From plot to prompts: a case study on interactive narratives in museums

Enrico Mensa[1,*,†], Rossana Damiano[1,†]

[1]Dipartimento di Informatica, Università di Torino

Abstract

In the cultural field, Artificial Intelligence can assist heritage professionals in the creation and development of stories; in addition, it can manage interactivity, monitoring the audience response in real time. In this paper, we introduce a case study on interactive narratives in museums carried out at the Egyptian Museum in Turin within the framework of the Italian project "Cultural Heritage Active Innovation for Next-Gen Sustainable Society (CHANGES)". The case study consists in the design and implementation of a digital platform that assists curators in writing interactive stories tailored to the assets and equipment available for a given venue. The platform supports the curator in mapping the museum collection onto the story events and generating the text and media for each event using Large Language Models (LLM).

Keywords

Interactive storytelling, Prompt Engineering, Museum narratives

1. Introduction

In the cultural field, Artificial Intelligence can support the creation of narratives in two main complementary ways. On the one hand, AI can support human creativity, assisting cultural heritage professionals in the creation and development of stories; on the other, it can manage interactivity, monitoring the audience response in real time to optimize the progress of the story according to some predefined path.

Concerning the creative process, an input can come from the availability of data in digital form, which could be fed to AI tools to create (or support the creation of) narratives that fit the needs of specific cultural venues. However, the creation of well crafted, solidly constructed stories is still out the reach of AI: stories, in fact, obey to a complex and articulated set of conventions and rules, as witnesses by the long tradition of story and script writing literature, from Propp to Polti and Campbell [1, 2, 3]. For the same reason, creating engaging stories is notoriously a demanding activity, which requires professional skills and resources. In the cultural contexts, this problem is exacerbated by the need to keep a close control on the cultural notions conveyed by the narrative.

In this paper, we introduce a case study on interactive narratives in museums carried out at the Egyptian Museum in Turin[1] within the framework of the Italian project "Cultural Heritage Active Innovation for Next-Gen Sustainable Society (CHANGES)", 2022-2025. Developed as part of the action on "Virtual technologies for museums and art collections (Spoke 4)", the case study consists in the design and implementation of a digital platform that assists curators in writing interactive stories tailored to the assets and equipment available for a given venue. Our approach assumes that the story plot is manually created and helps the curator to map the museum collection onto the story events and generate the text and media for each event using Large Language Models (LLM) as a way to boost and support the creative process.

3rd Workshop on Artificial Intelligence for Cultural Heritage (AI4CH 2024, https://ai4ch.di.unito.it/), co-located with the 23nd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2024). 26-28 November 2024, Bolzano, Italy

*Corresponding author.

†These authors contributed equally.

enrico.mensa@runito.it (E. Mensa); rossana.damiano@unito.it (R. Damiano)

[1]https://www.museoegizio.it/

This paper is structured as follows. After discussing the role of storytelling in cultural contexts in Section 2, we describe the abstract architecture of our system in Section 3. Section 4 illustrates the creation of prompts from an annotated story events, exemplified in Section 4.3. Conclusion and future work end the paper.

2. Storytelling in Cultural Heritage: background and state of the art

The numerous advantages of storytelling have been explored in various disciplinary fields in the last decades. In the field of communication, storytelling is considered a particularly effective and easy to remember format for conveying information, thanks to the intrinsic compactness of the message [4, 5]. In cognitive and social psychology, it has been recognized that stories have the specific function of transmitting the values of a culture [6]. In media aesthetics, the ability of narratives to engage the public has been related to the process of identification with the characters [7, 8]. In narratology, narrative structures have been studied in order to extract patterns of meaning (motifs, tropes, situations) recurring between genres and cultures [2, 9] and media [3].

