



CULTURAL COMPLEXITY AND COMPUTATIONAL APPROACHES

Toward a Systemic and Critical
Understanding of Cultural Ecosystems

30 June, 1-2 July 2026 MÁLAGA, SPAIN

ABSTRACTS BOOKLET



COMPLEXHIBIT



UNIVERSIDAD
DE MÁLAGA



CÁTEDRA^{5G}
Telefónica | UMA



COLECCIÓN
DEL
MUSEO
RUSO



Complexity is a general property of cultural systems across historical periods. Cultural life has long unfolded through dense interdependencies among institutions and patronage regimes, media and material infrastructures, markets and circuits of exchange, publics and interpretive communities, and the norms and policies that govern cultural legitimacy.

Across time —whether in early modern correspondence networks, nineteenth-century print cultures, or today's platform-mediated environments— these interacting forces constitute cultural ecosystems whose dynamics are non-linear, multi-scalar, and historically stratified. They generate emergent phenomena such as canon formation and marginalization, reputational cascades, diffusion and imitation, cycles of attention, unequal circulation, and durable asymmetries of access and preservation.

SUMMARY

1	Modeling Time, Space, and Complexity: Simulating Prehistoric Societies as Event Systems. Juan A. Barceló, Joaquim Fort, Albert García Piqué, Olga Palacios, Ramón Buxó, Jaume Noguera.
2	<i>The Interweavings Between Politics and Culture in Multiplatform Publication Practices of a Right-Wing-Blog: A Russian Example of Whiteness Discourse and How to Investigate its Complexity.</i> Elena Hamidy.
3	Mapping the 20th Century Architectural Ecosystem through Journals: a retrospective and prospective reflection on Bidiria. Federico Deambrosis and Andrea Gritti.
4	Constructing the Canon: Museums, Exhibitions and Data Analysis in the Processes of Cultural Legitimation of Pictorial Abstraction. Carmen Padrun.
5	Capturing Ephemeral Complexity: A Linked Open Data Approach to Theatre Productions. Vivien Wolter.
6	From Live In Your Head: When Attitudes Become Form to Linked Data: a methodological proposal to describe exhibitions and processual contemporary artworks through the OntoExhibit ontology. María Serena Matarrese and Nuria Rodríguez Ortega.
7	<i>Hotspots, Bypasses, and Semantic Drift: Mapping Cultural Knowledge Through Wikipedia's Biographical Network.</i> Pachalis Agapitos, Juan Luis Suárez and Gustavo Ariel Schawrtz.
8	<i>German Democratic Republic in Context: Reducing the Complexity of Historical Data through Visualization.</i> Linda Freyberg and Josefine Wähler.
9	<i>Mapping Cultural Complexity in Networks of Writers Over the Last Three Centuries.</i> Pachalis Agapitos, Gabriel Cwilich and Gustavo Ariel Schawrtz.
10	The Order of Indeterminacy: Complexity, Uncertainty, and the Black Box — The Limits of Modelling. Arantzazu Saratzaga Arregi.
11	Modeling an Intermedial Ecosystem: A Computational Approach to Gendered Canon Formation in 1930s Madrid. Irene Palencia Mora.
12	Modelling Visual Complexity in Performance Photography: Period, Venue, and Photographic Mediation at the Festival d'Avignon. Giacomo Alliaa.
13	Regrouping the Avant-Garde: A Multiscalar Network Approach to European Artist Groups (1905–1915). Teresa Kamencek, Michaela, Silvia Miksch and Raphael Rosenberg.
14	Historical Frames. Interacting with historical change and complexity in cultural heritage ecosystems. Sofia and Charles van den Heuvel.
15	Fragile Networks: Latent Comprehension and the Hermeneutics of LLM-Based Citation Extraction in Art History. Andrea Alfarano, Eduardo Trabattoni and Dario Negueruela del Castillo.
16	<i>Computational Hermeneutics of Museum Discourse: Modeling Bias and Interpretive Framing in Curatorial Texts.</i> Mahnoush Pourhosseini Akbarie, Shiva Talaei Koochehbagh and Paola Pisano.
17	Charting Contemporaneity: Pathways through Exhibitions, Museums, and Markets. Clarissa Ricci.

18	Ten years after MUSA16: Advances and Challenges in Compiling a Corpus for the Study of Translated Museum Texts. Jorge Leiva Rojo.
19	Exhibition Mise en Place. Gender and Transnationality as Key Variables for Analyzing the Art Ecosystem. Bárbara Romero Ferrón.
20	"Dense regions of complexity": Intrinsic Dimension and the Neural Geometry of Art-Critical Knowledge in Large Language Models. Ángel M ^a Lumbreras.

Modeling Time, Space, and Complexity: Simulating Prehistoric Societies as Event Systems

By Juan A. Barceló (Universitat Autònoma de Barcelona); Joaquim Fort (Universitat de Girona); Albert García Piqué (Universitat Autònoma de Barcelona); Olga Palacios (Durham University); Ramón Buxó (Museu d'Arqueologia de Catalunya); and Jaume Noguera (Universitat Autònoma de Barcelona).

The study of prehistoric social systems requires taking seriously the problem of non-linearity, contingency, and multidimensionality in past human dynamics. Archaeological data, rich in spatial extension but temporally fragmented, rarely provide linear narratives. Instead, they register discontinuities, accelerations, and stagnations that result from multiple interacting mechanisms—demographic, environmental, technological, and symbolic. The challenge is thus not only empirical but conceptual: how to translate fragmentary evidence into coherent explanatory models without reducing complexity to mere complication, that is, to the accumulation of variables and parameters that obscure rather than clarify social processes. This paper argues for moving from complicated models—those that attempt to include every conceivable variable—to complex dynamic simulations that foreground interaction, feedback, and temporal evolution.

To achieve this, we adopt a chrono- and geo-referenced temporal event perspective in modelling prehistoric societies. Rather than representing history as an undifferentiated timespan divided into arbitrary phases, our goal is to represent specific events and processes that can be temporally and spatially located. This allows simulation frameworks to operate directly on observable data—such as radiocarbon dates, site locations, and artefactual assemblages—while exploring the generative mechanisms that produced them. The concept of the event here is broader than simple occurrence: it refers to a unit of transformation, an intersection of agents, material conditions, and opportunities that leaves a trace in the archaeological record. Modelling such events through computational methods enables a dynamic rather than static understanding of prehistory.

In this work, we will compare three distinct types of computational modelling—reaction-diffusion, agent-based, and network models—applied to a shared archaeological and geographical case: the origins and diffusion of farming among prehistoric populations of the Iberian Peninsula approximately 6000 years ago. This period, corresponding to the transition from foraging to farming economies, represents an ideal test case for computational simulation. It includes clear spatial patterns, well-established chronological anchors, and theoretical debates about the mechanisms of change: was the spread of farming primarily driven by demic expansion, cultural transmission, environmental adaptation, or social connectivity? By applying the same dataset and chronological

framework across three modelling approaches, the comparison seeks to evaluate how different conceptualizations of social interaction lead to distinct reconstructions of historical process.

The first class of models, reaction–diffusion simulations, treats social or demographic change as a continuous process that evolves over a discretized landscape according to differential rates of population movement and innovation transfer. The mathematical structure of these models, derived from physical systems, makes them particularly effective for exploring large-scale diffusion patterns and equilibrium dynamics. In the Iberian case, the reaction–diffusion approach helps visualize how early farming communities might have spread across ecological gradients, integrating constraints such as topography or resource distribution. However, its strength—continuous regularity—can also be a limitation when social decisions, communication, or cooperation introduce discontinuities that cannot be captured by diffusion alone.

The second modelling approach, agent-based simulation, introduces decision-making and behavioral variability at the level of individual or group agents. Each agent acts according to rules related to subsistence, mobility, reproduction, and interaction, generating complex macroscopic outcomes through iteration and feedback. In modelling early Iberian farmers and hunter-gatherers, agent-based models allow us to explore how local interactions and cultural learning could produce emergent regional trajectories. They make visible the non-linearity of social transformations—how small changes in preferences, environmental stress, or intergroup contact can amplify into large differences in settlement distribution or social structure. However, this flexibility increases the risk of overparameterization, where too many behavioral rules produce a complicated but uninterpretable system. Our chrono- and geo-referenced event perspective serves as a methodological constraint, anchoring simulation rules to empirical evidence and keeping complexity analytically productive.

The third approach, network modelling, conceptualizes prehistoric societies as systems of relationships rather than spatially continuous populations. Nodes represent settlements, communities, or individuals, and edges signify potential paths of communication, exchange, or mobility. Network topologies—whether centralized, modular, or scale-free—strongly influence the rate and direction of innovation diffusion. In the context of early Iberian farming, network models help evaluate the structure of connectivity among coastal, inland, and highland communities. They also make it possible to compare how changes in network density or node importance affect the tempo of agricultural adoption. Because networks evolve through time, integrating event-based chronology provides insight into how connections appear, persist, and dissolve, thus linking social structure to temporal dynamics.

Across these three approaches, a key conceptual issue arises: the distinction between explicit and emergent time. In most archaeological modelling, time is treated explicitly—as a sequence of measured intervals derived from chronometric data. This explicit time provides the reference frame necessary to calibrate and compare models. Yet many processes that drive social transformation unfold in emergent time: the internal temporalities produced by interaction delays, adaptation rates, or feedback cycles within the system itself. Emergent time is not imposed from outside but arises from the simulation's internal logic; it marks how the system experiences its own development. Bridging explicit chronological time and emergent dynamic time is essential to move from static reconstruction to processual explanation. In our approach, events serve as the interface between these two temporalities: each archaeologically dated event anchors emergent system dynamics to observable evidence.

This comparative framework thus provides both a methodological and theoretical contribution. Methodologically, it demonstrates how chrono- and geo-referenced datasets can constrain different types of computational models, allowing cross-validation between simulated trajectories and archaeological chronologies. Theoretically, it illustrates how complexity emerges from interaction and adaptation rather than from adding variables. The use of multiple modelling paradigms clarifies what aspects of social dynamics are robust across representations and which depend on specific assumptions. By exploring how similar empirical data behave under continuous (reaction–diffusion), discrete (agent-based), and relational (network) formulations, we gain insights into the invariants of prehistoric change.

In preliminary results, all three models reproduce recognizable features of the Neolithic expansion in the Iberian Peninsula—coastal concentration, inland delay, and regional differentiation—but they do so for different reasons. The reaction–diffusion model attributes spatial heterogeneity to environmental resistance and diffusion coefficients, while the agent-based model relates it to cultural adoption thresholds and interaction frequency. The network model, in contrast, suggests that emergent centralities and transient isolation among nodes can replicate the same pattern. Such multiplicity of mechanisms reinforces the idea that prehistoric processes are not determined by single causes but by networks of interdependent feedbacks. It also emphasizes the importance of exploring not one “best-fitting” model but a family of plausible dynamic reconstructions.

Ultimately, our study advocates a shift in how computational archaeology conceptualizes time, process, and data. Rather than using models merely to reproduce observed distributions, we view them as laboratories for testing how interactions generate temporally organized outcomes. The chrono- and geo-referenced event perspective offers a framework for integrating explicit archaeological time—anchored in measurable

data—with emergent system time—produced by interactions and adaptations. This synthesis enables us to simulate prehistory not as a static arrangement of sites but as a living field of evolving dynamics. The broader implication is that computational modelling, when properly constrained and conceptually grounded, can transform archaeology into an experimental science of past human complexity.

The Interweavings Between Politics and Culture in Multiplatform Publication Practices of a Right-Wing-Blog: A Russian Example of Whiteness Discourse and How to Investigate its Complexity.

By Elena Hamidy (University of Giessen).

The relationship between culture and politics is complex: Especially after the spread of culture wars in the worldwide public discourses beyond the USA (Bures 2023), cultural matters such as language, civilization and history serve as arguments in right-wing populism and in the justification of military aggression (e.g. Russia's invasion of Ukraine as in Gaufman 2025, Laruelle et al. 2023 and Laruelle 2025). The classical understanding of right-wing rhetoric classifies it as "a thin-centered ideology, addressing only a limited set of issues", with a "restricted morphology" (Mudde and Rovira Kaltwasser 2017, 6–21). Therefore, it is considered to be manipulative yet simplistic (Miglietta et al. 2023). From this viewpoint, a capacity for a more sophisticated view of complex historical and cultural factors is seen as a guarantee of a more ambivalent and critical perspective (Contu et al. 2024). In this presentation, I question such judgements by highlighting the inner and in-between complexity of multiplatform right-wing blogging practices while describing the complex intersections between political and cultural issues as a matter not of content, but of media practice. In this regard, I follow Milani's proposal to investigate the mediatization of far-right practices as the tension between "the capillarity of fascism" and "its semiotic complexity", which allows it to take "properties of a rhizome". (Milani 2024, 209)

The use of social media and especially the consumption of content from online video platforms proved to be a predictor of "cultural backlash" and support for right-wing parties across different European countries (Verboord et al. 2025). Interestingly, the proliferation of right-wing ideas in Russia is not only a matter of state ideology, as the national and racist rhetoric of white supremacy offers a common frame for both political radicals and oppositional liberal intellectuals (Djagalov 2021). In contrast to state propaganda, non-state actors lack resources and rely broadly on social media to gain attention. This study contributes to an understanding of such practices by investigating the publication and argumentation strategies of the oppositional conservative blogger Evgenii Ponasenkov, under whose name a set of online video channels and social media blogs is referenced. A famous media persona, the so-called "Maestro", he features multiple memes and uses a professionally orchestrated environment, managed by his team members, to convey messages to his "adepts". His elitist self-presentation as a historian, "man of science" and "culture warrior" is not limited to specific cultural topics

but combines civilizational discourses with humorous commentaries on current political and migration issues.

RQ1: In what direct and indirect ways are cultural objects and issues used to advance the ideas of white supremacy and support political claims?

RQ2: What types of cultural

topics and objects are presented, and how is their selection motivated?

RQ3: How can the complexity of multiplatform content strategies be analytically reduced to identifiable categories without losing sight of their performative and processual character?

The investigation is based on collected blog data; by using computational methods and content mapping, I aim to establish a model for analyzing complex interactions, content re-shares, and content composition across blogs with different affordances. To analyze attention strategies in multiplatform settings, one must simultaneously acknowledge and transgress the level of single-platform interactions – which necessarily adds complexity to the methodological setting. An analysis of their typical features suggests that discovering patterns requires methodological alignment and reduction. In this sense, I will demonstrate ways of establishing observations from both platform-specific and cross-platform perspectives by examining tactics of producing redundancies, content re-use and re-sharing and cross-posting.

The complex multiplatform interactions enable the mobilization of resources from multiple social media platforms (differently censored) and the establishment of an interplay of racist arguments across political and cultural topics. To grasp the tactics beyond the complexity of the total media interplay, it is necessary to build a clear vocabulary of explicit and implicit meanings. Not only comparisons, analogies, embeddings, narrative framings (Schilk 2025) and hate speech (Matamoros-Fernández and Farkas 2021) are used in politicizing cultural and historical facts. Some forms of ambient digital racism (Agudelo and Olbrych 2022) are less obvious on the content level: building common context (van Dijk 1993), bricolage of found and self-produced objects, semantic neighborhoods and norm-compliant cover messages (Saul 2024). To evoke associations and support the level of arguments, a whole framework of associations is built that advocates for the ideas of white supremacy and racist civilizational discourse.

