Data as Expression

Jaana Okulov

To cite this article: Jaana Okulov (2022) Data as Expression, Journal of Aesthetics and Phenomenology, 9:2, 191-207, DOI: 10.1080/20539320.2022.2150465

To link to this article: https://doi.org/10.1080/20539320.2022.2150465

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 21 Jun 2023.

Submit your article to this journal

Article views: 50

View related articles

View Crossmark data
Data as Expression
Jaana Okulov
The Department of Art and Media, Doctoral candidate, Aalto University

ABSTRACT
In machine learning literature, the concept of expression is seldom addressed as a general term in relation to information construction and data. The term more often carries a narrower meaning and refers to human bodily, emotional, or artistic dimensions that are recorded to train a model. However, this paper discusses a view in which expression is understood as any human sensory realm that can bring incidents explicit through the act of “pressing out,” as the early etymology of expression indicates. Also, all the sensory data that are used to train a machine learning model and the data the trained model gives as an output are considered meaningful through their expressive mediality—a form that is actively produced and can therefore be subjected to critical phenomenological analysis. This paper contrasts nonverbal expressions with categorical or linguistic expressions and asks, “What do eye movements express when they are used in training a machine learning model? What kind of expression arises in linguistic models? What could be considered aesthetic data (and what would it express)?”

For philosopher John Dewey, instinctive reactions in human behavior that for example exhibit mere discharge of an emotion should be separated from purposeful expressions. However, in machine learning, the two Deweyan positions (instinctive and intentional) collapse, as artificial expressions are purely simulations of learned logic. Therefore, the phenomenological question of this paper is traced back to the original stimulus and its human annotators. In this paper, philosopher Don Ihde’s experimental phenomenology explains how appearances can be attended to without subsuming them under any assumptions, whereas art theory—arising especially from philosopher Dieter Mersch’s thinking—provides an understanding of the mediality of expression. This paper introduces examples from machine learning that are not generally considered expressive but are used for regular tasks, such as object detection, and provides an alternative approach from aesthetics and artistic research that understands these modalities as expressive. If data are understood as expressive, it can be critically assessed how current machine learning models constitute knowledge.

1 Introduction
The question of expression is essential for artistic practices; it designates how a material is articulated through sculptural gestures, how a musical idea is shaped by the dancer’s
movements, and how phenomena appear from carefully assorted words in a text piece or performance. However, in machine learning literature, the more nuanced aspects of the concept often seem to be missing; the term is frequently used in a narrow sense when training and input data are discussed, for example, when certain facial expressions guide the detection of emotional categories in affective computing. The whole question of phenomena represented in the training data or by the trained machine learning model is seldom reflected through the concept of expression.

If expression were understood to have a stronger influence on knowledge production, the ethical dimensions of information would become accessible. The theoretical potential of the concept for machine learning becomes visible in the following questions: What kind of stimuli have led to the emergence of information? (Information is not born of nothing, there is always a stimulus causing the expression. This is well understood in artistic practices that emerge from a reflective approach). Through what kind of expressive medium is the stimulus translated into information? (Different modalities are expressive in different ways). How has someone chosen to express something as information – is intentionality a necessary condition for knowledge production? This paper acknowledges that interdisciplinary reflection is required to answer these questions in theory and in practice.

Expression can refer to a simple pattern appearing from sensory signals across a period of time, or it can signify an established association that an intelligent agent forms between incidents that are otherwise unrelated. Through expression, temporal and spatial relationships are crafted out of noise. In that sense, and in the classical expression theories formulated through the history of aesthetics by Eugène Véron (L’Esthétique, 1878), Leo Tolstoy (What Is Art?, 1898), Benedetto Croce (Estetica come scienza dell’espressione e linguistica generale, 1902), Curt Ducasse (The Philosophy of Art, 1929), and R. G. Collingwood (The Principles of Art, 1938), expression can refer to both what is perceived (observed and interpreted to appear) and what is explicit (expressed through a verbal description or other sensory modality in relation to the phenomenon). Therefore, the expression theory arising from the philosophy of art establishes two subjects in an aesthetic and artistic experience: the expresser and the recipient.

