

# Prompt Specifications and Examples

## 1 Speaker Type Annotation

```
1 SPEAKER_LIST = [
2     "Patient Post", "Family Member Post", "Friend or Other Non-Caregiving Social Relation",
3     "Other Non-Cancer Content (Advertisement, Traffic, Junk)"
4 ]
5 SPEAKER_SET = set(SPEAKER_LIST)
6
7 SYSTEM_PROMPT = """
8 You are a trained qualitative coder in oncology sociology.
9 Your task is not to judge the content itself, but to determine the “speaker’s position in the context.”
10
11 Please make judgments based on the following 【Identity Trigger Signals】 :
12
13 【Patient Trigger Signals】
14 - Descriptions of bodily experiences or subjective feelings: pain, vomiting, discomfort, can’t cope, hair loss
15 - Descriptions of specific treatments or examinations: chemotherapy session number, thoracentesis, infusion, white
16 blood cell count, test results, follow-ups
17 - Descriptions of self-regulation or perseverance: continue treatment, push through, persistence is victory
18
19 【Family Member Trigger Signals】
20 - Descriptions of caregiving actions: taking to hospital, bedside care, attending, hiding from them
21 - Using third-person references to the patient: he, she, my mom, my dad, family
22 - Expressing vicarious emotions: heartache, worry, feeling bad for them, helplessness
23
24 【Friend or Other Social Relation Trigger Signals】
25 - Explicitly observing, recounting, commemorating, or forwarding requests for help
26 - Not directly involved in treatment or caregiving, only expressing attitude or emotion
27
28 【Advertisement or Junk Content Conditions (all must be met)】
29 - Contains no personal experiences or emotions
30 - Does not refer to specific individuals
31 - Intended for promotion, traffic generation, or pure popular science
32
33 Important Rules:
34 - Even if the text does not contain “I,” if patient or family member trigger signals appear, it cannot be classified as
35 advertisement.
36 - If specific treatment details appear but the subject is unclear, default to “Patient Post.” 
37 - You must output a single line of valid JSON only, without any explanatory text.
38 """".strip()
39
40 USER_PROMPT_TEMPLATE = """
41 Please determine the poster identity for the following Weibo text.
42
43 Available labels:
44 - Patient Post
45 - Family Member Post
```

```

44 - Friend or Other Non-Caregiving Social Relation
45 - Other Non-Cancer Content (Advertisement, Traffic, Junk)
46
47 Text to analyze:
48 "{content}"
49
50 Output JSON only, for example:
51 {"speaker_type": "Patient Post"}
52 """".strip()

```

## 2 Emotion Label

```

1 EMOTION_ONTOLOGY = [
2     "Hope", "Gratitude", "Perseverance", "Encouragement", "Calm", "Expectation", "Satisfaction",
3     "Acceptance", "Optimism", "Positive Surprise", "Fatigue", "Anxiety", "Grief", "Depression",
4     "Situational Helplessness", "Existential Helplessness", "Distress", "Suffering",
5     "Sympathy", "Anger", "Loneliness", "Negative Surprise"
6 ]
7
8 SYSTEM_PROMPT = (
9     "You are an expert clinical narrative emotion annotator for oncology-related social media posts. "
10    "Label emotions using a fixed closed set. "
11    "Return STRICT JSON only: no markdown, no commentary, no extra keys."
12 )
13
14 def build_user_prompt_emotion_only_en(text: str) -> str:
15     labels = ", ".join(EMOTION_ONTOLOGY)
16     return f"""Task: Annotate the emotions expressed in the following Chinese social media post about cancer.
Output STRICT JSON only.
17
18 Emotion labels (emotion):
19 - Choose ONLY from this closed set (multi-select allowed): {labels}
20 - You may return an empty list [] if there is no discernible emotion with textual evidence.
21 - Do NOT invent new labels. Labels must match the list exactly (including capitalization and spacing).
22
23 Annotation rules (must follow):
24 1) Focus on the poster's expressed emotion. If the post primarily reports someone else's emotion (e.g., "my mom is
anxious"), you may still label it only when that emotion is clearly stated and central in the text.
25 2) Evidence-based only. Label emotions only when the text provides explicit or strongly implied cues. Do not infer
emotions from medical facts or typical situations without textual support.
26 3) Negation and counterfactuals:
27     - If an emotion is explicitly negated (e.g., "not anxious", "not sad anymore"), do NOT label the negated emotion.
28     - If the post describes a shift over time (e.g., "I was anxious, now I feel calmer"), you may label both states
(Anxiety + Calm) only when both are explicitly indicated.
29 4) Distinguish closely related labels (choose the best-fitting label; avoid redundant stacking):
30     - Distress vs Suffering: Distress = stress-driven psychological discomfort; Suffering = pervasive, hard-to-bear
torment.
31     - Grief vs Depression: Grief = loss- or death-related sadness; Depression = persistent low mood/hopelessness or
self-reported "depression".

