



HAL
open science

CRYPTOZOOLOGY UNDER STUDY: How terminology deals with folklore

Ana I M Soares, João M. A. R. P. Gomes, Karen N Montero, Natacha G A Alvarenga, Sara M de Almeida, Margarida V Ramos

► **To cite this version:**

Ana I M Soares, João M. A. R. P. Gomes, Karen N Montero, Natacha G A Alvarenga, Sara M de Almeida, et al.. CRYPTOZOOLOGY UNDER STUDY: How terminology deals with folklore. Humanities Going Digital Teaching, training & research experiences (HUGOD 2023), Feb 2023, Lisbonne, Portugal. , 2023. hal-04169978

HAL Id: hal-04169978

<https://hal.science/hal-04169978>

Submitted on 24 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

CRYPTOZOOLOGY UNDER STUDY: HOW Terminology DEALS WITH FOLKLORE

Ana I. M. Soares ² | João M. A. R. P. Gomes ² | Karen N. Montero ¹ | Natacha G. A. Alvarenga ¹ | Sara M. De Almeida ¹ | Margarida V. Ramos ³

¹ Graduating in Licenciatura em Tradução, Faculdade de Ciências Sociais e Humanas, Universidade NOVA de Lisboa

² Graduating in Licenciatura em Ciências da Linguagem, Faculdade de Ciências Sociais e Humanas, Universidade NOVA de Lisboa

³ CLUNL – Centro de Linguística da Universidade NOVA de Lisboa

INTRODUCTION

This study focuses on the systematisation of the terms and concepts from the domain of cryptozoology to propose (i) a typology of cryptid entities and, (ii) to develop a corpus-based glossary in the context of Portuguese folklore. To accomplish our goals, we have compiled a domain-specific corpus from scratch in order to semi-automatically extract domain-knowledge information using a natural language processing (NLP) tool. The linguistic analysis of the terminological data, namely terms, definitions, and definitional contexts identified in the corpus, allowed us to organise some concepts of this domain according to their form and stage of evolution.

