

# VAST ANNOTATION METHODOLOGY NARRATIVES



# Preface

**Authors:** Sara Graça da Silva (NOVA), Alba Morollón Díaz-Faes (NOVA), Diogo Marques (NOVA), Theodore Grammatas (NKUA), Maria Dimaki Zora (NKUA), Kaiti Diamantakou (NKUA), Giulia Giannini (UMIL), Francesco Barreca (UMIL)

**Contributors:** Georgios Petasis (NCSR-D), Carla Sofia Ribeiro Murteira (NOVA)

**Design:** Galateia Iatraki, Maria Dagioglou

**Project: VAST: Values Across Space & Time**

**CC BY 4.0**

**2023 VAST**

<https://www.vast-project.eu/>

Reproduction is permitted provided the source is acknowledged.

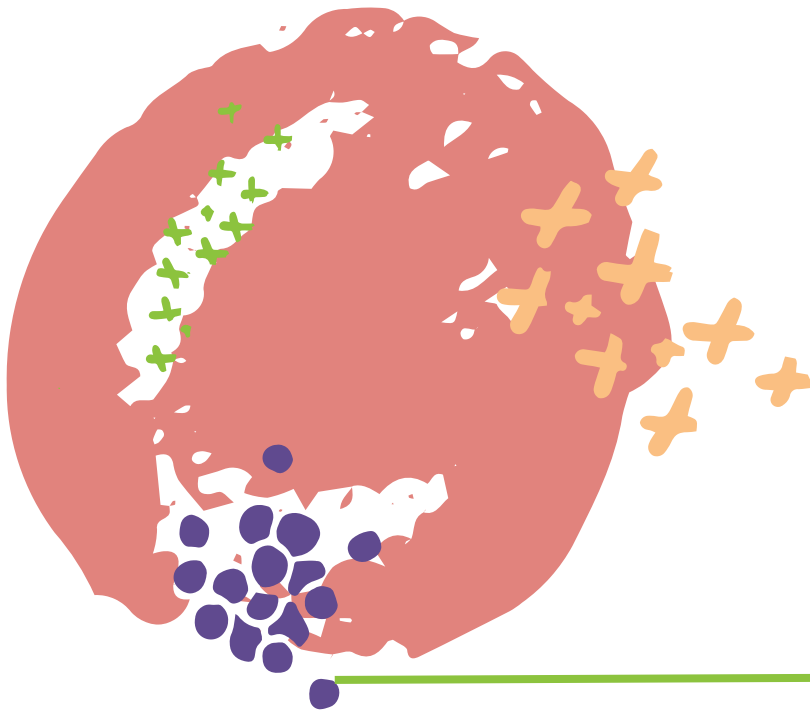


The VAST project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement **No 101004949**. This educational guide reflects only the view of the authors and the European Commission is not responsible for any use that may be made of the information it contains.

<sup>1</sup> Folklife and Ethnological Museum of Macedonia-Thrace, Greece

# Table of contents

<b>What does annotation mean in the VAST project?</b>	p. 1
<b>Annotation methodology</b>	p. 3
<u><b>1. Before you start</b></u>	p. 4
<u><b>2. Working with labels</b></u>	p. 4
<u><b>3. Annotation Steps</b></u>	p. 7
<u><b>4. Important rules to follow</b></u>	p. 9
<u><b>5. Q&amp;A</b></u>	p. 11



