rather than replacing it. A compact summary of the codebook is provided in Appendix A.

3.2 Typology of Use

The first dimension corresponds to a typology of use that qualitatively describes, through nominal tags, the type of operation delegated to GenAI in the composition of popular music. Here we expand on the "use of AI" tag proposed by Pons et al. (2025) through a structure that divides the labels into three domains: (i) vocal transformation[9], which encompasses vocal cloning and timbre transfer in vocal tracks; (ii) music generation, which refers to the use of symbolic or raw audio systems for the generation of musical elements; and (iii) lyric generation, referring to the use of text-based models for the creation of lyrics. In turn, the first two domains include, in addition to the operational tags OT (exclusive), the assignment of discursive tags DT (non-exclusive).[10]

The operational tags indicate whether work with GenAI was conducted in each domain and the objective pursued. The discursive tags capture how this use is publicly framed, thereby clarifying the ambiguous terminology mentioned in the introduction.

3.2.1 Operational tags OT (exclusive: a single code must be assigned to each domain)

Table 1: Typology of use, operational tags (OT)

Domain	Tag (code)	Assignment criterion
Vocal transformation	No vocal transformation (V0)	Vocal track does not exist or has not been transformed by GenAI.
Vocal transformation	Substitution (V1)	GenAI replaces the timbre of the vocal track without targeting a specific identity recognizable to listeners.
Vocal transformation[11]	Impersonation (V2)	GenAI simulates a specific and identifiable vocal identity other than the creator's own.
Vocal transformation	Self-cloning (V3)	The creator uses an AI model trained on their own voice.
Music generation	No music generation (M0)	No musical element, instrumental or vocal, has been generated by GenAI.
Music generation	Component generation (M1)	GenAI generates part of the musical components of the piece.
Music generation	Full-track generation (M2)	GenAI generates the full set of musical elements in the piece.
Lyric generation	No lyric generation (L0)	Lyrics do not exist or have not been generated by GenAI.
Lyric generation	With lyric generation (L1)	The lyrics of the piece have been generated wholly or partially by GenAI.

Lyric generation includes only two labels, without the partiality/totality distinction that does apply in the domain of musical generation, since the amount of lyrics generated will be measured on the contribution axis.

[7] Uses of GenAI for production, distribution, or artistic concepts fall outside the scope of this study.
[8] For more information on these uses, see Galuszka (2024) and Feffer et al. (2023).
[9] In this domain, transformation is central: vocal tracks generated by systems that create complete music are evaluated in the music generation domain, since singing also falls into that category.
[10] In the lyric domain, we consider only operational tags, as these were designed as "constructed types" in the sense of Bailey (1994). Based on the analysis of multiple projects, unlike the musical and vocal domains, in this one we found no terminology or narratives that went beyond the "lyric generation" code proposed by Pons et al. (2025).
[11] The typology's domain architecture reflects currently dominant configurations of practice; refining it for emerging ones, such as impersonation through direct generation without vocal transformation, is a line of future work.

Musical generation, on the other hand, will be broken down in greater detail along that same axis (§3.3).

3.2.2 Discursive tags DT (non-exclusive: more than one tag may be assigned per domain)

Table 2: Typology of use, discursive tags (DT)

Typically associated domain (code)	**Tag (code)**	**Assignment criterion**
Vocal transformation (V2)	Fake Cover (FC)	A work based on a third-party composition whose vocal timbre is modified by GenAI to simulate the performance of an artist or character who did not originally sing it. The music may or may not be modified. e.g., Freddie Mercury *AI Covers*[12]
Vocal transformation (V2[13])	Fake Featuring (FF)	A work whose vocal timbre is modified by GenAI to emulate a fictitious collaboration between two or more artists or characters. e.g., FlowGPT - *DEMO #5: nostalgIA*[14] (2023)
Vocal transformation (V2)	Musical Necromancy (MN)	GenAI is used to simulate the voice of a deceased singer. e.g., Caro Molina - *Maldito Galán* (Remake) (2024)[15]
Vocal transformation (V2 and V3)	De-aging (DA)	GenAI digitally rejuvenates the artist's voice. The substituted base voice may belong to the same singer or to a third party. e.g., Randy Travis - *Where That Came From* (2024)[16]
Vocal transformation (V1 and V3)	Vocal Expansion (VE)	GenAI expands the vocalist's capacities, for example by widening range, altering timbre, or extending phonetic abilities, enabling multilingual performance (Pons et al., 2025). e.g., MIDNATT - *Masquerade* (2023)[17]
Music generation (M1 and M2)	AI Artist (AIA)	Authorship is assigned to digital musical performers with no physical existence. The music, including instruments and voices, is generated mostly or entirely by GenAI. e.g., Breaking Rust - *Walk My Walk*[18] (2025)
Music generation (M1 and M2)	Style-Transfer Reversion (STR)	A work based on a third-party composition whose musical arrangement is generated by GenAI in order to reinterpret it in a genre different from the original. The vocal track may or may not be transformed. e.g., Fake Music (*Funk AI Covers*)[19]
Cross-domain (V2, M1 and M2)	Apocryphal Release (AR)	The work is presented as a new release by an inactive artist or band. GenAI is used for vocal impersonation and/or music generation that replicates their style.