As this paper does not discuss information as an interactive element between two actors but rather as a knowledge-production process in a system (or an individual) related to machine learning, the two positions intertwine. The recipient and expresser become the two roles a human or an algorithm plays when translating stimuli into computable information. Human expressions are an essential part of supervised machine learning, and they shape information in a process called annotating. In annotating, a person provides a definition for a phenomenon they perceive, and this description is used to train a machine learning model. Similarly, when an algorithm forms associations between incidents in its training data to establish information, the expressive form of information is simultaneously created. In both examples, the original stimuli’s features collapse with their expressions. For example, a collage of visual features becomes a reductive interpretation—a label, a classification, or a category—that only partially represents the original phenomenon. Therefore, the question of expression should be central to information construction in any machine learning model, as it determines the model’s ability to draw information from stimuli.
Interestingly, a certain kind of objectivity is still assumed in the machine learning classifying practices. For example, when human annotating is used for object detection it might not be questioned whether the annotator can objectively express the essence of the stimulus, if the chosen label for the task is correct, or whether the data’s informative nature remains unchanged in the process.\textsuperscript{5} A critical attitude to the construction of information is a good starting point for anyone dealing with data in practice or thinking about data’s assumed objectivity in theory. In Claude Shannon’s foundational work on information, data communication systems are clarified through three technically developed steps that include a source of data, a communication channel, and a receiver.\textsuperscript{6} Although the theory introduces information as an engineering problem and addresses how a message can be identified from noise in a technical sense, it creates space for a third element between the expresser and the receiver, as specified earlier from an artistic perspective. In Shannon’s theory, the term medium is used to clarify what is meant by communication channels. For Shannon, medium is a carrier of information—for example, radio frequencies or a beam of light that can transmit a signal from a transmitter to a receiver.\textsuperscript{7} A medium always operates in the presence of noise. Therefore, the emergence of information in a system is relational to the uncertainty that is transmitted via the medium; for example, in the case of entropy, information is missing.\textsuperscript{8} Following this, in Shannon’s theory, information is relational to the capacity of a medium to translate a message. Information is not constituted by its possible semantics; it is the uncertainty and noise—the entropy-like expressions the medium carries that measure information.\textsuperscript{9} To shed light on the human aspect of an expressive medium, I will briefly turn to John Dewey’s thinking. Although Dewey is not the most well-known author on mediality, his theory clarifies the immediacy and intentionality that expressions carry in their mediating form in human experience. According to Dewey, the relationship between expression and the medium is intrinsic.\textsuperscript{10} For example, a smile functions both as the medium and the expression of joy only if the smile is intended; an unintentional act lacks a medium. From the standpoint of the one acting, they simply give away the experience; enraged person is raging and not expressing rage.\textsuperscript{11}

While artistic expression is not only tied to an artist expressing their inner emotions but also the carefully chosen and crafted materials determine expression for Dewey, according to theorist Krzysztof Guzalski, artworks have mainly been understood as mediators of the artist’s affective experiences in the history or the philosophy of art.\textsuperscript{12} Art was given an instrumental role in this debate when feelings were positioned external to the artworks. Although the instrumentalization of art fails to grant art its intrinsic value, from an interdisciplinary perspective concerning data, it is quite striking that art as a medium was inherently acknowledged as a subjective construct through which the artist’s internal attitudes, thoughts, and affections can become explicit. Artistic gestures enabling something to appear exist because someone has taken an ethical, bodily, and intellectual relation to a phenomenon. Dewey’s understanding of the role of the medium is useful in reflecting on the informative nature of data. The modality and choice of the medium in the data used in training a machine learning model is essential in shaping its expression.

Counterintuitively, from an artist’s perspective, in data society, data is sometimes given a neutral position. It can be referred to as the raw and low-level variables that must be interpreted to obtain meaningful information.\textsuperscript{13} Consequently, data is considered
a neutral and meaningless state of information\textsuperscript{14} and as data ethics researchers Miceli Milagros et al. describe data-related work is accepted neutral, despite its strong interpretative character resulting from human intervention.\textsuperscript{15} The belief about data’s neutrality assumes that it is not expressive. In titling this paper “Data as Expression,” I aim to challenge the question of data neutrality in relation to the concept of expression and arguing that data is always expressive. This provides data with an ethical position and the reader with an opportunity to unravel what is being expressed with data.

2 Data annotations as the expressive force of algorithms

Dewey discusses how energy is transformed into thoughtful action. Through this transformation, an act becomes expressive and does not remain a mere discharge of an emotion.\textsuperscript{16} Dewey does not consider unintended bursts of feelings or other physical impulses that are directly executed (such as sneezing) as expressive; expressive acts need time to be constructed. In relation to art and self-expression, he describes expression as “a prolonged interaction of something issuing from the self.”\textsuperscript{17} However, reflecting the Deweyan definition of expression in machine learning models, the two perspectives seem to collapse; artificial expressions are neither instinctive nor intentional, but an optimization of a learned logic. All human expressions, even those occurring without intention when training a machine learning model, are indifferent to a machine, statistically speaking; unintentional acts are algorithmically transformed into intentional ones, as both are “just” data for the model. This results in the inclusion of implicit biases in the training data that equally influence how the model behaves when in use.\textsuperscript{18} Although most studies on machine biases concentrate on supervised learning, in which human users label data, unsupervised methods—such as data clustering where data is grouped according to similarities in its expressive patterns without human intervention\textsuperscript{19}—also incorporate human biases.\textsuperscript{20} Therefore, data in general can be considered essentially expressive, because as artificial intelligence ethics researchers Simone Fabbrizzi et al. conclude, there are no bias-free datasets.\textsuperscript{21} To describe how information emerges from human influence, this section focuses on supervised data.