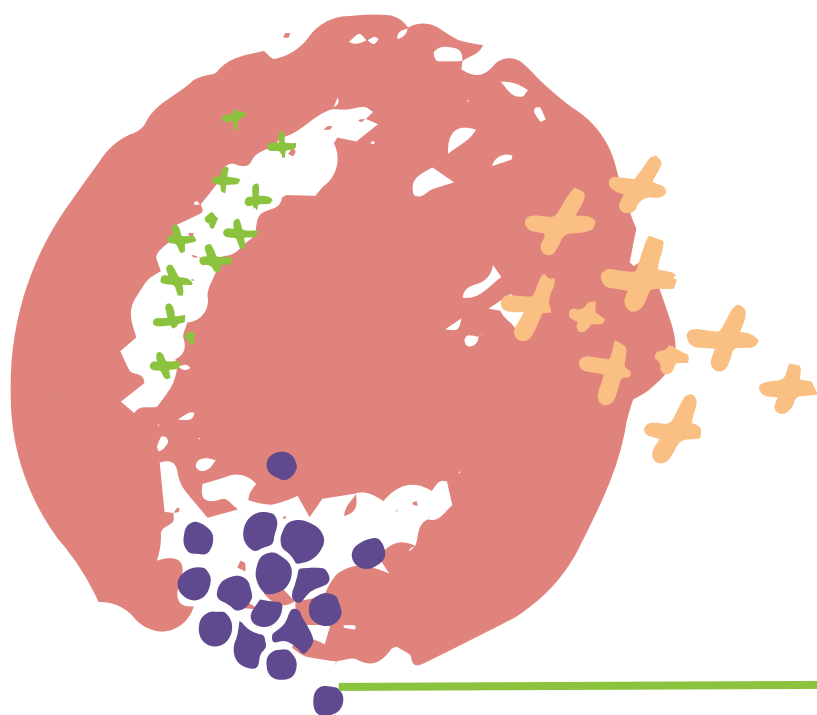
**What does annotation mean  
in the VAST project?**

The VAST project (Values Across Space and Time - <https://www.vast-project.eu>) has studied the transformation of moral values across space and time, with particular emphasis on core European values that constitute the pillars of European society. To this end, the project has been divided into three distinct pilots that focus on different foundational texts:

- Pilot 1: Values in ancient Greek tragedies;
- Pilot 2: Values in texts of 17th century about natural philosophy;
- Pilot 3: Values in European folktales.

One way to identify values in these texts is through annotation. “**Annotation**” refers to the **matching of defined labels to text chunks, or segments**. In VAST, the annotation is done manually by an annotator through the use of VAST annotation tool, i.e. the Ellogon Annotation Platform (<https://annotation.vast-project.eu>)

What you are asked to do, as an annotator, is to read the given texts, identify values, and tag them using the defined labels. The following section describes the annotation methodology.



## Annotation methodology

# 1. Before you start

Annotation can be a difficult task. To ensure optimal performance, please follow these recommendations:

- Ensure you are in a quiet and calm environment with minimal distractions.
- Read the whole text before you annotate.
- Familiarise yourself with the labels provided.
- When annotating, use your first impression and do not try to determine whether there is a single correct label to match to a text segment.



## 2. Working with labels



The labels available in the tool are divided into **three categories** for the annotator to choose from. **Do not attempt to exhaustively use all labels**, just try to match segments of text to labels as you read by using your intuition.



**Category I:**  
Key/Main/Concepts/  
Values/Ideas

*These labels correspond to concepts and values recognised as fundamental by the European Union.*

**I. KEY / MAIN CONCEPTS / VALUES / IDEAS**

Democracy	Dialogue	Equality	Freedom
Human Dignity	Human Rights	Justice	Peace
Progress	Reason		



**Category II:**  
Expanded/Variou  
Content

*These labels correspond to an expanded list of relevant concepts or values that are commonly found across the texts under consideration.*


Cooperation	Curiosity	Democratic State and Equality	Demonstrable Truth
Empathy	Evidence	Evolution	Free Thinking
Experimentation	Female Emancipation and Autonomy	Female Speech	Generosity
Freedom of Expression	Gender Equality	Integrity	Kindness
Honesty	Ingenuity	Love/Marital Love	Loyalty
Knowledge	Love among Siblings	Piety	Power of Motherhood
Non-Discrimination	Objectivity	Reason/Political Corre...	Reliability
Punishment	Purity	Reward	Rule of Law
Research Freedom	Respect of Majority Decisions	Social Altruism and Self-sacrifice	Social/Political Functio...
Science for Public Good	Tolerance	Truth	
Solidarity			
Validation			



