

Annotation in DH (annDH)

Workshop at  **ESLLI 2018**
August 6-17, 2018, Sofia, Bulgaria

Ontology-driven Annotation of Literary Texts

Thierry Declerck
Multilingual Technologies Lab
DFKI GmbH
Saarbrücken, Germany

Background

- This presentation is based on the results of a series of Bachelor/Master theses and software projects conducted by students of the Computational Linguistics Department of the Saarland University. Next slide lists some of the papers that describe those works.

Selected List of Publications

- Antonia Scheidel, Thierry Declerck. APfML - Augmented Proppian fairy tale Markup Language. In: Sándor Darányi, Piroska Lendvai (eds.): First International AMICUS Workshop on Automated Motif Discovery in Cultural Heritage and Scientific Communication Texts: Poster session, Vienna, Austria, Szeged University, Szeged, Hungary, 10/2010
- Thierry Declerck, Antonia Scheidel, Piroska Lendvai. Proppian Content Descriptors in an Integrated Annotation Schema for Fairy Tales. In: Language Technology for Cultural Heritage. Selected Papers from the LaTeCH Workshop Series, Theory and Applications of Natural Language Processing, Pages 155-169, Springer, Heidelberg, 2011
- Nikolina Koleva, Thierry Declerck, Hans-Ulrich Krieger. An Ontology-Based Iterative Text Processing Strategy for Detecting and Recognizing Characters in Folktales. In: Jan Christoph Meister (ed.): Digital Humanities 2012 Conference Abstracts, Pages 467-470, Hamburg, Germany, Hamburg University Press, University of Hamburg, Hamburg, Hamburg, 7/2012
- Thierry Declerck, Nikolina Koleva, Hans-Ulrich Krieger. Ontology-Based Incremental Annotation of Characters in Folktales. In: Kalliopi Zervanou, Antal van den Bosch (eds.): Proceedings of the 6th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH 2012), Pages 30-35, Avignon, France, ACL, Association for Computational Linguistics (ACL), 209 N. Eighth Street Stroudsburg, PA 18360. USA, 4/2012
- Christian Eisenreich, Jana Ott, Tonio Süßdorf, Christian Willms, Thierry Declerck. From Tale to Speech: Ontology-based Emotion and Dialogue Annotation of Fairy Tales with a TTS Output. In: Proceedings of ISWC 2014, Riva del Garda, Italy, Springer, 10/2014
- Thierry Declerck, Antónia Kostová, Lisa Schäfer. Towards a Linked Data Access to Folktales classified by Thompson's Motifs and Aarne-Thompson-Uther's Types. In: Proceedings of Digital Humanities 2017, Montréal, QC, Canada, ADHO, 8/2017
- Thierry Declerck, Lisa Schäfer. Porting past Classification Schemes for Narratives to a Linked Data Framework. In: Apostolos Antonacopoulos, Marco Büchler (eds.): Proceedings of DATECH2017, Göttingen, Germany, ACM, 6/2017
- Thierry Declerck, Anastasija Aman, Martin Banzer, Dominik Macháček, Lisa Schäfer, Natalia Skachkova. Multilingual Ontologies for the Representation and Processing of Folktales. In: Anca Dinu, Petya Osenova, Cristina Vertan (eds.): Proceedings of the First Workshop on Language technology for Digital Humanities in Central and (South-)Eastern Europe, Pages 20-24, Varna, Bulgaria, INCOMA Ltd, Shoume, 9/2017
- Matthias Lindemann, Stefan Grünewald, Thierry Declerck. Annotation and Classification of Locations in Folktales. In: Andrew U. Frank, Christine Ivanovic, Francesco Mambrini, Marco Passarotti, Caroline Sporleder (eds.): Proceedings of the Second Workshop on Corpus-Based Research in the Humanities, Vienna, Austria, Gerastree Proceedings, GTP 1., Academy Corpora of the Austrian Academy of Science, Sonnenfelsgasse 19, 1010 Wien, Austria, Vienna, 1/2018

Background (2)

- In this talk I will give a special focus to two topics that have been presented in the annDH workshop
 - “Added Value of Coreference Annotation for Character Analysis in Narratives”, presented by *Melanie Andresen* and Michael Vauth.
 - “An Extended Hermeneutic Cycle” presented by Heike Zinnsmeister and Sandra Kübler in their introduction to the workshop and also by Janis Pagel et al., “A Unified Annotation Workflow for Diverse Goals”. For both cases our focus is on trying to specify what can be the “theory” that can be (in)validated by annotations.
- Overall, our aim is to investigate how Computational Linguistics AND Semantic Web technologies can help for the annotation of literary texts, with a focus on folk tales. The main technology we are dealing with in this talk is given by ontologies (in the IT sense).

Iterative and incremental Interaction between
Computational Linguistics and a Domain Ontology
for the Detection and Mark-Up of Characters in
Folktales

(Bachelor Work by Nikolina Koleva)

Ontology as a semantic Resources for detecting and storing Characters of Folk Tales

- We developed an ontology for the formal representation of some tales, giving a lot of place to the description of family relations, since this is an important topic in folk tales. (Theory?)
 - Studying the use of ontologies for the persistent storage of referential elements of tales, and for a subset of co-reference resolution task, together with the text data (annotations), not dealing (yet) with anaphora resolution.
 - Studying the relation between Computational Linguistics and Ontologies for knowledge-based text analysis
- The ontology models concepts and the relations between them, as well as individuals and their properties. (Theory?)
 - The ontology was created with the Protégé editor and we used the Web Ontology Language OWL for modelling the domain

A Screenshot of the Definition of the Class “Mother”, in the uninstantiated Ontology

The screenshot displays a web browser interface for an ontology. The address bar shows the URL `monnet.owl (http://www.dfki.de/lt/monnet.owl)`. The main interface has several tabs: **Active Ontology**, **Entities**, **Classes**, **Object Properties**, **Data Properties**, **Individuals**, **OWLviz**, and **DL Query**. The **Classes** tab is selected, showing an **Asserted class hierarchy** on the left and **Annotations: Mother** and **Description: Mother** on the right.

Asserted class hierarchy: Mother

- Thing
 - Abstract
 - Animal
 - BodyOfWater
 - Bodypart
 - Event
 - Human
 - Object
 - Plant
 - SocialAbstraction
 - Multiple
 - Single
 - Child
 - Parent
 - BiolParent
 - Father
 - Mother**
 - Stepparent
 - Relative
 - Sibling
 - Speaker
 - Supernatural
 - Time

