Thymele : A Collaborative Digital Lexicon for Performing Arts in the Digital Age

Eleni PAPALEXIOU(1), Avra XEPAPADAKOU(2),
Costas VASSILAKIS(3), Vassilis POULOPOULOS(4)

(1) University of the Peloponnese, (2) University of Nicosia,
(3) University of the Peloponnese, (4) University of the Peloponnese

DOI : https ://doi.org/10.17184/eac.10176

Abstract : The *Thymele* platform is a pioneering digital tool designed to support the creation and management of a comprehensive lexicon for digital performing arts. It leverages collaborative methodologies, advanced technological tools, and innovative frameworks to enhance research and documentation. By integrating multimedia content, interconnectivity among terms and AI/ML for ontology recognition, Thymele addresses the complexities of performing arts research in the digital era. This paper explores the theoretical foundations, technological implementations, and the broader impact of Thymele on digital humanities. It aims to highlight how the platform contributes to advancing research methodologies, ensuring semantic consistency, and fostering collaborative knowledge construction in the field of performing arts.

Keywords : Performing Arts, Digital Humanities, Digital Lexicography, Onto-terminology, Linked Open Data, Multimodality

1 Introduction

1.1 Contextualizing digital transformation in performing arts research

Over the past three decades, digital media have redrawn the creative and analytical landscape of the performing arts. Networked stages, real-time motion capture, and algorithmic composition blur traditional boundaries between theatre, music, and dance. Born-digital artifacts – live-streamed rehearsals, VR scenography, social-media dramaturgies – now constitute primary research material. The COVID-19 pandemic accelerated this shift, forcing artists and institutions to adopt online formats almost overnight (Quillet 2022). As a result, scholars confront heterogeneous datasets that combine text, image, sound, and code; established philological tools no longer suffice. Digital humanities (DH) methods – corpus analysis, linked data modeling, machine

learning – have therefore become integral to performing arts documentation and analysis (Bardiot 2021, 31–40; Papalexiou 2020, 376–423), yet their full potential is limited without a stable, domain-specific vocabulary and interpretation.

1.2 The need for structured dictionaries in digital humanities

DH scholarship depends on interoperable metadata: concepts must carry the same meaning across archives, databases, and analytical tools. In the performing arts this requirement is acute, because works are multimodal, collaborative, and ephemerally documented (Papalexiou and Xepapadakou 2022, 147–194; Doerr, Bekiari, and Le Bœuf 2008, 7–8). Existing Greek glossaries are dated, art-form specific, or locked in print; international resources rarely adapt terms to Greek morphology or cultural context (Christofidou 2017). The absence of a standardized lexicon hampers:

- precise annotation of archival records;
- aggregation of performance data across institutions;
- consistent teaching materials in university and conservatory programs;
- public access to Greek cultural heritage via multilingual search portals.

A structured, updatable dictionary that embeds each term in an explicit ontology is therefore essential for credible DH workflows in the field.

1.3 Purpose and scope of the *Thymele* project

Thymele addresses this gap by designing an open-access multimedia dictionary for the performing arts. Named after the altar at the center of the ancient Greek orchestra (Chantraine 1944, 449), the project aspires to become a focal point for contemporary performance studies. Funded by the Hellenic Foundation for Research and Innovation and coordinated by the University of the Peloponnese, a consortium of twelve institutions will publish approximately 600 peer-reviewed headwords in Greek, each mapped to English, French, German, and Italian equivalents.[1] Every entry will provide:

- a preferred Greek term grounded in clear word-formation principles;
- a concise definition and brief encyclopedic note;
- corpus-based usage examples;
- audio-visual documentation;
- semantic links expressed through a CIDOC-inspired ontology.

Technically, *Thymele* integrates a relational database for lexicographic data, a triple store for ontological assertions, and an asset repository for multimedia files. Optical

[1] The research project *Thymele, A Dynamic Open-access Multimedia Dictionary for the Performing Arts* is implemented in the framework of Hellenic Foundation for Research and Innovation call "Basic Research Financing (Horizontal Support of All Sciences)" under the National Recovery and Resilience Plan "Greece 2.0" funded by the European Union – NextGenerationEU (H.F.R.I.).
Duration: 12 March 2024 – 31 December 2025.
Host Institution: University of the Peloponnese, Department of Theatre studies.
Collaborating Institutions:
Academy of Athens, Research Center for Scientific Terms and Neologisms: Research Center for Modern Greek Dialects;
National Center for Scientific Research "Demokritos", Institute of Informatics and Telecommunications (*Thymele* (n.d.)).
Project Website: https://thymele-project.uop. gr/. The *Thymele* Lexicon Website: http://thymele-lexicon.gr/\\

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

Character Recognition (OCR) and deep-learning models convert legacy PDFs into machine-readable text, feeding both the term-selection corpus and citation manager (Kaddas *et al.* 2023). Generative language models suggest additional links, which human editors validate before publication (Doumanas *et al.* 2025). All content passes through a workflow of authoring, editing, and curation overseen by thematic work groups.

By anchoring Greek terminology within an interoperable, multilingual infrastructure, *Thymele* seeks to stabilize scholarly discourse, enhance archival interoperability, and strengthen the visibility of Greek research in the European DH arena (Xepapadakou, Papalexiou, and Manolessou 2025).