The main findings offer an ambivalent picture in which complexity unfolds over time within the dimension of practice, which is at once organized in the sense of ritualized actions and appears spontaneous and chaotic at specific moments. Multi-platform interactions enable the productive embedding of artifacts and the drawing of resources from different platforms, overcoming platforms' censorship, navigating across segments of the audience, and re-sharing and re-using content to keep people engaged. The playful orchestration of

these resources evokes the impression of plentiful complexity and abundance. In my analysis, I seek a balanced way of reducing this complexity by mapping it to the level of categories and simultaneously conceptualizing it not as a static composition, but as a performance of different types of content unfolding over time.

Cited works

Agudelo, Felipe I.; Olbrych, Natalie (2022): It's Not How You Say It, It's What You Say: Ambient Digital Racism and Racial Narratives on Twitter. In: *Social Media + Society* 8 (3).

DOI: 10.1177/20563051221122441.

Bures, Eliah (2023): The Intellectual as Culture Warrior. In: *Fascism* 12 (1), S. 1–26. DOI: 10.1163/22116257-bja10055.

Djagalov, Rossen (2021): Racism, the Highest Stage of Anti-Communism. In: *Slavic rev.* 80 (2), S. 290–298. DOI: 10.1017/slr.2021.83.

Gaufman, Elizaveta (2025): "The Last White Country": Mapping the Far-Right Ideology in Russia during the Full-Scale Invasion of Ukraine. In: *Javnost - The Public* 32 (1), S. 19–32. DOI: 10.1080/13183222.2025.2469029.

Laruelle, Marlene; Grek, Ivan; Davydov, Sergey (2023): Culturalizing the Nation: A Quantitative Approach to the Russkii/Rossiiskii Semantic Space in Russia's Political Discourse. In: *Demokratizatsiya: The Journal of Post-Soviet Democratization* (1), S. 3–27.

Laruelle, Marlène (2025): *Ideology and meaning-making under the Putin regime*. Stanford, California: Stanford University Press.

Matamoros-Fernández, Ariadna; Farkas, Johan (2021): Racism, Hate Speech, and Social Media: A Systematic Review and Critique. In: *Television & New Media* 22 (2), S. 205–224. DOI: 10.1177/1527476420982230.

Miglietta, Anna; Molinengo, Giorgia; Rizzo, Marco (2023): Endorsing populism to cope with ambiguity? The role of the need for closure, self-deception, and personal values in advocating populist attitudes. In: *Personality and Individual Differences* 203, S. 112031. DOI: 10.1016/j.paid.2022.112031.

Milani, Tommaso M. (2024): Why Everybody Wants to Be a Fascist and Why We Should Study Language to Understand It. In: *Journal of Right-Wing Studies* 2 (2). DOI: 10.5070/RW3.41981.

Mudde, Cas; Rovira Kaltwasser, Cristóbal (2017): *Populism. A very short introduction*. Oxford, New York, NY: Oxford University Press.

Saul, Jennifer Mather (2024): Racist Dogwhistles. In: Jennifer Mather Saul (Hg.): *Dogwhistles and Figleaves*. Oxford University Press, S. 34–69.

Schilk, Felix (2025): The Metapolitics of Crises: How the New Right Weaponises Narratives to Mainstream Far-Right Ideology. In: *Int J Polit Cult Soc.* DOI: 10.1007/s10767-025-09519-3.

van Dijk, T. A. (1993): Analyzing racism through discourse analysis: Some methodological reflections. In: John H. Stanfield und Rutledge M. Dennis (Hg.): Race and ethnicity in research methods. Newbury Park, Calif.: Sage Publications, S. 92–134.

Verboord, Marc; Heikkilä, Riie; Janssen, Susanne; Purhonen, Semi (2025): Media use, cultural orientations and support for radical right-wing parties: comparing nine

European countries. In: Information, Communication & Society, S. 1–20. DOI: 10.1080/1369118X.2025.2536159.

Mapping the 20th Century Architectural Ecosystem through Journals: a retrospective and prospective reflection on Bidiria.

By Federico Deambrosis (Politecnico di Milano) and Andrea Gritti (Politecnico di Milano).

Digitale delle Riviste Italiane di Architettura / Towards a Digital Library of Italian Architectural Periodicals) project aimed to verify the possibility and potential of a new retrospective approach to 20th-century design-centred journals. Funded by the Italian Ministry of Research through a competitive call, the project was designed to produce a prototype to be developed in an eventual second phase. This prototype consists of a query system that allows various searches on a database grouping the issues of seven Italian periodicals published in 1959. These are "Casabella continuità", "L'architettura: cronache e storia", "La casa. Quaderni di architettura e critica", "Edilizia Moderna", "Zodiac", "Urbanistica", and "Civiltà delle Macchine". As an informed observer will realise at first glance, this is a heterogeneous sample, comprising well-known and almost forgotten journals that differ in their periodicity, target audience, subjects, style, and mode of communication. Equally glaring are some absences, starting with "Domus", which we excluded in this phase as it is perhaps the only Italian architectural periodical with a digital archive accessible via the web, albeit rather conventional in its search methods. BiDiRIA can be understood as a dual observatory. On the one hand, it aimed to study architecture and its evolutions and hybridisations through the traces left on the pages of periodicals; on the other, it was also extremely interested in journals as autonomous objects, themselves the outcome of design thoughts. As such, it provides a test-bed to critically assess both the potential and the limits of computational approaches when applied to complex cultural sources.

The construction of the prototype began with the digitisation of the issues of the diverse journals. Each article or, more precisely, each item (considering by this term also the editorials, news, reviews and the other kind of columns that populated the periodicals of those years) corresponds to a file. Journals were thus "exploded", atomising the logics that originally guided the construction of indexes or the structuring of monographic issues. Each item is in fact visualised in the prototype as a dot, contributing to the construction of the cloud of articles published in the seven journals during the semester of reference. Each point contains a link to a copy of the item and chromatically reveals its relevance to the criteria used to query the contents. BiDiRIA thus enables the use of periodicals to map architectural discourse. These maps are all the more reliable the more the subjective elements involved in their generation are minimised. This is why journals

with diverse orientations and approaches were chosen. In other words, the prototype allows access to implicit quantitative elements, such as the number of references to the production of a given foreign country or the space devoted to specific building types or construction materials. The maps, of course, also highlight other themes: the critical fortune of individual designers, the consistency of individual authors' work, or unexpected typological correspondences. The criteria introduced in the query column of the prototype are: "Journal", "Category", "Authors", "Designer/Artist", "Geography", "Function", "Scale", "Representation", "Building Types", and "Building Elements".

Each search chromatically highlights a set that meets the selected criteria; within it, it is possible to obtain information on each article and, eventually, visualise it. The last two criteria (Types and Elements) proved to be problematic. In many cases, the results are affected by the information introduced during the pre-classification of the buildings or projects described. While a "bridge" or an "urban plan" can be objectively identified, "multi-storey building" is a typological label that can be applied to very different cases. Similar situations occur when focusing on construction materials and techniques. In other words, the journals we analysed seem to resist traditional cataloguing based on predefined categories. On the other hand, it is precisely these criteria that make it possible to explore in depth the more strictly architectural aspects of the classified information, thus representing the most challenging frontier to be reached. The tension between analytical ambition and the need for synthesis that underlies the research project generates divergences that are difficult to reconcile. When the former prevails, categories become numerous and articulated across multiple levels, pursuing a potentially inexhaustible complexity. When the latter prevails, on the contrary, a level of simplification is introduced that proves inadequate to describe an extraordinarily articulated set of experiences and projects.

These questions (one could sum up as "To what extent should research be left free, structuring it by keywords, or how much, instead, can it benefit from predefined parameters? How many parameters and which ones? Can parameters really be objective, or do they inevitably end up guiding and influencing the research and its outcome?") point to a fundamental tension between the need for structured data and the irreducible complexity of architectural discourse, and have been on our minds throughout the long period between the end of the first phase and the closing of applications for the second. The technological acceleration that had taken place in this interval suggested moving beyond a query system based on manually extracted features, and on criteria and facets by complementing it with other technologies that allow more detailed information to be extracted automatically. These include machine learning techniques for pattern detection and semantic clustering, which reduce the need for rigid, predefined classification

schemes. This new, flexible structure will also make it possible to achieve different outcomes, addressed to research as well as to divulgation. Journals still appear as extraordinarily rich materials which, to date, have been just partially explored by scholars. Re-leafing through them, a mostly forgotten story emerges. It becomes evident how historiography operated exclusions and choices that, in an incremental process, widened a cone of shadow that still dominates our 'collective memory'. At a time when new syntheses are being attempted from specific perspectives (such as the environmental one), showing the need to rethink the historiographical canon, periodicals are probably among the most reliable sources from which to begin such a weighty task. Nevertheless, these sources are multilayered and complex; they often resist simplification. Dismembering the files and exploring them with digital tools will hopefully reveal new perspectives and possible narratives. These operations, which can greatly benefit from AI, could also yield another outcome: by separating content and reaggregating it in new forms, new teaching and divulgation formats centred on architecture will likely emerge. In this sense, computationally assisted readings of journals do not simply extend existing methods, but call for a critical redefinition of how architectural history can be constructed from complex cultural ecosystems such as architectural journals.

Constructing the Canon: Museums, Exhibitions and Data Analysis in the Processes of Cultural Legitimation of Pictorial Abstraction.

By Carmen Padrun (University of Málaga).

The aim of our proposal is to analyse the process by which the style of pictorial abstraction has gained legitimacy within the art world and society, considering both its origins and its historical development. We start from the premise that the history of abstraction has not been constructed solely based on artistic practices themselves, but through a network of institutional, curatorial and academic discourses that have established a particular narrative as hegemonic. In this context, we advocate the need to integrate the use of databases and current digital technologies as a standard practice, as these tools enable the systematic and efficient handling of large volumes of information, optimising resources and opening up new analytical possibilities for art historiography.

This research proposes to examine how a specific discourse on pictorial abstraction has been disseminated, consolidated, and naturalised, which we inherit both culturally and socially. To this end, it considers the observation of various quantifiable factors, such as the number of exhibitions dedicated to a particular artist, the museums that have organized the greatest number of shows centred on them, as well as the chronology of these exhibitions. The analysis of these variables allows us to examine the implementation and legitimation of an official narrative regarding the emergence of abstract poetics and the pre-eminent role attributed to certain artist within this process. Major museum institutions dedicated to modern art —such as the Museum of Modern Art and the Solomon R. Guggenheim Museum in New York, the Centre Pompidou in Paris, Tate Modern in London, and the Stedelijk Museum in Amsterdam, among others— have played a decisive role in the construction of an authorised narrative that has been widely disseminated in society, directly influencing both artistic practice and its reception.

The cultural interpretation of pictorial abstraction and its history is thus deeply conditioned by the exhibition policies and curatorial discourses promoted by these institutions. Through a continuous process of feedback between museums and academic bodies, certain features and figures have received sustained emphasis, contributing to their historiographical centrality. This mechanism of selection and valorization simultaneously entails the marginalisation of other elements, which come to occupy a secondary position within the dominant narrative. Such exclusion does not in fact stem

from a lack of artistic relevance, but rather from specific contingent factors linked to political, economic, social, and cultural contexts.

From this theoretical framework, we propose a methodological approach grounded in the use of statistical tools and probability analysis, integrated within a critical perspective characteristic of art history. The starting point consists in the systematic counting of exhibitions devoted to pictorial abstraction corresponding to the period of the Historical Avant-Gardes in the most influential museums in the field of modern art. The initial dataset focuses on exhibitions whose curatorial discourse explicitly addresses both the origins of abstraction and the identification of its principal representatives. This information would be gathered from libraries, archives, and digital platforms of the museums under study, with particular emphasis on the existence—or the potential creation—of unified databases that allow for the efficient organisation and cross-referencing of data.

In an initial phase, descriptive statics would be applied in order to classify and organize the information obtained, identifying general trends and significant variations over time. In a second phase, the use of probability analysis would make it possible to estimate the relative centrality of a particular artist or discourses within the set of exhibitions analysed, as well as to measure the frequency with which certain names or approaches appear associated with the genesis of abstraction. The analysis would be supported by statical tables and graphical representations that facilitate the visualisation of the data and their comparison across different institutions and historical periods. The ultimate aim is, on the one hand, to underscore the importance and power of the museum institution in the interpretation and dissemination of particular historical narratives and, on the other, to highlight the need for a critical reassessment of the established canon surrounding the history of pictorial abstraction in the context of the *Historical Avant-Gardes*. By identifying which artists or approaches have received greater attention from cultural legitimising agents, it becomes possible to trace more precisely the contours of the dominant discourse, as well as to detect its gaps and silences, which in recent decades have served as the point of departure for new theoretical contributions that challenge the hegemonic account of the origins of abstraction.

Among these recent revisions, particular emphasis has been placed on the revaluation of practices and figures traditionally considered marginal, such as the abstractions of artists like František Kupka or Mikalojus Konstantinas Čiurlionis, as well as the mediumistic abstractions of the late nineteenth and early twentieth centuries, in which figures such as Hilma af Klint and Georgiana Houghton stand out. Likewise, the historiographical explanation of the pictorial development of abstraction has, for example, been strongly

conditioned by exhibition models that take styles such as Impressionism or Cubism as inevitable precedents. Since the landmark exhibition *Cubism and Abstract Art*, held at MoMA in 1936, this approach has shaped numerous subsequent exhibitions, including *Monet and Abstraction* (Museo Nacional Thyssen-Bornemisza, 2010), *From Impressionism to Abstract Art* (Guggenheim Museum Bilbao, 2004-2005), *Cézanne, Picasso, Mondrian: A New Perspective* (Kunstmuseum Den Haag, 2009-2010), *Picasso & Abstraction* (Royal Museums of Fine Arts of Belgium, 2022-2023), and *From Cézanne to Malevich: From Arcadia to Abstraction* (Museum of Fine Arts, Budapest, 2021-2022). These exhibitions consolidate a narrative that situates the remote origins of abstraction in Impressionism or in Post-Impressionist styles such as that of Cézanne, sometimes extending this temporal framework even further, as in *El Greco and Modern Painting* (Museo del Prado, 2014) or *Origins of Abstraction (1800-1914)* (Musée d'Orsay, 2003-2004).

Beyond its function as an exhibition space, the twentieth-century museum thus became established as an agent that produces discourses capable of shaping the course of art and its history. As Gleen D. Lowry, one of the most influential directors of MoMA, noted in *Inventing Abstraction, 1910-1925*, this shift in paradigm was closely linked to the figure of Alfred H. Barr Jr., who conceived of the museum as an experimental laboratory and employed a language akin to that of science in order to legitimise a vision of modern art in which abstraction appeared as the logical culmination of preceding movements. This approach responded to a clear aim: to justify the institutional commitment to pictorial abstraction as the structuring axis of the discourse proposed regarding the development of modern art (Frankel, 2012:361).

We consider this type of historiographical analysis essential for broadening and refining our understanding of the history of art and its social impact. The study of pictorial abstraction as a paradigmatic case allows us to reflect on the mechanisms of cultural legitimation and on the active role of institutions in the construction of knowledge. In this sense, we advocate the critical incorporation of technological tools and quantitative methodologies as a means of streamlining the analysis of large datasets, optimizing research resources, and opening new avenues for a more comprehensive reassessment of traditional historiographical narratives.