**Category III:  
Opposite Concepts**

*The third category contains concepts in opposition, which you should use only when you find two binary concepts or values in direct conflict. This means that whenever a concept appears alone, without its oppositional concept, you should not use these labels.*

Appearance vs Reality	Arrogance vs Modesty
Autonomy vs Heteronomy	Clarity vs Ambiguity
Deceptiveness vs Honesty	Demonstration vs Revelation
Evidence vs Authority	Exclusion vs Respect of Minorities and Marginalised Communities
Faithful vs Unfaithful	Female Oppression and Devaluation vs Female Emancipation and Autonomy
Fate/Destiny vs Individual Law	Gender Equality vs Submission
Freedom of Speech vs Censorship	
Freedom vs Slavery	



### 3. Annotation steps



Values can be **explicit** or **implicit** in a text, and you should try to look for both kinds. To identify them, the annotation process is divided into **two distinct steps**.

#### 1. Annotate explicit references to values

First, you should look for **explicit or literal references** to labels in-text. Very often we find words or terms in the text that correspond to a label directly, or that are derived from that word/label. What you need to look for in this step is the presence of relevant words or terms in the text.

##### **Example 1:**

*In the following excerpt from Little Snow White by Brothers Grimm (1857) we find the explicit reference to the label "**Punishment**" from Category II:*

*Afterwards, as they came out of the church, the older one was on the left side, and the younger one on the right side, and then the pigeons pecked out the other eye from each of them. And thus, for their wickedness and falsehood, they were **punished [label: Punishment]** with blindness as long as they lived.*

##### **Example 2:**

*Another example is found in The Jew in the Thorns by Brothers Grimm (1857), where we find an explicit reference to the label "**Reward**":*

*As he was being led away, the Jew screamed after him, "You good-for-nothing. You dog of a musician. Now you will receive your well earned **reward. [label: Reward]**"*

**In this case, the word in brackets should be annotated with "Reward", even though context reveals this reward is actually a punishment.**



## 2. Annotate implicit references to values

The second step is more subjective, as you need to annotate text segments according to their **implicit connection to a label**. After completing step one, you will be more familiar with the text, which will help you in its interpretation. However, you are still encouraged to annotate following your first impressions, without trying to determine whether a label is the correct one.

### **Example:**

The following example is taken from Brothers Grimm's *Fitcher's Bird* (1857). Following the protagonist's journey, who is now free of her husband's control, we annotate the relevant segment of the last sentence with the value from Category I, "**Freedom**":

*When the man returned home he immediately demanded the keys and the egg, and when he was unable to detect any trace of blood on them, he said, "You have passed the test. You shall be my bride."*

*He now had no more power over her [label: Freedom] and had to do whatever she demanded.*

### **NOTE:**

Note that part of the excerpt from Brothers Grimm's *The Jew in the Thorns* (1857), which we annotated with "Reward" in the previous step, is annotated with "Punishment" at this stage, since we are taking into consideration the context:

*As he was being led away, the Jew screamed after him, "You good-for-nothing. You dog of a musician. Now you will receive your well earned reward. [label: Punishment]"*

## 4. Important rules to follow



A text segment can often include several components which must be annotated separately. Other times, it is better to annotate longer segments. The following rules will help you **identify the boundaries** of your annotations.

### 1. Relevance rule

You should include **all words that are relevant to identify a label**. This means that all relevant content should be included in the annotation, but also that all the unnecessary elements should be left out. If you are unsure about whether to annotate a segment or not, please leave it out.

#### Example:

The example below is taken from Brothers Grimm's *Juniper Tree* (1857):

"Mother," said the little daughter, "is brother not to have one too?" This made the woman angry,

Labels: Equality, Empathy

Labels: Generosity vs Greed, Right vs Wrong

In the excerpt, we find a short exchange between mother and daughter. Marlene, the daughter, is kind and wants to share her food with her step-brother. We annotate her compassionate question, "is brother not to have one too?", with the label from Category I, "Equality", and the label from Category II, "Empathy".

The villain in the story is her mother, who always favors her biological daughter over her adopted son. We annotate both Marlene's question and the mother's reaction with Category III labels, "Generosity vs. Greed" and "Right vs. Wrong".

