

GOMME: A Generic Ontology Modelling Methodology for Epics

Udaya Varadarajan

Documentation Research and Training Centre (DRTC), Indian Statistical Institute, Bangalore Centre, India
Department of Library and Information Science, University of Calcutta, Kolkata, India
E-mail: udayav@drtc.isibang.ac.in

Mayukh Bagchi

Department of Information Engineering and Computer Science (DISI), University of Trento, Trento, Italy
E-mail: mayukhbagchi@drtc.isibang.ac.in

Amit Tiwari*

Documentation Research and Training Centre (DRTC), Indian Statistical Institute, Bangalore Centre, India
Department of Library and Information Science, University of Calcutta, Kolkata, India
E-mail: amittiwari@drtc.isibang.ac.in

M. P. Satija

Department of Library and Information Science, GNDU, Amritsar, India
E-mail: satija_mp@yahoo.com

ABSTRACT

Ontological knowledge modelling of epic texts, though being an established research arena backed by concrete multilingual and multicultural works, still suffers from two key shortcomings. Firstly, all epic ontological models developed till date have been designed following ad-hoc methodologies, most often combining existing general purpose ontology development methodologies. Secondly, none of the ad-hoc methodologies consider the potential reuse of existing epic ontological models for enrichment, if available. This paper presents, as a unified solution to the above shortcomings, the design and development of GOMME - the first dedicated methodology for iterative ontological modelling of epics, potentially extensible to works in different research arenas of digital humanities in general. GOMME is grounded in transdisciplinary foundations of canonical norms for epics, knowledge modelling best practices, application satisfiability norms, and cognitive generative questions. It is also the first methodology (in epic modelling but also in general) to be flexible enough to integrate, in practice, the options of knowledge modelling via reuse or from scratch. The feasibility of GOMME is validated via a first brief implementation of ontological modelling of the Indian epic *Mahabharata* by reusing an existing ontology. The preliminary results are promising, with the GOMME-produced model being both ontologically thorough and competent performance-wise.

Keywords: ontology development methodology, GOMME methodology, modelling epics, *Mahabharata* ontology, ontology reuse methodology, digital humanities

Received: July 21, 2022
Accepted: January 10, 2023

Revised: December 2, 2022
Published: March 30, 2023

*Corresponding Author: Amit Tiwari
 <https://orcid.org/0000-0003-4147-3931>
E-mail: amittiwari@drtc.isibang.ac.in



All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

The most cited definition of ontology is that it is “an explicit specification of a shared conceptualization” (Gruber, 1995). Ontology development methodologies are practices that focus on the life cycle of ontologies, their development process, and the techniques, tools, and languages employed to create them (Gómez-Pérez et al., 2004). The research arena of Knowledge Representation (KR) has witnessed a plethora of knowledge modelling methodologies exploiting a diverse array of KR formalisms including semantic nets (Lehmann, 1992), frames (Lassila & McGuinness, 2001), ontologies (Guarino et al., 2009; Phoksawat et al., 2019) and now knowledge graphs (Bagchi & Madalli, 2019; Ehrlinger & Wöß, 2016). Such methodologies can be generic (Fernández-López et al., 1997; Noy & McGuinness, 2001) or tailor-designed for specialized domains such as for health (Das & Roy, 2016), machine translation (Mahesh et al., 1996), and smart cities (Espinoza-Arias et al., 2019). Tailor designed or ad-hoc methodologies most often combine an existing general methodology or methodologies. They are created as and when necessary, in particular, for modelling knowledge in diverse arenas of digital humanities such as folklore (Abello et al., 2012), narrative information (Damiano & Lieto, 2013; Swartjes & Theune, 2006), and epics and mythology (Hyvönen et al., 2005; Iqbal et al., 2013; Syamili & Rekha, 2018).

1.1. Motivation

As mentioned in the introduction, the ad-hoc approaches (for instance, in Syamili & Rekha, 2018) inherit the architectural characteristics and advantages of the combinations of established methodologies. But they suffer from three crucial shortcomings specific to modelling epic texts and other genres in digital humanities in general. Firstly, as observed from the literature study, there exists no methodology with a focus on modelling and representing epics (Greene, 1961) and on works in digital humanities in general. A generic ontology development methodology will not be satisfactory, because modelling epics requires classificatory and ontological elegance and should adhere to well-established literary cardinals (similarly for other literary genres). Secondly, the existing models across the domains in digital humanities do not discuss potential reuse of existing ontological formalizations. Thirdly, in accordance with dynamism in knowledge organization and representation (Giunchiglia et al., 2014; Kent, 2000), none of the approaches for modelling domains in

digital humanities stress iterative knowledge development and modelling at the methodological level itself. This is important to ensure that the model is constantly perfected from the feedback mechanism within. The motivation of the present work lies in the above shortcomings.

1.2. Novelty

In response to the aforementioned challenges, this work develops and proposes GOMME. The novelty of the work lies in being the first dedicated methodology for iterative ontological modelling of literary epics. The other novel factors of the work are: Firstly, GOMME is grounded in the novel characteristic design cardinal of re-purposing and flexibility, which makes the methodology (or appropriate fragments of it) completely customizable and extensible for modelling domains in (digital) humanities and social sciences beyond epics. Secondly, the methodology is not only grounded in knowledge modelling best practices and application satisfiability norms, but also adheres to canonical norms for epics (Greene, 1961) which makes it innately suitable for modelling epics while, at the same time, preserving (some of) its unique literary features (detailed in Section 3). Thirdly, the novel fact that GOMME, in practice, offers the flexibility of not only modelling an epic from scratch but also is natively grounded in the ontology reuse paradigm. It thereby accommodates reuse and potential enrichment of any formal or conceptual model of an epic, if available.

The remainder of the paper is organized as follows: Section 2 details the state-of-the-art frameworks in modelling literary works and narrative information, and highlights the research gaps. Section 3 describes, in fine detail, the foundations and steps of the GOMME methodology. Section 4 presents the proof of the methodology via a brief case study of modelling the Indian epic *Mahabharata*, showing the feasibility and advantages of the GOMME methodology. Section 5 discusses some notable implications of the work and Section 6 concludes the paper with a discussion of future prospects of the present work.

2. LITERATURE REVIEW

Ontology development methodologies are “activities that concern with the ontology development process, the ontology life cycle, and the methodologies, tools, and languages for building ontologies” (Gómez-Pérez et al., 2004). For over two decades, research has been happening in the domain of methodologies for ontology development. For the current work, the most commonly reused

and acclaimed methodologies were studied. Ontology Development 101 (Noy & McGuinness, 2001) is an initial guide for amateur ontology developers; NeON methodology (Suárez-Figueroa et al., 2015) focuses on building large-scale ontologies with collaboration and reuse; and DILIGENT (Vrandečić et al., 2005) is a methodology that lays emphasis on ontology evolution and not on initial ontology designing. TOVE (Gruninger & Fox, 1995a) highlights ontology evaluation and maintenance. ENTERPRISE (Uchold & King, 1995) discusses the informal and formal phases of ontology construction sans identification of the ontology concepts. Yet Another Methodology for Ontology (YAMO) (Dutta et al., 2015) is a methodology for ontology construction for large-scale faceted ontologies. The principles of facet analysis and analytico-synthetic classification guide the YAMO methodology. The methodology was used in the construction of ontology for food.

Following this, grounded in the general methodologies as aforementioned, a literature search was conducted to identify specific methodologies for narrative/literary ontological models, via the models developed in the similar

domains. These works are similar in scope to our objective of examining the state-of-the-art for modelling epics. The existing models do not follow any standard, matured ontology development methodologies. Instead, they have ad-hoc methodologies. Varadarajan and Dutta (2021a) conducted a study of such models (Table 1 details the same). The study discovered a lack of any methodology dedicated to the development of narrative/literary ontologies.

