

---

# Moral Framing – Annotation Guidelines

---

*Anonymous Project*  
*Version 1.0*  
*12.05.2025*

## Abstract

This document describes the annotation of moral frames in German political text. Our schema uses a frame-based description of morality, capturing moral values and moral acts and goals, together with the actors who participate in the frame (i.e., the moral roles of the frames). In addition to moral frames and roles, we annotate their moral foundations, following Moral Foundations Theory [Haidt et al., 2009, Graham et al., 2009]. Our schema also borrows ideas from the Narrative Policy Framework [Shanahan et al., 2017].

## Contents

<b>1</b>	<b>Background</b>	<b>3</b>
<b>2</b>	<b>Overview of the Classification Scheme</b>	<b>4</b>
2.1	Annotation procedure . . . . .	5
<b>3</b>	<b>Identifying moral framing</b>	<b>8</b>
3.1	What counts as moral framing? . . . . .	8
3.2	Which frames to annotate? . . . . .	8
3.3	Tests . . . . .	10
<b>4</b>	<b>Annotation span</b>	<b>12</b>
4.1	When do we annotate the whole clause instead of a VP? .	14
4.2	Coordination . . . . .	14
4.3	Complement and attributive clauses . . . . .	15
<b>5</b>	<b>Annotation of moral roles</b>	<b>16</b>
5.1	Hero . . . . .	16
5.2	Villain . . . . .	17
5.3	Victim . . . . .	18
5.4	Beneficiary . . . . .	18
5.5	Antecedent . . . . .	19

<b>6</b>	<b>Annotation of Moral Foundations</b>	<b>21</b>
6.1	Annotation perspective . . . . .	21
6.2	Care . . . . .	22
6.3	Equality and Proportionality . . . . .	22
6.4	Loyalty . . . . .	23
6.5	Authority . . . . .	23
6.6	Purity . . . . .	24
6.7	Liberty . . . . .	25
6.8	General Morality . . . . .	25

# 1 Background

## Moral Foundations Theory

MFT is a descriptive, pluralist theory of morality, developed in the area of social psychology [Haidt et al., 2009, Graham et al., 2013]. In contrast to monist theories that explain morality in terms of one single principle or dimension, *right-wrong*, MFT believes that the concept of morality is based on more than one such dimension, or foundation. According to MFT, these foundations have been developed during evolution as responses to several adaptive challenges, e.g., the emergence of the PURITY foundation has been driven by the need to avoid pathogens. Moral foundations are seen as intuitions or feelings rather than conscious judgments which is in contrast to other moral theories that describe moral intuitions as “strong, stable, immediate moral beliefs” [Sinnott-Armstrong et al., 2010] or as moral judgments [McMahan, 2000].

MFT assumes at least five moral intuitions that can be divided into *binding* foundations (ingroup LOYALTY, respect for AUTHORITY, and PURITY) and *individualising* foundations (CARE and FAIRNESS). Newer work has proposed that ideas of fairness can be based on different notions of justice, and has further divided the FAIRNESS foundation into EQUALITY and PROPORTIONALITY [Atari et al., 2023] where EQUALITY favours an equal distribution of opportunities and resources while PROPORTIONALITY prefers a distribution in proportion to an individual’s merit or contribution.

MFT explains inter-personal differences of moral values by assuming the existence of an “innate draft of the moral mind” that is later revised by experience and cultural influences [Graham et al., 2013, p. 9]. This makes MFT particularly interesting for comparative analyses of moral values across time and cultures (see, e.g., [Xie et al., 2019, Wu et al., 2023, Hämmerl et al., 2023]).

We consider the following set of Moral Foundations below (also see the MFT website<sup>1</sup>):

Care	is related to our long evolution as mammals with attachment systems and an ability to feel (and dislike) the pain of others. It underlies the virtues of kindness, gentleness, and nurturance.
Equality	is defined as “Intuitions about equal treatment and equal outcome for individuals.”
Proportionality	is defined as “Intuitions about individuals getting rewarded in proportion to their merit or contribution.”
Loyalty	is related to our long history as tribal creatures able to form shifting coalitions. It is active anytime people feel that it’s “one for all and all for one.” It underlies the virtues of patriotism and self-sacrifice for the group.
Authority	was shaped by our long primate history of hierarchical social interactions. It underlies virtues of leadership and followership, including deference to prestigious authority figures and respect for traditions.
Purity	was shaped by the psychology of disgust and contamination. It underlies notions of striving to live in an elevated, less carnal, more noble, and more “natural” way (often present in religious narratives). This foundation

---

<sup>1</sup> <https://moralfoundations.org/>

Label	Description
<i>Moral frames</i>	
<i>Moral_Value</i>	Moral value
<i>Moral_Act_or_Goal</i>	Moral actions, goals, or other desirable targets
<i>Immoral_Value</i>	Immoral value
<i>Immoral_Act_or_Goal</i>	Immoral, irresponsible or otherwise negatively portrayed act
<i>Political_Act_Or_Goal</i>	Policy issue, concrete political goal or proposal
<i>Moral roles</i>	
<i>Hero</i>	Person/group/organisation performing or persuing a commendable, responsible or praiseworthy act or goal
<i>Villain</i>	Person/group/organisation performing an act portrayed as immoral, irresponsible or otherwise dispreferred
<i>Victim</i>	Person/group that is harmed or disadvantaged by an action
<i>Beneficiary</i>	Person/group that benefits from an action or goal

Table 1: Inventory of moral frames and roles.

underlies the widespread idea that the body is a temple that can be desecrated by immoral activities and contaminants (an idea not unique to religious traditions). It underlies the virtues of self-discipline, self-improvement, naturalness, and spirituality.

Liberty is about the feelings of reactance and resentment people feel toward those who dominate them and restrict their liberty. Its intuitions are often in tension with those of the authority foundation. The hatred of bullies and dominators motivates people to come together, in solidarity, to oppose or take down the oppressor.

Our annotation guidelines make no claims about the psychological validity of the MFT (as this is not our area of expertise). Instead, we consider the moral foundations as a descriptive model for annotating moral framing in political communication.

## 2 Overview of the Classification Scheme

Table 1 gives an overview of our classification scheme. We annotate situations that are framed as (im)moral, using the two roles *Moral\_Act\_or\_Goal* and *Immoral\_Act\_or\_Goal*. The roles can belong to one of the moral foundations CARE, EQUALITY, PROPORTIONALITY, LOYALTY, AUTHORITY, PURITY and LIBERTY. We assign each *Moral\_Act\_or\_Goal* and *Immoral\_Act\_or\_Goal* the corresponding MF label.

The moral roles *Hero*, *Villain*, *Victim* and *Beneficiary* are universal and can be used with all moral foundations.

## 2.1 Annotation procedure

Annotation tool As annotation interface, we use the INCEpTION<sup>2</sup> [Klie et al., 2018], a browser-based tool.

### 2.1.1 (Im)moral values, goals, actions and policies

As a first step, the coders are asked to read the whole speech, focussing on the overall argument presented in the speech. We assume that a political argument will include values, goals and actions that are presented as desirable (praiseworthy) as well as ones deemed to be undesirable (blameworthy). In addition, political actors talk about concrete political proposals, such as laws, regulations and proposals for legislation. Once the annotators have understood what the speech is about, they are asked to annotate the praiseworthy and reprehensible values, goals and actions mentioned in the speech. For that, we use the labels *Moral\_Value*, *Moral\_Act\_or\_Goal*, *Immoral\_Value*, *Immoral\_Act\_or\_Goal* and *Political\_Act\_Or\_Goal*.

#### Values vs Acts or Goals

We consider moral values as abstract concepts that describe persistent preferences guiding moral behaviour, in short, how to distinguish right from wrong. Moral acts or goals, on the other hand, are typically far more concrete and are meant to take place at some point in time (past, present, future). They can be guided by moral values (but not *vice versa*).