2 Theoretical Framework

2.1 Conceptual basis for dictionary creation in the performing arts

The performing arts constitute a research field where meaning is embodied as much as it is verbalized. A choreographic term, for instance, is inseparable from the kinesthetic pattern it names; a stage device acquires significance through its visual and acoustic presence. Any lexicographic project that seeks to serve the discipline must therefore integrate three conceptual strands. First, it has to be multimodal: definitions that remain exclusively verbal cannot convey the gestural, sonic, or spatial qualities that give many concepts their force (Lew 2010). Digital delivery solves this limitation by allowing images, sound clips, and video extracts to accompany text. Second, it must rest on ontoterminology, an approach that treats terms and the conceptual relations between them as two sides of the same coin (Roche 2012, 2627). By modeling hypernyms, hyponyms, roles, and processes in an explicit ontology, the dictionary provides a machine-readable scaffold for archival metadata and linked data services. Third, the work must be empirically grounded. A purpose-built corpus of scholarly articles, reviews, and program notes supplies frequency counts and authentic collocations, ensuring that each headword reflects actual usage rather than idiosyncratic preference (Gatto 2014). Bringing these strands together allows the dictionary to document current practice while steering it toward terminological consistency.

2.2 The need for a common dictionary

Greek discourse on theatre, dance, music, performance art, and their digital hybrids is fragmented along institutional and historical lines. Separate scholarly traditions have imported English terms independently, often translating them on the fly and without reference to Greek word-formation norms. The result is polysemy – performance may render παράσταση, επιτέλεση, or εκτέλεση, or even περφόρμανς [performance] – and synonymy, as τοποσυναφής, τοποαναφορικός, σάιτ-σπεσίφικ, τοποειδικός, ιδιοτοπικός, and τοποειδής [site specific] circulate simultaneously. Such variability hampers cataloging in archives, complicates data aggregation across repositories, and confuses students who meet different labels in adjacent modules. International resources offer rich content yet rarely adapt to Greek morphology or cultural context, while most local glossaries date from the 1990s and exist only in print (Papalexiou *et al.* 2025). A dictionary that assembles, verifies, and standardizes terminology

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

across all performing-arts subfields therefore becomes a prerequisite for credible DH workflows, interoperable metadata, and coherent higher-education curricula.

2.3 Collaborative knowledge implementation

No single scholar commands expertise across the full breadth of contemporary performance practice; hence *Thymele* adopts a distributed, peer-validated model of knowledge production. More than eighty specialists, drawn from twelve institutions, operate within thematic work groups that mirror the ontology's top-level classes (e.g. theatre, dance, music, multimedia). Each lemma passes through three successive roles: authoring, where subject experts draft the entry and propose multimedia illustrations; editing, during which disciplinary peers refine wording, verify sources, and align the text with lexicographic style; and curation, where an interdisciplinary team harmonizes definitions with the overarching ontology and resolves cross-domain inconsistencies. All actions are logged in a version-control layer so that the evolution of every entry remains transparent and reversible (Humbley 2018, 320). This workflow emulates journal peer-review standards while preserving the agility required for continuous online publication.

2.4 Need for a "crowdsourcing-based" digital dictionary

While expert review guarantees scholarly rigor, long-term sustainability demands wider participation. *Thymele* therefore embeds crowdsourcing mechanisms that allow practitioners, students, and informed members of the public to contribute without diluting quality. Registered users can suggest new terms, upload illustrative material, or flag potential errors through a moderated interface. Suggestions enter a quarantine queue, visible only to editors, where automated alerts assign them to the appropriate thematic group. Reputation points accrue to contributors whose input is consistently accepted, creating an incentive structure that rewards careful work. In parallel, post-publication annotation lets readers enrich an entry with additional bibliography or contextual notes, again subject to editorial approval (Ooi 2018, 690). By coupling open contribution channels with tiered permissions and audit trails, *Thymele* harnesses collective intelligence while maintaining the reliability required of a scholarly reference tool.

3 Platform Overview

The *Thymele* platform operates on a robust cloud-based infrastructure designed to support multiple users simultaneously. This architecture enables seamless collaboration among researchers, artists, and scholars, allowing them to contribute to and refine lemmas in real time. The multi-user environment is structured to provide different levels of access and interaction, ensuring that roles such as editors, contributors, and reviewers are dynamically assigned to maintain a balanced and coherent lexicon development process.

Collaboration is at the heart of *Thymele*, with a lemma implementation process that emphasizes collective input and iterative refinement. Each lemma undergoes a thorough review cycle where users can suggest edits, add multimedia content, and discuss

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

conceptual nuances. This collaborative methodology not only enriches the quality of the content but also ensures that the lexicon continues to reflect the diverse perspectives inherent in the performing arts.

In addition, the platform is designed to accommodate a wide range of multimedia formats, including video, images, and audio files, to provide a richer context for each lemma. By integrating these elements, *Thymele* enhances the research experience, allowing users to explore and understand performing arts concepts in a more holistic and immersive way.