We now focus on briefly elucidating some of the above works. Constructed on the basis that the narrative situation (Klarer, 2013) needs characters and objects which form a larger story once connected, the ontology by Mulholland et al. (2004) provides intelligent support for the exploration of digital stories to encourage heritage site visits. Hyvönen et al. (2005) develop a framework and national ontology for representing the culture and heritage of Finland on the Semantic Web. Developed as an upper ontology, it is based on the Finnish General Thesaurus. The National Library of Finland is the competent authority for its maintenance. The ontology by Nakasone and Ishizuka (2006) is constructed with the generic aspects of

Table 1. Selected ontology models for study

Ontology name	Purpose	Domain	Ontology design methodology
Ontology for story fountain (Mulholland et al., 2004)	To describe stories and related themes	Literature	Not available
Finnish national ontology (FinnONTO) (Hyvönen et al., 2005)	To express the national Finnish Thesaurus as ontologies	Cultural heritage	Not available
Ontology model for storytelling (Nakasone & Ishizuka, 2006)	To build a coherent event centric generic storytelling model	Literature	Not available
The archetype ontology (AO) (Damiano & Lieto, 2013)	To link the various resources in the archive through narrative relations	Cultural heritage	Not available
Quran ontology (Iqbal et al., 2013)	To develop ontology that provides contextual information for the correct interpretation of Quran	Religious text/Epic	Mixed existing methodologies
ODY-ONT (Khan et al., 2016)	Explicit representation of the story found in any text	Epic	Not available
Transmedia fictional worlds ontology (Branch et al., 2017)	To represent the elements of the fictional world, connecting characters, places, and events	Literature	Not available
BK Onto (Yeh, 2017)	To capture biographical information	Literature	Not available
Ontology for Greek mythology (Syamili & Rekha, 2018)	To develop an ontology for Greek mythology	Mythology/Epic	Mixed methodologies
Drammar ontology (Damiano et al., 2019)	Represent drama elements independent of media and task formally	Literature	NeOn

storytelling as its founding philosophy. The purpose of a domain independent model was to provide coherence to the events in the story. The Archetype Ontology (Damiano & Lieto, 2013) is built to explore a digital archive via narrative relations among the resources. To query and retrieve contextual information along with Quran verses, an ontology was developed by Iqbal et al. (2013). A merging of various methodologies was adopted to model this ontology. To represent narration in a literary text, the ODY-Onto (Khan et al., 2016) was constructed. The developed ontology is part of a system constructed for querying information from literary texts. Transmedia ontology (Branch et al., 2017) allows users to search for and retrieve the information contained in the transmedia world. The ontology will help in inferring connections between transmedia elements such as characters, elements of power associated with characters, items, places, and events. The ontology contains 72 classes and 239 properties. Biographical Knowledge Ontology (BK onto) was created to capture biographical information. The ontology was deployed in the Mackay Digital Collection Project Platform (<http://dml.csie.au.edu.tw>) for linking “the event units to correlate to the contents of external digital library/archive systems so that more diverse digital collections can be presented in StoryTeller system” (Yeh, 2017). Ontology for Greek Mythology (Syamili & Rekha, 2018) developed an ontology for the relationship among Greek mythological characters. For ontology development, a combination of the existing methodologies was reused. The Drammar ontology (Damiano et al., 2019) was developed to represent the elements of drama independent of the media and task. Drama, as a domain, is evolving, but there is a concrete manifestation of drama such as in screenplays, theatrical performances, radio dramas, and movies. The top four classes of dramatic entities are (1) DramaEntity, the class of the dramatic entities, i.e., the entities that are peculiar to drama; (2) DataStructure, the class that organizes the elements of the ontology into common structures; (3) DescriptionTemplate, which contains the patterns for the representation of drama according to role-based templates; and (4) ExternalReference, the class that bridges the description of drama to commonsense and linguistic concepts situated in external resources.

Observations from the literature study indicate that the generic ontology development methodologies involve domain identification, term collection, relationship and attributes establishment, representation, and documentation. These methodologies are for ontology development from scratch. Rarely do such methodologies integrate

the reusing of existing ontologies. Similar is the case with theoretically established methodologies.

The methodologies used for ontology construction in narrative/literary domains, on the other hand, are ad-hoc and assume previous knowledge in ontology construction. The steps in the methodologies often start with the ontology population. Further investigation found that the ontology development methodologies for narrative/literary domains lack an iterative process and that these methodologies are mostly linear. A linear methodology is unidirectional and does not include traversing in a feedback loop, therefore leaving very little space for improvement. Keeping these research gaps in mind, the following section will detail our proposed GOMME methodology.

3. THE GOMME METHODOLOGY

This section elucidates the GOMME methodology in detail. In Section 4, a first brief implementation of GOMME for modelling the Indian epic *Mahabharata* is detailed. The *Mahabharata*, the longest epic in the history of world literature, was originally written by Ved Vyasa in Sanskrit. It comprises around 100,000 stanzas (Lochtefeld, 2002). The *Mahabharata* narrates the story of a kingdom called Hastinapur, where two brothers’ families, Pandavas and Kauravas, fought for control over the kingdom. Each of the sides was directly backed up by several other kingdoms. In the end, Vyasa illustrated the war between the two sides, and the Pandavas won with the help of Lord Krishna. Several connecting mythological stories make *Mahabharata* more comprehensive (Rajagopalachari, 1970). The stages of GOMME are discussed as follows.

3.1. Initial Scoping Exercise

The initial scoping exercise (see Fig. 1) marks the inception for modelling epic texts in the context of the GOMME methodology. The step is an exercise in consolidating a set of objective-focused design guidelines for initiating, scoping, and informing the steps of the methodological architecture in an all-encompassing manner. A non-exhaustive candidate set of such design concerns was identified and discussed as follows: Primary design concern is scopic review, i.e., to achieve a shared convergence in delineating the existing formal modelling attempts in the chosen (epic) domain, culminating in research gaps yet to be addressed. The next was addressing feasibility and novelty of research ideation, which involves brainstorming and federating the theoretical novelty and implementational feasibility of the proposed research

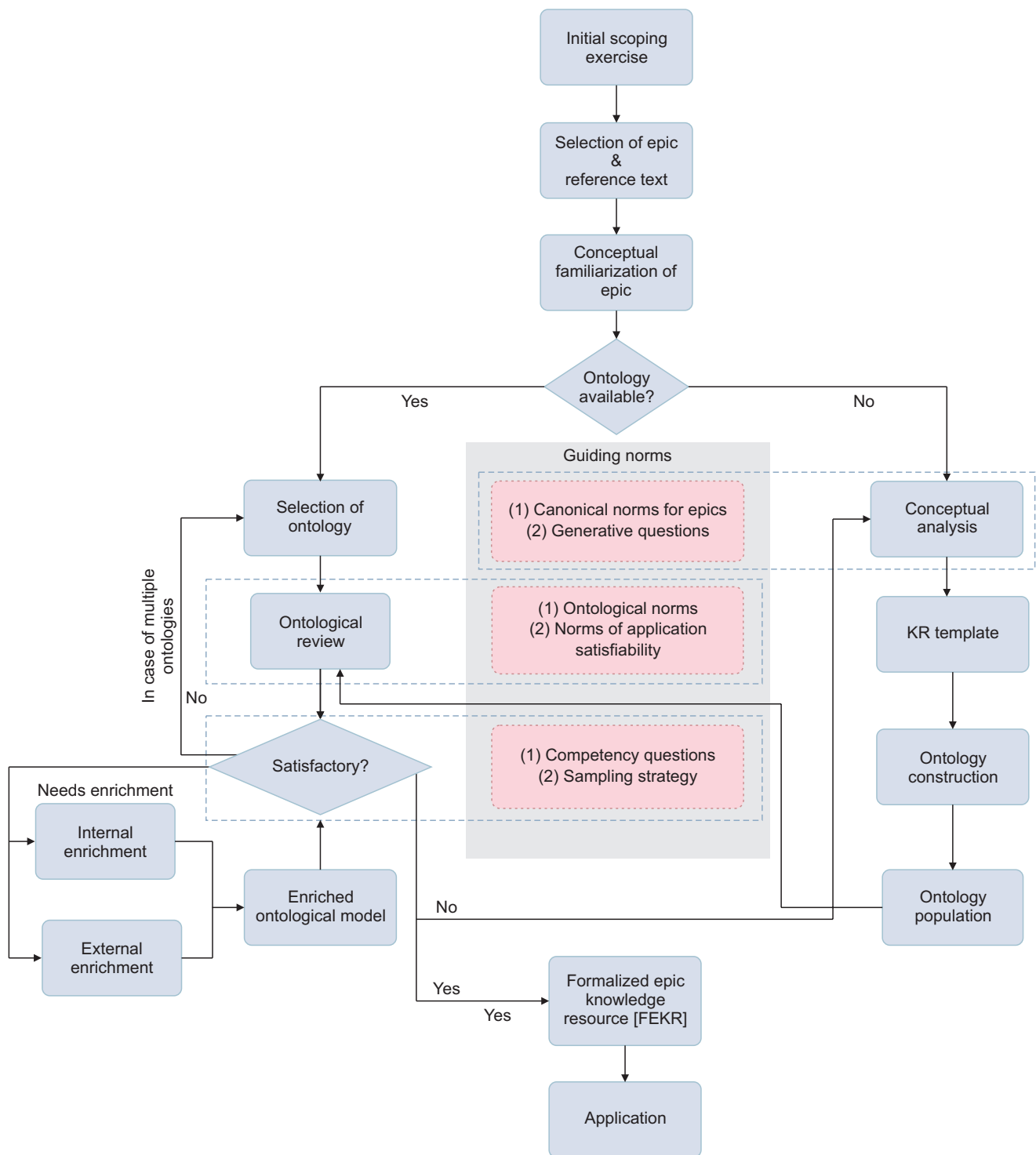


Fig. 1. The GOMME methodology. KR, knowledge representation.

ideation in sync with the research scope as understood in the scopic review. The third design concern was related to technologies and optimization, which is an attempt at a broad identification of conceptual tools (often transdisci-

plinary) and technological paradigms required for optimal effectuation of the architecture. The next design concern was architectural agility, meaning to ground the overall architectural flow in the notion of technological agility, thus

realizing it as not only methodologically a-technological per se but also incremental, adaptive, and scalable (for instance, via maturity models; see Tiwari & Madalli, 2021 for a recent study). The final concern was purposing and extensibility, i.e., to incorporate the scope for repurposing the architectural components (or the architecture in its entirety), thereby facilitating its extension in varied contexts situated within the scope of knowledge modelling.