*Moral\_Value* and  
*Immoral\_Value*

For the *Moral\_Value* / *Immoral\_Value* class, we consider

- moral values expressed through abstract concepts  
(*liberty, freedom, tradition*)
- immoral values expressed through abstract concepts  
(*injustice, oppression*)
- concrete objects or entities that symbolise (im)moral values  
(*the statue of liberty, the wall of shame*)

Example	Label
<i>Abstract concepts (typically NPs)</i>	
westliche Freiheit	<i>Moral_Value</i>
unsere Demokratie	<i>Moral_Value</i>
Rechtsstaatlichkeit	<i>Moral_Value</i>
die traditionelle Familie	<i>Moral_Value</i>
Ungerechtigkeit	<i>Immoral_Value</i>
Unfrieden in der Region	<i>Immoral_Value</i>
<i>Objects/Entities (NPs)</i>	
die Freiheitsstatue	<i>Moral_Value</i>
deutsches Kolonialzeitverächtdenkmäl	<i>Immoral_Value</i>
die kommunistische Schandmauer	<i>Immoral_Value</i>
Corona-Schandmal	<i>Immoral_Value</i>

*Moral\_Act\_or\_Goal*

For the *Moral\_Act\_or\_Goal* class, we consider

<sup>2</sup> <https://inception-project.github.io>

- moral or praiseworthy acts or goals  
(*protecting the children, social redistribution, upholding the rule of law*)

For German examples, see below:

Example	Label
<i>Actions (NPs)</i>	
die Befreiung der Menschheit aus der Tyrannei	<i>Moral_Act_or_Goal</i>
die Rettung des Planeten	<i>Moral_Act_or_Goal</i>
<i>Actions (VPs)</i>	
die Menschheit von der Tyrannei der Not befreien	<i>Moral_Act_or_Goal</i>
unseren Planeten heilen	<i>Moral_Act_or_Goal</i>
die Welt auf den Pfad der Nachhaltigkeit bringen	<i>Moral_Act_or_Goal</i>
<i>States describing behaviour (VPs)</i>	
Merkel <b>ist stets zuverlässig</b>	<i>Moral_Act_or_Goal</i>
<i>Requests (S)</i>	
Der Anwalt sollte nicht teurer sein als die Schuld.	<i>Moral_Act_or_Goal</i>
Umweltaspekte müssen dabei stärker in den Fokus rücken	<i>Moral_Act_or_Goal</i>
Wir wollen, dass <b>unsere Schulen sichere Orte sind</b>	<i>Moral_Act_or_Goal</i>

*Immoral\_Act\_or\_Goal*

For the *Immoral\_Act\_or\_Goal* class, we consider

- immoral or blameworthy acts or goals (*violating human rights, undermining the right to asylum, deceiving the German people*)

See below for German examples:

Example	Label
<i>Actions (NPs)</i>	
<b>Masseneinwanderung in unsere Sozialsysteme</b>	<i>Immoral_Act_or_Goal</i>
<b>der Verrat an den Pendlern</b>	<i>Immoral_Act_or_Goal</i>
<b>die Ausbeutung der Arbeiterschaft</b>	<i>Immoral_Act_or_Goal</i>
<i>Actions (VPs)</i>	
Die FDP hat <b>die Pendler verraten</b>	<i>Immoral_Act_or_Goal</i>
<b>die Arbeiterschaft ausbeuten</b>	<i>Immoral_Act_or_Goal</i>
<i>States describing behaviour (VPs)</i>	
Merkel <b>ist bis ins Mark korrupt</b>	<i>Immoral_Act_or_Goal</i>
<i>Blameworthy situations (S)</i>	
Da kommt das Fressen vor der Moral.	<i>Immoral_Act_or_Goal</i>
Die Renten in Deutschland sind viel zu niedrig.	<i>Immoral_Act_or_Goal</i>

Point of view

The question of what should be considered as dispreferred or blameworthy is, of course, highly subjective.  
**We always annotate the speaker's point of view (and not our own).**

The example below shows actions that are praiseworthy from the view of the speaker but might be blameworthy for many others. However, we always code the speaker's point of view.

- (1) **Atommülltransporte schützen**<sub>*Moral\_Act\_or\_Goal*</sub>, **gewalttätige Demonstranten neutralisieren**<sub>*Moral\_Act\_or\_Goal*</sub>, **Grenzen kontrollieren und schützen**<sub>*Moral\_Act\_or\_Goal*</sub> und **abgelehnte Asylbewerber abschieben**<sub>*Moral\_Act\_or\_Goal*</sub>:

*Context:* Wer diese Aufgaben in Deutschland engagiert umsetzt, wird automatisch zum Feindbild der Grünen, und dass das so ist, beweist der hier vorgelegte Gesetzentwurf.

#### *Political\_Act\_Or\_Goal*

For the *Political\_Act\_Or\_Goal* class, we consider

- proper names of policy issues / political goals (*the COVID-19 vaccination mandate, the climate tax, the animal welfare tax*)
- very concrete political actions or goals (*den Soli abschaffen, die AKWs abschalten, den CO<sub>2</sub>-Ausstoß verringern*).

**Attention:** Instances like the ones below do not refer to concrete policies and are thus treated as *Moral\_Act\_or\_Goal*, not as *Political\_Act\_Or\_Goal*:

- Wir müssen **dafür sorgen, dass die Rentner endlich mehr Geld im Portemonnaie haben**<sub>*Moral\_Act\_or\_Goal*</sub>.
- Wir **fordern mehr Unterstützung für die von den Maßnahmen betroffenen Arbeitgeber**<sub>*Moral\_Act\_or\_Goal*</sub>.
- Unsere Aufgabe ist es, endlich wieder **für krisensichere Verhältnisse in Bezug auf die Familien und Arbeitsmarktpolitik zu sorgen**<sub>*Moral\_Act\_or\_Goal*</sub>.
- Wir müssen **mehr fürs Klima tun**<sub>*Moral\_Act\_or\_Goal*</sub>.
- Wir müssen **das Pariser Klimaabkommen einhalten**<sub>*Political\_Act\_Or\_Goal*</sub>.

**Please note** that *Political\_Act\_Or\_Goal* behaves differently from the moral frames. We expect that policies are often discussed in a differentiated manner, often without the speaker taking a clear position (see below).

- Ich verstehe ja, dass man unruhig ist, weil nichts passiert, und zwar in allen Punkten. Bei **der Tierwohlkennzeichnung**<sub>*Political\_Act\_Or\_Goal*</sub> weiß keiner, was kommt. (speaker: Greens)
- Es ist auch absurd, **das Gebäudeenergiegesetz**<sub>*Political\_Act\_Or\_Goal*</sub> aufzugeben. Es kommt zu spät und will zu wenig; das ist unsere Kritik als Linke.
- Das, was Sie **Einwanderungsgesetz**<sub>*Political\_Act\_Or\_Goal*</sub> nennen, ist doch nur eine Zusammenfassung des Bisherigen. (speaker: FDP)

(i) In contrast to the moral frames, *Political\_Act\_Or\_Goal* does not require the speaker's perspective: In the example below, the goal is not the one expressed by the speaker but one promoted by another party. The speaker's own stance towards this goal remains unclear. We therefore do not annotate any moral roles for this particular instance below.

- Die FDP fordert **eine Coronaelternzeit, die einen Rechtsanspruch auf Arbeitszeitreduzierung mit entsprechendem Kündigungsschutz und staatliche Verdienstaufschüttung für Eltern beinhaltet**<sub>*Political\_Act\_Or\_Goal*</sub>.

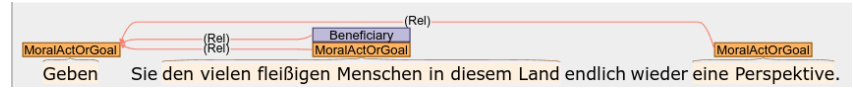
(ii) In contrast to our moral frames, the *Political\_Act\_Or\_Goal* label does not have a positive/negative polarity as it is often used without a strong moral framing (but we can infer this information from the roles linked to the frame, if the frame is morally loaded):

- **Meine Partei**<sub>Hero</sub> fordert **eine Coronaalternzeit, die einen Rechtsanspruch auf Arbeitszeitreduzierung mit entsprechendem Kündigungsschutz und staatliche Verdienstausschüttung für Eltern beinhaltet**<sub>Political\_Act\_Or\_Goal</sub>.
- Und ausgerechnet zu diesem Zeitpunkt **verhängte**<sub>Political\_Act\_Or\_Goal</sub> **die Regierung**<sub>Villain</sub> in einer Nacht-und-Nebel-Aktion **einen beispiellosen Lockdown**<sub>Political\_Act\_Or\_Goal</sub> über Deutschland.