3.1 Technological implementation

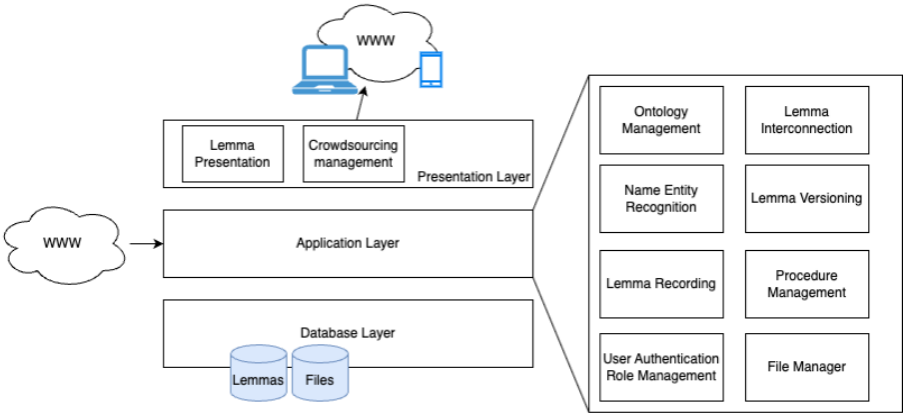


Figure 1: System Architecture

The back end of *Thymele* follows a three-tier model – data layer, application layer, and presentation layer – deployed on a virtualized server hosted by the University of the Peloponnese.

3.1.1 Data layer

All lexicographic micro-data – definitions, usage examples, bibliography – are serialized in a relational schema. As the system employs a large number of users having multiple and different roles, each action of the users is recorded. The data layer records connections to files and external information that are recorded as part of the lemmas.

3.1.2 Version control and provenance

Every edit is stored as a delta, similar to a Git commit, recording user, timestamp, and diff. The UI exposes an entry history pane where authorized users can compare any two revisions and roll back to a previous state.

3.1.3 Role management and dynamic grouping

Access control is enforced by a custom authentication and authorization system built on the Laravel Breeze. Roles are usually not static but they can change across the

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

different stage of a lemma creating a dynamic environment for the users. As this can occur during the lifetime of a lemma the system records each access right of the users to be able to serve only information that is related to their access history to the platform. This flexibility keeps the workflow agile without compromising accountability.

3.1.4 Scalability and maintenance

Kubernetes horizontal pod autoscaling spins up extra replicas of stateless services when CPU or memory thresholds exceed 70%. Stateful components – PostgreSQL, Blazegraph, MinIO – run in active-replica configurations with nightly offsite backups. Continuous Integration/Continuous Deployment (CI/CD) pipelines execute unit tests, ontology consistency checks, and license-compliance scans before any new container reaches production.

3.2 Process of lemma generation

The life cycle of a *Thymele* lemma follows a staged workflow that combines distributed authorship with layered quality control. Each stage is mapped to specific user roles, and every transition is logged by the version-control engine described in 3.1. Figure 2 visualizes the entire process; the text that follows explains the human and automated actions behind each step.

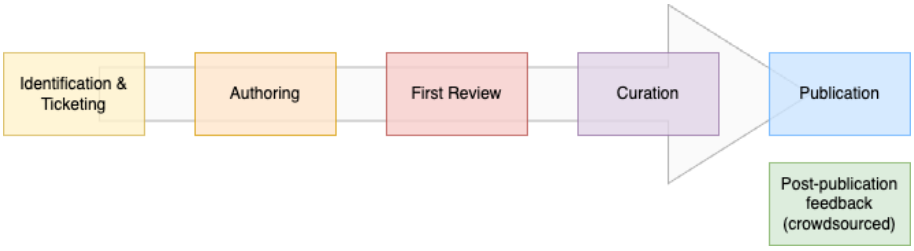


Figure 2: Lemma Generation Process

3.2.1 Workflow stages

- 1. Identification and ticketing

  A prospective term enters the system in one of two ways:

  1.1. automated term-extraction flags a frequent collocation in the corpus (Macken, Lefever, and Hoste 2013);

  1.2. a registered contributor submits a proposal *via* the Suggest term form.

  In both cases the system creates a ticket that is routed to the appropriate Thematic Work Group (TWG).

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

- 2. Authoring (status: Draft)

  Two or three authors within the TWG claim the ticket and complete the structured entry template: headword, grammatical information, definition, brief encyclopedic note, translation equivalents, corpus citations, initial ontology links, and at least one multimedia asset. The platform performs automated checks for field completeness, duplicate headwords, and license compatibility of uploaded media.

- 3. First blind review (status: In Review I)

  The entry passes to a pair of editors – senior scholars from the same or a cognate TWG – who:

  3.1. verify factual accuracy and bibliographic references;

  3.2. refine wording for clarity and adherence to lexicographic style;

  3.3. validate preliminary ontology relations;

  3.4. accept, modify, or reject AI-suggested links.

  Anonymous inline comments are sent back to authors; the cycle may repeat until all issues are resolved.

- 4. Curation (status: In Review II)

  An interdisciplinary curator team checks cross-domain coherence:

  4.1. harmonizes terminology with parallel entries;

  4.2. ensures consistency of multilingual equivalents;

  4.3. assigns or confirms a persistent identifier;

  4.4. approves the final set of ontology triples.

  Upon curator approval the entry advances to pre-publication.

- 5. Publication (status: Published)

  The lemma becomes publicly visible on the portal and is exported to the RDF triple store. Persistent URIs enable citation (e.g. thymele.org/lemma/0123). The system generates JSON-LD and IIIF manifests so that external platforms can ingest the data.

- 6. Post-publication feedback (status: Live Updates)

  Registered users – practitioners, students, general readers – may annotate the entry, suggest further bibliography, or upload additional media. All submissions enter a moderation queue handled by editors; accepted changes create a new version node in the history graph. Thus, each lemma remains a living document while preserving a full audit trail.

