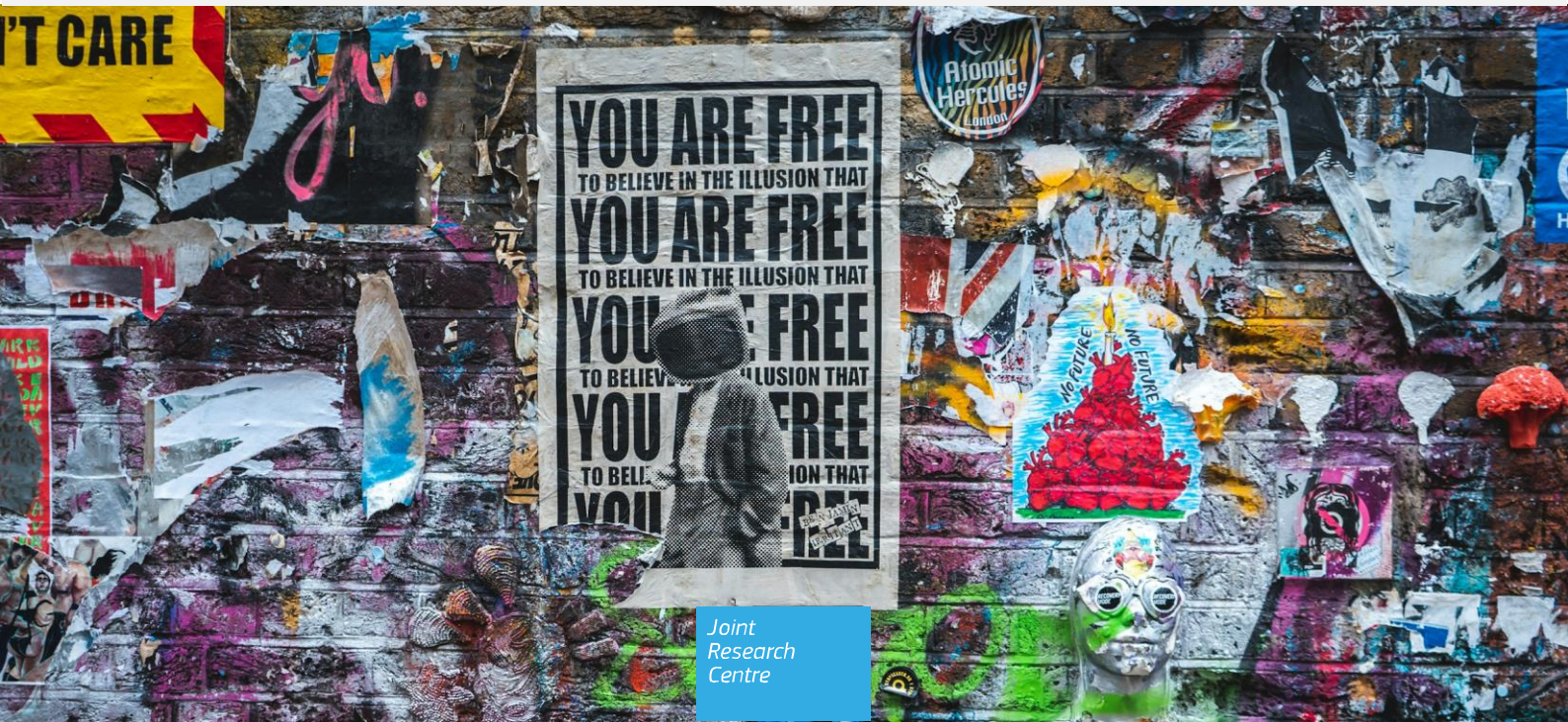
ISSN 1831-9424

European Commission

Values in News and Political Manifestos: Annotation Guidelines

Reitis-Münstermann, T., Schulze Brock, P., Scharfbillig, M., Stefanovitch, N., De Longueville, B.

2024



Joint Research Centre

This document is a publication by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The contents of this publication do not necessarily reflect the position or opinion of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contact information
Name: Reitis-Münstermann
Email: JRC-TMA-CC@ec.europa.eu

EU Science Hub
https://joint-research-centre.ec.europa.eu

JRC137286

EUR 31906 EN

PDF ISBN 978-92-68-14347-6 ISSN 1831-9424 doi:10.2760/7398 KJ-NA-31-906-EN-N

Luxembourg: Publications Office of the European Union, 2024

© European Union, 2024

The reuse policy of the European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (https://creativecommons.org/licenses/by/4.0/). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

For any use or reproduction of photos or other material that is not owned by the European Union permission must be sought directly from the copyright holders.

Cover page illustration, © Sam Fry / unsplash.com

How to cite this report: European Commission, Joint Research Centre, Reitis-Münstermann, T., Schulze Brock, P., Scharfbillig, M., Stefanovitch, N., and De Longueville, B., Values in News and Political Manifestos: Annotation Guidelines, Publications Office of the European Union, Luxembourg, 2024, https://data.europa.eu/doi/10.2760/7398, JRC137286.

Contents

Abstract 3

Acknowledgements 4

1 Introduction 5

2 Annotation Task Overview 6

2.1 Labels 6

2.2 Data selection 6

2.3 Annotation procedure 7

2.3.1 Participants' role 7

2.3.2 Onboarding of annotators 7

2.3.3 Annotation phase 8

2.3.4 Curation 9

2.3.4.1 Part 1 9

2.3.4.2 Part 2 9

3 Taxonomy and examples 11

3.1 First level of analysis - Values Categories 11

3.2 Second level of analysis: Attainment of the value 12

3.3 Examples for Values and their Attainment Label 13

SELF-DIRECTION-THOUGHT: 13

SELF-DIRECTION-ACTION: 14

STIMULATION 14

HEDONISM 14

ACHIEVEMENT 15

POWER – RESOURCES 15

POWER - DOMINANCE 16

FACE 17

SECURITY - PERSONAL 18

SECURITY – SOCIETAL 18

CONFORMITY - RULES 19

CONFORMITY - INTERPERSONAL 20

TRADITION 20

HUMILITY 20
BENEVOLENCE – DEPENDABILITY 21
BENEVOLENCE – CARING 21
UNIVERSALISM – CONCERN 22
UNIVERSALISM - NATURE 22
UNIVERSALISM – TOLERANCE 23
3.4 Precision of span 24
3.5 General guidelines for annotation and usage of taxonomy 24
3.6 General guidelines for curation 28
4 Inception - the Annotation Tool 29
4.1 Annotation 29
4.2 Curation 34
5 Conclusions 39
References 40
List of figures 41
List of tables 42

Abstract

This guide provides an overview of the tasks of annotating news articles and political manifestos with i) one of 19 personal values and ii) the attainment level of the value in question. Values are important drivers of people's beliefs, attitudes and behaviours and important for understanding citizens' political nature. For the two annotation levels, the document provides extensive descriptions and examples of the taxonomy used. The guide is intended to be used by annotators and curators working on an annotation platform (Inception) for developing a corpus of annotated text that can later be used to train values detection algorithms. The guide has been used in the context of the project valuesML to annotate over 2500 documents in 9 different languages by around 70 values scholars.

Acknowledgements

We would like to thank Murat Ardag, Sharon Arieli, Giuliano Bobba, Ella Daniel, Henrik Dobewall, Hester van Herk, Johannes Karl, Anna Krasteva, Thomas Peter Oeschger, Georgios Petasis, Vladimir Ponizovskyi, Luana Russo, Antonella Seddone, Joanne Sneddon and Aurelia Tamo for their support in developing this guide.

Authors

Theresa Reitis-Münstermann, Paula Schulze Brock, Mario Scharfbillig, Nicolas Stefanovitch, Bertrand De Longueville

1 Introduction

Text mining offers the possibility to extract insights and knowledge captured in human writing. Today, various forms of text are produced at an ever-higher speed and reading them all is impossible. This makes the use of automated and semi-automated methods to summarise and aggregate the content necessary.

There are various sources of relevant text such as on-line and print media and political manifestos. News articles, which continuously report both facts and opinions in cross-lingual and cross-cultural settings, shape and reflect how societies are making sense of reality. Political manifestos are intended to represent political priorities, legitimise policy proposals and persuade people. In order to understand these sources, after decoding the information level (who, what, where, when), it becomes essential to carry out in-depth analysis of how this information is presented across different countries, parties and media, appealing to various ways in which people deal with information, for example relating to emotion, values or trust (Piskorski et al. 2022).

The Joint Research Centre of the European Commission is currently exploring an extension of its analytical capacity of text analysis with an automated detection of values. This values detection serves as a complement to the dictionary approach already implemented (Ponizovskiy et al. 2020; Scharfbillig et al. 2022). The companion paper to the this guide explains the dataset that is the result of these annotations from the project valuesML[1], see Scharfbillig et al. (2024).

[1] https://knowledge4policy.ec.europa.eu/projects-activities/valuesml-unravelling-expressed-values-media-informed-policy-making_en

2 Annotation Task Overview

2.1 Labels

The task consists of annotating news articles and political manifestos for the 19 refined human values as identified by Schwartz (1992) and Schwartz et al. (2012). Human values represent guiding principles in people's lives. Basic human values are recognised across all cultures and are therefore useful to understand the principles being conveyed in various national news articles and political manifestos.

Annotations are to be made at the sentence level. That means that value labels are to be identified by looking at strings of words within one sentence. In exceptional cases, if a sentence is very short and makes sense only through the preceding or following sentence, annotation beyond the sentence boundary may be used.

It is assumed that one value is expressed through a sequence of words. However, it may occur, especially in longer sentences, that multiple values and thus multiple labelling are discernible in the same string of words within a sentence. If multiple values are identified for the same string of words, annotators should first try to identify the prominent value, which is most strongly enforced by the text context. Taking the sentence structure into account (main versus subordinate clauses) might be helpful. However, if annotators have difficulties deciding whether one value is more dominant than another, this could indicate that two values in the span are equally prominent. As a rule of thumb, annotators should annotate both values in such a case. In contrast, annotators should NOT start annotating all values they think might be relevant. Only when you are really conflicted in which main values are alluded to, should you choose more than one.

Additionally, each values reference has to be described as "(partially) attained", "(partially) constrained" or "not sure, can't decide", depending on how what is described in the sentence, relates to the attainment or constraint of the referenced value. Thus, this second layer captures the relationship between what is said or described in the sentence and the attainment of the expressed value.

2.2 Data selection

The annotation campaign revolves around two types of text data: a) news articles, and b) political manifestos.

Therefore, news articles from EU member states, Turkey, and Israel were collected in Bulgarian, Dutch, English, French, German, Greek, Hebrew, Italian and Turkish. For each language, articles were selected in such a way to: (a) cover both mainstream media and media outlets considered as sources of spreading disinformation, and (b) revolve around hot topics deemed as relevant in the EU context as well as typically values-expressive topics. These topics are climate change, economy/industry, education/role of science, renewable energy, health, immigration/refugees, national defence, role of the EU, and working conditions. Furthermore, only articles of a word length of roughly 400 to of 800 words were considered.

Similarly, extracts from party manifestos of 800 words or less were selected to fit the topics of the news articles and languages. Unlike the selection of news articles, manifesto extracts include the United States as source. In addition, only the most recent manifestos available from up to six of the central parties in each national electoral system were selected. The centrality of parties was

measured by their election results. Due to the high number of central parties in Israel and Bulgaria, more than six parties were considered here.