More German examples are shown below.

Example	Label
<i>NPs</i>	
die Tierwohlabgabe	<i>Political_Act_Or_Goal</i>
der Solidaritätszuschlag	<i>Political_Act_Or_Goal</i>
die Einführung der Reichensteuer	<i>Political_Act_Or_Goal</i>
<i>Actions (VPs)</i>	
<b>die Tierwohlabgabe einführen</b>	<i>Political_Act_Or_Goal</i>
endlich <b>den Soli abschaffen</b>	<i>Political_Act_Or_Goal</i>
<b>die Bremse bei der Solarenergie aufheben</b>	<i>Political_Act_Or_Goal</i>
<b>die Steuern für die kleinen Leute erhöhen</b>	<i>Political_Act_Or_Goal</i>

Group mentions like "fleißige Bürger“, "kriminelle Ausländer“, "islamistischen Mörderbanden“ are not encoded as *Moral\_Act\_or\_Goal* or *Immoral\_Act\_or\_Goal* but can be role fillers for moral frames (see section 5 on moral roles below).



### 3 Identifying moral framing

#### 3.1 What counts as moral framing?

There is no straightforward answer to this question, and previous work has reported negative IAA for the binary distinction of whether a sentence includes moral language or not [Shahid et al., 2020]. Our definition of what we consider as *Moral\_Value* / *Moral\_Act\_or\_Goal* and *Immoral\_Value* / *Immoral\_Act\_or\_Goal* is as follows.

As moral framing, we consider statements that describe

- violations of a social norm;
- goals or intentional acts that have the potential to harm people (or fail to provide help) or that accept that others are harmed while pursuing a goal or performing an action.

The concept of harming someone is not restricted to physical harm but might include the violation of someone's feelings or interests through disloyalty or betrayal, restricting someone's rights or freedom or denying someone equal opportunities.

#### 3.2 Which frames to annotate?

Act of God vs  
moral framing

We distinguish acts of God from moral framing, e.g., example 2a describes an event where people experienced harm. However, the event does not



include any intentional action by a human actor. We thus do not annotate this instance. In contrast, example 2b does include moral framing: Here the speaker blames the Social Democratic Parte (SPD) for not protecting to the people in Ahrtal, thus causing harm.

- (2) a. Der Sturm hat enorme Schäden angerichtet.
- b. Das Komplett-Versagen **der SPD**<sub>Villain</sub> hat im Ahrtal **zahlreiche Tote gefordert**<sub>Immoral\_Act\_or\_Goal</sub>.

#### Evaluation

For copula sentences that include evaluation, it is sometimes not clear what should be considered as the moral frame (see Example 3). For those sentences, we do not encode the evaluative expression but, instead, the topic/theme/target of the evaluation:

- (3) **Ihre Finanzpolitik**<sub>Immoral\_Act\_or\_Goal</sub> ist *zutiefst korrupt*<sub>Evaluation</sub>.
- (4) **Ihre korrupte Finanzpolitik**<sub>Immoral\_Act\_or\_Goal</sub> ist *ein Schlag ins Gesicht der Bürger*<sub>Evaluation</sub>.

Sometimes the evaluation is part of the annotated span, e.g., for attributive adjectives. For those cases, the evaluation is part of the moral frame.

- (5) **Ihre korrupte Finanzpolitik**<sub>Immoral\_Act\_or\_Goal</sub> werden wir nicht hinnehmen.

In the next example, the target of the evaluation is a person and can thus neither be *Immoral\_Value* nor *Immoral\_Act\_or\_Goal*. While *bis ins Mark korrupt sein* can also be interpreted as an evaluation, it also describes the immoral behaviour of the target (*Merkel*). We therefore treat these cases as follows:

- (6) Merkel **ist bis ins Mark korrupt**<sub>Immoral\_Act\_or\_Goal</sub>.

#### Precision and clarity of moral frames

Frames that we consider for annotation should be precise and clear. We only annotate frames that:

- explicitly carry a precise, morally loaded statement (e.g., to (not) help someone, to steal someone’s freedom) or goal
- leave it perfectly clear to the hearer/reader how roles are assigned or would be assigned in the case of a hypothetical act being realized (e.g., who is to blame/praise, who is the victim/beneficiary). The roles do not have to be explicitly filled in the text.

In short, we restrict our annotations of moral frames to clear and comprehensible (im)moral statements.

#### Empty phrases

We do not annotate platitudes and empty phrases (de: Floskeln) such as the following:

- sich ehrlich machen

- der Moment der Wahrheit
- zur Ehrlichkeit gehören

Likewise, we only annotate statements focussing on human incompetence (see example 7 below) or human misconduct when they can be assigned to a specific moral foundation. If that is not the case, we consider them as an evaluative statement without moral relevance.

- (7) An Verkehrsminister Scheuer kann man sehen, wie Sie es schaffen, sich im Zeitraffer in selbst geschaffenen Problemen festzufahren.

However, we do annotate frames that encode acts with a clear moral connotation or loading. The table below shows examples of actions that we annotate and those that we do not annotate.

We don't annotate:	We do annotate:
	<i>Acts</i>
von etwas ablenken	<b>etwas verdecken</b> <i>Immoral_Act_or_Goal</i>
seiner Verantwortung nicht gerecht werden	<b>etwas leugnen</b> <i>Immoral_Act_or_Goal</i>
einen kühlen Kopf bewahren	<b>Taschenspielertricks</b> <i>Immoral_Act_or_Goal</i>
das Gegenteil bewirken	<b>etwas aufdecken</b> <i>Moral_Act_or_Goal</i>
	<b>etwas ans Licht bringen</b> <i>Moral_Act_or_Goal</i>
	<i>Goals</i>
Den öffentlich-rechtlichen Rundfunk ...	
neu gestalten	<b>freier gestalten</b> <i>Moral_Act_or_Goal</i>
	<b>diverser gestalten</b> <i>Moral_Act_or_Goal</i>
	<b>feministischer gestalten</b> <i>Moral_Act_or_Goal</i>

### 3.3 Tests

The following questions should help the annotators to decide whether or not a text conveys moral framing.

1. Does the speaker assign blame to a person or group?

⇒ *Immoral\_Value* / *Immoral\_Act\_or\_Goal*

*Examples:*

- (8) a. Wir verurteilen die Pläne **der Regierung***Villain*, **die Schulen zu schließen***Immoral\_Act\_or\_Goal*  
b. Heute diskutieren wir den Antrag der SPD, die Schulen zu schließen. (*no moral framing*)

2. Does the speaker frame a political goal, action or moral concept as blameworthy?

⇒ *Immoral\_Value* / *Immoral\_Act\_or\_Goal*

*Examples:*

- (9) a. Der Fall Greta Thunberg ist von höchster Symbolkraft für **die wahnhafte Klimarettungspolitik***Immoral\_Act\_or\_Goal*.  
b. Klimapolitik ist mehr als nur Energiepolitik. (*no moral framing*)

3. Does the speaker frame a political goal, action, behaviour or moral concept as a violation of a social norm?  $\Rightarrow$  *Immoral\_Value* / *Immoral\_Act\_or\_Goal*

Examples:

- (10) a. **Homosexualität**<sub>*Immoral\_Value*</sub> ist unnatürlich und moralisch falsch<sub>*Evaluation*</sub>.  
 b. Homosexualität wird in einigen Ländern mit dem Tode bestraft. (*no moral framing*)  
 c. Sie haben in Ihrer Presserklärung zwar zu Recht auf das himmelschreiende Unrecht hingewiesen, dass in einigen Ländern **Homosexuelle mit dem Tode bedroht**<sub>*Immoral\_Act\_or\_Goal*</sub> werden.

