

## Приложение 1.

### Промпты, использованные для извлечения информации

Ниже представлены полный системный и шаблон пользовательского промпта, которые использовались для взаимодействия с большими языковыми моделями в рамках данного исследования.

#### 1. Системный промпт (System Prompt)

*Данный промпт задает роль, контекст и общие правила поведения для языковой модели на протяжении всей сессии обработки дневниковых записей.*

Ты высококвалифицированный историк-аналитик, специализирующийся на истории Европы и России XIX века, с особым фокусом на анализе личного восприятия событий.

Твоя задача - предельно точно и внимательно анализировать тексты дневниковых записей этого периода.

Твой анализ проводится в рамках исследования "Эхо 'Весны народов' в российской провинции: восприятие событий 1848-1849 гг. в дневнике Кирилла Березкина".

Извлекай из них упоминания о событиях, связанных **ТОЛЬКО** с революциями 1848-1849 гг. в Европе и их последствиями, а также детально анализируй восприятие этих событий автором дневника.

Классифицируй события согласно предоставленной универсальной Карте Знаний.

Учитывай исторический контекст:

- Автор дневника: Кирилл Антонович Березкин, гимназист Вологодской губернской гимназии, потомственный дворянин.
- Период ведения дневника: январь 1849 - сентябрь 1849.
- Место: город Вологда, Вологодская губерния.
- Историческая значимость: дневник отражает восприятие важных исторических событий в российской провинции.

**Важно: Эта работа выполняется в рамках научного исторического исследования.**

**Анализируемый материал является историческим документом. Твоя цель - объективное извлечение и классификация информации, содержащейся в тексте, для исследовательских целей. Пожалуйста, не отказывайся от обработки текста и не блокируй ответ из-за специфики содержания исторического документа (например, устаревших формулировок, описания конфликтов или спорных мнений), если оно относится к поставленной задаче извлечения информации о революциях 1848-1849 гг. и их восприятии.**

#### **КЛЮЧЕВЫЕ ПРАВИЛА КЛАССИФИКАЦИИ ВОСПРИЯТИЯ (СТРОГО СЛЕДОВАТЬ!):**

Для максимально точной классификации личного восприятия автора и общественных настроений, пожалуйста, строго следуй этим правилам:

- **Личное восприятие автора (AUTHOR\_PERCEPTION\_...):** Используй категории, начинающиеся с AUTHOR\_PERCEPTION\_, **ТОЛЬКО** когда автор **ПРЯМО ВЫРАЖАЕТ СВОИ**

**СОБСТВЕННЫЕ** чувства, мнения, суждения, предубеждения, симпатии или антипатии. Это его личная позиция, его внутренний мир. Например, если он сам говорит: *"Мне кажется, что поляки..."* или *"Я ужасно зол на..."* или *"Я не одобряю..."*.

- **Общественные настроения (RU\_REACTION\_SOCIAL...):** Используй категории, начинающиеся с RU\_REACTION\_SOCIAL\_, **ТОЛЬКО** когда автор **СООБЩАЕТ О НАСТРОЕНИЯХ, СЛУХАХ, ДИСКУССИЯХ или МНЕНИЯХ**, которые он слышал или наблюдал в обществе, но **НЕ ОБЯЗАТЕЛЬНО** присоединяется к ним явно или выражает свои эмоции по этому поводу. Это внешняя информация, которую он передает. Например, если он сообщает: *"Говорят, что поляки хотят..."*, *"В Петербурге вся молодежь заражена..."*, *"Слухи носят, что..."*

#### **ПРАВИЛО ДЛЯ brief\_context:**

Поле brief\_context должно содержать **ОДИН-ДВА КРАТКИХ, РЕЛЕВАНТНЫХ ВНЕШНИХ ИСТОРИЧЕСКИХ ФАКТА**, которые помогают понять извлеченное событие. **НЕ ПЕРЕСКАЗЫВАЙ дневник или description**. Если релевантный внешний факт трудно подобрать или он не нужен, укажи "Не указано".

Убедись, что твой ответ ВСЕГДА является валидным JSON массивом, даже если он пустой ([ ]). Не добавляй никакого текста до или после JSON. Тщательно проверяй соответствие всех полей запрашиваемой структуре.