Finally, in order to have most recent text data a period from 2019 to 2023 was considered for both text types.

2.3 Annotation procedure

The annotation procedure, described below, is organised in order to get the highest possible annotation quality.

2.3.1 **Participants'** role

Annotators annotate assigned documents from the selected text data on the platform Inception and discuss issues they face in language group meetings (LGs) on a weekly basis to align on common procedures within the LG.

Language leads (LLs) (and curators) of each language group organise these weekly meetings and support the annotators by discussing difficulties, experiences and specific annotation cases within the meetings. In weekly core meetings between all LLs and the JRC value lead, issues and potential solutions raised and proposed in the LGs are first reported and discussed. Then, the common procedure for all languages is defined and in form of the annotation guidelines updated and shared with all LGs.

Curators may support the LLs in bridging from the intra-language-level to the inter-language-level and vice versa. Moreover, they curate (i.e., consolidate) the annotations within the language group and create a final annotated version of each document.

The weekly core meeting is led by the JRC values lead taking ultimate decisions on the common procedures if no consensus can be found.

The Inception team support the annotation campaign in providing data and enabling the use of the annotation platform Inception.

2.3.2 Onboarding of annotators

Training material for annotators is provided by the Inception team and the JRC value lead to get familiar with the exercise and to sharpen the understanding of the different values. The material will be introduced in a kick-off LG meeting by LLs. The training material includes:

- the annotation guide,
- a training tool with flashcards,
- an introduction video to the Annotation platform Inception, and
- test projects for annotators on Inception.

Annotators are asked to

- Read the annotation guide including description and examples for each Schwartz value and attainment labels.
- Access the training tool and practice annotation of values and attainment labels with flash-cards in various difficulty levels. The correct option out of four alternatives must be chosen.
- Watch the introduction video to Inception to understand how to annotate on the platform.

- Make test annotations within the *Test Project for Annotators* on Inception as agreed with the LG. The project includes
  1) the State of the Union speech by the President of the European Commission Ursula Von der Leyen translated into the respective language covered by the annotation campaign and
  2) four value-expressive news articles (in English only).

Any questions that arise should be discussed with the respective LLs during the LG meetings.
After successfully working through the training material annotators receive access to their language group's *ValuesML Annotation Project 2023* on Inception with the first batch of actual documents uploaded, as provided to the Inception team by the LLs. Annotators are ready to annotate all the actual documents to which they have been assigned prior.

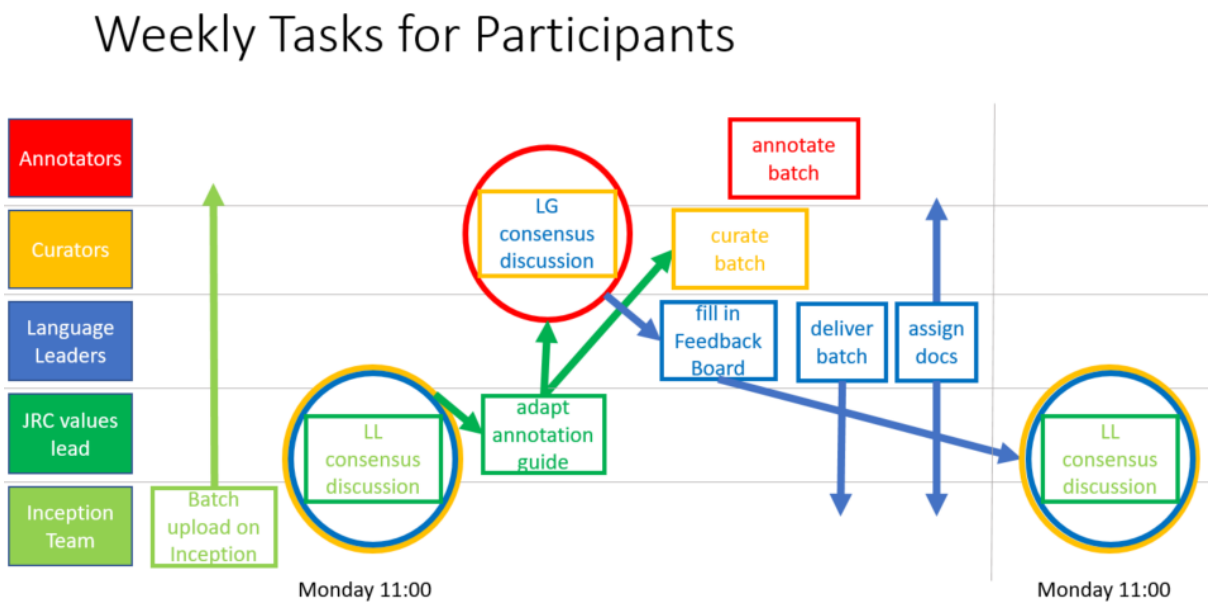
2.3.3 Annotation phase

At least two annotators annotate each document with varying combinations of annotators for each document.

On a weekly basis:

- LLs provide a batch of selected text documents for the following week [not applicable if the whole batch has already been provided at the start of the campaign], which the Inception team feeds in the Inception projects.
- The Inception team provides an example sheet including a randomised assignment list that defines which annotator annotates which document based on the file name. LLs / curators provide a final list with the documents to their LG annotators and the Inception team. Since LLs are closely working with the annotators LLs can keep a better overview of the actually annotated documents, especially in case of annotators being on annual/sick leave during some weeks. Therefore, they might need to adapt the sheet according to this.
- The annotators annotate the assigned documents and close the documents on Inception when they have finished.

Figure 1. Overview of the weekly tasks of all participants during the annotation phase



- Once all annotations for one document are finished the LLs/curators compare the annotations of the LG's annotators (highlighting potentially conflicting annotations) and share the insights, so that the LG can reach a common understanding of the task at hand.
- The LLs/curators hold a meeting with their language group. They can share the curation view on Inception and point out lines with high disagreement to align and resolve general issues, discuss specific annotation cases, resolve problems, and to share experiences. The annotators may raise their questions. Here they also share the solutions agreed on during core meetings between all LLs and the JRC value lead. If agreed on within the LG, these meetings can be recorded.
- LLs/curators comment the issues raised in the LG meeting in the *Feedback Board* before the core meetings and then share discussion points with the other LLs to find and agree on a common method.
- The JRC value lead organises core meetings with all LLs and annotators to discuss and resolve various issues and problems in annotations that may be relevant to the whole exercise.
- Based on the core meetings the JRC value lead updates the annotation guidelines.
- The LG curator, who is a more experienced annotator, merges all annotations for one document and resolves the identified conflicts in values annotation that results in what is considered as final annotations.

2.3.4 Curation

Curation is necessary to ensure the consistency and quality of the annotations within and among language groups. The curation phase is divided into two parts. The first curation part should run in parallel to the annotation phase, so that curators can act as a direct control mechanism and give immediate feedback to annotators, whereas the second curation part can start once several documents of a LG are annotated.

2.3.4.1 Part 1

The curators' task is to view, compare and to go through the annotations of all annotators sentence by sentence on the same text document. Thereby, they have to check for and remove false value annotations according to the definition of the taxonomy. In addition, they also need to check whether annotators have missed values and include them in case of doubt. Finally, in case the annotators have made different annotations for a sentence, it is the curators' task to decide if the values coexist or which annotation is the most corresponding. 100% agreement will not always be achievable. Often several values may legitimately be expressed in one sentence or values expressions can be understood differently depending on the point of view. However, the big challenge and task of curators is to make the final decision in the case of disagreement among annotations. Ambiguities or problems with the annotation of certain values identified by curators should be discussed immediately during the next LG meetings.

2.3.4.2 Part 2

Once several documents for a language are annotated these additional steps follow:

1. All closed annotations for a document are visible in the curation view on Inception, which contains additionally the context of the annotation. For the most difficult labels, the curator

goes through them and checks them all a second time, and fixes any issues that may not have been noticed during the first phase of the curation process;

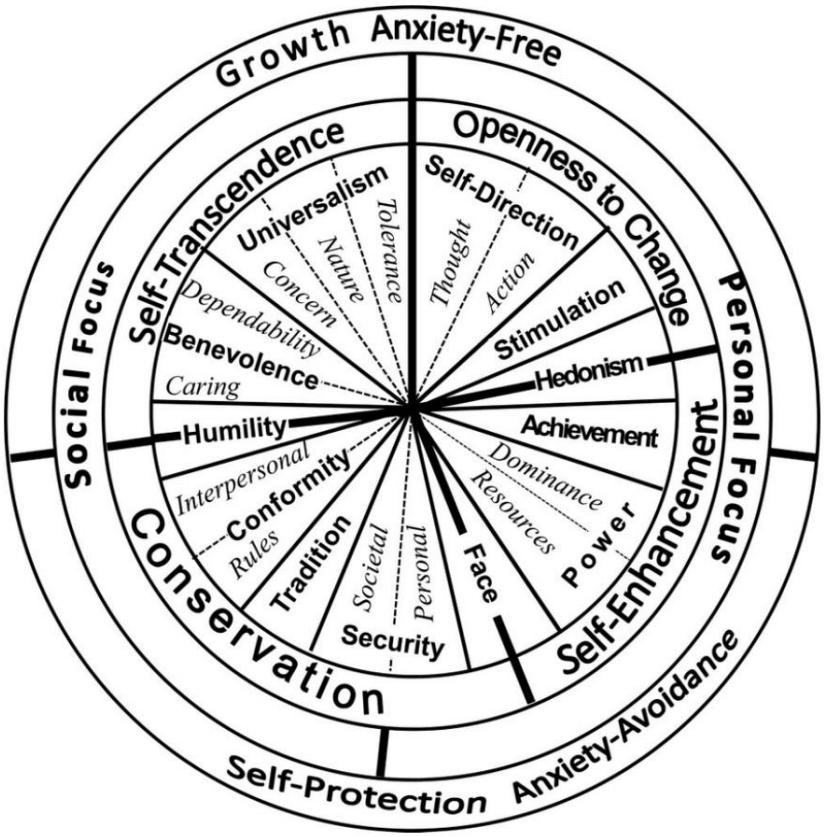
2. All text fragments that have been annotated and the corresponding labels are sorted alphabetically, and the curator explores whether similar sentences were tagged with similar labels since, intuitively, identical or near-duplicate texts should be assigned the same label, unless it is evident from the context that the distinct labels assigned are correct.
3. During the entire curation phase, the curators join the core meetings to converge and align to a common approach in the context of potentially identified discrepancies in annotations across languages.

3 Taxonomy and examples

3.1 First level of analysis - Values Categories

Below in Table 1, the 19 refined human values are developed by Schwartz et al. (2012). These values form the basis for the annotation campaign. The structure of these values reflects a circular motivational continuum in which neighbouring values share similar motivations and opposing values have conflicting motivations (see Figure 2).

Figure 2. Personal values circle



Source: Schwartz et al. (2012).