Looking specifically at the relationships between cultural heritage and narrative heritage (also intended in its meaning of intangible heritage), it can be observed that cultural objects often refer to narrative elements, such as mythological characters and deities (e.g. Ariadne, Horus), events and actions (e.g. leaving, fighting, donating), objects (e.g. the ball of yarn, the scales), and places (e.g. Arcadia, Rome). The intrinsic narrative component of cultural objects can be exploited by curators to create narratives with which visitors can relate on an emotional level, accessing through them the objects to which they refer. Such stories typically have varying degrees of "narrativity", ranging from a simple sequential ordering of objects to the construction of a set of chronological and causal relationships between narrative episodes. In the field of cultural heritage communication and dissemination, recent findings suggest that the experience of interactive narratives adds a sense of agency to the visitor's narrative experience positively affecting learning: "Agency, in particular, can indeed play a key role in transformative learning." [10].

Finally, it is necessary to highlight the difference between narrative content and its realization. The same narrative content, understood as a semantic representation of a set of narrative elements and the relationships (temporal, causal, thematic, etc.) between them, can be conveyed to the public through different media and languages [11], producing narratives that can be directly enjoyed by the public through a specific format, normally typical of a specific communication medium: the textual format of a biographical profile, the visual format of a chronological one on an information panel, the video format of a documentary in a thematic room. From a computational point of view, this distinction is relevant because it enables the separation of the representational component of a story, which can be manipulated at a symbolic level by a formal system, from its expressive component, which can be realized with different means and with different degrees of automation. For example, the work by [12, 13] exemplifies the creation of narratives that allows the audience to explore archaeometric findings in a VR application.

3. System Architecture

The architecture of the platform keeps the story editing process distinct from the story delivery (Figure 1), similarly to the approach described by [14]. The process starts from the Templates Editor and the Annotation Editor. The first tool is designed to assist curators in the development of *Narrative templates* (or story templates) which define the various scenes of a story. In each scene the curator can directly specify elements such as characters, objects and background by selecting them in the Knowledge Base (KB), or leave them as free 'variables' which will then be populated automatically at a later stage. The Annotation Editor, on the other side, is a tool designed for the annotation of museum elements. The annotation provides both semantic information regarding the museum item (e.g., its name, use, ownership, etc.) and technical details of the asset (e.g., its a video, image, etc.). The Story Editor then

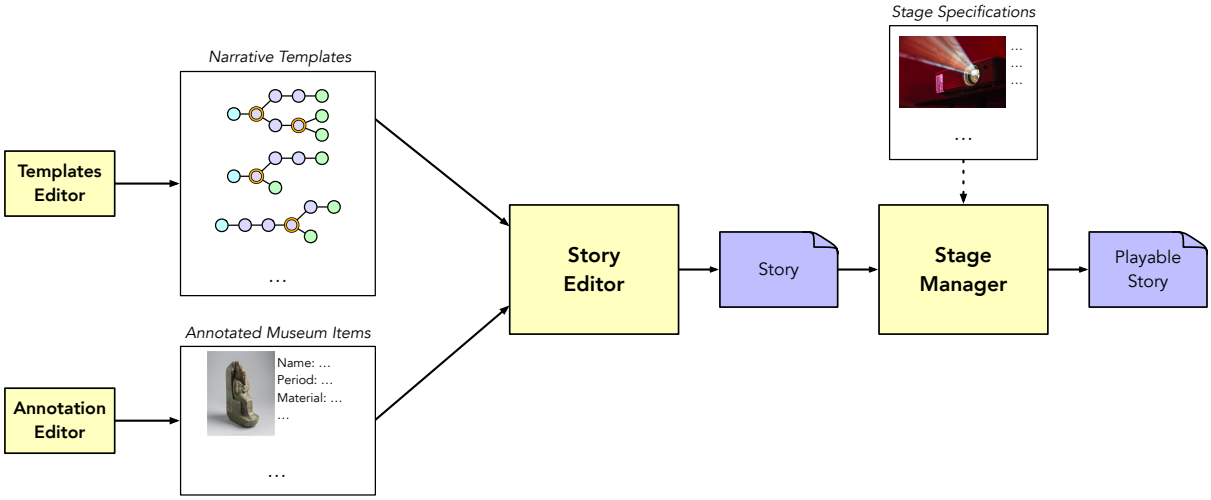

Figure 1: The System Architecture. The Story Editor allows the curator to create a complete story by filling a narrative template with the narrative entities from the museum collection. The Stage Manager converts the story specification into an object which can be played with the equipment of a specific exhibit.