The above-mentioned design concerns are indicative of generic aspects which might be considered in the initial scoping exercise and are completely flexible in allowing added tailored formulations for specific research contexts, for instance, if we extend for other categories of works in the digital humanities. The methodology recommends the involvement of the entire Project Development Team (PDT), by which we mean the team responsible for end-to-end coordination of the GOMME methodology for a particular project, for the initial scoping exercise.

3.2. Selection of Epic and Reference Text

An epic usually transgresses commonplace literary praxis and might dwell in conventions as varied as proposition, invocation, enumeration, or in medias res (Cooren, 2015). Therefore, mindful selection of an epic and a related reference text is crucial for robust development and evolution of the proposed methodological architecture. The process of deciding on an epic must involve the following conventions in varying degrees. The conventions are cultural affinity, literary expertise, and computational amenability. By cultural affinity, it is recommended to select an epic which embodies, to a certain extent, cultural coherence and rootedness with (at least some of) the PDT members. Literary expertise indicates that the selection of an epic should also factor in the availability of requisite literary expertise, without which an inceptive understanding and subsequent conceptual analyses will prove knotty; and computational amenability implies that the selected epic should have the ability to elicit, engineer, actualize, and manipulate its specific representation computationally. The convention is non-trivial; for instance, the Buddhist epic *Mahavamsa* written in the Pali language, a language for which literary expertise is scant, thus potentially has low computational amenability.

Similarly, the approval of a standard reference text(s) for the selected epic should also be based on certain non-normative norms, the core of which are scholarly standing, contributor foothold, equilibrium of comprehensiveness, and scholarly currency. Scholarly standing means that the chosen reference text is expected to have high standing

amongst relevant scholarly communities. Book reviews, exegeses, and bibliographic citations can be some parameters cumulatively determining a text's scholarly authority. The contributor foothold refers to the scholarly reputation of the contributors (editors, translators, etc.) to the reference text. In general, it is best to select the work of those contributors who are noted to have proven expertise and intellectual acumen in the chosen epic. The equilibrium of comprehensiveness addresses the selection of reference text in such a fashion that it exhibits a healthy equilibrium between brevity of message, description, and all-encompassing coverage of the crux of the epic storyline. It is in line with our focus on achieving a well-balanced knowledge model of the epic as opposed to any model which is constrained, pragmatically, due to detailed literary analysis. Lastly, scholarly currency of the selected reference text should be considered. Ideally the text should be in vogue amongst relevant scholarly communities. Details such as edition number and reprint number can be good markers in assessing the scholarly currency of such a text.

3.3. Conceptual Familiarization of Epics

The next requirement is for an inceptive, conceptual familiarization (see Fig. 1) of the epic. Such an exercise should involve a preliminary understanding of the storyline at the idea plane (Ranganathan, 1967). It should also involve a preliminary understanding of the activities involved in the epic in a non-rigid fashion and intellection of the implications that might arise while giving shape to its knowledge model. The primary cardinal aspect which informs the current architectural component is exploratory intuition. It skims through the chosen reference text in order to arrive at holistic, rapid insights about various entities and activities they shape and inform, as embedded within the plot of the epic. The other cardinal aspect is a reflective discussion that involves discussing, reflecting, and coalescing individual insights gained to arrive at a shared understanding. This will be the pivotal groundwork from the requirements perspective as the architecture steps into its more formal components.

The recommendation is to model the conceptual familiarization effort as exhaustive sets of Competency Questions (CQs) (Grüninger & Fox, 1995b), possibly in collaboration with potential users of varying scholarly expertise in the chosen epic.

3.4. Ontology Availability

Once a shared comprehension of the plot of the epic is attained, it is pertinent to first search for available on-

tological formalizations of the same epic broadly aligned with intermediate research needs. This should be in conformance with, but not exclusively, the ontology reuse paradigm (Carriero et al., 2020; Pinto & Martins, 2000; Simperl, 2009). The sources considered for searching the availability of such formal ontologies can be from amongst:

- Cross-domain ontology repositories which might host ontologies on epics appropriately exposed via requisite metadata annotations, and/or
- Relevant research paper(s), which might outline the skeletal framework and developmental nuances of an ontology for the selected epic, and/or
- Research content platforms, which might be websites hosting such ontologies developed within the scope of a specific research project or research grant.

Depending on the availability of the ontology (i.e., available or not available) there can be only two outcomes of this decision block. The first is the availability of ontologies (YES). This means that one or more ontological formalizations of the epic are available in conformance to our overall research requirements as understood from conceptual familiarization, i.e. as further detailed in Section 3.4.1. The second outcome is availability of ontologies (NO). This means that no formal ontological representation of the chosen epic is found; in that case we start activity as further detailed in Section 3.4.2.

3.4.1. Availability of Ontologies (YES)

In case one or more formal ontologies related to the chosen epic are available and found, the methodology prescribes the following steps sequentially:

3.4.1.1. Selection of Ontology.

This step concerns the selection of an ontology from amongst the available ontologies found, to be considered as an input to the next step. The methodology does not stipulate any convention for the selection at this stage (in case of multiple available ontologies). It might be chosen at random or as per moving criterion set collaboratively by the PDT. In the case of the availability of a single ontology relating to the chosen epic, the task is trivial. The motivation for this step and related subsequent steps is to model our ontology following the direct reuse paradigm in ontology reuse (Carriero et al., 2020).

3.4.1.2. Ontological Review.

After selecting the ontology from the previous step, this step reviews the selected ontology. The principal intuition behind this step is to check whether the selected ontology conforms to established standards in formal ontology engineering and exhibits requisite ontological commitment (Guarino & Welty, 2002). It also involves a general qualitative examination of the ontology to ascertain whether it is suitable for the application for which it is developed. There are two guiding norms for a comprehensive review, namely, ontological norms and norms of application satisfiability.

The former is the first set of guidelines. It is a mix of the OntoClean best practices (Guarino & Welty, 2002) and Ranganathan's classificatory principles (Ranganathan, 1967). Ontological norms are for reviewing sound ontological conformance. During the ontology review, firstly, we examine the classes and properties of the selected ontology as per their context, such as whether a particular axiomatization of an object property should really be such and not a data property specific to the current requirements, etc. Secondly, with respect to OntoClean, a lightweight ontological analysis can be carried out by examining the ontology elements, namely, classes, object properties, and data properties. This analysis should be from the perspective of essence, rigidity, identity, and unity as appropriate. Thirdly, and most importantly, a guided validation of the subsumption hierarchies in the ontology following Ranganathan's principles is highly recommended. Finally, the exact determination of the mix depends on the requirements of the PDT with respect to the targeted level of ontological conformance.

The latter is the second set of guidelines prescribing an assessment of the selected ontology's suitability for the application scenario for which it is modelled or enhanced. The concrete norms for this step are best left to be articulated by the PDT. Some general norms for such an activity which can be reused and improved upon are:

- Whether the considered ontological schema is amenable to be utilized as a classification ontology (Zaihrayeu et al., 2007), a descriptive ontology for metadata management applications (Satija et al., 2020), or a domain-linguistic ontology, accounting in the objectives of the application scenario;
- Whether the ontology is aligned with any standard, upper ontology in the specific instance of data integration (Giunchiglia et al., 2021a) as an application scenario;

- Whether it reuses concepts from any general-purpose, core, or domain-specific ontological schemas which are considered standard or de facto standard in the ontology engineering research community.

3.4.1.3. Satisfactory Examination.

After the review of intermediate ontology, the methodological flow now concentrates on an active assessment of its suitability to the project requirements (Gómez-Pérez, 2004). To that end, there can be three possible outcomes, which are Yes, No, and Needs Enrichment.

If the reviewed ontology satisfies the examination, i.e., case of Yes, the intermediate ontology is fully aligned with the ontological commitment as expected from the project requirements (modelled as CQs). In this case, the intermediate ontology is: (i) rechristened as the Formalized Epic Knowledge Resource (FEKR) (see 3.5), and (ii) it is production-ready, to be exploited in applications such as management of semantic annotations.