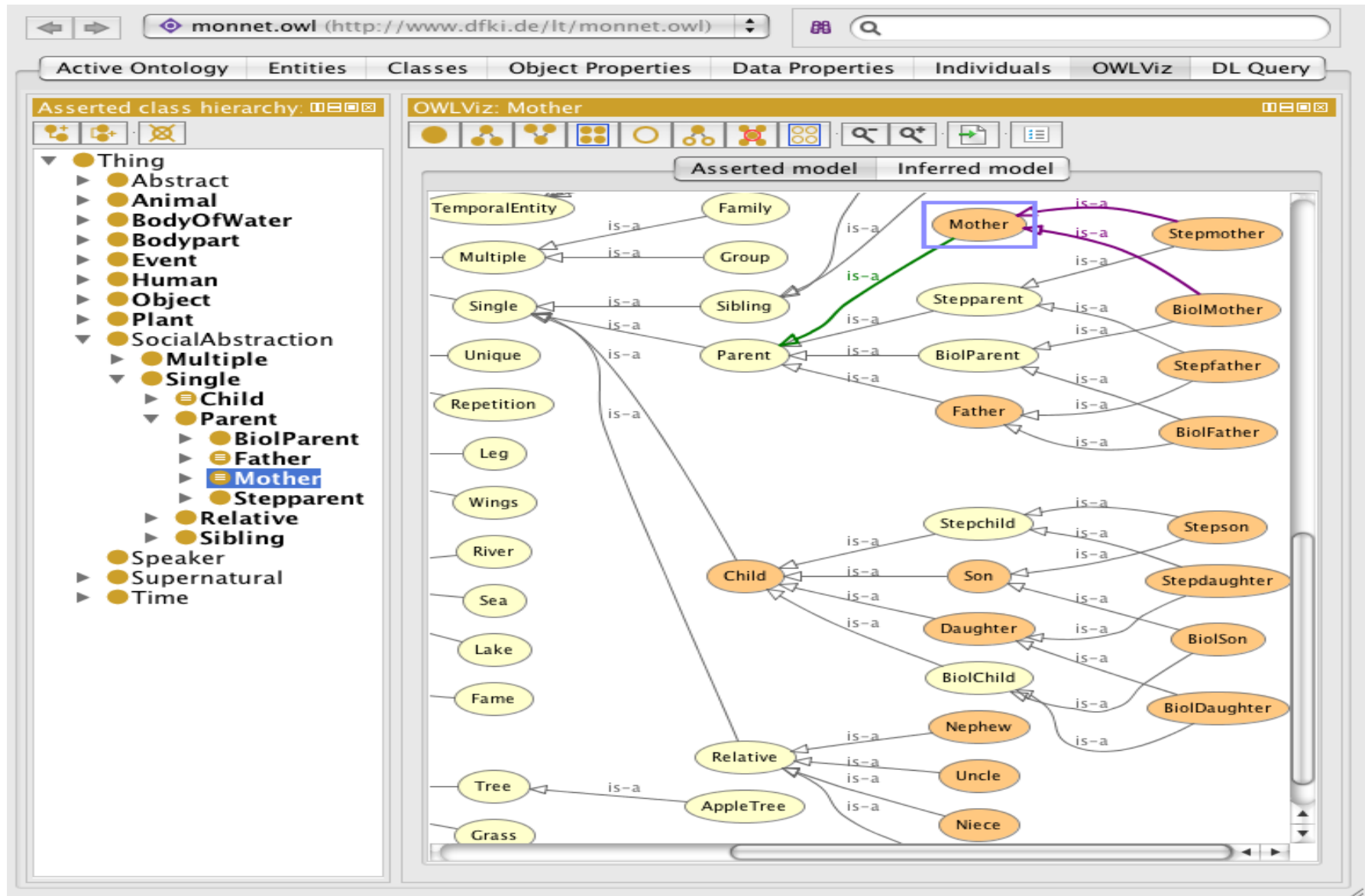
Annotations: Mother

- Annotations +
- language** "Майка"@bg
- language** "Mutter"@de
- language** "Mother"@en
- language** "Мать"@ru
- comment** "The class of mother is a subclass of parent and designates all mothers. Whether the mother is biological or stepmother is unspecified."

Description: Mother

- Equivalent classes +
- Parent** and hasGender value "f"
- hasGender value "f" and hasChild min 1 Thing
- Superclasses +
- Parent**
- Inferred anonymous superclasses
- Members +
- Disjoint classes +

Class Hierarchy



A Screenshot of the object_property “hasChild”

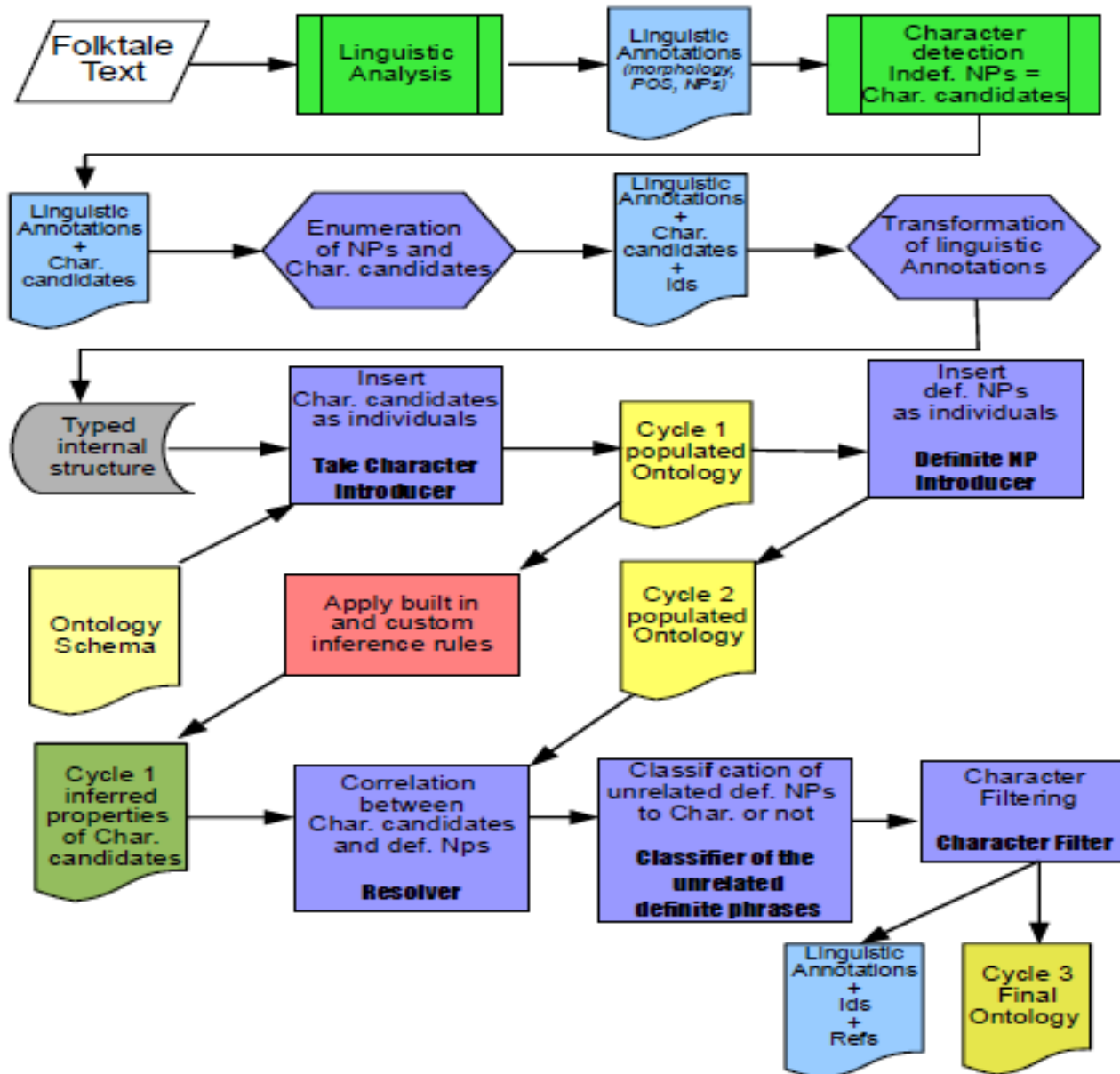
The screenshot displays a web-based ontology editor interface for the file `monnet.owl` (URL: `http://www.dfki.de/lt/monnet.owl`). The interface is divided into several panes:

- Object properties: hasChild:** A tree view of the ontology's object properties. The `hasChild` property is selected and expanded, showing sub-properties: `hasBiolChild` and `hasStepChild`. Other visible properties include `hasAunt`, `hasHusband`, `hasMember`, `hasNephew`, `hasNiece`, `hasOwner`, `hasParent`, `hasPart`, `hasParticipant`, `hasSibling`, `hasSubevent`, `hasTime`, `hasUncle`, `hasValidPeriod`, `hasWife`, and `isLocation`.
- Characteristics:** A list of checkboxes defining the property's characteristics:
 - Functional
 - Inverse functional
 - Transitive
 - Symmetric
 - Asymmetric
 - Reflexive
 - Irreflexive
- Description: hasChild:** A panel showing the property's domain and range:
 - Domains (intersection):** `Family or Parent`
 - Ranges (intersection):** `Child`
 - Equivalent object properties:** `hasParent`