Capturing Ephemeral Complexity: A Linked Open Data Approach to Theatre Productions.

By Vivien Wolter (Trier Centre for Digital Humanities, Trier University).

Theatrical productions represent a particularly complex form of cultural event. As an artistic practice, they become visible and perceptible only through their performances, and yet the production as a concept extends beyond any single performance, remaining recognisable across multiple performances through the consistency of its scenographic and structural features (cf. Balme, 2021).

A production consists of several interdependent dimensions: space and architecture, bodies and movement, objects and props, lighting, text, direction, and the audience, whose presence actively shapes the event. None of these dimensions alone determines the character of a production. Meaning and form arise from the interplay of all these constituent elements. In this respect, theatre productions can be understood as paradigmatic examples of cultural systems whose properties cannot be reduced to the sum of their parts.

At the same time, theatrical productions are ephemeral. When a production closes, it survives only through indirect traces: director's prompt books, photographs, video recordings of varying completeness, programme booklets, and theatre reviews (cf. Brincken & Enghart, 2008). These materials serve primarily as carriers of information about an immaterial event. This creates a fundamental methodological challenge for theatre studies: to what extent can something that becomes perceptible only through its performances, and that resists stable documentation, be made accessible to systematic, comparative analysis?

This research approach examines to what extent the complexity of theatrical productions can be formally modelled using Linked Open Data and what it means to translate a transitory cultural event into structured data. The process of developing and implementing such a model simultaneously raises fundamental questions about which types of cultural knowledge can be formally represented and which cannot.

Existing computational approaches within theatre studies have developed considerably over the past decade, encompassing a range of national and international projects. A survey of these initiatives reveals a consistent pattern: they predominantly process metadata relating to productions and the persons involved, contextualise archival

materials in relation to practitioners and performance histories, and facilitate the digitisation of physical theatre materials such as playbills and programme booklets. Modelling theatre data presents particular challenges: the concepts associated with theatre are highly variable, the range of practitioners and their functions is diverse, and the relevant sources are heterogeneous, all in a field where shared standards for data formats and controlled vocabularies have yet to be established (cf. Probst & Pinto, 2020).

What remains largely absent from these approaches is a structured representation of the production itself. No project has been identified that models, alongside data on personnel, performance dates, and venues, the features of what actually takes place on stage: set design, costumes, props, spatial configurations, and forms of audience interaction. This research addresses that gap by developing a data model that places precisely these production-level features at the centre of formalisation.

The data model developed as part of this research captures both the organisational and scenographic dimensions of theatre productions. On the organisational level, it represents venues and institutions, people involved, performance dates, and the underlying dramatic work. On the scenographic level, it models individual roles and their connections to costume components, spatial configurations, stage elements, and forms of audience interaction. A central methodological concern is not only to capture entities but also the relationships between them, for example how a costume is linked to a role or how audience interaction is structured within the production concept. The aim is to make productions analytically accessible and comparatively queryable both retrospectively and from multiple perspectives.

The model was implemented on a dedicated Wikibase instance¹, building on the data structures of the WikiProject Performing Arts, with the explicit intention of extending rather than replacing existing approaches. The Wikibase environment enables flexible property definitions, complex SPARQL queries across production characteristics, and potential integration into existing LOD ecosystems. Three productions of Ferdinand von Schirach's *Terror* served as iterative case studies, selected because the play's fixed dramatic framework and built-in audience vote on the defendant's guilt or acquittal provide a productive basis for developing a generalisable model while accommodating considerable variation in staging decisions across different productions.

This research asks what can be known and modelled about a production through its traces, not how the live experience of a performance can be reconstructed, which remains irreducibly singular. The aim is to capture scenographic and structural features of productions in order to archive them, make them accessible and machine-readable, and enable forms of analysis not previously supported by available sources.

This research constitutes an initial contribution to a largely uncharted area of theatre data modelling. The model presented should be understood as a foundation for further research rather than a comprehensive solution. The following observations aim to identify the structural challenges that any LOD-based approach to modelling theatre productions will need to engage with.

A first set of challenges concerns the availability of sources. The materials available for modelling theatre productions are heterogeneous, fragmentary, and vary greatly in their informational content. They range from director's prompt books and video recordings to programme booklets, theatre reviews, and theatre-pedagogical materials. Access to these materials is often restricted by institutional constraints, inconsistent archiving practices, and copyright considerations. The data underlying this model were collected exclusively from freely Wikibase instance, available at: <https://mtm.uni-trier.de/wiki/Introduction>, available, non-copyright-protected sources to ensure transparency and traceability, and to make visible which data is actually accessible under these conditions.

A second set of challenges concerns the ontological infrastructure. Existing semantic standards, including the Wikidata schema, structures already developed by the WikiProject Performing Arts, CIDOC CRM, and FRBRoo, provide substantial frameworks for theatre and cultural heritage data, but were not developed with production-specific entities in mind. Properties suitable for representing stage elements, costume components, or interactive forms of audience participation are not present in existing vocabularies or ontologies. The development of new properties within the Wikibase instance was therefore necessary, but this complicates the integration of the dataset into broader LOD ecosystems; how to model production-specific entities in ways compatible with existing standards remains an open question for future research.

A third challenge concerns the interpretive nature of data collection. Modelling a theatre production is an inevitably interpretive act. Sources at the reception level, such as theatre reviews and production photographs, reflect the subjective perceptions of their authors and cannot straightforwardly be treated as objective accounts of a production. Decisions regarding how a stage element is classified, how an ambiguous spatial configuration is represented, or how contradictory sources are reconciled are shaped by the researcher's theoretical assumptions and disciplinary perspective.

These limitations define the scope of the present research and simultaneously constitute its most productive open questions. To what extent can the complexity of a theatrical production be formally represented? What is lost in the formalisation, and what becomes newly visible? Where do new biases arise in the modelling process? Which production features can be meaningfully captured through structured data, and where do the model's categories prove too imprecise or too rigid?

This research approach does not (yet) offer definitive answers to these questions. It proposes a framework for engaging with them more precisely, and demonstrates through concrete implementation that such an endeavour is methodologically feasible and analytically productive. The model is presented as a foundation: precise enough to address the research question and open enough to be expanded.

Cited works

Balme, C. (2021). Einführung in die Theaterwissenschaft [Introduction to Theatre Studies] (ESV basics, 6th revised and expanded edition). Schmidt, Erich.

Brincken, J. von & Enghart, A. (2008) Einführung in die moderne Theaterwissenschaft (Einführungen Germanistik) [Introduction to modern theatre studies. Introductions german studies). Darmstadt: Wiss. Buchges.

Probst, N., & Pinto, V. (2020). Re-Collecting Theatre History. Theaterhistoriografische Nachlassforschung mit Verfahren der Digital Humanities [Re-Collecting Theatre History. Theatre historiographical estate research using digital humanities methods]. In B. Wihstutz & B. Hoesch (Eds.), Neue Methoden der Theaterwissenschaft [New methods in theatre studies] (Vol. 133, pp. 157–179). Transcript Verlag.
<https://doi.org/10.1515/9783839452905-008>

From Live In Your Head: When Attitudes Become Form to Linked Data: a methodological proposal to describe exhibitions and processual contemporary artworks through the OntoExhibit ontology.

By Maria Serena Matarrese (Sapienza Università di Roma – Università Cattolica del Sacro Cuore di Milano) and Nuria Rodríguez Ortega (Universidad de Málaga).

In 1969 the curator Harald Szeemann presented at the Kunsthalle Bern an exhibition destined to become legendary, entitled *Live In Your Head: When Attitudes Become Form* – Works – Concepts – Processes – Situations – Information and imagined a space of great versatility that could also accommodate the extreme case of works that could not be exhibited permanently or statically. The new processual art therefore legitimately rejected the iconic image, although its attitude toward the event showed that it aspired to the non-representative image. The show, which included works by 69 international artists described as not object makers, «keine Objektmacher» (Szeemann, 1969, p. 5) was extensively documented. It was believed, in fact, that photography, and the accumulation of images, could ensure the preservation of works characterized by extreme obsolescence.

In recent years the art historian Beatrice von Bismarck has drawn attention to the «disappearance of the exhibited pieces». The photographic documentation of the works highlighted, in fact, the impossibility «to do justice, through illustration alone, to the quality of the artistic works within the context of the exhibition» and «to account for the relationships between the elements or to convey the spatio-temporal experience», ultimately resulting in the loss of «the continuation of the artistic working process» (Bismarck, 2013, p. 179, our translation). This limitation would highlight a defect of memory linked to the ephemeral nature of the works and to their visual recording. This suggests, in fact, that there was not a perfect match between the nature of the work, based on duration and process, and the technical means of documentation, which instead tended toward a logic of archiving and fixation.

Today, thanks to the Digital Humanities, we have the opportunity to open up new perspectives for the study and analysis of data. In particular, through specific ontologies, we are able to describe and query large bodies of information. But what happens when we are confronted with a contemporary artwork in which the image tends to dissolve into the anti-formal, the conceptual or the processual, and the exhibition is no longer just a place of display but becomes a device for understanding and interpretation?

A valuable conceptual framework for addressing this need is OntoExhibit, an ontology developed as part of a research program initiated in 2015 at the University of Málaga, encompassing three projects: the Exhibitium Project (2015–2016), the Andalex Project (2017–2022), and the Complexhibit Project (2022–2025). CIDOC-CRM and FRBRoo were used in particular to define certain classes and properties, after which the process continued more freely in order to adapt the model to the specific needs of art exhibitions. The most interesting aspect is precisely that «OntoExhibit regards the art exhibition field as a complex cultural ecosystem comprising heterogeneous actors, both human and non-human, engaged in an entanglement of discursive and social practices embodied in material relationships» (Rodríguez-Ortega, 2024, pp. 128-130).

This research, starting from the case of the 1969 exhibition seeks to outline a methodological proposal for the ontological representation of exhibitions and processual contemporary artworks that can be connected with the materials at our disposal. Using the OntoExhibit ontology, the aim will be to reflect not only on the specific requirements of the exhibition under analysis, but also on the broader needs that arise from artworks, which require specific attention to the artistic process.

Szeemann's intuition to conceive the exhibition as a laboratory rather than a mere display space proves to be an excellent example for the application of OntoExhibit, which understands the exhibition as a complex system. Finally, the possibilities offered by the ontology for the representation of process-based artworks demonstrate that, especially in the extreme case of a work that exists only within the temporality of an exhibition, it becomes essential to approach the understanding of the artwork not as a single, isolated object, but as an ensemble of interrelated entities.

Cited works

von Bismarck B., Harald Szeemann et l'art de l'exposition in "Perspective. La revue de l'INHA", Vol. 1, 2013, pp.176-182.

Rodríguez-Ortega, N., Contours of Knowledge: Epistemological Implications of Semantic Models in the Representation of the Art Exhibition Domain through the Lens of the OntoExhibit Ontology in "Život umjetnosti", 114 (1), 2024, pp. 122–147. <https://doi.org/10.31664/zu.2024.114.06>

Szeemann H. (curated by), When Attitudes Become Form. Works – Concepts – Processes – Situations – Information, Kunsthalle Bern, 22 March-27 April 1969. Exhibition catalogue with texts by S. Burton, G. Muller, J. A. Murphy, H. Szeemann and T. Trini, Stämpfli + Cie Ltd., Bern 1969.

Hotspots, Bypasses, and Semantic Drift: Mapping Cultural Knowledge Through Wikipedia's Biographical Network.

By Pachalis Agapitos (Donostia International Physics Center| CSIC-UPV/EHU); Juan Luis Suárez (CulturePlex Lab, Western University London) and Gustavo Ariel Schawrtz (Donostia International Physics Center| CSIC-UPV/EHU).

Cultural interactions are shaped by a complex web of relationships between individuals, places, and ideas. These relationships are inherently difficult to observe directly, and even harder to quantify at scale, because the everyday connections that constitute cultural life rarely leave a single, coherent record. What survives instead are fragmented traces, mediated through historical documentation, interpretation, and preservation. This problem becomes even more difficult the further we go to the past. However, in recent decades, the expansion of digital knowledge infrastructures has made it possible to study these mediated traces in new ways. Online encyclopedias, in particular, provide large-scale, systematically organised records of how knowledge about the past is curated and interconnected in the present. Wikipedia, the largest of these platforms, offers a corpus of biographical articles that document notable individuals while also encoding relationships between them through hyperlinks. These links do not only represent direct historical interactions but also they reflect the cumulative editorial decisions of contributors who interpret, select, and organise information about the past. As such, Wikipedia does not provide access to cultural interactions as they occurred, but to a structured representation of how those interactions are understood and recorded. Using Wikipedia as a proxy of interactions between cultural entities, this study performs a longitudinal analysis of Wikipedia biographies of notable individuals across 25 centuries. The research questions are the following: 1) What are some broad historical patterns that emerge from the interactions between people and fields of human activity (FHAs) across time? 2) Do these interactions change or remain stable across time? 3) Lastly, how does the interpretation of results change when we take into account that what we measure is not direct historical connection but what contemporary Wikipedia editors document and interconnect? To address these questions, the study draws on two methodological traditions: network data science and computational linguistics. Network analysis provides the formal tools to construct a directed and weighted network of biographical links, with particular attention to the interaction of scientists and philosophers across centuries. Complementing this structural perspective, computational linguistics provides word embeddings, a technique that transforms the textual content of biographical pages into

numerical representations, making it possible to map the cultural semantic landscape that emerges from editorial choices. While the network analysis focuses primarily on scientists and philosophers, the embeddings analysis extends to all biographical pages in the English Wikipedia, thereby capturing a broader picture of how cultural knowledge is organised.

The results of our analysis are twofold but they are connected to each other since both approaches reflect how the knowledge is organised in Wikipedia and how the cultural landscapes are being constructed by contemporary editors in a non-centralised manner. The network analysis reveals specific “hotspots” of connections between the examined FHAs as well as a topological bypass that happens between modern scientists and scientists of the Middle Ages¹; a bypass that it is not possible to detect unless we put the incoming and the outgoing connections next to each other. The embeddings analysis reveals the trajectory of the cultural semantic landscape of scientists and philosophers across 25 centuries which is characterised by clear clustering of broad historical periods and gradual change. In other words, biographies tend to connect to concepts, people, ideas, or places that are similar to ones people from the previous century tend to connect but at the same time slightly different; a finding that could reflect how knowledge is produced. These findings matter on two levels. Methodologically, they show that combining network analysis with word embeddings exposes structural features, such as the topological bypass between modern and medieval scientists, that neither approach reveals on its own. Conceptually, they reframe what large-scale analyses of Wikipedia can tell us: not how cultural interactions unfolded historically, but how they are being collectively reconstructed by contemporary editors, an increasingly consequential process given Wikipedia’s role in shaping public knowledge of the past.

¹ The term “scientists of the Middle Ages” is used here retrospectively to refer to individuals that were born during the period of Middle Ages but are presented as scientists in the English Wikipedia.

German Democratic Republic in Context: Reducing the Complexity of Historical Data through Visualization.

By Linda Freyberg (Leibniz Institute for Research and Information in Education) and Josefine Wähler, (Leibniz Institute for Research and Information in Education, Warschauer).

The application of digital tools and methods, including data visualization, necessarily requires a high standard of data quality. Historical data, however, are characteristically messy: the vast majority of archival sources remain undigitized, and the data tend to be heterogeneous, ambiguous, contradictory, incomplete, and incoherent, frequently lacking both standardized formatting and adequate semantic description on a metadata level.

