

# The Technical Appendix

## Anonymous submission

### Definitions of Personality Models

- **Myers–Briggs Type Indicator (MBTI):** The MBTI categorizes personality into four dimensions. Extraversion (E) vs. Introversion (I): Extraverts are outgoing and energized by social interactions, while Introverts are reserved and energized by solitude. Sensing (S) vs. Intuition (N): Sensors focus on present, concrete information, valuing practicality, whereas Intuitives are imaginative and future-oriented, valuing abstract ideas. Thinking (T) vs. Feeling (F): Thinkers base decisions on logic and fairness, prioritizing objectivity, while Feelers base decisions on personal values and the impact on others, prioritizing harmony. Judging (J) vs. Perceiving (P): Judgers prefer structured and organized lives, liking plans and decisiveness, while Perceivers prefer flexibility and spontaneity, liking to keep their options open. Each MBTI type is defined by a combination of four cognitive functions, which can be either introverted (i) or extraverted (e). Extraverted Sensing (Se): Focuses on the present moment and physical reality, highly attuned to sensory experiences. Introverted Sensing (Si): Relies on past experiences and memories, valuing tradition and consistency. Extraverted Intuition (Ne): Sees patterns and connections, focusing on future possibilities and abstract ideas. Introverted Intuition (Ni): Focuses on internal insights and foresight, seeing underlying meanings and future potentials. Extraverted Thinking (Te): Organizes and structures the external world, prioritizing logic and efficiency. Introverted Thinking (Ti): Analyzes and categorizes information internally, valuing logical consistency and understanding. Extraverted Feeling (Fe): Prioritizes harmony and social values, focusing on the needs and feelings of others. Introverted Feeling (Fi): Values personal beliefs and feelings, making decisions based on inner values and ethics.
- **Big Five Personality Traits:** The Big Five model describes personality using five broad traits. Openness to Experience: High openness involves imagination and insight, while low openness involves practicality and routine. Conscientiousness: High conscientiousness is characterized by organization and dependability, while low conscientiousness is characterized by spontaneity and flexibility. Extraversion: High extraversion includes sociability and assertiveness, while low extraversion (in-

troversion) includes reserve and solitude. Agreeableness: High agreeableness involves trust and altruism, while low agreeableness involves skepticism and competition. Neuroticism: High neuroticism involves emotional instability and anxiety, while low neuroticism involves emotional stability and calmness.

- **Enneagram:** The Enneagram classifies personality into nine types, each representing different motivations and fears. Type 1: The Reformer, driven by a need for perfection. Type 2: The Helper, driven by a need to be loved. Type 3: The Achiever, driven by a need for success. Type 4: The Individualist, driven by a need for uniqueness. Type 5: The Investigator, driven by a need for knowledge. Type 6: The Loyalist, driven by a need for security. Type 7: The Enthusiast, driven by a need for variety and fun. Type 8: The Challenger, driven by a need for control. Type 9: The Peacemaker, driven by a need for harmony. A 2w3 individual is likely to be more ambitious, charming, and goal-oriented than a typical Type 2. They still seek to help others but are also motivated by a desire for success and recognition.
- **Instinctual Variants:** The Instinctual Variants theory describes three primary instinctual drives influencing behavior. Self-Preservation (SP): Focuses on safety, health, and comfort. Social (SO): Focuses on relationships, status, and community. Sexual (SX): Focuses on intimacy, attraction, and one-on-one connections. For instance, an 8w7 with a Sexual variant, is highly charismatic and seeks intense and passionate connections with others. He or she is bold and assertive, often focusing his or her energy on building strong, impactful relationships.

### Definitions of Relations

Human social networks are complex and multifaceted. By categorizing relations, we can better understand the dynamics and nuances of how people interact with each other. Different types of relations provide context for interactions, which is crucial for analyzing social behaviors and patterns, improving social network analysis, and applying this knowledge across various fields and applications. Table 1 and Table 2 provide a structured approach to understanding the complex web of relations that individuals navigate. By categorizing these relations into social and emotional types, we

Relations type	Description
Family	Parents (grandparents) and children, siblings, etc.
Friendship	Based on common interest, mutual respect and affection, but not related to the blood.
Romantic	Based on emotional attraction and include dating, marriage, etc.
Professional	Formed in a work environment, such as colleagues, superiors and subordinates, etc.
Social	Formed in a broader social context, such as neighbors, club members.
Academic	Formed in an educational setting, such as between teachers and students, classmates.
Online	Established in online spaces or through social media platforms.

