

ТЕХНИЧЕСКОЕ ЗАДАНИЕ / ИКТ-РАЗРАБОТКА И ПИЛОТ

# Разработка высокопроизводительной системы раннего выявления резонансных инфоповодов и динамической оценки информационных рисков

|                      |                                                                                                                            |
|----------------------|----------------------------------------------------------------------------------------------------------------------------|
| Заказчик             | ТОО «IMAS GROUP»                                                                                                           |
| Результат            | Рабочий прототип раннего предупреждения: risk-score, объяснения, дедупликация тревог, API, проверка на историческом потоке |
| Приоритет            | Опережение простых пороговых алертов без избыточных ложных тревог и с контролем ошибок                                     |
| Предварительный срок | 18 недель (предложение к согласованию)                                                                                     |
| Основание            | Отдельный перспективный AI/ML-компонент; использует готовую категоризацию и совместим с внешним модулем кластеризации      |

Предмет — независимая система оценки вероятности и срочности информационного резонанса по динамике инфоповодов. Мультиязычная категоризация публикаций уже существует и не входит в объем разработки. Кластеризация событий создаётся отдельной задачей: MVP рисков работает с согласованным потоком event\_id и тестовыми событиями.

## 1. Вводные разделы

| №     | Элемент шаблона                           | Содержание и требования                                                                                                                                                                                                                              |
|-------|-------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1.1.1 | Наименование                              | Разработка высокопроизводительной системы раннего выявления резонансных инфоповодов и динамической оценки информационных рисков (далее — Система).                                                                                                   |
| 1.1.2 | Краткая характеристика области применения | Мониторинг информационной повестки СМИ и социальных сетей для PR, GR, служб коммуникаций, ситуационных центров и управленческой аналитики; выявление начинающихся резонансных событий и приоритизация информационных рисков для конкретного проекта. |
| 1.2.1 | Документ-основание                        | Публичная технологическая задача ТОО «IMAS GROUP» на Astana Hub. Техническое задание устанавливает условия разработки MVP и опытной апробации без разглашения внутренних данных и систем.                                                            |
| 1.2.2 | Дата утверждения ТЗ                       | После согласования технических, финансовых и правовых условий с выбранным исполнителем.                                                                                                                                                              |
| 1.2.3 | Орган утверждения                         | Уполномоченное лицо ТОО «IMAS GROUP».                                                                                                                                                                                                                |
| 1.3.1 | Наименование объекта                      | Независимый программный движок событийных сигналов: поток публикаций и event_id, история динамики, проектный профиль риска, ранжирование, объяснение, уведомления, обратная связь аналитика и API.                                                   |
| 1.3.2 | Заказчик / разработчик                    | Заказчик — ТОО «IMAS GROUP»; исполнитель определяется по результатам отбора. Контакты и права сторон устанавливаются договором.                                                                                                                      |
| 1.3.3 | Сроки                                     | Предварительно 18 недель: постановка и данные 3, MVP моделей/базисов 6, потоковое API и интерфейс сигналов 4, независимая оценка 3, пилот/передача 2. Итоговые сроки согласовать.                                                                    |
| 1.3.4 | Финансирование                            | Поэтапный договор на MVP и пилот. Смета PoC утверждается после рассмотрения технических предложений, не является предложением оплаты до заключения договора.                                                                                         |

| №     | Элемент шаблона                | Содержание и требования                                                                                                                                                                                                                                                                                                    |
|-------|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1.4.1 | Перечень терминов и сокращений | Инфоповод — динамический набор сообщений об одном событии; risk-score — вычисляемая приоритетность проверки, не доказанный будущий ущерб; lead time — время опережения базового обнаружения; P@K — доля релевантных среди K верхних тревог; event-time replay — воспроизведение потока без просмотра будущих данных.       |
| 1.5.1 | Объект                         | Мультиязычный поток событий с разными темпами упоминаний, динамикой тональности, источниками и метриками взаимодействия. Риск оценивается относительно профиля клиента, отрасли, темы и исторического фона.                                                                                                                |
| 1.5.2 | Цель                           | Разработать программный модуль, который выявляет потенциально резонансные события раньше простого порога частоты, объясняет важность сигнала, снижает шум повторных тревог и эффективно работает в непрерывном потоке.                                                                                                     |
| 1.5.3 | Методы                         | Статистический baseline, анализ изменений и сезонности, ML/гибридное ранжирование, event-time воспроизведение хронологий, сравнение P@K/lead time/FAR, ablation-пробы сигналов, мониторинг изменений качества, нагрузочные тесты.                                                                                          |
| 1.5.4 | Новизна                        | Технологическая новизна проверяется в комплексном механизме: ранняя событийная динамика + клиентская релевантность + мультиязычные признаки + объяснимые причины риска + фильтрация повторных тревог + адаптация к различным базовым уровням активности. Новизна не приписывается существующей тематической категоризации. |
| 1.5.5 | Эффективность                  | Доля действительно полезных тревог в top-K, ложные тревоги на проект/сутки, медианное опережение baseline, P95 задержки сигнала, устойчивость при пропусках признаков и стоимость 1 млн публикаций/событий.                                                                                                                |