If the reviewed ontology does not satisfy the examination, i.e., case of No, the intermediate ontology fails mostly in its alignment to the ontological commitment as expected from the project requirements. In this case, there are two routes prescribed by the GOMME methodology. The first is, in case of multiple available ontologies of the epic, the methodological flow loops back to (3.4.1.1) to consider the suitability of the next ontology as decided by the PDT. The second is, in case this ontology was the single artefact considered in (3.4.1.1), the methodological flow shifts to the conceptual analysis step, which is a crucial step to design a new ontology of the chosen epic as per the expressed ontological commitments.

If the reviewed ontology partially satisfies the examination, i.e., in case of needing enrichment, the intermediate ontology is satisfactorily aligned as expected from the project requirements but still needs to be enriched in order to attain maximal coverage. The recommended modes of enrichment are external and internal.

- External enrichment is where an in-depth conceptual familiarization of the epic is collaboratively arrived at and compared with both the ontological commitment modelled in the intermediate ontology and the project requirements. Subsequently, the difference in ontological commitment in terms of classes, object properties, data properties, and other relevant axioms are modelled and integrated in the intermediate ontology.
- Internal enrichment is where possible modelling

bugs in the ontology such as mispositioning of classes, object properties and data properties, overloading of property restrictions, or *is-a* overloading are resolved.

Finally, the inclusion of external and/or internal enrichment results in an enriched ontological model which again undergoes the suitability test. The iteration continues till the ontological commitment of the enriched model maximally converges with that of the project requirements, as deemed appropriate by the PDT. Once the enriched model achieves maximal convergence, it is deemed as a FEKR which is application ready.

Most importantly, the assessment as to whether the intermediate ontology is satisfactory or not can be determined by evaluating the model against CQs (defined in Section 3), formalized as SPARQL queries (DuCharme, 2013) or even via appropriate sampling survey approaches such as stratified sampling (Parsons, 2014). The strategy is to keep enriching the ontology till it achieves maximal coverage of the project requirements. The exact threshold leading to the determination of maximal, satisfactory, or unsatisfactory alignment as from above is in essence an experiential learning issue for each PDT and is still an open question to be stated generically, if at all. We plan to widely adopt and popularise GOMME in the coming years towards iterative formal modelling of different epics and literary works, each of whose thresholds we expect to be different as of now.

3.4.2. Availability of Ontologies (NO)

Orthogonal to (3.4.1), if no formal ontology related to the chosen epic was found, or, in the case where a single ontology was found but proved unsuitable, the GOMME methodology prescribes the following sequential approach (see Fig. 1) to design an ontology from scratch for the chosen epic:

3.4.2.1. Conceptual Analysis.

The essence of this step is to build a detailed, in-depth, and formally documented shared conceptual understanding of the epic on top of the initial lightweight conceptual familiarization. We find literary justification for such in-depth analysis for knowledge modelling from the following canonical norms for epics which we derived from the genre-theoretic interpretation of epics (Greene, 1961): expansiveness, distinctions, tonality, and feeling.

Expansiveness is the literary quality of an epic to “extend its own luminosity in ever-widening circles” (Greene,

1961). While distinctions are what exist between the “director and executor of action” (Greene, 1961), in other words, not only factoring in actions but also actors and actants (Ranganathan, 1967) specific for knowledge modelling are considered. Tonality and feeling are needed to capture the “heroic energy, superabundant vitality which charges character and image and action” (Greene, 1961), and the fact remains that causation is much less important for epic modelling than tonality or feeling.

For concretely performing the conceptual analysis, the work focuses on the following two dimensions. First is elicitation via Generative Questions (GQs), which is the conceptual extraction of ontology elements, viz. classes, object properties, and data properties (subsuming actors, actants, and actions) from an in-depth and shared understanding of the epic. Evidently, this approach goes beyond a mere preliminary familiarization with its storyline. The recommendation is to record the information elicited as formal documentation which is compatible with the PDT members. It is important to note that to achieve true conceptual elicitation grounded in expansiveness, distinctions, tonality, and feeling of the epic, we exploit GQs (Vosniadou, 2002), first proposed in knowledge modelling by Bagchi (2021a). GQs reflect conceptualization rooted in human cognition (Giunchiglia & Bagchi, 2021) and cognitive theory of mental models (Johnson-Laird, 1983). These are questions directed at the incremental, dynamic understanding of knowledge domains being modelled, and thus “cannot elicit ready-made solutions” (Bagchi, 2021a). At each iteration of knowledge modelling, GQs “challenge the existent mental model of the domain and improve it” (Bagchi, 2021a). The crucial distinction between GQs with CQs (Grüninger & Fox, 1995b) is the fact that while the latter is focused on adjudicating the competency of the final knowledge model, the former is focused on the iterative conceptual mental model development of the knowledge domain being modelled, and consequently, also of the final knowledge model.

In addition to grounding the conceptual analysis in the canonical norms for epics and GQs, GOMME also recommends utilising and extending the same formal documentation by modelling ontology elements as per standard top-level ontological distinctions. This renders

the subsequent formal ontological design and commitment more concrete. For example, the conceptual entities can be characterised as endurants or perdurants, depending on whether the entity is wholly present in time or not (Gangemi et al., 2002). This becomes non-trivial, for instance, when we wish to capture the different facets of an individual in an epic via the roles (Masolo et al., 2004) they play in different sub-contexts of the epic.

3.4.2.2. Knowledge Representation Template.

Alongside the documentation of conceptual analysis, the GOMME methodology prescribes the design of a KR template (see Fig. 2). A KR template aspectually captures the conceptual analysis recorded structurally. It records the actors, actants, and actions via the properties encoded in the contextual recording of actors and actants. The design of such a template should mandatorily (but not only) include the following:

- An enumeration of all the relevant characters of the epic which the PDT wants to model and integrate with the final FEKR, cumulatively grounded in their incremental generation via GQs and inclusion via CQs from the competency perspective.
- The definition, primary, filial, or contextual identity of the characters enumerated in the above step as aptly determined by the PDT.
- The secondary relations among each such character in the epic, thus recording a major proportion of the actions on which the epic is built. A crucial observation is the fact that secondary relations among characters can be potentially many, but the primary definition is only one (all being determined in sync with the objectives of the PDT).

3.4.2.3. Ontology Construction.

Once the design and population of the KR template are completed, the methodological flow concentrates on developing the formal ontological schema from the documented conceptual analysis. The key observation here is that the ontological schema, at this level, focuses on modelling the top-level classes and properties. It postpones the population of the schema with entities from the KR tem-

Sl. No:	Characters	Primary Definitions (of each character)	Secondary Relation (among characters)	
48	Dhritarashtra	kingOf Hastinapur	sonOf Vichitravirya and Ambika	performed VaishnavaSacrifice
49	Draupadi	wifeOf Pandavas	daughterOf Drupada	
50	Drona	teacherOf Kauravas and Pandavas	killed Abhimanyu	
51	Drupada	fatherOf Draupadi	kingOf Panchala	
52	Duhsasana	brother of Duryodhana	killedBy Bhima	
53	Durdhara	brotherOf Duryodhana	killedBy Bhima	

Fig. 2. An example of a fragment of a KR template. KR, knowledge representation.

plate to the later stage to keep distinct the schema layer and data layer at the methodological level itself (Giunchiglia et al., 2021b). The ontological schema is constructed via any state-of-the-art open source ontology editor (such as Protégé) and adheres to the generic ontology development steps of designing the formal hierarchies following:

- Class subsumption hierarchy, wherein the top-level classes for the ontological schema of the epic are modelled,
- Object property hierarchy, wherein the properties interrelating the classes are modelled, with the constraint that each such object property should mandatorily have a domain and range axiomatized,
- Data property hierarchy, wherein the attributes appropriate to be captured and axiomatized for classes are modelled, each having a class as its domain and a datatype as its range.

3.4.2.4. Ontology Population.

Given the availability of the KR template and the formal ontological schema, the current step maps the entity-level data of the KR template to the formal ontological schema developed. There are three options depending upon the nature of the ontology developed and the scope within which the PDT is working, which can be leveraged to perform ontology population. The first option is the manual mapping of data from the KR template to the ontology. This is apt for epics which are limited in scope from the perspective of entities. If the epic is considerably expanded in its scope in terms of data or entities, the second option can be to exploit appropriate plugins (such as cellfie for Protégé) to map the data to their corresponding ontology concepts. This mapping can be via definition of appropriate transformation rules. In the case of an epic/literary work of extensive magnitude (big data, in popular parlance) which requires extended man-hours of work and very complex transformation rules for data mapping, state-of-the-art semi-automatic data mapping tools such as Karma (Knoblock et al., 2012) can be leveraged as the third option. The concrete output of this step is a unified knowledge model encoding both the (semantic) ontological schema and the epic entities mapped to their semantically corresponding concepts in the schema.