Miceli et al. explain how annotations are sense-making processes in machine learning, in which data are transformed into an understandable form for the machine.\textsuperscript{22} Annotations can be seen as a form of guidance for the machine learning model, so that it can perceive the world in a certain way, recognize features that resemble an algorithmically known face, or label objects with their assigned names. The annotation process begins by framing the media used to train the model. In the visual context, and more specifically in object detection, the annotating process often contains two steps: image segmenting and labeling. In segmenting, the annotator separates the recognized objects according to their locations; in labeling, the segmented objects are given names.\textsuperscript{23} While the stimulus used sets the initial boundaries for the information to arise as a learning output, annotations are reflections of those boundaries, or of yet another layer of boundaries that preserve the phenomenal aspect of the incident. If an image is reflected and annotated through a mode of perception that establishes only the object’s categories, after labeling it will no longer include descriptive information, for example, about the object’s aesthetics. To say that something is required from the data means that stimuli and their annotations are gathered to serve an expressive purpose.
The selected annotation categories are often explicitly determined by tech companies’ priorities\textsuperscript{24}. However, the process is also implicitly influenced by the values and beliefs of all actors in the process who, for example, have already decided what type of data is needed. Consequently, some information is acknowledged and included as data in the process, while other is silently excluded. According to the study of Miceli et al., data annotators perceive the labels provided by clients and reaffirmed by managers as correct and self-evident, and hardly ever question them, although the annotation process is known to follow current dominant worldviews and existing social hierarchies: By naming a phenomenon, its existence is reified, and the category is naturalized as something objective and neutral.\textsuperscript{25} Hence, data expresses and reinforces existing values, preconceptions, and hierarchies at the societal level through annotations. Therefore, data’s expressiveness has a concrete and direct impact on society when trained models are utilized for various perceptual tasks.

To give an example of two databases annotated with different mediums that expose completely different modes of expression, I will describe the aesthetics and contents of the IAPS (The International Affective Picture System)\textsuperscript{26} and CAT2000 (A Large Scale Fixation Dataset for Boosting Saliency Research)\textsuperscript{27} image datasets. The two datasets are used for different purposes and include different annotation methods. I chose the datasets randomly to demonstrate how phenomena can bleed from images into expressive data. The datasets are publicly available and can be more thoroughly and visually explored (IAPS can be downloaded upon request for research purposes) by the readers of this paper. However, as sharing these materials can compromise their value as a research stimuli, I will not include them here, but will do my best to describe their contents textually.

The IAPS dataset is used in psychological studies and for affective computing, and it consists of 9,941 images that include perceivable theme categories such as violence, death, food, abstract patterns, tools, animals, landscapes, porn, portraits of people, and different social situations.\textsuperscript{28} In comparison, the CAT2000 dataset has 4,000 images inside 20 folders, including categories such as cartoon, action, social, pattern, and random, used for saliency modeling to compute fixation maps, for example, for object detection.\textsuperscript{29} The aim of the first dataset is to arouse emotions, the second dataset seems to seek the expressivity of gaze in how the eyes move when gazing images of different categories. The datasets aesthetically reveal their distinct purposes; the IAPS dataset includes images showing dramatic cinematic scenes that are high in contrast and loud in affective expressiveness, placing the affective objects in the center. Conversely, the CAT2000 dataset exhibits a more neutral-seeming collage of images with more variety in their styles, resolutions, and framings, even including images that are upside down, jumbled, and randomly framed, as well as looking like they were taken by accident. In both cases, the different mediums used for the annotations can be described as expressive. In IAPS, annotators used verbal concept scales, such as “pleasant”/“unpleasant” and “calm”/“excited,” to express the affectivity of the images. Other documents in the IAPS zip file folder include, for example, brain imaging studies and physiological measurements that encompass other modalities considered expressive mediums for emotion. Unfortunately, and expectedly, the folder does not contain studies in which aesthetic expressions would have been used to denote the affective arousal caused by the images.
CAT2000 images have a distinguishable impact on how the gaze behaves in viewing: the eyes seek familiar faces, fixate reactively on disturbing instances, look for patterns, try to recognize objects, or to make sense of something vague, and wander aimlessly. The aesthetic and semantic contents in the images establish sensible rhythms for the gaze. In general, images annotated using gaze can be assumed through perceptual studies to express the cognitive mode one is adjusting when viewing images, and an object-based mode of perception, such as human vision, uses high-level representation and categories to adjust the gaze. If the perceiver is not a trained art viewer, they fixate mainly on the most semantically salient objects in a scene, while perceptually trained viewers use more global perceptual strategies. Images arouse certain kinds of gaze expressions that reflect the underlying thinking; therefore, if the task is not well controlled, the annotations are prone to show implicit biases caused by the viewers interpreting the most salient incidents in the images according to their background knowledge, assumptions, interests, history, and even psychological conditions. Expertise in artistic practices is in knowing what kind of sensations different stimuli induce and curating how the eyes move across a scene. Therefore, by reflecting on the free-viewing research protocols behind many saliency models, such as the one in the CAT2000 dataset, from an artist’s perspective and understanding individual influences on gaze, it is clear that there is nothing “free” in free-viewing. Curated images induce very distinct gaze expressions that are by no means neutral or objective.