[12] Available at https://youtube.com/playlist?list=PLYpJH96O5prj6XySGuCi-xpCD3ICR_wzy. All links to works cited in this article were last reviewed in August 2026.
[13] Although most FFs emulate celebrities through impersonation (V1), there are cases of self-cloning (V2), such as Yumi Matsutoya's song *Call Me Back* (2022), where the artist at age 68, simulates a duet with her own cloned 1970s-era vocal timbre.
[14] Available at https://youtu.be/XIg9fNaw0rE.
[15] Available at https://youtu.be/X5EnFMaFw1I.
[16] Available at https://open.spotify.com/track/6JhrtLjmZTOqo4Nhv6sdmM.
[17] Available at https://youtu.be/l19AEmRvh-Y.
[18] Available at https://open.spotify.com/track/17HsyeI0Eq06n3XRntZ4DT.
[19] Available at https://youtube.com/playlist?list=PLT7ibESsX4yfgVdX1ROPpAGtE1nINJrFN.

		e.g., The Lost Tapes of the 27 Club: NirvanaAI - *Drowned In The Sun* (2021)[20]
Cross-domain (usually V2 and M2)	Humorous/ Parodic Content (HP)	The work is framed as parody or comedic content, generally seeking circulation or virality on social media. It may include humorous musicalizations of social media posts or impersonations of figures outside the music industry. e.g., Glorb - *The Bottom* (2023)[21]
Cross-domain (usually V1, V3 and M1)	GenAI-Enhanced Creativity (GEC)	The piece is publicly framed by its authors as a practice in which GenAI stimulates or expands human creativity, generally through partial or selective uses of generated material. e.g., YACHT - *(Downtown) Dancing* (2019)[22]

These discursive tags are non-exclusive because, in the analysis of pieces, we have observed that they often overlap and combine. For example, the case assigned as MN (Molina, 2024) also qualifies as FF, since Caro Molina simulates, through impersonation, a duet between herself and the late singer Selena. Furthermore, if it had been publicly announced as an official featuring, it could also have been assigned AR in combination with the others.

3.3 Contribution Axis

The second dimension, the contribution axis (or output axis), aims to quantitatively measure the proportion of material generated by GenAI compared to the material contributed by the human creator in each piece. We adopt the approach of Huang et al. (2020) and Morris et al. (2024), in which the piece is broken down into parameters. In our model, we consider nine: melody (MEL), harmony (HAR), bass line (BL), rhythm/groove (RHY), form/structure (FOR), instrumentation/arrangement (INS), vocal performance (VP), instrumental performance (IP), and lyrics (LYR). Formally, we define the set of parameters (P) as:

$$P = \{\mathrm{MEL}, \mathrm{HAR}, \mathrm{BL}, \mathrm{RHY}, \mathrm{FOR}, \mathrm{INS}, \mathrm{VP}, \mathrm{IP}, \mathrm{LYR}\}$$

$$|P| = 9.$$

Unlike the aforementioned studies, which assign descriptive labels to each parameter, we propose a numerical approach based on scores that allows us to assign, at the end of the measurement, a level of GenAI contribution (C_{AI}) to the piece as a whole.

3.3.1 Parameter-based coding

Each parameter $i \in P$ receives a GenAI contribution score (AIS) on a five-point ordinal scale: $\mathrm{AIS}(i) \in \{0, 1, 2, 3, 4\}$, where 0 = no contribution; 1 = minimal contribution; 2 = equal contribution; 3 = majority contribution; 4 = total contribution. The human contribution score (HS) per parameter is derived as a complementary variable:

$$\mathrm{HS}(i) = 4 - \mathrm{AIS}(i)\,. \qquad [1]$$

3.3.2 Treatment of NA and ND

A parameter may be coded as NA (not applicable) when it is irrelevant to the case (e.g., vocal performance in an instrumental piece) or as ND (not determinable) when it applies but the available evidence does not allow for a reliable score to be assigned. Based on this coding, we define two key sets:

$$\mathrm{AP} = 9 - n_{\mathrm{NA}} \qquad [2]$$

$$\mathrm{MP} = \mathrm{AP} - n_{\mathrm{ND}}\,, \qquad [3]$$

where AP (applicable parameters) corresponds to the total number of parameters relevant to the piece, and MP (measured parameters) to the subset of AP effectively coded with a value on the 0–4 scale; n_{NA} and n_{ND} denote the count of parameters coded as NA and ND, respectively.

As a sufficiency condition, the axis requires that at least 50% + 1 of the applicable parameters have been effectively measured. Otherwise, it is reported as ND in its entirety.