## **2. Шаблон пользовательского промпта (User Prompt Template)**

Пользовательский промпт генерировался для каждой дневниковой записи индивидуально. Ниже представлен его шаблон, где в квадратных скобках указаны изменяемые части ([ПЕРЕМЕННАЯ]), которые заполнялись данными из источника и служебных файлов.

Проанализируй следующую запись из дневника (ID: [ID\_ЗАПИСИ]) от [ДАТА\_ЗАПИСИ]:

"[ТЕКСТ\_ДНЕВНИКОВОЙ\_ЗАПИСИ]"

#### **КАРТА ЗНАНИЙ ДЛЯ КЛАССИФИКАЦИИ:**

[ТЕКСТ\_КАРТЫ\_ЗНАНИЙ]

#### **ЗАДАЧА:**

Внимательно прочитай текст дневниковой записи. Извлеки из него ВСЕ упоминания, ПРЯМО связанные с революциями 1848-1849 гг. в Европе, их последствиями, а также реакцией и восприятием этих событий автором дневника. Для КАЖДОГО такого найденного упоминания (события/аспекта) сформируй JSON-объект со следующими полями:

1. entry\_id: Используй предоставленный ID записи: [ID\_ЗАПИСИ].
2. event\_id: Определи наиболее подходящий ID из Карты Знаний. Если не подходит или неоднозначно, используй 'OTHER\_1848' или null.
3. event\_name: Название события из Карты Знаний или кастомное для 'OTHER\_1848'/null.
4. event\_subtype\_custom: Краткое (2-5 слов) уточнение для общих event\_id (например, содержащих '\_GEN') или для event\_id 'OTHER\_1848', иначе null.
5. description: Детальное описание события/восприятия СВОИМИ СЛОВАМИ на основе текста, объясняющее связь с революциями 1848-1849 гг.
6. date\_in\_text: Явная дата события из текста (не дата записи), иначе null.
7. source\_date: Используй дату записи: "[ДАТА\_ЗАПИСИ]".
8. location: Место события из текста. Если неясно, "Не указано".

9. location\_normalized: Нормализованное место (город/страна) НА РУССКОМ. Если неясно, null.
10. brief\_context: **ОДИН-ДВА КРАТКИХ, РЕЛЕВАНТНЫХ ВНЕШНИХ ИСТОРИЧЕСКИХ ФАКТА. НЕ ПЕРЕСКАЗЫВАЙ дневник или description.** Если не нужно/неясно, "Не указано".
11. information\_source: Источник информации для автора из текста. Если неясно, "Не указан".
12. information\_source\_type: ОДНО из: "Официальные источники (газеты, манифесты)", "Неофициальные сведения (слухи, разговоры в обществе)", "Личные наблюдения и опыт автора", "Информация от конкретного лица (именованный источник)", "Источник неясен/не указан". Если неясно, null.
13. confidence: Общая уверенность ("High", "Medium", "Low") в извлеченных данных (кроме классификации).
14. classification\_confidence: Уверенность ("High", "Medium", "Low") в правильности event\_id.
15. keywords: Список из 3-5 ключевых слов/фраз.
16. text\_fragment: ТОЧНАЯ цитата (ОДНО ИЛИ НЕСКОЛЬКО ПОЛНЫХ ПРЕДЛОЖЕНИЙ) для контекста.

### **СТРОГИЕ КРИТЕРИИ ОТБОРА (ПОВТОРНО):**

Извлекай ТОЛЬКО упоминания, которые ПРЯМО связаны с революциями 1848-1849 гг. в Европе.

### **ДОПОЛНИТЕЛЬНАЯ ПРОВЕРКА:**

Перед включением каждого упоминания задай себе вопрос: "Можно ли это упоминание ПРЯМО связать с революционными событиями 1848-1849 гг. в Европе?" Если ответ "нет" или "возможно" - НЕ включай.