3.5. Formalized Epic Knowledge Resource

When the knowledge model as output (via 3.4.1.3) undergoes ontological review and subsequent to the review, there can be three possibilities: (i) the model passes the

satisfactory examination and is rechristened as FEKR (case of direct reuse), or (ii) the model undergoes ontological enrichment and is re-sent for ontological review (case of reuse via enrichment), or, (iii) in the worst case, if the designed model is majorly misaligned with the expressed ontological commitment (which we envision to be rare), the methodological flow loops back to conceptual analysis (see 3.4.2.1; see also Fig. 1). Further, as a crucial aspect of the flexibility of GOMME, the PDT can venture for further iterations of (suitable fragments of) GOMME with respect to developing a more fine-grained representational model of the FEKR, which would not only entail modelling more fine-grained CQs but will also depend on the practical requirements and updates to the project that the PDT is contributing to.

3.6. Application

The output of the GOMME methodology - the FEKR - can be exploited and extended in a variety of application scenarios in primarily, but not only, social sciences and digital humanities. Given the crucial characteristics of architectural agility, repurposing, and extensibility of the GOMME methodology, the FEKR developed for a specific epic can be taken as the starting point for developing an all-encompassing Knowledge Organization Ecosystem (Bagchi, 2021b) for narrative and literary domains (Varadarajan & Dutta, 2021a, 2021b) for intricacies specific to modelling narration which innately maps to GOMME. Further, FEKRs modelled via the GOMME methodology can be exploited in the back-end semantic infrastructure (Bagchi, 2019) of chatbots (Bagchi, 2020) built, for instance, for digital humanities research and popularisation or folk literature.

4. THE MAHABHARATA CASE STUDY - BRIEF HIGHLIGHTS

The principal focus of this paper is to introduce the first dedicated methodology - GOMME - for knowledge modelling for epics and more generally for digital humanities. Additionally, this section presents brief highlights of the project on (classificatory) modelling of the Indian epic *Mahabharata* (Rajagopalachari, 1970) and a first implementation of the methodology. A fuller implementational discussion of the Mahabharata Ontology modelled following GOMME (henceforth referred to as MO-GOMME) remains the subject of a future paper.

We first brief the inceptual phase of the GOMME methodology for modelling the *Mahabharata*, namely

the steps of initial scoping exercise, selection of the epic and reference text, conceptual familiarization of the epic, and investigation of ontology availability as from Section 3 above. The interdisciplinary PDT comprised four in-

dividuals accommodating diverse a priori competencies such as literary expertise, ontology methodology development, involvement in hands-on knowledge organization and representation projects, and expertise in qualitative

Table 2. Sample competency questions

Sl. No.	Competency questions (randomly ordered)	Questions by
1	How many sons did Arjuna have?	AP
2	What was the name of the last son of Ganga?	AP
3	Who was Abhimanyu's teacher?	AP
4	Which Rishi cursed King Pandu?	AP
5	What was the name of Arjuna during Agyatvasa (exile)?	KMP
6	Who was Gandhari?	KMP
7	Of which kingdom was Drupad the king?	KMP
8	Who wrote the <i>Mahabharata</i> ?	KMP
9	What was the name of Arjuna's Bow?	KPT
10	What was the name of Pandu's second wife?	KPT
11	Who was Shakuni?	KPT
12	What was the name of the Lord Krishna's Conch?	KPT
13	Which Yagna was organised by Yudhishtira?	KPT
14	Who was Keechak?	MP
15	Who killed Ghatotkacha?	MP
16	Who designed the Chakravyuha (circular maze)?	MP
17	Who was the second commander of Kauravas?	MP
18	Who was the first commander of the Kaurava Army?	RA
19	Who was the first commander of the Pandava Army?	RA
20	For how many days was the Mahabharata war fought?	RA
21	Who killed Duryodhana?	RA
22	Who was Kripacharya?	RA
23	In which kingdom did Pandavas spend their Agyatvasa?	RPU
24	What was the name of Krishna's Armour?	RPU
25	Who killed Keechak?	RPU
26	Who was the teacher-trainer of Karna?	RPU
27	Whose name was Sairandhri?	VKM
28	Who was the father of Karna?	VKM
29	Where was the Mahabharata war fought?	VKM
30	Who was the teacher-trainer of Kauravas and Pandavas?	VKM

and quantitative evaluation. The initial scoping exercise, over five dedicated two-hour sessions, produced two guiding design insights. Firstly, the PDT conducted a survey of major existing knowledge modelling methodologies for epic, narrative, and literary works and also examined the concrete conceptual/knowledge models produced (summarised in Section 2). The PDT converged on the highlight that there was no dedicated methodology for modelling epics and GOMME was a proposed candidate solution. Secondly, the PDT decided to adopt GOMME not only as it is targeted to model epics, but also to test the intricacies concerning the key features of architectural agility and feasibility of repurposing and extensibility that it highlights.

Regarding the selection of the epic and the reference text, the PDT, after three dedicated two-hour sessions, naturally converged on the idea to model *Mahabharata* given the pre-eminent standing of the epic in the literary and folk cultural tradition of the entire Indian subcontinent and as the baseline for countless other literary works of repute. The PDT chose the one-volume *Mahabharata* by C. R. Rajagopalachari (1970) in English as the reference text, broadly following the norms listed in step 3.2 of Section 3 above. The PDT discussed the reference text over ten dedicated two-hour sessions for the conceptual familiarization step and set up controlled interactions with three categories of participants (seven participants in total) - novice, aware, and expert - to elicit the kinds of questions they would wish the *Mahabharata* ontology modelled following GOMME to answer. The concrete

output was a manually documented list of their questions modelled as CQs after minimal manual pre-processing (only when appropriate) for abiding by scholarly norms (see Table 2). Table 2 enlists a random snapshot of thirty CQs and abbreviations of the survey participants' names, i.e., AP, KMP, KPT, MP, RA, RPU, and VKM. The abbreviations are used for anonymizing the survey practitioners' identities. The same abbreviations are used in Table 3 as well. Some of the questions asked, being descriptive and/or repeated by participants, were filtered out. Finally, the PDT, over two dedicated two-hour sessions, decided to reuse the *Mahabharata* ontology (<https://sites.google.com/site/ontoworks/ontologies>) developed by the Ontology-based Research Group of Indian Institute of Technology (IIT), Madras (henceforth referred to as MO-IITM), following the recommended grounding of the GOMME methodology in the ontology reuse paradigm.

We now briefly elucidate the ontological review and the satisfactory examination process that the PDT performed with respect to MO-IITM. The step concerning the selection of the reused ontology (step 3.4.1.1) was trivial for the discussed project as MO-IITM was the only existing ontology concerning the epic *Mahabharata*. The ontological review following our proposed ontological norms of MO-IITM was conducted by the PDT over five dedicated two-hour sessions, and the exercise yielded many crucial highlights. We only cite three general observations given the focus of the paper. The first observation is the systematic and pervasive violation of tried-and-tested classificatory principles (Ranganathan, 1967) in many sub-fragments of

Table 3. Satisfactory assessment (via CQs) for MO-IITM

Questions by	Questions	Repetition	Unique questions	Answered	% Answered	Unanswered
MP	67	1	66	31	46.9697	35
KPT	37	5	32	18	56.25	14
AP	23	1	22	8	36.3636	14
RPU	17	3	14	5	35.7143	9
RA	20	1	19	7	36.8421	12
KMP	19	2	17	8	47.0588	9
VKM	20	2	18	10	55.5556	8
Total	203	15	188	87	46.2765	101
Descriptive questions	12	0	12	0	0	12
Considered	191	15	176	87	49.4318	89

CQs, competency questions.

the ontology at different levels of abstraction, thus rendering the overall axiomatization of the model not only ontologically unwell-founded (Guizzardi et al., 2002) but also classifiatorily ungrounded (Giunchiglia & Bagchi, 2021; Satija, 2017) to a certain extent. The second observation is the fact that MO-IITM was not supported by openly available documentation of any kind such as a research paper or webpage. As a result, it is not grounded in the state-of-the-art philosophically inspired and ontologically well-founded best practices such as rigidity (Guarino & Welty, 2002). Finally, as a deviation from the two aforementioned critical observations, the PDT post-ontological review was of a unanimous opinion that MO-IITM could be clearly classified as a classification ontology (Zaihrayeu et al., 2007) given its axiomatization.

After the completion of the ontological review, the PDT, over six dedicated two-hour sessions, actively engaged to determine the satisfiability of MO-IITM and refine it, if and as necessary. The assessment of the satisfiability was a curated process in which the coverage of the ontological commitment targeted for MO-GOMME (modelled via CQs as requirements) was examined vis-à-vis the axiomatization of the ontological commitment formalized in MO-IITM. Table 3 displays the results of the satisfactory assessment. It incorporates the statistical summary of the questions asked by the three categories of survey participants: novices, aware, and experts. From all the participants a total of 203 questions were asked out of which 12 questions were descriptive and 15 were repeated. Therefore, $203 - (12 + 15) = 176$ unique questions were considered for the ontology appraisal process. Each of the questions was manually checked against MO-IITM. The results show that the ontology could answer 87 questions, i.e. 49.4% of the questions, that was considerably close to the rough threshold of 50% that the PDT assumed. The PDT decided to proceed with the reuse and enrichment of the MO-IITM.