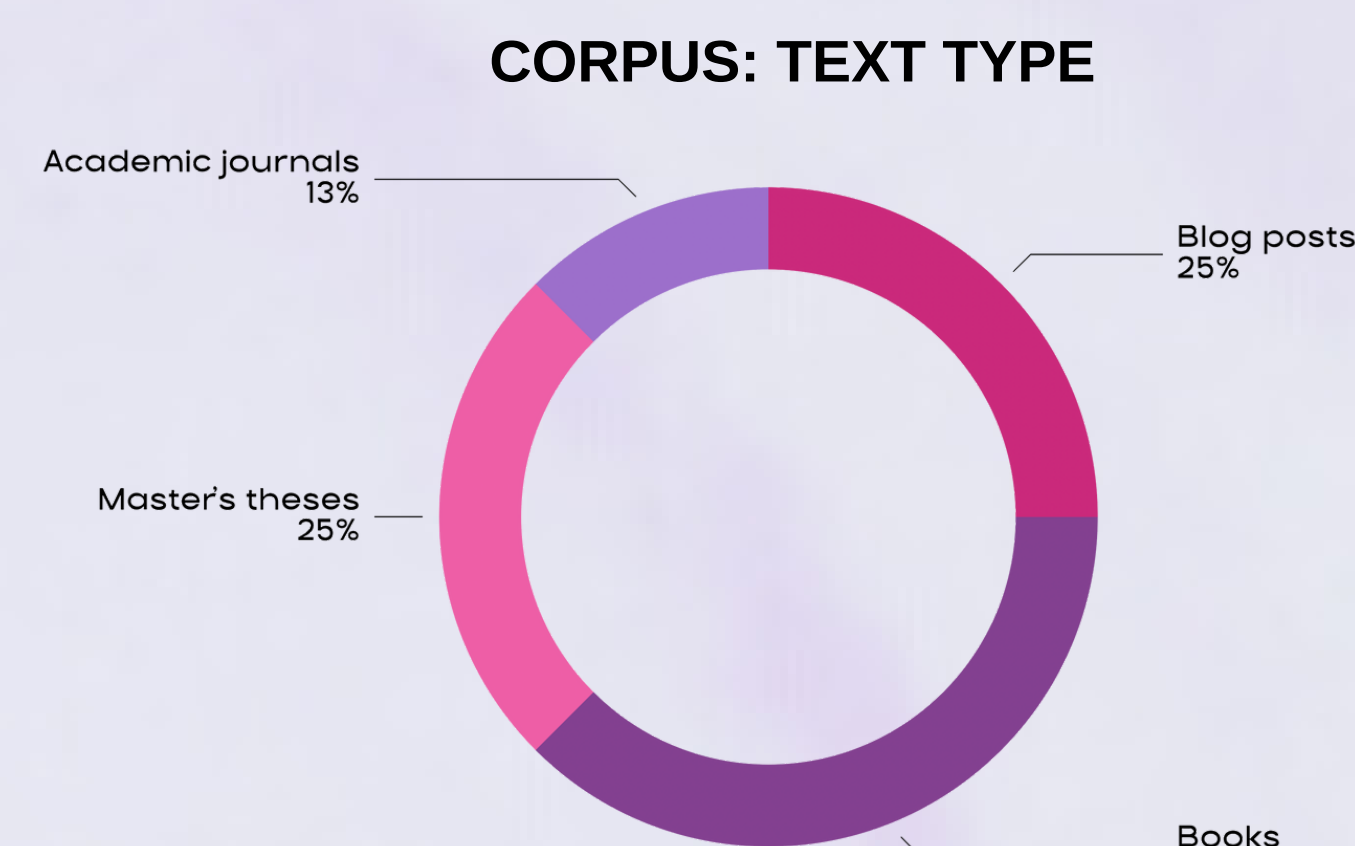
1. DOMAIN UNDER FOCUS

According to the Merriam-Webster dictionary, cryptozoology is the study and search for animals, especially legendary animals (such as Sasquatch), usually to evaluate the possibility of their existence. In other words, it is a field of interest that studies unknown, legendary or extinct animals whose current existence is disputed or unsubstantiated, especially those popular in folklore. These entities are referred to as cryptids, a term coined by the subculture. Not following the scientific method, it is considered a pseudoscience by mainstream science and is neither a branch of zoology nor folklore studies. Cryptozoology was founded in the 1950s by zoologists Bernard Heuvelmans and Ivan T. Sanderson.

2. METHODOLOGY

I. Corpus

We first compiled a domain-specific corpus with texts written in Portuguese. The typology of texts includes books, master's theses, academic journals, and texts copied from Web sites, all in digital format and available without paywalls, from sources as recent as 2021 to sources dating back to 1882.



II. Terminological data extraction

In order to semi-automatically explore the corpus, we used AntConc (NLP tool) to identify and extract the most frequent terms co-occurring with contextual definitions – i.e. definitions found in context. The criteria for selecting terms were (i) the highest frequency (see Table 1) and (ii) the possibility of extracting definitions of the concepts they designate (see Table 2). Therefore, for the purposes of this study, the terms whose concepts are not defined or explained in the corpus were not considered.

Candidate term	Frequency	Candidate term	Frequency	Candidate term	Frequency	Candidate term	Frequency
mouras	399	sereias	53	caca	8	mouras construtoras	3
diabo	361	dragão	37	basílico	6	balborinho	2
bruxas	186	alicórnio	15	corritório	5	sereia invertida	2
mourinhos	135	homem das sete dentaduras	15	hira	5	hipogrifo	1
mouras encantadas	114	grifo	14	fadas marinhas	4	maira fiandeira	1
lobisomem	83	almas penadas	13	centauro	3	mouras tendeliras	1
fadas	76	diaba	9	esfinge	3	sereia caprina	1