fully instantiates a narrative template by fixing the variables that were left free by the curator. The output of this phase is a Story, namely a story graph where nodes (the story scenes) are populated with specific narrative characters and objects taking part into specific events and the transition to the next node is specified as a function of the audience input. Finally, the last element of the pipeline is the Stage Manager, which matches the populated story structure onto the available media assets in the platform (as provided by the Annontation Editor) and convert it into a complex digital object (i.e., an aggregate of scripts and resources for a given game engine) that can be played on the infrastructure available at the museum (the Stage) where the audience can enjoy it.

In detail, the repository of narrative templates consists of a set of pre-defined story templates, sort of "narrative archetypes" [15], whose function is to simplify the creation of stories by providing narrative patterns designed to meet the goal of engaging the audience of the museum. The notion of surveying and cataloguing patterns in storytelling and drama is rooted in a long tradition, spanning from ethnology [9] and classical studies (Highet 1949) to semiotics [1] and scriptwriting [2, 3]. Archetypal representation of stories and characters are pervasive in iconology, from Warburg's Bilderatlas to Iconclass, connecting narratology with art. The goal of these initiatives is to systematize and reuse narrative knowledge to study stories across cultures, ages and media, but also gather a reliable source of knowledge for creating compelling, engaging narratives from tested schemes and examples. (see also [16] for an application of patterns to improvisational theater).

The repository of narrative templates and the annotation of the museum items relies on the vocabulary for describing stories encoded in a Story Ontology, namely, a formal encoding of the notion of story, its parts, and the relations over them, expressed through a machine-readable, standard language [17, 18]. The purpose of the ontology is to provide a formal, unambiguous account of the narrative universe, shared by the human actors in the project and the software architecture that will support the creation and delivery of the stories in a realization environment. The story ontology formalizes the key elements of a narrative, including its entities (events arranged in a plot and enacted by characters, with roles such as protagonist and antagonist and accompanying emotions) and the relations over them (characters' actions causing effects in the story world and occurring before or after other events), thus providing the vocabulary through which stories are described in the system, from creation to delivery. By giving a formal semantics to the description of actual stories, they become open to the application automated reasoning that can, for instance, classify them into story types, verify their properties against the ontology, and support the generation of new stories from the existing ones. As mentioned before, the vocabulary provided by the Story Ontology is employed to encode the narrative entities of both narrative templates and annotated museum items, in order to ensure that the properties of latter match

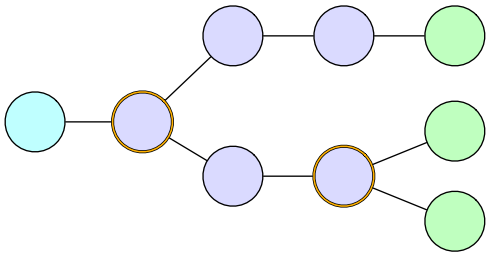

Figure 2: An example story graph. The start node is represented in blue, the end nodes in green. Decision nodes are characterized by multiple continuations.

the narrative context provided by each node in the template.

As an example of narrative entities which may be related to the domain of our case study, the Egyptian Museum in Turin, take the site called Geblein, situated in Upper Egypt. The archaeological area of Geblein is formed by two hills; on one of the two hills there are the remains of the Egyptian city of Per Hathor (house of Hator - goddess of joy, love, motherhood and beauty). On the other hill there are the remains of the city Inty-shuy. Given this context, the annotation of the location may include the following narrative entities: Characters, such as the goddess Hator, the ancient inhabitants of the two cities, Objects of common use in everyday life found in the site; Places, such as the temple of Hator, or the houses and buildings of the two cities. The archaeological context may include Characters, such as the archaeologists and local inhabitants working on the excavation, and the linguists expert in ancient languages and Places such as the natural elements present in the territory, the village and, last but not least, the archaeological excavation.