The enrichment, in accordance with GOMME (see 3.4.1.3), was performed via two distinct modes of external enrichment and internal enrichment. External enrichment concentrated on integrating, within MO-IITM, the missing classes (for example, Role), object properties (for example, hasSpouse, hasSibling) and data properties (for example, hasSkills, gender, etc.). This would mitigate the difference between the axiomatization already encoded in the MO-IITM and the ontological commitment (as explicated by the CQs) that it should encode. The internal enrichment was performed to fix modelling bugs within MO-IITM which the PDT found to be pervasive (a de-

tailed description is omitted given the scope of the paper).

During the enrichment process, certain observations were made. Firstly, we cite the case of rigidity from OntoClean (Guarino & Welty, 2002) as also outlined in the ontological review. The MO-IITM (as also most state-of-the-art ontologies) did not ground its axiomatization in ontology best practices, and thus failed to distinguish between roles such as charioteer, which is non-rigid and spatio-temporally boxed (Masolo et al., 2004), from the objects that play such roles in a particular spatio-temporal frame. Secondly, we cite a major classification mistake in the MO-IITM (as from Ranganathan, 1967), exemplified in Giunchiglia and Bagchi (2021). The canon of modulation states, quoting Ranganathan (1967), is “...one class of each and every order that lies between the orders of the first link and the last link of the chain” (p. 176), which enforces the rule that a classification chain “shouldn’t have any missing link” (Giunchiglia & Bagchi, 2021). The PDT found missing links to be pervasive, especially in the class and object property hierarchies of the MO-IITM. A case in point is that of the inconsistent modelling of relationships in the MO-IITM, wherein father and mother were classified into the hypernym *parent*, whereas husband and wife remained as free-floating object properties without any hypernym. The PDT refined the inconsistency by modelling a novel relationship matrix (see Fig. 3) following the canon of modulation in spirit. Thirdly and unexpectedly, the PDT also found the MO-IITM to be inconsistent in another major knowledge modelling best practice, that of mandatory domain-range assertion in ob-

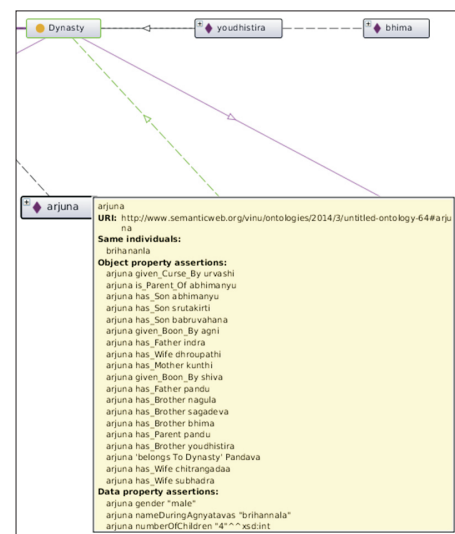


Fig. 3. An example of a relationship matrix.

ject properties, which were subsequently corrected by the PDT as appropriate.

Finally, subsequent to the enrichment process, the PDT conducted a second satisfactory examination of the now enriched MO-IITM, and consequently rechecked the outcome as from Table 3. Table 4 presents the results from the re-checking exercise. The exercise included 203 questions and after elimination of 21 repeated and 14 descriptive questions, 168 questions were considered for the assessment. The PDT found that the enriched MO-IITM could answer 149 questions in total, i.e. around 89% of the CQs, thus exhibiting a major leap from the raw

MO-IITM. Based on the t-test of the above sample, the PDT concluded that the curated enrichment is statistically significant. The t-value -5.40 and its corresponding p-value of 0.00083 meant that even at a significance level of 0.00083, the curated result is significant. Thus, the PDT decided to rechristen the enriched MO-IITM as the outcome (FEKR) of the case study and renamed it MO-GOMME. Fig. 4 presents a snapshot of MO-GOMME, the developed FEKR. It illustrates a glimpse of the classes, object properties, data properties, and instances of the enriched *Mahabharata* ontology MO-GOMME. As a further note, given the stress of the case study on practically

Table 4. Satisfactory assessment (via CQs) for MO-GOMME

Questions by	Questions	Repetition	Unique questions	Answered	% Answered	Unanswered
MP	65	5	60	48	80	12
KPT	39	3	36	29	80.5556	7
AP	24	2	22	17	77.2727	5
RPU	19	2	17	14	82.3529	3
RA	18	4	14	11	78.5714	3
KMP	17	3	14	13	92.857	1
VKM	21	2	19	17	89.4737	2
Total	203	21	182	149	81.8681	33
Descriptive questions	14	0	14	0	0	14
Considered	189	21	168	149	88.6905	19

CQs, competency questions.

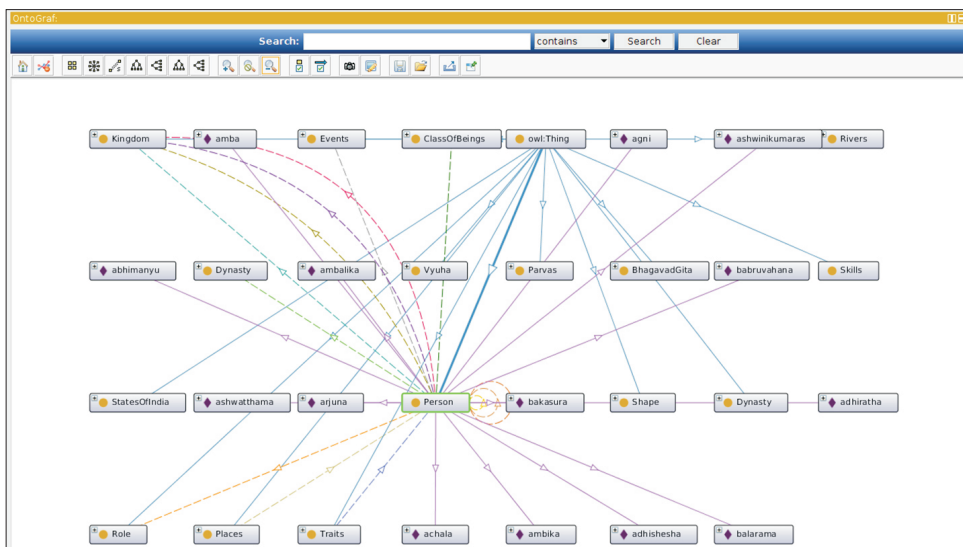


Fig. 4. Snapshot of MO-GOMME.

implementing GOMME, the PDT did not venture for further iterations of GOMME (which would have obviously required modelling more fine-grained CQs).

5. RESEARCH IMPLICATIONS

The principal research implication of GOMME is that it is the first dedicated methodology for modelling epics in particular, and literary works in general, from the perspective of digital humanities (Berry, 2012; Gold, 2012). All of the existing methodologies for modelling works in digital humanities focus on creating knowledge models such as ontologies and XML schemas from scratch. One of the crucial practical implications of this work is in linking entities in the developed FEKR to the corresponding texts of the epics. For example, we can link entities of the developed FEKR such as characters and places to open knowledge graphs such as DBpedia (Auer et al., 2007) or Wikidata (Vrandečić & Krötzsch, 2014), based on their dereferenceable availability. A case in point is the *Mahabharata* character Arjuna, who can be unambiguously linked to the Wikidata ID Q27131348. Subsequently, we can link such entities in the FEKR with the relevant text/documents/papers that mention the entities. Similar work was done for Book I of the *Odyssey* (Khan et al., 2016). The current work also contributes to information retrieval on epics of any scale, region, or religion by exploiting the concepts of the developed FEKR as elements of a base metadata scheme for epics. For example, the class Event is a metadata element that describes events which happen in an epic. Object properties such as main_Character_In is a metadata element that describes the characters involved in the event. This, we envision, is an initiation towards creating a dedicated metadata schema for annotating epics.

The methodology addresses the likes of researchers in diverse research arenas such as computer science, theology, library and information science, and literary studies. The knowledge model such as FEKR, in the backend of the question answering system and the metadata for the epics, will help in querying and retrieving the information concerning their domain. A line of research in computer science can be to develop a semi-automatic system that helps in classifying the entities of an epic and generating the FEKR. Semi-automation is not only a limitation of the current work but also extremely difficult for epics due to the lack of dedicated state-of-the-art frameworks in the corresponding arena of research in multilingual natural language processing.