Established quality metrics for metadata include completeness, accuracy, conformance to expectations, logical consistency and coherence, accessibility, timeliness, and provenance (Király, 2015, pp. 3-8). Gaps and uncertainties with respect to these criteria significantly increase the complexity of the data, a problem further compounded by the multimodality of the source material (Freyberg, 2024). One practical response is to adopt a pragmatic threshold of 'tolerable uncertainty', weighing the resource-intensive demands of data cleaning against the accuracy required for the specific research purpose and its representational goals. Bias constitutes a further critical challenge, particularly in the context of historical data, and calls for contextualization within the specific cultural and historical framework.

Data visualization has become an established method within the Digital Humanities as an access to complex data sets in order to provide guidance and visibility. We want to embrace visualization as a powerful method to present research results and above all as an explorative epistemic tool by triggering cognition and enabling new ideas (Morini et al. 2025, Freyberg 2021, 77ff./Stjernfelt 2007, 7), while dealing with 'messy data'. We specifically address the need for a humanities approach to the

graphical expression of historical data with more nuanced forms capable of representing ambiguity, uncertainty and complexity (Drucker 2011, 2; 9).

Our central research question is how messy historical data can be rendered accessible through visual means in order to provide overviews, reveal patterns, and make semantic relations within a dataset visible while reducing overall complexity.

Dataset: Pedagogical Lectures as heterogeneous sources

The historical dataset we focus on are the so-called Pedagogical Lectures (PL) as a distinctive and institutionalized feature of the educational culture and further teacher training in the German Democratic Republic (GDR). The PL are state-initiated written practice reports in which experienced teachers from all school types and subject areas, as well as other pedagogical personnel, reflect on and discuss their individual pedagogical experiences, classroom observations, and problem-oriented solutions for implementing the GDRs curricular requirements. The PL were presented, discussed, awarded at school, district and national levels, and subsequently circulated as so-called “experiences of the best”. At their peak, thousands were submitted annually, making the PL one of the largest practitioner-driven repositories of pedagogical knowledge in the GDR (Wähler & Hanke, 2020).

The corpus held at the BBF | Research Library for the History of Education at DIPF in Berlin comprises several thousand awarded lectures spanning nearly the full period of the GDR's existence (here 1961-1989). As a primarily text-based historical source, the PL are highly heterogeneous: they vary in length, structure, subject matter, school type, authorship, and regional provenance. Many are accompanied by supplementary audio-visual teaching materials or objects.

The lectures still exist in their original form as physical typescripts, with varying degrees of legibility and preservation. About $\frac{1}{3}$ have been digitized and enriched with selected metadata, though not consistently. This renders the corpus a paradigmatic case of the complex 'messy historical data' described above: rich in content and historically significant, yet resistant to straightforward computational processing and further research.

The PL thus represents both a challenge and an opportunity for Digital History Research. Their sheer volume makes manual analysis insufficient, while their heterogeneity and incompleteness demand a careful approach to any quantitative or visual analysis.

Visualization Concepts

Main entities of this historical source are: author/s (individual teachers, authors collectives), institutions (schools, school types, administrative districts), temporal metadata (such as years of submissions and presentation; signature), thematic/subject classifications (school subjects, pedagogical topics, ideological categories), geographic information (GDR districts and counties), textual content (full-text, abstracts, keywords when available). After the digitization and a proper formal and semantic description of the material, we propose the following approaches on the data expressed by visual forms:

- Overview and corpus structure: temporal distribution of contributions across decades, geographic spread across GDR districts, and distribution across school types and subjects

- Combination of timelines, maps and topics - Thematic clustering and semantic relations - Networks and topic clouds - Circulation, transfer and mobility of pedagogical knowledge - Hybrid visualizations combining maps and topics In our presentation, we will introduce these approaches through initial visual sketches and invite discussion of their viability and feasibility for research purposes. Our dataset is exemplary for historical data that not only generates cross-disciplinary and subject-area research questions, but for which new modes of visualization may serve as a productive entry point for further in-depth analysis. We seek to exchange ideas regarding visual approaches that go beyond already established forms such as maps, networks, and timelines.

Cited work

Drucker, J. (2011). "Humanities approaches to graphical display", in: Digital Humanities Quarterly, 5 (1).

Freyberg, L. (2021). Ikonizität der Information. Die Erkenntnisfunktion struktureller und gestalteter Bildlichkeit in der digitalen Wissensorganisation. Dissertation, Humboldt-Universität zu Berlin, <https://doi.org/10.18452/23813>.

Freyberg, L. et al. (2024). "Dealing with Uncertainty in Digital History Research - Reflections on the Research Data Process", in: In Karajgikar, J., Janco, A., & Otis, J. (2024). DH2024 Book of Abstracts. Zenodo. 192-194, <https://doi.org/10.5281/zenodo.13761079>.

Király, P. (2015). A Metadata Quality Assurance Framework, <http://pkiraly.github.io/metadata-quality-project-plan.pdf>.

Morini, F., Garretón, M., Pomerance, J., Zeissig, N., Guenther, S. d., Thomet, F., Freyberg, L., Kyriazis, I., Scholz, A. & Dörk, M. (2025). "Critical interactivity. Exploration and narration in data visualization", in: IEEE Computer Graphics and Applications, 45(3), 58-72. doi: 10.1109/MCG.2025.3544684.

Stjernfelt, F. (2007). Diagramatology. An Investigation on the Borderlines of Phenomenology, Ontology, and Semiotics. Dordrecht: Springer.

Wähler, J.; Hanke, M.-A. (2020). "Pacemaker report". GDR pedagogical innovators and the collection of Pädagogische Lesungen, 1952-1989, in: Paedagogica Historica. International Journal of the History of Education 58 (2022) 1, 66-83, <https://www.tandfonline.com/doi/full/10.1080/00309230.2020.1796720>.

Mapping Cultural Complexity in Networks of Writers Over the Last Three Centuries.

By Pachalis Agapitos (Donostia International Physics Center| CSIC-UPV/EHU), Gabriel Cwilich (Yeshiva University) and Gustavo Ariel Schawrtz (Donostia International Physics Center| CSIC-UPV/EHU).

Culture cannot be fully understood without facing complexity. Just as complex systems exhibit emergent behaviour based on how their elements are interconnected, the relationships among cultural entities give rise to emergent knowledge that is hardly imaginable from the individual components. In this context, the framework of complex networks provides the appropriate mathematical tools to unveil and make meaningful the knowledge hidden in cultural networks. Revolutionary ideas, cultural breakthroughs, or paradigm shifts emerge from complex interactions among many people, concepts, and objects that structure social, historical, and cultural spaces. These interactions are difficult to quantify, and the growing number of cultural elements to consider would eventually make it hard, if not impossible, to fully understand cultural dynamics using traditional tools. To overcome these limitations, several quantitative methodologies have been proposed recently (Schich et al., 2014; Goldfarb et al., 2015), some of which use Wikipedia as a suitable source for data mining and quantitative cultural research.

Recently, it has been shown that, in addition to the explicit knowledge presented in Wikipedia articles, a vast amount of implicit learning emerges from the dense network (~173M connections in the English version) of internal links that connect people, ideas, and works. This extensive network can be converted into a meaningful network of knowledge, using a metric based on the normalised Google distance (NGD) to measure the structural relatedness between each pair of Wikipedia entries (Schwartz, 2021). Using this methodology, it has been possible to quantitatively analyse the cultural network that relates Pablo Picasso, Albert Einstein and James Joyce. This approach also successfully addressed a multiscale analysis, allowing characterisation of individual nodes (degree and betweenness centrality, participation coefficient), clusters (size, density, openness) and the whole network (modularity, assortativity matrix). The same approach has also been applied to study the relationships among art, science and philosophy in Europe in the 17th century (Miccio et al., 2025).

In this work, we extend and deepen the proposed methodology to study the networks of prominent writers over the last three centuries. In particular, we analyse six writer networks comprising those born in the first and second halves of the 18th, 19th and 20th centuries.

The relatedness between each pair of writers has been calculated using the NGD. The resulting networks for each period have been sparsified using the so-called disparity filter, retaining only the most statistically significant links. Using a new clustering method based on link removal, we clearly identify specific clusters, such as British poets, comic writers, Italian authors and science fiction writers, among others. In addition, we perform a spectral analysis, using the Laplacian matrix, to quantify structural differences across periods and, more interestingly, to understand how the structure of writer networks has evolved over the last three centuries. Finally, the algebraic connectivity obtained from the spectral analysis indicates how fast

information diffuses in these networks. The results align with well-established literary and historical knowledge of the studied periods and, more importantly, provide new insights into how networks of writers grow and evolve.

Thus, by combining ideas from knowledge discovery in databases and complex network theory, we revealed the emergence of collective cultural dynamics and identified subtle connections among cultural elements on Wikipedia. It is important to emphasise that this approach is not limited to the Wikipedia corpus. The use of natural language processing techniques allows this methodology to be extended to any database, provided it contains sufficient information. Hence, this approach provides new insights into the structure of cultural networks, reveals previously unknown characteristics, and enhances quantitative studies in the humanities.

Cited works

Goldfarb, D., Merkl, D., & Schich, M. (2015). Quantifying cultural histories via person networks in Wikipedia. arXiv. <https://doi.org/10.48550/arXiv.1506.06580>

Miccio, L. A., Agapitos, P., Gámez-Pérez, C., González, F., Suárez, J. L., & Schwartz, G. A. (2025). Wikipedia as a cultural lens: A quantitative approach for exploring cultural networks. *Humanities and Social Sciences Communications*, 12, Article 462.

Schich, M., Song, C., Ahn, Y.-Y., Mirsky, A., Martino, M., Barabási, A.-L., & Helbing, D. A. (2014). Network framework of cultural history. *Science*, 345(6196), 558–562.

Schwartz, G. A. (2021). Complex networks reveal emergent interdisciplinary knowledge in Wikipedia. *Humanities and Social Sciences Communications*, 8, Article 127.

The Order of Indeterminacy: Complexity, Uncertainty, and the Black Box — The Limits of Modelling.

By Arantzazu Saratzaga Arregi (Frieburg im Bresigau).

Her research focuses on philosophies of complexity, cybernetics, systems theory, media philosophy, environmental philosophy, and matrixial philosophy, particularly on theories of operational closure, entropy, self-organization, opacity, and environmental embedding. Selected publications include *Matrixiale Philosophie* (transcript, 2019) and the forthcoming *Eine Einführung: Theorien der Selbstorganisation* (transcript/UTB, 2026).

Uncertainty and indeterminacy constitute a central component of the new worldview that emerged with the turn toward complexity (Cilliers 1998; Gershenson & Heylighen 2005). Complexity refers not only to a multiplicity of elements and interactions that resist calculation, but also to contingency, chance, and processes that cannot be fully determined. It designates an organized form of uncertainty. Complexity, therefore, is not merely a problem of quantity or computational capacity; rather, it marks the epistemological limits of observation and modelling themselves.

The paper examines this connection between complexity and uncertainty and conceptualizes it as a double bind of complexity research. Computational models, simulations, and algorithmic systems acknowledge uncertainty, emergence, and nonlinearity as constitutive properties of complex systems, while simultaneously attempting to reduce them through calculation, transparency, and modelling. In doing so, they produce precisely the forms of order that subsequently appear as objective descriptions of reality. Complexity reduction thus functions less as a neutral mode of inquiry than as an epistemic operation of producing order.

From this perspective, complexity should be understood less as an ontological property of the world than as an effect of specific modelling procedures. Within unstable and nonlinear processes, models determine the patterns, relations, and parameters of order through which “complexity” becomes visible and analysable in the first place. The question therefore shifts from the assumption of a given order to the epistemic conditions of its production. Complexity does not simply designate a condition of the world, but also the limits of the instruments through which the world is described. Models do not merely represent reality; they operate as performative procedures that generate visibility, order, and relevance. This constitutes the epistemological paradox of complexity research: it

attempts to calculate uncertainty without acknowledging that its own procedures are grounded in irreducible indeterminacy.

This double bind becomes particularly evident in the problem of the black box. In complexity research and contemporary AI debates, opacity is often treated as a deficit to be overcome through transparency, explainability, and interpretability. The “decoding” of the black box thus functions as an epistemic strategy: whatever becomes visible, calculable, and explainable is regarded as epistemic progress. Yet complex social, biological, and technical systems resist linear and causally closed explanatory models. The reduction of the black box therefore does not eliminate opacity; it merely displaces it. Visibility itself generates new zones of non-knowledge. The demand for complete transparency thus produces a paradoxical effect: the more complex systems are subjected to explanation, the more their limits of observability become apparent.

The black box is therefore not merely a metaphor for technical opacity, but the expression of a fundamental epistemological problem. Every act of observation operates selectively and necessarily produces a domain of the unobservable. Complexity research often responds to this problem through the expansion of data, computational capacities, and modelling techniques, without reflecting on the epistemological premises underlying these very procedures. It is precisely here that the double bind emerges: uncertainty is acknowledged as constitutive while simultaneously being treated as something that must be overcome. The result is a continuous production of new forms of opacity within systems that promise transparency.

Methodisch verbindet der Beitrag Komplexitätsforschung mit informations- und entropietheoretischen Ansätzen. Im Rückgriff auf Shannon, Szilárd und Brillouin sowie auf Arbeiten von Wicken, Landsberg, Gray, Blair und Gleick wird Information nicht als Lösungskonzept verstanden, sondern als Problembegriff. Information reduziert Ungewissheit, ohne sie vollständig zu eliminieren. Bereits Szilárds Verbindung von Information und Entropie sowie Brillouins Negentropie-Prinzip machen deutlich, dass jede Ordnung auf statistischen Wahrscheinlichkeiten basiert und einen irreduziblen Rest des Nichtbestimmbaren voraussetzt. Information erscheint damit nicht als Gegenbegriff zur Ungewissheit, sondern als operative Form ihrer Bearbeitung. Die Geschichte der Informationstheorie zeigt somit, dass Wissen niemals außerhalb von Entropie, Wahrscheinlichkeit und Unsicherheit operiert, sondern sich gerade aus deren Verarbeitung konstituiert.

Methodologically, the paper combines complexity research with approaches drawn from information theory and entropy theory. Drawing on Shannon, Szilárd, and Brillouin, as well as on the works of Wicken, Landsberg, Gray, Blair, and Gleick, information is understood

not as a solution concept, but as a problem concept. Information reduces uncertainty without ever fully eliminating it. Szilárd's connection between information and entropy, together with Brillouin's principle of negentropy, already demonstrates that every form of order is based on statistical probabilities and presupposes an irreducible remainder of indeterminacy. Information therefore appears not as the opposite of uncertainty, but as its operative mode of processing. The history of information theory thus reveals that knowledge never operates outside entropy, probability, and uncertainty, but is constituted precisely through their negotiation.

The paper intervenes in this debate through an epistemological perspective: the black box should not simply be reduced or dissolved, but integrated as a constitutive component of complex knowledge. Uncertainty thus appears not as a deficit, but as a condition of modelling itself. Complexity does not merely imply that prediction is difficult; rather, it indicates that no model can fully capture sudden and unexpected changes of state. Nonlinearity, emergence, and contingency therefore mark the limits of an epistemological paradigm grounded in complete calculability and transparency.