4. AI assisted story editing

As described in the previous section, the narrative template repository is designed to help curators create new variations of each story from its template. A story template defines a series of nodes (scenes) that follow one another (Figure 2). We refer to these nodes as template nodes belonging to a template story. We distinguish two types of nodes, that we describe in the following.

Narrative nodes. Narrative nodes are "simple" nodes, which present an outcome or a generic scene and therefore always and only have one subsequent node. They are made up of a series of fields, each of which defines a list of variables and constants. Variables will be automatically bound in the future to a specific value and to this end each variable is defined with a 'type' which restricts the legitimate values that the variable can assume. For example, the *Background* field could have a variable `B` that defines a generic background with type `Open_Space`. Alternatively, it is possible to directly assign a value to `B`, such as `TempioLuxor` which is a specific temple contained in the KB. In this case, `B` is a constant. It is important to note that variables in the i-th node can also recall variables from the $0 \dots (i-1)th$ previous nodes: this is a fundamental requirement to allow continuity. The time and climate in which the scene takes place can also be specified, together with the identifier for the next node in the graph.

The fields we decided to adopt are the following:

- Main characters
- Background characters
- Objects
- Background
- Time
- Weather
- Next node

In addition to these values, each template node also defines:

- Values: what values are represented in this node?
- Tones: what is the narrative tone of the node?
- Summary: a overall description of the node.

Finally, a template node defines a prompt template, that will be fed to a LLM of choosing at the end of the instantiation phase. The prompt template is defined in the template node in terms of its variables, e.g., "Write a story in which `MC`... " where `MC` is the variable corresponding to the main character. The tone field could also be exploited here to vary the tone of the resulting text.

Decision nodes. Decision nodes have the same skeleton as narrative nodes, however, the *next node* field is replaced by N next nodes, each with its own textual description. In particular, a question is posed to visitors who can choose between N paths.

An expansion to the framework could involve the implementation of a third type of node, *contextual decision nodes*. Visitors could be allowed to make decisions that do not change the plot of the story, but could potentially influence the appearance of certain characters, objects or backgrounds. This introduction poses complications for the stage manager that should then be able to dynamically allocate different resources depending on the decisions of the visitors.

4.1. Story instances

Once the template is fully developed, it can be instantiated. From one template, many variations (instances) of it can be generated. The curator who decides to generate a story can choose a template and proceed with an instantiation. This instantiation is developed in two steps, performed node by node.

Variable binding. Each free variable of the node is bound: values are extracted randomly from the KB in accordance with the associated type. Naturally, the binding of a variable `x` affects every further node that refers to `x` in one of its fields.

Text generation. Variables in the prompt template can now be replaced with bound values so that the prompt can be fed to an LLM. The result is then inserted within the node, for a subsequent review by the curator.

4.2. Playable Story

Once the instantiation attempt is finished, the multimedia elements associated to each bound variable are retrieved, in accordance with their role (e.g., for the `temple_02` background, a 3D render of the aforementioned temple is taken - or it is generated with a generative system). The various scenes are then composed, complete with texts/images/sounds/videos and submitted to the curator for a final correction. Alternative versions of the texts can be re-generated, while the multimedia elements can be replaced (always preserving the indications of the template node). The ability to manually edit automatically generated content is fundamental: curators must have the complete control on the final product in order to ensure that the LLMs do not produce conflicting or irrelevant texts. Once this process is finished, the instantiated story is now "playable".

4.3. Story Example

In this section we provide an example of a narrative node and show its instantiation process. The choosen story narrates the thief of funerary equipment and Figure 3 illustrates the interface for editing the template of narrative `node#3`.

