

Porting past Classification Schemes for Narratives to a Linked Data Framework

Full Paper

Thierry Declerck
DFKI GmbH – Multilingual Technologies
3 Stuhlsatzenhausweg
Saarbrücken 66123, Germany
declerck@dfki.de

Lisa Schäfer
Saarland University – Computational Linguistics
P.O. Box 15 11 50
Saarbrücken 66041, Germany
lisas@coli.uni-saarland.de

ABSTRACT

In this paper we give an overview on a number of achieved and on-going efforts dealing with porting to the Linked Data framework electronic versions of past classification schemes in the field of folktale narratives. Three of those schemes are in the field of folktales, including: (1) the work by Vladimir Propp on the Morphology of the Folktale, (2) the Stith Thompson's Motif-Index of Folk-Literature, and (3) the Aarne-Thompson-Uther classification system of types of international folktales. We are recently also considering the work by Georges Polti on categorizing the dramatic situations ("The thirty-six Dramatic Situations"), enlarging thus our focus on folktales to other literary genres.

Our aim is primarily to make those past schemes available in a formal representation system – implemented as a set of integrated OWL ontologies – in order to transform them in machine-readable data sets. This way, we are additionally supporting a cross-linking of distinct elements included in those influential classification schemes, and we are also able to populate the resulted integrated ontology with (elements of) stories.

CCS CONCEPTS

- **Information systems** → **Semantic web description languages**;
- **Applied computing** → **Arts and humanities**;

KEYWORDS

Classification Schemes for Folktales, Ontologies, Linked Data

ACM Reference format:

Thierry Declerck and Lisa Schäfer. 2017. Porting past Classification Schemes for Narratives to a Linked Data Framework. In *Proceedings of DATeCH2017, Göttingen, Germany, June 01-02, 2017*, 5 pages.
<https://doi.org/http://dx.doi.org/10.1145/3078081.3078105>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

DATeCH2017, June 01-02, 2017, Göttingen, Germany

© 2017 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

ACM ISBN 978-1-4503-5265-9/17/06...\$15.00

<https://doi.org/http://dx.doi.org/10.1145/3078081.3078105>

1 INTRODUCTION

In this paper we summarize past and on-going work, which is building on the top of different software projects and bachelor or master theses conducted at Saarland University and at DFKI. The goal of those efforts was to develop annotation schemes that support efficient access to topics of interest in folktales for their inclusion in applications. The resulting different layers of annotation for folktales have been mediated by a formal representation of the annotation elements in an ontological framework, using for this the OWL¹ W3C standard.

Following this line of work on annotations mediated by ontological models we started to consider a related formalization of past and influential classification schemes in the wider field of narratives. Considering first past classification systems for folktales, we started by formalizing the work by Vladimir Propp on the Morphology of the Folktale [5], before moving to a full ontological model of extended classification systems that were existing in a format very close to taxonomies, as the Stith Thompson's Motif-Index of Folk-Literature [7] and the Aarne-Thompson-Uther classification system of types of international folktales [8].

More recently we started to investigate the porting to an ontology of the "The thirty-six Dramatic Situations" suggested by Georges Polti [4] on categorizing dramatic situations, enlarging thus our focus beyond the folktale literary genre.

The remainder of this paper is concerned with first a brief description on our past work dealing with annotation schemes and showing which kind of application can result from our ontology mediation of annotation layers. We describe then the ontologization work of the systems of Stith Thompson and Aarne-Thomson-Uther. Finally we present in some details our on-going work with the ontologization of Solti's classification of dramatic situations.

2 ONTOLOGY-DRIVEN MEDIATION OF ANNOTATIONS FOR FOLKTALES

A first approach dealt with the design of a fine-grained annotation schema for representing the so-called functions and characters of the folktales introduced by Vladimir Propp [5]. The resulting annotation scheme is presented and discussed in [6]. Building on this work, an automated linguistic analysis of tales was developed. The goal was not only to automatically detect characters of the tales, but also to provide for a co-reference analysis such that the actions in which the characters are involved can be fully specified,

¹OWL stands for "Web Ontology Language", see <https://www.w3.org/OWL/>.

and thus helping for an automated detection of Proppian functions, together with the involved personages. Results of the analysis are stored in a database, which has been further developed onto an ontological framework [3], leading to our first formal representation of a classification scheme. The ontological representation allows also to apply generalizations for the specification of the characters (human vs animals, or supra natural etc.).