3.3.3 Calculation of the contribution level

When the sufficiency condition is met, the relative contribution of GenAI is obtained by summing the scores assigned to the measured parameters and normalizing them by the maximum possible score for those same parameters:

$$C_{\mathrm{AI}} = \frac{\sum_i \mathrm{AIS}(i)}{4 \times \mathrm{MP}}\,. \qquad [4]$$

The resulting value C_{AI} ranges from 0 to 1 and reflects GenAI's contribution to the parameters that were actually coded. ND parameters are excluded from the calculation, preventing the lack of information from being confused with a lack of contribution from GenAI. The count of NDs per item is retained as an input for the transparency level

[20] Available at https://youtu.be/t4DtnAYOHZQ.
[21] Available at https://youtu.be/gbbUWU-0GGE.
[22] Available at https://open.spotify.com/track/71RGbE1giYCvt1lovLGDut.

(§3.5).

3.3.4 Level Assignment

Finally, the normalized value C_{AI} is mapped to five discrete contribution levels, according to the following thresholds:

Table 3: Levels of GenAI contribution

Level	Range of C_{AI}	Interpretation
C_0	$C_{\mathrm{AI}} = 0$	Human composition[23]
C_1	$0 < C_{\mathrm{AI}} \leq 0.3$	Human composition with generative additions
C_2	$0.3 < C_{\mathrm{AI}} \leq 0.6$	Hybrid human/GenAI composition
C_3	$0.6 < C_{\mathrm{AI}} \leq 0.85$	Curated/modified artificial composition
C_4	$0.85 < C_{\mathrm{AI}} \leq 1$	Artificial composition

3.3.5 Considerations regarding the contribution axis

Three points should be clarified. First, contribution does not equate to authorship: this axis measures GenAI's relative contribution to the resulting musical/lyrical material, not intellectual property. The style-transfer reversion (STR) of *Toxicity* (System of a Down) into the funk genre, published by user Fake Music (2025)[24], would score low on C_{AI} and, consequently, high on human contribution. This is because the original harmony, melody, structure, and lyrics are preserved from the seed material with considerable fidelity; however, this largely human contribution belongs to the original composers, not to the creator of the reversion. The axis describes this distribution of contribution without prejudging questions of authorship, which belong to a different analytical domain.

Second, a high C_{AI}, while implying low human contribution, does not imply that human labor was also scarce. The axis measures the material present in the output, not the intensity of the process that produced it. *Relentless Doppelganger* (Dadabots, 2019)[25], a death metal livestream generated by an unconditional model (Pons et al., 2025), would be classified as an artificial composition with a score of C_4. However, its creation involved considerable human effort in curating the dataset, training the model, and the overall design of the project. Though outside our corpus of singles, this case illustrates creative work the contribution axis does not capture, motivating the integration axis below.

Third, equal parameter weighting and the Table 3 thresholds are provisional conventions. The thresholds are deliberately asymmetric: upper levels are narrower because stronger claims, such as labeling a piece an artificial composition, warrant stricter conditions. Equal weighting follows parsimony, complemented by the NA mechanism (§3.3.2), which excludes irrelevant parameters (an a cappella track codes instrumental performance and instrumentation as NA). Genre-sensitive weights and empirical threshold calibration remain as future work: both require a larger corpus of pieces coded with this instrument, which this prototype does not yet provide.

3.4 Integration Axis

Our third dimension, the integration (or process) axis, seeks to measure the degree of human-machine coupling during the creation of the piece: the imbrication that the human creator establishes with the generative system to produce the work. As a starting point, we draw on the research of Vear, who analyzes the act of *musicking* (Small, 1998) in live GenAI performances, recontextualizing it here toward the creation of pieces fixed in studio recordings. In particular, we revisit his distinction between "concurrent," "collaborative," and "co-creative" interactions (2021) and his approach centered on guiding sub-questions including how the GenAI system "interrelates with the human musician" (Vear et al., 2023). Unlike Vear, we opt for a numerical approach based on scores and mapping to levels of imbrication; in contrast to the contribution axis, the facets here operate not as complementary variables, but as sums of scores assigned to key questions. This axis is organized into two facets: structural integration (SI) and procedural integration (PI).

3.4.1 Structural Integration (SI)

Structural integration measures human-machine imbrication as defined by the creator's relationship with the construction, modification, and training of the generative system. It is assessed through two questions, each with three scoring options (Table 4). The first examines the model's authorship, ranging from the use of unmodified third-party systems to the development of a proprietary model from scratch. The second question addresses training, ranging from no training to training or fine-tuning with the creator's own material.

Table 4: Structural integration (SI)

Question (code)	Options	Score
How does the author relate to the construction of the system used? (S1)	An unmodified third-party model is used.	0
	An existing open model is modified or adapted.	1
	A model is developed from scratch. (its own architecture, not derived from an existing system).	2

[23] C0 is included for completeness and as a diagnostic marker: no corpus item should land here, so its appearance signals a review of inclusion criteria or coding; it also enables extension to control corpora without GenAI.
[24] Available at https://open.spotify.com/track/0TfERe8GKTHuKEuCQGvI3P
[25] Available at https://www.youtube.com/watch?v=JF2p0Hlg_5U

How does the author relate to the training of the system? (S2)	No training or fine-tuning is performed on the model used.	0
	The model is trained or fine-tuned with third-party data.	1
	The model is trained or fine-tuned with the creator's own material.	2

The scale prioritizes interventions that intensify the coupling between creator and system. Thus, Holly Herndon's *Eternal* (2019)[26], produced with Spawn, a model of her own authorship developed alongside her team and trained with Herndon's voice and her ensemble (Friedlander, 2019), would score 4. In contrast, *How Was I Supposed to Know?* by the AI artist Xania Monet (2025)[27], publicly presented as a song set to music using the GenAI Suno platform (CBS Mornings, 2025), would illustrate a case of zero structural integration.