The paper aims to develop a model of knowledge that does not exclude opacity, unobservability, and indeterminacy, but integrates them productively. In contrast to contemporary discourses on AI, transparency, and algorithmic rationality, it proposes that complexity should not be understood only under the condition of its reducibility, but as a challenge for an epistemology capable of operating with black boxes. In this sense, the black box is understood not merely as an obstacle, but as an indication of the fundamental limits of all observation. The paper thus argues for a shift from a paradigm of control toward an epistemological model that recognizes uncertainty and opacity as constitutive conditions of knowledge.

Cited works

Atlan, Henri (1979): *Entre le cristal et la fumée. Essai sur l'organisation du vivant*. Paris: Seuil.

Blair, Ann (2021): „Communication, Computation, and Information“. In: Paul Duguid / M. T. Hancher / W. J. Turkel (Hg.): *Information: A Historical Companion*. Princeton: Princeton University Press, 65–72.

Brillouin, Léon (1962): *Science and Information Theory*. 2. Aufl. New York: Academic Press.

Cilliers, Paul (1998): *Complexity and Postmodernism: Understanding Complex Systems*. London/New York: Routledge.

Foerster, Heinz von (1997): *Understanding Understanding: Essays on Cybernetics and Cognition*. New York: Springer.

- Foerster, Heinz von (2009): *Ethics and Second-Order Cybernetics*. Wien/New York: Springer.
- Gershenson, Carlos / Heylighen, Francis (2005): „How Can We Think Complex?“. In: K. Richardson (Hg.): *Managing Organizational Complexity*. Greenwich: Information Age Publishing, 47–62.
- Glanville, Ranulph (2012): „Black Boxes“. In: Ranulph Glanville: *The Black Box*, Volume 1: *Cybernetic Circles*. Wien: edition echoraum, 417–430.
- Glaserfeld, Ernst von (1995): *Radical Constructivism: A Way of Knowing and Learning*. London: Falmer Press.
- Gleick, James (2011): *The Information: A History, a Theory, a Flood*. New York: Pantheon Books.
- Gray, Robert M. (2011): *Entropy and Information Theory*. 2. Aufl. New York: Springer.
- Landsberg, Peter T. (1984): „Entropy and the Unity of Knowledge“. In: *Entropy and Information Theory*, 1–15.
- Luhmann, Niklas (1990): „Sthenographie“. In: *Soziologische Aufklärung 5: Konstruktivistische Perspektiven*. Opladen: Westdeutscher Verlag, 119–137.
- Luhmann, Niklas (1991): *Soziologie des Risikos*. Berlin/New York: de Gruyter.
- Luhmann, Niklas (2017): *Einführung in die Systemtheorie*. Heidelberg: Carl-Auer.
- Maturana, Humberto R. / Varela, Francisco J. (1987): *Der Baum der Erkenntnis: Die biologischen Wurzeln menschlichen Erkennens*. Bern/München: Scherz.
- Mitchell, Melanie (2009): *Complexity: A Guided Tour*. Oxford: Oxford University Press.
- Morin, Edgar (2008): *On Complexity*. Cresskill, NJ: Hampton Press.
- Parrish, Julia K. / Edelstein-Keshet, Leah (1999): „Complexity, Pattern, and Evolutionary Trade-Offs in Animal Aggregation“. In: *Science* 284 (5411), 99–101.
- Shannon, Claude E. / Weaver, Warren (1949): *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
- Szilárd, Leó (1929): „Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen“. In: *Zeitschrift für Physik* 53, 840–856.
- Szilard, Leo (1964 [1929]): „On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings“. In: *Behavioral Science* 9 (4), 301–310.
- Wicken, Jeffrey S. (1987): „Entropy and Information: Suggestions for a Common Language“. In: *Philosophy of Science* 54 (2), 176–193.

Modeling an Intermedial Ecosystem: A Computational Approach to Gendered Canon Formation in 1930s Madrid.

By: Irene Palencia Mora (Universidad Nacional de Educación a Distancia, UNED)

The Second Spanish Republic (1931–1936) represented a volatile and historically stratified cultural ecosystem. While the period was characterized by professional mobilization for women and important legal advances—such as the right to vote in 1931 and the Law of Divorce in 1932—it was simultaneously governed by deep-seated aesthetic rules that functioned as systemic filters, dictating norms that rendered certain types of art invisible. This paper conceptualizes the musical life of 1930s Madrid not merely as a series of concerts, but as a complex system of interdependencies among journalistic infrastructures (the press), patronage regimes, and gendered hierarchies inherent in music.

The research aims to unveil the power dynamics that led to the marginalization of prominent cultural figures such as María Rodrigo, Adela Anaya, and Rosa García Ascot, arguing that their exclusion was not an accidental omission but a systemic output of a specific cultural logic. During this period, European music criticism inherited from the 19th century elevated the concept of “absolute music” (abstract, intellectual, and coded as masculine) as the most desirable form of artistic production (Dahlhaus, 1978; Chua, 1999). Conversely, referential and intermedial musical forms—especially those linked to literature, functionality, and coded as feminine—were relegated to a secondary status (McClary, 1991). The aesthetic sublime in music was defined as the wholly autonomous work, thought to be self-referential and transcendent, distanced from sensuousness and contact with other arts. By treating these aesthetic dichotomies as interacting forces, it is possible to model how critical bias in dailies such as *El Sol* and *ABC Madrid* functioned as a regulative mechanism that governed visibility and access to the canon.

Consequently, the research questions are:

1. How can the systemic marginalization of women in music be modeled as a non-linear dynamic within the intermedial ecosystem of the Second Spanish Republic?
2. Specifically, what does a formal representation of the press-mediated intersections between music and literature reveal about the reinforcing patterns of exclusion that sustain androcentric canons?

To render this complexity analytically tractable, the study employs a relational modeling approach using the Heurist platform. The cultural ecosystem of 1930s Madrid is modeled through three integrated hermeneutic axes:

- Musical axis: The formalization of entities such as performers, organizers, scenic spaces, and musical works (including full programs) to map circuits of exchange within the city.
- Referential and intermedial axis: The mapping of links between musical compositions and literature—classifying literary works, authorship, genre, and themes—to identify how intermediality influenced a work's reputation and the surrounding discourse. The classification and closed vocabularies are based on intermedial theory (Wolf, 1999, 2015; Rajewsky, 2005; Gil González and Pardo, 2018).
- Discursive axis: The compilation of various types of media coverage (announcements, notices, theater listings, and music reviews) regarding concert life in Madrid from digitized media. Following transcription, Natural Language Processing (NLP) and semantic tagging are applied to identify the rhetoric of valuation and critical bias.

Structuring these data into a relational ontology allows us to move beyond simple measurement. The objective is to create a traceable model where specific critical reviews are linked to a work's visibility, enabling a simulation of how gendered discourse regarding the canon effectively de-platforms female creators and the specific artistic genres associated with them. This approach situates the long-standing question of the interchange between gender and genre (Harris and Nochlin, 1977; Gilbert and Guber, 1980; Citron, 1993) within a cultural ecosystem where disciplines are not isolated but work together to generate the inclusions and exclusions that formed the contemporary canon.

The evaluation strategy involves comparative analytics: contrasting the frequency and sentiment of reviews for male-authored absolute music against female-authored intermedial music. It also examines the professional roles occupied by women (e.g., as performers rather than organizers, composers, or critics) and the frequency with which these roles are mentioned in the press. By formalizing uncertainty—such as anonymous reviews or missing program notes—the study ensures that computational claims are grounded in a critical understanding of the archive's limitations. This is essential given that not all historical data can be compiled, particularly due to the disparate nature of

events and music halls. The model does not replace interpretation; rather, it provides an empirical map of the unequal circulation that defined the era.

Ultimately, this study demonstrates that computational forms of inquiry are highly productive for examining complex cultural objects that emerge from multiple traditions

and are shaped by varying artistic rules, media shifts, and discursive authority. By modeling 1930s Madrid as a cultural ecosystem, the durable asymmetries of the Spanish canon can be visualized and interpreted. This approach offers a scalable framework for understanding how ideological tensions at the intersection of media, gender, and art constitute the non-linear dynamics of cultural memory.

Modelling Visual Complexity in Performance Photography: Period, Venue, and Photographic Mediation at the Festival d'Avignon.

By Giacomo Alliaa (Université de Rennes 2).

Performance photography offers one of the richest visual records of contemporary performing arts, but it is also a highly mediated source: photographs encode not only performances, but also venue architecture, lighting, framing, archival selection, and institutional conventions of documentation. At the Festival d'Avignon, one of the premiere venues for the contemporary performing scene, for instance, official photographer Christophe Raynaud de Lage has produced hundreds of thousands of images over two artistic directions, first with Hortense Archambault and Vincent Baudriller from 2004 to 2013, and then with Olivier Py from 2014 to 2022. Within the field of performing arts analytics (Bardiot, 2021), such a collection represents a wealth of visual traces allowing us to study how performances have been visually documented, selected, and represented over time.

However, because of the mediated nature of performance photography, a computational analysis of such a collection needs to model this complexity to avoid over-claiming an analysis of performing arts aesthetics, in line with the inquiries that are informing research in the computational study of visual culture (Manovich, 2020; Arnold and Tilton, 2023). Specifically, I ask how photographic traces of performance encode the interaction between two periods of artistic direction, venue-specific spatial conditions, and evolving conventions of performing arts representation at the Festival d'Avignon. The paper contributes a framework for distinguishing between global period effects, venue-mediated visual regimes, and artefacts of photographic or archival mediation.

The study examines 34,000 photos personally selected by Raynaud de Lage for archiving at the Bibliothèque Nationale de France. After removing near-duplicate images with perceptual hash similarity and limiting the analysis to venues sufficiently represented in both periods, I arrive at a dataset of 21,000 photographs. Genre is excluded from the statistical modelling because available labels are promotional, inconsistent, and conceptually unstable in contemporary/postdramatic performance.

On this controlled dataset, I compute several features to model the complexity of the photographic mediation of performing arts through different, complementary angles. On the one hand, I extract visually interpretable features, both at the level of the image itself and based on 3D human poses estimated relative to the camera with SAM 3D Body (Yang et al., 2026). On the other hand, I model each image as a high-dimensional feature vector

using the foundation vision model DINOv3 (Siméoni et al., 2025). These embeddings are used as a contrastive example to the idiosyncratic features described above, assessing whether they can encode period or venue.

For each interpretable visual feature, I compare three generalised linear models: a period-only model, an additive period-plus-venue model, and a period-by-venue interaction model; with performance-level clustered standard errors to account for repeated images from the same performance. The interaction model is consistently preferred using AIC weights, suggesting that period-related visual changes are not global shifts across the archive, but are expressed differently

across venues. The Jardin de la Vierge du lycée Saint-Joseph appears here as a distinct venue, presenting strong evidence of a close and compact visual composition, while the Cour d'Honneur du Palais des Papes is the opposite, with many detected bodies, smaller body scale, and high spread. Overall, the Olivier Py period shifts toward images with more but smaller bodies, and is significantly darker. These changes, however, are not consistent across venues, contrasting a simpler global period effect.

For the multi-dimensional feature vectors, I conduct a performance-aware cross-validation classification task to predict period and location, both globally and within each instance of the controlling variable. Grouping images by performance is particularly important to avoid leakage of images from the same performance across train and test sets. To strengthen the results, random permutations are performed to statistically test the classification results obtained. Both period and venue are predicted substantially above the chance baseline. Within-venue and within-period classifications further test whether these signals persist after holding one metadata dimension constant. The embeddings are able to predict the period within 9 of the 13 venues, reaching 0.941 balanced accuracy for the Jardin de la Vierge, highlighting yet again its distinct visual regime. Similarly, within each period, location can be reliably predicted (with balanced accuracies of 0.418 for the first period, and 0.401 for the second).

Lastly, I inspect the computational findings through specific visualisations to contextualise the model outputs (Manovich, 2020). As an example, Figure 1 shows the DINOv3 features of photographs taken at the Jardin de la Vierge projected through UMAP (McInnes et al., 2018), coloured by period. It reveals that in the second period the venue undergoes a major shift with the appearance of ivy on the walls, producing an important visual shift that might explain part of the distinctive regime measured at this particular location. Additionally, even though other metadata dimensions such as genre could not be modelled, they can still further inform the results obtained. Indeed, all performances at the Jardin are part of an ongoing series, Sujets à Vif, which explicitly mixes disciplines, thus complicating the simpler reading that this venue is architecturally different.

In conclusion, the paper argues for a model of performing arts analytics in which computational methods of distant viewing such as classification and statistical modelling, are critically balanced with close reading to make cautious claims under complex conditions of archival mediation and entangled aesthetic variables.

Acknowledgements

Funded by the European Union (ERC, STAGE, grant agreement no. 101097091). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Figure 1: UMAP visualisation of the DINOv3 visual embeddings for images taken at the Jardin de la Vierge, coloured by period (blue: Archambault and Baudriller; orange: Olivier Py).

Cited works

Arnold, T., & Tilton, L. (2023). Distant Viewing: Computational Exploration of Digital Images. The MIT Press. <https://doi.org/10.7551/mitpress/14046.001.0001>

Bardiot, C. (2021). Performing Arts and Digital Humanities: From Traces to Data. Hoboken: ISTE / Wiley.

Broadwell, P., Rau, M., & Wiles, S. (2025). Stylistic Analyses of Human Pose in Theatrical Performances: Computational and Historical Frameworks. *Anthology of Computers and the Humanities*, 3, 1482-1493. <https://doi.org/10.63744/GxQbSDPPyqVL>

McInnes, L., Healy, J., & Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.

Oiva et al. (2024) Soviet View of the World. Exploring Long-Term Visual Patterns in "Novosti dnia" Newsreel Journal (1945-1992). In *Journal of Cultural Analytics*, vol 9, issue 4.

Raynaud de Lage, C., and Gachet, L. (2022). L'oeil présent: Photographier le Festival d'Avignon au risque de l'instant suspendu. Festival d'Avignon Théâtre(s).

Siméoni et al. (2025) DINOv3. ArXiv preprint arXiv:2508.10104

Tilton et al. (2025) Seeing Formalism or Formal Viewing: Computational Formalism for the Analysis of Visual Media Forms and Contexts. In *Journal of Cultural Analytics*, vol 10, issue 3.

Yang, X. et al. (2026) SAM 3D Body: Robust Full-Body Human Mesh Recovery, arXiv preprint arXiv:2602.15989.

Regrouping the Avant-Garde: A Multiscalar Network Approach to European Artist Groups (1905–1915).

By Teresa Kamencek, (CReA Lab, University of Vienna); Velitchko Filipov (CVASt, TU Wien); Michaela Tuscher (CVASt, TU Wien); Silvia Miksch (CVASt, TU); Raphael Rosenberg (CReA Lab, University of Vienna).