The decision to use an ontological framework for encoding our annotation schemes turned out to be very useful, since further descriptions of distinct elements of a tales could be easily integrated. So for example the work described in [2] considered the detection of sentiments expressed by the characters of the tales. Such sentiments (“joy”, “happiness”, “sadness” etc.) could be added in a straightforward manner to instances of characters within the time span in which they occur in the tales.

But more central to the work described in [2] is the recognition and segmentation of dialogue structures in the tales; this step is essential in order to know who is “speaking” to whom, as an anchor for building a Text-to-Speech (TTS) system applied to a folktale. This concerns also the detection of the text passages in which a narrator is “telling the story”. The work summarized in [2] mainly addresses the issue of adding such a TTS functionality to the automatic analysis of the text, as provided by the work described in [3]. The TTS system accesses the instances of the characters in the populated ontology, and can retrieve the information on sentiment encoded there and correspondingly model the voice output of the various characters. Details on annotations and ontologies developed in the context of the work described in this section, as well as the code and an example of the output of the used Text-to-Speech system are available in a Bitbucket repository².

3 ONTOLOGIZATION OF PAST CLASSIFICATION SCHEMES

Central resources for the classification of folktales have been made (partially) available on the World Wide Web. We are considering in this submission two such resources: The “Motif-index of folk-literature” [7] and the “Types of International Folktales” [8]. The first resource, which we abbreviate in this submission as TMI, is available as an on-line resource³. The second resource builds on former work by Antti Aarne [1] and Stith Thompson. This classification system was extended by Hans-Jörg Uther (see [8]), and in the following we are using the acronym ATU for referring to this resource. Recently a large amount of the ATU data has been made available on-line and offers edition and annotation facilities for tales in multilingual versions⁴.

Our work consisted in extracting from those on-line resources, which are stored in different formats, classification relevant information and to re-organize them in two interrelated ontologies, using for this the W3C standards OWL⁵, RDF(s)⁶ and RDF⁷. The aim was to make those classification resources interoperable and to support additionally by this formal representation of the metadata

²<https://bitbucket.org/ceisen/apftml2repo>

³https://sites.ualberta.ca/~urban/Projects/English/Motif_Index.htm.

⁴<http://www.mftd.org/>

⁵See <http://www.w3.org/TR/owl-semantics/>.

⁶See <http://www.w3.org/TR/rdf-schema/formoredetails>.

⁷See <https://www.w3.org/RDF/> for more details.

Table 1: A few motifs from Motif-index of folk-literature

Motif-id	Motif-name
A	Mythological motifs
A1	Identity of creator
A1.1	Sun-god as creator
A1.2	Grandfather as creator
A1.3	Stone-woman as creator
A1.4	Brahma as creator
A2	Multiple creators
A2.

access to folktales annotated by those classification systems in the context of the Linked Open Data framework⁸.

3.1 Thompson’s Motif Index (TMI)

A folktale motif can be defined as a “repeated story element, e.g., a character, an object, an action, or an event that can be found in several stories”⁹. In TMI all motifs are organized in a tree structure, so that each motif has a more abstract class that describes a span of subordinated motifs. One motif entry consists of a motif-id, motif name, motif description (optional), and references to literature where it occurs. We focused for the time being only on motif-ids and names. Table 1 provides for an example of few motifs illustrating the tree structure and hierarchy of TMI.

In the following example, the reader can see the formal representation of a TMI element, “K1010: Deception through false doctoring”:

```
:K1010
  rdf:type :K ;
  rdf:type :Motif ;
  rdf:type owl:Class ;
  :linkFromTMItoATU
    <http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-atu-ontology#1136> ;
  rdfs:comment "\"Index K1010 of TMI\""@en ;
  rdfs:label "\"Deception through false doctoring.\""@en ;
  rdfs:subClassOf :K ;
```

The reader can see in this example how the linking to the ATU data is realized by the property “:linkFromTMItoATU”. This way we can easily collect all the motifs that can be part of a folktale type. Our TMI ontology contains 46,950 motifs, also organized in sub-class hierarchy, as can also be seen in Figure 1.