In this example, the main character is defined via the constant `MC_1`, which is actually inherited by `node#1`. Its assigned value is `Ekhnaton`, which is the protagonist of this story (and a specific element of the KB). Three background characters are defined as variables: two have the type `noble_person`

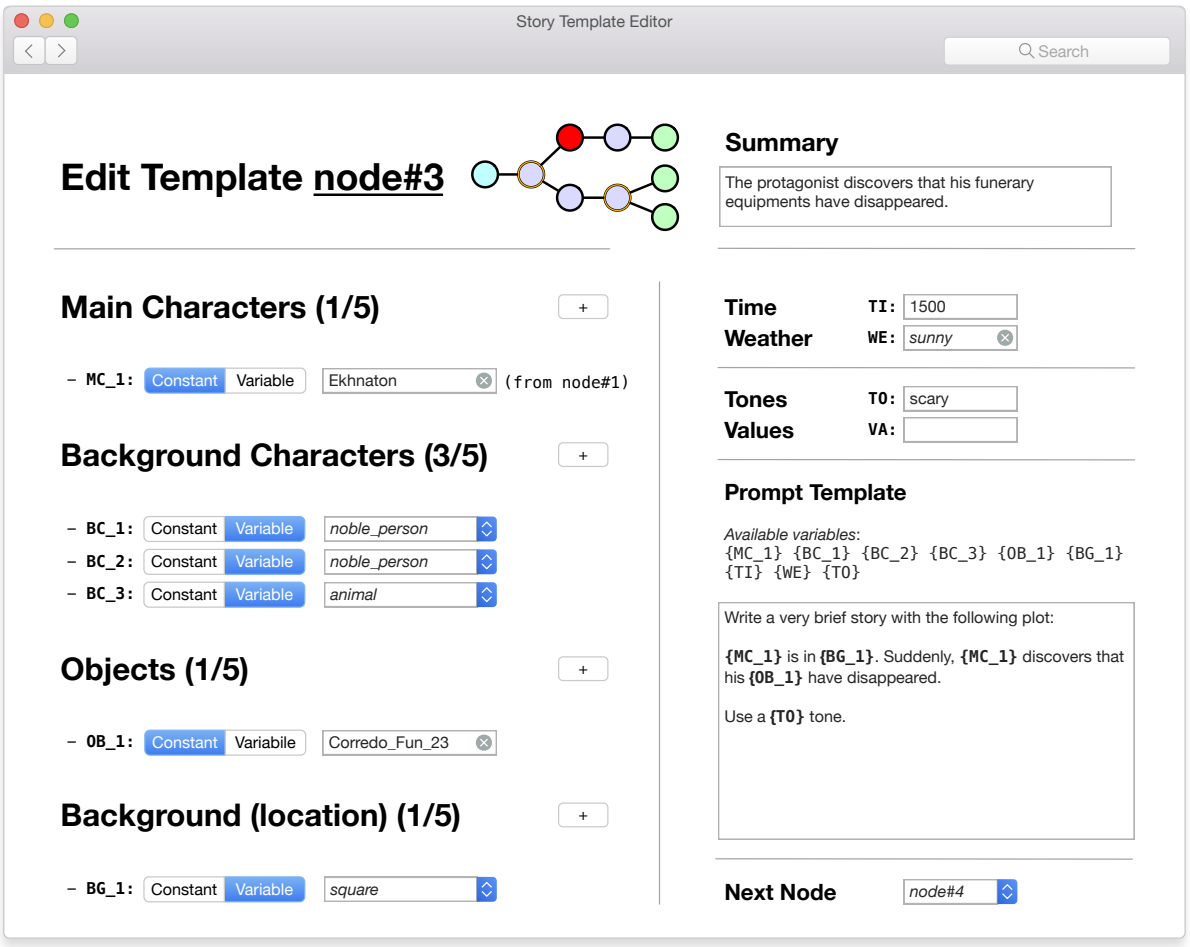

Figure 3: The Template Node editor. The panel on the left allows selecting the narrative entities appearing in the node; the box on the right contains the controls for setting the LLM and the generated prompt text.

and one the type `animal`. The key object of the scene is again a constant, `Corredo_Fun_23`, which is a specific instance of funerary equipment in the KB. The background is a variable of type `square`. The prompt template is compiled using the appropriate variables which will be replaced during the instantiation process. Finally, the next node is `node#4`.