In example 10a, the speaker states the view that homosexuality is a violation of a social norm. While example 10b could also be interpreted as reporting a violation of a social norm that is punished by death in some countries, the statement does not include any moral framing from the speaker's point of view. In contrast, example 10c clearly frames death penalty for homosexuals as an act of injustice and is thus annotated as *Immoral\_Act\_or\_Goal*, with *homosexuals* coded as the victim of this frame.

4. Does the speaker describe a person or group as praiseworthy?  $\Rightarrow$  *Moral\_Value* / *Moral\_Act\_or\_Goal*

Examples:

- (11) a. **die NATO**<sub>*Hero*</sub> ist auch **ein Wertebündnis**<sub>*Moral\_Value*</sub>  
 b. Putin würde sich drüber freuen, wenn die Türkei die NATO verlässt. (*no moral framing*)  
 c. **Wir**<sub>*Hero*</sub> **kämpfen dafür, dass es in diesem Land wieder gerecht zugeht**<sub>*Moral\_Act\_or\_Goal*</sub>.  
 d. Wir werden für Ihren Antrag stimmen. (*no moral framing*)

5. Does the speaker describe a concept, goal or action as praiseworthy?  $\Rightarrow$  *Moral\_Value* / *Moral\_Act\_or\_Goal*

Examples:

- (12) a. **Die Linke in Deutschland**<sub>*Hero*</sub> ist solidarisch mit **diesem Aufstand für mehr Demokratie**<sub>*Moral\_Act\_or\_Goal*</sub>.  
 b. Wir müssen uns **für eine Verkehrswende starkmachen**<sub>*Moral\_Act\_or\_Goal*</sub>.  
 c. Die Grünen fordern seit Jahren eine Verkehrswende. (*no moral framing*)

Please note that example 12c can be interpreted as praiseworthy if the speaker is a member of the green party and in favour of the Verkehrswende. If it is said by a member of another party, this sentence is merely descriptive.

## 4 Annotation span

This section describes how we determine the frame spans of moral frames for a number of frequent linguistic constructions.

### Minimal span principle

We annotate the minimal span that expresses the meaning of a moral frame. In the example below, the span in Ex. 13a is sufficient to capture the frame meaning. In other cases, adjuncts may have to be included in the span to capture the meaning of the frame (Ex 13b).

- (13) a. Eine Vormachtstellung in den sozialen Medien und im Bildungswesen darf nicht dazu missbraucht werden, **die politische Willensbildung einseitig zu beeinflussen**<sub>Immoral\_Act\_or\_Goal</sub>.
- b. Sie sind **angeklagt wegen des Versuchs, ein Verfassungsurteil in diesem Land umzusetzen**<sub>Immoral\_Act\_or\_Goal</sub>

For the annotation of moral frames that are VPs, we consider the core arguments, i.e., direct/indirect objects + verb (finite full verb or infinite verb or participle) as part of the frame span:

- **die Schwachen in unserer Gesellschaft besser schützen**

We do not include modifier PPs etc. in the frame span (or only in cases where the removal of the modifier would change the meaning of the moral frame). However, modifiers are often role fillers, as shown in the example below.

- **die sozialen und finanziellen Auswirkungen gerade auch für ärmere Familien**<sub>Beneficiary</sub> **abzufedern**.

### Passive

In passive sentences the subject is included in the frame span.

- Mit Sorge nehmen wir wahr, dass **Journalist\*innen** in vielen Teilen der Welt zunehmend durch staatliche Institutionen und Amtsträger **in ihrer Freiheit beeinträchtigt** werden.

### Auxiliaries

- We do not include finite auxiliaries or modal verbs in the annotation span:
  - Wir müssen **die Schwachen in unserer Gesellschaft besser schützen**
  - Wir werden **die Schwachen in unserer Gesellschaft besser schützen**
  - Wir haben **die Schwachen in unserer Gesellschaft besser geschützt**
  - damit **die Schwachen in unserer Gesellschaft besser geschützt sind**
- We do include finite full verbs in the annotation:
  - Die Linke **schützt die Schwächsten in unserer Gesellschaft**

- Wir **schützen die Schwächsten in unserer Gesellschaft**
- Wir **setzten uns ein für die Schwächsten in unserer Gesellschaft**
- weil die Linke **sich für die Schwächsten in unserer Gesellschaft einsetzt**

Relative clauses are included in the frame span.

- (14) **Die traditionelle Familie, die aus Vater, Mutter und Kindern besteht**<sub>Moral\_Value</sub>, muss endlich wieder zum Leitbild für unsere Gesellschaft werden.
- (15) In **normalen Familien, in denen ein Elternteil zuhause bleibt, um sich um die Kinder zu kümmern**<sub>Moral\_Value</sub>, können solche Probleme besser aufgefangen werden.
- (16) Speziell **der emotionale Kreuzzug gegen das Automobil, den Sie ja alle betreiben und der von den Grünen und von den Linken befeuert wird**<sub>Immoral\_Act\_or\_Goal</sub>, ist unverantwortlich für eine moderne Industriegesellschaft.

PP attachment is often hard to resolve. In the example below, it is not clear whether "in muslimischen Staaten" (in Muslim states) should be attached to the verb "stattfinden" (take place), to the coordinated NP "Verfolgung, Vertreibung und Ermordung" (persecution, expulsion and murder), or rather to the modifier PP "von Christen" (of Christians). We select the most probable reading and encode it as the frame span.

- (17) a. Mit der **Verfolgung von Christen**<sub>Immoral\_Act\_or\_Goal</sub> in muslimischen Staaten findet seit Jahren **eine ethnisch-religiöse Säuberung**<sub>Immoral\_Act\_or\_Goal</sub> statt, und das christlich geprägte Europa **schaut zu**<sub>Immoral\_Act\_or\_Goal</sub>.
- b. Mit der **Verfolgung von Christen in muslimischen Staaten**<sub>Immoral\_Act\_or\_Goal</sub> findet seit Jahren **eine ethnisch-religiöse Säuberung**<sub>Immoral\_Act\_or\_Goal</sub> statt, und das christlich geprägte Europa **schaut zu**<sub>Immoral\_Act\_or\_Goal</sub>.

Negation can shift the polarity of the label for both, NP and VP frames:

- (18) a. Das ist **differenzierte Rechtspolitik**<sub>Moral\_Act\_or\_Goal</sub>.
- b. Das ist doch **keine differenzierte Rechtspolitik**<sub>Immoral\_Act\_or\_Goal</sub>.
- c. So sieht **ein fairer Umgang mit dem Partner**<sub>Moral\_Act\_or\_Goal</sub> aus.
- d. Das ist **kein fairer Umgang mit einem Partner**<sub>Immoral\_Act\_or\_Goal</sub>.
- e. Wir **fordern Unterstützung**<sub>Moral\_Act\_or\_Goal</sub> für die von den Maßnahmen betroffenen Arbeitgeber.
- f. Die FDP **fordert keine Unterstützung**<sub>Immoral\_Act\_or\_Goal</sub> für die von den Maßnahmen betroffenen Arbeitgeber.

Embedded values

(Im)moralActOrGoals often include moral values that are violated (see below: "das Auskunftsrecht und das Recht von Betroffenen" (the right to information and the right of those affected) is framed as a moral value that should be protected). However, we do not annotate the embedded values if they are part of a moral or immoral act or goal. Instead, we only encode the *Moral\_Act\_or\_Goal* or *Immoral\_Act\_or\_Goal*.

- (19) Das kann doch nicht Ihr Ernst sein, **das Auskunftsrecht und das Recht von Betroffenen auf Berichtigung oder Löschung ihrer Daten auszusetzen**<sub>Immoral\_Act\_or\_Goal</sub>.

Placeholder-*es* For constructions with a placeholder-*es*, we encode the main verb, the placeholder and the embedded clause that is linked to the placeholder:

- Sie haben **es versäumt, Kindern und Jugendlichen eine Perspektive zu geben**<sub>Immoral\_Act\_or\_Goal</sub>
- Die Regierung **versäumt es**<sub>Immoral\_Act\_or\_Goal</sub> mal wieder, **Kindern und Jugendlichen eine Perspektive zu geben**<sub>Immoral\_Act\_or\_Goal</sub>

Nominalisations

Typically, (Im)MoralActOrGoals have the syntactic structure of a VP. We do annotate nominalisations that describe moral framing as (Im)MoralActOrGoal (and not as (Im)MoralValue), when their meaning is equivalent to that of a VP expressing the same act or goal.