3.4.2 Procedural Integration (PI)

This facet measures human-machine imbrication in the very process of musical composition and production with the generative system. Five key questions guide the scoring (Table 5), addressing different dimensions of the creative process with GenAI: (i) conditioning input; (ii) intra-system iteration; (iii) extra-system transformation; (iv) multi-system orchestration; and (v) performative dialogue.

Table 5: Procedural integration (PI)

Question (code)	Options	Score
How does the author instruct or condition the system in order to obtain a result? (P1)	General text prompt, with no explicit musical structure, including minimal or trivial indications.	0
	Non-performative structural conditioning, such as complete lyrics, harmonic progressions, formal instructions, or third-party vocal/instrumental takes.	1
	Incipient performative input, such as a short audio or MIDI file by the creator: humming, a motif, or a rhythm/groove idea.	2
	Structured performative input, such as a full demo or mock-up in audio or MIDI, or a complete vocal take.	3
How does the author explore and/or refine outputs within the system? (P2)	No iteration beyond the system minimum.	0
	Intra-system interaction: sampling multiple generations and/or using refinement functions such as inpainting, section regeneration, extend, etc.	1
How does the author transform generative material outside the system? (P3)	Direct publication of the output, or only minor technical modifications such as gain or equalization.	0
	Moderate musical editing, such as local arrangement changes, looping, sampling, or note adjustments.	1
	Substantive transformation, such as reharmonization, reorchestration, or timbral reworking.	2
How does the author articulate the different GenAI systems used in the piece? (P4)	A single system generates the core material.	0
	Two or more systems with distinct roles integrate their contributions into the final piece.	1
How is human–GenAI performative dialogue established? (P5)	No human performance is recorded over the generated material.	0
	Human performance executes material proposed by the system, for instance by rerecording musical elements or generated lyrics.[28]	1
	Human performance adds original creative material over the output, through voice or instruments using non-generated musical or	2

[26] Available at https://open.spotify.com/track/3SBhkamFAVooQtvDNz4ZJb
[27] Available at https://open.spotify.com/track/5k6KdTBo6Pio5fonm9QZY6
[28] Although P5 focuses on performative-musical dialogue, we include the vocal interpretation of generated lyrics, recognizing that singing lyrical material constitutes a form of performative appropriation of the output.

lyrical material built in relation to the generated material.

At the extremes of this facet, we can consider *(Downtown) Dancing* by YACHT (2019), produced using open-source GenAI tools trained on the group's previous catalog, with iterative exploration of outputs, the articulation of multiple generative systems, and human interpretation of the resulting material (Willmott, 2019), illustrating a case of high procedural integration. In contrast, *Be Inspired* by Papi Lamour (2025)[29], where the author wrote the lyrics and delegated musical and vocal creation to a GenAI tool (Reilly, 2025), without significant transformation or human-machine performative dialogue illustrates low procedural integration.

3.4.3 Level Assignment

In this axis, given that all questions can be formulated independently of the system used and the type of resulting work, the NA category is not considered. However, as on the contribution axis (§3.3.2), a question may be coded as ND when there is insufficient evidence to answer it reliably.

For each facet to be calculable, a sufficiency condition equivalent to 50% + 1 of answered questions must be met. Structural integration SI, considering only two questions, requires that both be answered. Procedural integration PI requires at least three of the five. Otherwise, the facet is reported as ND.

Once sufficiency is verified, both facets are calculated normalized to the interval [0, 1]. In SI, the normalized score corresponds to the sum divided by the maximum possible (4):

$$\mathrm{SI}_n = \frac{S_1 + S_2}{4} \; . \qquad [5]$$

In procedural integration, the normalized score is obtained by summing the scores assigned to the answered questions and dividing by the maximum possible score for those same questions:

$$\mathrm{PI}_n = \frac{\sum_i \mathrm{P}_i}{\mathrm{PI}_{\max}} \; , \qquad [6]$$

where the sum covers the questions actually answered and $\mathrm{PI}_{\max}$ corresponds to the sum of the maximum possible scores for those questions. For example, if P5 (maximum value = 2) is coded as ND, PImax = 7 and the normalization is calculated over the four answered questions.

Finally, the total integration TI is calculated only when both facets have an assigned value. In that case, it is obtained as the average of the normalized values of each facet:

$$\mathrm{TI} = \frac{\mathrm{SI}_n + \mathrm{PI}_n}{2} . \qquad [7]$$

If any of the facets is reported as ND, TI is also reported as ND.