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

3.2.2 User roles and permissions

Thymele distinguishes five escalating levels of access, implemented through Keycloak's role-based tokens:

- Contributor – may propose new terms or comment on published entries; cannot alter the authoritative text.
- Author – drafts and revises lemmas within an assigned TWG; may upload media and propose ontology links.
- Editor – reviews and edits drafts from the same or related TWGs; approves AI suggestions; returns entries for revision.
- Curator – oversees inter-domain consistency, assigns persistent identifiers, validates multilingual mappings, and authorizes publication.
- Moderator – administrates user accounts, resolves disputes, and can roll back any version in exceptional cases.

Role assignments are dynamic. A senior student might hold author privileges in one TWG while acting as contributor in another; an external specialist can receive temporary editor status limited to a single lemma. All role changes are time-stamped and reversible.

3.2.3 Provenance and citation

At each transition the system commits a delta: user, timestamp, changed fields, and an optional comment. Authors and external researchers can therefore cite a stable revision (e.g. v 7) or reference the "living" entry, satisfying the FAIR principles of traceability and reproducibility.

The structured, role-aware workflow described above ensures that every lemma benefits from collective expertise without sacrificing responsibility, and that the dictionary can expand continuously while maintaining scholarly reliability.

3.3 Interconnectivity and categorization

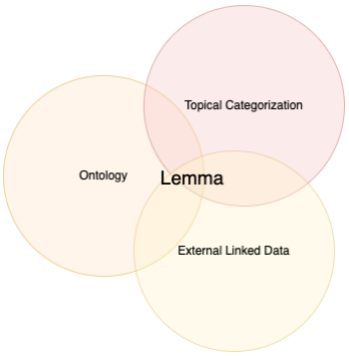


Figure 3: Lemma Interconnectivity

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

Interconnectivity lies at the core of *Thymele's* knowledge model. Each lemma inhabits three nested Webs of relation – internal ontology, topical categorization, and external linked data – so that users can move seamlessly from a single definition to broader conceptual or archival contexts.

3.3.1 Ontological links inside the dictionary

Every entry is expressed as an RDF entity that inherits properties from a CIDOC-inspired ontology for the performing arts. Four basic predicates structure the internal network:

- broader Term (hypernym) – e.g. *Kabuki* *Asian theatre*
- narrower Term (hyponym) – e.g. *immersion* *immersive theatre*
- related Term (associative) – e.g. *dramaturgy* *dramaturg*
- role In (participant/event relation) – e.g. *stage director* *production*

The public portal visualizes these links as an expandable graph.

3.3.2 Hierarchical and faceted categorization

To aid non-specialist navigation, every lemma is placed in a lightweight hierarchical tree of seven disciplinary domains:

- Stage & Production – stagecraft, lighting, sound, costume, theatre architecture, stage machinery etc.
- Theatre – dramatic genres, acting techniques, directing methods, dramaturgy.
- Music & Music-Theatre – opera, musical theatre, concert performance, singing.
- Dance – choreographic forms, movement vocabularies, somatic practices.
- Hybrid Forms – visual, digital, multimedia, and other cross-disciplinary instantiations.
- Applied Forms – community theatre, educational drama, therapeutic, and corporate applications.
- Paratheatrical Forms / Popular & Youth Culture – street performance, cosplay, live-action role-play, nightclubbing, fan practices.

In the next development phase, we plan for each lemma to be enriched with three transversal facets that will enable lateral browsing:

- Function (e.g., profession, technique, aesthetic movement);
- Medium (text, sound, image, movement, digital environment);
- Period/Region (Ancient, Early Modern, Contemporary, etc.).

Coupling the hierarchical tree with these faceted tags will allow users to drill down into specific sub-fields, such as Dance → Movement vocabularies, while also retrieving cross-cutting lists, for example all lemmas tagged "digital environment", across the entire dictionary.

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

3.3.3 Automated detection of outbound links

A background crawler scans each lemma's definition and bibliography, applies Named Entity Recognition (NER), and matches the detected entities to extract named entities and link them to lemmas. In parallel it performs a first-level check against databases like Wikipedia or DBpedia in order to obtain external links that have details about each of the entities extracted.

3.3.4 Benefits for research and discovery

- Contextual depth – scholars can trace a concept's genealogy across art forms and periods without leaving the platform.
- Semantic search – query expansion *via* ontology relations retrieves material that keyword search alone would miss.
- Interoperability – external links allow archives and digital exhibits to incorporate *Thymele* URIs as controlled vocabulary terms, aligning Greek resources with broader European infrastructures.

Through this triple-layer architecture – ontology, categorization, and Linked Open Data (LOD) – *Thymele* transforms a static dictionary into a navigable, extensible Web of knowledge that mirrors the inherently networked character of the performing arts.

3.4 Multi-level collaboration

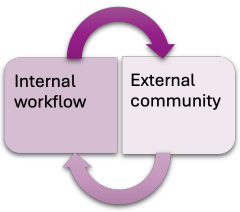


Figure 4: Multi-Level Collaboration

Thymele operates on the premise that reliable terminology emerges when expert curation and open participation reinforce one another. The platform therefore supports two nested collaboration layers – an internal scholarly layer that delivers authoritative content and an external community layer that keeps the resource responsive to evolving practice.

3.4.1 Internal collaboration

The internal layer involves five interlocking bodies, each with clearly defined responsibilities and communication channels.

- The Core Scientific Team (CST) sets editorial policy, maintains the ontology, and arbitrates disputed definitions.