During the instantiation process for this node (which happens during the instantiation of the whole template story), the variables for the background characters and the background are bound to specific elements of the KB: this search is based on the type of each variables, which is therefore employed as a filter. These newly bound values, together with the already defined `Ekhnaton` and `Corredo_Fun_23` constants are then replaced in the prompt template, which is then fed to a LLM for the text generation. Note that the tone (variable `TO`) is explicitly used in this prompt template.

5. Conclusion and Future work

This paper illustrates the design of a framework for the development of interactive narratives with the support of Generative AI models within the context of interactive museum narratives. In our approach the story is still handcrafted so to allow museum professionals to preserve historical accuracy, but the generation of all textual content for these stories is demanded to LLMs.

The development and testing of this framework will allow for the investigation of many research questions. The first issue is the appropriateness of off-the-shelf LLMs: are current LLMs suited for the task? LLMs are mostly trained on *written* stories rather then interactive ones. We are not searching for dialogues and detailed descriptions of the scenery (since the scene is directly shown to the curator via

the stage manager) and so the adherence to the prompts is of the utmost relevance in this application. The development of specific benchmarks will shed light on this issues.

A final interesting aspect has to do with the overall dimensionality of a story. How many decision nodes vs narrative nodes are appropriate? How many variables can be defined in a node? Again, these aspects require careful testing and evaluation. Also, we plan to extend the same approach to the generation of other media from scene specification.

References

[1] V. Propp, Morphology of the Folktale, University of Texas Press, 1968.

[2] G. Polti, Les trente-six situations dramatiques, Mercure de France, Paris, 1895.

[3] J. Campbell, The hero with a thousand faces, volume 17, New World Library, 2008.

[4] N. Gershon, W. Page, What storytelling can do for information visualization, Communications of the ACM 44 (2001) 31–37.

[5] R. Kosara, J. Mackinlay, Storytelling: The next step for visualization, Computer 46 (2013) 44–50.

[6] J. Bruner, The narrative construction of reality, Critical Inquiry 18(1) (1991) 1–21.

[7] G. Currie, Narrative and the Psychology of Character, Journal of Aesthetics and Art Criticism 67 (2009) 61–71.

[8] N. Carroll, Beyond Esthetics: Philosophical Essays, Cambridge University Press, New York: Cambridge, 2001.

[9] S. Thompson, Motif-Index of Folk-Literature, Volume 4: A Classification of Narrative Elements in Folk Tales, Ballads, Myths, Fables, Mediaeval Romances, Exempla, Fabliaux, Jest-Books, and Local Legends, volume 4, Indiana University Press, 1955.

[10] D. Petousi, A. Katifori, K. Servi, M. Roussou, Y. Ioannidis, Interactive digital storytelling in cultural heritage: the transformative role of agency, in: International Conference on Interactive Digital Storytelling, Springer, 2022, pp. 48–67.

[11] J. D. Bolter, R. Grusin, Remediation: Understanding new media, Mit Press, 2000.

[12] V. Murtas, V. Lauro, V. Lombardo, Virtual archaeology in a multi-platform and multi-lingual setting, in: Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization, 2023, pp. 422–426.

[13] V. Lombardo, V. Lauro, V. Murtas, S. Goud, Merging archaeological site recreation and museum exhibition, in: International Conference on Interactive Digital Storytelling, Springer, 2023, pp. 68–84.

[14] R. Damiano, V. Lombardo, G. Monticone, A. Pizzo, Studying and designing emotions in live interactions with the audience, Multimedia Tools and Applications 80 (2021) 6711–6736.

[15] R. Damiano, V. Lombardo, A. Lieto, D. Borra, Exploring cultural heritage repositories with creative intelligence. the labyrinth 3d system, Entertainment Computing 16 (2016) 41–52.

[16] V. Vranić, A. Vranić, Drama patterns: Extracting and reusing the essence of drama, in: Proceedings of the 24th European Conference on Pattern Languages of Programs, 2019, pp. 1–9.

[17] R. Damiano, V. Lombardo, A. Pizzo, The ontology of drama, Applied Ontology 14 (2019) 79–118.

[18] R. Damiano, A. Lieto, et al., Ontological representations of narratives: a case study on stories and actions, Open Access Series in Informatics (2013) 76–93.