6. CONCLUSION AND FUTURE WORK

The current work contributes towards a novel methodology by not only allowing the PDT to model epics/literary works from scratch but also by factoring in the possibility of reusing via enrichment of existing conceptual or ontological formalizations of epics. Apart from the mere fact that GOMME adds to the basket of existing methodologies, it is also, as far as our know-how, the only theoretically enhanced methodology grounded in transdisciplinarity. This is due to the crucial grounding of GOMME in the state-of-the-art norms of epics, ontological well-founded best practices, and tried-and-tested classificatory principles. It is also the first methodology to uniquely accommodate the flexible combination of CQs and GQs to model ontological commitments. It illustrates the feasibility and advantages of the methodology via a first brief implementation of modelling the Indian epic *Mahabharata*. Although the current version of GOMME is focused on modelling epics, in future its potential extensibility for modelling works in a wide variety of other related disciplines such as different arenas of digital humanities, literature, religion, and social sciences will be explored. We envision three immediate future streams of research which we will pursue in upcoming dedicated research papers: (i) the development of a novel theory of knowledge representation grounded in the design foundations which are implicit in GOMME, (ii) modelling different works in different arenas of digital humanities via GOMME, and (iii) potentially creating a dedicated methodology suite for modelling literary knowledge in different arenas of digital humanities and social sciences. In addition, real-world oriented solutions, such as question-answering systems dedicated to epics grounded in (semantic) data management plans (Gajbe et al., 2021), entity disambiguation, and linking in the wild for digital humanities (Frontini et al., 2015; Plirdpring & Ruangrajitpakorn, 2022), are much valued in the domain of digital humanities. GOMME, via its flexible knowledge management infrastructure, can partially (as of now) facilitate their realization.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the late Mr. Keshav Prasad Tiwari and Mr. Mahendra Pandey for their efforts in CQs formulation. Additionally, we wish to thank the editor, copy-editor, and anonymous reviewers for their constructive comments and suggestions.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

- Abello, J., Broadwell, P., & Tangherlini, T. R. (2012). Computational folkloristics. *Communications of the ACM*, 55(7), 60-70. <https://doi.org/10.1145/2209249.2209267>
- Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2007, November 11-15). DBpedia: A nucleus for a web of open data. In K. Aberer, K. S. Choi, N. Noy, D. Allemang, K. I. Lee, L. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, & P. Cudré-Mauroux (Eds.), *The Semantic Web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference* (pp. 722-735). Springer.
- Bagchi, M. (2019). *A knowledge architecture using knowledge graphs*. (Master's thesis). Indian Statistical Institute, Bangalore, India.
- Bagchi, M. (2020). Conceptualising a library chatbot using Open source conversational artificial intelligence. *DESID-OC Journal of Library & Information Technology*, 40(6), 329-333. <https://doi.org/10.14429/djlit.40.6.15611>
- Bagchi, M. (2021a). A large scale, knowledge intensive domain development methodology. *Knowledge Organization*, 48(1), 8-23. <https://doi.org/10.5771/0943-7444-2021-1-8>
- Bagchi, M. (2021b). Towards knowledge organization ecosystem (KOE). *Cataloging & Classification Quarterly*, 59(8), 740-756. <https://doi.org/10.1080/01639374.2021.1998282>
- Bagchi, M., & Madalli, D. P. (2019, February 26-28). Domain visualization using knowledge cartography in the big data era: A knowledge graph based alternative. In K. R. Patnaik, & K. Satpathi (Eds.), *International Conference on "Future of Libraries"* (pp. 1-22). Indian Institute of Management.
- Berry, D. M. (2012). Introduction: Understanding the digital humanities. In D. M. Berry (Ed.), *Understanding digital humanities* (pp. 1-20). Palgrave Macmillan.
- Branch, F., Arias, T., Kennah, J., Phillips, R., Windleharth, T., & Lee, J. H. (2017). Representing transmedia fictional worlds through ontology. *Journal of the Association for Information Science and Technology*, 68(12), 2771-2782. <https://doi.org/10.1002/asi.23886>
- Carriero, V. A., Daquino, M., Gangemi, A., Nuzzolese, A. G., Peroni, S., Presutti, V., & Tomasi, F. (2020). The landscape of ontology reuse approaches. In G. Cota, M. Daquino, & G. L. Pozzato (Eds.), *Applications and practices in ontology design, extraction, and reasoning* (pp. 21-38). IOS Press.
- Cooren, F. (2015). In medias res: Communication, existence, and materiality. *Communication Research and Practice*, 1(4), 307-321. <https://doi.org/10.1080/22041451.2015.1111075>
- Damiano, R., & Lieto, A. (2013). Ontological representations of narratives: A case study on stories and actions. In M. A. Finlayson, B. Fisseni, B. Löwe, & J. C. Meister (Eds.), *2013 Workshop on Computational Models of Narrative* (pp. 76-93). Schloss Dagstuhl.
- Damiano, R., Lombardo, V., & Pizzo, A. (2019). The ontology of drama. *Applied Ontology*, 14(1), 79-118. <https://doi.org/10.3233/AO-190204>
- Das, S., & Roy, S. (2016). Faceted ontological model for brain tumour study. *Knowledge Organization*, 43(1), 3-12. <https://doi.org/10.5771/0943-7444-2016-1-3>
- DuCharme, B. (2013). *Learning SPARQL: Querying and updating with SPARQL 1.1*. O'Reilly.
- Dutta, B., Chatterjee, U., & Madalli, D. P. (2015). YAMO: Yet another methodology for large-scale faceted ontology construction. *Journal of Knowledge Management*, 19(1), 6-24. <https://doi.org/10.1108/JKM-10-2014-0439>
- Ehrlinger, L., & Wöß, W. (2016, September 12-15). Towards a definition of knowledge graphs. In M. Martin, M. Cuquet, & E. Folmer (Eds.), *SEMANTICS Leipzig 2016*. Committee of SEMANTICS 2016.
- Espinoza-Arias, P., Poveda-Villalón, M., García-Castro, R., & Corcho, O. (2019). Ontological representation of smart city data: From devices to cities. *Applied Sciences*, 9(1), 32. <https://doi.org/10.3390/app9010032>
- Fernández-López, M., Gómez-Pérez, A., & Juristo, N. (1997, March 24-26). Methontology: From ontological art towards ontological engineering. *Proceedings of the Ontological Engineering AAAI-97 Spring Symposium series* (pp. 33-40). AAAI.
- Frontini, F., Brando, C., & Ganascia, J. G. (2015, June 1). Semantic web based named entity linking for digital humanities and heritage texts. In A. Zucker, I. Draelants, C. F. Zucker, & A. Monnin (Eds.), *Proceedings of the First International Workshop Semantic Web for Scientific Heritage at the 12th ESWC 2015 Conference* (pp. 77-88). HAL Open Science.
- Gajbe, S. B., Tiwari, A., Gopalji, & Singh, R. K. (2021). Evaluation and analysis of data management plan tools: A parametric approach. *Information Processing & Management*, 58(3), 102480. <https://doi.org/10.1016/j.ipm.2020.102480>
- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., & Schneider, L. (2002, October 1-4). Sweetening ontologies with DOLCE. In A. Gómez-Pérez, & V. R. Benjamins (Eds.), *EKAW 2002: Knowledge engineering and knowledge management: ontologies and the semantic web* (pp. 166-181).