Around 1900, artist groups provided the most important framework for artistic development in Europe. Groups like Der Blaue Reiter, Jack of Diamonds, the Futurists, and the Moderne Kunstkring established themselves in quick succession in various cities across Europe, making their presence felt through manifestos, institutional structures, periodicals, and, above all, through collaborative exhibitions. So far, art history has adopted two research approaches. On the one hand, only a handful of those groups were examined from a monographic perspective, focusing on individual prominent members, their biographies, and stylistic development.² On the other hand, there have been some exhaustive lists of the individuals in these groups, especially by Christoph Wilhelmi, whose three-volume work compiles membership lists for artist groups across Europe dating from 1900 onwards.³ Wilhelmi's lists, however, remain a reference work and have not yet been utilised for comparative analysis. We address this gap by asking: What were the benefits of joining an early 20th-century avant-garde artist group? How did these groups interact through their exhibition activities and how did they change in terms of time and space? To what extent can their formation and activity be explained by similarities among their members? We draw on the Database of Modern Exhibitions⁴ (DoME), representing the

² See, for example: Doris Fässler (Ed.), *Der Moderne Bund. Beginn der Moderne in der Schweiz*, Luzern 2011; Agnes Husslein-Arco (Ed.), *Hagenbund. Ein europäisches Netzwerk der Moderne 1900 bis 1938*, München/Wien 2014; Matthias Mühling/Annegret Hoberg/Anna Straetmans (Eds.), *Gruppendynamik - Der Blaue Reiter*, Berlin 2021; Iris Müller-Westermann, *German Expressionism. The artist group Brücke and the beginnings of Modernism*, München 2024.

³ Christoph Wilhelmi, *Künstlergruppen in Deutschland, Österreich und der Schweiz seit 1900. Ein Handbuch*, Stuttgart 1996; Christoph Wilhelmi, *Künstlergruppen im östlichen und südlichen Europa seit 1900. Ein Handbuch*, Stuttgart 2001; Christoph Wilhelmi, *Künstlergruppen in West- und Nordeuropa einschließlich Spanien und Portugal seit 1900. Ein Handbuch*, Stuttgart 2006.

⁴ <https://exhibitions.univie.ac.at/> [last accessed on 05/05/2026].

most comprehensive dataset of modern art exhibitions available for this period.⁵ Based on exhibition catalogues, DoME covers over 13,200 artists and around 1,300 exhibitions with over 207,000 catalogue entries, organised in 100 cities by 219 organisations. The project ArtVis: Dynamic Network Perspectives on Digital Art History develops open-access network tools⁶ to visualise the connectivity of exhibition data from DoME, making the relationships between artists and exhibitions comparable and comprehensible for the first time. In order to apply the same quantifiable metrics to the more complex phenomenon of artist groups, we are developing a multiscalar, three-level methodology:

I. We have extracted Wilhelmi's membership lists and integrated them with DoME. This initial step allowed us to expand our graph database with expert knowledge to include the entity "Group" and the relationship "member of" (Fig. 1), which, in turn, augmented the scope of our query possibilities. We are now able to investigate where and how frequently groups have exhibited, which members have exhibited, how often, and which artist groups were connected by artists with multiple memberships (Fig. 2). Where possible, we further added information drawn from primary sources and secondary literature regarding the founding dates of the groups, their city of origin, and the duration of members' affiliations. We will be using network visualisation to provide a base for macro-level overview of the dynamic changes in the network and connectivity of these associations between 1905 and 1915. In addition to facilitating analyses of formation processes and membership flux, this further refinement lays the foundation for future micro-level approaches to questions of style and artistic strategy.

II. We then perform data-driven clustering on all the artists associated with a historically declared group. In accordance with Kappelhof, we define a cluster as any densely interconnected subgroup within an overall network.⁷ The identification of these subgroups allows us to map them against the self-defined artist groups. We measure similarity based on demographic characteristics, as well as on geography and exhibition activity. We also include the temporal component (1905-1915), leading to 214 similarity matrices in total. This allows us to choose which ones to combine and to assign weights for each one to prioritise certain attributes or years.

⁵ Christina Bartosch et al., *The Database of Modern Exhibitions (DoME): European Paintings and Drawings 1905-1915*, in: Katryn Brown (Ed.), *The Routledge Companion to Digital Humanities and Art History*, New York 2020, 423–434.

⁶ <https://artvis.cvast.tuwien.ac.at/> [last accessed 05/05/2026].

⁷ Peter Kappelhof, *Cliquenanalyse. Die Bestimmung von intern verbundenen Teilgruppen in sozialen Netzwerken*, in: Franz Urban Pappi (Ed.), *Methoden der Netzwerkanalyse*, Berlin/Boston 1987, 39–63.

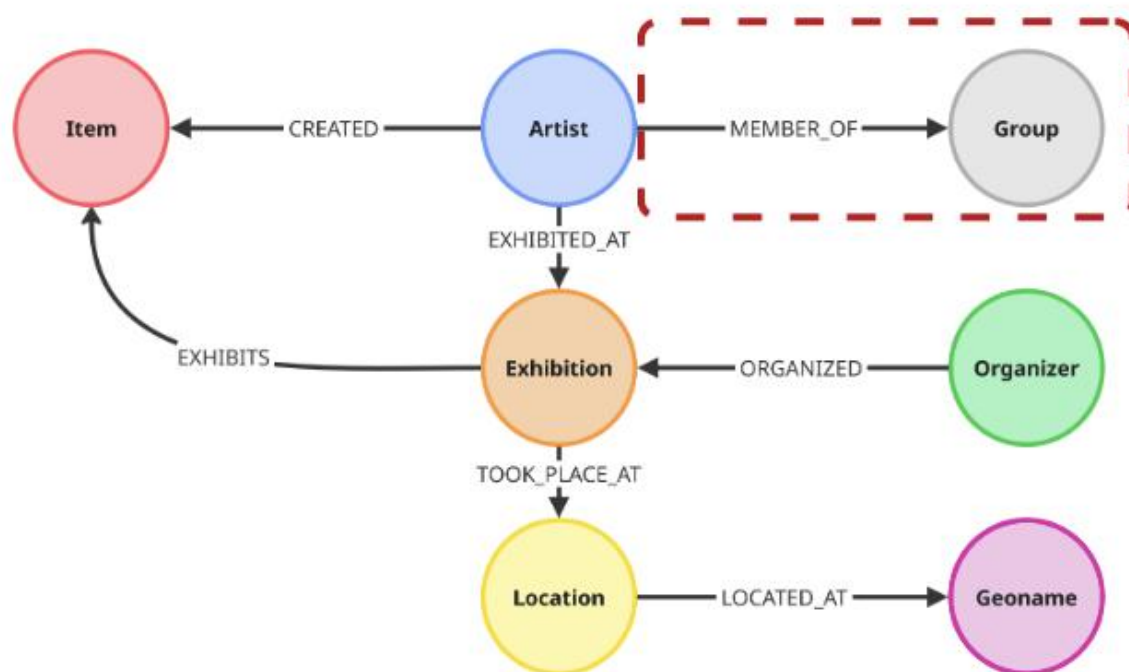


Fig. 1: The database schema showing the incorporated entity “Group” and membership “member of”, connecting the group membership data to exhibition catalogue data.

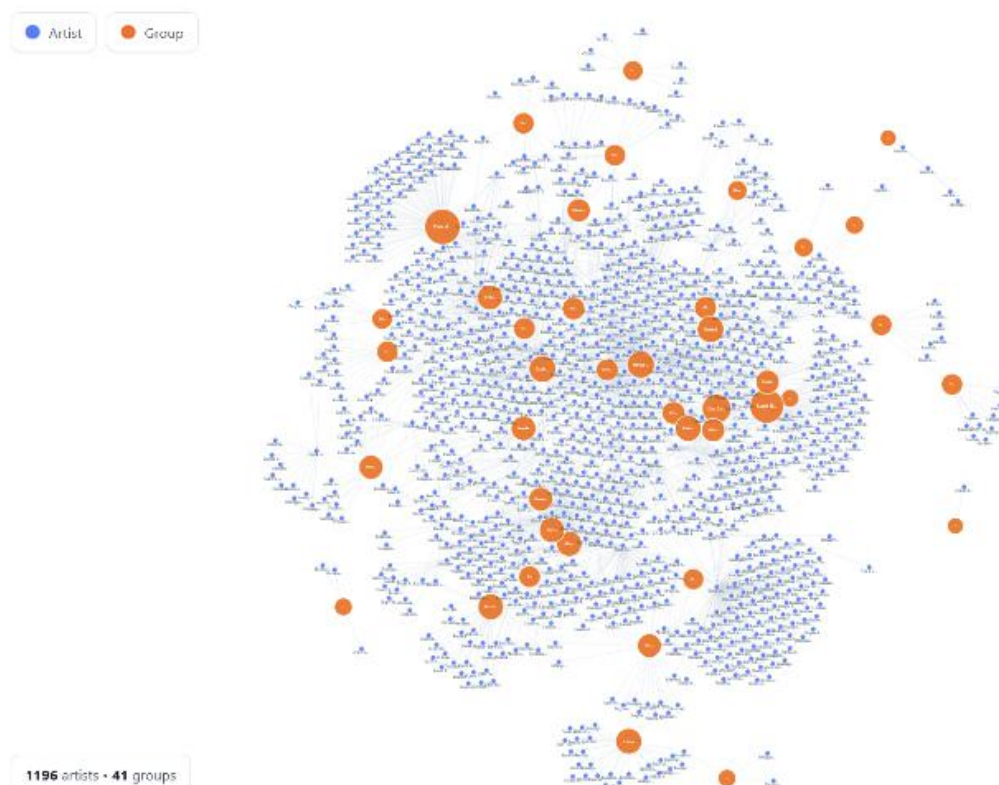


Fig. 2: Network visualisation of all artist groups (in orange) recorded in DoME, along with their members (in blue). The nodes representing the groups are of varying sizes according to the number of their members.

III. We will then return the computationally identified clusters back to the art-historical record for interpretive analysis. For example, we focus on cases where the data-driven analysis reveals mismatches; i.e. groups whose members are demographically or geographically dissimilar, or conversely, artists who cluster together computationally but never formally organised. A case in point is *Der Blaue Reiter*, whose members show remarkably low demographic similarity, proving that their alliance was driven by shared strategic or aesthetic goals rather than proximity. In this way, we can explore not only which groups formed along lines of mere similarity, but what the deviations from these patterns reveal about the social, strategic, and aesthetic motivations behind group formation.

This three-level approach allows us to assess the connectivities and evolution of modern artist groups over time and space, as well as their impact on members' careers. For example, we found that group membership led to a significantly higher rate of average exhibition participation and, consequently, greater artistic visibility.⁸ By rendering the complexity of these groups visible and measurable for the first time, we lay the groundwork for a fundamentally revised account of group formation as one of the defining phenomena of early European modernism.

⁸ For example, our analysis shows that group members participated in an average of 7.6 exhibitions, compared to just 2.8 for non-members.

Historical Frames. Interacting with historical change and complexity in cultural heritage ecosystems.

By Sofia Baroncini (Leibniz Institute of European History, IEG) and Charles van den Heuvel (Knowledge & Art Practices Huygens Institute, Amsterdam Netherlands).

The value of Cultural heritage (CH) artefacts primarily arises from the socio-cultural systems in which they are embedded. Here, they acquire functions, meanings, and symbolic relevance. Nevertheless, semantic representations of cultural heritage often privilege the description of their material aspects while the complexity of the broader cultural contexts in which artworks traverse remain difficult to formalize. Despite current developments attempt to formalize socio- cultural aspects of culture, (see, Art and Architectural Argumentation Ontology (AAAO))⁹, there is a need to include aspects of uncertainty, conflicting interpretations, provenance of heterogenous sources in conceptualizations and models to inquire and represent the complexity of past cultural heritage knowledge.

In this sense, following Drucker (2011), Piotrowski (2023), Pasqual (2024), and Vitali & Pasqual (2025), we consider structured data about a reconstructed past as the interpretative result of scholars ("capta") in which a narrative is applied to make sense of the source data.

This contribution reflects on the implications of formally representing such interpretive knowledge by discussing the "Historical Frame class" through the recently released Cultural Heritage Interactions ontology (CHint)¹⁰. The model attempts to embed the CH objects in the broader socio-cultural contexts they traverse during their existence by representing not only their functions and meanings, but also related contextual socio-cultural traits and immaterial practices reconstructed by the scholarship (van den Heuvel & Baroncini, 2025; Baroncini & van den Heuvel, 2026). The following research question will be addressed: how can semantic models contribute to analysing and representing reconstructed socio-cultural knowledge while preserving provenance and uncertainty and acknowledging conflicting interpretations? More specifically, the contribution examines what it means to formalize forms of "historical common-sense" that are neither objective nor directly observable, but emerge from the interpretive work performed by scholars while making sense of heterogeneous evidence.

⁹ <https://www.sari.uzh.ch/en/ordea/aaao.html>

¹⁰ <https://github.com/SofiBar/CulturalHeritageHistoricalContext>

The mentioned Historical Frame class was developed by building upon frame theories in linguistics, AI, information technology, and art history. It functions as a structure “framing” CH object(s) in a context, namely, a portion of the spacetime continuum, and relating them with traits recognized as relevant for their understanding. Frames can include or overlap with further ones, constituting themselves a network to observe objects at different levels of granularity and group relevant traits. In this way, artworks can be analysed in relation to each other, other sources, immaterial CH manifestations (e.g., celebrations), and the traits of cultural ecosystems in which they acquired meaning. At the same time, the model explicitly acknowledges the instability of these reconstructions. Historical frames are not intended as objective representations of past societies, but as formalized aggregations of interpretive knowledge.

From this perspective, the ontology becomes the means to gather and bridge existing information about common contexts shared by CH objects of different nature (material, immaterial). It potentially becomes an epistemic means to critically examine and represent the common-sense knowledge emerging from the domain, highlighting canons and assumptions, while preserving complexity and uncertainty rather than concealing them behind apparently neutral data structures.

Cited works

Baroncini, S., & van den Heuvel, C. (2026). The Historical Framing Problem: Temporal Modeling of Interactions between Tangible and Intangible Cultural Heritage [accepted]. <https://semantic-web-journal.net/content/historical-framing-problem-temporal-modeling-interactions-between-tangible-and-intangible-0>

Beretta, F. (2024). Semantic Data for Humanities and Social Sciences (SDHSS): An Ecosystem of CIDOC CRM Extensions for Research Data Production and Reuse. In T. Wolf, Technische Universität Braunschweig, P. Schiele, Stiftung LEUCOREA Lutherstadt Wittenberg, F. Beretta, Laboratoire de recherche historique Rhône-Alpes, CNRS / Université de Lyon, C. Erlinger, Zentral- und Hochschulbibliothek Luzern, M. Görmar, & Herzog August Bibliothek Wolfenbüttel, Professorale Karrieremuster Reloaded. Entwicklung einer wissenschaftlichen Methode zur Forschung auf online verfügbaren und verteilten Forschungsdatenbanken der Universitätsgeschichte. HTWK Leipzig / OA-HVerlag. <https://doi.org/10.33968/9783966270502-05>

Drucker, J. (2011). Humanities Approaches to Graphical Display: Digital Humanities Quarterly 5(1) <https://doi.org/10.63744/r4ysrh7ae534>

Pasqual, V. (2025). The critical inquiry in humanities knowledge graphs: Challenges, methods and innovations [Doctoral Thesis, Alma Mater Studiorum - Università di Bologna]. <https://doi.org/10.48676/unibo/amsdottorato/12142>

van den Heuvel, C., & Baroncini, S. (2025, October 16). Opening Boxes: Contextualizing a 3D Scholarly Edition of the Restoration of the 12th Century Reliquary Shrine of St Servatius with Ontologies of Uncertainty [Conference Presentation]. Digital Art History – Methods, Practices, Epistemologies V Critical Approaches to Sources in (Digital) Art History, Zagreb. <https://www.ipu.hr/article/en/2017/digital-art-history-methods-practices-epistemologies-v>

Vitali, F., & Pasqual, V. (2025). The Representation of Historical Uncertainties as the Outcome of Competing and Incompatible Certainties. In Models of Data Extraction and Architecture in Relational Databases of Early Modern Private Political Archives (Edizioni Ca' Foscari-Venice University Press, pp. 61–77). Fondazione Università Ca' Foscari.