Table 1: Descriptions of Social Relations

Relations type	Description
Fondness	A positive emotion characterized by a person’s fondness for another.
Jealousy	Unhappy and angry because someone has something that you want.
Aversion	A negative emotion, referring to a feeling of disfavor towards someone.
Pity	A feeling of sadness for someone else’s difficult situation.
Respect	Admiration felt or shown for someone that you believe has good ideas or qualities.
Hostility	An unfriendly or unkindness towards someone or something.
Envy	A discontented feeling when a person desires what someone else has.
Gratitude	An emotion of being thankful for someone else’s help or kind actions.

Table 2: Description of Emotion Relations

can better analyze and predict personality dynamics in various contexts (Collins and Sroufe 1999; Emmons and McCullough 2004).

### Data Alignment Algorithm

The details of data alignment algorithm are as follows:

#### Algorithm 1: Scripts and Subtitles Matching

**Input:** *Script, Subtitles*

**Output:** Updated subtitles with speaker names

```

1: dial&speakers ← empty
2: threshold ← 0.8
3: for scene in Script do
4:   for Dials in scene do
5:     Extract speaker and dial from Dials
6:     dial&speakers ← speaker, dial
7:   end for
8: end for
9: for subtitle in Subtitles do
10:  match_score ← 0
11:  match_speaker ← Null
12:  for line in subtitle do
13:    for speaker, dial in dial&speakers do
14:      score ← Similar(subtitle, dial)
15:      if score ≥ match_score then
16:        Update match_score and match_speaker
17:      end if
18:    end for
19:    if match_score ≥ threshold then
20:      Update line with match_speaker
21:    end if
22:  end for
23:  Update subtitle
24: end for
25: return Updated Subtitles

```

several scenes according to the coherence in language of camera, instead of randomly clipping in a certain time period. This segmentation is guided by explicit scene transition cues found in movie scripts, such as “*CUT TO:*” or scene location indicators. For TV show scripts, which might lack uniform scene transition markers, we identify scene changes by detecting pauses exceeding 3 seconds between utterances.

2. *Match the utterance* This algorithm is rooted in the comparison of utterances from original scripts and subtitles based on a similarity threshold. If the similarity between a pair of utterances meets or exceeds this threshold, the character’s name is accurately associated with the utterance.
3. *Rematch with the slide window* Basically, the content in scripts is slightly different with the subtitles, because the director may have improvised on the set. Thus, we introduce a slide window algorithm to evaluate the utterance-level similarity. As shown in Algorithm 2, we set a window to slide over the script and, for each utterance, compare the content inside the window with each subtitle entry to get the similarity of the paragraph in the window.

### References

- Collins, W. A.; and Sroufe, L. A. 1999. *Capacity for Intimate Relationships: A Developmental Construction*, 125–147. Cambridge Studies in Social and Emotional Development. Cambridge University Press.
- Emmons, R. A.; and McCullough, M. E. 2004. *The Psychology of Gratitude*. Oxford University Press. ISBN 9780195150100.

1. *Preprocess the raw data* Firstly, we divide the scripts into

---

**Algorithm 2: Slide Window Matching**

---

**Input:** *Script, Subtitles*

**Output:** Updated subtitles

```
1: window_size  $\leftarrow$  10
2: threshold  $\leftarrow$  0.8
3: matches  $\leftarrow$  empty_list
4: for  $i \leftarrow 0$  to  $\text{Len}(\text{Script}) - \text{window\_size}$  do
5:   window  $\leftarrow$  slice(scriptTokens,  $i$ ,  $i + \text{window\_size}$ )
6:   match_score  $\leftarrow$  0
7:   for  $j \leftarrow 0$  to  $\text{Len}(\text{Subtitles}) - 1$  do
8:     score  $\leftarrow$  Similar(window, Subtitles[ $j$ ])
9:     if score  $\geq$  match_score then
10:       Update match_score
11:     end if
12:   end for
13:   if match_score  $\geq$  threshold then
14:     matches  $\leftarrow$  Subtitles[ $j$ ]
15:   end if
16: end for
17: return Updated Subtitles with matches
```

---