Etymologically, annotations are situated in observations and remarks, but in machine learning, the annotators’ task is often to classify data by assigning meaning to its content using verbal labels. This is the case in the largest and one of the most widely used visual databases for object detection, ImageNet. ImageNet was created with human annotators labeling image content according to their semantic object-level categories. “We are currently in the biggest experiment of classification in human history,” as researcher Kate Crawford puts the huge demand and consumption of today’s tech industry for annotated data and specifies, those data have lifecycles that always reflect the current time. Therefore, the way stimuli are attended change the kind of information they can provide; if the task is to classify data, the outcome is an expression that results from a mode of perception that enables classifying. On a perceptual level, the expressive possibilities of annotations in these datasets are often perceptually reduced to categorical methods of world building. Consequently, it can even be argued that many current machine learning models express objectified phenomena.

In studies of perceptual psychology, this mode of perception can be called object-based, language-based, or category-based attention. In psychology, the use of the concept can be summarized as attention guided by developed concepts that are available for perception because the mind has learned of their linguistic existence. This results from a partial mode of attending that has reduced and merged featural, qualitative information into object representations to establish stable semantics. According to researchers Ophelia Deroy and Charles Spence, however, before infants grasp any linguistic concepts, they might perceive the world as deeply multimodal. This means that perception does not establish objects in the mind but resonates across all sensory dimensions, similar to synesthesia. The skill of suspending objectifying thought is partially retained in adulthood; however, if it is not practiced, its intervention in categorical thought weakens. I call this mode of attention that expresses other than
object-based perception aesthetic according to Dieter Mersch’s theory, and attempt to explain with arguments developed through art theory and my own artistic practice how data could exceed its object-based nature to reach dynamic phenomena with expression that exist beyond discursiveness. The next section describes this perceptual position.

3 Perceptual conditions for aesthetic expression

Philosopher Ladislav Hejdánek states how objective thinking ignores the non-objective nature of incidents that all thinking includes by ignoring, unreflecting, or even unseeing the non-objective nuances of phenomena. His term nepředmětný is translated as non-objective; however, it means “a distance taken toward the notion of object and toward the objectifying tendency of thought.” While Hejdánek locates the possibility of non-objectifiable events in theology and religious experiences, I argue that all events have the potential to emerge through a multitude of appearances; the prevailing attentional mode designates their nature. Or, as philosopher Don Ihde puts it, “Each appearance appears in a certain way, always relationally to my degree and type of attention,” hence “[E]ach seer sees what he already believes is ‘out there.’” For more object-oriented directions of postphenomenology this argument that objects appear only through a subjective human mind is anthropomorphic. However, for Ihde, there is no ontological priority for human beings, as we appear through objects (i.e., mediating technology) when embodying them. To clarify, although this paper is dedicated to the division of the object-based and the aesthetic, not only are just objectifying thought and aesthetic thought contesting it but other attentional modes can also induce an accumulation of alternative perspectives to a phenomenon. For example, with space-based attention a person perceives events through their spatiality and can locate them with eye movements.

