Human Systems Engineering and Design (IHSED 2026), Vol. 237, 2026, 253–261
https://doi.org/10.54941/ahfe1008165

Graphic Design History Lab: Teaching Critical AI Literacy Through Archival Practice in Design Education

Valentina Nitti[1] and Elena Cavallin[2]

[1]University Iuav of Venice, Santa Croce, 191, 30135 Venezia VE, Italy
[2]Polytechnic of Turin, Corso Duca degli Abruzzi, 24, 10129 Torino TO, Italy

ABSTRACT

This article reports on an experimental seminar held within the Master's Degree Programme in Product, Communication and Interior Design at University Iuav of Venice, conceived as an initial effort to formalise an AI-assisted procedure for cataloguing graphic design artefacts. Forty-seven students catalogued a corpus of computing and technology magazine covers spanning the 1970s through the 2000s, using a web platform that queries AI models via API. The seminar opened with a theoretical lecture introducing the conceptual tools for visual analysis, after which students built a cataloguing schema by hand. This sequence addresses a specific disciplinary problem: graphic design lacks both a systematic methodology for visual analysis and cataloguing standards adequate to its materials. AI-driven completion of the schema followed, testing whether the tool was viable in this setting. Pre/post Likert questionnaires and the resulting CSV files reveal three notable effects: students saw their initial scepticism about unrevised AI output confirmed in practice; they recalibrated their self-assessment around the distinction between description and interpretation; and they came to recognise, in retrospect, the critical value of having built the schema manually beforehand. The findings suggest that this sequence offers a replicable pedagogical model for introducing computational tools into humanistic disciplines that lack established methodological infrastructures.

Keywords: Graphic design history, AI for the humanities, Cataloguing, Human–AI collaboration, Digital humanities

INTRODUCTION

Graphic design history still occupies an unsettled position within the humanities. Since entering Anglophone academic debate in the 1970s as a branch of the broader field of design history, it has had to build its disciplinary legitimacy through ongoing dialogue with neighbouring fields such as art history, architectural history, and the history of material culture (Woodham, 1995; Fallan, 2010; Margolin, 1992). In most institutions, graphic design history remains at the margins of practical training, treated as a complement to professional education rather than as an autonomous field of inquiry (Poynor, 2011; Triggs, 2011). One consequence is the absence of a developed analytical framework for visual artefacts, an absence the field has yet to address systematically.

Received May 7, 2026 Revised June 13, 2026 Accepted June 24, 2026 Available online September 25, 2026

Against this backdrop, the difficulty of working with broad and heterogeneous archival corpora, often without adequate cataloguing infrastructure, becomes especially clear. The problem is sharply visible in the work of Master's students approaching historical research for the first time, when they have to build inventories of magazine covers, posters, or ephemeral printed material that share no descriptive standards. From this situation arose the question that motivated our seminar: is it possible to develop a method for analysing visual artefacts in graphic design history?

Recent technological developments make the question more specific: can computational tools be introduced into a discipline that does not yet have a systematic methodology of its own?

The digital humanities have already shown that computational methods can open new perspectives on visual culture (Wevers & Smits, 2020; Manovich, 2020). Computer vision and vision-language models (VLMs) now make it possible to analyse corpora too large for close reading. The potential is particularly relevant for graphic design history, which works on serial visual materials organised diachronically, sequences of magazine covers being a prime example. Yet using these tools raises questions that are not only technical but also epistemological: what can AI observe in a graphic artefact, and, more importantly, what can it not understand?

This article describes a seminar held within the Master's Degree Programme in Product, Communication and Interior Design at University Iuav of Venice, in which students of graphic design history were introduced to AI models for the automatic cataloguing of computing and technology magazine covers from the 1970s to the early 2000s. The aim was to develop skills in analysing graphic artefacts alongside a domain-specific critical AI literacy (Long & Magerko, 2020): the ability to evaluate what computational tools can and cannot do in a field whose methodological foundations are still being defined.

RELATED WORKS

Most computational studies of visual artefacts come out of art history and visual studies. Two strands of research are particularly relevant here, both of them attentive to serial artefacts and to how those artefacts change over time.