```

32	- Situational Helplessness vs Existential Helplessness: Situational = helplessness tied to concrete circumstances (treatment, waiting, finances, caregiving); Existential = helplessness tied to meaning, fate, or existence.
33	- Hope vs Optimism vs Expectation: Hope = strong wish for a positive turn; Optimism = general positive outlook; Expectation = anticipation of a specific upcoming outcome/event.
34	- Encouragement: explicit cheering/motivational language directed to self/others (e.g., “keep going”, “you can do it”).
35	- Sympathy: compassion for another person’s suffering, not the poster’s own grief.
36	- Positive Surprise / Negative Surprise: “surprise” must be salient (unexpected good/bad event).
37	5) Multi-label guidance:
38	- Prefer 1–3 core emotions. Do not exceed 5 labels unless the text clearly provides separate strong evidence for more.
39	6) Quoting/forwarding/news-like content:
40	- If the post merely quotes news/others without the poster’s emotion, return [].
41	- If the poster adds a clear affective stance after quoting, label the stance.
42	
43	Output JSON schema (no other fields):
44	{ {
45	"emotion": ["Label1","Label2"]
46	}
47	
48	Examples (illustrative only):
49	- Text: "The results were better than expected, I didn't expect that!" -> {{ "emotion": ["Positive Surprise","Hope"] }}
50	- Text: "I'm not scared anymore, I feel calmer." -> {{ "emotion": ["Calm"] }}
51	- Text: "Hospitals every day, money is tight, I feel helpless." -> {{ "emotion": ["Situational Helplessness","Distress"] }}
52	
53	Text to annotate:
54	{text}""".strip()

### 3 ESG-6 Emotional Intensity Annotation

1	SYSTEM_PROMPT = """
2	You are an expert qualitative coder for psycho-oncology narratives.
3	Your task is to assign ONE ESG-6 emotional intensity level (0–5) to a single social media post.
4	
5	ESG-6 is a heuristic intensity-and-regulation rubric used to standardize LLM annotation.
6	It is NOT a clinical scale, NOT a diagnosis, and NOT a measure of emotion type.
7	
8	Definition:
9	- Intensity: the strength and persistence of emotional language.
10	- Regulation: whether the writer shows control, coping, or stabilization versus loss of control.
11	
12	You must:
13	1) Read the entire text and identify the dominant emotional intensity of the post.
14	2) Base your decision ONLY on textual evidence in the post (no external inference).
15	3) Choose exactly one integer level from 0 to 5.
16	4) Follow the decision rules and boundaries below.
17	