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

- The Core group of lexicographers – aligned with the major sub-fields of the discipline (theatre studies, dance studies, musicology, multimedia arts, etc.), this team drafts and revises the bulk of the entries. Its members are drawn from a broad research community and were selected for their specialized expertise so that the vast range of the performing arts is covered as fully as possible. Regular online meetings (roughly once a month) allow the lexicographers to present newly completed lemmas, surface conceptual overlaps, and coordinate next steps with the Task Force (TF) and the CST.
- The Lexicographic–Ontological Team (LOT), hosted by the Academy of Athens, reviews linguistic accuracy, applies Greek word-formation rules, and verifies multilingual equivalents. Its members provide asynchronous feedback inside the platform's inline comment system and meet quarterly with CST and TF leads to update style guidelines.
- The Technological Development Team (TDT) maintains the infrastructure but also participates in content discussions whenever ontology extensions are required. Continuous-integration dashboards broadcast build status and ontology-consistency checks to all editors, allowing nontechnical members to see in real time whether a change introduces conflicts.
- The Task Force (TF) – a nine-member panel that operates as the day-to-day engine of the project. TF members run the first technical and conceptual checks on every draft, mentor lemma authors, and coordinate clusters of junior researchers who specialize in particular sub-fields. TF meets weekly online, maintaining a Kanban board that tracks each lemma from ticket to publication. Its bridging role is invaluable for keeping work flowing smoothly between authors, curators, and technologists.

Version control guarantees provenance: every save creates a commit that names the authoring user and links to an issue number in the shared Kanban board. As a consequence, disagreements can be resolved by inspecting the exact wording or link that triggered a concern, and alternative proposals can coexist as branches until a decision is reached.

3.4.2 External collaboration and community engagement

Beyond the scholarly core, *Thymele* invites contributions from practitioners, students, archivists, and the wider public. Two participation modes are available:

- Guided contributions. After creating a free account, users can suggest new terms, submit multimedia examples, or point out inconsistencies. A wizard prompts them for minimum documentation (source, context, license). The submission enters the *Community Queue*, visible to the relevant TWG and tagged with a preliminary confidence score produced by an automated content checker (plagiarism scan, file-type validation, license verification).
- Post-publication annotation. Every published lemma features a public "Comment & Improve" panel. Suggestions posted here may range from a single bibliographic reference to a lengthy usage note. Editors review comments in

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

batches; accepted items create a new minor revision (e.g. v 8 → v 9) and credit the contributor on the entry's provenance tab.

To foster informed participation, the project organizes quarterly open workshops that cover basic principles of performing-arts terminology and demonstrate the online authoring interface. Workshop recordings become part of an openly licensed tutorial series, so potential contributors can self-train at any time.

3.4.3 Governance and conflict resolution

All collaborators – internal or external – are subject to a code of conduct that emphasizes respectful dialogue, source transparency, and acknowledgement of intellectual property. Disputes follow a three-step process: informal discussion inside the platform, mediation by the LOT, final arbitration by the CST. The decision record, along with the underlying arguments, is archived for future reference so that similar cases can be handled consistently.

3.4.4 Long-term sustainability

Multi-level collaboration is also a sustainability strategy. Because editors and curators can delegate targeted tasks to motivated community members, the dictionary can grow and update without proportionally increasing staff costs. Conversely, the open layer benefits from the scholarly layer's quality controls, which maintain the dictionary's reputation and motivate continued voluntary engagement.

Through this two-tier structure *Thymele* balances openness with rigor: expert teams ensure conceptual stability, while the wider community supplies fresh material and practical insight, turning the dictionary into a living resource that matures alongside the performing arts themselves.

3.5 Example lemma generation

To illustrate how the micro-structure described in 3.1–3.4 materializes inside the platform, Table 2 reproduces – in textual form – the entry for the ανοικείωση (defamiliarization), the preferred Greek equivalent for the German term *Verfremdung (estrangement).*