The first is the *distant viewing* programme developed by Arnold and Tilton (2019; 2023). Their work on film posters and large photographic collections sets out a methodological framework that alternates large-scale computational analysis with close critical reading, and explicitly articulates the line between what computer vision systems can detect and what requires human interpretation. The second is *Cultural Analytics*, developed by Manovich and colleagues (2020). Their analyses of the Soviet photography magazine *Sovetskoe Foto* and of *Time Magazine* (Manovich & Douglass, 2009; Tifentale, Manovich, & Indaco, 2018) trace patterns of visual transformation across long time spans. For graphic design specifically, however, no comparable studies appear to exist, only scattered, unsystematic attempts.[1]

[1]A precedent that begins to undertake this work is the study conducted at University Iuav of Venice in collaboration with the Levi Foundation (Lancia, 2025), which tested the AI-assisted cataloguing of nineteenth-century musical score covers, reaching an accuracy of 80–85% in the extraction of structured fields such as publisher, composer, and place of publication. This result demonstrated the possibility of structuring tools and methodologies in combination with GenAI.

QUESTIONS ON GRAPHIC DESIGN ARTIFACTS

Before turning to cataloguing, it is worth clarifying what we mean by an artefact in graphic design. By graphic artefact, or, as Anceschi (1988) calls it, communicative artefact, we mean an "intentional product, [which] performs functions (whether considered as responses to needs, as use values, as solutions to problems, or, more neutrally, as modalities or opportunities of functions)" (Anceschi, 1988, p. 23). These are not just any objects, but materials produced for specific purposes of transmission and meaning-making. Graphic artefacts also carry an intrinsic ambiguity: they can be approached from different perspectives depending on the discipline that studies them. The book is a telling example. For literary history, it is the container of a text whose interest lies in content, authorship, and circulation. For product design history, it is studied as a manufactured object (D'Ambrosio, Grimaldi & Lenza, 1983). For graphic design history, it is studied through its cover and layout: through the visual organisation of the page, the relationship between text and image, the typography, and the grid system that structures reading. These three perspectives on the same object, each with its own approach and questions, together define the plural identity of the graphic artefact.

Problems of Ambiguity

This ambiguity is reflected in the difficulty of cataloguing graphic artefacts within the domain of cultural heritage preservation. Existing institutional tools do not produce records that meet the needs of researchers in graphic design. The ICCD (Istituto Centrale per il Catalogo e la Documentazione) records were designed for objects such as paintings, sculptures, and architectural artefacts, objects that can be described through categories such as identified author, definite location, documented provenance, and verifiable state of conservation. None of these assumptions transfers directly to graphic design artefacts, which are mostly anonymous or collective, produced serially, without a stable location, and often without process documentation. They therefore pose challenges that traditional archival institutions are not equipped to address (De Smet & De Bondt, 2012).

Systematising Analysis

In the absence of established norms for analysing graphic artefacts, and given the need to give students a critical lens, we built a theoretical foundation rooted in Jeremy Aynsley's observation: "Illustration and typography had long and interesting histories of their own, but in the nineteenth century these had been combined, in many instances prompted by [the] changing function of word and image within an industrialised society. The poster, combining word and image as a visual form, as well as mass-circulation magazines, newspapers and books all required new kinds of organization and 'design' for print" (Aynsley, 2012, p. 25). On this view, the graphic artefact is the spatial composition of writing and image, and so can be broken down into individual elements that can be isolated and analysed. Sergio Polano's studies (n.d.) align with this view: he locates both image and writing within the domain of graphics, dividing them in turn into "artificial" and "natural" categories. Alongside this taxonomy of elements, and on the axis of interpretive depth,

lies the iconological hierarchy of Erwin Panofsky (1939), which reads the image across three progressive levels: the *pre-iconographic*, concerned with literal description of the visual elements present; the *iconographic*, concerned with the cultural decoding of shared symbols and references; and the *iconological*, concerned with deep, contextual interpretation rooted in the historical structures of thought of an era (Yadav et al., 2025). As Yadav and colleagues note, computer vision models operate reliably only at the first, pre-iconographic level; the cultural competence required at the iconographic and iconological levels remains beyond them. This was the analytical framework presented to students as the basis for structuring the schema given to the AI model.

STUDY DESIGN

Objectives

The seminar had a methodological aim: to test the feasibility of an ad hoc human–AI collaboration framework for the analysis and cataloguing of graphic artefacts. The test was carried out with a sample of students, so that the verification process would also serve to build critical awareness of computational tools in humanistic research.

The Choice of the Corpus: Computing Magazines

Before any computational mediation, students were given a corpus of magazine covers on technology and computing, drawn from the digital collections of the Internet Archive[2] and spanning the 1970s to the early 2000s. This period coincides with the birth and gradual democratisation of the personal computer as a mass-market object, and specialist magazines record that transformation directly: in their design choices, in the audiences they imagined, and in the visual vocabulary used to represent and sell technology. The gradual sedimentation of visual stereotypes across these decades, combined with the thematic coherence of the corpus, makes these materials well suited both to comparison across titles and to long-term analysis of variation. Each student worked with an average of 200 images and was asked to consider explicitly which information was relevant for describing such artefacts.