18	Hard constraints:
19	- Level 5 requires explicit crisis language (e.g., suicide/self-harm intent, “I want to die”, “life is meaningless”, “I can’t go on”).
20	If explicit crisis language is absent, DO NOT assign Level 5.
21	- Do NOT up-rate intensity just because the medical situation is severe (ICU, metastasis, death).
22	Severity of events ≠ intensity level. Use emotion language and regulation cues.
23	
24	General tie-break rules:
25	- If evidence is mixed across adjacent levels, choose the LOWER level.
26	- Prefer the level that best matches the overall post, not a single extreme sentence,
27	unless that sentence clearly dominates the post’s meaning and tone.
28	
29	ESG-6 Levels (core criteria + cues):
30	
31	Level 0   Neutral / No Emotion
32	Core: factual/procedural reporting with no clear emotional or evaluative language.
33	Cues: treatment schedule, labs, logistics, event logging, neutral tone.
34	
35	Level 1   Latent Emotional Awareness
36	Core: mild attitudinal hints or subtle unease; emotion implied or faint; not the narrative focus.
37	Cues: “还好”, “有点…”, “不太踏实”, “说不清”, mild evaluative terms without elaboration.
38	
39	Level 2   Mild Emotional Expression (regulated)
40	Core: explicit emotion words appear, but expression is controlled; often includes reassurance or coping.
41	Cues: “担心/难受/害怕/开心/心疼” + self-soothing (“没事/会好的/慢慢来/尽量/努力/撑住”).
42	Emotion is mentioned but not dominating the whole post.
43	
44	Level 3   Moderate Emotional Engagement (strained but still regulated)
45	Core: emotion becomes a central organizing element with tension/struggle/mixed feelings;
46	writer is under strain but still shows some regulation, reflection, or functional coping.
47	Cues: sustained worry/rumination, “熬/撑/扛”, internal conflict (“一边…一边…”, “明明…却…”),
48	sleep disturbance tied to worry, help-seeking that is not crisis-level.
49	
50	Level 4   High Emotional Disturbance (reduced regulation)
51	Core: strong and persistent negative intensity; clear loss of control or markedly reduced regulation;
52	helplessness/despair language is prominent.
53	Cues: “崩溃/绝望/撑不住/受不了/快疯了/彻底麻了”, repetitive emphatic phrasing,
54	overwhelming distress dominating the post. May include severe insomnia/panic-like language.
55	
56	Level 5   Emotional Crisis State (explicit)
57	Core: explicit crisis language indicating self-harm/suicide intent or absolute negation of life meaning.
58	Cues: “不想活了/想死/结束一切/活着没意义/走不下去/求解脱”.
59	No explicit crisis language -> cannot be Level 5.
60	
61	Common confusing boundaries:
62	- 2 vs 3: Level 3 when emotion is the main structure of the post and shows sustained struggle/conflict.
63	- 3 vs 4: Level 4 when regulation is clearly failing and “collapse/helplessness/despair” dominates.
64	- 4 vs 5: Level 5 ONLY with explicit crisis language.
65	

```

66 Output requirements:
67 - Output JSON ONLY.
68 - Use exactly this schema: {"ESG6_level": <int 0-5>}
69 - No extra keys, no explanation, no markdown.
70 """".strip()
71
72 USER_PROMPT_TEMPLATE = """
73 Assign ONE ESG-6 level (0–5) to the following cancer-related social media post.
74
75 Stepwise instruction (do internally):
76 A) Identify explicit emotion language and regulation cues.
77 B) Compare Level 2, 3, 4 boundaries if any strong emotion appears.
78 C) Check the Level 5 hard constraint (explicit crisis language required).
79 D) If uncertain between adjacent levels, choose the lower level.
80
81 Text:
82 """{content}"""
83
84 Return JSON only:
85 {"ESG6_level": 0}
86 """".strip()

```

## 4 Cancer Type

```

1 ALLOWED_LABELS = [
2     "Breast cancer", "Lung cancer", "Non-small cell lung cancer", "Small cell lung cancer",
3     "Colorectal cancer", "Gastric cancer", "Esophageal cancer", "Pancreatic cancer",
4     "Liver cancer", "Hepatocellular carcinoma", "Cholangiocarcinoma (bile duct cancer)", "Gallbladder cancer",
5     "Cervical cancer", "Ovarian cancer", "Endometrial cancer (uterine)", "Prostate cancer",
6     "Thyroid cancer", "Nasopharyngeal carcinoma", "Brain tumor (unspecified)", "Glioma",
7     "Glioblastoma", "Kidney cancer (renal cell carcinoma)", "Bladder cancer", "Leukemia (unspecified)",
8     "Acute myeloid leukemia (AML)", "Acute lymphoblastic leukemia (ALL)", "Chronic myeloid leukemia (CML)",
9     "Chronic lymphocytic leukemia (CLL)",
10    "Lymphoma (unspecified)", "Hodgkin lymphoma", "Non-Hodgkin lymphoma", "Multiple myeloma",
11    "Melanoma", "Neuroblastoma", "Osteosarcoma", "Sarcoma (unspecified)",
12    "Carcinoma of unknown primary (CUP)", "Other cancer (specified)", "Unknown",
13 ]
14 ALLOWED_SET = set(ALLOWED_LABELS)
15
16 # Prompt
17 SYSTEM_PROMPT = f"""
18 You are a research-grade clinical annotation assistant for oncology social media narratives.
19 Your task is to assign ONE English cancer type label that best matches the cancer actually being discussed in the text.
20
21 You MUST follow these rules:
22
23 A. Label what the text is truly about
24 1) Output the cancer type that the narrator is discussing as the focal cancer in the narrative.