- Springer.
- Giunchiglia, F., & Bagchi, M. (2021, September 11-18). Millikan + Ranganathan – From perception to classification. In E. M. Sanfilippo, O. Kutz, N. Troquard, T. Hahmann, C. Masolo, R. Hoehndorf, & R. Vita (Eds.), *CAOS 2021: 5th Workshop on Cognition And Ontologies*. Springer.
- Giunchiglia, F., Bocca, S., Fumagalli, M., Bagchi, M., & Zamboni, A. (2021a). iTelos -- purpose driven knowledge graph generation. *arXiv*, arXiv:2105.09418. <https://doi.org/10.48550/arXiv.2105.09418>
- Giunchiglia, F., Dutta, B., & Maltese, V. (2014). From knowledge organization to knowledge representation. *Knowledge Organization*, 41(1), 44-56. <https://doi.org/10.5771/0943-7444-2014-1-44>
- Giunchiglia, F., Zamboni, A., Bagchi, M., & Bocca, S. (2021b, June 6). Stratified data integration. In D. Chaves-Fraga, A. Dimou, P. Heyvaert, F. Priyatna, & J. Sequeda (Eds.), *Knowledge Graph Construction 2021*. CEUR.
- Gold, M. K. (2012). *Debates in the digital humanities*. University of Minnesota Press.
- Gómez-Pérez, A. (2004). Ontology evaluation. In S. Staab, & R. Studer (Eds.), *Handbook on ontologies* (pp. 251-273). Springer.
- Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological Engineering: With examples from the areas of knowledge management, e-commerce and the semantic web*. Springer.
- Greene, T. (1961). The norms of epic. *Comparative Literature*, 13(3), 193-207. <https://doi.org/10.2307/1768997>
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, 43(5-6), 907-928. <https://doi.org/10.1006/ijhc.1995.1081>
- Grüninger, M., & Fox, M. S. (1995a). *Methodology for the design and evaluation of ontologies*. Paper presented at IJ-CAI'95, Montreal, Canada.
- Grüninger, M., & Fox, M. S. (1995b). The role of competency questions in enterprise engineering. In A. Rolstadås (Ed.), *Benchmarking — Theory and practice* (pp. 22-31). Springer.
- Guarino, N., & Welty, C. (2002). Evaluating ontological decisions with Ontoclean. *Communications of the ACM*, 45(2), 61-65. <https://doi.org/10.1145/503124.503150>
- Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology? In S. Staab, & R. Studer (Eds.), *Handbook on ontologies* (pp. 1-17). Springer.
- Guizzardi, G., Herre, H., & Wagner, G. (2002, October 7-11). On the general ontological foundations of conceptual modeling. In S. Spaccapietra, S. T. March, & Y. Kambayashi (Eds.), *Conceptual modeling — ER 2002* (pp. 65-78). Springer.
- Hyyönen, E., Valo, A., Komulainen, V., Seppälä, K., Kauppinen, T., Ruotsalo, T., Salminen, M., & Ylisalmi, A. (2005). Finnish national ontologies for the semantic web -- Towards a content and service infrastructure. *International Conference on Dublin Core and Metadata Applications* (pp. 219-222). Dublin Core.
- Iqbal, R., Mustapha, A., & Yusoff, Z. M. (2013). An experience of developing Quran ontology with contextual information support. *Multicultural Education & Technology Journal*, 7(4), 333-343. <https://doi.org/10.1108/METJ-03-2013-0009>
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.
- Kent, R. E. (2000, July 10-13). The information flow foundation for conceptual knowledge organization. In C. Beghtol, L. Howarth, & N. J. Williamson (Eds.), *Dynamism and stability in knowledge organization: Proceedings of the Sixth International ISKO Conference* (pp. 111-117). International Society for Knowledge Organization.
- Khan, A. F., Bellandi, A., Benotto, G., Frontini, F., Giovannetti, E., & Reboul, M. (2016, September 18-20). Leveraging a narrative ontology to query a literary text. In B. Miller, A. Lieto, R. Ronfard, S. G. Ware, & M. A. Finlayson (Eds.), *The 7th Workshop on Computational Models of Narrative (CMN 2016)* (pp. 10:1-10:10). Schloss Dagstuhl.
- Klarer, M. (2013). *An introduction to literary studies*. (3rd ed.). Routledge.
- Knoblock, C. A., Szekely, P., Ambite, J. L., Goel, A., Gupta, S., Lerman, K., Muslea, M., Taheriyani, M., & Mallick, P. (2012, May 27-31). Semi-automatically mapping structured sources into the semantic web. In E. Simperl, P. Cimiano, A. Polleres, O. Corcho, & V. Presutti (Eds.), *ESWC 2012: The semantic web: research and applications* (pp. 375-390). Springer.
- Lassila, O., & McGuinness, D. (2001). *The role of frame-based representation on the semantic web*. Linköping University Electronic Press.
- Lehmann, F. (1992). Semantic networks. *Computers & Mathematics with Applications*, 23(2-5), 1-50. [https://doi.org/10.1016/0898-1221\(92\)90135-5](https://doi.org/10.1016/0898-1221(92)90135-5)
- Lochtefeld, J. G. (2002). *The illustrated encyclopedia of hinduism*. Rosen Publishing.
- Mahesh, K., Helmreich, S., & Wilson, L. (1996). *Ontology development for machine translation: Ideology and methodology*. New Mexico State University.
- Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., & Guarino, N. (2004, June 2-5). Social roles

- and their descriptions. In D. Dubois, C. Welty, & M. A. Williams (Eds.), *Proceedings of the 9th International Conference on the Principles of Knowledge Representation and Reasoning* (pp. 267-277). AAAI Press.
- Mulholland, P., Collins, T., & Zdrahal, Z. (2004, January 13-16). Story fountain: Intelligent support for story research and exploration. In J. Vanderdonck (Ed.), *Proceedings of the 9th International Conference on Intelligent User Interfaces* (pp. 62-69). ACM.
- Nakasone, A., & Ishizuka, M. (2006, December 18-22). Storytelling ontology model using RST. In T. Nishida (Ed.), *2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology* (pp. 163-169). IEEE.
- Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Stanford University.
- Parsons, V. (2014). Stratified sampling. In N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, & J. Teugels (Eds.), *Wiley StatsRef: Statistics reference online*. John Wiley & Sons.
- Phoksawat, K., Mahmuddin, M., & Ta'a, A. (2019). Intercropping in rubber plantation ontology for a decision support system. *Journal of Information Science Theory and Practice*, 7(4), 56-64. <https://doi.org/10.1633/JISTaP.2019.7.4.5>
- Pinto, H. S., & Martins, J. P. (2000, March 20-22). Reusing ontologies. *AAAI 2000 Spring Symposium on Bringing Knowledge to Business Processes* (pp. 77-84). AAAI.
- Plirdpring, P., & Ruangrajitpakorn, T. (2022). Using ontology to represent cultural aspects of local products for supporting local community enterprise in Thailand. *Journal of Information Science Theory and Practice*, 10(1), 45-58. <https://doi.org/10.1633/JISTaP.2022.10.1.4>
- Rajagopalachari, C. (1970). *Mahabharata*. Diamond Pocket Books.
- Ranganathan, S. R. (1967). *Prolegomena to library classification*. (3rd ed.). Ess Ess Publications.
- Satija, M. P. (2017). Colon Classification (CC). *Knowledge Organization*, 44(4), 291-307. <https://doi.org/10.5771/0943-7444-2017-4-291>
- Satija, M. P., Bagchi, M., & Martínez-Ávila, D. (2020). Metadata management and application. *Library Herald*, 58(4), 84-107. <https://doi.org/10.5958/0976-2469.2020.00030.2>
- Simperl, E. (2009). Reusing ontologies on the semantic web: A feasibility study. *Data & Knowledge Engineering*, 68(10), 905-925. <https://doi.org/10.1016/j.datak.2009.02.002>
- Suárez-Figueroa, M. C., Gómez-Pérez, A., & Fernández-López, M. (2015). The NeOn methodology framework: A scenario-based methodology for ontology development. *Applied Ontology*, 10(2), 107-145. <https://doi.org/10.3233/AO-150145>
- Swartjes, I., & Theune, M. (2006). A fabula model for emergent narrative. In S. Göbel, R. Malkewitz, & I. Iurgel (Eds.), *Technologies for interactive digital storytelling and entertainment* (pp. 49-60). Springer.
- Syamili, C., & Rekha, R. V. (2018). Developing an ontology for Greek mythology. *Electronic Library*, 36(1), 119-132. <https://doi.org/10.1108/EL-02-2017-0030>
- Tiwari, A., & Madalli, D. P. (2021). Maturity models in LIS study and practice. *Library & Information Science Research*, 43(1), 101069. <https://doi.org/10.1016/j.lisr.2020.101069>
- Uschold, M., & King, M. (1995). *Towards a methodology for building ontologies*. University of Edinburgh.
- Varadarajan, U., & Dutta, B. (2021a, September 18-20). Towards development of knowledge graph for narrative information in medicine. In B. Villazón-Terrazas, F. Ortiz-Rodríguez, S. Tiwari, A. Goyal, & M. Jabbar (Eds.), *KG-SWC 2021: Knowledge graphs and semantic web* (pp. 290-307). Springer.
- Varadarajan, U., & Dutta, B. (2021b). Models for narrative information: A study. *arXiv*, arXiv:2110.02084. <https://doi.org/10.48550/arXiv.2110.02084>
- Vosniadou, S. (2002). Mental models in conceptual development. In L. Magnani, & N. J. Nersessian (Eds.), *Model-based reasoning: Science, technology, values* (pp. 353-368). Springer.
- Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledge base. *Communications of the ACM*, 57(10), 78-85. <https://doi.org/10.1145/2629489>
- Vrandečić, D., Pinto, S., Tempich, C., & Sure, Y. (2005). The DILIGENT knowledge processes. *Journal of Knowledge Management*, 9(5), 85-96. <https://doi.org/10.1108/13673270510622474>
- Yeh, J. H. (2017, May 13-17). StoryTeller: An event-based story ontology composition system for biographical history. In T. H. Meen, S. D. Prior, & A. D. K. T. Lam (Eds.), *Proceeding of the 2017 International Conference on Applied System Innovation (ICASI)* (pp. 1934-1937). IEEE.
- Zaihrayeu, I., Sun, L., Giunchiglia, F., Pan, W., Ju, Q., Chi, M., & Huang, X. (2007, November 11-15). From web directories to ontologies: Natural language processing challenges. In K. Aberer, K. S. Choi, N. Noy, D. Allemang, K. I. Lee, L. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, & P. Cudré-Mauroux (Eds.), *The semantic web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference* (pp. 623-636). Springer.