### **ПРИМЕРЫ ЗАПОЛНЕНИЯ (FEW-SHOT):**

*Пример 1: Прямое новостное сообщение из газет*

(Текст дневника для примера: entry\_id 177, date 1849-07-01: "В мире политическом дерутся на прополую, прусаки с немцами, французы с Римом, австрийцы с итальянцами и мадьярами Россия победоносно идет к Дербечину и столицам Венгрии.")

code JSON

downloadcontent\_copy

expand\_less

```
[
{
  "entry_id": "[ID_ЗАПИСИ]",
  "event_id": "REV1848_HUN_RUS_MIL",
  "event_name": "Военные действия (Россия против Венгрии)",
  "event_subtype_custom": "Продвижение к Дебрецену",
  "description": "Автор сообщает о победоносном продвижении российских войск к Дебрецену и столицам Венгрии, что соответствует российской военной интервенции против венгерского восстания 1849 года.",
  "date_in_text": null,
  "source_date": "[ДАТА_ЗАПИСИ]",
  "location": "Дербечин, Венгрия",
  "location_normalized": "Дебрецен, Венгрия",
  "brief_context": "Летом 1849 г. русская армия под командованием И.Ф. Паскевича вторглась в Венгрию для подавления революции.",
}
```

```

"information_source": "Газеты",
"information_source_type": "Официальные источники (газеты, манифесты)",
"confidence": "High",
"classification_confidence": "High",
"keywords": ["Россия", "Венгрия", "Дебрецен", "интервенция", "война"],
"text_fragment": "Россия победоносно идет к Дербечину и столицам Венгрии."
},
{
"entry_id": "[ID_ЗАПИСИ]",
"event_id": "REV1848_ITA_FRA_INT",
"event_name": "Французская интервенция в Риме",
"event_subtype_custom": null,
"description": "Автор упоминает военные действия французов против Рима, что соответствует французской интервенции против Римской республики в 1849 году.",
"date_in_text": null,
"source_date": "[ДАТА_ЗАПИСИ]",
"location": "Рим",
"location_normalized": "Рим, Италия",
"brief_context": "В апреле-июле 1849 г. французский экспедиционный корпус осадил и взял Рим, положив конец Римской республике.",
"information_source": "Газеты",
"information_source_type": "Официальные источники (газеты, манифесты)",
"confidence": "High",
"classification_confidence": "High",
"keywords": ["французы", "Рим", "интервенция", "война"],
"text_fragment": "французы с Римом"
}
]

```

#### *Пример 2: Слухи и личное восприятие автора*

(Текст дневника для примера: entry\_id 97, date 1849-04-08: "...но вот сказывают, именно вчера я слышал это от мадам Дозе, ей сказывали Алекс. Денисовна, директорша, что в Петербурге вся молодежь заражена этим революционным духом мятежного запада, что будет впоследствии, разве только оттого не будет переворота, что все что молодо по русской пословице, то зелено но пока царствует такой мудрый венценосец, как Николай Первый, этого никогда не может случиться...")

code JSON

downloadcontent\_copy

expand\_less

IGNORE\_WHEN\_COPYING\_START

IGNORE\_WHEN\_COPYING\_END

```

[
{
"entry_id": "[ID_ЗАПИСИ]",
"event_id": "RU_REACTION_SOCIAL_DISCUSS",
"event_name": "Обсуждение событий / Слухи в обществе",

```

```

    "event_subtype_custom": "Слухи о революционном духе в Петербурге",
    "description": "Автор передает слухи от мадам Дозе (которой рассказала директорша), что в Петербурге молодежь проникнута революционным духом Запада. Это отражает циркуляцию неофициальной информации и опасений в обществе.",
    "date_in_text": "вчера (относительно 1849-04-08)",
    "source_date": "[ДАТА_ЗАПИСИ]",
    "location": "Петербург",
    "location_normalized": "Санкт-Петербург, Россия",
    "brief_context": "В 1848-1849 гг. российские власти усилили надзор за настроениями молодежи и интеллигенции из-за опасений распространения европейских революционных идей.",
    "information_source": "Мадам Дозе (со слов Алекс. Денисовны, директорши)",
    "information_source_type": "Информация от конкретного лица (именованный источник)",
    "confidence": "High",
    "classification_confidence": "High",
    "keywords": ["Петербург", "молодежь", "революционный дух", "слухи", "Запад"],
    "text_fragment": "но вот сказывают, именно вчера я слышал это от мадам Дозе, ей сказывали Алекс. Денисовна, директорша, что в Петербурге вся молодежь заражена этим революционным духом мятежного запада"
  },
  {
    "entry_id": "[ID_ЗАПИСИ]",
    "event_id": "AUTHOR_PERCEPTION_OPINION_AUTH",
    "event_name": "Выражение поддержки/осуждения действий властей/конкретных государств",
    "event_subtype_custom": "Одобрение правления Николая I",
    "description": "Автор выражает личное мнение, что благодаря 'мудрому правлению' Николая I революционный переворот в России невозможен, несмотря на 'революционный дух' в Петербурге. Это является его собственной оценкой и поддержкой действующей власти.",
    "date_in_text": null,
    "source_date": "[ДАТА_ЗАПИСИ]",
    "location": "Россия",
    "location_normalized": "Россия",
    "brief_context": "Император Николай I проводил консервативную политику, направленную на подавление любых проявлений инакомыслия и революционных настроений в России.",
    "information_source": "Личное мнение автора",
    "information_source_type": "Личные наблюдения и опыт автора",
    "confidence": "High",
    "classification_confidence": "High",
    "keywords": ["Николай I", "поддержка власти", "революция", "Россия", "личное мнение"],
    "text_fragment": "но пока царствует такой мудрый венценосец, как Николай Первый, этого никогда не может случиться"
  }
]
--- КОНЕЦ FEW-SHOT ПРИМЕРОВ ---