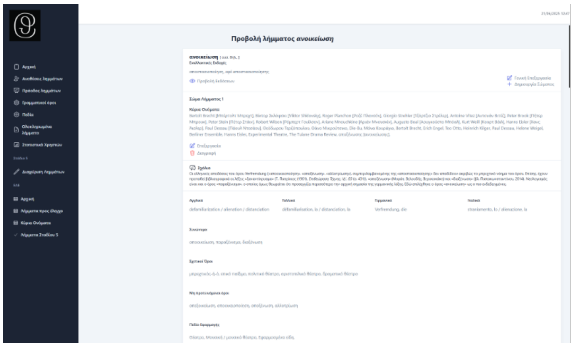


Fig. 5a

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

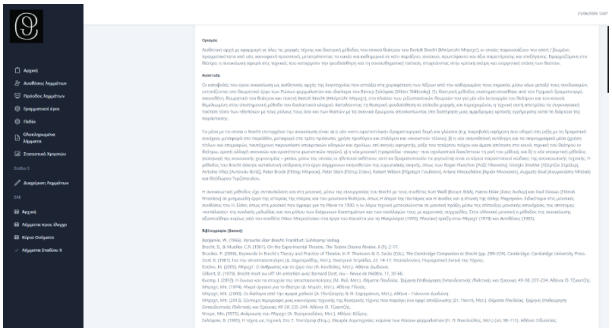


Fig. 5b

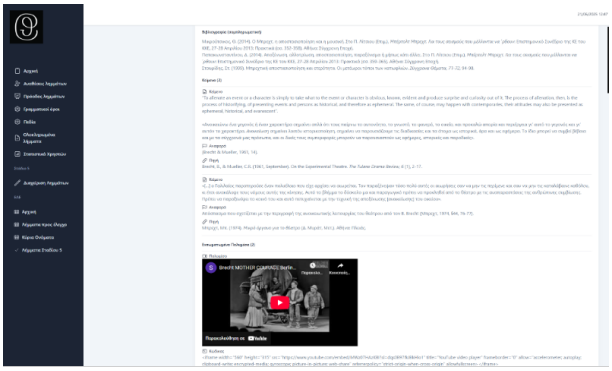


Fig. 5c

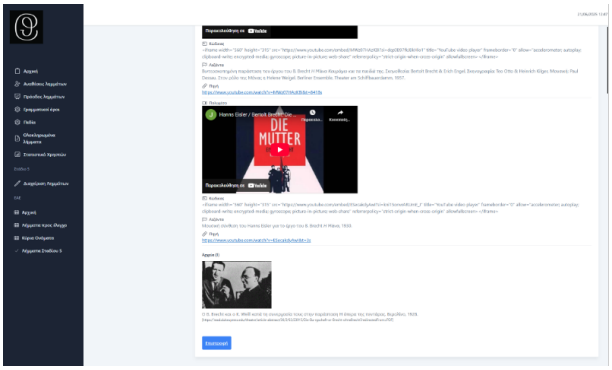


Fig. 5d

The lemma ανοικείωση [distanciation Verfremdungseffekt] (backend)

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588



FIG. 6a

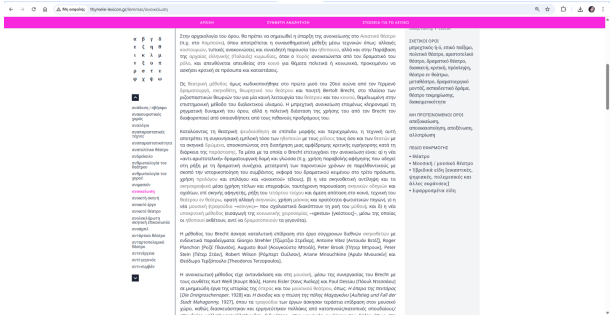


FIG. 6b



FIG. 6c

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588



Fig. 6d



Fig. 6e

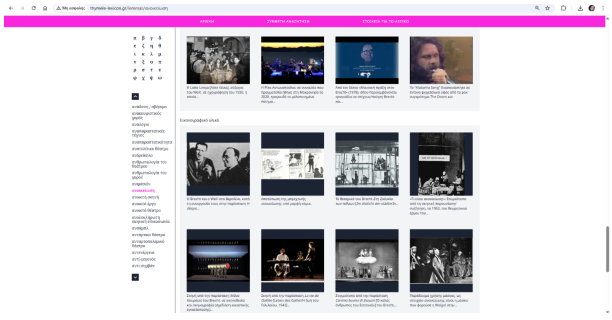


Fig. 6f

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

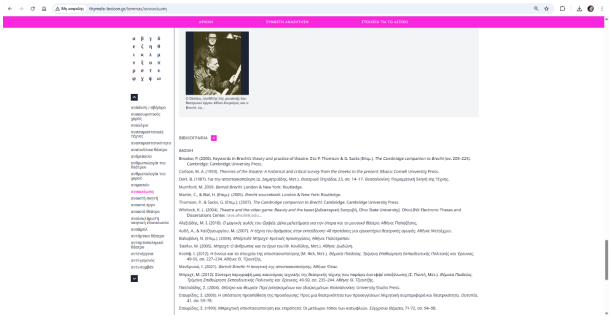


Fig. 6g



Fig. 6h

The lemma ανοικείωση [distanciation , Verfremdungseffekt] (frontend)

Key points demonstrated by the snapshot:

- All mandatory micro-fields appear in a single, scrollable pane; collapsible accordions reveal optional sections (multimedia, citations, notes).
- Ontological relations (broader/narrower/related) are stored as RDF triples and visualized as clickable graph edges; external authority links are generated automatically and then validated by editors (see 3.3.3).
- The “Revision history” panel at the bottom records every commit, enabling users to cite either the living entry or a frozen version for long-term reproducibility.

This example shows how textual data, multimedia evidence, and semantic links converge inside a single lemma, turning what would traditionally be a linear dictionary entry into a node within an extensible knowledge network.

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

4 Impact and Innovation

4.1 Contribution to performing arts research

Thymele delivers the first fully documented, Greek-centered vocabulary that spans the entire spectrum of theatre, dance, music, performance art, and hybrid/digital forms. By aligning each term with corpus evidence and ontological relations, the project supplies a reference grid that can anchor scholarly argument, metadata design, and pedagogical practice. Researchers gain a stable set of concepts for describing creative processes – e.g. meta-opera, immersive scenography, algorithmic composition – without resorting to ad hoc calques. Archives can index holdings with persistent *Thymele* identifiers rather than free-text tags, making it feasible to conduct federated queries across disparate catalogs. Because every term is multilingual, Greek material also becomes more discoverable in international portals, thereby inserting local scholarship into global conversations on performance theory and history (Marinou 2025, 4).

The dictionary's multimodal layer further enriches research possibilities. Embedded video clips, audio files, and high-resolution images allow scholars to juxtapose textual definitions with sensory evidence, an approach suited to embodied art forms where meaning is often encoded in gesture, timbre, or spatial design. Early classroom trials at two Greek universities indicate that students who consult *Thymele* report higher confidence in applying specialist terms during studio critiques and written assignments, suggesting that the resource shortens the gap between theoretical discourse and artistic practice (unpublished pilot data).