Custom Inference Rules applied to Ontology Elements

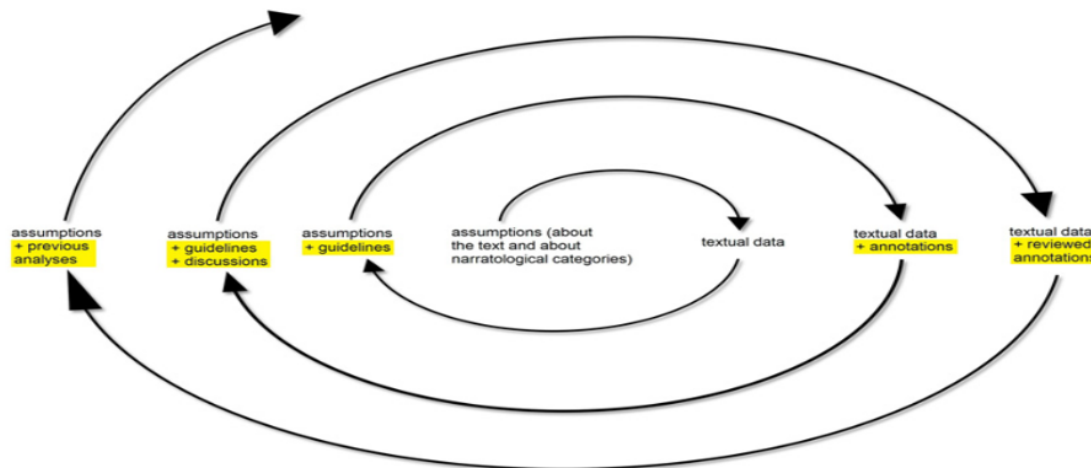
1. $\text{hasParent}(\text{?x}, \text{?x1}), \text{hasParent}(\text{?x}, \text{?x2}), \text{hasParent}(\text{?y}, \text{?x1}), \text{hasParent}(\text{?y}, \text{?x2}), \text{hasGender}(\text{?x}, \text{"f"}), \text{notEqual}(\text{?x}, \text{?y}) \Rightarrow \text{Sister}(\text{?x})$
2. $\text{Daughter}(\text{?d}), \text{Father}(\text{?f}), \text{Son}(\text{?s}) \Rightarrow$
 $\text{hasBrother}(\text{?d}, \text{?s}), \text{hasChild}(\text{?f}, \text{?s}), \text{hasChild}(\text{?f}, \text{?d}),$
 $\text{hasSister}(\text{?s}, \text{?d})$

Workflow of the ontology-based Algorithm for the Detection, Recognition and Annotation of Characters in Folk Tales



Relation between our workflow and the “Hermeneutical Cycle”

- The iterative and incremental cyclic form of the workflow is very similar to the one of the “Hermeneutic Circle” – mutatis mutandis -- mentioned by Zinsmeister & Kübler in the introduction of the workshop or by Pagel et al. (A Unified Annotation Workflow for Diverse Goal), also in this workshop.



Gius and Jacke (2017)

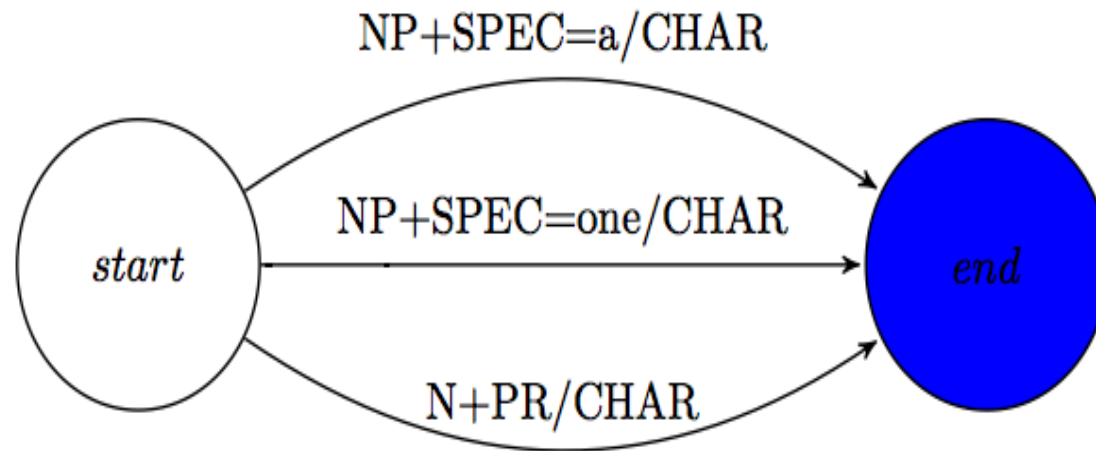
Grammar for 1st Population Cycle – detecting Indefinite NPs or NEs

- FST Code

Main = :Char | :PropName;

Char = <E>/<CHAR (<NP+SPEC=a> | <NP+SPEC=one>)
<E>/>;

PropName = <E>/<CHAR <N+PR> <E>/>;



Resulting (linguistic) Annotation, enumerating the detected Characters

```
<text>
<s id="S1" tokstart="tok1" tokend="tok17">
<clause id="C1" tokstart="tok1" tokend="tok9">
<w pos="EX" id="tok1">There</w>
<w pos="VBD" id="tok2">lived</w>
<chunk cat="NP" id="ph1" tokstart="tok3"
tokend="tok9">
<chunk cat="NP" id="ph2" ref="ch1"
tokstart="tok3" tokend="tok5">
<w pos="DT" id="tok3">an</w>
<w pos="JJ" id="tok4">old</w>
<w pos="NN" id="tok5">man</w>
</chunk>
<w pos="CC" id="tok6">and</w>
<chunk cat="NP" id="ph3" ref="ch2"
tokstart="tok7" tokend="tok9">
<w pos="DT" id="tok7">an</w>
<w pos="JJ" id="tok8">old</w>
<w pos="NN" id="tok9">woman</w>
</chunk>
</chunk>
</clause>

```

```
<w pos="$PUNCT" >;</w> <clause id="C2"
tokstart="tok10" tokend="tok17">
<w pos="PRP" id="tok10" ref="ph1">they</w>
<w pos="VBD" id="tok11">had</w>
<chunk cat="NP" id="ph4" tokstart="tok12"
tokend="tok17">
<chunk cat="NP" id="ph5" ref="ch3" tokstart="tok12"
tokend="tok13">
<w pos="DT" id="tok12">a</w>
<w pos="NN" id="tok13">daughter</w>
</chunk>
<w pos="CC" id="tok14">and</w>
<chunk cat="NP" id="ph6" ref="ch4" tokstart="tok15"
tokend="tok17">
<w pos="DT" id="tok15">a</w>
<w pos="JJ" id="tok16">little</w>
<w pos="NN" id="tok17">son</w>
</chunk>
</chunk>
</clause>
<w pos="$.">.</w>
•</s>
•</text>

```

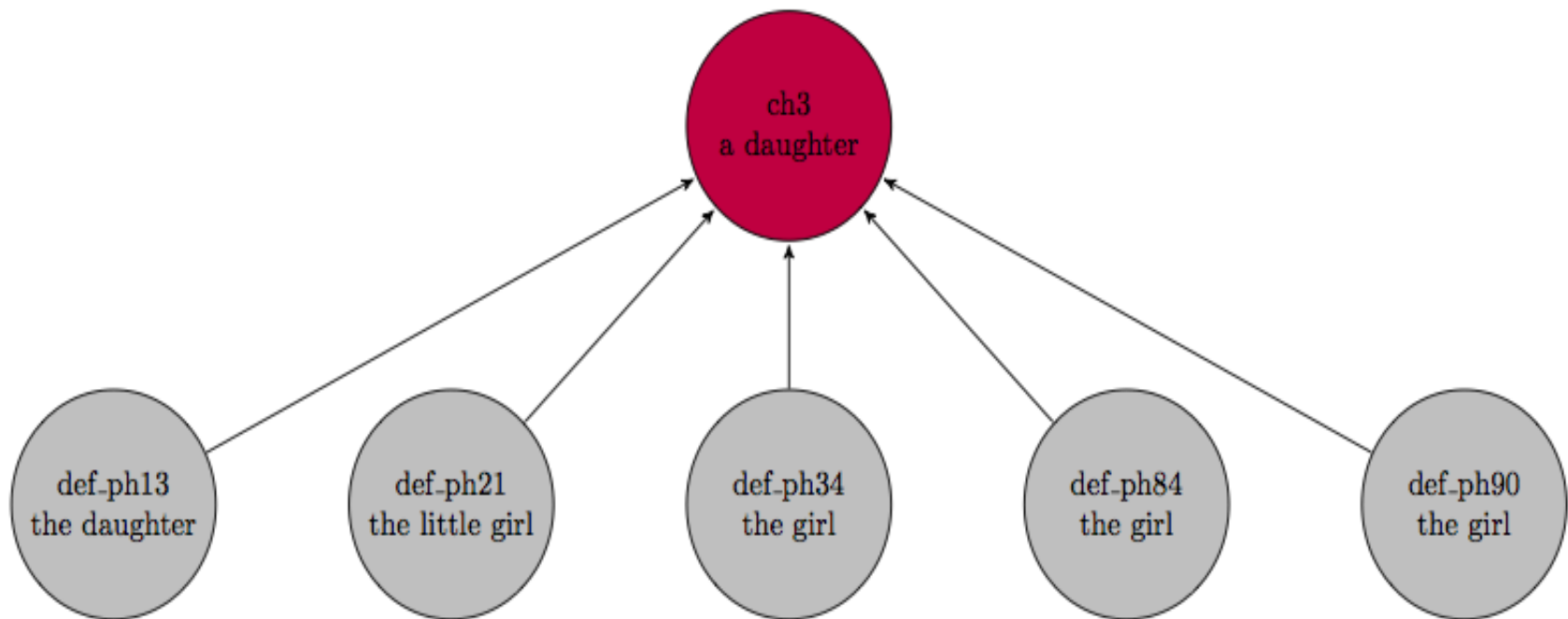
Screenshot of the Ontology after the first Population and running the Reasoner for the third Tale Character