- **die Spendenabzocker**<sub>Immoral\_Act\_or\_Goal</sub>
- **Spenden abzocken**<sub>Immoral\_Act\_or\_Goal</sub>

#### 4.1 When do we annotate the whole clause instead of a VP?

Often, moral actions are expressed as a VP where the subject fills the role of the *Villain* or *Hero*. For instances where the subject cannot be assigned any role but is part of the (im)moral act or goal, we encode the whole sentence as a moral frame:

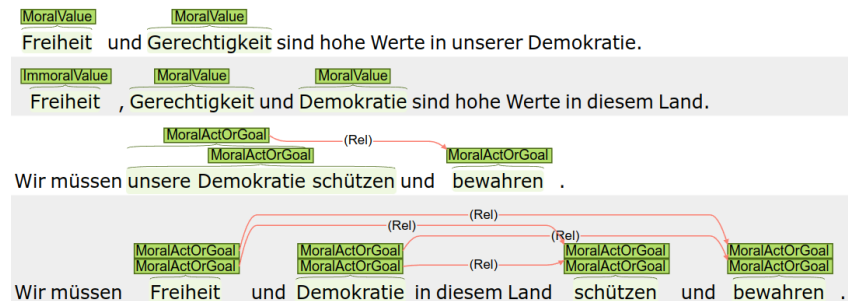
- VP: **Unsere Kultureinrichtungen**<sub>Villain</sub> **kümmern sich nicht ausreichend um Kinder aus schwierigen sozialen Verhältnissen**<sub>Immoral\_Act\_or\_Goal</sub>.

vs.

- **Kultureinrichtungen**<sub>Moral\_Act\_or\_Goal</sub> müssen **barrierefrei sein**<sub>Moral\_Act\_or\_Goal</sub>.

#### 4.2 Coordination

Coordination can result in annotation spans for moral values or actions where the different conjuncts evoke different moral foundations (e.g.: Freiheit (freedom) → Liberty; Gerechtigkeit (justice) → Equality). We therefore annotate each conjunct as a separate frame span (see examples below).



We do **not** split the annotation if the coordinated NP is not the head of the NP but embedded in a modifying NP, PP, or an apposition. In that case, we annotate the coordination as one frame only (see examples below):

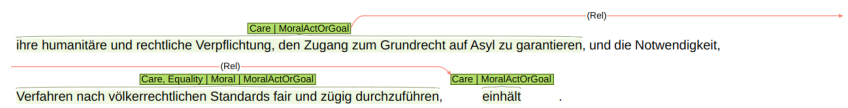
- (20) Treten Sie **unserem Bündnis für kulturelle Bildung und Vermittlung sowie Medienkompetenz**<sub>Moral\_Act\_or\_Goal</sub> bei.
- (21) Dies ist **ein Angriff auf die Vertrauenswürdigkeit und die Integrität der IT**<sub>Immoral\_Act\_or\_Goal</sub>.
- (22) Sie **lässt Personen unerkannter Identität rein – Vergewaltiger, Mörder, Terroristen inklusive**<sub>Immoral\_Act\_or\_Goal</sub>

We also do not split the frame annotations for the following cases:

- gender-neutral language via double mention (Ex. 23)
  - cases in which the conjuncts clearly represent one frame such that a separation would render the frame unintelligible (Ex. 24)
- (23) Hier wurden ja **viele Kolleginnen und Kollegen attackiert**<sub>Immoral\_Act\_or\_Goal</sub>.
- (24) Solche Menschen **sagen das Eine und tun das Andere**<sub>Immoral\_Act\_or\_Goal</sub>.

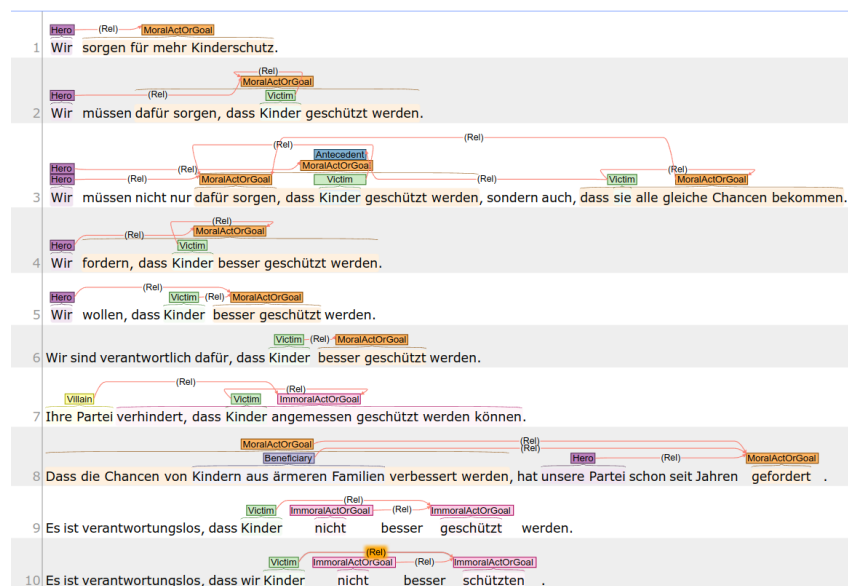
### 4.3 Complement and attributive clauses

Complement and attributive clauses are part of the annotated span. See example below:



*that*-complement clauses

The examples below illustrate the treatment of *that*-complement clauses



Label	Description
<i>Moral roles</i>	
<i>Hero</i>	Person/group/organisation performing or persuing a commendable, responsible or praiseworthy act or goal
<i>Villain</i>	Person/group/organisation performing an act portrayed as immoral, irresponsible or otherwise dispreferred
<i>Victim</i>	Person/group that is harmed or disadvantaged by an action
<i>Beneficiary</i>	Person/group that benefits from an action or goal

Table 2: Inventory of moral roles with examples.

## 5 Annotation of moral roles

The annotation of moral roles is inspired by the Narrative Policy Framework [Shanahan et al., 2017] and considers four different (im)moral roles, *Hero*, *Beneficiary*, *Victim* and *Villain* (see Table 2).

Annotation span for roles

For role annotation, we always consider the largest NP, including modifiers (Ex. 25b) and complement clauses (Ex. 25c). Coordinated conjuncts are annotated as separate roles (Ex. 25d).

- (25) a. **Die NATO**<sub>Hero</sub> hat (**Europa**<sub>Beneficiary</sub> **gegen die militärische Bedrohung durch (die Sowjetunion)**<sub>Villain</sub> **geschützt**)<sub>Moral\_Act\_or\_Goal</sub>.
- b. **Die siegreiche, tapfere NATO**<sub>Hero</sub> hat (**das Europa der Herzen**<sub>Beneficiary</sub> **gegen die militärische Bedrohung durch (die bis an die Zähne bewaffnete Sowjetunion)**<sub>Villain</sub> **geschützt**)<sub>Moral\_Act\_or\_Goal</sub>.
- c. **Die NATO, die schon immer gern mit dem Feuer spielt**<sub>Hero</sub> hat (**Europa**<sub>Beneficiary</sub> **gegen die militärische Bedrohung durch (die Sowjetunion)**<sub>Villain</sub> **geschützt**)<sub>Moral\_Act\_or\_Goal</sub>.
- d. **Die NATO**<sub>Hero</sub> hat (**Europa**<sub>Beneficiary</sub> **und den Rest der Welt**<sub>Beneficiary</sub> **gegen die militärische Bedrohung durch die Sowjetunion**<sub>Villain</sub> **geschützt**)<sub>Moral\_Act\_or\_Goal</sub>.