METHODOLOGY

Participants

Forty-seven students enrolled in the Master's Degree Programme in Product, Communication and Interior Design at University Iuav of Venice[3] took part in the seminar. The institution trains designers with a substantial historical and theoretical grounding, and the Master's communication track already includes teaching in graphic design history.

[2]Internet Archive, https://archive.org, is a non-profit digital library of internet sites and other cultural artifacts in digital form.

[3]In particular, the activity was carried out within the Communication curriculum, during the Graphic Design History lessons taught by Prof. Fiorella Bulegato, on April 8th, 2026, with a follow-up discussion session held on May 6th, 2026.

The Task

Students were asked to carry out AI-assisted cataloguing of a corpus of computing and technology magazine covers drawn from the Internet Archive. The task involved (a) applying a cataloguing schema to a corpus and (b) using AI tools critically as aids in completing the schema's fields.

The Tool

The platform built specifically for the seminar offers a batch cataloguing interface that fills out a CSV record, using the empty fields prepared by the students, across the entire corpus, producing a structured CSV file. AI models are accessed through OpenRouter via API.

Protocol

The seminar took place over a single intensive day, in the following sequential phases:

1. **AI literacy.** A lecture on descriptive vocabulary and the composition of the graphic artefact gave students the conceptual basis they needed both to recognise the categories useful for analysing a visual artefact and to evaluate the AI's output critically afterwards.
2. **Acquisition of the corpus**. A demonstration of the Jupyter notebook used to download covers from the Internet Archive. Students working on the same collection received its URL and ran the download independently.
3. **Critical analysis of the artefact.** A joint discussion of how to break down the elements of a visual artefact (see Section 3) into a workable cataloguing schema.
4. **Building the schema.** Students examined the covers in their own collection and identified the categories needed to describe them systematically.
5. Introduction to the platform. A walkthrough of the platform and the steps to follow.
6. **Pre-exercise questionnaire.** Before the autonomous work session, each student completed the pre-exercise questionnaire individually (20 Likert items across 5 thematic blocks).
7. **Cataloguing.** An autonomous work session, completed from home.
8. **Post-exercise questionnaire and submission of results.** At the end of the session, each student completed the post-exercise questionnaire and submitted their output: the CSV file of the cataloguing.

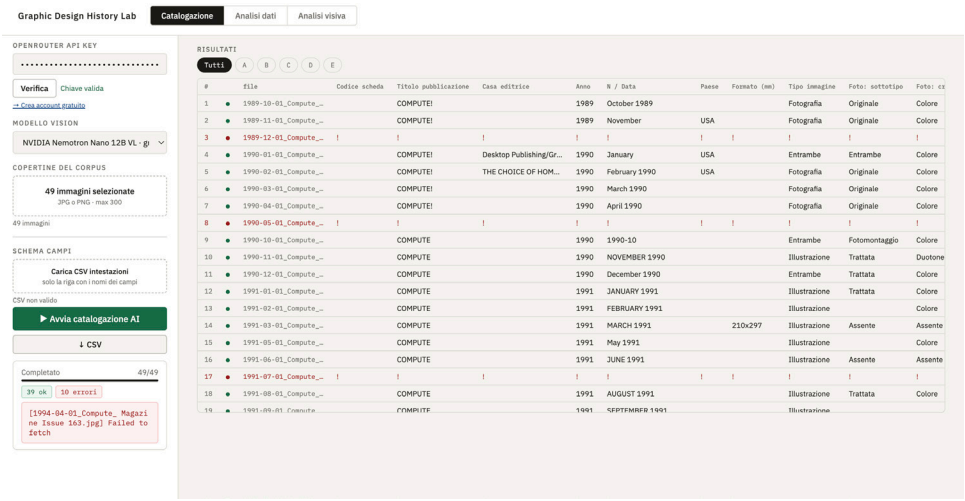

Figure 1: The cataloguing interface of the platform: workflow controls on the left, batch results on the right, with rows in red marking failed queries.

RESULTS

The data collected fall into two sets: (a) the pre- and post-exercise questionnaires (Likert scales 1–5), and (b) the CSV files produced by the AI cataloguing of each student (about 200 records per student).

Pre/Post Test Results

The questionnaires were analysed by checking the internal reliability of the scales with Cronbach's alpha and by running a Wilcoxon signed-rank test on paired samples for the pre/post comparison.