Recently a cooperation was started with the project “Digital Breadcrumbs of Brothers Grimm”¹⁰, which is among others dealing with a fine-grained manual annotation of motifs in versions of tales, also across languages and cultures, that are marked as belonging to the same folktale type. It was straightforward to integrate the first results of this excellent annotation initiative in our ontological model, also adding new motifs suggested by the annotators.

⁸<http://linkeddata.org/>

⁹[https://en.wikipedia.org/wiki/Motif_\(folkloristics\)](https://en.wikipedia.org/wiki/Motif_(folkloristics))

¹⁰See <http://www.etrapp.eu/digital-breadcrumbs-of-brothers-grimm/> for more details.

3.2 Aarne-Thompson-Uther Folktales Types (ATU)

A folktale type can be described as a main story line that can be found in several cultures. The parts of this story line can refer to specific story elements also known as motifs. A folktale type is therefore a bigger unit than a motif. As can be seen in example 1 just below, each entry in the ATU classification consists of a type id (“6”), its label (“Animal Captor Persuaded to Talk”) and a text summarizing the typical events occurring in this type of folktale. At the end of this “script”, a link to a corresponding Thompson’s Motif-Index is provided (“[K561.1]”). Finally (and optionally), types can be indicated, with which the current type is usually combined.

(Example 1):

```
6~Animal Captor Persuaded to Talk.~
A fox (jackal, wolf) catches a chicken
(crow, bird, hyena, sheep, etc.)
and is about to eat it.
The weak animal asks a question and
the fox answers.
Thus he releases the prey and it escapes.
~K561.1
```

An example of our encoding of a higher level type of ATU is given just below in example 2:

(Example 2):

```
<http://www.semanticweb.org/tonka/ontologies/
2015/5/tmi-atu-ontology#ANIMAL_TALES_(1-299)>
rdf:type owl:Class ;
rdfs:comment "ANIMAL_TALES_(1-299)"@en ;
rdfs:label "ANIMAL TALES"@en ;
rdfs:label "TIERMAERCHEN"@de ;
rdfs:subClassOf :ATU ;
```

Here we model the ATU section on “Animal Tales”, which is ranging from the Type number “1” to the number “299”. In this example, we can also see how multilingualism is addressed, marking with “@en” or “@de” the language of the labels that can be associated with the folktale type. Example 3 is then showing a further sub-classification of the Animal Tales: “Wild Animals”.

(Example 3):

```
<http://www.semanticweb.org/tonka/ontologies/
2015/5/tmi-atu-ontology#Wild_Animals_1-99>
rdf:type owl:Class ;
rdfs:comment " Wild Animals 1-99" ;
rdfs:label "Wild Animals"@en ;
rdfs:label "Wildtiere"@de ;
rdfs:subClassOf <http://www.semanticweb.org/
tonka/ontologies/2015/5/tmi-atu-ontology#
ANIMAL_TALES_(1-299)> ;
```

Example 4 displays then our representation of what we can consider a specific type, and not a range of types, as this is the case for “Animal Tales” etc. In example 4 we include also the text of ATU describing the typical content of such a tale type, within the property “rdfs:isDefinedBy”. We can see that within this text, optional characters are included “fox (jackal, wolf)”. This is an important feature that can ease the automated classification of folktales with respect to a type number. And last, we introduce the property “:linkToTMI”

as the inverse property of “:linkFromTMIToATU” for establishing the two ways relation between TMI and ATU.

