**Приложение 1. 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 1. JSON-объекты, сгенерированный моделью Gemini-2.5-Pro для 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."  }  ] |