Cronbach's alpha was acceptable (≥ .70) for all thematic blocks except one: the "Disciplinary agency post" scale, which dropped to $\alpha = .333$, against $\alpha = .768$ for its pre-exercise counterpart. The use of AI for the task pulled participants' positions apart along the four dimensions of agency: leading the process, making autonomous interpretive choices, evaluating the schema in retrospect, and deciding what role to assign to AI.

Pre/post comparisons were carried out on three pairs of items, selected on the basis of strict conceptual equivalence: each pair measures the same construct first in the anticipation phase (pre) and then in the lived-experience phase (post).

The most striking result concerns output quality. Students entered the exercise already sceptical that AI output would be directly usable without revision (D4 pre, $M = 1.81$, $SD = 1.014$), and the experience confirmed and quantified that scepticism through the reported frequency of corrections (D4 post, $M = 3.66$, $SD = 1.089$), $W = 16.5$, $p < .001$.

The second pair concerns cataloguing time (D1 pre, $M = 3.83$, $SD = 0.868$; D1 post, $M = 3.28$, $SD = 1.210$), with $W = 525$, $p = .054$, marginally above threshold because of eight ties. The result is consistent with an expectation that was already fairly realistic before the task.

The third pair, on the distinction between description and interpretation (B4 pre, M = 3.91, SD = 0.830; B2 post, M = 2.87, SD = 0.947), gave $W = 621$, $p < .001$. The drop in the post score is best read as a recalibrated self-assessment: the exercise made the real complexity of that distinction, confidently asserted in abstract terms beforehand, visible in retrospect.

Cataloguing Records: Results

The sample comprises 47 cataloguing schemas drawn up by the students as a preliminary instrument for model-assisted annotation. On average, each schema contained 26.2 fields (median 21; range 5–54), but exact overlaps reduce the sample to 22 effectively distinct structures: the same template circulating among multiple students roughly halves the observed variety. In every schema, with only minor lexical differences, three descriptive elements recur, identifying data of the artefact (name, year and month of publication); a generic Image entry; and a generic Text entry. These constitute a shared "minimum identification record" corresponding to Panofsky's first level of analysis. Distance from the critical framework discussed in class emerges instead in how the categories are subdivided: 40% of schemas do not distinguish photograph from illustration, keeping a single Image entry. Technical disciplinary vocabulary appears in only a minority of schemas, and is concentrated in those that try to push description beyond a purely inventorial level.

The post-exercise questionnaire included two open-ended questions designed to elicit students' perceptions of the AI's role in cataloguing and analysis in an unstructured way. The first asked them to comment on possible shortcomings of the tool ("Is there anything you wished the AI had done differently?"); the second asked them to identify operations they would have preferred to carry out themselves ("Is there anything you would have preferred to do yourself?").

Responses to the first question fall into three groups. The most common concern technical and operational issues with the platform: limits on batch size per session, intermittent errors, slow uploads, and the need to restart the process when results came back incomplete. A second group concerns inconsistencies in data normalisation: heterogeneous formats within the same column (dates in alternative formats, mixed upper- and lower-case), separate treatment of linguistic variants of the same entity, and failure to recognise duplicates. The more articulate responses identified interpretive limits and hallucination: cases in which the AI filled fields with data not derivable from the image, and difficulty distinguishing photography from illustration.

Responses to the second question reveal a polarisation worth noting. Some students reported no sense of lost agency, crediting the tool with sparing them an especially burdensome stage of manual labour. A more reflective subset, however, identified the areas where they would have preferred to work on their own, not in cataloguing, but in analysis. Cataloguing was acknowledged as the task at which AI is most effective, since the systematic application of a schema to a large corpus is described as "exactly the type of task in which the AI excels, freeing time for critical reflection." One student

wrote: "I would have preferred to deal directly with the ambiguous cases, because they are the most interesting from a design point of view, and the AI tends instead to normalise them often."

Alongside the questionnaire analysis, we examined the system's outputs on the records generated. The picture matches what emerged from the questionnaires: the issues students reported are present in the outputs we examined. It is worth noting, however, that despite these specific difficulties most fields of the descriptive schema are filled in accurately and coherently, providing a first level of cataloguing usable as the basis for subsequent manual verification. The difficulties recur on categories that demand fine visual discrimination and field-specific design competence, for example, the correct attribution of a typeface to the serif or sans-serif class.

CONCLUSION

The seminar described here is a first attempt at formalising an AI-assisted procedure for cataloguing graphic design artefacts.

The platform-related limitations encountered (batch size, data normalisation, slow uploads) are not properties of AI as such, but consequences of the architectural choice to rely on open-source models accessible via OpenRouter, models less performant than their top-tier commercial counterparts. That choice was driven by pedagogical considerations and should be revisited in subsequent iterations, so that the analysis stage can distinguish the limits of the tool from those of the user's competence.