4.2 Methodological innovation in digital lexicography and digital humanities

Thymele advances lexicographic methodology on three fronts. First, it integrates corpus-based evidence with an explicit ontology inside the same authoring environment; term selection, definition writing, and semantic linking are therefore interdependent tasks rather than separate work packages. Second, the project demonstrates how multimodal lexicography can move beyond simple image thumbnails: IIIF services deliver deep-zoom stage photographs, while adaptive streaming makes rehearsal footage accessible even on low-bandwidth connections. This design acknowledges that, in the performing arts, a lighting cue or a movement motif may function as primary data equal in status to written description.

Third, *Thymele* is one of the first humanities dictionaries to embed machine-learning services directly into the editorial workflow (Kaddas *et al.* 2023; Doumanas *et al.* 2025). OCR pipelines convert legacy PDFs into searchable text; NER proposes candidate lemmas; a link-discovery engine suggests new ontology edges. Editors remain in control – no automatic change reaches the public instance, but preliminary metrics show that AI assistance halves the time required to finalize an entry and reduces omission errors in inter-lemma linking by roughly one-third. These findings contribute empirical evidence to ongoing debates about the productive coexistence of Large Language Models (LLMs) and expert curation in the humanities (Francis-Landau, Durrett, and Klein 2016).

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

The platform's technical stack – API gateway, triplestore, IIIF asset server – implements the FAIR principles in a lexicographic setting. Each lemma is findable *via* a persistent URI, accessible through open protocols, interoperable as JSON-LD/RDF, and reusable under a Creative Commons license. In doing so, Thymele offers a blueprint for how terminology resources can participate in the European Open Science Cloud (EOSC) without extensive re-engineering.

4.3 Scalability and cross-disciplinary transfer potential

Although tailored to the performing arts, *Thymele's* architecture is intentionally generic. Term templates are stored as JSON schemas; swapping the performing-arts ontology with, for example, a heritage-craft or film-studies ontology would require configuration rather than rewrites. The modular micro-services – OCR, NER, link discovery – communicate through message queues, so additional language models or domain-specific authority files can be plugged in with minimal disruption.

Three concrete extension scenarios are already under negotiation. First, a consortium of museology departments is testing *Thymele's* authoring workbench for a bilingual terminology of exhibition design, exploiting the existing multimedia fields to document floor plans and lighting plots. Second, a national library pilot will re-use the triple store and API layer to expose its controlled vocabularies as LOD, with *Thymele's* role system adapted to library catalogers. Third, an EU-funded network on intangible cultural heritage is exploring the platform as a shared terminological hub that could federate glossaries on oral traditions, festive events, and traditional craftsmanship (Ciciliato 2020).

Scalability also applies to language coverage. Because multilingual equivalents live in a dedicated table, adding Spanish or Russian involves curating the new column rather than altering the core schema. The crowdsourcing interface will allow native speakers to contribute candidate equivalents that editors can then validate. Over time, *Thymele* can therefore evolve into a Pan-European reference point while preserving Greek as its semantic anchor.

In sum, *Thymele's* impact extends well beyond its 600 initial entries. By coupling rigorous scholarship with open technological standards and participatory governance, the project not only stabilizes terminology in a rapidly evolving artistic field but also offers a transferable model for other disciplines wrestling with similar challenges of multilingualism, multimodality, and digital sustainability.

5 Conclusion

Thymele arrives at a moment when the performing arts expand into hybrid, digitally inflected territories while Greek terminology remains fragmented and uneven. This article has outlined the conceptual rationale, collaborative workflow, and technical architecture of a platform designed to redress that imbalance. By combining corpus-based evidence, ontology-driven modeling, and a tiered crowdsourcing mechanism, *Thymele* is producing an initial set of 600 peer-reviewed entries that can underpin archival metadata, university curricula, and international research agendas.

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

Methodologically, the project advances digital lexicography by integrating rich-media assets, version-controlled editing, and AI-assisted services within a single authoring environment. Compliance with FAIR principles – persistent identifiers, open protocols, RDF/JSON-LD export, Creative Commons licensing – ensures that every lemma can circulate across the wider ecosystem of DH infrastructures. The dual collaboration model preserves scholarly rigor while inviting sustained participation from practitioners, students, and the public.

Future work will extend both linguistic and disciplinary reach. Planned activities include adding more languages (e.g. Spanish, Russian), enlarging the reference corpus for stronger statistical validation, and adapting the software framework to adjacent domains such as museum exhibition design and intangible cultural heritage (Martin *et al.* 2020, 112). In these ways *Thymele* aspires to function not merely as a static dictionary but as a transferable blueprint for the creation, interconnection, and long-term stewardship of specialized terminologies in the digital age.

References

Bardiot, Clarisse. 2021. *Arts de la scène et humanités numériques: Des traces aux données.* London: ISTE Editions.

Chantraine, Pierre. 1944. *Dictionnaire étymologique de la langue grecque: Histoire des mots.* Paris: Klincksieck.

Christofidou, Anastasia. 2017. "Neología ke Orología stin Akadimía Athinón Neology and Terminology at the Academy of Athens." In *Greek Language and Terminology*, edited by Kostas Valeontis, 25–38. Athens: Academy of Athens.

Ciciliato, Vincent. 2020. *Arts et pratiques de l'archivage numérique: Collecter, cataloguer, cartographier.* Saint-Étienne: Publications de l'Université de Saint-Étienne.

Doerr, Martin, Chryssoula Bekiari, and Patrick Le Bœuf. 2008. "FRBRoo, a Conceptual Model for Performing Arts." Paper presented at the *CIDOC Annual Conference*, Athens, 15–18 September.