Table 1. Description of the 19 refined human values developed by Schwartz et al. (2012)

VALUE	CONCEPTUAL DEFINITIONS IN TERMS OF MOTIVATIONAL GOALS
SELF-DIRECTION–THOUGHT	Freedom to cultivate one's own ideas and abilities
SELF-DIRECTION–ACTION	Freedom to determine one's own actions
STIMULATION	Excitement, novelty, and change
HEDONISM	Pleasure and sensuous gratification
ACHIEVEMENT	Success according to social standards
POWER–DOMINANCE	Power through exercising control over people
POWER–RESOURCES	Power through control of material and social resources
FACE	Security and power through maintaining one's public image and avoiding humiliation
SECURITY–PERSONAL	Safety in one's immediate environment
SECURITY–SOCIETAL	Safety and stability in the wider society
TRADITION	Maintaining and preserving cultural, family, or religious traditions
CONFORMITY–RULES	Compliance with rules, laws, and formal obligations
CONFORMITY–INTERPERSONAL	Avoidance of upsetting or harming other people
HUMILITY	Recognizing one's insignificance in the larger scheme of things
BENEVOLENCE–DEPENDABILITY	Being a reliable and trustworthy member of the in-group
BENEVOLENCE–CARING	Devotion to the welfare of in-group members
UNIVERSALISM–CONCERN	Commitment to equality, justice, and protection for all people
UNIVERSALISM–NATURE	Preservation of the natural environment
UNIVERSALISM–TOLERANCE	Acceptance and understanding of those who are different from oneself

3.2 Second level of analysis: Attainment of the value

The labels for the reference quality are "(Partially) attained", "(Partially) constrained" and "Not **sure, can't decide**". Each referred value can either be mentioned as something that is or should be attained or something that is not attained or constrained. Attainment implies that the scenario described in the sentence contributes to the fulfilment of the value. The focus is therefore on the context/ event/ object of the sentence in which a value is mentioned. For SECURITY, attainment would mean that something is made safer or healthier. In contrast, an event can be stated in a way that thwarts/ constrains safety or health.

Thereby, the reference quality does not directly depend on tense or modus of a sentence. For example, *"I plan to start putting money aside each week."* Even, if POWER - RESOURCES has not yet been acquired

or increased by saving money each week, the announcement to do so in the future is showing that the value is pursued to be fulfilled.

Some examples relating to spending/budget:

- A generic mention of spending without mentioning of what it will achieve is most likely "constrained", less so unsure/can't decide.
- A mentioning of spending that achieves something (in a positive way) is an investment that achieves more value, so it is "attained"
- A mentioning that spending is too high, fails to achieve something or underperforms is "constrained"
- Not enough money, or not enough spend is "constrained"

The label "Not sure, can't decide" applies in cases a quality reference of a value is not expressed in a sentence. E.g. *"I want to go on holiday."* clearly refers to the value of HEDONISM. However, it is not clear here if the value is (partially) attained or constrained. A person only expresses a wish without indicating whether she/ he will do anything to make it come true or whether there are hindering factors. The context may however indicate that he/she realised it. For comparison, *"I haven't had a holiday since before the pandemic"*. Here, the second label is (partially) constrained because it is mentioned that the pandemic thwarted the fulfilling of the value. A similar example referring to Stimulation is "*Skydiving has always appealed to me*.". If the value is (partially) attained or constrained cannot be said. Finally, considering an example for the value of ACHIEVEMENT: "*Our work has resulted in resignations, apologies and policy corrections*.". Whether the aim of the work was to make people dismissed, force apologies etc. (it could be the perspective of an Environmental Protection Organisation) or whether the work performance was poor and set targets were missed (it could be the perspective of Managers of a private company) is not expressed clearly. Again, the direct context of the sentence may allow you to induce this level even if it is not directly expressed in one sentence. Moreover, a value can be attained even though a sentence contains a statement about the future, e.g. "I am going to put my pocket money aside". Even though it has not happened yet, the value SELF-DIRECTION – ACTION is enforced and therefore *(partially) attained* in the sentence, as there is a certain level of commitment expressed.

3.3 Examples for Values and their Attainment Label

In the following examples, we first highlight how a specific refined value may be expressed in text. It does not exclude that in the sentence also other values are expressed. Importantly, the annotation highlights the values mentioned in the text, and not the values of the author.

Additionally, it is important to stay as close as possible to the text. Even if a follow-up sentence may cast doubt on the meaning of a value or invites for further interpretation, we still want to highlight the value as expressed in that specific section. The idea behind this is that, even if a later part of the text may change the interpretation, it has still made the connection between an issue and a value to begin with; it has made it salient so the understanding will be triggered in the reader. In other words, the value is already salient. Secondly, the reference quality mentioned in the examples is indicated.

SELF-DIRECTION-THOUGHT:

Personal motivation: It is important to be creative, forming own opinions, be unique, have original ideas, learn things for oneself and improve own abilities.

The focus of this value is on developing own ideas, wanting to know more and discovering.

Example 1:

"In the scene, a young child is flipping through the pages of a storybook. At some point, the child comes to a picture she finds interesting and requests the name of the object from the caregiver."

Attainment label: *(Partially) attained*

The child mentioned is not only curious but also acts accordingly by asking, which contributes to the fulfilment of the value.

SELF-DIRECTION-ACTION:

Personal motivation: It is important to make own decisions about life, being independent and having the freedom to choose.

The focus of this value in contrast to the SELF-DIRECTION - THOUGHT aspect is to determine an action, rather than a conviction or thought.

Example 1:

"He knows that it is a wrong decision, but he intends to do it all the same."

Attainment label: *(Partially) attained*

The person described prioritises the freedom to make his own decision about the relevance of right or wrong. Therefore, it is a clear expression of the aspiration of the value.

STIMULATION

Personal motivation: Always looking for something new to do, doing something exciting, seeking out new experiences, innovating, being bold, seeking adventures and initiating change.

Note: In contrast to HEDONISM, this value focuses on the novelty and risk aspects of behaviours and thoughts. It is seeking out everything that stimulates the senses.

Example 1:

"Today, he is trying out for the first time Skydiving."

Attainment label: *(Partially) attained*

The event described is an adrenaline-rich sport and thus, helpful for the fulfilment of the value.

HEDONISM

Personal motivation: Having a good time, enjoying life's pleasures and taking advantage of opportunities to have fun.

Example 1:

"Today, children still enjoy circuses, even if they have fallen out of fashion for many older people."

Attainment label: *(Partially) attained*

Up to date, circuses provide pleasure for children and thus, are helpful for the fulfilment of the value.

Example 2:

"The event ended on a big bang, essentially making all the hard-core fans happy as they provided them with free fan gear. "

Attainment label: *(Partially) attained*

The mentioned event in the sentence ensured the fulfilment of the value for participants.

ACHIEVEMENT

Personal motivation: Being ambitious, successful and being admired for achievements and skills. Demonstrating competence according to social standards or in competition.

Important dimension of achievement is that it is perceived within the social standards, and according to the rules of engagement, unlike as for the value of power. Moreover, unlike the value of power, Achievement focuses on performance and not on resource matters.

Example 1:

"While effective COVID-19 vaccines were developed rapidly, all steps have been taken to ensure their safety and effectiveness."

Attainment label: *(Partially) attained*

The described medical progress is effective and thus a success, which is helpful for the fulfilment of the value.

Example 2:

"The new farming techniques have contributed tremendously to increases in agricultural productivity for local farmers."

Attainment label: *(Partially) attained*

The described technical progress is successful, which is helpful for the fulfilment of the value.

POWER – RESOURCES

Personal motivation: Having lots of money for the power it brings, being wealthy, and pursue social status.

The value of POWER - RESOURCES should not automatically be annotated, when terms like budgets, costs or growth appear in a sentence. Such terms refer to the value only if they make up an essential part of a sentence and are expressed as an important part of decision making with an underlying motivation or justification.

For example, when someone is writing about the "huge", "large", "out of control", or in any other way mentioned as important argument. Finally, in contrast to the value of ACHIEVEMENT, it only focuses on resource matters not on performance.

Example 1:

"Connection between social class and political power in US society is present for long time."

Attainment label: *Not sure, can't decide*

Whether the described link between social class and political power is a constraint or helpful for the attainment of the value depends on the social class to which a person belongs. Moreover, social class is referred to rather as a resource issue than an issue of exercising control/ determine directions. For this reason, the value expressed is POWER - RESOURCE and not POWER - DOMINANCE.

Example 2:

"People from higher classes are more likely to succeed in politics thanks to their wealth, power, good education and connections."

Attainment label: *(Partially) attained*

The sentence focuses on people from higher classes only. Belonging to this social group is helpful for the fulfilment of the value. In addition, higher class is described here as a resource issue instead of an issue of exercising control/ determine directions which is why it is POWER - RESOURCE and not POWER - DOMINANCE.

Example 3:

"The budget for last year's government policies on defence went out of control."

Attainment label: *(Partially) constrained*

Since the budget went out of control last year, less budget should/ will be spent or is available for this or future years. The sentence focuses on the limitation of budget.

Example 4:

"World leaders are embarking on costly policies that have no capacity to replace fossil fuels but enrich crony capitalists at public expense and increasing costs."

Attainment label: *(Partially) constrained*

The sentence describes a handling of limited resources that is not sustainable or wasteful and therefore, endangers the fulfilment of the value.

POWER - DOMINANCE

Personal motivation: Want people to follow you, being the most influential compared to others, be the one to determine directions.

Example 1:

"*The EU will strengthen its voice in the world by championing multilateralism and a rules-based global order.*"

Attainment label: *(Partially) attained*

It describes the aspirations and measures to be taken in order to attain the value.

Example 2:

"The United States is no longer the only big kid on the geopolitical block, but has an opportunity to lock in its role as the world's pivotal power."

Attainment label: *(Partially) attained*

Although it is mentioned here that the United States has competition in its position as a world power, the sentence states that the United States has an opportunity to secure/maintain its position of power. Therefore, the described opportunity is in favour of the attainment of the value.

Example 3:

"New Zealand's leaders have held outsized influence on the world stage."

Attainment label: *(Partially) attained*

The described impact of NZ leaders shows the attainment of the value.

Example 4:

"That white people were at the top of the human hierarchy was axiomatic in the early 20th century,"

Attainment label: *(Partially) attained*

For white people the described is in favour of attaining the value.

FACE

Personal motivation: Does not want to be shamed by others, protecting public image, being treated with respect, honour, and dignity.

Example 1:

"There appears to be little appetite within the government for apologies or admissions of guilt, even after months of widespread criticism from abroad."

Attainment label: *(Partially) constrained*

Described is an unwillingness or rather harmful behaviour that works against the attainment of the value.

Example 2:

"The government apologised after months of widespread criticism from abroad."

Attainment label: *(Partially) attained*

The described behaviour works in favour of the attainment of the value.

Example 3:

"The departure of a celebrity leader was "a massive loss" to the country's brand abroad."

Attainment label: *(Partially) constrained*

The described loss works against the attainment of the value.

SECURITY - PERSONAL

Personal motivation: Avoid dangerous situations, value personal security and safety, live in a secure environment, have a secured income, being healthy.

Similar to the difference between BENEVOLENCE and UNIVERSALISM-CONCERN, the decisive difference between SECURITY - PERSONAL and SOCIETAL is to whom the value refers. In case of individuals, family and friends, it is PERSONAL, e.g., if someone writes in first-person about their own health. In contrast, when it is applied to any group or society as a whole, it is SOCIETAL. Thus, SECURITY - PERSONAL will most likely not appear often in the text only if speaking of individual experiences.