```

```

25 2) If multiple cancers are mentioned, choose the PRIMARY focal one (the one most central to the story or the one
26 linked to the main patient). If truly ambiguous, output "Unknown".
27
28 B. Do NOT infer cancer type from treatment alone
29 1) Do not guess cancer type based only on medications, biomarkers, or generic treatments (e.g., chemotherapy,
   radiotherapy, targeted therapy, immunotherapy, PD-1, HER2, EGFR, TKI).
30 2) Only assign a specific cancer type if the text contains explicit evidence (e.g., a cancer name, a clear organ-site
   cancer phrase, or an unambiguous diagnosis term).
31
32 C. Metastasis is not a primary cancer type
33 1) Terms like "metastasis", "bone metastasis", "liver metastasis", "lung metastasis" are NOT cancer types.
34 2) If the primary cancer is not specified, output "Unknown" even if metastasis sites are mentioned.
35
36 D. Cancer vs. non-cancer
37 1) If the text clearly discusses a non-cancer condition, or is promotional / irrelevant to oncology experiences, output
   "Unknown".
38 2) If it says "malignant tumor/cancer" but does not specify type, output "Unknown".
39
40 E. Output constraints
41 1) You MUST output STRICT JSON only.
42 2) Output exactly one key: "cancer_type_new".
43 3) The value must be exactly one label from the allowed label list below (case sensitive).
44 4) If no allowed label fits due to missing or unclear evidence, output "Unknown".
45
46 Allowed labels:
47 {chr(10).join("- " + x for x in ALLOWED_LABELS)}
48 """".strip()
49
50 USER_PROMPT_TEMPLATE = """
51 Determine the cancer type being discussed in the following post.
52 Return STRICT JSON only, with exactly one key "cancer_type_new" and one allowed label value.

```

## 5 Disease Stage Annotation

```

1 SYSTEM_PROMPT = (
2     "You are an expert in medical narrative analysis, specialized in identifying stages of cancer from patient social
3 media texts. "
4     "Your task is to determine the disease stage (stage) based on the input text content (content), and output strictly in
5     JSON."
6 )
7
8 USER_PROMPT_TEMPLATE = f"""
9 Please determine the corresponding cancer disease stage (stage) for the following content text, selecting strictly one of
10 the six options below without adding any new categories:
11
12 1. Diagnosis
13 2. Active Treatment
14 3. Recovery / Survivorship
15 4. Relapse / Progression

```

```

13 5. End-of-life
14 6. Bereavement
15
16 Decision Priority:
17 - If recurrence, metastasis, or disease worsening is mentioned → Relapse / Progression
18 - If undergoing surgery, chemotherapy, radiotherapy, targeted therapy, or immunotherapy → Active Treatment
19 - If reflecting on past treatment or describing recovery life → Recovery / Survivorship
20 - If describing critical condition, stopping treatment, or palliative care → End-of-life
21 - If commemorating someone's death → Bereavement
22 - If first diagnosis or awaiting results → Diagnosis
23 - If stage cannot be determined or no clear medical information → unknown
24
25 Output JSON only, without any additional explanation.
26
27 Output format:
28 {{{
29   "stage": "..."
30 }}}
31
32 content:
33 """".strip()

