

Technical Appendix / Техническое приложение

Conte Synthetic Research Desk Demo - внутренний пакет доверия

Сформировано: 2026-06-29 21:21:41 +03. Цель: зафиксировать schema, prompts, scoring formula, model settings, validation checks, версии панели, дату запуска, пробелы по token cost и логику request count.

Исходные артефакты

Роль	Путь	Есть	Изменён	SHA short
JSON schema	schemas/respondent.schema.json	да	2026-06-29 18:27:11 +03	126f27631b93d35f
Prompt генерации респондента	prompts/03_respondent_generator.md	да	2026-06-29 18:27:31 +03	d1c076ec4334114f
Panel config	data/respondents/panel_config.yaml	да	2026-06-29 16:48:37 +03	a4bda112e5219139
Валидированная панель	data/respondents/respondents_50.validated.json	да	2026-06-29 18:38:42 +03	8c24dfc53a7be83a
Study YAML	data/studies/study_003_conte_underwear_campaign_demo.yaml	да	2026-06-29 20:00:10 +03	a29c2c75e75fe4dd
Interview traces	data/outputs/conte_demo/interview_traces.jsonl	да	2026-06-29 20:06:50 +03	71b8cd2f27bacb5f
Score matrix	data/outputs/conte_demo/respondent_score_matrix.csv	да	2026-06-29 20:06:50 +03	973fa19c451c37aa
Message ranking	data/outputs/conte_demo/message_ranking.csv	да	2026-06-29 20:06:50 +03	4d10297f70ed6d7f
Segment heatmap	data/outputs/conte_demo/segment_message_heatmap.csv	да	2026-06-29 20:06:50 +03	2aecc5a27d643a48
Objection map	data/outputs/conte_demo/objection_map.csv	да	2026-06-29 20:06:50 +03	00f462ed4420b818
Creative matrix	data/outputs/conte_demo/creative_matrix.csv	да	2026-06-29 20:06:50 +03	243ac202082f5bcb
Segment recommendations	data/outputs/conte_demo/segment_recommendations.csv	да	2026-06-29 20:06:50 +03	8a80621c790e09b3

JSON schema

Schema: schemas/respondent.schema.json. SHA-256: 126f27631b93d35f9393b1d6f2f710c85705181a10c0fb32278cbbc88c5c174d.

Обязательные поля верхнего уровня: id, segment, demographics, location, shopping_behavior, category_behavior, brand_perception, decision_logic, media_behavior, research_behavior.

Строгость: additionalProperties=false в объектах респондентов; рейтинги ограничены; enum ограничивают segment, gender, city type, category behavior, channels, purchase drivers, deal breakers и research behavior.

Известный разрыв: prompt требует proof needed, но respondent.schema.json не хранит respondent-level proof_needed. Study trace хранит proof_needed на уровне respondent-message answer.

Prompts

Prompt генерации респондента: prompts/03_respondent_generator.md. SHA-256: d1c076ec4334114f9cad16f9be6ca85776fdd483ee1bb3e3faec5a02bfb02517.

Prompt требует реалистичных русскоязычных потребителей из Беларуси и явно сохраняет skeptical, price-sensitive, competitor-loyal и non-Conte-friendly respondents.

Фрагмент prompt

03 Respondent Generator

You generate synthetic respondents for Conte research segment by segment.

Use:

- `data/brand\_context/conte\_brand\_profile.md` as brand context.
- `data/respondents/panel\_config.yaml` as panel and segment configuration.
- `schemas/respondent.schema.json` as the required output schema.

Task

Generate one realistic synthetic respondent at a time for the requested segment.

The caller will run this prompt 5 times per segment, across these 10 segments:

1. `practical\_everyday\_women`
2. `office\_hosiery\_users`
3. `young\_style\_seekers`
4. `mothers\_family\_buyers`
5. `price\_sensitive\_mass\_buyers`
6. `quality\_loyalists`
7. `body\_comfort\_sensitive`
8. `men\_basic\_underwear\_buyers`
9. `marketplace\_first\_buyers`
10. `brand\_switchers`