The screenshot shows the OWL browser interface for the ontology 'monnet.owl' (http://www.dfki.de/lt/monnet.owl). The 'Individuals' tab is active, and the individual 'ch3' is selected. The interface is divided into several panels:

- Individuals:** A list of individuals from ch1 to ch19. 'ch3' is highlighted.
- Annotations:** Shows a label 'a daughter' for individual 'ch3'.
- Description:** Shows the class hierarchy for 'ch3', including 'Daughter', 'Thing', 'Girl', and 'Sister'.
- Property assertions:** Shows object property assertions for 'ch3': 'hasBrother ch4', 'hasParent ch1', 'hasParent ch2', and 'hasSibling ch4'. It also shows data property assertions: 'sole true' and 'hasGender "f"'. Negative assertions are also visible but empty.

Second CL cycle: Mapping Definite NPs to already stored Indef-Def Nps: co-ref. resolution

model for the assignment of the referring definite noun phrases to the already identified characters (including the indices of the analysed phrases)



Screenshot of the Ontology after a Character Filtering Step and running the Reasoner for the third Tale Character.

The screenshot displays the 'monnet.owl' ontology viewer interface. The 'Active Ontology' tab is selected, and the 'Individuals' panel on the left shows a list of characters. The 'ch3' character is highlighted, and its details are shown in the main area. The 'Annotations: ch3' panel shows a label 'a daughter'. The 'Description: ch3' panel shows the character is a 'Daughter', 'Thing', 'Girl', and 'Sister'. The 'Property assertions: ch3' panel shows various relationships, including 'hasBrother' and 'hasParent'.

Individuals: ch3

- indef_ph24
- indef_ph29
- indef_ph35
- indef_ph36
- indef_ph42
- indef_ph5
- indef_ph61
- indef_ph7
- indef_ph8
- indef_ph9
- ch1
- ch10
- ch12
- ch15
- ch16
- ch19
- ch2
- ch20
- ch21
- ch3
- ch4
- def_ph10
- def_ph11
- def_ph12
- def_ph13
- def_ph14
- def_ph15
- def_ph16
- def_ph17
- def_ph18

Annotations: ch3

Annotations +

label "a daughter"

Description: ch3

Types +

- Daughter
- Thing
- Girl
- Sister

Same individuals +

- def_ph25
- def_ph56
- def_ph60
- def_ph12
- def_ph18

Different individuals +

Property assertions: ch3

Object property assertions +

- hasBrother ch4
- hasBrother def_ph41
- hasBrother def_ph17
- hasParent def_ph6
- hasParent ch1
- hasParent ch2
- hasSibling ch4
- hasSibling def_ph41
- hasSibling def_ph17

Data property assertions +

- sole true
- hasGender "f"
- hasNumber "s"

Negative object property assertions +

Negative data property assertions +

IDs of text spans are thus included => Storage of annotations

Checking the Validity of the Approach

Limited Scenario (run on 2 tales) and “very tiny” evaluation “study” on Gold Standard for one tale (The “Magic Swan Geese”).

	acc. to GoldStandard is character	acc. to GoldStandard is not character
acc. to the tool is character	8	1
acc. to the tool is not character	3	7

The precision amounts to 88%; the recall to 73%; and the value of the balanced F-measure is 80%. A result obtained on the base of a very simple algorithm, making use of more sophisticated ontology technologies. Promising.

The wrongly detected character is due to the presence of an oven as a character (a “helper” in Proppian terms) and the “real” oven in the house of Baba Yaga. Missed characters are due partly to the lack of data in the ontology.

Some comparisons to the work by Melanie Andresen & Michael Vauth

- Contrary to the novels:
 - In tales we have very few named entities. All could be listed in a gazetteer (example: “Baba Yaga”). Most characters are introduced by an Indefinite-NP, but can be sometimes introduced by a Definite-NP if we have a prototypical character (“the wolf”).
 - Tales are short texts with a limited number of characters.
- We didn’t yet implement our heuristic rules for anaphora resolution and we expect with this step another significant increase in the detection of occurrences of characters
- In both studies we can see the added-value of co-reference resolution for the better (automated?) interpretation of characters in narratives

Possible Extensions of the ontology-based Approach to the novel type of literary work

- Considering the work “Added Value of Coreference Annotation for Character Analysis in Narratives”, presented by *Melanie Andresen* & Michael Vauth in this workshop, one idea was to look if there are kind of “casts” available, so that the (main) characters are known before starting the detection of co-reference expressions.
- Those “casts” can in fact be “transformed” into the (first type of) ontology we developed for the folk tales, along the line of gender, age, family relations etc. We assume that this can offer support for improving still a bit the work of co-reference detection of characters.
- Many thanks to Heike Zinnsmeister for putting me on this track :-)

“Cast” for “Corpus Delicti” can be taken from [https://en.wikipedia.org/wiki/The_Method_\(novel\)](https://en.wikipedia.org/wiki/The_Method_(novel))

- Mia Holl is a 34-year-old Biologist and is the main character of the novel. After the death of her brother, Moritz Holl, she gets lonely and depressed. Throughout the book she becomes a rebel against the government. Her dedication to the METHODE is not strong enough to be actively against it. The name 'Mia Holl' comes from the name 'Maria Holl', a woman who was thought to be a witch in the 17th century.
- Moritz Holl is Mia's 27 year old brother, who even before the story starts commits suicide. He loves nature but is also an independent rebel who wants nothing more than freedom. Because of his complex thoughts he couldn't find a person to talk to and therefore created the Ideal beloved, who he passed on to Mia just before his death. The name 'Moritz Holl' comes from another character named 'Max' in another Novel from Juli Zeh.
- The Ideal Beloved is a fictional character who helps Mia through the difficult time after the death of her brother. She has the same ideology and thoughts as Moritz and could be his ghost. After Mia's emotional wound is healed she disappears, since her quest is done.
- Bell is a public prosecutor and a follower of the METHOD. He is a follower of the METHOD and is always in a conflict with Sophie, a young judge.
- Lutz Rosentreter is Mia's lawyer in the process of proving her brother's innocence. He is against the METHOD, since he lost the love of his life because of the government system. Since then he is trying to get revenge.
- Heinrich Kramer is Mia's opponent in the Novel. He comes across as a nice Gentleman who has a huge influence on the METHOD and wants to explain the system to everyone. In reality he is a fanatic who can only see his goal.

What is needed for this Step

- Moving from the small scale ontology created for the co-reference detection and annotation of characters in folk tales to a fully fledged biographical ontology, containing not only family relationships, but also professional activities etc. So that expressions like “the lawyer” can be associated with the correct character in the novel (Lutz Rosentreter).
- We have such an ontology (<http://www.dfki.de/It/onto/trendminer/BIO/biography.owl>). With this the detection and annotation of characters in novels could be supported by a knowledge-base.
- As a “by-product”: the full set of co-reference annotations can lead to the creation of a biography of fictional characters.