With a phenomenological method—or postphenomenological, as Ihde calls it—to be discussed later in the section, the normative modes of perception can be broadened. In general, phenomenology begins with empirical observation directed at the whole field of possible experiential phenomena, which are observed with no predefining assumptions or beliefs. Ihde’s postphenomenology is a pragmatically enriched phenomenology, a specific perception technique that invites phenomenological variations from incidents. However, as this paper’s aim is to create room for “other thought” as a perceptual alternative for those modes of attention that many machine learning models exhibit, leaning toward the aesthetic aspects of perception supports this goal. I will begin by positioning my thoughts near psychology’s discoveries and argue that aesthetic attention resembles the infant’s ability to attend to the world multimodally, in such a way that perceptual phenomena can arise from featural and affective information in the absence of objects. Aesthetics establish the foundation for knowledge; being able to investigate sensory phenomena and their possibilities through their relationships to the body creates an essential space for linguistic communication. Furthermore, language itself emerges from aesthetic investigations, where sounds are affectively, bodily, and informationally tested and adjusted. According to Dewey, the utterances of a baby can be considered expressive exactly when they become intentional. However, I consider the probing of possibilities as expressive and deeply essential for aesthetics. I hypothesize that aesthetic perception is an evolutionally old and essential mechanism not for just humans but for all organisms, as it informs the body about its expressive potential related to the stimuli.
Mersch describes similar positioning as a reflexive aesthetic thought that is “other thought” or “other than thought” and differentiates it from discursive knowledge of argumentations and propositional knowledge of statements. According to Mersch, aesthetics brings “attentiveness to nuances, to details, to fragile and often overlooked marginal perceptual phenomena and their vibrations.” For Dewey, thinking directly in terms of colors, tones, and images differs operationally from thinking in words, as asking about what phenomena mean in the sense of language “is to deny their distinctive existence.” In Dewey’s theory, aesthetic thought is immediately linked to the expression of meaning with visible and audible qualities. Similarly, Mersch discusses how artistic practices can make something appear. However, unlike Dewey, Mersch argues that in art the medium does not expose but exhibits phenomena as art’s modus is in showing. Therefore, the knowledge that emerges with aesthetic reflexivity leads to the capturing of a phenomenon in its medium, a medium that “reflects the perceivable through perception and the experiential through experience.” To allow Mersch’s words to resonate with my own, the expressive data resulting from aesthetic attention is a form of practice beyond discursiveness and objectifying language. While Hejdaček’s, Mersch’s and Dewey’s theories provide ways to understand what constitutes aesthetic attention and expression, Ihde’s phenomenological method and hermeneutic rules describe how one can adopt a perceptual mode (or perceptual modes, plural?) that reveal incidents through the multitude of their appearances, from openness to phenomena variations. However, to discuss them as perceptual conditions for aesthetic expression, some of Ihde’s steps require extra contemplation to pivot toward the aesthetic theories of Mersch and Dewey; I will try to do this without losing the essence of either theorist’s thought.

According to Ihde’s first hermeneutical rule, phenomena should be attended as they appear. The phenomenological method acknowledges the multiplicity of available perceptual phenomena and works as a device to discover them while understanding that the variations of phenomena cannot be exhausted. As mentioned, aesthetic phenomena should not be overruled by objectifying language; when attending to a multitude of phenomena aesthetically, determining incidents with static verbal labels should be avoided. This links to Ihde’s second hermeneutical rule: “Describe, do not explain.” For Ihde, to explain means to create a theory, idea, concept, or construction that attempts to go beyond phenomena. Although Ihde does not directly call linguistic categorizations explanations, he states that beliefs tend to reduce and limit the full range of appearances. I argue that objectifying thought does the same: it establishes presumptive categories that prevent other aspects of phenomena from being seen.

With the third hermeneutical rule—“horizontalize or equalize all immediate phenomena”—any beliefs about incidents are suspended, which also avoids arranging the observations into a hierarchy. Only those aspects of the phenomena that are immediate are observed carefully. In an aesthetic sense, this is an attentional mode that opens the experiencer to qualities and features of experience that resonate with them, or what Mersch would consider creating the oscillation of the phenomena. The fourth hermeneutical rule states, “seek out structural or invariant features of the phenomenon.” The structural features and emerging repetitive patterns of phenomena constitute their appearances; in aesthetics and artistic practices, this can mean, for example, investigating the incident’s visual possibilities in how it resonates multimodally, exploring auditory dimensions with sound, or mapping bodily phenomena through dance. An
investigatory relationship with phenomena is essential both for experimental phenomenology and for aesthetic inquiries leading to aesthetic expression. The next section will further explain how data can attain an aesthetic dimension.

4 Aesthetic data

In art and art theory, the concept of notation is more often in use than its etymological sibling annotation. Notations often refer to a kind of nonverbal translation in which an event is represented by symbols or signs. Musical notation is the best-known system of this kind; it translates musically informative features, such as pitch and tempo, into visual notes. Fontana Mix, an experimental music score by composer John Cage, is an interesting exception to classical music notation. Its aim is not to represent music; instead, the piece can be played in a number of different ways (baroque music notation, too, was less specific and left more room for individual expression). It brings an element of chance to the score. The score is annotated with randomly distributed points and a grid on a transparent film that is placed on top of curved line drawings each time it is used to change how the piece is performed (for example, pitch, tone, and volume change accordingly). The work’s notational aim comes close to the philosophy of Ihde’s hermeneutic rules: performing the piece gives birth to variations of phenomena rather than to static definitions.

A more structured notational method, similar to musical notation, was developed for dance by Rudolf Laban. In Labanotation, a vertical symbolic system is used to document the movements of a dancer/s using visual logic; symbols and their locations relative to the central line of the stave describe different body parts, their directions, flow, dynamics, and inner attitude, which refers to a movement’s expressive or functional value. While musical notations leave more room for individual expression of the scripted notes, Labanotation allows the expressive intentions to be captured in more detail. They are noted in colors to describe whether a movement is executed in a high or low position—accents that show the momentary increase or decrease in speed—and whether the movement makes a sound. Symbols mark the gravity of a movement and its release, and wavy lines describe vibrations with different tensions. Dance notation and movement analyst Ann Hutchinson Guest describes labanotation as a language in which the grammar defines the relationship of the movement words to each other, while the words serve a given function in a movement sentence. The basic elements of this language are nouns, verbs, and adverbs. Curator and writer Henrik Folkerts makes a similar claim in his essay “Keeping Score: Notation, Embodiment, and Liveness” calls notational devices languages that connect the material and disciplinary knowledge of performance, architecture, linguistics, or mathematics. Notations produce description, transmission, and signification to be enacted, or executed. It seems that the core of these nonverbal languages is their ability to record an event in a form that can be read, studied, and later expressed again. However, what should be highlighted concerning aesthetic expression is that the notational medium itself may not carry the original quality of the phenomenon but functions as a trace. These kinds of notations provide access to data processing on a cognitive level but should not be characterized as aesthetic data because phenomena themselves have been arrested in the process.