Inputs

The caller must provide:

```
```json
{
 "respondent_id": "R001",
 "segment_id": "practical_everyday_women",
 "segment_config": {},
 "existing_respondent_summaries": []
}
```

Use `existing\_respondent\_summaries` to avoid clones. Different respondents in the same segment must vary by age, city type, income, shopping channel, purchase frequency, category needs, Conte relationship, competitor set, barriers, and interview answer style.

## Output Rules

Return valid JSON only.

Return a single JSON object only, not markdown, not an array, not commentary.

The object must validate against `schemas/respondent.schema.json`.

Use enum values exactly as defined in the schema. Do not translate enum keys into Russian.

Do not add fields outside the schema. Strict JSON means strict. No "just one helpful extra field" nonsense.

## Respondent Requirements

В PDF только фрагмент. Полный текст есть в data/outputs/conte\_demo/technical\_appendix.md и source file.

## Формула скоринга

$0.25 * \text{relevance\_score} + 0.20 * \text{trust\_score} + 0.25 * \text{purchase\_intent\_score} + 0.15 * \text{emotional\_resonance\_score} + 0.15 * \text{distinctiveness\_score} - 0.15 * \text{objection\_strength\_score}$

Raw weighted index, not normalized to 10. Higher is better; lower objection strength improves relative performance.

Измерения: relevance, trust, purchase intent, emotional resonance, distinctiveness и objection strength. Objection strength вычитается.

## Модель и настройки: model used, temperature/settings

Слой	Реализация	Модель	Настройки	Статус аудита
Генерация респондентов	src/generate_respondents.py	default gpt-4o-mini	JSON object, temperature 0.7, max 3 retries/respondent	actual mode не embedded в panel artifact
Offer study runner / запуск offer study	src/run_study.py	default gpt-4o-mini	JSON object, temperature 0.65, max 3 retries/respondent	usage не логируется
Conte campaign demo / текущий демо-слой	src/run_conte_underwear_campaign_demo.py	без LLM	детерминированные local score rules	current demo outputs generated locally
Экспорт отчётов/PDF/deck	src/export_*	без LLM	local file generation	0 OpenAI requests

OPENAI\_API\_KEY был доступен при генерации appendix: yes. Значение не выводится.

## Стоимость токенов и запросы

Пункт	Значение	Комментарий
Расход токенов	Не записан	OpenAI response usage object не сохраняется.
Стоимость токенов	Не аудирется	Нет usage log и model-price snapshot.
Запросы генерации респондентов	50 calls; до 150 с retries	Один API call на респондента при OpenAI mode.
Запросы offer study	50 calls; до 150 с retries	Один API call на респондента при OpenAI mode.
Conte campaign demo outputs	0 OpenAI calls	Детерминированный local simulation layer.

## Проверки валидации

Проверка	Статус	Деталь
Валидация JSON schema	PASS	0 ошибок схемы
Общее число респондентов	PASS	50 респондентов
Уникальные respondent_id	PASS	50 уникальных id
Число сегментов	PASS	10 сегментов
По 5 респондентов на сегмент	PASS	{'body_comfort_sensitive': 5, 'brand_switchers': 5, 'marketplace_first_buyers': 5, 'men_basic_underwear_buyers': 5, 'mothers_family_buyers': 5, 'office_hosiery_users': 5, 'practical_everyday_women': 5, 'price_sensitive_mass_buyers': 5, 'quality_loyalists': 5, 'young_style_seekers': 5}
Критичные поля респондентов	PASS	0 пустых/отсутствующих критичных полей
Разброс знания Conte	PASS	{1: 2, 2: 9, 3: 13, 4: 16, 5: 10}
Разброс online/offline поведения	PASS	{'offline_first': 18, 'online_first': 18, 'omnichannel': 14}
Полнота study trace	PASS	250 строк trace; 0 неполных
Полнота score matrix	PASS	250 строк score; 0 пропусков score
Message ranking	PASS	5 сообщений в рейтинге
Segment x message heatmap	PASS	10 строк сегментов
Поле proof_needed на уровне панели	WARN	Промпт требует это поле, study traces его содержат, respondent schema его не хранит.

## Версии панели

Артефакт	Доказательство версии
Panel config	a4bda112e5219139 / 2026-06-29 16:48:37 +03
Validated panel	8c24dfc53a7be83a / 2026-06-29 18:38:42 +03
Study trace	71b8cd2f27bacb5f / 2026-06-29 20:06:50 +03

Семантического поля версии панели пока нет; текущая версия = file hash + modified timestamp.

## Снимок рейтинга сообщений

Rank	Message	Weighted	Purchase intent	Objection
1	body_comfort_fit_proof	6.74	7.34	4.16
2	belarusian_quality	6.01	6.18	3.66
3	confidence_without_overpaying	5.86	6.86	5.66
4	everyday_comfort	5.47	6.26	5.66
5	invisible_under_clothes	5.18	5.16	5.66

## Что исправить до внешнего клиентского использования

Добавить run\_metadata.json на каждый pipeline run: model, temperature, seed, git commit, command, start/end time, OpenAI request id, token usage и cost snapshot. Добавить respondent-level proof\_needed или выровнять prompt. Держать synthetic outputs как hypotheses, пока NT DSP или real survey data их не валидируют.