Added-Values of the Ontology-Based Approach

- Characters (and other elements) in literary works can be uniquely stored (with an URI) and referred to from annotations within the text but also from other texts, also in a multilingual context (“Le Petit Chaperon Rouge”, “Rotkäppchen”, “Little Red Riding Hood”, “Красная Шапочка”)
- Easy versioning of text and annotation
- Changes in the model leading straight away to new annotations.
- Possibility to link to text external knowledge bases in the context of the Linked Data cloud

End of Part 1

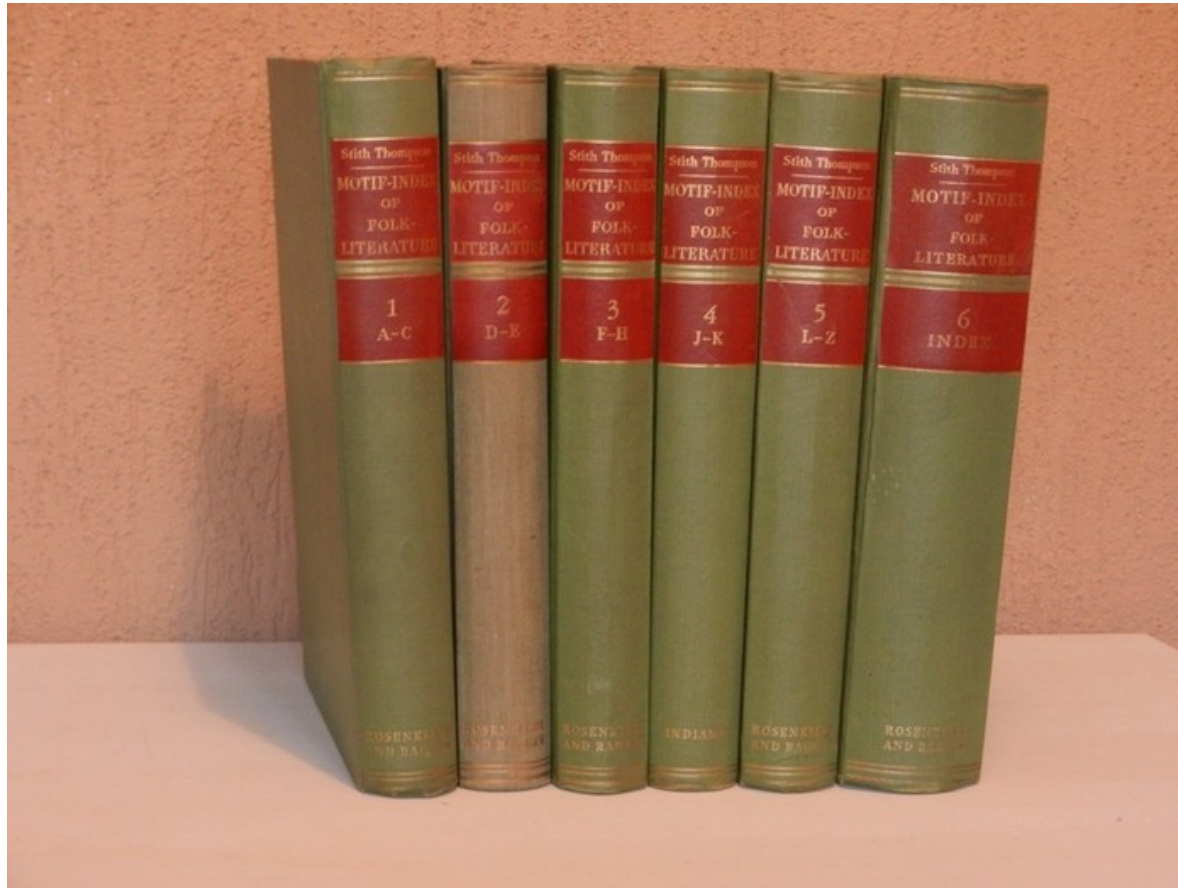
Part 2

Integrated Ontologies for the Classification of Folk Tales

The starting Point: Two classical Classification/Indexing schemes

- Two well-known classification/indexing systems used by folklorists (**Theory?**):
 - TMI - Thompson-Motif-Index of Folk-Literature
 - ATU - Aarne-Thompson-Uther classification of tale types
- Both of them are available as printed sources, or as online resources in html or pdf format. Since the two systems are related to each other, our aims are to:
 - organize them in one ontology with appropriate references,
 - make the resulting ontology available online,
 - *implement a web interface for SPARQL querying, and*
 - *implement an automatic classifier of texts based on statistical approach.*

TMI



Motif-index of folk-literature, a classification of narrative elements in folk-tales, ballads, myths, fables, mediaeval romances, exempla, fabliaux, jest-books and local legends. Helsinki, Academia scientiarum fennica, 1932-1936. 6 volumes. Folklore Fellows Communications, no 106-109, 116-117
Revised and enlarged edition. Bloomington ; London, Indiana university press, 1955-58. 6 volumes

TMI as a Web/HTML Resource

← ⓘ | https://sites.ualberta.ca/~urban/Projects/English/Motif_Index.htm | 🔍 TMI alberta motif index | ☆ | 📄 | 📁 | 🏠 | ↶ | ☰

S. Thompson. Motif-index of folk-literature : a classification of narrative elements in folktales, ballads, myths, fables, mediaeval romances, exempla, fabliaux, jest-books, and local legends.

Revised and enlarged edition. Bloomington : Indiana University Press, 1955-1958.

[Restore frame](#)

Grant support: INTAS project 05-1000008-7922, ÐÒÒÈ #06-06-80-420a, ÐÒÒÈ #07-06-00441-à

- [A. Mythological Motifs](#)
- [B. Animal Motifs](#)
- [C. Motifs of Tabu](#)
- [D. Magic](#)
- [E. the Dead](#)
- [F. Marvels](#)
- [G. Ogres](#)
- [H. Tests](#)
- [J. the Wise and the Foolish](#)
- [K. Deceptions](#)
- [L. Reversals of Fortune](#)
- [M. Ordaining the Future](#)
- [N. Chance and Fate](#)
- [P. Society](#)
- [Q. Rewards and Punishments](#)
- [R. Captives and Fugitives](#)

[Main page](#)

S. Thompson. Motif-index of folk-literature : a classification of narrative elements in folktales, ballads, myths, fables, mediaeval romances, exempla, fabliaux, jest-books, and local legends.

Revised and enlarged. edition. Bloomington : Indiana University Press, 1955-1958.

Grant support: INTAS project 05-1000008-7922, ÐÒÒÈ #06-06-80-420a, ÐÒÒÈ #07-06-00441-à

A. MYTHOLOGICAL MOTIFS

†A0-†A99. **Creator.**

A0. †A0. **Creator.**--For a general bibliography of creation myths, see Alexander N. Am. 278 n. 15. For bibliographies of North American Indian mythologies arranged by areas, see Thompson Tales 272 n. 1; **Feilberg Skabelses og Syndflodssagn; Jewish: Neuman.--Mexican Indian: (Tarascan) Alexander Lat. Am. 85, (Zapotecan) ibid. 87; Guarayú: Métraux RMLP XXXIII 147; Polynesia: Dixon 21 n. 47; Hawaiian: Beckwith Myth 42; Mono-Alu: Wheeler 28, 66f., 70; Easter Is.: Métraux BMB CLX 313; Marshall Is.: Davenport Folk Tales 221f.; Tahiti: Henry Ancient Tahiti 335ff.; New Hebrides: Codrington II 365.--Armenian: Ananikian 20; African: Werner African 127ff., **Frobenius and Fox, (Loango): Pechuël-Loesche 267; Hindu: Penzer I 10; Buddhist myth: Malalasekera II 338; Icel.: Boberg, MacCulloch Eddic 326; Irish myth: Cross.

A1. †A1. *Identity of creator.*

A1.1. †A1.1. *Sun-god as creator.*--Egyptian: Müller 69; Persian: Carnoy 260.

A1.2. †A1.2. *Grandfather as creator.*--S. Am. Indian (Paressi): Métraux BBAE CXLIII (3) 359, (Guarayú): Métraux RMLP XXXIII 147.

A1.3. †A1.3. *Stone-woman as creator.*--Paressi: Métraux BBAE CXLIII (3) 359.

A1.4. †A1.4. *Brahma as creator.*--Buddhist myth: Malalasekera II 338.