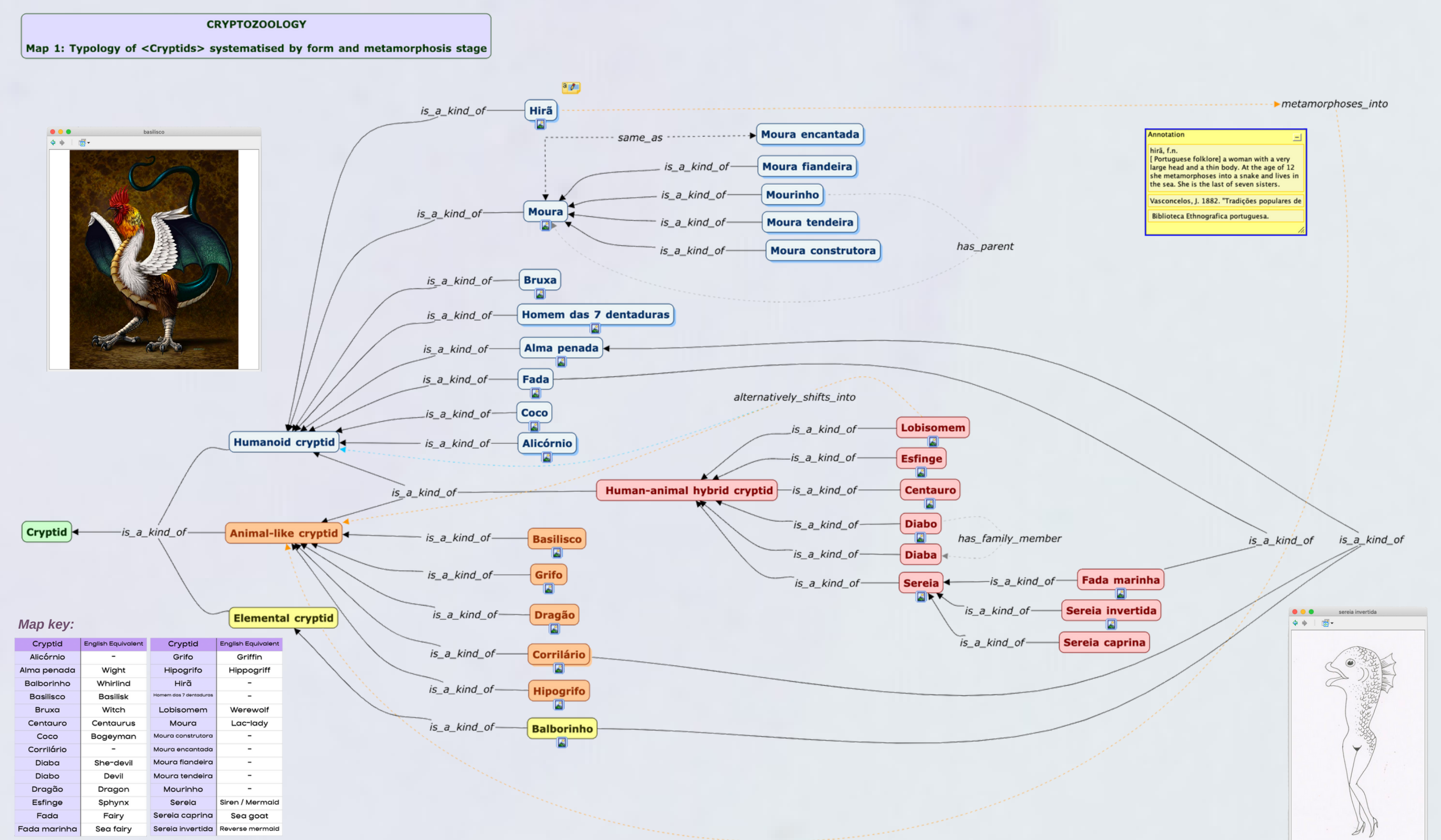
Table 1: Candidate terms extracted from the corpus

III. Terminological data systematisation

Analysing the terminological data extracted from the corpus, namely definitions and definitional contexts, i.e., contexts rich in knowledge information useful for writing definitions, allowed us to organise some domain concepts (see Table 2). The concepts were hierarchically systematised using CmapTools (see Map 1) according to the type of form (and metamorphosis if applicable) of the entities they denote, such as <humanoid cryptid>, <human-animal hybrid cryptid>, <animal-like cryptid>, and <elemental cryptid>. Some sub-types of these concepts are related to each other through non-hierarchical associative relations (see ISO/FDIS 1087, 2019 (E)), depending on their family relationship, shape-shifting property (see <Werewolf> in Map 1) or form evolution over time (see <Hirã> in Map 1).

Candidate term	Context 1	Context 2	Context 3
lobisomem [werewolf]	essa peessoa transforma-se depois em lobisomem , tomando por algum tempo a figura de um lobo , dum jumento , dum bode ou de outro qualquer animal	Ainda na Epopeia de Gilgamesh poderia estar o primeiro relato do que viria a dar origem à lenda do lobisomem, homem que se transforma em lobo .	pode aparecer o lobisomem sob a forma de qualquer animal, por exemplo o lobo , o cão , o bode , etc., e esta forma não é indiferente para a pessoa que tem a desdita de o encontrar, pois conforme é crença o pior lobisomem e o que faz piores malefícios é o que tem a forma de homem com pés de cabrito ou cavalo
sereia [siren]	mulher atraente com cauda de peixe e que encanta suas vítimas , tanto pela beleza quanto pelo seu canto	encantados pelo "canto das Sereias", os incautos marinheiros precipitavam-se no mar para irem ao seu encontro; estas, [...] não resistiam a devorá-los , cometendo assim o crime da antropofagia .	já no Séc. XIX, as Sereias foram definitivamente reabilitadas pela literatura infantil dinamarquesa, tornando-se personagens de boa índole e fazendo esquecer a natureza perversa que originalmente tiveram.
fada [fairy]	as fadas tomadas em si mesmas são apenas mulheres gentis ; o que é fantástico é a natureza quando obedece às fadas	mulheres que tem o poder de transformar tudo, e de se transformarem a si, por meio da varinha do condão	as mouras podem facilmente ser associadas a fadas, pois surgem encantadas em serpentes de longos cabelos louros[...] formosas e sedutoras. [...] já semelhança das fadas que concedem desejos , as mouras concedem recompensas a quem as desentencar
fada marinha [sea fairy]	Nos Açores[...] ainda existe a crença nas Fadas marinhas , ou sereias , que vêm pentear-se à praia.	Gil Vicente allude também às fadas marinhas , que parecem ser o mesmo que as Sereias	Gil Vicente, o poeta que mais se inspirou do maravilhoso popular, traz nos seus autos as fadas marinhas . E aquelas fadas que tem as ribeiras de verde pintadas

