Thiago Hersan, Cássia Hosni, Giselle Beiguelman, *Meta-Acervos* and New Interpretations Emerging from Error

In October 2025, the FAPESP Thematic Project "Digital Collections and Research", an initiative aimed at developing methodologies and a theoretical-conceptual framework for museum collections and digital documents, launched the *Meta-Acervos* platform. Using computer vision models and techniques, a subfield of Artificial Intelligence research primarily concerned with the interpretation and extraction of semantic information from visual data, the platform functions as a navigation interface for Brazilian museums. It articulates a system for the analysis, processing, and organization of open-access image collections, with a specific emphasis on drawings and paintings. By applying AI to existing online data, the interface enables filtering by institution, technique (drawing or painting), chronology, and biological elements (including search terms related to fauna, flora and human forms). Navigation allows for the visualization of works through timelines, chromatic scales, or latent space, as well as algorithmically-curated selections defined by clustering procedures. Users may also generate search results as mosaics or compositions for archiving purposes. For a more detailed description of the system, the algorithms employed and its interface, refer to: *Meta-Acervos*: A Navigator for Online Museums.

In this visual essay for the issue dedicated to Flusserian thought and artificial intelligence, we highlight specific aspects of the development of *Meta-Acervos* within the realm of technical images, focusing on the system's interpretive errors and the potential for new readings suggested by its outputs.

While trees or human silhouettes yield higher accuracy in the platform's search results, figures and objects that do not conform to patterns recognized by the models undergo machinic interpretations, and are associated with parameters that assign meanings that diverge from the user's original intent. These search and visualization results reveal the fragility of AI models when dealing with modern and contemporary art, particularly, but not limited to, non-figurative works. Such works are processed by the statistical calculations of the apparatus and assigned new meanings shaped by the biases inherent in the datasets used to train these models and by the ways in which the models synthesize associations within the data.

The figures presented in this visual essay are the result of searching for works containing the objects "ox", "hands", "palm tree", and "birds', using the Meta-Acervos interface. The specific works used in each of these compositions were selected to present, in a concise manner, a sample of the variety of styles and periods analyzed by the tool. The images were placed side by side, in

chronological order, in order to make the errors explicit, qualify their magnitude, and highlight the correlation and tension between the works, particularly between figurative and abstract ones.
For instance, when queried for "ox", the system identifies the torso in Domingos Sequeira's *Retrato de D. João VI, rei de Portugal* (1802) as the animal. Similarly, searches for "hands" reveal an epaulette on the figure of Dom Pedro I, and not infrequently, feet. Such is the case with *Nu Deitado*, by Anita Malfatti (1925), making evident that the apparatus treats the extremities of the human body as variations within a single morphological repertoire.

Among images categorized under "palm trees" the interpretation of Luiz Rodolfo Annes's figurative drawing stands out, wherein the pubic region is misread as a potential tropical plant. When the fauna and flora filter is used to find works with "birds," *Meta-Acervos* produces a kind of semantic metamorphosis, transforming objects composed of circles and lines into birds, correlating with flight any figure with a center and suspended stem, as seen in the work by Honoré Marius Bérard (1914).

What becomes evident in this processing is the model's mode of reading: the descriptor does not merely query an archive of predefined visual descriptions, but operates as a mechanism that displaces graphic fragments into a probabilistic grammar of forms.
Technically, this effect is linked to the adoption of OWLv2, an open-vocabulary model that operates in a zero-shot regime, articulating, on the one hand, the proposal of "object regions" in the image and, on the other, the verification of compatibility between these regions and the queried term. From the standpoint of interpretation, however, the result cannot be reduced to mere "error" since the model (here understood as an apparatus) installs a third term between the user's intention and the represented content. In a sense, it institutes a mediating instance that, while classifying, simultaneously rewrites the symbolic.

Despite not having written explicitly about artificial intelligence systems as we know them today, there are two main concepts from Flusser's works that we think are pertinent when considering the current landscape of technical image production, perception and reception afforded by these technologies and illustrated here through our exploration of *Meta-Acervos*.
Categorization as technical image: The systematic processes of tagging, indexing, cataloging, and describing image collections in scale generate their own set of technical images. Like photocopies of technical drawings and diagrams, but operating at a different scale, the classification and visualization of images through these programmed systems produce images that are themselves "computations of concepts". Just like "what appears in the photograph are the categories of the camera which ensnare the cultural conditions like a net with a limited view through its mesh" (Flusser 2013), what appears here in these annotations and classifications are also situated cultural artifacts of the apparatus that created them.

Premature decodification: humans are not the only ones in the business of "demonstrating the significance of a symbol" anymore. Whether these machinic predictive interpretations are correct, consistent or productive is another story altogether, and perhaps even irrelevant here. What matters is the presence of this new agent that quite literally, in the most Flusserian of terms, *informs* meaning by "creating improbable combinations of elements and imprinting them upon objects" (Flusser 2013).

In this way, a Flusserian reading of these systems as apparatus already always highlights their potential to act against their programmed intent. By "playing against the program", these misreadings, along with the images that they create, do not conform to an indexical regime of evidence, but rather to something closer to art, with its procedures of reconditioning and reworking what is already given. The errors enable alternative readings that go against expectations of certainty established by the computer vision apparatus, and also against traditional hierarchies and categories of art history, creating an opening for non-functionary uses of the technology.

What matters in this discussion, therefore, is not the proposition of an increasingly correct and precise system of taxonomy (even though refinements are possible), but the recognition that *Meta-Acervos* makes visible the distances and openings between systems of programmed classification and human interpretation. Errors in search results (such as birds where there are graphic marks, or feet where hands were sought) make explicit a form that computer vision imposes on contemporary images. Which traits are imposed by and within this apparatus? Which details are discarded or valorized? It is in this friction, where the user must decide whether to follow or resist the apparatus's suggestion, that a critical field opens up. A field less obvious than the rhetoric of supposed algorithmic "precision" and that offers the possibility of diverting, tensioning, and reprogramming the very regime of visibility that sustains automated cataloging. In this regard, our visual essay exposes the rules of a computer vision system by highlighting its errors and assigning to it a quality of ambiguity and interpretation not predicted by its program. From a Flusserian perspective, this "smuggling of human intention" into the program, and the act of playing with information that is not predicted within its logic, point to larger questions of freedom, or what we today might call agency, in a world dominated by apparatuses (Flusser 2013).

Acknowledgements

This study was financed by the São Paulo Research Foundation (FAPESP), Brasil. Process Number 2022/05946-9 and 2023/16018-8.

The opinions, hypotheses, conclusions and recommendations expressed in this material are the responsibility of the authors and do not necessarily reflect the views of FAPESP.

Note: While ChatGPT Pro (OpenAI) was used to support the translation and revision of the text, the authors are fully responsible for the content and ideas expressed.

Reference

Flusser, V. (2013). Towards a Philosophy of Photography. London: Reaktion Books.