A2. †A2. *Multiple creators.*

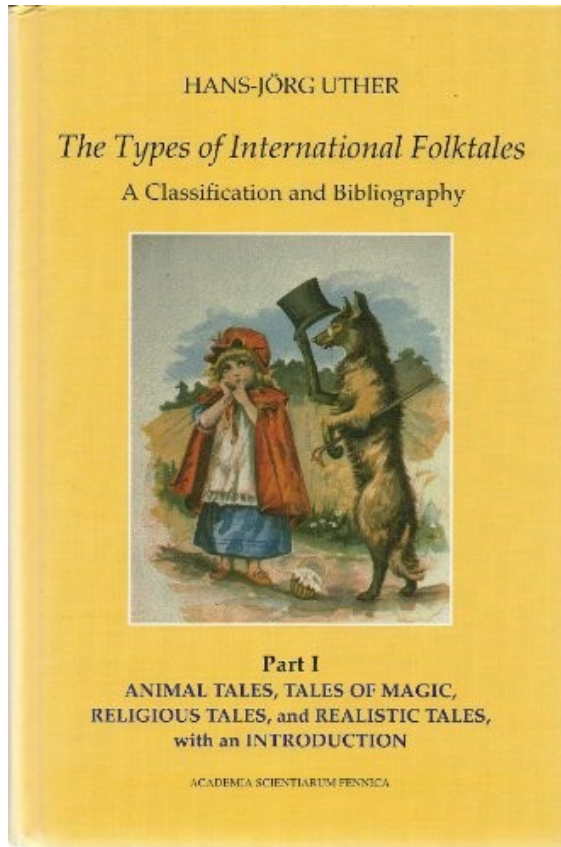
A2.1. †A2.1. *Three creators.*--Icel.: Boberg, MacCulloch Eddic 327.--Oceanic: Dixon 24; Hawaii: Beckwith Myth 42.

A2.2. †A2.2. *First human pair as creators.* (Cf. †A1270.) Chinese: Eberhard FFC CXX 115 No. 70.

A3. †A3. *Creative mother source of everything.*--India: Thompson-Balys.

A5. †A5. *Reason for creation.*

ATU



Uther, Hans-Jörg. 2004. *The Types of International Folktales: A Classification and Bibliography*.
Based on the system of Antti Aarne and Stith Thompson.
FF Communications no. 284–286.
Helsinki: Suomalainen Tiedeakatemia. Three volumes.

ATU partially available on-line

<http://www.mftd.org/index.php?action=atu>

Multilingual Folk Tale Database

[Home](#) • [Browse Stories](#) • [Classification](#) • [login](#)

Aarne-Thompson-Uther Classification of Folk Tales

There are many different folk tales in the world, but many tales are variations on a limited number of themes. The classification of Aarne-Thompson and later by Uther, is intended to bring out the similarities between tales by grouping variants of the same tale under 1

Below is the full tree of the ATU classification. Click on a title to see all the stories within that class.

- [ANIMAL TALES](#) 1-299
 - [Wild Animals](#) 1-99
 - [The Clever Fox \(Other Animal\)](#) 1-69
 - [Other Wild Animals](#) 70-99
 - [Wild Animals and Domestic Animals](#) 100-149
 - [Wild Animals and Humans](#) 150-199
 - [Domestic Animals](#) 200-219
 - [Other Animals and Objects](#) 220-299
- [TALES OF MAGIC](#) 300-749
 - [Supernatural Adversaries](#) 300-399
 - [Supernatural or Enchanted Wife \(Husband\) or Other Relative](#) 400-459
 - [Wife](#) 400-424
 - [Husband](#) 425-449
 - [Brother or Sister](#) 450-459
 - [Supernatural Tasks](#) 460-499
 - [Supernatural Helpers](#) 500-559
 - [Magic Objects](#) 560-649

Example of typed Folk Tale in the Multilingual ATU Database

Multilingual Folk Tale Database

[Home](#) • [Browse Stories](#) • [Classification](#) • [login](#)

Лисичка-сестричка и волк

Александр Афанасьев

Жил себе дед да баба. Дед говорит бабе: «Ты, баба, пеки пироги, а я поеду за рыбой». Наловил рыбы и везет домой целый воз. Вот едет он и видит: лисичка свернулась калачиком и лежит на дороге. Дед слез с воза, подошел к лисичке, а она не ворохнется, лежит себе как мертвая. «Вот будет подарок жене», — сказал дед, взял лисичку и положил на воз, а сам пошел впереди. А лисичка улучила время и стала выбрасывать полегоньку из воза все по рыбке да по рыбке, все по рыбке да по рыбке. Пovyбросала всю рыбу, и сама ушла.

«Ну, старуха, — говорит дед, — какой воротник привез я тебе на шубу». — «Где?» — «Там, на возу, — и рыба и воротник». Подошла баба к возу: ни воротника, ни рыбы, и начала ругать мужа: «Ах ты, старый хрен! Такой-сякой! Ты еще вздумал обманывать!» Тут дед смекнул, что лисичка-то была не мертвая; погоревал, погоревал, да делать-то нечего.

А лисичка собрала всю разбросанную по дороге рыбу в кучку, села и ест себе. Навстречу ей идет волк: «Здравствуй, кумушка!» — «Здравствуй, куманек!» — «Дай мне рыбки!» — «Налови сам, да и ешь». — «Я не умею». — «Эка, ведь я же наловила; ты, куманек, ступай на реку, опусти хвост в прорубь — рыба сама на хвост нацепляется, да смотри, сиди подольше, а то не наловишь».

Волк пошел на реку, опустил хвост в прорубь; дело-то было зимою. Уж он сидел, сидел, целую ночь просидел, хвост его и приморозило; попробовал было приподняться: не тут-то было. «Эка, сколько рыбы привалило, и не вытащишь!» — думает он. Смотрит, а бабы идут за водой и кричат, завидя серого: «Волк, волк! Бейте его! Бейте его!» Прибежали и начали колотить волка — кто коромыслом, кто ведром, чем кто попало. Волк прыгал-прыгал, оторвал себе хвост и пустился без оглядки бежать. «Хорошо же, — думает, — уж я тебе отплачу, кумушка!»

А лисичка-сестричка, покушавши рыбки, захотела попробовать, не удастся ли еще что-нибудь стянуть; забралась в одну избу, где бабы пекли блины, да попала головой в кадку с тестом, вымазалась и бежит. А волк ей навстречу: «Так-то учишь ты? Меня всего исколотили!» — «Эх, куманек, — говорит лисичка-сестричка, — у тебя хоть кровь выступила, а у меня мозг, меня больней твоего прибили; я насилу плетусь». — «И то правда, — говорит волк, — где тебе, кумушка, уж иди; садись на меня, я тебя довезу». Лисичка села ему на спину, он ее и понес. Вот лисичка-сестричка сидит, да потихоньку и говорит: «Битый небитого везет, битый небитого везет». — «Что ты, кумушка, говоришь?» — «Я, куманек, говорю: битый битого везет». — «Так, кумушка, так!»

«Давай, куманек, построим себе хатки». — «Давай, кумушка!» — «Я себе построю лубяную, а ты себе ледяную». Принялись за работу, сделали себе хатки: лисичке — лубяную, а волку — ледяную, и живут в них. Пришла весна, волчья хатка и растаяла. «А, кумушка! — говорит волк. — Ты меня опять обманула, надо тебя за это съесть». — «Пойдем, куманек, еще поконаемся, кому-то кого достанется есть». Вот лисичка-сестричка привела его в лес к глубокой яме и говорит: «Прыгай! Если ты перепрыгнешь через яму — тебе меня есть, а не перепрыгнешь — мне тебя есть». Волк прыгнул и попал в яму. «Ну, — говорит лисичка, — сиди же тут!» — и сама ушла.

Идет она, несет сказочку в ладках и просится к мужику в избу: «Пусти лисичку-сестричку перекусить»... «У



Information

Author: **Александр Афанасьев** - 1855

Original version in **Russian**

Source: **Народные Русские Сказки** (nr. 001)

Country of origin: **Russia**

Story type: **The theft of fish (ATU 1)**

Translations

There are no translations available for this story

[Add a translation](#)

ATU Textfile

1 *The Theft of Fish*. (Including the previous Types 1* and 1**.) A fox (hare, rabbit, coyote, jackal) lies in the road pretending to be dead. A fisherman throws him on his wagon which is full of fish (cheese, butter, meat, bread, money). The fox throws the fish out of the wagon [K371.1] and jumps down after them [K341.2, K341.2.1].

A wolf (bear, fox, coyote, hyena) tries to imitate this and pretends to be dead, too. The fisherman catches him and beats him [K1026]. Cf. Types 56A, 56B, and 56A*.