Table 2: Examples of definitional contexts extracted from the corpus



Map 1: Proposal for a typology of <Cryptids> according to their form and evolution

3. TERMINOLOGICAL RESOURCE

With the outcomes of our terminological work, we developed a glossary using Lemony, an online dictionary publisher (see Fig. 1). The terms included in this resource are those we have found co-occurring with definitional contexts. Each entry of the glossary is provided with contexts extracted from the corpus, related terms, and equivalents in English and Spanish sourced from online dictionaries.

lobisomem s. m.

[grafia do séc. XIX **lobishomem**]

homem que se transforma em lobo

fonte *Suriadakis de Melo, J. (2015). O Bestiário Fantástico de Murilo Rubião [Tese de Mestrado, Universidade do Estado do Amazonas]*

exemplo (...) *essa pessoa transforma-se depois em lobisomem, tomando por algum tempo a figura de um lobo, dum jumento, dum bode ou de outro qualquer animal (...)*

exemplo (...) *pode aparecer o lobisomem sob a forma de qualquer animal, por exemplo o lobo, o cão, o bode, etc., e esta forma não é indiferente para a pessoa que tem a desdita de o encontrar, pois conforme é crença o pior lobisomem e o que faz piores malefícios é o que tem a forma de homem com pés de cabrito ou cavalo (...)*

fonte *Pedroso, C. (1988). Contribuições para uma mitologia popular. Lisboa: Publicações Dom Quixote*

termos relacionados **lobisomem fêmea**

fonte *Pedroso, C. (1988). Contribuições para uma mitologia popular. Lisboa: Publicações Dom Quixote*

imagem ilustrativa

em **werewolf**

fonte *Cambridge Dictionary. <https://dictionary.cambridge.org/dictionary/english/werewolf>*

cast **hombre lobo**

fonte *Diccionario de la Real Academia Española. <https://dle.rae.es/hombre?m=form#CpCYH7A>*



Fig. 1: The glossary's entry for "lobisomem", on Lemony

Fig. 1 represents the entry of "lobisomem" [werewolf]. The definition is our proposal, based on the definitional contexts we have found in the corpus (see Table 2).

4. FINAL REMARKS

(1) While searching for equivalent terms in English and Spanish, we have observed that some terms designating <Cryptids> are common to different cultures, while others may refer to a slightly different concept given the description of different characteristics in the contextual definitions, or to a completely different concept when designated by an apparent loanword, e.g. "alicórnio" and "unicorn". As far as we have been able to observe, while most of the concepts in this domain have remained unchanged over time, e.g. <Sphinx>, some have stabilised, e.g., <Werewolf>, a concept that was once defined with different combinations of characteristics as in <Wolf>, <Donkey>, <Goat>, among others, thus with different descriptions of human-animal forms. The same happens with the concepts <Mermaid>, <Siren> and <Sea Fairy>, the latter having evolved from the medieval tradition as a malevolent entity to a gentle one according to 19th-century Danish literature. These observations were made possible by a comparative analysis between contextual definitions found in contemporary texts (20th and 21st century) and older texts (19th century). (2) Cryptozoology is not an exact science with well-defined boundaries, which rises some difficulties to organise concepts whose characteristics vary inter- and intra-culture. Some doubts remained unsolved and require more study, such as the case of "alicórnio" and "unicorn": are these true loanwords? (3) With this study, we intended to demonstrate that terminological methods are useful for systematising domain knowledge information, even in the context of folklore - a non-scientific field of interest in which some concepts are perceived differently (or not) across cultures.

Acknowledgements: This study was carried out within the framework of the terminology course unit of two bachelor programmes (1st cycle) at the NOVA FCSH: Translation, and Language Sciences.



Co-funded by the Erasmus+ Programme of the European Union