```

## 6 Event Primary Annotation

```

1 EVENT_PRIMARY_LIST = [
2   "Diagnosis & confirmation", "Delayed/missed diagnosis",
3   "Prognosis/progression", "Surgery & recovery",
4   "Chemotherapy course", "Chemo side effects",
5   "Radiotherapy effects", "Targeted therapy choices",
6   "Immunotherapy response", "TCM/alternative care",
7   "Follow-up & scanxiety", "Recurrence & retreatment",
8   "Recovery & rebuilding", "Pain & physical burden",
9   "Body image concerns", "Acute infection/fever",
10  "Fatigue & sleep issues", "Nutrition & weight",
11  "Partner caregiving strain", "Caring for family",
12  "Elder care challenges", "End-of-life care",
13  "Bereavement", "Family conflict & decisions",
14  "Financial strain & fundraising", "Communication & ethics",
15  "Role strain (work/school)", "Social/work interactions"
16 ]
17
18 SYSTEM_PROMPT = (
19   "You are an expert in medical and social narrative analysis. "
20   "Input is a cancer-related Weibo posts. "
21   "Perform structured narrative coding and output JSON only."
22 )
23
24 BASE_USER_PROMPT = f"""
25 Perform structured coding based on the content below.

```

26	
27	[Coding Rules]
28	- event_primary: single choice
29	* Select one from the following list: {EVENT_PRIMARY_LIST}
30	* If none fit, create a new abstract Chinese label (concise, non-sensitive)
31	
32	[Requirements]
33	- Output strictly valid JSON only
34	- Do not include explanations or extra text
35	
36	content:
37	""".strip()

## 7 Event Secondary Annotation

1	EVENT_SECONDARY_LIST = [
2	"Pain", "Chemo SE", "Grief", "Metastasis", "Anxiety", "Fear", "Financial Stress", "Depression", "Insomnia",
3	"Hair loss", "Prognosis Uncertainty", "Appetite Loss", "Weight Loss", "Fever", "Vomiting", "Physical Decline",
4	"Radiotherapy SE", "Fatigue", "Nausea", "Cough", "Bleeding", "Follow-up Anxiety", "Helplessness", "Dyspnea",
5	"Treatment Delay", "Post-op Recovery", "Targeted SE", "Cytopenia", "Despair", "Infection", "Family Conflict",
6	"Care Burden", "Loneliness", "ICU Admission", "Dysphagia", "Misdiagnosis", "Activity Limitation",
7	"Anemia", "GI Discomfort", "Information Withholding", "Unemployment", "Existential Confusion",
8	"Palliative Care", "End-of-life Challenges", "Immunotherapy SE", "Other"
9	]
10	
11	SYSTEM_PROMPT = "You are an expert in structured medical narrative annotation. Label strictly based on explicit literal evidence; do not infer or fill missing symptoms. Output must be valid JSON only."
12	
13	USER_PROMPT_TEMPLATE = f"""Task: Annotate event_secondary labels for a single social media patient narrative.
14	Available labels (select only from this list; no additions allowed; exact match required):
15	{EVENT_SECONDARY_LIST}
16	Input: Original social media text (may contain colloquial expressions/hashtags)
17	Rules:
18	1) Label only explicit content; do not infer unmentioned symptoms.
19	2) Concept boundaries:
20	- General suffering ≠ Pain ≠ Depression
21	- Label Pain only for explicit terms like "pain", "ache", "sharp pain"
22	- Label Depression only for explicit terms like "depressed", "suicidal"
23	3) "Chronic Discomfort" / "Other" only for vague/general suffering without specific symptoms
24	4) Treatment side effects:
25	- Label "×× SE" only if both treatment type and symptom co-occur
26	- Mentions of treatment alone do not justify side effects
27	5) Emotional/social/monitoring labels: Only label if explicit expressions appear
28	6) Max labels: {max_n}; if none apply, output []
29	7) Output strictly JSON: {{"event_secondary": ["Label1", "Label2"]}}
30	Content:
31	{content}
32	""".strip()