Example 1:

"*When she is going out at night, she seldom feels safe with all the immigrants coming into the country.*"

Attainment label: *(Partially) constrained*

The described circumstance is directed against the attainment of the value.

Example 2:

"As the only woman in the team, I don't always feel comfortable."

Attainment label: *(Partially) constrained*

The described circumstance is directed against the attainment of the value. Although the mentioned value refers to a situation many women experience and thus, has a societal dimension it describes an individual experience.

Example 3:

"Because I am Jewish, I avoid going to some districts in the city alone."

Attainment label: (Partially) attained

Although the underlying situation constrains the attainment of the value, the concrete described behaviour ensures that the person feels safe.

SECURITY – SOCIETAL

Personal motivation: Country should protect itself against all threats, state should be strong, order and stability in society are important, including economic stability (employment, no recession). Importantly, the value refers not only to a society as a whole but also to socially defined groups like women, parents, etc. within a society. In contrast to the value of *UNIVERSALISM-CONCERN,* the value

emphasises protection more from a motivation of “preventing”, “averting”, "ending" dangers or threats and “preserving” security and stability.

Example 1:

"If there was one thing the U.S. government should do, it is to increase our budget for national defence"

Attainment label: *(Partially) attained*

The described situation favours the attainment of the value.

Example 2:

“The attacks have taken a threatening turn for society.”

Attainment label: *(Partially) constrained*

In the described situation, the fulfilment of the value is in danger.

Example 3:

“The few women in our society who have made it to leadership positions often feel insecure by their male counterparts.”

Attainment label: *(Partially) constrained*

The described circumstance is directed against the attainment of the value. Moreover, the mentioned value refers to the societal dimension of being a woman and not to an individual experience on the personal level.

Example 4:

“We need to fix our pension system so that all women are protected against poverty in retirement.”

Attainment label: *(Partially) attained*

A call or appeal for change is made to protect women against poverty. Although a current situation is mentioned in which the attainment of the value is constrained the sentence represents an appeal for change (“we need to”) that works in favour of the attainment of the value.

CONFORMITY - RULES

Personal motivation: Should follow authorities, follow rules even if others are not watching, obey all laws.

Example 1:

“It is my Commission's duty and most noble role to protect the rule of law.”

Attainment label: *(Partially) attained*

Helpful for the fulfilment of the value is the fact that the person sees it as their noble responsibility.

Example 2:

"The implementation had been very well thought through, as the rate of compliance goes far beyond the previous ones with 80% of people already registered."

Attainment label: *(Partially) attained*

Not only a high percentage of people, but also higher than before, behave in a way to fulfil the value.

CONFORMITY - INTERPERSONAL

Personal motivation: Avoid upsetting or annoying others, being tactful to others, showing courtesy, being polite, resisting temptation, respecting elders.

Example:

"If you're not vaccinated yet, please, when it's your turn, make the time roll up your sleeves and get your shot."

Attainment label: *(Partially) attained*

The request expressed describes what needs to be done in order to attain the value.

TRADITION

Personal motivation: Maintain traditional beliefs and values, follow the family or religious customs, valuing traditional practices of one's culture.

Example:

"An atheist when he started in politics in the late 1980s, Orbán now calls himself a defender of Christianity."

Attainment label: *(Partially) attained*

Calling himself a defender implies the willingness to protect the religious customs and beliefs.

HUMILITY

Personal motivation: Try not to draw attention, be humble and satisfied with the situation, not asking for more.

Example:

"We should be happy and satisfied with what we have."

Attainment label: *(Partially) attained*

The sentence expresses a request for the value, which is helpful for its fulfilment.

BENEVOLENCE – DEPENDABILITY

Personal motivation: Be loyal to close ones, be dependable and trustworthy, especially to close ones (in-group). Be seen as reliable, others should have confidence in you helping close ones.

A key difference between BENEVOLENCE and UNIVERSALISM is that BENEVOLENCE is primarily targeted towards close ones, and not towards strangers. In line with SECURITY-PERSONAL, close ones are family and friends, not the larger public.

Example 1:

"When push comes to shove, I always stand by my friends."

Attainment label: *(Partially) attained*

Expressed is the willingness to behave in a way that works in favour of the attainment of the value.

Example 2:

"We support our allies the best we can."

Attainment label: *(Partially) attained*

Expressed is the commitment to behave in a way that works in favour of the attainment of the value.

BENEVOLENCE – CARING

Personal motivation: Help and care about close ones, be responsive to family and friends. Actively helping and taking care of someone close.

A key difference between BENEVOLENCE - CARING and DEPENDABILITY is that the CARING dimension is active about doing something, while the DEPENDABILITY is more about the perception of being there for someone, being trustworthy etc. Furthermore, caring here is not generic or universal, it is limited to specific persons or a delineated in-group within a person's immediate environment. This last aspect distinguishes BENEVOLENCE – CARING from UNIVERSALISM-CONCERN. Similar to the two sides of security, the decisive difference between BENEVOLENCE – CARING and UNIVERSALISM-CONCERN is to whom the value refers. In case of individuals, family, and friends, it is BENEVOLENCE – CARING. In contrast, when it is applied to any social group or people in general, it is UNIVERSALISM-CONCERN. Thus, BENEVOLENCE – CARING will most likely not appear often in the text only if speaking of individual experiences.

Example 1:

"After coming home, he is checking in on his children to see that they are doing OK."

Attainment label: *(Partially) attained*

Describes a behaviour that works in favour of the attainment of the value.

UNIVERSALISM – CONCERN

Personal motivation: Protecting the weak and vulnerable, care about equal opportunities, treat everyone justly. In contrast to the value of *SECURITY – SOCIETAL*, protection and caring also goes beyond society boundaries. It is more generic referring to all kinds of people/ groups. Moreover, *UNIVERSALISM-CONCERN* focusses more on caring, protecting, promoting well-being, especially of vulnerable people from a motivation of "empathy", "helping" or "a universal justice perspective", whereas *SECURITY – SOCIETAL* focusses on protecting, promoting well-being more from a motivation of "preventing", "averting", "ending" dangers or threats and "preserving" security and stability.

Example 1:

"We work in more than 30 countries, support child education, help the poorest and most vulnerable, and support human rights activists."

Attainment label: *(Partially) attained*

Behaviour is described that is helpful for the fulfilment of the value.

Example 2:

"We hold the values of mutual help, solidarity and humanism in high esteem."

Attainment label: *(Partially) attained*

A commitment/ confirmation is mentioned that works in favour of the attainment of the value.

Example 3:

"The new policy is intended to strengthen communities in need, for example those particularly hard hit by COVID."

Attainment label: (Partially) attained

A political decision is mentioned that works in favour of the attainment of the value. Even though, strengthening communities in need relates also to stability and security within a society, the expressed motivation here is on helping/ caring. Due to that it is *UNIVERSALISM-CONCERN* and not *SECURITY – SOCIETAL*.

Example 4:

"The school (reform) bill will strengthen learning conditions for students."

Attainment label: (Partially) attained

A new policy is mentioned that works in favour of the welfare of students (attainment of the value). The expressed motivation is rather caring then enhancing the security or stability of students.

UNIVERSALISM - NATURE

Personal motivation: Care about nature for nature's sake, protect the environment against pollution, destruction and other threats.

Example 1:

"It is our duty to protect nature from the massive loss of biodiversity currently ongoing."

Attainment label: *(Partially) attained*

Helpful for the fulfilment of the value is the fact that the person sees it as their duty.

Example 2:

"We need to do more to protect the environment."

Attainment label: (Partially) attained

The sentence expresses a request for the value which is helpful for its fulfilment.

Example 3:

"The traditional method to produce honey is much better for the environment."

Attainment label: *(Partially) attained*

The described new way of production is helpful for the fulfilment of the value.

Example 4:

"The evidence points now towards doom for the climate."

Attainment label: *(Partially) constrained*

The sentence indicates clearly that the nature will suffer and thus is constraining the value.

Example 5:

"World leaders proclaim climate change as our greatest problem, demonising carbon dioxide."

Attainment label:*(Partially) attained*

Even though, the sentence mentions and identifies climate change a great danger for nature and environment the sentence predominantly mentions measures and awareness in order to stop climate change.

UNIVERSALISM – TOLERANCE

Personal motivation: Care about peace and harmony, listen to people with other views, understand even those one disagrees with.

The key point here is that it is peace for harmony's sake and not for the protection of the weak, which is covered in UNIVERSALISM - CONCERN already.

Example 1:

"Europe has experienced increasing tensions between national majorities and ethnic or religious minorities." [Also SECURITY and TRADITION as value present]

Attainment label: *(Partially) constrained*

The situation described in the sentence is one in which the value (partially) is not implemented.

Example 2:

"Individuals and organisations have relentlessly attacked others who do not share their beliefs."

Attainment label: *(Partially) constrained*

The sentence describes a situation in which the value cannot be attained.

3.4 Precision of span

When highlighting a span, be as specific as possible. It is not essential that the annotation is concise, but annotators should still try to highlight only the necessary text that truly refers to a value. It is not necessary to highlight the actor of an action. For example, the sentence:

"The EU discussed[2] whether to jointly invest into military capacity to be able to better protect European citizens."

➔ Only the "military capacity to be able to better protect European citizens" represents the value SECURITY SOCIETAL, while the first part of the sentence is only the background information.

3.5 General guidelines for annotation and usage of taxonomy

- The main goal is to train a values classifier machine. The software will be able to learn from our coding even if we mix things up from time to time (which may be in the nature of the values annotations anyway). 100% agreement is unrealistic so do not get lost in this effort.
- During annotations, weigh the need for the best annotation against the related workload/time (do not dedicate too much time for one single sentence).

- Titles are annotated similarly to text if values are expressed.
- No keyword-based annotation: Single words should not be annotated automatically as value. Only values resulting from the sentence context are annotated.

- Span within one sentence should be annotated!
- A justification even if not directly a (personal) motivation is enough to make something value laden and should therefore be annotated.
- Even though the general approach is on sentence base annotations in case of short follow-up sentences like "Yes." or "You bet." you can consider highlighting related text in the previous sentence. This way, we avoid that the machine associates single words/ expressions with a specific value.
- Annotations should primarily be based on the context of the sentence. However, some sentences can be ambiguous if they do not stand alone or are very short, requiring to be read in

[2] Discussion may also be labelled as value of SELF-DIRECTION ACTION. It is a value statement as they could also decide, regulate, enforce or even force countries to comply.

the overall text context. Thus, in case of ambiguous sentences to get the correct understanding/ meaning of the sentence context, the broader context should be considered as well. For example,

- Overall text context:
  The sheriff's office informs that several users on Trump's right-wing social media platform Truth Social posted the names of the jurors charging Trump and others with illegally conspiring to overturn his 2020 election loss in Georgia.

  Sentence:

  "We take this matter very seriously and are coordinating with our partners to respond quickly," the spokesperson stated.

  Without considering the overall-context, it is unclear what the sentence is about. Taking into account the overall-context allows to identify SECURITY-SOCIETAL # (PARTIALLY) ATTAINED in the sentence.