Fragile Networks: Latent Comprehension and the Hermeneutics of LLM-Based Citation Extraction in Art History.

By Andrea Alfarano, Eduardo Trabattoni and Dario Negueruela del Castillo (University of Zurich, UZH; Switzerland Max Planck Institute).

Large Language Models are entering the evidentiary infrastructure of the humanities, from named- entity extraction in archival and curatorial texts (van Hooland et al., 2015; Luthra et al., 2024) to the LLM-assisted mining of bibliographic references for citation-network reconstruction, a task long pursued in the humanities (Gingras, 2010; Colavizza, 2018). Tasks that once required years of manual annotation can now be approached at scale, with implications for how the discipline studies circulation, canon formation, and its own memory. Yet the reading instrument itself hallucinates (Huang et al., 2023), and the entropy-based methods we rely on to flag its hallucinations turn out, in this setting, to fail at chance. The damage, when it occurs, is not at the level of the individual edge but at the level of the structure the network was built to reveal.

Citation networks are paradigmatic instances of the complex cultural systems this conference fore- grounds: non-linear, multi-scalar, path-dependent, and emergent. In our setting, an LLM extracts bibliographic references from art-historical articles; these references form the nodes of the citation network, and one reference citing another forms an edge. The network's analytical value rests not on individual edges but on higher-order properties such as centrality, community structure, broker- age, and clustering (Borgatti et al., 2009), which encode patterns of canon formation, reputational cascades, and durable asymmetries of visibility and circulation. Graph theory has long shown how small perturbations propagate through complex topologies: long-range shortcuts in small-world networks, percolation thresholds, the disproportionate weight of hubs in scale-free networks, the volatility of modularity under noise. A modest fraction of fabricated edges can in principle shift centrality rankings, redraw community boundaries, and alter inferences about diffusion and influ- ence. The distortion is silent, because the network metrics themselves cannot tell the topology from its hallucinated double.

How prone current LLMs are to producing such fabricated edges is an empirical question. We evaluate the Gemma 4 family across model sizes from 2 to 31 billion parameters (Gemma Team, 2026) on three citation-extraction datasets: CEX (Cioffi, 2022; Pagnotta, 2024), EXCITE (Hosseini et al., 2019), and LinkedBooks (Colavizza and Romanello, 2017), using AUROC for the binary classification of extracted citations as faithful or hallucinated. We ask where, if anywhere, an emergence of latent comprehension makes itself legible,

and we test this at two distinct sites. The first lies inside the model: as it generates, the distribution from which it draws each token may reveal an emergence of latent comprehension, and with it the trace of hallucination (Malinin and Gales, 2021). The second works from the outside: the model is asked to reflect on its own generation in natural language — to assign a verbalised confidence, qualify, hedge or commit (Lin et al., 2022). We find that hallucination is not marginal: a non-trivial share of extracted citations does not correspond to any reference in the source, introducing edges with no provenance. Hallucination rates decrease with model scale but do not vanish at the upper end, excluding the consolation that the problem is an artefact of under-resourced models. The more consequential finding concerns detection. No emergence of latent comprehension can be read from the entropy of the token distribution over the generated reference span: discriminative power is indistinguishable from a coin toss (AUROC 0.55). The reason is structural: when extracting from a long, well-structured source, the task itself forces the model into confident outputs regardless of whether they correspond to anything in the source. Reflection fares substantially better (AUROC 0.77), yet remains far from reliable.

The remaining gap between 0.77 and reliable detection motivates a third direction: probing the latent space directly, in search of an emergence of comprehension deeper than the surface that can distinguish recovery from fabrication (Azaria and Mitchell, 2023). Our hypothesis is that this latent-space legibility, established in prior work on factual generation, survives in the more constrained regime of extractive tasks over long, structured inputs — precisely the regime in which surface entropy signals collapse to chance. This reframing shifts the register from asking whether the model is uncertain to asking what it means, internally, for a citation to be recovered from a source rather than fabricated from the model's prior knowledge. Such a latent-space orientation is not alien to humanistic method but its extension. In Panofsky's iconological reading of intrinsic meaning (Panofsky, 1955), Carpo's analysis of computational design (Carpo, 2011, 2017), and Paglen's confrontation with operational images (Paglen, 2014, 2019), art history has always interpreted how meaning is constituted. Probing the model's internal trajectory is a hermeneutic act: it treats the LLM not as a black box but as a system whose meaning-making must be critically interpreted.

The implication is not that LLMs should be excluded from art-historical knowledge-graph construction, but that their outputs cannot be treated as transparent evidence. If fabricated citations can enter the network confidently and evade the detection methods currently available, responsible use requires what Daston and Galison (2007) would recognize as a new epistemic virtue for the computational humanities: not the old objectivity of the manual index, but an auditability capable of distinguishing recovery from invention before those inventions sediment into the evidentiary infrastructure of the

discipline. We argue for a critical digital humanities practice built around provenance, uncertainty signals that remain meaningful at the topological level, and the recognition that machine-extracted cultural networks, whether bibliographic, urban, or visual, share the same systemic fragility.

Cited works

Azaria, A., and Mitchell, T. (2023). The internal state of an LLM knows when it's lying. Findings of EMNLP 2023, 967–976.

Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892–895.

Carmo, M. (2011). *The Alphabet and the Algorithm*. Cambridge, MA: The MIT Press.

Carmo, M. (2017). *The Second Digital Turn: Design Beyond Intelligence*. Cambridge, MA: The MIT Press.

Cioffi, A. (2022). Data for Testing and Evaluating References Extraction and Parsing Tools. Zenodo. <https://doi.org/10.5281/zenodo.6182066>.

Colavizza, G. (2018). *The Intellectual Organisation of History*. Doctoral dissertation, EPFL. <https://doi.org/10.5075/epfl-thesis-8537>.

Colavizza, G., and Romanello, M. (2017). Annotated references in the historiography on Venice: 19th–21st centuries. *Journal of Open Humanities Data*, 3. <https://doi.org/10.5334/johd.9>.

Daston, L., and Galison, P. (2007). *Objectivity*. New York: Zone Books.

Gemma Team (2026). *Gemma 4 Technical Report / Model Card*. Google DeepMind.

Gingras, Y. (2010). Mapping the structure of the intellectual field using citation and co-citation analysis of correspondences. *History of European Ideas*, 36(3), 330–339.

Hosseini, A., Ghavimi, B., Boukhers, Z., and Mayr, P. (2019). EXCITE – A toolchain to extract, match and publish open literature references. In 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL), 432–433. <https://doi.org/10.1109/JCDL.2019.00105>.

Huang, L., Yu, W., Ma, W., et al. (2023). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*. arXiv:2311.05232.

Lin, S., Hilton, J., and Evans, O. (2022). Teaching models to express their uncertainty in words. *Transactions on Machine Learning Research*.

- Luthra, M., Todorov, K., Jeurgens, C., and Colavizza, G. (2024). Unsilencing colonial archives via automated entity recognition. *Journal of Documentation*, 80(5), 1080–1105.
- Malinin, A., and Gales, M. (2021). Uncertainty estimation in autoregressive structured prediction. *International Conference on Learning Representations (ICLR)*.
- Paglen, T. (2014). Operational images. *e-flux journal*, 59 (November). <https://www.e-flux.com/journal/59/61130/operational-images/>.
- Paglen, T. (2019). Invisible images: Your pictures are looking at you. *Architectural Design*, 89(1), 22–27. <https://doi.org/10.1002/ad.2383>.
- Pagnotta, O. (2024). CEX Project — Dataset and Gold Standard. Zenodo. <https://doi.org/10.5281/zenodo.10535653>.
- Panofsky, E. (1955). *Meaning in the Visual Arts: Papers in and on Art History*. Garden City, NY: Doubleday.
- van Hooland, S., De Wilde, M., Verborgh, R., Steiner, T., and Van de Walle, R. (2015). Exploring entity recognition and disambiguation for cultural heritage collections. *Digital Scholarship in the Humanities*, 30(2), 262–279.

Computational Hermeneutics of Museum Discourse: Modeling Bias and Interpretive Framing in Curatorial Texts.

By Mahnoush Pourhosseini Akbarie (Sapienza University of Rome), Shiva Talaei Koochehbagh (Sapienza University of Rome) and Paola Pisano (Turin University).

Museums increasingly position themselves as inclusive institutions committed to accessibility, plurality, and public engagement. Yet exhibition texts and online collection descriptions often continue to reproduce implicit interpretive hierarchies, Eurocentric framings, euphemistic narratives, and asymmetries of cultural visibility. While critical museum studies and postcolonial heritage scholarship have extensively examined such issues qualitatively, computational approaches capable of systematically modeling interpretive bias in curatorial discourse remain underdeveloped. This paper addresses the following research question: how can computational methods contribute to the identification, formalization, and analysis of bias mechanisms embedded in museum texts without reducing interpretation to purely quantitative classification?

The study presents a small-scale but methodologically explicit computational analysis of curatorial language using a pilot corpus derived from the Metropolitan Museum of Art's Open Access online collection. The corpus consists of museum object descriptions annotated for bias presence, bias type, interpretive mechanism, severity, and explanatory rationale. Rather than conceptualizing bias as a binary attribute, the project treats curatorial discourse as a layered interpretive system in which institutional authority, historical framing, omission, valuation, and linguistic abstraction interact to shape cultural meaning. The corpus, therefore, operationalizes bias not merely as problematic terminology, but as a structured interpretive phenomenon embedded within broader regimes of representation and cultural legitimacy.

The methodological framework combines qualitative annotation with lightweight Natural Language Processing (NLP) analysis. Annotation categories include phenomena such as value hierarchization, provenance omission, euphemistic violence, civilizational framing, decontextualization, and asymmetrical attribution of agency. Each annotation is accompanied by structured explanatory reasoning intended to preserve interpretive traceability and make explicit the connection between claims, evidence, and annotation decisions. This emphasis on traceability responds directly to current concerns in computational humanities regarding transparency, reproducibility, and epistemic accountability in AI-assisted cultural analysis.

The analytical workflow proceeds in three stages. First, the annotated corpus is examined quantitatively to identify the distribution and co-occurrence of bias mechanisms across object descriptions. Second, NLP-based linguistic analysis investigates how bias manifests through specific discursive features, including evaluative adjectives, passive constructions, abstraction patterns, omission structures, and markers of institutional authority. Third, qualitative close reading is integrated with computational outputs to contextualize how linguistic framing contributes to

broader historical and cultural narratives. Rather than replacing interpretation, computational analysis functions here as a means of rendering interpretive patterns systematically visible across the corpus.

The study intentionally adopts a modest corpus scale to foreground methodological rigor and interpretive transparency over large-scale automation. This design choice aligns with emerging approaches in digital humanities that advocate computationally assisted interpretation grounded in historically situated reading practices rather than purely predictive modeling. The paper, therefore, positions computational analysis as a form of “computational hermeneutics,” where formal annotation and NLP methods support reflective examination of cultural discourse while preserving ambiguity, contextual nuance, and interpretive plurality.

Preliminary findings indicate that bias frequently emerges not through overtly exclusionary language but through subtler mechanisms of narrative framing, including selective contextualization, asymmetrical agency assignment, and implicit normalization of institutional perspectives. The results also demonstrate that structured annotation schemas can provide reproducible frameworks for modeling interpretive phenomena that are traditionally treated as resistant to formalization.

By combining computational analysis, critical heritage studies, and museum discourse analysis, this paper contributes to ongoing discussions surrounding responsibility, interpretive authority, and epistemic governance in digital cultural heritage. More broadly, it argues that small, carefully curated datasets can serve as productive methodological laboratories for developing computationally tractable yet critically grounded approaches to cultural interpretation. The project ultimately proposes a transferable framework for studying museum discourse as a complex interpretive system in which language, institutional authority, and historical representation remain deeply entangled.

Charting Contemporaneity: Pathways through Exhibitions, Museums, and Markets.

By Clarissa Ricci (University of Bologna).

How can contemporaneity be understood and studied? What allowed modern-contemporary to become a value that structured the art system in different contexts?

The first argument is that cultural complexity is best understood through diverse, interdisciplinary methodological approaches. Archives and documents often serve as starting points, offering clues to possible pathways, as Paul Ginsborg suggests. Computational approaches can then test and refine historical models and pose bolder questions about art ecosystems.

Therefore, the presentation addresses how the value of contemporary art formed in the early XX century, through the development of networked institutional infrastructures and computational analysis in art archival investigations. More specifically, the Venice Biennale's Sales Department is taken as a case study to outline an investigative pathway that considers sales and exhibition infrastructures as nodes in a system and examines their relationships. Based on a network analysis of the Venice Biennale's Sales Department between 1895 and 1942, the presentation shows the Biennale's contribution to museum collections and the development of the art market in Italy. The sales registers also reveal how this system developed and how museum acquisitions served as a core legitimising mechanism. We found, for example, that acquisitions depended on political and cultural decisions.

Furthermore, analysing nodes within the art system, especially museum collection formation, from the perspective of large-scale exhibitions enables us to test the impact of such biennial-type exhibitions on the broader art system and to clarify their contributions to art infrastructures and the affirmation of contemporary art.

How can we explain the apparent increase in acquisitions alongside a decrease in sales? Does analysing museum acquisitions suggest that museums functioned as 'covered' stakeholders within the exhibition network, and what was their contribution to the final budget?

Investigating the ties between the Venice Biennale and museums could reveal models applicable to other biennial sales datasets. The aim is to show that art's value can be studied as a living, interconnected system in which markets, museums, and biennials form dynamic networks rather than isolated events, while setting a methodological approach

that integrates different methods to clarify how institutional networks shape artistic ecosystems.

Ten years after MUSA16: Advances and Challenges in Compiling a Corpus for the Study of Translated Museum Texts.

By Jorge Leiva Rojo (University of Málaga)

The main aim of this paper is to advance the study of the translation of museum texts, a relatively unexplored field to date despite its relevance (Liao, 2018, p. 46, 2025, p. 2; Neather, 2024, p. 1). For the analysis of the translation of this type of text, corpus linguistics has been relied upon, with the purpose of using the texts that make up the corpora to conduct the most systematic analysis possible—not only of the terminology, but also of other subjects, such as identifying which texts are translated most frequently or which diatopic variety of Spanish is most commonly used.

As has been mentioned, this study makes use of corpus linguistics, a field that deals primarily with text corpora, which can be defined as “a large collection of authentic texts that have been gathered in electronic form according to a specific set of criteria” (Bowker & Pearson, 2002, p. 9). Two text corpora will be presented. The first, called MUSA22 (a name derived from “museums,” “USA,” and the final year of compilation, 2022), is a parallel corpus, that is, a corpus composed of “[s]ource texts in one language plus translations into one or more other languages” (McEnery & Hardie, 2012, p. 19), which has been aligned at the sentence level and consists exclusively of museum texts from institutions in the United States—originally written in English—and their corresponding translations into Spanish. This corpus has undergone three previous versions prior to the current one, which will be briefly discussed below. The first version of the corpus, named MUSA16, consisted of 151 bitexts—that is, English originals and their corresponding Spanish translations—sourced from the websites of 21 museums, all located in New York City. The total volume was just over 380,000 words in the English subcorpus and 429,000 words in the Spanish subcorpus.