When a visual character denotes a hand, it is equal to the semantic label “hand” in verbal language. Both representations resemble the object-based nature of the phenomenon,
although their mediums differ. Thus, there isn’t always a significant difference in how different mediums carry information; more essential is what kind of meanings they convey. Although I have defended the view where linguistic labels fail in representing phenomena holistically, this is where I form a counterargument to further specify what I mean: It is not the choice of medium that constitutes “the aesthetics” of aesthetic data, but (following Mersch’s aesthetic theory) an expressive form that can maintain phenomena in its medium. Languages, whether visual, verbal or any other kind, can fail or achieve this through the modes of perception they mobilize. For a person who can both dance and read notations, a symbol that was previously a static label of a body part moving in a certain way can become aesthetic data if the quality of the note combines with the aesthetic quality of the real dance gesture. When a symbol is perceived, the body executes the act (the symbol affects the body by invoking its movement potential; in other words, the symbol is embodied) without the need to further cognitively process it.

The capacity of the human body to embody a stimulus to establish aesthetic data could be represented as follows: the graph below shows how a stimulus is annotated nonverbally. The notation begins by choosing stimuli. In Figure 1, the chosen stimulus is a color change, but it can be any phenomenon. When the stimulus is seen, a bodily sensation arises that describes how the body attunes to the perception of that stimulus; its energy can rise or drop, and the muscles can stiffen or relax. This sensation is taken as an affective motif for the expression performed in a different medium. In this example, the expressive modalities are vocalizations and line drawings; however, any modality, such as the gaze or bodily movement, can be relevant. As a note about the influence of the medium, if a stimulus is changed—for example, to dynamic sounds—then the perceptual

Figure 1.
modality would switch to hearing accordingly. Multimodal expressions and their temporal/spatial relations to stimuli can be saved as aesthetic data. If digitally produced and processed, this data can be used for aesthetic or affective computing; it enables the quantifying qualia without the intervention of categorical language.

To further develop this, I will turn back to the question of language and ask: what languages can avoid object-centeredness to bring forth phenomena’s dynamic natures (in the example above, expressive language took an artistic form; more specifically, the sensations were expressed in auditive and visual languages), and can some verbal labels be considered expressive in an aesthetic sense? Asemic writing is a visual genre of poetry that lacks semantic content. It suggests a language and can be understood as writing, but forces the reader to “fail” to understand it semantically as it has no verbal form. Asemic writing can have letterlike symbols and lines arranged similar to writing, but according to writer and artist Michael Jacobson, it is intentionally illegible, abstract, and wordless. Poet Geof Huth insightfully unravels what is retained when the semantics of a language are lost: “Writing does not just contain semantic information. It also contains aesthetic information (when seen as a shape or image) and emotional information (such as a graphologist would analyze).” Therefore, when asemic writing eliminates linguistic semantics, it brings the emotional and aesthetic content to the foreground. Asemic writing shows in an interesting way how language continues to exist and be readable in other ways even after its semantic content has been emptied. Although Huth differentiates between aesthetic and affective content, in many cases they appear simultaneously and rely on shared expressive forms; a joyfully bouncing line is the phenomenon’s aesthetic and affective appearance. Therefore, removing linguistic semantics amplifies its other qualities, which activate not the mind’s cognitive processing but its aesthetic and affective imagination.

To perceive aesthetically means to be open to the qualities of an experience and to let them resonate in the body. Consequently, aesthetic data can be read as having embodied characteristics—as discussed at the beginning of this section. In neurological studies, some words that describe actions have been noted to induce bodily attunement; words such as “grasp” or “reach” which activate the motor cortex similarly to its activation while executing these gestures although their actual, complete execution would be inhibited. It can therefore be argued that the kinds of words that contain action potential concretely resonate in the body and succeed in maintaining the dynamics of the original phenomenon in a verbal medium. Can these verbal descriptions of dynamic words be considered expressive in the aesthetic sense? I would say that they cannot, at least not alone, as for example, labels in machine learning. However, in a multimodal context, they could function as embodied phenomena when they cause a primitive urge to act in relation to appearances and impressions. What could describe aesthetic data better than an urge arising from a stimulus?