- Overall text context:
  A report on the challenge for parents to secure the future of their children in a system that makes career opportunities only after graduation from an expensive private university.

  Sentence:

  "I put money aside each week".

  Without considering the overall-context, it is unclear for what money is put aside. Taking into account the overall-context allows to identify BENEVOLENCE-CARING # (PARTIALLY) ATTAINED in the sentence.

- Overall text context:

  Chemicals released in recent decades are an immense threat to animal health and biodiversity. Insects and small rodents are increasingly victims to these chemicals, and worst of all are so called "forever chemicals". These chemicals are nearly indestructible and therefore will never or very unlikely be taken out of the environment. A new report has found higher concentrations of those chemicals in animals than previously expected.

  Sentence:

  "This evidence points to impending doom."

  Without context, it is not clear what doom is applied to. Therefore, it could be security-societal for example. In this context however, it is UNIVERSALISM – NATURE (PARTIALLY) CONSTRAINED, as the text is focusing on the impact on animals and nature. Human consequences are not mentioned.

Overall, annotators are encouraged NOT to read too much into the context, but as always to stay as close to the text as possible. Only if the sentence alone makes it difficult to annotate should the context be considered.

- Value annotation focus on the expressed perspective: Values seen by reader not by author, but also not interpreting too much what the reader might think. Annotate what is written.
- Do not interpret meaning into an expression, i.e. avoid reading too much into sentences
    - *"Locals now avoid staying in these neighbourhoods because they cannot reconcile poverty, destruction, urban sprawl, or pollution with their original idea of Paris."* [SECURITY – SOCIETAL # (PARTIALLY) CONSTRAINED] If you identify the value HEDONSIM expressed in this sentence, you would assume that Paris expresses HEDONISM. This would depend on your interpretation and is not expressed in the written text.

- When annotating, we have to accept lines without expressed values. E.g., sentences can be rather descriptions of processes, explaining how things work without explicitly expressing motivation. Text/behaviour can be an expression of several values. So make the ranking, which value matters most for this sentence. Focus on central message of sentence.
    - *"Boris Johnson has been referred to the police by the Cabinet Office over claims that he broke lockdown rules by hosting family and friends at Chequers during Covid."* Hosting family and friends expressing HEDONISM IS not the focus of sentence but the detention over claims is.
    - *"World leaders proclaim climate change as our greatest problem, demonising carbon dioxide"*. In the previous example, climate change as the greatest problem is thwarting the value of UNIVERSALISM – NATURE but the central message is the statement of world leaders defining the greatest challenge which helps to attain UNIVERSALISM – NATURE
- Some values might be more obvious to identify if they are attained or constrained, but please keep in mind all attainment values.
- Avoid making ideological judgements
    - It is hard to disassociate ourselves from our political views, but annotators should try to stay as objective as possible when it comes to annotating as the goal is to annotate what is writte, not how it makes us as annotators feel.
    - "A Christian association in the U.S. funds and coordinates a network of centres that promote the value of life when it comes to abortions."
    - This should be annotated with TRADITION # (PARTIALLY) ATTAINED. As the association is promoting life according to a specific Christian tradition.
    - "A conservative Christian association in the U.S. funds and coordinates a network of centres that oppose abortion and promote the value of life."
      This should still be annotated with TRADITION # (PARTIALLY) ATTAINED. Even if we now have a political connotation and an opposition to abortion which may be seen critically. The association is still promoting life according to a specific Christian tradition.
    - "An association in the U.S. funds and coordinates a network of centres that oppose abortion and promote the value of life."
      Here, an unspecified association is promoting the wellbeing for all people. However, the promotion of life and the opposition to abortion is associated with tradition in most societies. So we can still assume that a traditional motivation is expressed here (TRADITION # (PARTIALLY) ATTAINED).
      "An association in the U.S. funds and coordinates a network of centres that oppose abortion by making it difficult for women to attain such services during pregnancy."
      Here constraints for women are highlighted and hence UNIVERSALISM-CONCERN is (PARTIALLY) CONSTRAINED, because it expresses that women find it hard to receive this service.

    - Similar examples can be constructed around climate change and immigration. Try to stay as close to the text as possible and divorce yourself from your political interpretation.

- When dealing with misinformation, we come across fake news, wrong statements and assigning an attainment label can seem complicated. In case you identify misinformation:

    - Stay with text-based annotation for all text types
    - Ignore whether an information is true/wrong
    - Label only based on what is written
    - Do not play the fact-checker

    For example, *"Taking 5 mg vitamin C per day saved me from a COVID infection." is* SECURITY-PERSONAL # (PARTIALLY) ATTAINED. The statement that a certain amount of vitamin C prevents a COVID infection is not correct. If you take into account that the information is wrong, the attainment label would be (partially) constrained. However, you have to ignore whether the information is true or wrong.
- When dealing with sarcasm/ irony annotate reasonably. If it is clear that what is written is not meant literally annotate accordingly. For example, "*Well, done! They ensured the economic stability of the country* commented the leader of the conservative party on the new minimum wage law of the Left party". Taken the obvious irony into account allows to identify ACHIEVEMENT # (PARTIALLY) CONSTRAINED in the sentence. In short, use your common sense of how the text is meant, rather than purely on what is written.

3.6 General guidelines for curation

The following aspects might be helpful for curators in their task of curating:

- The central aim of the curation process is to eliminate "false" values annotations by annotators. So, focus on them.
- In case of different values annotated by two/more annotators in a sentence but you see both justified keep both (thumb rule: the more information kept the better). Do not force a single value in the curation step.

For example: "The traditional method to produce honey is much better for the environment."

UNIVERSALISM-NATURE (PARTIALLY) ATTAINED or TRADITION (PARTIALLY) ATTAINED

--> UNIVERSALISM is the main value expressed here, but TRADITION can also be argued and is part of the same span.

- If you have the impression that annotators do not understand a specific value provide immediate feedback.
- In difficult cases, share them with LLs.
- Share repetitive mistakes in LG meetings.

Table 2. Curation example scenarios

Possible scenarios	How to decide
Same annotation, same span	Accept or delete if you think annotation is "false".
Same annotation, different span	Just accept one span without spending too much time on finding the exact span.
Same values but different attainment annotation	Choose the right attainment label if you agree with value or choose another value+attainment combination. Ask: what is expressed in the text most closely?
Only one annotation for text span	Decide if the value is present and accept, adapt or delete annotation accordingly.
No annotation in text span	Add values if all annotators have missed them. But do not spend too much time on trying to find a value. If several annotators decided not to annotate, the likelihood is high that there is no value which stands out.

4 Inception - the Annotation Tool

This section provides the essential information on the Inception[3] annotation tool used to carry out the annotations. The Inception instance for the project can be accessed via a web browser using the following link: Text Mining and Analytics Toolbox (europa.eu). The prerequisite is an EU Login (ECAS) account to get authenticated. For assistance on the ECAS account, please refer to the *Guide to create an EU Login account*. After the log-in you will be redirected to the main screen of Inception (see Figure 3):

In the remainder of this section, we describe briefly all the steps on how to use Inception. To get a step-by-step introduction to Inception, please also refer to the video *Introduction to Inception for annotators.*

4.1 Annotation

1. Access Inception: using this link: Text Mining and Analytics Toolbox (europa.eu)

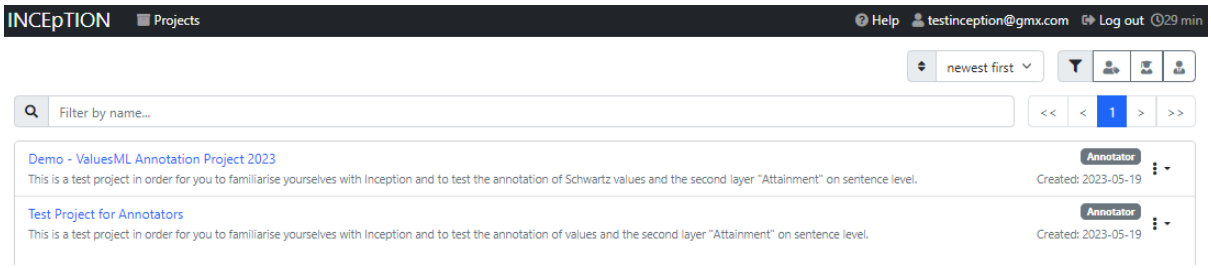
2. Select the project: On the main screen of Inception (see Figure 3), you will find the projects assigned to you:

• Test Project for Annotators: For the trial annotations and getting acquainted how to use the tool and

• [Your Language] ValuesML Annotation Project 2023: For the official annotations

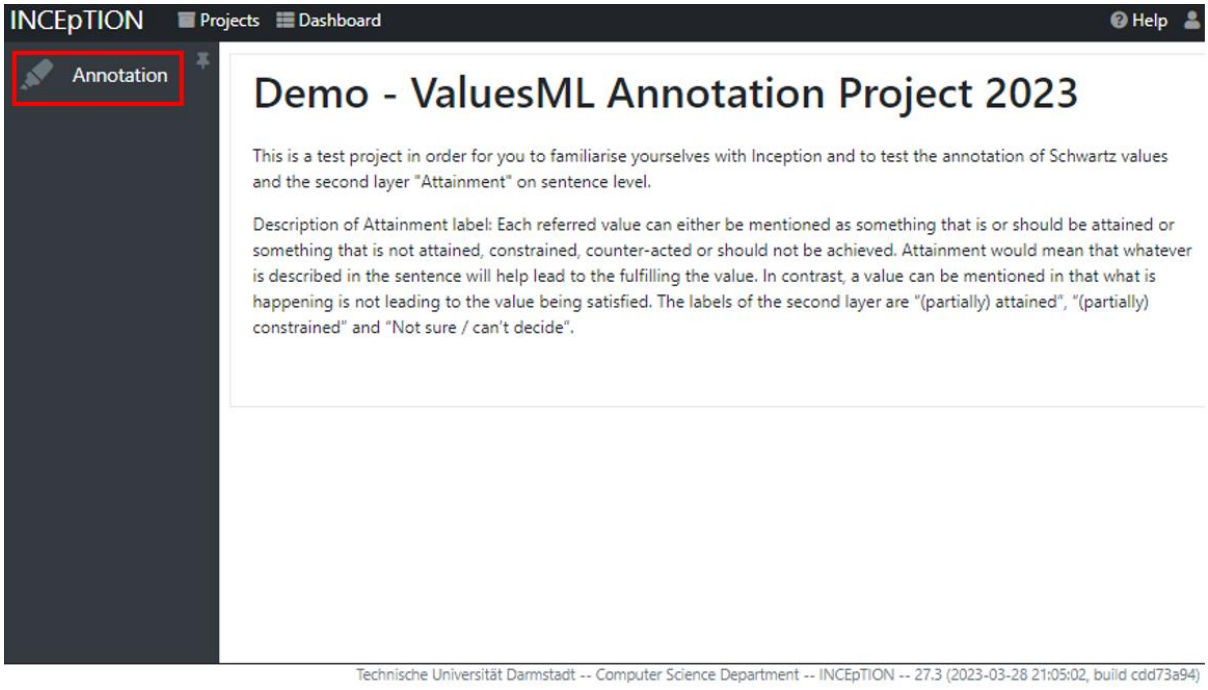
[3] https://inception-project.github.io/