(Example 4):

```
<http://www.semanticweb.org/tonka/ontologies/
2015/5/tmi-atu-ontology#6>
rdf:type :Type ;
rdf:type owl:Class ;
rdf:type owl:NamedIndividual ;
:linkToTMI <http://www.semanticweb.org/tonka/
ontologies/2015/5/tmi-atu-ontology#K561.1> ;
rdfs:comment "\"Type 6 of ATU\""@en ;
rdfs:isDefinedBy "Animal Captor Persuaded to Talk.
A fox (jackal, wolf) catches a chicken (crow, bird,
hyena, sheep, etc. ) and is about to eat it. The weak
animal asks a question and the fox answers. Thus he
releases the prey and it escapes . Cf. Types 20D*,
122A, 122C, 122B*, 227, and 227*."@en ;
rdfs:label "\"Animal Captor Persuaded to Talk\""@en ;
rdfs:subClassOf <http://www.semanticweb.org/tonka/
ontologie/2015/5/tmi-atu-ontology#
The_Clever_Fox_(Other_Animal)_1-69> ;
```

Our ontology contains 2802 elements for the ATU domain. We added also a class called “AaTh” for encoding the elements of the former version of ATU that are not included anymore in ATU, but the latter is containing references to those obsolete or updated types. We have a linking mechanisms for expression the relation between AaTh types and ATU types.

3.3 Generation of the Ontology

The OWL, RDF(s) and RDF representation for the ontology was generated automatically from both the HTML code of TMI and ATU, responding to few design decisions we had to take. We kept the hierarchical organization of both TMI and ATU but we also went for a double representation: the hierarchy structure of the IDs is represented as an OWL subclass hierarchy, but all terminal nodes (leaves of the tree) are represented as both an instance of a class we call “Motif” (for TMI) and “Type” (for ATU). This reflects our intuition that what Thompson called a motif or ATU a type is in most of the cases the content of the terminal nodes of the classification systems, while the non-terminal nodes are more to be considered as abstraction helping in the taxonomic structures. Figures 1 and 2 are displaying screen shots taken from the graphical interface we are using for visualizing the ontologies, in this case the TopBraid composer¹¹

4 TOWARDS AN ONTOLOGY FOR THE “THIRTY-SIX DRAMATIC SITUATIONS”

“The Thirty-Six Dramatic Situations is a descriptive list which was created by Georges Polti to categorize every dramatic situation that might occur in a story or performance.”¹². Examples of situations are “1. Supplication”, “2. Deliverance”, or “3. Crime pursued by vengeance”. Related to those situations are the central “elements” participating to them. In the same order as above we have the

¹¹<http://www.topquadrant.com/tools/IDE-topbraid-composer-maestro-edition/>.

¹²https://en.wikipedia.org/wiki/The_Thirty-Six_Dramatic_Situations. The 36 situations and their main elements are also listed in this Wikipedia page, but there are numerous other Webpages presenting the list of situations.

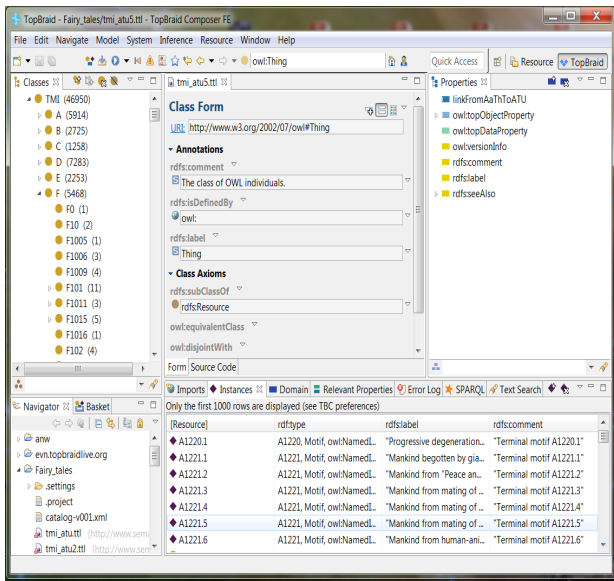


Figure 1: A screen shot of the graphical interface to the TMI model.