The three values SIn, PIn and TI are mapped to five discrete levels, following the same threshold scheme. In Table 6, "I" represents any of these values. The levels I_2, I_3, and I_4 encompass the three modes of human-machine imbrication proposed by Vear (2021) (concurrent, collaborative, and co-creative), recontextualized here as gradations of a quantitative continuum.

Table 6: Levels of human–GenAI integration

Level	Range of I	Interpretation
I_0	$\mathrm{I} = 0$	Null integration
I_1	$0 < \mathrm{I} \leq 0.3$	Instrumental integration
I_2	$0.3 < \mathrm{I} \leq 0.6$	Concurrent integration
I_3	$0.6 < \mathrm{I} \leq 0.85$	Collaborative integration
I_4	$0.85 < \mathrm{I} \leq 1$	Co-creative integration

3.4.4 Considerations on the integration axis

Four clarifications should be made. First, in SI, the higher score assigned to training with one's own creations does not reflect an ethical assessment of the data's origin (a dimension that merits its own discussion), but rather the fact that the creator's material becomes a constitutive part of the system, establishing a degree of imbrication that is not achieved when training data is external to their practice.

Second, SI and PI remain analytically independent facets because they operate on distinct moments and objects: SI captures work on the system as technical infrastructure; PI, work with the system during the creative process. A high score in one facet does not imply a high score in the other. This independence allows for the description of diverse configurations of human-machine integration. In total integration, both facets weigh equally: each represents significant human work with the system, and any purpose-specific reweighting belongs to the use of the instrument rather than its design. Table 6 thresholds follow the provisional logic of §3.3.5.

Third, the terms "collaboration" and "co-creation" presuppose a certain attribution of agency to the generative system, a philosophically debatable issue.[30] Our taxonomy does not aim to resolve this ontological question; it adopts the terminology in use in the literature on musical GenAI to describe observable gradations of human-machine coupling, reserving the discussion on the nature of that agency for future work.

Finally, the coding of the axis depends critically on publicly available information about the creative process. Since many projects do not document their workflow in detail, applying the axis may require frequent use of ND. This limitation is addressed in the transparency level, discussed below.

[29] Available at https://open.spotify.com/track/6RsHrLwr9ARTwE13lQv4mQ
[30] Weaver (2025) uses the term "sense of collaboration" rather than collaboration in the strict sense to describe his own experience composing with GenAI.

3.5 Level of Evidence and Transparency

Applying this taxonomy goes beyond the analysis of the work as a sound object: it requires documentary sources on the creative process, such as interviews, public statements, or press releases. Therefore, before coding a piece, we must establish a level of evidence (E) regarding the use of GenAI, according to the criteria presented in Table 7. Only cases E1–E3 proceed to the coding process.

Table 7: Levels of evidence (E)

Level	Label	Interpretation
E_0	Not verifiable	There are insufficient sources.
E_1	Confirmed	There is a direct statement by the creator or team.
E_2	Substantiated	Verifiable external sources document the use.
E_3	Disputed but substantiated	Public denial coexists with verifiable external evidence.

Once this filter is passed, the typology of use is applied. Every eligible piece must receive at least one operational code demonstrating the use of GenAI in one of the three domains. If the coding simultaneously yields V0, M0, and L0, this indicates an absence of verifiable use or a measurement error, so the piece does not proceed to the quantitative axes.

Once the presence of GenAI is verified, the contribution and integration axes are calculated according to their own rules of sufficiency and handling of ND. Finally, the level of documentation transparency (T) is calculated as the proportion of parameters and questions effectively determined relative to the applicable total:

$$T = 1 - \frac{\mathrm{ND}_{tot}}{A}, \qquad [8]$$

where ND_{tot} corresponds to the total number of undeterminable codings and A to the total number of applicable parameters and questions in the quantitative axes. The value T, ranging from 0 to 1, is interpreted on five levels (Table 8).

Table 8: Levels of transparency

Level	Range of T	Interpretation
T_0	$T = 0$	Null transparency
T_1	$0 < T < 0.35$	Low transparency
T_2	$0.35 \leq T < 0.65$	Medium transparency
T_3	$0.65 \leq T < 1$	High transparency
T_4	$T = 1$	Total transparency

Based on this methodological framework, the following section presents a pilot application of the taxonomy.

4 PILOT APPLICATION

4.1 Case Selection

For illustrative purposes, we selected three key cases to which we applied our taxonomic system. In addition to meeting the corpus criteria (§3.1), these cases had to, on the one hand, have generated media attention (measured by the number of plays, all exceeding one million on at least one platform) and, on the other hand, provide confirmed evidence of GenAI usage (E_1). Furthermore, we ensured that the examples were heterogeneous in their use of GenAI, in order to generate a sample that covers diverse configurations within the taxonomy.

The selected examples are: *Daddy's Car*[31] by Benoît Carré and François Pachet (2016); *DEMO #5: nostalgIA*[32] by FlowGPT (2023); and *Let Go, Let God*[33] by Xania Monet (2025).