A later version of the corpus, MUSA19, involved a partial search for additional texts in museums that were already part of the corpus; the increase in the number of texts or words in the subcorpora was therefore minimal (21 bitexts and just over 30,000 words in each subcorpus). The following year’s version, MUSA20, consisted of a complete update of the entire subcorpus, which made it possible to add all the new texts (plus some older ones not previously located) that the museums had translated into Spanish between 2016 and 2020 and published on their websites. This resulted in the inclusion of translations

from five additional museums in New York City, increasing the corpus size by 40% compared to the previous version. The figures for this version of the corpus are 26 museums, 304 bitexts and just over 760,000 and 840,000 words in the English- and Spanish-language subcorpora, respectively.

The corpus truly took on a national scale with the current version, called MUSA22, as it incorporated museums from across the country. To expand the corpus, starting in May 2020, the following actions were carried out:

- 51 museums and art centers in the state of Massachusetts were surveyed, which allowed the identification of texts at 16 museums, although the number of bitexts was small—only 81.
- The websites of the 20 most-visited museums in the United States in 2018 (TEA & AECOM, 2019, p. 73) were reviewed, which significantly increased the number of texts in the corpus, given that, to a greater or lesser extent, all of these museums had material translated into Spanish on their websites.
- Messages were sent to Museum Junction, the forum of the American Alliance of Museums, which made it possible to establish contact with 18 additional museums that offered Spanish translations on their websites. 43 of the museums already included in the corpus were contacted directly to request their collaboration in providing translated material in addition to what was already available on their websites.
- In July 2021, the websites of all the museums in the corpus were checked again to identify new bitexts and to ensure that the last date of text retrieval from the web was the same for all of them.
- Finally, over the course of nearly a year (July 2021–June 2022), corpus management tasks were carried out: file management, alignment of the new records, and complete compilation of the corpus.

The complete data for MUSA22, the current version of the parallel and aligned corpus, show that it has already reached a considerable size, as it comprises 78 museums from 19 U.S. states and the nation's capital, Washington, D.C., with a total of 1,375 bitexts, which amounts to 2.7 million words in the English subcorpus and 3.1 million in the Spanish subcorpus.

Next, a comparable corpus was compiled. This type of corpus can be defined as that which consists of “original texts in one, two, or more languages with similar content, e.g., on the same subject matter” (Laursen & Arinas Pellón, 2012, p. 48). This corpus, named MUSA_COMP, comprises texts originally written in Spanish—it is, therefore, monolingual—and originating from six of the most representative diatopic varieties of the Spanish language: Mexico, Colombia, Spain, Argentina, Costa Rica, and Peru. Regarding

the volume of words in the comparable corpus, a volume of approximately three million words was established for each of the subcorpora, with the aim of ensuring that each subcorpus was balanced (Hardie, 2015, pp. 505–506). As a result, this corpus comprises a total of 2,863 files from 270 museums, art centers, and institutions, amounting to a total of 18.2 million words. In light of the compilation results, it is reasonable to hypothesize that the corpus size is adequate to serve as a validation tool for translations from English to Spanish in the MUSA22 corpus.

With regard to the main difficulties encountered in the compilation of the first corpus, MUSA22, the following are noted: first, locating the texts, since not everything produced by U.S. museums is translated into Spanish, which in many cases means having to start with the translated text and then track down the original text. As a result, this parallel corpus is much smaller than a comparable one if the same number of hours were invested to compile it. Second, graphic standardization was a time-consuming task, given that 44% of the bitexts were PDF files, which required effort to clean up the texts prior to alignment. Third, the alignment process—which is simple but tedious—was complicated by the fact that the software used, ABBYY Aligner 2.0 (ABBYY Software, 2011), ceased to offer support in June 2023, which complicates any future expansion of this corpus. Finally, the fact that, when compiling the corpus, Sketch Engine only allows the upload of a single file (in this case, a TMX file) results in the deletion of all repetitions of matching sentences in the corpus and, furthermore, does not allow for the identification of the bitext from which a given occurrence in the corpus originates.

In the case of the comparable corpus—a task that was initially considered secondary—it became apparent during the process that compiling it would be more difficult than expected. The main obstacles identified were that attempts at automatic text extraction were unsuccessful, so this process had to be carried out almost entirely by hand. Similarly, there were quite a few instances of duplicate sentences in the corpus, so it was necessary to use software to search for and remove them from the corpus. In any case, the greatest difficulty in the compilation process was undoubtedly locating the texts, with situations varying widely: while in countries such as Spain the volume of texts published by museums online is very high, in other countries—as was the case with Peru—the online presence of museums (and, consequently, the publication of texts) is much more limited.

Cited works

ABBYY Software. (2011). ABBYY Aligner 2.0 (1.0.6.59). ABBYY Software.

Bowker, L., & Pearson, J. (2002). Working with specialized language: A practical guide to using corpora (Vol. 24, Issue 4). Routledge. <https://doi.org/10.1093/applin/24.4.553>

- Hardie, A. (2015). Corpus linguistics. In K. Allan (Ed.), *The Routledge handbook of Linguistics* (pp. 502–515). Routledge.
- Laursen, A. L., & Arinas Pellón, I. (2012). Text corpora in translator training. *The Interpreter and Translator Trainer*, 6(1), 45–70. <https://doi.org/10.1080/13556509.2012.10798829>
- Liao, M.-H. (2018). Translating multimodal texts in space: A case study of St Mungo Museum of Religious Life and Art. *Linguistica Antverpiensia*, 17, 84–98.
- Liao, M.-H. (2025). *Translation on display: Multilingual texts in multimodal museum space*. Routledge. <https://doi.org/10.4324/9781003545286>
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- Neather, R. (2024). *Translating for museums, galleries and heritage sites*. Routledge. <https://doi.org/10.4324/9780203702505>
- TEA, & AECOM. (2019). 2018 Theme Index and Museum Index: The Global Attractions Attendance Report. <http://www.teaconnect.org/pdf/TEAAECOM2013.pdf>

Exhibition *Mise en Place*. Gender and Transnationality as Key Variables for Analyzing the Art Ecosystem.

By Bárbara Romero Ferrón (Leuphana University).

This paper proposes moving beyond the basic statistical analyses typically used to address transnationality and gender in the exhibition ecosystem, putting forward instead a broader analytical toolkit, including but not limited to graph theory and network analysis, capable of capturing structural dynamics that simple counts cannot reveal. In this presentation, we introduce a set of complementary analyses, applied to the same corpus, each designed to surface a different layer of complexity surrounding transnationality and gender, moving from descriptive statistics toward approaches that can account for circulation, position, persistence, and relation within the system.

The corpus analyzed was developed across a sequence of successive research projects (Exhibitium 2015 to 2016, Andalex 2017 to 2021, Complexhibit 2021 to 2026). It comprises 12,536 exhibitions held in Andalusia between 2010 and 2023, in both provincial capitals and peripheral localities, and registers 14,002 unique artists.

In contrast to the percentual counts that dominate previous studies on female presence or internationalization, this work draws on network analysis alongside other complementary methods to examine not only how many foreign agents or women participate, but how they circulate, who they connect with, and what structural position they occupy over time. We refer to this broader set of tools, which goes beyond a count or a piece of descriptive statistics toward analyses able to extract the underlying complexity of the two variables under study, as the *mise en place* of the analysis itself, the basic apparatus that must be put in place before the structural logics of the system can become visible. The transnationality analysis reveals a system that is apparently open, with peripheral micro platforms channeling underrepresented international circuits, yet crossed by internal symbolic borders. The gender analysis shows that, although women represent more than a third of all artists, their structural centrality is lower and their connections skew mostly toward men, reproducing a center periphery logic even within the female subnetwork.

Taken together, these analyses show that openness and equity cannot be properly assessed through percentages or simple aggregate figures alone. Only by setting up this more complex analytical *mise en place*, one capable of attending to circulation, position and connection rather than mere presence, can the structural logics that organize

transnationality and gender within the exhibition ecosystem be properly identified, pointing toward the need to redistribute relational capital across the system rather than simply rebalancing its numbers.

"Dense regions of complexity": Intrinsic Dimension and the Neural Geometry of Art-Critical Knowledge in Large Language Models.

By Ángel M^a Lumbreras (University of Málaga).

This paper starts from a theoretical and methodological hypothesis: cultural complexity should not be understood merely as an accumulation of data, textual difficulty, or documentary richness, but as a relational structure made up of proximities, distances, condensations, displacements, and dense regions of meaning. In the case of an art-historical corpus, concepts do not operate as isolated units, but as positions within a historically situated constellation. From this perspective, terms such as *disegno*, *grazia*, or *bellezza* do not possess only lexical value; rather, they acquire meaning through the relations they establish with one another. The aim of this research is to analyse what happens to this relational complexity when an artistically situated corpus is processed through the internal layers of a large language model.

The paper is situated within Nuria Rodríguez-Ortega's reflections on techno-concepts, understood as co-productions between technoscientific or computational rationality and human thought/imagination (Rodríguez-Ortega, 2022). Within this framework, *n*-dimensional, vectorial, and latent spaces are not merely technical instruments, but concepts that transform the epistemological conditions from which culture is interpreted. The notion of vector space makes it possible to think of cultural objects as numerical and spatial configurations in which complexity appears simultaneously as a cultural phenomenon —historical, semantic, and relational density— and as a mathematical structure —dimensionality, distribution, density, and spatial organisation. For this reason, distance, density, form, and dimension are not innocent metaphors, but operations that reorder what can be seen, compared, and interpreted.

This perspective makes it possible to address some of the problems involved in the incorporation and retrieval of art-historical concepts within LLMs. More specifically, in retrieval-augmented generation systems (RAG), the presence of documentary fragments does not guarantee that the generated claims are actually grounded in them (Lewis et al., 2020), since a source may be thematically relevant without functioning as effective evidence. Similarly, in fine-tuned models, a corpus may be absorbed as style or vocabulary without ensuring the preservation of its deeper conceptual relations. Moreover, fine-tuning with new knowledge can generate tensions between factual learning, parametric memory, and hallucination (Gekhman et al., 2024). In both cases, the

problem is not only whether the model cites, retrieves, or imitates the corpus, but how it reorganises it.

The paper takes as its case study Giorgio Vasari's *Torrentiniana*, published in 1550, and analyses an operative selection of ten artistic concepts: *disegno*, *invenzione*, *imitazione*, *maniera*, *grazia*, *naturale*, *bellezza*, *proporzione*, *patria*, and *fama*. The experiment does not seek to demonstrate that an LLM—in this case, Llama—“understands” Vasari, nor that it can faithfully reproduce a Vasarian voice. On the contrary, it begins from a distinction between semantic-distributional understanding, functional understanding, and phenomenological-historical understanding. Language models can produce functionally competent responses and organise semantic relations on the basis of distributional patterns, but this does not imply a form of historical or phenomenological understanding equivalent to that of humans (Bender & Koller, 2020; Harnad, 1990). Whereas human categories are formed through cultural temporalities, social practices, discursive traditions, bodily memories, and historiographical institutions, machinic representations are organised as positions within activation spaces optimised for linguistic prediction.

On the basis of this difference, the central question is not whether Llama understands Vasari, but what form a Vasarian conceptual constellation takes within Llama. To answer this question, the experiment unfolds in two phases. First, a reference geometry is constructed from the Vasarian corpus itself, independently of the model being analysed. This external geometry makes it possible to establish a distance matrix and an ordinal consensus matrix among the selected concepts, avoiding a comparison that would depend on the very same representations later being evaluated. Secondly, the Vasarian contexts are introduced into Llama-3.2-3B-Instruct through a forward pass, extracting the model's internal representations layer by layer. The analysis does not focus on textual generation, but on the internal transformation of the conceptual constellation.

Methodologically, the study employs metrics designed to compare relational forms rather than isolated meanings: Spearman correlation, Pearson correlation, neighbourhood overlap, pairwise ordinal deformation, Procrustes analysis, local density, and intrinsic dimension. The choice of ordinal metrics is especially important because the model's internal distances and the corpus-derived distances do not belong to the same scale. The analysis therefore focuses on the order of relations: which pairs remain close, which ones contract or expand, and in which layer the Vasarian structure can be most clearly recognised.

The results show that the Vasarian space is not destroyed by the model, but neither is it preserved intact. In certain layers, a partial, non-random, and selectively deformed

preservation of some conceptual relations emerges. The nineteenth layer of the neural network constitutes the point of greatest ordinal preservation of the corpus structure, whereas other layers display different properties, such as greater dimensional complexity, lower global deformation, stronger conceptual separation, or late-stage reorganisation. This plurality prevents us from speaking of a single “correct” layer and instead requires us to understand the model as a trajectory of representational transformations.

The most significant result is the identification of a dense region of complexity formed by *imitazione, proporzione, bellezza, and grazia*. In the geometry of the corpus, these concepts do not merely constitute a series of neighbouring pairs, but a higher-order configuration, a kind of irreducible conceptual surface. Its interest is not only mathematical, since it points to a recognisable area of Vasari’s critical vocabulary and allows us to observe how an art-historical relation can take on geometric form. Within Llama, this region does not disappear, but it is deformed. These transformations indicate an internal reorganisation of conceptual relations within the model.

The final hypothesis is that the analysis of internal geometries can offer an intermediate criterion between documentary traceability and stylistic imitation. It does not, by itself, solve the problems of truthfulness, semantic drift, false authorial voice, or lack of grounding, but it makes it possible to diagnose them at another level: that of relational fidelity. Knowing the form a corpus takes within a model allows us to ask whether generation preserves, contracts, expands, or destroys the conceptual relations that historically structure that corpus. In this way, internal geometry becomes a critical tool for studying the difference between human understanding, functional competence, and the machinic representation of artistic knowledge.

Cited works

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. En Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 5185–5198). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.463>

Engels, J., Michaud, E. J., Liao, I., Gurnee, W., & Tegmark, M. (2025). Not all language model features are one-dimensionally linear. International Conference on Learning Representations.

Fierro, C., Dhar, R., Stamatiou, F., Garneau, N., & Søgaard, A. (2024). Defining knowledge: Bridging epistemology and large language models. En Proceedings of the

2024 Conference on Empirical Methods in Natural Language Processing (pp. 16096–16111). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/2024.emnlp-main.900>

Gekhman, Z., Yona, G., Aharoni, R., Eyal, M., Feder, A., Reichart, R., & Herzig, J. (2024). Does fine-tuning LLMs on new knowledge encourage hallucinations? En Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (pp. 7765–7784). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/2024.emnlp-main.444>

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1–3), 335–346. [https://doi.org/10.1016/0167-2789\(90\)90087-6](https://doi.org/10.1016/0167-2789(90)90087-6)

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474.

Rodríguez-Ortega, N. (2020). Inteligencia artificial y campo del arte. *pArAdigmA. Revista Universitaria de Cultura*, 23, 32–51.

Rodríguez-Ortega, N. (2022). Techno-concepts for the cultural field: n-dimensional space and its conceptual constellation. *Multimodal Technologies and Interaction*, 6(11), 96.
<https://doi.org/10.3390/mti6110096>