Adjectives

can also fill a moral role (Ex. 26a). Please note that we do not annotate subword components as roles (e.g., "familienfreundlich" (family-friendly), "bürgernah" (citizen-oriented), "ausländerfeindlich" (anti-foreigner), etc., see Ex. 26b)

- (26) a. (**(die (russische))**<sub>Villain</sub> **Aggression gegen (die Ukraine)**<sub>Victim</sub>)<sub>Immoral\_Act\_or\_Goal</sub>
- b. (**(die ausländerfeindliche Politik (der AfD))**<sub>Villain</sub>)<sub>Immoral\_Act\_or\_Goal</sub>

### 5.1 Hero

The Hero is the agent who is n performing or persuing a commendable, responsible or praiseworthy act or goal. We restrict the annotation of villains to sentient human beings but also consider metonymic expressions.



### 5.1.1 When do we assign the Hero role?

We only annotate the Hero role when the agent fulfills this role, but not in hypothetical situations where a person or group is asked to do something without intending to do it. For illustration, consider the next example where "The new federal government" does not actually assume the Hero role:

- (27) The new federal government must defend human rights and the right to asylum. (request by a member of the opposition)

The same applies to the next example. Here, too, we do not regard the "we" as a Hero, because the sentence merely expresses an unspecific wish without any concrete prospect of realisation:

- (28) We want a humanitarian admission policy.

We also do not annotate the Hero role in hypothetical statements, suggestions and proposals:

- (29) Groups or associations can pledge to support refugees, helping people through resettlement and resettlement opportunities.  
(30) We suggest doing X.  
(31) With this program, the German government would help many people.

However, if the (im)moral action is formulated as a promise, concrete action or demand, then we can assign the Hero role, as shown in the following examples:

- (32) **We<sub>Hero</sub>** will again **set up more humanitarian admission programs as well as quotas from EU states with external borders**<sub>Moral\_Act\_or\_Goal</sub>.  
(33) **We<sub>Hero</sub>** are committed to ensuring that **((more people)<sub>Beneficiary</sub> receive protection)**<sub>Moral\_Act\_or\_Goal</sub>.

## 5.2 Villain

The Villain is the agent who is (potentially) causing harm to a person or entity by performing an act that is portrayed as immoral, irresponsible or otherwise dispreferred. We restrict the annotation of villains to sentient human beings but also consider metonymic expressions.

### 5.2.1 When do we assign the Villain role?

For a Villain to be annotated, the person/group must be explicitly mentioned or identifiable from context. For illustration, consider the next example which describes a hypothetical situation. Here we do not assign a Villain role, since there is no real-life agent performing an immoral act and the agent of the immoral act is not identifiable.

- (34) Wer **X<sub>Immoral\_Act\_or\_Goal</sub>** tut, sollte nicht in Deutschland bleiben dürfen.

In the next example, however, the person/group who is blamed to have committed/is planning to commit an immoral act is explicitly addressed by the speaker and thus clearly identifiable. We consider the intention to commit an immoral act as sufficient for assigning the Villain role.

- (35) And there **you**<sub>Villain</sub> want to **build a penitentiary**<sub>Immoral\_Act\_or\_Goal</sub>.

### 5.3 Victim

The Victim is the being or entity that is put at risk, endangered, disadvantaged or injured. We restrict the annotation of victims to sentient beings but also consider metonymic expressions.

#### 5.3.1 When do we assign the Victim role?

We only encode the Victim role when the victim status is explicitly expressed (see examples below). The distinction between *Victim* and *Beneficiary* usually depends on the semantics of the predicate (see the difference between "die Opfer schützen" (protect the victims) and "die Opfer unterstützen" (support the victims)).

- (36) Sie wollen ((den Menschen)<sub>Victim</sub> **ihre Auto und damit ihre Freiheit wegnehmen**)<sub>Immoral\_Act\_or\_Goal</sub>.
- (37) damit ((die Opfer)<sub>Victim</sub> **geschützt**)<sub>Moral\_Act\_or\_Goal</sub> werden
- (38) **(die Not (der Geflüchteten))**<sub>Victim</sub> **lindern**)<sub>Moral\_Act\_or\_Goal</sub>
- (39) **((Viele Flüchtlinge))**<sub>Victim</sub> **gefoltert**<sub>Immoral\_Act\_or\_Goal</sub> wurden

### 5.4 Beneficiary

The Beneficiary is the being or entity that benefits from an action or goal. We restrict the annotation of beneficiaries to sentient beings but also consider metonymic expressions.

#### 5.4.1 When do we assign the Beneficiary role?

We only annotate the Beneficiary role when the beneficiary status is explicitly expressed (see examples below). The distinction between *Victim* and *Beneficiary* usually depends on the semantics of the predicate (see the difference between "die Opfer schützen" and "die Opfer unterstützen").

- (40) **((Opfer von Rassismus))**<sub>Beneficiary</sub> **unterstützen**)<sub>Moral\_Act\_or\_Goal</sub>
- (41) Wir wollen **((den Menschen))**<sub>Beneficiary</sub> **und der Umwelt in Deutschland helfen**)<sub>Moral\_Act\_or\_Goal</sub>.

Complex example from the political manifesto of Die Linke (2021) on the topic of migration:

”Zivilgesellschaftliche Gruppen, die sich gegen Rassismus, Antisemitismus, Homo- und Transfeindlichkeit, Antiziganismus, religiösen Fundamentalismus, antimuslimischen Rassismus und für mehr Demokratie engagieren, sowie Flüchtlingsräte, migrantische Verbände, selbstverwaltete Beratungsangebote und die Selbstorganisation von Migrant\*innen wollen wir durch ein Demokratieförder gesetz stärker und endlich dauerhaft fördern.“

English translation:

”We want to provide stronger and lasting support for civil society groups that are committed to combating racism, anti-Semitism, homophobia and trans-hostility, antiziganism, religious fundamentalism, anti-Muslim racism and for more democracy, as well as refugee councils, migrant associations, self-managed advisory services and the self-organisation of migrants through a democracy promotion law.“

Moral frames that we consider for the example above:

- **Frame 1:** ((Zivilgesellschaftliche Gruppen, die sich gegen Rassismus, Antisemitismus, Homo- und Transfeindlichkeit, Antiziganismus, religiösen Fundamentalismus, antimuslimischen Rassismus und für mehr Demokratie engagieren)<sub>Beneficiary</sub> stärker und endlich dauerhaft fördern)<sub>Moral\_Act\_or\_Goal</sub>.
- **Frame 2:** ((Flüchtlingsräte)<sub>Beneficiary</sub> stärker und endlich dauerhaft fördern)<sub>Moral\_Act\_or\_Goal</sub>.
- **Frame 3:** ((migrantische Verbände)<sub>Beneficiary</sub> stärker und endlich dauerhaft fördern)<sub>Moral\_Act\_or\_Goal</sub>.
- **Frame 4:** (selbstverwaltete Beratungsangebote stärker und endlich dauerhaft fördern)<sub>Moral\_Act\_or\_Goal</sub>.
- **Frame 5:** (die Selbstorganisation von (Migrant\*innen)<sub>Beneficiary</sub> stärker und endlich dauerhaft fördern)<sub>Moral\_Act\_or\_Goal</sub>.

We do **not** annotate *Selbstorganisation* (self-organisation) and *Beratungsangebot* (advisory services) with the label *Beneficiary*.

## 5.5 Antecedent

To resolve coreferences for roles filled by pronouns or underspecified NPs (typically including a demonstrative pronoun, e.g., “diese Menschen” (these people), “diese 80-Jährigen” (these 80-year-olds), we also annotate their antecedents and link the role to its antecedent by drawing a relation from the role to the antecedent.

As antecedent, we chose the closest NP that disambiguates the pronoun or NP. We do not annotate coreference chains with more elements.

Please note that we only annotate antecedents for roles, not for frames. The two examples below illustrate this. In the first example, we annotate the antecedent of the Villain and link the two entities. In the second example, however, the antecedent “sie” (them) is not part of a role but of the moral frame. We therefore do not annotate the antecedent:

- (42) Haben **Menschenrechtsverletzer**<sub>Antecedent</sub> keine Konsequenzen zu befürchten, dann **verletzen**<sub>Immoral\_Act\_or\_Goal</sub> **sie**<sub>Villain</sub> erneut

**Menschenrechte**<sub>Immoral\_Act\_or\_Goal</sub>.

- (43) **Menschenrechte**<sub>Immoral\_Act\_or\_Goal</sub> werden überall auf der Welt mit **Füßen getreten**<sub>Immoral\_Act\_or\_Goal</sub>. Wir müssen **sie besser schützen**<sub>Moral\_Act\_or\_Goal</sub>.