4.2 Cases and Results

In *Daddy's Car*, Carré and Pachet used Flow Composer, a system they developed themselves trained on a dataset of melodic lines from works composed by third parties (Lead Sheet Database). Based on conditioning with songs by The Beatles, they generated numerous melodic lines and chords in the band's style. After selecting fragments, they completed the composition and handled the arrangement, orchestration, and performance themselves. The lyrics were also written by Carré. This is a pioneering example of popular music in which the authors have publicly highlighted the new creative possibilities of these systems (Pachet et al., 2021; Carré, n.d.).

In *DEMO #5: nostalgIA*, Chilean producer Mauricio Bustos, creator of FlowGPT, used the Kits.ai platform to apply timbre transfer to his own vocal performance and simulate a fictional collaboration between Bad Bunny, Justin Bieber, and Daddy Yankee. Bustos uses GenAI solely for vocal impersonation, writing the lyrics, singing the original vocal track, and delegating the musical backing to human producers (McGowan, 2024).

In *Let Go, Let God*, poet Telisha Jones, creator of Xania Monet, uses the GenAI platform Suno to set her own lyrics to music. By entering them into the system, Jones automatically generated all the audio—both musical and vocal elements—from a text prompt (CBS Mornings, 2025).

Based on this publicly available information, we obtained the following results in our taxonomy:

[31] Available at https://www.youtube.com/watch?v=LSHZ_b05W7o
[32] Available at https://youtu.be/XIg9fNaw0rE
[33] Available at https://open.spotify.com/track/0in5pWf8oEzZRjl7nl3wJO

Table 9: Taxonomy applied to key cases

Axis - dimension / case	Daddy's Car – B. Carré & F. Pachet	DEMO #5: nostalgIA – FlowGPT	Let Go, Let God – X. Monet
E	E_1	E_1	E_1
OT	V0, M1, L0	V2, M0, L0	V0, M2, L0
DT	GEC	FF, AR	AIA
C_{AI}	C_1	C_1	C_4
SI	I_3	I_0	I_0
PI	I_3	I_3	I_1
TI	I_3	I_2	I_1
T	T_4	T_3	T_3

Detailed scores for each case are provided in Appendix B.

5 CONCLUSIONS

This study proposed a hybrid multidimensional taxonomy to describe the phenomenon of GenAI in popular music with greater precision than current terminology, and to systematize a documentation framework that future projects can adopt to report their own creative process. The taxonomy does not intend to make value judgments, though it offers a framework upon which to articulate specific ethical discussions if required.

The instrument is presented as a prototype. Applying it to larger corpora will allow us to identify ambiguities and refine parameters, questions, and thresholds. Rather than proposing a closed system, we aim to advance a methodological approach: one that is multidimensional and employs hybrid quantitative-qualitative definitions.

Every categorization loses information, and ours is no exception: reducing a creative process to a set of levels omits the singular texture that a detailed account preserves. We assume that cost as the counterpart of a different scale of analysis: like a world map that shows every country but loses the streets of each city, this taxonomy does not compete with ethnographic approaches but complements them. That scale is where its usefulness lies: it allows musicology to analyze large corpora with defined categories and offers rights management organizations a way to operationalize criteria such as "significant human contribution," measuring contribution and integration separately.

While we are convinced that multidimensionality is the path to a more complete understanding of the phenomenon, this does not exclude partial and flexible use: each dimension can be activated, deactivated, or reweighted according to the user's purpose, operating less as a fixed instrument than as an extensible template, adaptable to other disciplines by substituting parameters and questions. An interactive coding tool implementing the taxonomy is also under development.

In a landscape as diverse as GenAI in music, reductionism is particularly harmful: it confines heterogeneous projects under a single label or a single linear dimension, hindering precise analysis and stigmatizing proposals that could expand the creative space. A flexible taxonomic system enables an analytical language that more fairly distinguishes these systems' specific contributions, facilitating more informed research. We hope this work contributes to more nuanced and rigorous discussions on the role of GenAI in contemporary music creation.

AUTHOR DECLARATIONS

This paper relied exclusively on publicly available secondary sources and therefore did not require ethics committee approval. Claude Opus 4.7 (Anthropic) and ChatGPT 5.4 (OpenAI) were used as writing assistance, and DeepL was used for the English translation. All conceptual decisions and all paper content are the sole responsibility of the author.

ACKNOWLEDGMENTS

This work was developed within the Master's Programme in Music, field of Music Technology, at the Universidad Nacional Autónoma de México (UNAM), with the support of a postgraduate scholarship from the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI). I thank my supervisor, Dr. Pablo Padilla, for his guidance throughout this project, and Dr. Jorge David García and my colleagues in the research seminar for their comments on earlier stages of the manuscript.