In some variants one animal (rabbit, fox) pretends to be dead in order to distract a man who is carrying a basket of food. Another animal (fox, wolf) steals the basket. (Previously Type 1*, cf. Type 223.) Or an animal makes a hole in the basket so that the contents fall out. (Previously Type 1**.)

Integrated Ontology (ATU-TMI)

The screenshot displays the ATU-TMI ontology editor interface. The top navigation bar includes 'File', 'Edit', 'View', 'Reasoner', 'Tools', 'Refactor', 'Window', and 'Help'. The address bar shows the URL 'tmi-ontology (http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-ontology#)'. The main window is divided into several panes:

- Class hierarchy:** A tree view on the left showing various classes. The selected class is '"What Should I Have Said (Done)?" Miscellaneous type'. The hierarchy includes categories like 'ANIMALS.', 'CAPTIVES AND FUGITIVES.', 'CHANCE AND FATE.', 'DECEPTIONS.', 'HUMOR.', 'MAGIC.', 'MARVELS.', 'MISCELLANEOUS GROUPS OF MOTIFS.', 'MYTHOLOGICAL MOTIFS.', 'OGRES.', 'ORDAINING THE FUTURE.', 'RELIGION.', 'REVERSAL OF FORTUNE.', and 'REWARDS AND PUNISHMENTS.'
- Class Annotations:** A pane on the right showing annotations for the selected class. It includes:
 - label:** '"What Should I Have Said (Done)?" Miscellaneous type'
 - comment:** '"Type 1696 of ATU"'
 - seeAlso:** '"Literal following of instructions about greetings."' (with a link icon)
 - isDefinedBy:** '"What Should I Have Said (Done)?" Miscellaneous type. (Including the previous Type 1696A*.) A mother tells her stupid son (man tells his wife) what he should have said (done) in a particular situation. The son follows the advice at the next opportunity, where it turns out to be inappropriate. He is punished (is told again what he should have done or said, and he follows that advice in the wrong circumstances, etc.) For example, the fool congratulates mourners and offers sympathy to a bridal couple . Cf. Types 1681A, 1681B, and 1691B.'
- Description:** '"What Should I Have Said (Done)?" Miscellaneous type'
- Equivalent To:** A plus sign icon.
- SubClass Of:** '"The Types of International Folktales Aarne-Thompson-Uther"'
- General class axioms:** A plus sign icon.
- SubClass Of (Anonymous Ancestor):** A plus sign icon.

Altogether 60.000 classes and instances. On-going multilingual extensions

Importing TMI Annotations as instances of our integrated Ontology

Cooperation with the BMBF Project: eTRAP – Digital Breadcrumbs of Brothers Grimm, Göttingen, <http://www.etrapp.eu/digital-breadcrumbs-of-brothers-grimm/>, importing Excel tables containing TMI annotations manually added to various versions of Snow-White, in various languages (Work and next slides by Lisa Schaefer, Uni Saarland).



Basic Framework

- Integration based on W3C standards: rdf, owl, rdfs, skos and skos-xl; and of Dublin Core (dc)
- dc for **annotation properties** (dc:title, dc:creator, dc:date, dc:source, dc:rights)
- skos and skos-xl for integrating the **words representing a motif** in a fairytale (skosxl:Label)

Extension of the Ontology

- Introduction of **new classes**:
 - **Tale** for specific fairy tales as representations (or instance) of an ATU type
 - **Tale collection** for the collection the specific tale is published in
 - **eTRAP_Motif** for all motifs introduced by the eTRAP-project (marked by preceding “e”) and for the terminal TMI motifs that became classes
 - Built-in **skosxl:Label** for representing the content of the cells of the Excel tables delivered by the Goettingen colleagues

Mapping from eTrap data to the Ontology (1)

1. The fairy tales

Concrete Tale as instance of class **Tale**; name = "title_author_year"

GER
Grimm_1812
VIAF: 187449723
Sneewittchen

The screenshot shows an ontology editor interface. At the top, a yellow header reads "Class hierarchy: Tale". Below this, a tree view displays the class hierarchy starting from "owl:Thing". The classes listed are: "AaTh", "ATU", "eTRAP_Motif", "Motif", "rdf:List", "skos:Collection", "skos:Concept", "skos:ConceptScheme", "skosxl:Label", and "Tale". The "Tale" class is highlighted in blue. Below the hierarchy, a purple header reads "Instances: Sneewittchen_Grimm_1812". Underneath, a list of instances is shown for the "Tale" class, including "Gold-Tree_and_Silver-Tree_Jacobs_1892", "Schneewittchen_Hahn_1864", "Sneewittchen_Grimm_1812" (highlighted in blue), "Sneewittchen_Grimm_1819", and "Sneewittchen_Grimm_1837". A red arrow points from the table on the left to the "Sneewittchen_Grimm_1812" instance in the list.

Mapping from eTrap data to the Ontology (2)

1. The fairy tales

Information as
dc annotation
properties

```
:Sneewittchen_Grimm_1812 a  
owl:NamedIndividual , :Tale ; [...]  
    dc:creator "Grimm"@en ;  
    dc:date "1812"^^xsd:integer ;  
    dc:language "ger" ;  
    dc:rights "*tba*"@en ;  
    dc:source "*tba*"@en ;  
    dc:title "Sneewittchen"@de .
```

Grimm_1812 VIAF: 187449723	Sneewittchen
----------------------------------	--------------



Annotations: Sneewittchen_Grimm_1812

Annotations +

- dc:title [language: de]
Sneewittchen
- dc:rights [language: en]
tba
- dc:creator [language: en]
Grimm
- dc:date [type: xsd:integer]
1812
- dc:language
ger
- dc:source [language: en]
tba

Mapping from eTrap data to the Ontology (3)

1. The fairy tales

Connection to **ATU type** via two inverse object properties:
represents and **isRepresentedBy**

Property assertions: Sneewittchen_Grimm_1812

■ **represents** 709

Property assertions: 709

■ **isRepresentedBy** Sneewittchen_Grimm_1812

```
:709 a owl:NamedIndividual , :Type ;  
      :isRepresentedBy  
:Sneewittchen_Grimm_1812 , [...] :Snow-  
Drop_Taylor_1823 [...] ;  
      :linkToTMI :D1311.2 , [...] .
```

```
:Sneewittchen_Grimm_1812 a  
owl:NamedIndividual , :Tale ; [...]  
      :represents :709 [...] .
```

Mapping from eTrap data to the Ontology (4)

1. The fairy tales

- Connection to fairy tale **collection** via object property **partOfCollection** and inverse **hasPart**
- Collections as instances of class **Tale_collection**; name = "author_year"

The screenshot shows a class hierarchy with **Tale** as a parent class and **Tale_collection** as a child class. Below this, a panel titled "Instances: Grimm_1812" shows a list of instances for the **Tale_collection** class: **Briggs_1970**, **Calvino_1956**, **Campbell_1958**, and **Grimm_1812**. The **Grimm_1812** instance is highlighted in blue.

The screenshot shows two panels of property assertions. The top panel, titled "Property assertions: Grimm_1812", shows an object property assertion for **hasPart** with the value **Sneewittchen_Grimm_1812**. The bottom panel, titled "Property assertions: Sneewittchen_Grimm_1812", shows an object property assertion for **partOfCollection** with the value **Grimm_1812**. Below these panels, a snippet of OWL code is visible: `:Sneewittchen_Grimm_1812 a owl:NamedIndividual , :Tale ; [...] :partOfCollection :Grimm_1812 ; [...].`

The screenshot shows a panel titled "Annotations: Grimm_1812" with a plus sign icon. It lists several annotations for the **Grimm_1812** instance: **dc:title** (language: de) with value ***tba***, **dc:rights** (language: en) with value ***tba***, **dc:creator** (language: en) with value **Grimm**, **dc:date** (type: xsd:integer) with value **1812**, **dc:language** with value **de**, and **dc:source** (language: en) with value ***tba***.