The following examples illustrate the annotation of antecedents for personal pronouns (ex. 44a), possessive pronouns (ex. 44b) and NPs (ex. 44c).

- (44) a. **Menschen mit Migrationsgeschichte**<sub>Antecedent</sub> sind nach Generationen noch häufiger von Armut, Ausgrenzung und Arbeitslosigkeit betroffen. **Sie**<sub>Victim</sub> werden im Bildungssystem systematisch benachteiligt.
- b. **Menschen mit Migrationsgeschichte**<sub>Antecedent</sub> sind nach Generationen noch häufiger von Armut, Ausgrenzung und Arbeitslosigkeit betroffen. **Ihre Kinder**<sub>Victim</sub> werden im Bildungssystem systematisch benachteiligt
- c. Auch mich belastet ganz besonders, was **die Menschen**<sub>Antecedent</sub> erdulden müssen, **die in Pflege-, Senioren-, Behinderteneinrichtungen leben**<sub>Antecedent</sub>. Dort, wo Einsamkeit ohnehin zum Problem werden kann, ist es in Zeiten der Pandemie und ganz ohne Besuch noch viel einsamer. [...] **Diese 80-, 90-Jährigen**<sub>Hero</sub> haben **unser Land aufgebaut**<sub>Moral\_Act\_or\_Goal</sub>.

Coordination We do not annotate the different constituents of coordinated antecedents as separate frames but consider them as one frame.

- **(Menschen mit Migrationsgeschichte, Geflüchtete und Geduldete)**<sub>Antecedent</sub> sind häufiger von Armut, Ausgrenzung und Arbeitslosigkeit betroffen. **Sie**<sub>Victim</sub> werden im Bildungssystem systematisch benachteiligt.
- **((Opfer von Rassismus und ihre Angehörige)**<sub>Beneficiary</sub><sub>Moral\_Act\_or\_Goal</sub> müssen besser **unterstützt**<sub>Moral\_Act\_or\_Goal</sub> werden

External antecedents In some rare cases, we find null-instantiations where the antecedent of a pronoun or NP can be inferred from context or world knowledge but is not overtly expressed in the document, thus making it impossible to link antecedent and moral role.

This is illustrated in the example below where the party AfD can be inferred from context as the antecedent of the pronoun “Sie” (you) but has not been overtly expressed in the document. For those cases, we insert the antecedent in the free text field of the annotation layer **ExternalAntecedent**.

- (45) Wir können – das will ich abschließend sagen – dem AfD-Antrag nicht zustimmen. **Sie**<sub>Villain</sub> **bieten den Leuten Steine statt Brot**<sub>Immoral\_Act\_or\_Goal</sub>.

## 6 Annotation of Moral Foundations








							
	Care	Equality	Proportionality	Liberty	Loyalty	Authority	Purity
<b>Intuition</b>	Avoiding emotional and physical damage to another individual	Equal treatment and equal outcome for individuals	Individuals getting rewarded/punished in proportion to their merit/contribution	Promoting individual freedom and fighting oppression	Cooperating with ingroups and competing with outgroups	Deference toward legitimate authorities and defense of traditions, stability, fending off chaos	Avoiding bodily and spiritual contamination and degradation
<b>Adaptive Challenge</b>	Protect and care for children	Reap benefits of two way partnerships		Keep dominant individuals in the group 'in check'	Form cohesive coalitions	Form beneficial relationships within hierarchies	Avoid contaminants
<b>Original triggers</b>	Distress or neediness expressed by child	Cheating, cooperation, deception		Bullying and constraining others	Threat of challenge to group	Signs of dominance and submission	Waste products, diseased people
<b>Key emotion</b>	Compassion	Anger, gratitude, guilt		Anger at oppression	Group pride, rage against traitors	Respect, fear	Disgust
<b>Relevant virtues</b>	Caring, kindness	Fairness, justice, trustworthiness		Freedom and self-determination, protection of victims	Loyalty, patriotism, self-sacrifice	Obedience, deference	Temperance, chastity, piety, cleanliness

Figure 1: Overview of Moral Foundations considered in our annotation schema (figure adapted from [Haidt, 2012, Graham et al., 2013]).

The main goal of the final annotation subtask is to capture the moral framing of arguments in political debates, based on Moral Foundations Theory (see Section 1 above). We annotate Moral Foundations (MFs) for each (im)moral frame. Annotators can assign at most two MFs to each frame.

### 6.1 Annotation perspective

Importantly, we **do not encode our own moral values** but always annotate the *moral framing* used by the speaker.

We distinguish **the speaker’s moral values** from **the moral framing** used to convey a certain political message. Although there might be a high correlation between the two, they do not necessarily align. We have no means to directly observe and measure the moral values of the speaker,<sup>3</sup> as it is not clear whether and to what extent politicians express their own values or adapt to the party line, or intentionally use a specific framing as a strategic device to achieve a certain effect. We can, however, observe and encode the moral framing in the speech.

**We do not encode the moral values that, from the annotator’s perspective, should be ascribed to the speaker but, instead, encode the moral framing used by the speaker in their speech.**

Below we give instructions and examples for the different Moral Foundations that we consider in the annotation.

<sup>3</sup> This is usually done with the help of psychological measurement tools like the Moral Foundations Questionnaire [Graham et al., 2011].

## 6.2 Care

The Care foundation centers around the following values:

- Empathy for others. An important aspect here is that empathy is experienced universally, for anyone, and not just for the in-group.
- Protecting others (again, universally, not limited to the in-group), reducing or preventing harm.

Below are examples that we annotate as CARE.

- (46)
- a. ensure security in our country
  - b. protect women and girls from assault
  - c. exemption decisions to protect our population

It is important to note that the values of care, empathy and protection apply to all people, not just members of the in-group. Consider ex. 46b above and the modified example below which we consider as an instance of the LOYALTY foundation:

“Protect our women and girls! Protect Germany.” LOYALTY

## 6.3 Equality and Proportionality

The two foundations EQUALITY and PROPORTIONALITY once have been considered as a single foundation named FAIRNESS. However, the concept of fairness can be conceptualised in different ways. For some fairness means that everyone should have the same benefits/rights/resources (EQUALITY), while others think that these benefits/rights/resources should be distributed according to a person’s performance (PROPORTIONALITY).

The following examples illustrate different nuances of FAIRNESS:

Example	Label
A society is fair when income and wealth are equally distributed among all people.	EQUALITY
A society is fair when hard-working people earn more than others.	PROPORTIONALITY
A society is fair when people from families with high social status enjoy privileges in their lives.	PROPORTIONALITY

### 6.3.1 Equality

Equality is about rights such as fairness and justice for everyone (including those who are not part of the in-group), regardless of their performance. It advocates equal opportunities for all and an equal distribution of resources.

Below we list frequent terms and concepts that we encode as EQUALITY.

- (47)
- a. Xenophobia, misogyny, racism, antiziganism (as an attitude/ideology; when expressed as an action, such as *hate against women*, *incitement against foreigners*, then we additionally assign the CARE foundation)
  - b. Racism (when expressed as ‘violent racism’ or ‘racist attack’, then we additionally assign the CARE foundation)
  - c. social injustice

- d. nationalism, a nationalist image of women and the family (when framed as an immoral value)
- e. the upholding of universal human rights, a society that realises human rights (additionally: CARE)
- f. human dignity

### 6.3.2 Proportionality

In this foundation, each person gets what they deserve. Chances and advantages are distributed, based on someone's performance, as is punishment. Penalties should be proportionate to the crime.

Below we list examples of frames that we annotate as PROPORTIONALITY.

- (48)
- a. A degree should also be reflected in your salary later on.
  - b. An eye for an eye, a tooth for a tooth.
  - c. The penalty should not exceed the damage caused.

## 6.4 Loyalty

This foundation centers around the values of loyalty, patriotism and self-sacrifice for the group.

- Protecting members of the in-group and prioritising their advantage over others.
- Avoiding harm or disadvantages for the in-group (including social division)

Below we list examples of frames that we annotate as LOYALTY.