REFERENCES

American Society of Composers, Authors and Publishers, Society of Composers, Authors and Music Publishers of Canada, & Broadcast Music, Inc. (2025, October 28). ASCAP, BMI and SOCAN announce alignment on AI registration policies. Retrieved April 2026, from https://www.ascap.com/press/2025/10/10-28-ai-registration-policies

Anderson, T. (2025, November 3). AI artist Xania Monet debuts on Adult R&B Airplay - A radio chart breakthrough. Billboard. Retrieved April 2026, from https://www.billboard.com/music/chart-beat/ai-artist-xania-monet-debut-adult-rb-airplay-chart-1236102665/

Avdeeff, M. (2025). Digital discourse and the human experiences of AI music: An ethnomusicological perspective. Ethnomusicology Forum, 34(2), 219–224.

Bailey, K. D. (1994). Typologies and taxonomies: An introduction to classification techniques. SAGE.

Baris, A. (2024). AI covers: Legal notes on audio mining and voice cloning. Journal of Intellectual Property Law & Practice, 19(7), 571–576.

Briot, J.-P., Hadjeres, G., & Pachet, F.-D. (2020). Deep learning techniques for music generation. Springer.

Broadcast Music, Inc. (n.d.). BMI and AI: What music creators need to know. Retrieved April 2026, from https://www.bmi.com/specials/bmi-and-ai

Browne, K. (2022). Who (or what) is an AI artist? Leonardo, 55(2), 130–134.

Bu-min, K., Whittemore, D., Michael, J., & Hitchhiker. (2023). Masquerade [Recorded by MIDNATT]. BigHit Music.

Bustos, M. (2023). DEMO #5: nostalgIA [Recorded by FlowGPT].

Carré, B. (n.d.). Daddy's Car: Flow Machines, The Beatles. Skygge. Retrieved April 2026, from https://www.skygge.fr/daddys-car-flow-machines-the-beatles/

Carré, B., & Pachet, F. (2016). Daddy's Car [Recorded by B. Carré]. Sony CSL.

CBS Mornings. (2025, November 5). Creator of AI artist says there's "a real person behind Xania" amid controversy [Video]. YouTube. Retrieved April 2026, from https://www.youtube.com/watch?v=EUBhjOToIC0

Cope, D. (1989). Experiments in Musical Intelligence (EMI): Non-linear linguistic-based composition. Interface, 18(1–2), 117–139.

Cross, A. (2025). Tool, collaborator, or participant: AI and artistic agency. British Journal of Aesthetics, 65(4), 535–551.

Dadabots. (2019). Relentless Doppelganger [Recorded by Dadabots].

Emerick, S., & Scott Sherrill, J. (2024). Where That Came From [Recorded by R. Travis]. Warner Music Nashville.

Evans, C. L., Bechtolt, J., Kieswetter, R., & Goodwin, R. (2019). (Downtown) Dancing [Recorded by YACHT]. On Chain Tripping. DFA.

Feffer, M., Lipton, Z. C., & Donahue, C. (2023). DeepDrake ft. BTS-GAN and TayloRVC: An exploratory analysis of musical deepfakes and hosting platforms. In Proceedings of the 2nd Workshop on Human-Centric Music Information Research (HCMIR).

Fernández, J. D., & Vico, F. J. (2013). AI methods in algorithmic composition: A comprehensive survey. Journal of Artificial Intelligence Research, 48, 513–582.

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. Business & Information Systems Engineering, 66(1), 111–126.

Friedlander, E. (2019, May 21). How Holly Herndon and her AI baby spawned a new kind of folk music. The FADER. Retrieved April 2026, from https://www.thefader.com/2019/05/21/holly-herndon-proto-ai-spawn-interview

Galuszka, P. (2024). The influence of generative AI on popular music: Fan productions and the reimagination of iconic voices. Media, Culture & Society, 47(3), 603–612.

Glorb. (2023). The Bottom [Recorded by S. & Mr Swags]. Underwata.

Herndon, H., & Dryhurst, M. (2019). Eternal [Recorded by H. Herndon]. On Proto. 4AD.

Herremans, D., Chuan, C.-H., & Chew, E. (2017). A functional taxonomy of music generation systems. ACM Computing Surveys, 50(5), 1–30.

Hiatt, B. (2025, July 5). AI "band" The Velvet Sundown officially confirm they're AI—and a "provocation." Rolling Stone. Retrieved April 2026, from https://www.rollingstone.com/music/music-features/ai-band-the-velvet-sundown-confirm-ai-1235379354/

Hiller, L. A., & Isaacson, L. M. (1957). Illiac Suite [Musical composition].

Huang, C.-Z. A., Koops, H. V., Newton-Rex, E., Dinculescu, M., & Cai, C. J. (2020). AI song contest: Human–AI co-creation in songwriting. arXiv [Preprint]. arXiv:2010.05388.

Instituto Nacional del Derecho de Autor. (2025, August 28). La SCJN resolvió que las obras creadas por IA no pueden registrarse como derechos de autor en México. Retrieved April 2026, from https://www.indautor.gob.mx/

Jones, T. (2025). How Was I Supposed to Know? [Recorded by X. Monet]. On Unfolded. Hallwood.

Jones, T. (2025). Let Go, Let God [Recorded by X. Monet]. On Unfolded.

Lamour, P. (2025). Be Inspired [Recorded by P. Lamour]. Quantum Beats.