Doumanas, Dimitrios, Andreas Soularidis, Dimitris Spiliotopoulos, Costas Vassilakis, and Konstantinos Kotis. 2025. "Fine-Tuning Large Language Models for Ontology Engineering: A Comparative Analysis of GPT-4 and Mistral." *Applied Sciences* 15 (4): 2146. https://doi.org/10.3390/app15042146

Dubatti, Jorge. 2012. *Introducción a los estudios teatrales: Propedéutica.* Buenos Aires: Atuel.

Dubatti, Jorge. 2021. *Teatro y territorialidad: Perspectivas de filosofía del teatro y teatro comparado.* Barcelona: Gedisa Editorial.

European Commission. 2020. "European Multilingualism Policy Statement." Brussels. Available at: https://europa.eu/multilingualism-policy (accessed 12 June 2025).

Francis-Landau, Matthew, Greg Durrett, and Dan Klein. 2016. "Capturing Semantic Similarity for Entity Linking with Convolutional Neural Networks." arXiv:1604.00734

Fuertes-Olivera, Pedro A., and Sven Tarp. 2014. *Theory and Practice of Specialised Online Dictionaries: Lexicography versus Terminography.* Berlin: De Gruyter.

Gatto, Maristella. 2014. *Web as Corpus. Theory and Practice.* London: Bloomsbury.

Humbley, John. 2018. "Specialized Dictionaries." In *The Routledge Handbook of Lexicography*, edited by Pedro A. Fuertes-Oliveira, 317–334. Abingdon: Routledge.

Kaddas, Panagiotis, Basilis Gatos, Konstantinos Palaiologos, Katerina Christopoulou, and Konstantinos Kritsis. 2023. "Text Line Detection and Recognition of Greek Polytonic Documents." In *17th International Conference on Document Analysis and Recognition (ICDAR 2023)*, International Workshop on Machine Learning (4th ed.), Lecture Notes in Computer Science 14194, 213–225. Cham: Springer.

Lew, Robert. 2010. "Multimodal Lexicography: The Representation of Meaning in Electronic Dictionaries." *Lexikos* 20: 290–306.

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN: 9782813005588, doi: 10.17184/eac.9782813005588

Macken, Lieve, Els Lefever, and Véronique Hoste. 2013. "TExSIS: Bilingual Terminology Extraction from Parallel Corpora Using Chunk-Based Alignment." *Terminology* 19 (1): 1–30. https://doi.org/10.1075/term.19.1.01mac

Marinou, Dionysia. 2025. "I omada pou organonei tin istoria tis technis [The team that organizes art history]." Τα Νέα/Ta Nea, 3 May 2025, 4–5. https://staff.uop.gr/docs/2025-05/Μαρίνου Δημοσίευμα Τα Νέα 03.05.2025.pdf

Martin, Paul, Barbara Magagna, Xiaogang Liao, and Ziheng Zhao. 2020. "Semantic Linking of Research Infrastructure Metadata." In *Towards Interoperable Research Infrastructures for Environmental and Earth Sciences*, edited by Ziheng Zhao and Mikael Hellström, 105–117. Lecture Notes in Computer Science 12003. Cham: Springer.

Ooi, Vincent B. Y. 2018. "Using the Web for Lexicographic Purposes." In *The Routledge Handbook of Lexicography*, edited by Pedro A. Fuertes-Oliveira, 684–700. Abingdon: Routledge.

Papalexiou, Eleni. 2020. "Towards a Model of Digital Narration of the Creative Process of Performance." *European Journal of Theatre and Performance* 2: 376–423.

Papalexiou, Eleni and Avra Xepapadakou. 2022. "Tracing the Past: The Archive of Romeo Castellucci and the Socìetas Raffaello Sanzio—A Contribution to the History of European Theatre." In *Creative Archives as Living Landscapes of Memory in the Digital Era*, edited by Eleni Papalexiou, 147–194. Athens: FagottoBooks.

Papalexiou, Eleni, Anastasia Christofidou, Io Manolessou, Avra Xepapadakou, Costas Vassilakis, Vassilis Poulopoulos, and Basilis Gatos. 2025. "Renforcer la Terminologie Grecque pour les Arts de la Scène: Le Cas du Dictionnaire *Thymèle*. Défis épistémologiques, enjeux méthodologiques et applications numériques." Paper presented at TOTh 2025, Université Savoie Mont-Blanc (forthcoming).

Quillet, François. 2022. *La scène mondiale en période de confinement*. Paris: L'Harmattan.

Roche, Christophe. 2012. "Ontoterminology: How to Unify Terminology and Ontology into a Single Paradigm." In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, 2626–2630. Istanbul.

Thymele. (n.d.) "Thymele." Available at: https://thymele-project.uop.gr/ (accessed 12 June 2025).

Xepapadakou, Avra, Eleni Papalexiou, and Io Manolessou. 2025. "Fostering Multicultural and Multilingual Terminology in the Performing Arts: The Case 'Thymele', a Greek-Based Innovative Dynamic Multimedia Dictionary." *Bulletin suisse de linguistique appliquée* 122. (forthcoming).

Published in : BRISTAULT CANOVA, Juliette ; CHANTRAINE BRAILLON, Cécile ; ŽUGIĆ BORIĆ, Anamarija (dir.), *Performing Arts in the Digital Age. Research, Archives, and Innovation*, Editions des archives contemporaines, ISBN : 9782813005588, doi : 10.17184/eac.9782813005588