- A
 - fugitives imploring the powerful for help against their enemies
 - assistance implored for the performance of a pious duty which has been forbidden
 - appeals for a refuge in which to die
- B
 - hospitality besought by the shipwrecked
 - charity entreated by those cast off by their own people, whom they have disgraced
 - expiation, the seeking of pardon, healing or deliverance
 - the surrender of a corpse, or of a relic, solicited
- C
 - supplication of the powerful for those dear to the suppliant
 - supplication to a relative in behalf of another relative
 - supplication to a mother’s lover, in her behalf

On the basis of this type of data, we designed a first version of an ontological model, which we present below, taking as an example the “Supplication” dramatic situation. The first code of our ontology we are presenting is introducing the class “Supplication”.

```
<http://tutorial-topbraidd.com/situationsd#1_Supplication>
  rdf:type owl:Class ;
  rdfs:comment "1. SUPPLICATION (To humbly petition).
  Elements: a persecutor, a humble petitioner, and a
  power in authority whose decision is doubtful" ;
  rdfs:comment "The Suppliant is chased, harmed or
  otherwise threatened by the Persecutor and begs
  for help from the Power in Authority." ;
  rdfs:label "Supplication"@en ;
  rdfs:subClassOf :Situations ;
  .
```

In the second example below, we introduce a sub-class we call “1_A”, modeling thus this way the first variant of the Supplication situation.

```
<http://tutorial-topbraidd.com/situationsd#1_A>
  rdf:type owl:Class ;
  rdfs:comment "Unlabeled_subclass_of_1.
  Adding the Letter A" ;
  rdfs:label "1_A"@en ;
  rdfs:subClassOf <http://tutorial-topbraidd.com/
  situationsd#1_Supplication> ;
  .
```

This is now th place to introduce an instance, the first one for the the class “1_A”. In this instance, we list also the participating elements.

```
<http://tutorial-topbraidd.com/situationsd#1_A_1>
  rdf:type <http://tutorial-topbraidd.com/
  situationsd#1_A> ;
  :hasSituationElements
  :Power_in_Authority_whose_decision_is_doubtful ;
  :hasSituationElements :enemy ;
  :hasSituationElements :fugitives ;
  rdfs:comment "Instance of 1_A" ;
  rdfs:label "Fugitives imploring the powerful
  for help against their enemies."@en ;
  .
```

The next instance is similar, only that is an instance of the second variant of the Supplication situation.

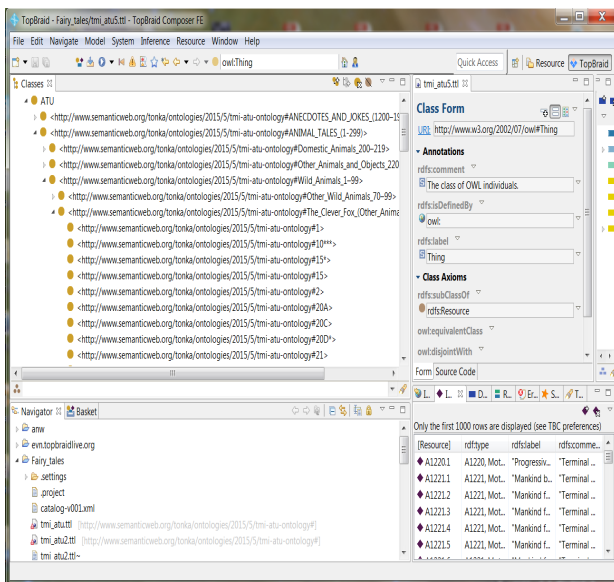


Figure 2: A screen shot of the graphical interface to the ATU model.

following elements “a persecutor; a suppliant; a power in authority, whose decision is doubtful,” “an unfortunate; a threatener; a rescuer” and “a criminal; an avenger”. To each situation a small description of the “plot” is given. So for for Supplication “The suppliant appeals to the power in authority for deliverance from the persecutor.”, etc. Additionally there are kinds of “sub-plots” or variants. Again focusing on the first situation, Supplication, we have the following variants:

```
<http://tutorial-topbraid.com/situationsd#1_B_1>
  rdf:type <http://tutorial-topbraid.com/
situationsd#1_B> ;
  :hasSituationElements :shipwrecked ;
  rdfs:comment "instance of class 1_B" ;
  rdfs:label "Hospitality besought by the
shipwrecked."@en ;
.
```

The elements of the thirsty-six situations are listed in a separate class hierarchy. Below we have the case of the Persecutor.

```
:Persecutor
  rdf:type owl:Class ;
  rdf:type sh:Shape ;
  rdfs:comment "A persecutor" ;
  rdfs:label "Persecutor"@en ;
  rdfs:subClassOf :Elements ;
.
```