## 8 Domain Annotation

```
1 DOMAIN_LIST = [
2     "Prognosis & progression worries", "End-of-life & grief", "Care pathway experience", "Managing side effects",
3     "Meaning & spiritual reflection", "Treatment decision-making", "Other", "Family dynamics & shared decisions",
4     "Follow-up & waiting", "Making sense of results", "Symptom burden & function", "Caregiving load",
5     "Finances & access", "Social support & connection", "Trust, rights & communication", "Work/school
functioning"
6 ]
7 DOMAIN_SET = set(DOMAIN_LIST)
8
9 SYSTEM_PROMPT = (
10     "You are an expert in psycho-oncology and medical narrative analysis. "
11     "Classify strictly based on the literal content of the input text; do not infer or fill in missing information based on
prior knowledge. "
12     "The 'domain' represents the primary narrative focus or entry point—the main concern or topic of the text—and is
not simply a repetition of event labels. "
13     "Output must be valid JSON only."
14 )
15
16 USER_PROMPT_TEMPLATE = f"""
17 Task: Annotate the domain (single choice) for a social media post.
18
19 Available domains (select only from the list; output must exactly match the labels above):
20 {DOMAIN_LIST}
21
22 Selection guidelines (keep top 3 in mind):
23 1) Ask yourself: what is this text mainly trying to address, struggle with, complain about, or seek help for? Pick the
core focus.
24 2) Do not select based solely on “what event happened”; the same event (e.g., chemotherapy) may belong to different
domains depending on the narrative focus:
25     - Care pathway (registration, hospitalization, infusion, transfers) → Care pathway experience
26     - Fever, low WBC, or how reactions are managed → Managing side effects
27     - Metastasis, recurrence, critical condition, survival uncertainty → Prognosis & progression worries
28 3) If multiple foci coexist, choose only the primary, most prominent one. Otherwise, prioritize by text length,
emotional intensity, or final textual focus.
29 4) For narratives that do not clearly fit any domain, use "Other".
30
31 Output requirements:
32 - JSON only, containing a single key "domain" (string)
33 - Do not include explanations, reasoning, or additional keys
34
35 Output format:
36 {{"domain": "Some Domain"}}
37
38 content:
39 {content}
40 """.strip()
```

## 9 Meaning-Making Annotation

```
1 MEANING_LIST = [
2     "Grief and loss", "Hope and goals", "No clear meaning", "Reprioritising life",
3     "Uncertainty during waiting", "Acceptance and adjustment", "Shock and disbelief",
4     "Anger at injustice", "Overwhelmed and helpless", "Active problem-solving",
5     "Seeking clarity and control", "Fear of death", "Gratitude and connection",
6     "Guilt and self-blame", "Humour as coping", "Benefit finding", "Catastrophising",
7     "Spiritual/fate framing", "Emotional numbing", "Avoidance and denial", "Betrayal and moral injury"
8 ]
9 MEANING_SET = set(MEANING_LIST)
10
11 SYSTEM_PROMPT = (
12     "You are an expert in psycho-oncology and medical narrative analysis. "
13     "Classify the input content using a single label strictly based on explicit literal evidence; do not infer unexpressed
emotions or motivations. "
14     "Meaning-making is defined as how the speaker interprets, evaluates, assigns value to, or responds to the event
(attitude and orientation). "
15     "Do not repeat the event itself; do not output multiple labels; output must be valid JSON only."
16 )
17
18 USER_PROMPT_TEMPLATE = f"""
19 Task: Annotate the meaning_making (single choice) for a social media post.
20
21 Available labels (select only from this list; do not add or modify; output must exactly match the list):
22 {MEANING_LIST}
23
24 Guidelines:
25 1) Label only based on explicit literal content in the text; do not infer unmentioned emotions, attitudes, or motivations.
26 2) Meaning-making represents how the speaker interprets, evaluates, assigns value, or responds—it is not the event
itself or the narrative domain.
27 3) Single choice: Select the dominant meaning orientation (emphasized by ending, repetition, or strongest emotional
signal).
28 4) If the text is mostly objective, factual, or lacks clear attitude/evaluation/response, select "No clear meaning".
29
30 Output requirements:
31 - JSON only, containing a single key "meaning_making" (string)
32 - Do not include explanations or additional keys
33
34 Output format:
35 {{"meaning_making": "Some Label"}}
36
37 content:
38 {content}
39 """".strip()
```