When manual and AI-assisted work are compared, the data confirm a real saving of time, though smaller in magnitude than participants had initially expected (D1, $p = .054$). Students' initial scepticism about directly usable output without revision (D4) was confirmed and quantified by experience, reinforcing the conclusion that AI cannot fully substitute for human interpretive work. A further interesting finding came from the open-ended responses, in which some students identified cataloguing, understood as the systematic application of a schema, as the task where AI is most effective, while reserving the analysis of ambiguous cases as the researcher's own ground.

Almost all students recognised, in the post phase, that building the schema by hand beforehand had made the difference in their critical engagement with the AI's output (E3 post). This suggests that the teaching sequence we adopted, manual schema construction, AI use, verification, can serve as a replicable pedagogical model in similar contexts.

The study's limitations point directly to next steps. The plan is to repeat the experiment with researchers already active in graphic design history, whose familiarity with the corpora is structurally greater, on a platform updated in response to the issues raised here.

ACKNOWLEDGMENT

The authors would like to acknowledge Fiorella Bulegato for hosting the seminar within her course.

REFERENCES

Anceschi, G. (1988). *Monogrammi e figure: Teorie e storie della progettazione di artefatti comunicativi* [Monograms and figures: Theories and histories of the design of communicative artefacts]. La Casa Usher.

Aynsley, J. (2012). Graphic design. In S. De Bondt & C. de Smet (Eds.), *Graphic design: History in the writing (1983–2011)* (pp. 21–35). Occasional Papers.

Arnold, T., & Tilton, L. (2019). Distant viewing: Analyzing large visual corpora. *Digital Scholarship in the Humanities, 34*(Suppl. 1), i3–i16. https://doi.org/10.1093/llc/fqz013

Arnold, T., & Tilton, L. (2023). *Distant viewing: Computational exploration of digital images*. MIT Press.

D'Ambrosio, G., Grimaldi, P., & Lenza, C. (1983). E se Gutenberg fosse stato un designer? [What if Gutenberg had been a designer?]. *Op. cit., 58*, 5–20.

De Smet, C., & De Bondt, S. (Eds.). (2012). Graphic design: History in the writing (1983–2011). Occasional Papers.

Fallan, K. (2010). Design history: Understanding theory and method. Berg.

Lancia, A. (2025). L'intelligenza artificiale al servizio degli archivi multisensoriali [Artificial intelligence at the service of multisensory archives]. In G. Clera (Ed.), *Nuove competenze digitali per il patrimonio culturale degli archivi* [New digital skills for the cultural heritage of archives] (1st ed., pp. 67–71). Edizioni Fondazione Levi.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. CHI Conference on Human Factors in Computing Systems. ACM.

Manovich, L. (2020). Cultural analytics. MIT Press.

Manovich, L., & Douglass, J. (2009). Timeline: 4535 Time Magazine covers, 1923–2009 [Research project]. Cultural Analytics Lab. https://lab.culturalanalytics.info/2016/04/timeline-4535-time-magazine-covers-1923.html

Margolin, V. (1992). Design history or design studies: Subject matter and methods. Design Studies, 13(2), 104–116.

Panofsky, E. (1939). *Studies in iconology: Humanistic themes in the art of the Renaissance*. Oxford University Press.

Polano, S. (n.d.). sp - mappa v.2.7 [Study document]. Sergio Polano Collection, Iuav Project Archive, University Iuav of Venice, Venice.

Poynor, R. (2011, January 10). Out of the studio: Graphic design history and visual studies. *Design Observer*.

Tifentale, A., Manovich, L., & Indaco, A. (2018). Exploring photography magazine Sovetskoe Foto, 1926–1991 [Research project]. Cultural Analytics Lab. https://lab.culturalanalytics.info/2018/07/exploring-photography-magazine.html

Triggs, T. (2011). Graphic design history: Past, present, and future. Design Issues, 27(1), 3–6.

Wevers, M., & Smits, T. (2020). The visual digital turn: Using neural networks to study historical images. Digital Scholarship in the Humanities, 35(1), 194–207.

Woodham, J. M. (1995). Reviewing design history. Design Issues, 11(1).

Yadav, S., Tilton, L., Antoniak, M., Arnold, T., Li, J., Pawar, S. M., Karamolegkou, A., Frank, S., An, Z., Rostamzadeh, N., Hershcovich, D., Belongie, S., & Shutova, E. (2025). *Evaluation of cultural competence of vision-language models* (arXiv:2505.22793). arXiv. https://doi.org/10.48550/arXiv.2505.22793