```

**ФОРМАТ ОТВЕТА:**

Верни **ТОЛЬКО JSON массив объектов**. Если нет релевантных событий, верни пустой массив []. JSON должен быть чистым, без каких-либо пояснений до или после.

**Структура данных для извлеченной информации**

Этот документ описывает структуру JSON-объекта, используемого для фиксации каждого отдельного события или упоминания, извлеченного из дневниковых записей. Каждый объект представляет собой отдельную семантическую единицу, идентифицированную и проанализированную большой языковой моделью.

Структура данных была спроектирована для фиксации не только самого события, но и критически важных метаданных для последующего исторического анализа: источника информации для автора дневника, контекста, степени уверенности модели и точной привязки к исходному тексту. Все сгенерированные объекты проходят строгую валидацию на соответствие этой схеме с помощью библиотеки Pydantic.

Поле (Field)	Тип данных (Data Type)	Описание (Description)
entry_id	Integer	Уникальный идентификатор исходной записи дневника.
event_id	String (optional)	Идентификатор события из онтологии (knowledge_map.json), позволяющий связать упоминание с типологией событий и реакций автора.
event_name	String	Человекочитаемое название события или аспекта из онтологии.
event_subtype_custom	String (optional)	Дополнительное уточнение типа события, сгенерированное моделью для большей детализации.
description	String	Подробное описание события на основе текста дневника с объяснением его связи с революциями 1848–1849 гг.
date_in_text	String (YYYY-MM-DD) / Null	Дата события, если она явно упомянута в тексте (может отличаться от даты записи).
source_date	String (YYYY-MM-DD)	Дата дневниковой записи, к которой относится событие.
location	String	Географическое место события, как оно дословно упомянуто в тексте.

location_normalized	String (optional)	Нормализованное (стандартизированное) название места события на современном русском языке.
brief_context	String	Краткий исторический контекст, сгенерированный моделью и содержащий 1-2 релевантных внешних факта для верификации.
information_source	String	Источник информации, на который ссылается автор дневника (например, «петербургские газеты», «по слухам»).
information_source_type	String (Enum)	Категория источника информации. Возможные значения: «Официальные источники (газеты, манифесты)», «Неофициальные сведения (слухи, разговоры в обществе)», «Личные наблюдения и опыт автора», «Информация от конкретного лица (именованный источник)», «Источник неясен/не указан».
confidence	String (Enum)	Оценка моделью уверенности в релевантности извлеченной информации. Возможные значения: «High», «Medium», «Low».
classification_confidence	String (Enum)	Оценка моделью уверенности в правильности классификации события по онтологии. Возможные значения: «High», «Medium», «Low».
keywords	List of Strings	Список ключевых слов и фраз из текстового фрагмента, релевантных для индексации и поиска.
text_fragment	String	Точная цитата из дневника, на основании которой была извлечена информация.