As the Persecutor can appear in different instantiations, we introduce corresponding instances in our ontology, like “enemy” (see below). Instances of situations can then point to those elements.

```
:enemy
  rdf:type :Persecutor ;
  rdfs:comment "Enemies as an instance of element
persecutor in situation 1_A_1" ;
  rdfs:label "ennemy"@en ;
.
```

Having now all those situations and elements of such in an ontology framework, we are starting to investigate on how to link those objects to classes and instances we have in the TMI and ATU ontologies

5 CONCLUSION AND FUTURE WORK

We summarized first our past work on the ontology mediated integration of annotation for folktales. As the use of ontologies for our work with folktales was getting increasingly relevant, we started also with the ontology generation of two widely used folktale classification systems, while offering a descriptive way for interlinking those resources. This ontology can be visualized and processed by standard OWL tools such as protégé¹³ or TopBraid (see Figures 1 and 2). The integrated ontology will be made publicly available in a Git repository after last quality controls, concerning mainly the linking between the two sub-ontologies, TMI and ATU.

We started also to model the “Thirty-six dramatic situations” with the aim of proposing a set of ontologies that can be used for multiple metadata annotation of literary work. Current work is on adding to the instances of the ontologies URLs of folktales that are marked with the corresponding numbers and so to allow access to those via Linked Data mechanisms. For this we are looking at the web page <http://www.mftd.org/> for ATU and we started a cooperation with the project “Digital Breadcrumbs of Brothers Grimm”¹⁴ for including the fine grained and multilingual motif annotations generated in this project in our TMI ontology.

6 ACKNOWLEDGMENTS

We would like to thank all the (former) students involved in the efforts described in this paper, all of them co-authoring papers on the corresponding topics, some of those included in the references list: Antonia Scheidel, Nikolina Koleva, Christian Eisenreich, Jana Ott, Tonio Süßdorf, Christian Willms, Tyler Klement and Antónia Košťová

We would also like to thank the colleagues involved in the “Digital Breadcrumbs of Brothers Grimm” project, for initiating a fruitful cooperation and offering a concrete application scenario for our ontologies.

REFERENCES

- [1] Antti Aarne. 1961. *The Types of the Folktale: A Classification and Bibliography*. The Finnish Academy of Science and Letters. Translated and Enlarged by S. Thompson. Second Revision (FFC 184).
- [2] Christian Eisenreich, Jana Ott, Tonio Süßdorf, Christian Willms, and Thierry Declerck. 2014. From Tale to Speech: Ontology-based Emotion and Dialogue Annotation of Fairy Tales with a TTS Output. In *Proceedings of ISWC 2014*.
- [3] Nikolinal Koleva, Thierry Declerck, and Hans-Ulrich Krieger. 2012. An Ontology-Based Iterative Text Processing Strategy for Detecting and Recognizing Characters in Folktales. In *Proceedings of the Digital Humanities 2012 Conference*.
- [4] Georges Polti. 1916. *The Thirty-six Dramatic Situations*. Trans., Lucile Ray, Editor Company.
- [5] Vladimir Propp. 1968. *Morphology of the folktale*. Trans., Laurence Scott. 2nd ed., University of Texas Press.
- [6] Antonia Scheidel and Thierry Declerck. 2010. APfML - Augmented Proppian fairy tale Markup Language. In *Proceedings of the First International AMICUS Workshop on Automated Motif Discovery in Cultural Heritage and Scientific Communication Texts*.
- [7] Stith Thompson. 1955-1958. *Motif-index of folk-literature: A classification of narrative elements in folktales, ballads, myths, fables, medieval romances, exempla, fabliaux, jest-books, and local legends*. Revised and enlarged edition, Indiana University Press.
- [8] Hans-Jörg Uther. 2014. *The Types of International Folktales: A Classification and Bibliography. Based on the system of Antti Aarne and Stith Thompson*. Suomalainen Tiedekatemia.

¹³<http://protege.stanford.edu/>

¹⁴<http://www.etrapp.eu/digital-breadcrumbs-of-brothers-grimm/>