Let me offer another example from my artistic experiments from 2017 to expand on how verbal descriptions can affect phenomena in practice. I was working with a simple color program on my computer where a randomly chosen digital color filled the whole display and tilted slowly toward another randomly chosen color, for example, from dark red to bright yellow or from light pink to a low-saturation blue in a smooth rhythm. After the color change, the program asked me to write down a word that would describe the phenomenon. I saved the written words, and the program continued to change the colors. At the beginning of the task, it was confusing and hard for me to describe this phenomenon with words when it lacked a definition. However, after a while, I began to
notice a logic in what I was seeing; I saw repetitive affective patterns in alternating color combinations. Although the colors and their pairings were randomly chosen by the computer and were the same only by chance, my descriptions began to repeat themselves more frequently. The words started to feel comfortable and even suitable for describing the phenomenon. Suddenly, I fell into a meditative flow state with this simple task.

The next day, I was eager to continue the practice but noticed that something was different—as if the phenomenon had died. I did not know what had changed, but the phenomenon no longer had any aesthetic impact on me, and felt obtrusively static; no flow state followed, only disappointment. When I reviewed the semantics of my written notes, I noticed something. In the flow state, I had written words in a very specific form. I had described the color changes as “kohoava” (Finnish and means “rising”), “kihtyvä” (Finnish and means “accelerating”), and “sulkeutuva” (Finnish and means “closing”), which is an active form and the VA-partisiipi (Finnish definition for present participle) of the word, while the next day the words had turned into “kohonnut” (Finnish and means “risen”), “kihtynyt” (Finnish and means “accelerated”), and “sulkeutunut” (Finnish and means “closed”), existing in their NUT-partisiipi (Finnish definition for past participle) form. It was as if some learning had occurred during the night, and I had determined the categorical definitions for the color phenomena. Static labels replaced the words that had previously been dynamic and affective, and I sensed in my body how the words no longer moved me. I believe that by defining the phenomenon and objectifying it with a certain form of language, I lost its affective and aesthetic quality.

During the task, another aspect also became obvious to me. The aesthetic dynamics were not tied to any specific static color (here the reader can associatively imagine a static color label such as “red” as a verbal rendition that leads to an objectified phenomenon) but existed relationally between the colors and caused sensations by arousing the body’s energetic state or by loosening its tension (instead of a static definition, the quale “redness” can be seen as a verbal form that can exist between different wavelength positions). It seemed that these affective features were somehow covert, but constituted the aesthetics of the color phenomenon from a strange intermediate position. According to researchers Shreyan Chowdhury and Gerhard Widmer, mid-level features can indeed bridge the semantic gap between low-level features and high-level descriptors, such as emotion. They are qualities that can be described with words such as “cold,” “dynamic,” “passionate,” “gentle,” “mechanical,” or “delicate.” Interestingly, using these mid-level features in machine learning can broaden the expressive possibilities of the machine. If a machine reads and plays music using music scores as their data, their expressions are perceived as static and “unexpressive,” whereas if they have been trained with data of individuals playing the music in relation to the scores labeled according to mid-level features, the learned expressivity of the machine can fool even human listeners into thinking the song has been performed by a human. It seems that aesthetic expressivity might lie in mid-level features. They can be qualities that bind low-level features together with an affective logic to establish semantic meanings during neurological processing but not yet arriving at the high cognitive end of semantic categorization. I would therefore suggest that the aesthetics that exist as mid-level features are phenomena enabled by low-level features through their shared affective relationality. Artistic articulations can directly express these qualitative features of phenomena through a universal, yet not fully comprehensively mapped, multimodal logic. Artistic expression exhibits expressing, and this is how, I argue, that data can have an aesthetic form.
5 Conclusion

Aesthetics, phenomenology, and artistic practices could play stronger roles in reshaping modern technologies, as they enable understanding information as a subjective construct. This paper offers expression as an auxiliary concept to relocate the discussion around ethical computing toward information construction from stimuli. By reflecting the expressiveness of machine learning models, it becomes possible to understand what kinds of stimuli cause the emergence of information, how an expressive medium influences its formation, and what the role of annotating is in what is represented as data. A label, classification, or category is always a reductive interpretation that can only partially encapsulate the original phenomenon, which is why the assumed objectivity of data should be critically reflected upon. This paper offers aesthetic thought as an alternative to a categorical mode of perception that is strongly present in current models of machine perception. Derived from Mersch’s theories, aesthetic thought can be understood as attentiveness to nuances and other overlooked marginal perceptual phenomena. Aesthetic attention emerges multimodally from the absence of objects, from featural and affective information, and can, in this sense, establish the foundation of knowledge, as it precedes high-level cognitive processing that uses categorical logic. While human annotations used as supervised data for machine learning often arise from associations between categorical language and other modalities, artistic notations can be established nonverbally, for example between sounds and dance gestures. Although these associations are not generally understood as information, this paper argues that they form a logic that can maintain the original phenomena in its medium. Ihde’s hermeneutic rules provide a way to understand how normative modes of perception can be broadened to establish this kind of nonverbal information. With experimental attentional techniques, perceptual variations are invited from phenomena. Central to this kind of attention is expression; when a knowledge production process occurs, no categorical language can be used, rather incidents are expressed in terms of colors, tones, and gestures resulting in aesthetic data, which is knowledge beyond objectifying language.