## 2. Continuity rule

Labels must only be used on **continuous segments**. Two segments related to the same label that have unrelated, or less relevant words in the middle must be annotated separately.

### Example:

The following example is taken from *The Jew In the Thorns* by Brothers Grimm (1857) and it contains several **implicit** and **explicit** references to the label of "Justice", marked with brackets. Instead of labelling the whole segment with "Justice", annotators should select the relevant segments individually.

*"Judge [Label: Justice] sir," he said, "Oh, woe is me! See how a godless man has robbed me and abused me on the open road. A stone on the ground would feel sorry for me. My clothes are ripped into shreds. My body is pricked and scratched to pieces. And what little I owned has been taken away with my purse -- genuine ducats, each piece more beautiful than the others. For God's sake, let the man be thrown into prison. [Label: Justice] "*

*The judge [Label: Justice] asked, "Was it a soldier who cut you up like that with his saber?" "God forbid," said the Jew. "He didn't have a naked dagger, but rather a blowpipe hanging from his back, and a fiddle from his neck. The scoundrel can easily be recognized."*

*The judge sent his people out after him. [Label: Justice] They found the good servant, who had been walking along quite slowly. And they found the purse with the money on him as well.*

## 3. Overlapping labels rule

When you identify several relevant labels in the same segment, **select them all**.

### Example:

The following example is taken from *The Girl Without Hands* by Brothers Grimm (1857). In it, the annotator identifies two separate labels from Category III. The excerpt describes the encounter between a faithful, pious girl and the devil, the personification of evil. As such, it is labelled with the labels of "Good vs. Evil" and "Faithful vs. Unfaithful".

*The miller's daughter was a beautiful and pious girl, and she lived the three years worshipping God and without sin. When the time was up and the day came when the evil one was to get her, she washed herself clean and drew a circle around herself with chalk. The devil appeared very early in the morning, but he could not approach her.*

*[Labels: Faithful vs. Unfaithful & Good vs. Evil ]*

## 5. Q&A



### Q: What do you mean by label?

A: We use "labels" to denote a set of predefined words that annotators need to match to parts of the text by means of highlighting and selection. These labels, in our project, refer to values, or concepts relating to values.

### Q: Can a word with the same root as a label also be an explicit reference to a value?

A: Yes, as it can be seen in the example in Step 1 of the Annotation steps (the verb "punished" is tagged with the label "Punishment"). Other examples could be the words "judge" or "justification", which would be tagged with the label "Justice".

### Q: Can I add labels that I don't see in the list?

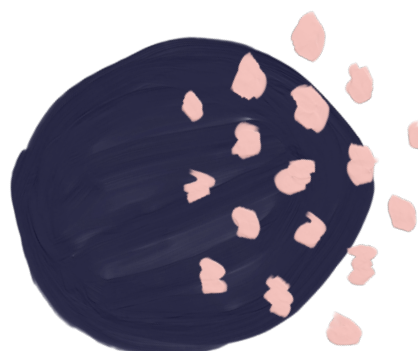
A: Yes, you can add your own labels on the Ellogon tool. However, use this with care, as it dilutes the focus of the activity and might result in annotations that do not contribute to the end goal of finding values present in the texts.

### Q: Could an identified explicit reference contradict the implicit reference in the same text chunk?

A: This is possible and its identification during annotation is particularly valuable.

### Q: Can I use an opposite concept from Category III when I find only one of its two elements?

A: No. Labels in this category are only to be used when they appear contiguously. That means you can only choose it when both concepts appear in the same segment side by side. For example, to select the opposite concepts "Arrogance vs Modesty", both labels need to be present. If there is only a reference to "Arrogance" in the text chunk, this particular label should not be chosen.





[HTTPS://WWW.VAST-PROJECT.EU/](https://www.vast-project.eu/)

CC BY 4.0

2023 VAST

Reproduction is permitted provided the source is acknowledged.



The VAST project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement **No 101004949**. This educational guide reflects only the view of the authors and the European Commission is not responsible for any use that may be made of the information it contains.