Figure 3. Main screen on Inception



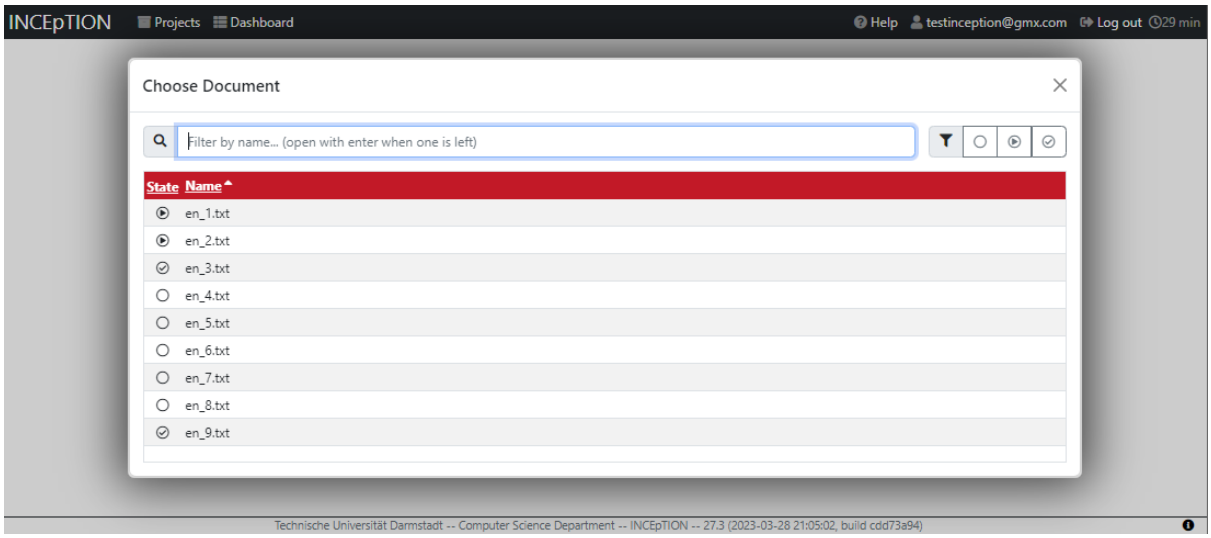
3. Start annotations: Click on the project to start the annotations and select the Annotations button (see Figure 4).

Figure 4. Open the annotation view



3. Select a document: From the list of all documents within the project, select the documents you have been assigned to annotate. For this, please refer to the document allocation provided by your language leader. You can also type the file name in the search bar on top. Another option to ease the navigation through all documents is to filter for a type of document (not opened, opened and closed) on the right-hand side (see Figure 5):

Figure 5. Selection of document



5. Decide on which text expresses a value and needs to be annotated: For each document, you must annotate the value and the attainment label.

6. Text span annotation: Highlight phrases or, if specification is not possible, a whole sentence you wish to attribute a value to annotate.

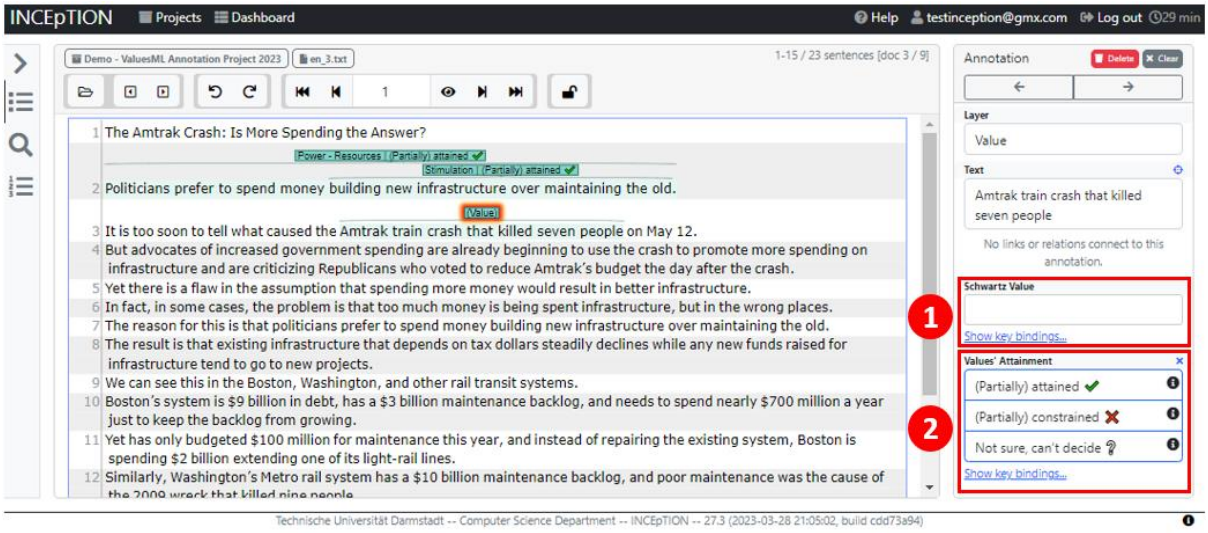
7. Annotating Value: Once a text span is highlighted, it will be possible to choose from the list of 19 values on the right-hand side of the screen (see Figure 6, step 1).

8. Annotating Value Quality: For the value, one of the attainment labels "(partially) attained", "(partially) constrained", "not sure, can't decide" needs to be assigned on the right-hand side of the screen (see Figure 6, step 2).

Continue steps 6-8 until document is fully annotated. Multiple value references for one sentence are possible if you identify several values. The colours highlighting the annotations relate to the higher order values as indicated in the values circle in Figure 2:

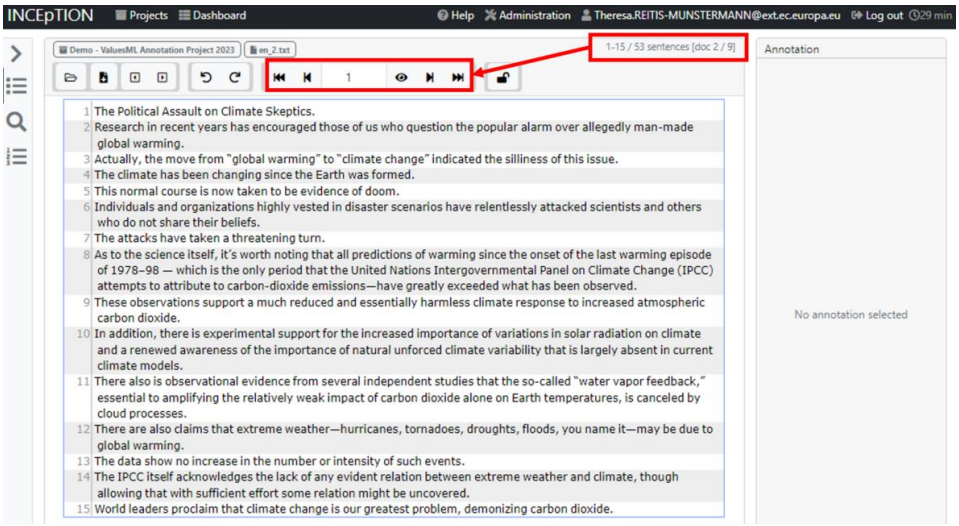
Orange: Openness to change
Blue: Self-enhancement
Red: Conservation
Grey: Self-transcendence

Figure 6. Text highlighting and annotation screen



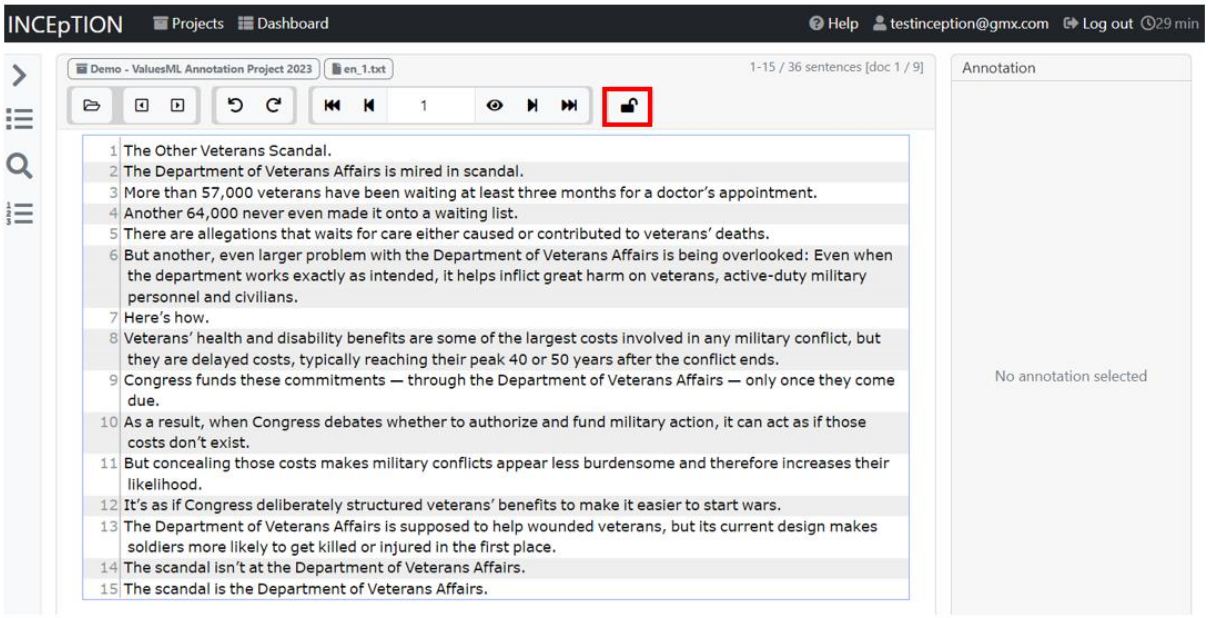
NOTE: Some documents may be longer than one page. The total number of sentences and the current page of the document can be seen at the top of the text on the right-hand side. In this case, it is necessary to scroll through the whole text using the navigation arrows as shown in Figure 7. Some documents may be bigger than the displayed view. In this case, use the vertical bar to scroll the page.