Matsutoya, Y., & Karuho, K. (2022). Call Me Back [Recorded by Y. Matsutoya & Y. Arai]. On Yuming BANZAI! Yumi Matsutoya 50th Anniversary Best Album. Universal Music LLC.

Mayring, P. (2014). Qualitative content analysis: Theoretical foundation, basic procedures and software solution. Retrieved April 2026, from https://www.ssoar.info/ssoar/handle/document/39517

McGowan, C. (2024, July 9). The AI artist who used Bad Bunny's voice—and shot to fame. Rest of World. Retrieved April 2026, from https://restofworld.org/2024/ai-song-flowgpt-bad-bunny/

Molina, C. (2024). Maldito Galán (Remake) [Recorded by C. Molina].

Morris, L., Leger, R., Newman, M., Burgoyne, J. A., Groves, R., Mangal, N., & Lee, J. H. (2024). HAISP: A dataset of human–AI songwriting processes from the AI Song Contest. In Proceedings of the International Society for Music Information Retrieval Conference (ISMIR).

Morris, L., Newman, M., Tang, X., Renee, S., Vélez Vásquez, M., Leger, R., & Ha Lee, J. (2025). Expanding the HAISP dataset: AI's impact on songwriting across two contests. In Proceedings of ISMIR 2025.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. European Journal of Information Systems, 22, 336–359.

Pachet, F., Roy, P., & Carré, B. (2021). Assisted music creation with Flow Machines: Towards new categories of the new. In E. Reck Miranda (Ed.), Handbook of artificial intelligence for music (pp. 485–520). Springer.

Pons, J., Zukowski, Z., Parker, J. D., Carr, C., Taylor, J., & Evans, Z. (2025). Music and artificial intelligence: Artistic trends. arXiv [Preprint]. arXiv:2506.11694.

Reilly, N. (2025, November 14). BBC Introducing criticised after championing "pitch perfect" AI-generated music. Rolling Stone UK. Retrieved April 2026, from https://www.rollingstone.co.uk/music/bbc-introducing-criticised-after-championing-pitch-perfect-ai-generated-music-55904/

Roberts, A. (2019, September 13). YACHT's new album is powered by ML + artists. Magenta. Retrieved April 2026, from https://magenta.withgoogle.com/chain-tripping

Robinson, K. (2023, November 10). Ghostwriter, The Mastermind Behind The Viral Drake AI Song, Speaks For The First Time. Billboard. Retrieved April 2026, from https://www.billboard.com/music/pop/ghostwriter-heart-on-my-sleeve-drake-ai-grammy-exclusive-interview-1235434099/

SCD. (2025, December 29). Cómo declarar tus obras en SCD a través de Servicios Web: Tutorial paso a paso [Video]. YouTube. Retrieved April 2026, from https://www.youtube.com/watch?v=nLsoe4YV_dg

Shin, S. I., Kim, J. S., Kim, J. B., Lee, O. K., & Yun, H. (2025). What do you feel about an AI singer created by generative AI? Exploring sentiment about an AI singer from a dual realism perspective. In Proceedings of the Annual Hawaii International Conference on System Sciences (pp. 503–512).

Small, C. (1998). Musicking: The meanings of performing and listening. Wesleyan University Press.

Tagg, P. (1982). Analysing popular music: Theory, method and practice. Popular Music, 2, 37–67.

Tankian, S., Malakian, D., & Odadjian, S. (2025). Toxicity [Recorded by J. A. Fake Music]. Fake Music Records.

Taylor, A. R., Garvey, T., & Taylor, D. (2025). Walk My Walk [Recorded by B. Rust].

The Lost Tapes of the 27 Club. (2021). Drowned In The Sun [Recorded by Over the Bridge].

Ünal, R., & Taylan, A. (2025). AI covers: Listener perspectives on a controversial production model between technology, copyright and ethical violations. Ankara Üniversitesi İlef Dergisi, 12(2), 216–254.

Vear, C. (2021). Creative AI and musicking robots. Frontiers in Robotics and AI, 8, Article 634692.

Vear, C., Benford, S., Avila, J. M., & Moroz, S. (2023). Human-AI musicking: A framework for designing AI for music co-creativity. AIMC 2023. [Preprint]. Retrieved from https://aimc2023.pubpub.org/pub/zd46ltn3

Ventayen, R. J. (2026, February 11). AI music covers and their influence on listeners' acceptance and preferences. SSRN. Retrieved April 2026, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6143926

Weaver, A. (2025). Sisters, not twins: Exploring artistic control and anthropomorphism through composing with a bespoke generative AI. AI & Society, 41, 1609–1621.

Willmott, B. (2019, September 12). YACHT: Sail of the century. Electronic Sound. Retrieved April 2026, from https://www.electronicsound.co.uk/features/long-reads/yacht-sail-of-the-century/

Zellner, X. (2025, November 4). How many AI artists have debuted on Billboard's charts? Billboard. Retrieved April 2026, from https://www.billboard.com/lists/ai-artists-on-billboard-charts/unbound-music/