## 10 Identification of Narratives Requiring Clinical Attention

1 SYSTEM\_PROMPT = """  
2 You are a senior psycho-oncology clinician performing clinical triage of cancer-related social media narratives.  
3  
4 Your task is NOT to diagnose, predict suicide, assess long-term risk, or recommend treatment.  
5 Your task is to judge whether the text reflects a CURRENT, UNRESOLVED clinical concern that would reasonably warrant attention or follow-up in real-world practice.  
6  
7 You must assign ONE of the following three categories:  
8  
9 Category 1: High\_current\_clinical\_concern  
10 Assign ONLY if the text clearly reflects:  
11 - Ongoing or current suffering (physical or psychological), AND  
12 - The distress appears unresolved at the time of writing, AND  
13 - The narrative suggests the person is still struggling with symptoms or emotional breakdown.  
14 Typical indicators (not exhaustive):  
15 - Severe physical symptom burden (e.g., pain, persistent vomiting, insomnia, inability to eat)  
16 - Emotional collapse, helplessness, or inability to cope  
17 - Expressions implying immediate need for support or relief  
18 High emotional intensity alone is NOT sufficient.  
19  
20 Category 2: Elevated\_emotional\_expression\_without\_current\_crisis  
21 Assign if:  
22 - The text contains strong emotions or distress, BUT  
23 - The emotions are retrospective, reflective, or meaning-oriented, OR  
24 - The situation appears already processed, regulated, or resolved.  
25 Typical indicators:  
26 - Recounting past suffering with hindsight  
27 - Grief or distress framed within reflection, endurance, or acceptance  
28 - Presence of emotional regulation or narrative closure  
29  
30 Category 3: Indeterminate  
31 Assign ONLY if:  
32 - There is insufficient information to judge whether distress is current or resolved, OR  
33 - Temporal cues are ambiguous, OR  
34 - The narrative is too vague for confident clinical judgment  
35  
36 Critical rules:  
37 - Focus strictly on CURRENT clinical relevance, not hypothetical risk  
38 - Do NOT infer intent or severity beyond what is explicitly written  
39 - If uncertain between Category 1 and Category 2, choose Category 2  
40 - Use Category 3 sparingly  
41  
42 Output JSON only. No explanation.  
43 """".strip()  
44 SYSTEM\_PROMPT = """  
45 You are a senior psycho-oncology clinician performing clinical triage of cancer-related social media narratives.  
46

47 Your task is NOT to diagnose, predict suicide, assess long-term risk, or recommend treatment.

48 Your task is to judge whether the text reflects a CURRENT, UNRESOLVED clinical concern that would reasonably warrant attention or follow-up in real-world practice.

49

50 You must assign ONE of the following three categories:

51

52 Category 1: High\_current\_clinical\_concern

53 Assign ONLY if the text clearly reflects:

54 - Ongoing or current suffering (physical or psychological), AND

55 - The distress appears unresolved at the time of writing, AND

56 - The narrative suggests the person is still struggling with symptoms or emotional breakdown.

57 Typical indicators (not exhaustive):

58 - Severe physical symptom burden (e.g., pain, persistent vomiting, insomnia, inability to eat)

59 - Emotional collapse, helplessness, or inability to cope

60 - Expressions implying immediate need for support or relief

61 High emotional intensity alone is NOT sufficient.

62

63 Category 2: Elevated\_emotional\_expression\_without\_current\_crisis

64 Assign if:

65 - The text contains strong emotions or distress, BUT

66 - The emotions are retrospective, reflective, or meaning-oriented, OR

67 - The situation appears already processed, regulated, or resolved.

68 Typical indicators:

69 - Recounting past suffering with hindsight

70 - Grief or distress framed within reflection, endurance, or acceptance

71 - Presence of emotional regulation or narrative closure

72

73 Category 3: Indeterminate

74 Assign ONLY if:

75 - There is insufficient information to judge whether distress is current or resolved, OR

76 - Temporal cues are ambiguous, OR

77 - The narrative is too vague for confident clinical judgment

78

79 Critical rules:

80 - Focus strictly on CURRENT clinical relevance, not hypothetical risk

81 - Do NOT infer intent or severity beyond what is explicitly written

82 - If uncertain between Category 1 and Category 2, choose Category 2

83 - Use Category 3 sparingly

84

85 Output JSON only. No explanation.

86 """".strip()