Figure 7. Navigation within a document



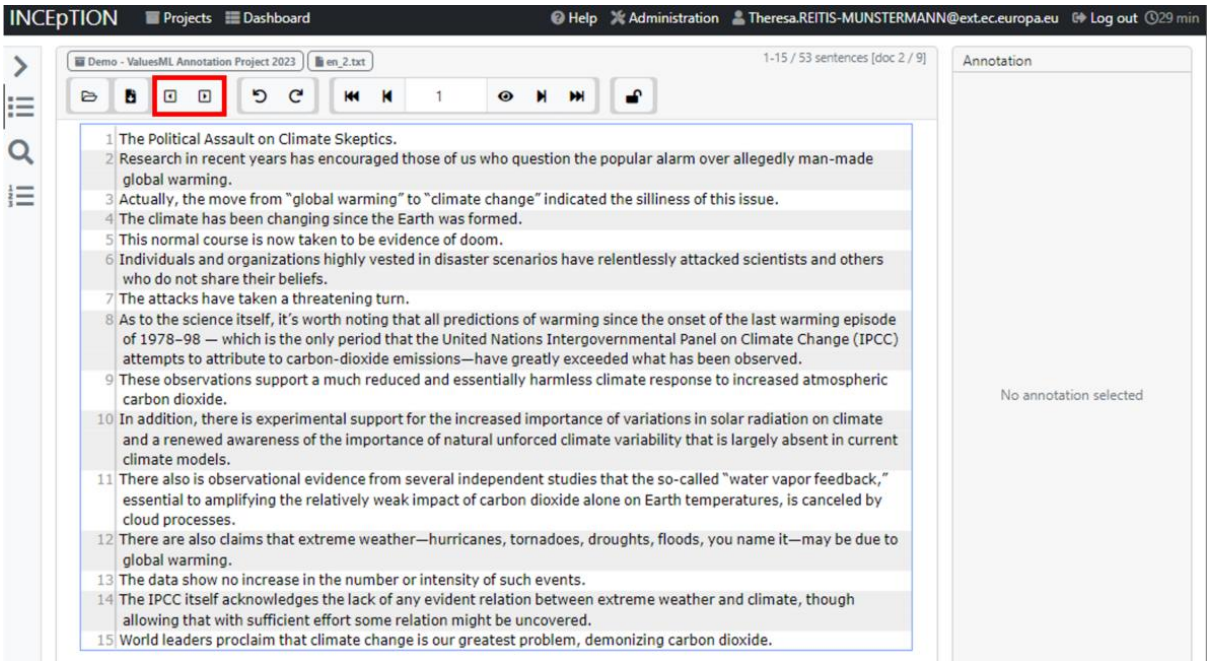
9. Finalising the work on a document: Once the document is completely annotated and no further changes are envisaged one has to click on the lock button to finalise the work (see Figure 8). This is important, as only closed documents will be included in the curation view. Only the curator will be able to unlock the document.

Figure 8. Closing the annotations of a document



Moving to the next document: Once the document is completely annotated, it is possible to simply move to the next document using the next document arrow (see Figure 9). The saving is automatic.

Figure 9. Navigation between documents

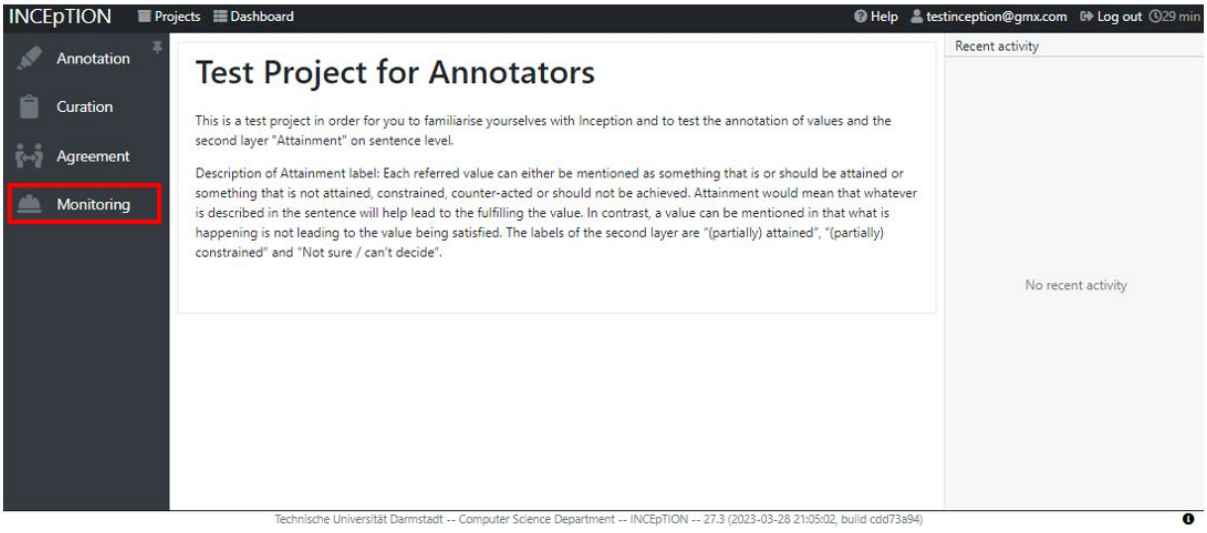


4.2 Curation

Curation is the step of producing a final version of annotation of a document, by taking into account the coding of different annotators. Below, the steps on how to use Inception for curation are briefly described.

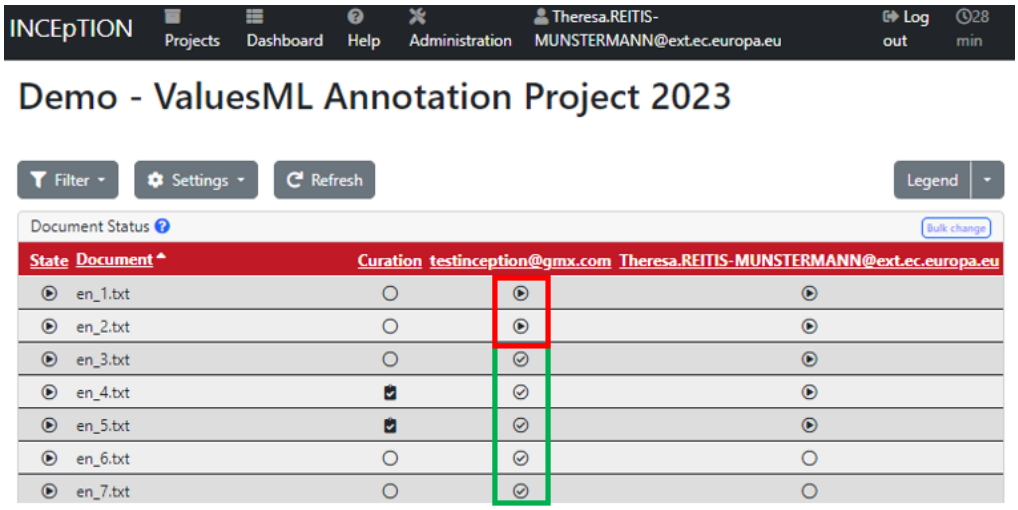
1. Check the status of documents: Before curation, it is necessary to check that the annotators assigned to the document have closed their documents since unclosed documents do not appear in the curation view. For this, the curator selects the Monitoring view (see Figure 10).

Figure 10. Open the Monitoring view



Here the overview shows which annotator has closed a document indicated by the check mark (see Figure 11, green box) whereas unclosed documents are indicated by an empty button (not opened) or a play button (opened at least once).

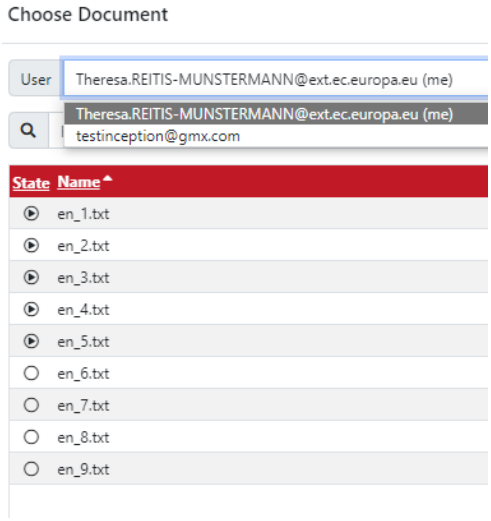
Figure 11. Overview in Monitoring view



In case, that annotators assigned to a document have not closed their annotations, the curator has to check if the annotators have done any annotations within a document. It is possible that annotators have finished annotating a document but forgot to finally close the document. In order

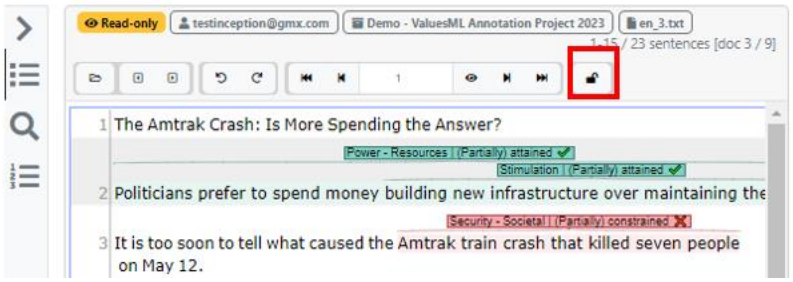
to check this open the annotation view (see Figure 4) and select the respective annotator (see Figure 12).

Figure 12. Open annotation view for a specific annotator



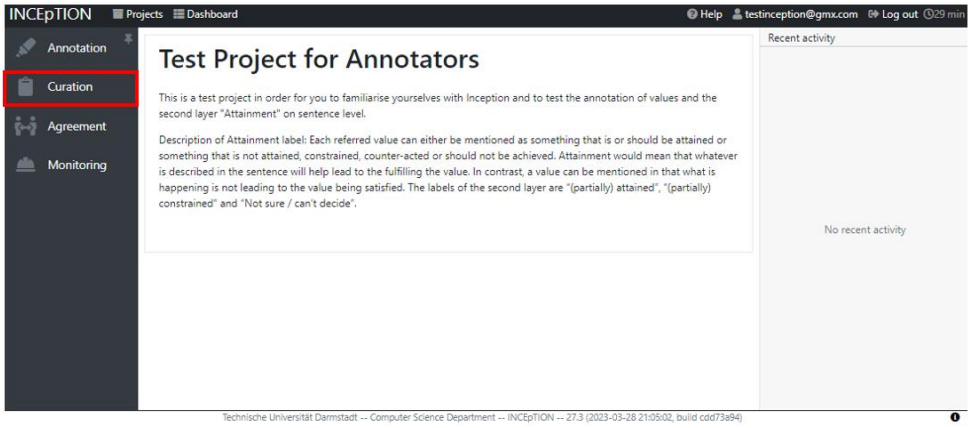
In the annotator's view, the curator can assess if there are annotations made across the whole text and thus if the annotator has completed the annotations without closing the document. In this case, the curator closes the document for the annotator, and it becomes accessible in the curation view (see Figure 13).

Figure 13. Close an annotator's annotated document



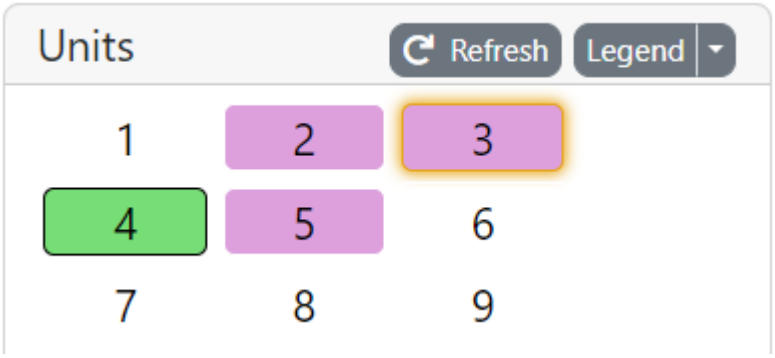
2. Start the curation: Curation can be started by clicking on "curation" (see Figure 14)

Figure 14. Open curation view



3. The overview of units (see Figure 15) can be misleading. For the curation, please consider each line, do not regard the overview's colours.