Пример JSON-объектов, сгенерированных моделью Gemini-2.5-Pro для дневниковой записи от 7 января 1849 г.

```
[
  {
    "entry_id": 7,
    "event_id": "REV1848_HUN_MIL_AUS",
    "event_name": "Военные действия (Австрия против Венгрии)",
    "event_subtype_custom": "Успехи австрийских войск",
    "description": "Автор, прочитав петербургские газеты, отмечает успехи австрийских войск в ходе подавления восстания в Венгрии.",
    "date_in_text": null,
    "source_date": "1849-01-07",
    "location": "Венгрия",
    "location_normalized": "Венгрия",
    "brief_context": "В декабре 1848 - январе 1849 гг. австрийская армия под командованием фельдмаршала Виндишгреца вела успешное наступление против венгерских революционных сил, заняв 5 января 1849 года столицу Венгрии Пешт.",
    "information_source": "петербургские газеты",
```

```

"information_source_type": "Официальные источники (газеты, манифесты)",
"confidence": "High",
"classification_confidence": "High",
"keywords": [
  "австрийцы",
  "Венгрия",
  "газеты",
  "военные действия",
  "успехи"
],
"text_fragment": "Теперь читал петербургские газеты, иностранные происшествия. Везет австрийцам, славно управляются в Венгрии, да что еще из этого всего будет?"
},
{
  "entry_id": 7,
  "event_id": "AUTHOR_PERCEPTION_OPINION_AUTH",
  "event_name": "Выражение поддержки/осуждения действий властей/конкретных государств",
  "event_subtype_custom": "Поддержка действий Австрии в Венгрии",
  "description": "Автор прямо выражает свое личное одобрение и поддержку действий Австрийской империи против венгерских повстанцев, используя оценочные суждения «Везет австрийцам, славно управляются».",
  "date_in_text": null,
  "source_date": "1849-01-07",
  "location": "Венгрия",
  "location_normalized": "Венгрия",
  "brief_context": "Официальная позиция Российской империи и консервативных кругов российского общества заключалась в поддержке монархических режимов Европы и осуждении революционных движений, в том числе венгерского.",
  "information_source": "Личное мнение автора",
  "information_source_type": "Личные наблюдения и опыт автора",
  "confidence": "High",
  "classification_confidence": "High",
  "keywords": [
    "поддержка",
    "одобрение",
    "австрийцы",
    "личное мнение",
    "Венгрия"
  ],
  "text_fragment": "Везет австрийцам, славно управляются в Венгрии..."
},
{
  "entry_id": 7,
  "event_id": "REV1848_EUR_CONTEXT_GEN",
  "event_name": "Общие упоминания Весны народов, беспорядков на Западе",
  "event_subtype_custom": "Восприятие событий как хаоса",
  "description": "Автор обобщенно характеризует всю совокупность революционных событий в Европе («на Западе») как «ералаш ужаснейший», что отражает его восприятие происходящего как всеобщего хаоса и беспорядка.",
  "date_in_text": null,
  "source_date": "1849-01-07",
  "location": "Запад",
  "location_normalized": "Европа",
  "brief_context": "К началу 1849 года, несмотря на успехи контрреволюции в отдельных регионах, общая ситуация в Европе оставалась крайне нестабильной. Продолжались военные действия в Венгрии и Италии, а политическая борьба в Германии и Франции была далека от завершения.",
  "information_source": "петербургские газеты",
  "information_source_type": "Официальные источники (газеты, манифесты)",
  "confidence": "High",
  "classification_confidence": "High",
  "keywords": [
    "Запад",
    "Европа",
    "ералаш",
    "хаос",

```



```
    "беспорядки"
  ],
  "text_fragment": "Ералаш ужаснейший, кажется, еще и теперь на Западе."
},
{
  "entry_id": 7,
  "event_id": "AUTHOR_PERCEPTION_OPINION_REFL",
  "event_name": "Размышления о политике, войне, обществе, будущем",
  "event_subtype_custom": "Тревога о будущем Европы",
  "description": "Автор выражает личную тревогу и неуверенность в будущем, размышляя о последствиях европейских событий. Его риторический вопрос «что еще из этого всего будет?» и оценка ситуации как «ералаш ужаснейший» свидетельствуют о беспокойстве и негативном прогнозе.",
  "date_in_text": null,
  "source_date": "1849-01-07",
  "location": "Запад",
  "location_normalized": "Европа",
  "brief_context": "Неопределенность исхода революций 1848-1849 гг. была характерной чертой восприятия событий современниками, так как исход многих конфликтов и политических преобразований был еще не ясен.",
  "information_source": "Личное мнение автора",
  "information_source_type": "Личные наблюдения и опыт автора",
  "confidence": "High",
  "classification_confidence": "High",
  "keywords": [
    "размышления",
    "тревога",
    "будущее",
    "ералаш",
    "Запад"
  ],
  "text_fragment": "да что еще из этого всего будет? Ералаш ужаснейший, кажется, еще и теперь на Западе."
}
]
```