Mapping from eTrap data to the Ontology (5)

2. The motifs

- Inserting of newly introduced motifs as instances of class **eTRAP_Motif**

D1300-D1379. Magic objects effect changes in persons	
D1364.	Object causes magic sleep
D1364.13.	Cloth causes magic sleep
D1364.13.1.	Lace causes magic sleep



Class hierarchy: eTRAP_Motif

owl:Thing
AaTh
ATU
eTRAP_Motif

Instances: eD1364.13.1

For: eTRAP_Motif

- eD1364.13.1
- eD1364.32.1

Annotations Usage

Annotations: eD1364.13.1

Annotations +

rdfs:label [language: en]
"Lace causes magic sleep"

rdfs:comment [language: en]
"eTRAP added motif eD1364.13.1"



Mapping from eTrap data to the Ontology (6)

3. Connection between fairy tales and motifs

- Realized by two object properties that are inverse to each other:
 - **containsMotif** for the linking from the concrete fairy tale to all motifs that it contains
 - **appearsInTale** for the linking from a motif to all fairy tales in which it appears

```
:Sneewittchen_Grimm_1812 a owl:NamedIndividual , :Tale ;  
    :containsMotif :D1163 , :D1163 , :D1310 , :D1311 , :D1311.2 , :D1364  
, :D1364.13 , :D1364.4 , :D1364.4.1 , :D1364.9 , :D1610 , [...] ,:eZA7 ; [...].
```

```
T10 a owl:NamedIndividual , :Motif , :T ; [...]  
    :appearsInTale :Bella_Venezia_Calvino_1956 , :Sneewittchen_Grimm_1812  
, [...] :Сказка_о_мертвой_царевне_и_о_семи_богатырях_Pushkin_1833 .
```

Mapping from eTrap data to the Ontology (7)

3. Connection between fairy tales and motifs

Property assertions: Sneewittchen_Grimm_1812

Object property assertions +

- containsMotif E21.3
- containsMotif eS322.2.4
- containsMotif T16.2
- containsMotif Q211.4
- containsMotif P322
- containsMotif eD1610.37
- containsMotif W195
- containsMotif eZA101
- containsMotif E422.1
- containsMotif Q414.4
- containsMotif eR131.1.1
- containsMotif eS119

Property assertions: E21.3

Object property assertions +

- appearsInTale Sneewittchen_Grimm_1840
- appearsInTale Sneewittchen_Grimm_1843
- appearsInTale Snow-Drop_Taylor_1823
- appearsInTale Sneewittchen_Grimm_1850

Description: containsMotif

Equivalent To +

SubProperty Of +

- owl:topObjectProperty

Inverse Of +

- appearsInTale

Mapping from eTrap data to the Ontology (8)

4. Inserting the words per motif

- Realized by `skosxl:Label`
- For every motif **one skosxl:prefLabel** and several **skosxl:altLabel** for every “verbalization” of the motif in a certain fairy tale
- **prefLabel**: the TMI or eTRAP motif itself
- **altLabel**: the “verbalization” of the motif, words accessible as value of data property **skosxl:literalForm**

Mapping from eTrap data to the Ontology (9)

4. Inserting the words per motif

- labels assigned via object properties **prefLabel** and **altLabel**
- labels themselves are instances of class **skosxl:label**
- connection between **prefLabel** as basic motif and **altLabels** as “realisations” of this motif via two inverse object properties as sub-properties of **skosxl:labelRelation: verbalizes** and **verbalizedAs**

Mapping from eTrap data to the Ontology (10)

4. Inserting the words per motif

- Example: Motif T16.2
- Label assertions and their classification as skosxl:label

Instances:

For: ● skosxl:Label

- ◆ T16.2
- ◆ T16.2_s_1812
- ◆ T16.2_s_1819
- ◆ T16.2_s_1823

Annotations Usage

Annotations: T16.2

Annotations +

[rdfs:label](#) [language: en]
"Man falls in love on seeing dead body of beautiful girl."

[rdfs:comment](#) [language: en]
"Terminal motif T16.2"

Object property assertions +

- [skosxl:prefLabel](#) T16.2
- [appearsInTale](#) Sneewittchen_Grimm_1850
- [skosxl:altLabel](#) T16.2_s_1958
- [skosxl:altLabel](#) T16.2_s_1850
- [skosxl:altLabel](#) T16.2_s_1840
- [skosxl:altLabel](#) T16.2_s_1956
- [skosxl:altLabel](#) T16.2_s_1843

Mapping from eTrap data to the Ontology (11)

4. Inserting the words per motif

- Example: Motif T16.2


```
:T16.2 a owl:NamedIndividual , :Motif , :T16 ;
      skosxl:prefLabel
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-
atu-ontology/label#T16.2> ;
      skosxl:altLabel
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-
atu-ontology/label#T16.2_s_1812> ,
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-
atu-ontology/label#T16.2_s_1819> , [...]
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-
atu-ontology/label#T16.2_s_1970> ; [...] .
```

Mapping from eTrap data to the Ontology (12)


4. Inserting the words per motif

- Example: Motif T16.2
- PrefLabel

Property assertions: T16.2

Object property assertions 

Data property assertions 

 **skosxl:literalForm** "Man falls in love on seeing dead body of beautiful girl"

Annotations: T16.2

Annotations 

[dc:creator](#) [language: en]

Stith Thompson

[rdfs:comment](#) [language: en]

The original label for motif T16.2

[dc:source](#) [language: en]

TMI - Motif-Index of Folk-Literature (1955-1958)

[skosxl:verbalizedAs](#)

 [T16.2_s_1812](#)

[skosxl:verbalizedAs](#)

 [T16.2_s_1819](#)

Mapping from eTrap data to the Ontology (13)

4. Inserting the words per motif

- Example: Motif T16.2 – PrefLabel

```
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-atu-ontology/label#T16.2> a skosxl:Label ;
    skosxl:verbalizedAs
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-atu-ontology/label#T16.2_s_1812> , [...]
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-atu-ontology/label#T16.2_s_1970> ;
    skosxl:literalForm "Man falls in love on seeing dead body of beautiful girl" ;
    dc:creator "Stith Thompson"@en ;
    dc:source "TMI - Motif-Index of Folk-Literature (1955-1958)"@en ;
    rdfs:comment "The original label for motif T16.2 "@en .
```

Mapping from eTrap data to the Ontology (14)

4. Inserting the words per motif

- Example: Motif T16.2
- an AltLabel

property assertions: T16.2_s_1812

Object property assertions +

Data property assertions +

■ **skosxl:literalForm** "Prinz, Sneewittchen, nicht satt an Schönheit sehen können"@de

Annotations
Usage

Annotations: T16.2_s_1812

Annotations +

[dc:creator](#)
"eTRAP"

[rdfs:comment](#) [language: en]
Words of the tale "Sneewittchen Grimm 1812" for the motif T16.2.

[dc:source](#) [language: en]
Digital Breadcrumbs of Brothers Grimm

[skosxl:verbalizes](#)
◆ [T16.2](#)

T. SEX	
T0-T99. Love	
T16. Man falls in love with woman he sees bathing	null
T16.2. Man falls in love on seeing dead body of beautiful girl	Prinz, Sneewittchen, nicht satt an Schönheit sehen können

Mapping from eTrap data to the Ontology (15)

4. Inserting the words per motif

- Example: Motif T16.2 – an AltLabel

```
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-atu-ontology/label#T16.2_s_1812> a owl:NamedIndividual ,
skosxl:Label ;
    skosxl:verbalizes
<http://www.semanticweb.org/tonka/ontologies/2015/5/tmi-atu-ontology/label#T16.2> ;
    skosxl:literalForm "Prinz, Sneewittchen, nicht satt an
Schönheit sehen können"@de ;
    dc:creator "\"eTRAP\"" ;
    dc:source "Digital Breadcrumbs of Brothers Grimm"@en ;
    rdfs:comment "Words of the tale \"Sneewittchen Grimm
1812\" for the motif T16.2."@en .
```

Current and future Work

- Extending the “ontologization” approach to other classical classification/indexing works in the field of folk tales
 - Done for Vladimir Propp: Morphology of the tale, Leningrad 1928
- Extending to other genres
 - We started the same approach for the “36 Dramatic Situations” (Polti, Georges. The Thirty-Six Dramatic Situations, original in French)
- Interlinking all those approaches, where appropriate, towards a digital repositories of “theories” for the analysis and annotations of literary texts.

Some Links

- Propp Ontology:
<http://www.dfki.de/lt/onto/narratives/Propp/>
- TMI Ontology:
<http://www.dfki.de/lt/onto/narratives/TMI/>
- The Software Project that lead to the TTS application: <https://bitbucket.org/ceisen/apftml2repo>
- The Software Project dealing with the Propp Ontology and the detection of locations:
<https://gitlab.com/csteffens/Folktales2016>

Thanks!

- Questions?