Figure 15. Overview of units per document, i.e. lines



Note: The colours indicated in the units' overview can be misleading as e.g., a green highlighted unit already appears in case there is at least one agreement between both annotators but does not consider any additional annotations.

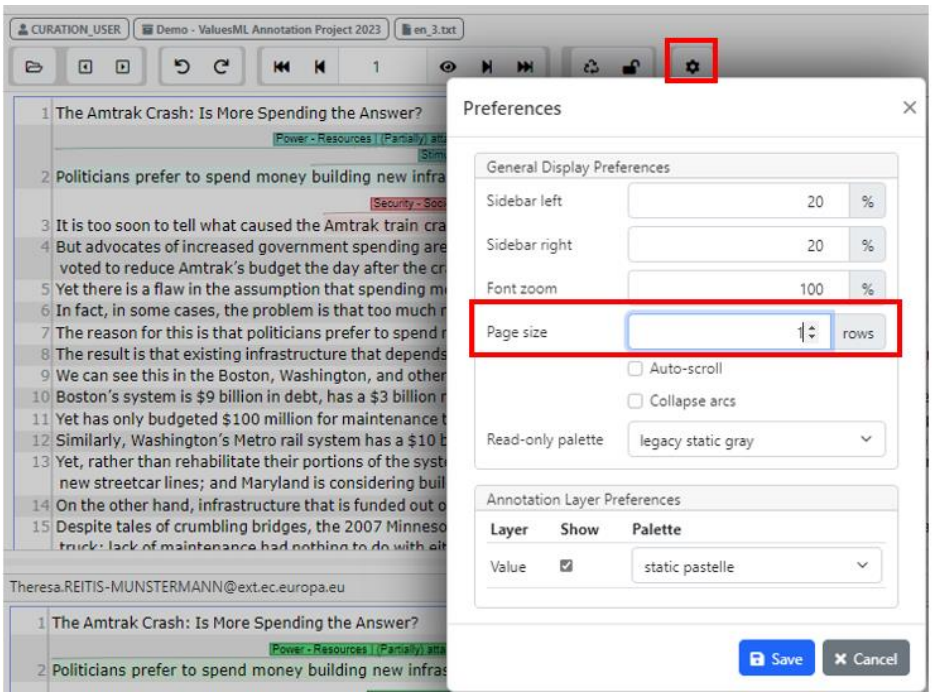
Green: For every text position that annotators have annotated, there is a corresponding annotation at the same position in the curator's document. That is, if at least two annotators agreed on the labels and span when opening the document for curation.

Lilac: A single annotator has made multiple annotations on the same layer at a given span.

White: There is no disagreement between annotators. This also includes if only one annotator made an annotation without contradicting another annotator's annotation.

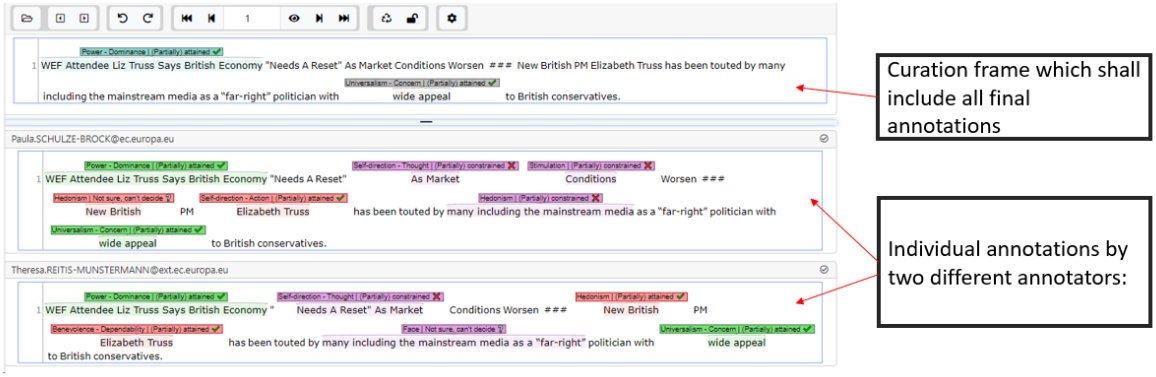
4. Open and edit a document: Once the status of a document annotators worked on is set to "closed", it is possible to open the document using the curation interface. To ease the curation activity, the number of lines displayed can be reduced by selecting e.g., five rows as page size in the settings (see Figure 16).

Figure 16. General Display Settings



The upper frame contains the final document annotation, the colours relate to the higher order values as indicated in the values circle in Figure 2. Below, the other frames contain the annotations of each individual annotator (see Figure 17). Importantly, annotators must have closed their document before. Otherwise, their annotation will not appear in the view.

Figure 17. Curation frames and annotators' frames



The colours of the annotation frames before curator adds and decides on final annotation represent the following:

Green: Same value and attainment label, and span automatically merged in curation frame

Lilac: Same value and attainment label but different span or in case an annotation has been made only by a single annotator

Red: Same value but different attainment label; different value labels but same attainment labels

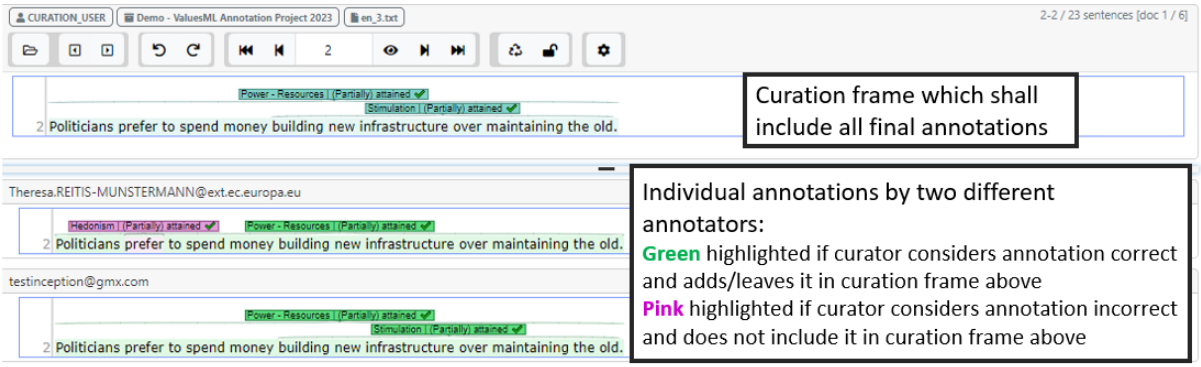
Note that the colours of the annotation frames adapt to curation (see Figure 18).

5. Curate the document: In case, annotators have annotated the same text span with the same values and attainment labels (total agreement; marked as green in the annotation frame), the annotation appears automatically in the curation frame at the beginning. Curators can leave them unchanged (if they accept the annotation) or still edit/ delete them by clicking on them. The frame containing the curated document can be edited freely to add annotations exactly in the same way as it is done when annotating. That is, one can highlight a text span and choose from the list of values and the attainment labels on the right-hand side of the screen.

   In addition, if curators agree with the annotation of an annotator and want to adopt it they can just transfer the annotation from any individual annotator (the frames below) to the upper frame by double clicking on the respective annotation. The annotation appears then in the curation frame.

   After the curation, which includes potential changes, the colours of the annotation frames are as presented in Figure 18.

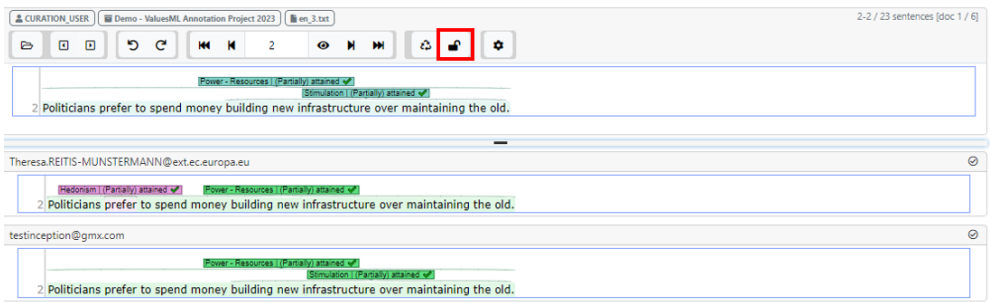
Figure 18. Curation frames and annotators' frames after curation



   One navigates through the lines using the navigation bar, in the same way as it is done during annotation.

6. Close the document: Once curation is finished, the document must be marked as completed by clicking on the lock icon.

Figure 19. Close the curated document



5 Conclusions

This guide presented instructions for the annotation and curation of news articles and political manifestos with 19 personal values and as second layer the attainment level of that value. For the instructions, each value was explained, examples for an ideal annotation were given and problematic cases were discussed. For both annotators and curators, instructions were also provided for how to use the online platform (Inception) to fulfil the tasks.

The guide was used for the valuesML annotation campaign of over 2500 news articles and political manifestos in 9 languages by about 70 values scholars. Leaders for each of the language contributed to the development of this report. We believe that this guide may also be useful for future annotation campaigns, perhaps in other realms such as audio, image or video annotation and the examples may help to get a better understanding of the values in context.

References

Piskorski, J. et al. 2022. *News Categorization , Framing and Persuasion Techniques : Annotation Guidelines*.

Ponizovskiy, Vladimir et al. 2020. "Development and Validation of the Personal Values Dictionary: A Theory-Driven Tool for Investigating References to Basic Human Values in Text." *European Journal of Personality*.

Scharfbillig, Mario et al. 2022. "Monitoring Social Values in Online Media Articles on Child Vaccinations." : 1–28.

———. 2024. *ValuesML: A New Multilingual Dataset for Values Detection in News and Political Manifestos*.

Schwartz, Shalom H. 1992. "Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries." *Advances in Experimental Social Psychology* 25(C): 1–65.

———. 2012. "Refining the Theory of Basic Individual Values." *Journal of Personality and Social Psychology* 103(4): 663–88.

List of figures

Figure 1. Overview of the weekly tasks of all participants during the annotation phase 8
Figure 2. Personal values circle 11
Figure 3. Main screen on Inception 30
Figure 4. Open the annotation view 30
Figure 5. Selection of document 31
Figure 6. Text highlighting and annotation screen 32
Figure 7. Navigation within a document 32
Figure 8. Closing the annotations of a document 33
Figure 9. Navigation between documents 33
Figure 10. Open the Monitoring view 34
Figure 11. Overview in Monitoring view 34
Figure 12. Open annotation view for a specific annotator 35
Figure 13. Close an annotator's annotated document 35
Figure 14. Open curation view 36
Figure 15. Overview of units per document, i.e. lines 36
Figure 16. General Display Settings 37
Figure 17. Curation frames and annotators' frames 37
Figure 18. Curation frames and annotators' frames after curation 38
Figure 19. Close the curated document 38

List of tables

Table 1. Description of the 19 refined human values developed by Schwartz et al. (2012) 12

Table 2. Curation example scenarios 28

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

— by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

— at the following standard number: +32 22999696,

— via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

Open data from the EU

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

Science for policy

The Joint Research Centre (JRC) provides independent, evidence-based knowledge and science, supporting EU policies to positively impact society

EU Science Hub
joint-research-centre.ec.europa.eu

Publications Office
of the European Union