Notes

1. Facial, EEG (electroencephalogram) and voice are the most popular modalities for emotion detection in machine learning. However other physiological signals such as electrocardiogram (ECG), electromyogram (EMG), facial electromyogram (FEMG), and galvanic skin response (GSR) are expressive in human affective experience. See Maithri et al., “Automated Emotion Recognition” for a review on current methods of emotion recognition through expressive modalities.
2. The classical expression theory according to Guzalski’s article, “Henryk Elzenberg.” See also Spackman, “Expression Theory of Art” for a review.
3. Miceli et al., “Between Subjectivity and Imposition.”
4. One example of annotating is the validation process in which a user is asked to prove that they are a human by choosing all the parts of an image that include a traffic light. Von Ahn et al., “ReCAPTCHA.”
5. Miceli’s et al., “Between Subjectivity and Imposition” article discusses the politics of annotating in depth.
7. The concept of medium is mentioned directly only when transmitting channels are discussed. Shannon, "A Mathematical Theory of Communication," 2; Shannon and Weaver, The Mathematical Theory of Communication, 34.


10. Dewey, Art as Experience, 64.

11. Ibid., 61.


13. See Chen et al., “Data, Information, and Knowledge” for a review of the definitions of information and data.


15. Miceli et al., “Between Subjectivity and Imposition.”


17. Ibid., 65.


19. Assent, “Clustering High Dimensional Data.”


22. Miceli et al., “Between Subjectivity and Imposition.”

23. Ibid.

24. Ibid.

25. Ibid.

26. Lang et al., “International Affective Picture System.”

27. Borji and Itti, “Cat2000.”

28. Lang et al., “International Affective Picture System.”


30. The early studies of psychologist Alfred L. Yarbus revealed that people use different perceptual strategies according to what they are looking for in an image. In the case of the Cat2000 dataset and its saliency maps, which were gathered in a free-viewing research setting, I argue that people still alter their perceptual strategies, although they are not explicitly stated. Different perceptual strategies lead to altered subjective perception. Yarbus, Eye Movements and Vision.

31. Hochstein and Ahissar, “View from the Top.”


33. Okulov, "Machine Attention.”

34. Miceli et al., "Between Subjectivity and Imposition.”


40. Deroy and Spence, “Are We All Born Synaesthetic.”

41. Ibid.

42. See Mersch, Epistemologies of Aesthetics.

43. Hejdánek, Nepředmětnost v myšlení; Miroslav, Philosophy en Noir.


45. Ihde, Experimental Phenomenology, 21, 30.

46. Lally, “Post-Phenomenology, Transduction, and Speculative Fabulations.”

47. Ihde, Bodies in Technology.

48. Ihde, Experimental Phenomenology, 16.

49. Ibid., xv.


52. Ibid., 52.

54. Ibid., 74.
56. Ibid., 46
57. The terms aesthetic data and expression are mine, but they are located in relation to Mersch’s thinking. Mersch, *Epistemologies of Aesthetics*, 48.
59. Ibid., 18.
60. Ibid., 23.
61. Ibid., 18.
62. Ibid., 20.
64. Ihde, *Experimental Phenomenology*, 22.
65. Ibid., 108
66. Cage “Fontana Mix.”
67. Ibid.
68. Hutchinson Guest, *Labanotation*.
69. Ibid.
70. Ibid., 14.
71. Folkerts, “Keeping Score.”
72. Ibid.
73. De Villo, “Definition of Asemic Writing.”
75. Jacobson. [untitled]
76. Huth, “Aemia Becomes You.”
77. Rizzolatti and Fabbri-Destro, “Mirror Neurons.”
80. Ibid.
81. Ibid.
82. Widmer, “Computer Accompanist.”

**Acknowledgments**

This work was supported by the KAUTE Foundation (The Finnish Science Foundation for Economics and Technology).

**Disclosure statement**

This author has no potential conflicts of interest to disclose.

**Notes on contributor**

*Jaana Okulov* is an interdisciplinary artist and a doctoral student at Aalto University in the Department of Art and Media. Okulov received their MFA in the Finnish Academy of Fine Arts in 2016 and completed basic and intermediate studies in psychology in 2017. Their research concentrates on the theme of attention in humans and machines and combines theoretical and empirical approaches from perceptual psychology, artistic research, aesthetics, and machine learning.
Bibliography


Okulov, J. “Machine Attention.” The 9th biennial Surveillance & Society conference of the Surveillance Studies Network, Erasmus University, Rotterdam, June 1, 2022


