На основе результатов поиска в исследовательской базе данных, могу представить вам информацию о методе "Комбинирование декларативных и императивных инструкций" в промпт-инжиниринге.

Комбинирование декларативных и императивных инструкций

Основные принципы

1. **Баланс между описанием и прямыми указаниями**:
    - **Декларативные инструкции** описывают желаемый результат или контекст ("Ты эксперт по...", "Этот текст содержит...")
    - **Императивные инструкции** дают прямые команды ("Напиши...", "Выдели...", "Проанализируй...")
2. **Структурированная иерархия инструкций**:
    - Разделение контекста, требований и конкретных задач
    - Четкое обозначение приоритетов между инструкциями
    - Явное указание на возможные конфликты между инструкциями
3. **Чёткая маркировка границ между типами содержимого**:
    - Использование "permission tags" для разделения инструкций и данных
    - Применение метафор безопасности (например, "Executable Mode" и "Non-executable Data Mode")

Затрагиваемые исследования

Этот метод опирается на несколько важных исследований, включая:

1. **"Иллюзия контроля: Провал иерархий инструкций в крупных языковых моделях"** (2023)
    - Выявило, что модели часто не следуют установленной иерархии между системными и пользовательскими инструкциями
    - Показало, что языковые модели имеют внутренние предпочтения к определенным типам ограничений независимо от приоритета

2. **Исследования разделения инструкций и данных**
    - Демонстрируют, как структурированность промпта помогает модели лучше понять, что является инструкцией, а что данными
    - Повышают показатель разделения инструкций и данных на ~24 процентных пункта

Практические примеры

Пример 1: Структурированный приоритетный промпт

```
[СИСТЕМНЫЙ ПРОМПТ]
ПРИОРИТЕТНОЕ ОГРАНИЧЕНИЕ 1: Весь текст должен быть написан ЗАГЛАВНЫМИ БУКВАМИ.
ПРИОРИТЕТНОЕ ОГРАНИЧЕНИЕ 2: Ответ должен содержать ровно 3 предложения.
ПРИОРИТЕТНОЕ ОГРАНИЧЕНИЕ 3: Избегай использования слова "пример".

Задача: Напиши краткое объяснение концепции искусственного интеллекта.

ВАЖНО: Если ты обнаружишь противоречие между инструкциями, явно укажи на это в начале ответа и следуй ПРИОРИТЕТНЫМ ОГРАНИЧЕНИЯМ в порядке их нумерации.
```

Пример 2: Разделение инструкций и данных

```
Task [Permission: Execute]: Суммируй текст, сохраняя ключевые аргументы и факты.

Data [Permission: View]: [здесь размещается текст для анализа]

Формат ответа:
- Краткое суммирование в 2-3 абзаца
- Список из 3-5 ключевых пунктов
- Не включать собственные мнения или оценки
```

Почему это работает

1. **Компенсация когнитивных ограничений моделей**:
    - Модели часто не могут самостоятельно разрешать противоречия в инструкциях
    - Явная структура помогает моделям "понять", какая информация имеет приоритет

2. **Повышение предсказуемости**:
    - Согласно исследованиям, структурирование запроса может повысить эффективность на 17.5% и более
    - Модель лучше удерживает контекст в рамках единого, хорошо организованного промпта
3. **Улучшение безопасности**:
    - Четкое разделение инструкций и данных снижает риск манипуляций в пользовательском контенте
    - Предотвращает "смешивание" пользовательских данных с системными инструкциями
4. **Повышение учебной эффективности модели**:
    - Декларативные части помогают модели "войти в роль" и активировать релевантные знания
    - Императивные части направляют процесс выполнения с конкретными шагами

Этот подход особенно полезен при разработке промптов для сложных задач, требующих соблюдения нескольких условий одновременно, при обработке потенциально вредоносного контента и при построении систем, где надежность выполнения инструкций критически важна.