- (49)
- a. **promote social division***Immoral\_Act\_or\_Goal*
  - b. **standing together against social division***Moral\_Act\_or\_Goal*
  - c. to protect our country
  - d. providing service for our safety
  - e. the security of our country
  - f. protect our borders / strong protection of the external borders
  - g. Germany first

## 6.5 Authority

This moral foundation is centered around values of leadership and followership, including deference to prestigious authority figures and respect for traditions.

- Respect for authorities, support for a strong state (country)
- To implement law and order, to fight criminality
- To maintain traditions and hierarchical societal structures
- Support for traditional values, such as respecting the elderly, good manners, obedience.

AUTHORITY frames are in line with the following statements from the *Authoritarianism Short Scale (KSA-3)*:

- We should take strong action against misfits and slackers in society.

- Troublemakers should be made to feel that they are not welcome in society.
- Rules in society should be enforced without pity.
- We need strong leaders so that we can live safely in society.
- People should leave important decisions in society to their leaders.
- We should be grateful for leaders telling us exactly what to do.
- Traditions should definitely be carried on and kept alive.
- Well-established behavior should not be questioned.
- It's always best to do things in the usual way.

**Please note:** Statements that question such values are **not** part of the AUTHORITY foundation but are encoded as LIBERTY.

- Ungehorsam gegen die eigene Regierung ist manchmal gerechtfertigt.  
Disobedience against one's own government is sometimes justified.

Below we list examples of frames that we annotate as AUTHORITY.

- (50)
- being a "trad wife" who makes healthy meals for her man, cleans, and decorates the house with embellishments
  - guarantee internal security as a central task of the state
  - demand strict punishment for false statements in the asylum procedure.
  - accept the rules, traditions and values of our society

## 6.6 Purity

This moral foundation is centered around values that aim at avoiding bodily and spiritual contamination and degradation. This includes foul language and any behaviour considered as immoral that is likely to trigger disgust in the viewer.

We also annotate PURITY for acts of desecration of religious/national symbols or when somebody feels disgust triggered by sexual preferences/fetishes/behaviour that deviate from the social norm, such as adultery, promiscuity, homosexuality, incest etc.

**Please note** that PURITY is not necessarily related to religion.<sup>4</sup>

- (51)
- We need to **protect God's creation***Moral\_Act\_or\_Goal*.
  - I hate it when people **burn our national flag***Immoral\_Act\_or\_Goal*.
  - Homosexuality***Immoral\_Act\_or\_Goal* is against nature.
  - Jewish cemetery desecrated***Immoral\_Act\_or\_Goal* again.
  - Therapeutic fasting: how to **detoxify your body***Moral\_Act\_or\_Goal*.

<sup>4</sup> [Crone and Laham, 2023] have identified measurement problems in the MFQ that "suggest that an average non-religious participant's Purity score are underestimated".



## 6.7 Liberty

This moral foundation includes moral values that promote individual freedom and fight oppression. This includes civil rights/liberties, the rights of citizens vis-à-vis the state, and also values of democratic/informational self-determination or self-determination in general. We also consider references to free democratic forms of government, including the protection thereof, as part of the LIBERTY foundation (in addition to EQUALITY).

- (52) a. strengthen civil rights on the Internet
- b. protection of privacy and free development of personality
- c. data protection and data privacy
- d. curtail people's fundamental rights
- e. the rights of the individual
- f. our open society
- g. the protection of our liberal democracies

**Please note:** Human rights are not part of the LIBERTY foundation but belong to the foundations CARE/EQUALITY, e.g.:

- (53) a. protect human rights at the external borders CARE
- b. the upholding of universal human rights CARE, EQUALITY
- c. a society that realizes human rights CARE, EQUALITY

The term “basic/fundamental rights” is ambiguous between references to civil rights (Bürgerrechte, LIBERTY) and references to human rights (Menschenrechte, CARE). We determine the adequate MF label, depending on the context.

## 6.8 General Morality

We use the label General-Morality for moral frames that are underspecified and cannot be clearly assigned to any MF (see examples below). This includes descriptions of Typical examples are selfish behavior that is only concerned with one's own advantage, trickery and taking advantage.

- honesty, dishonesty
- to trick, cheat (but: cheat the German people LOYALTY)
- to defame someone  
(but: discriminate against/denigrate minorities EQUALITY)
- to promote propaganda, promote vested interests
- to spread an ideology, represent ideological interests
- to protect protect art and culture  
(but: protect traditions AUTHORITY)
- make sure that everything is done properly

## References

- [Atari et al., 2023] Atari, M., Haidt, J., Graham, J., Koleva, S., Stevens, S. T., and Dehghani, M. (2023). Morality beyond the weird: How the nomological network of morality varies across cultures. *Journal of Personality and Social Psychology*, 5(125):1157–1188.
- [Crone and Laham, 2023] Crone, D. L. and Laham, S. M. (2023). Clarifying measurement issues with the purity subscale of the moral foundations questionnaire in christian and nonreligious participants. *Social Psychological and Personality Science*, 14(7):845–853.
- [Graham et al., 2013] Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., and Ditto, P. H. (2013). Chapter two - moral foundations theory: The pragmatic validity of moral pluralism. In Devine, P. and Plant, A., editors, *Advances in Experimental Social Psychology*, volume 47, pages 55–130. Academic Press.
- [Graham et al., 2009] Graham, J., Haidt, J., and Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5):1029–1046.
- [Graham et al., 2011] Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Spassena, K., and Ditto, P. H. (2011). Moral foundations questionnaire (mfq).
- [Haidt, 2012] Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Pantheon.
- [Haidt et al., 2009] Haidt, J., Graham, J., and Joseph, C. (2009). Above and below left-right: Ideological narratives and moral foundations. *Psychological Inquiry*, 20(2-3):110–119.
- [Hämmerl et al., 2023] Hämmerl, K., Deiseroth, B., Schramowski, P., Libovický, J., Rothkopf, C., Fraser, A., and Kersting, K. (2023). Speaking multiple languages affects the moral bias of language models. In Rogers, A., Boyd-Graber, J., and Okazaki, N., editors, *Findings of the Association for Computational Linguistics: ACL 2023*, pages 2137–2156, Toronto, Canada. Association for Computational Linguistics.
- [Klie et al., 2018] Klie, J.-C., Bugert, M., Boullosa, B., de Castillo, R. E., and Gurevych, I. (2018). The inception platform: Machine-assisted and knowledge-oriented interactive annotation. In *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations*, pages 5–9. Association for Computational Linguistics. Event Title: The 27th International Conference on Computational Linguistics (COLING 2018).
- [McMahan, 2000] McMahan, J. (2000). Moral intuition. In -, H. L., editor, *The Blackwell Guide to Ethical Theory*, pages 92–110. Blackwell.
- [Shahid et al., 2020] Shahid, U., Eugenio, B. D., Rojecki, A., and Zhelleva, E. (2020). Detecting and understanding moral biases in news. In *The First Joint Workshop on Narrative Understanding, Storylines, and Events*, pages 120–125, Online. Association for Computational Linguistics.
- [Shanahan et al., 2017] Shanahan, E., Jones, M., Mcbeth, M., and Radaelli, C. (2017). The narrative policy framework. In Weible, C. and Sabatier, P., editors, *The Theories of the Policy Process*, pages 173–213. Boulder, CO: Westview Press.

- [Sinnott-Armstrong et al., 2010] Sinnott-Armstrong, W., Young, L., and Cushman, F. (2010). 246Moral Intuitions. In *The Moral Psychology Handbook*. Oxford University Press.
- [Wu et al., 2023] Wu, W., Wang, L., and Mihalcea, R. (2023). Cross-cultural analysis of human values, morals, and biases in folk tales. In Bouamor, H., Pino, J., and Bali, K., editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 5113–5125, Singapore. Association for Computational Linguistics.
- [Xie et al., 2019] Xie, J. Y., Ferreira Pinto Junior, R., Hirst, G., and Xu, Y. (2019). Text-based inference of moral sentiment change. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4654–4663, Hong Kong, China. Association for Computational Linguistics.