## 2. Назначение, цели, объект

| №     | Элемент шаблона            | Содержание и требования                                                                                                                                                                                                                                                                                                                                                                                                  |
|-------|----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2.1.1 | Назначение системы         | Выявлять и ранжировать эволюционирующие информационные события по вероятности резонанса и срочности проверки; предоставлять понятные источники сигнала, оценки уверенности и инструменты обратной связи аналитика.                                                                                                                                                                                                       |
| 2.1.2 | Показатели достижения цели | Предлагаемые ориентиры MVP для согласования: Precision@10 $\geq 0,70$ на эталонном наборе ситуаций; медианное опережение частотного baseline $\geq 30$ мин при сопоставимом уровне ложных тревог; не более 1 ложной тревоги на проект в сутки на согласованных спокойных периодах; P95 публикация→оценка $\leq 120$ с при наличии входного event_id и доступных признаков. Пороги утверждаются после эталонной разметки. |
| 2.2   | Назначение разработки      | Создать внедряемое ИКТ-решение для операционного предупреждения и управленческой аналитики. Продукт не заменяет аналитика, не даёт юридических заключений и не обещает предсказывать реальные события или величину ущерба.                                                                                                                                                                                               |
| 2.3   | Характеристика предмета    | Поток с нерегулярными всплесками, повторными публикациями, задержанными материалами, ночными/праздничными колебаниями, различными географиями и языками, заказчиками с несовпадающими интересами и разным допустимым уровнем тревог.                                                                                                                                                                                     |
| 2.4   | Существующее состояние     | Уже доступны мультиязычная тематическая категоризация, базовые признаки публикаций и метаданные — они рассматриваются как входные данные, без повторной разработки. Мультиязычная кластеризация событий разрабатывается отдельной задачей; для автономного MVP                                                                                                                                                           |

| №   | Элемент шаблона         | Содержание и требования                                                                                                                                                                                                                                                                                                                                                                |
|-----|-------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     |                         | подаются test/event_id или события от адаптера.                                                                                                                                                                                                                                                                                                                                        |
| 2.5 | Научная цель / гипотеза | Гипотеза: совместная оценка скорости роста, ускорения, разнообразия независимых источников, изменения тональности и клиентского контекста при настройке по историческому фону даёт больше полезных ранних предупреждений, чем статический частотный порог, при контролируемом числе ложных тревог. Опровергается/подтверждается сравнением на хронологиях без утечки будущих сведений. |

### 3. Требования к результату / системе

| №      | Элемент шаблона                        | Содержание и требования                                                                                                                                                                                                                                                                                           |
|--------|----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3.1.1  | Структура и функционирование           | Поток публикаций и событий → загрузка уже имеющихся категорий и признаков → оконная оценка изменений → scoring/ранжирование по клиенту → группировка и подавление повторных тревог → API/дашборд/уведомления → экспертная обратная связь и журнал версий. Слой событий допускает подключение отдельного кластера. |
| 3.1.2  | Показатели назначения                  | Воспроизводимые тесты для 100 тыс., 1 млн и 5 млн публикаций/сутки с числом событий, отражающим многократные упоминания; устойчивое вычисление для нескольких независимых клиентских проектов; P95 задержки и стоимость обработки на фиксированном оборудовании.                                                  |
| 3.1.3  | Надёжность                             | Идемпотентность повторных событий, восстановление после перезапуска, учёт запоздавших публикаций, ограничение частоты алертов, отсутствие неконтролируемой эскалации при отказе внешних ML-компонентов.                                                                                                           |
| 3.1.4  | Безопасность                           | Разграничение клиентских профилей, тем и каналов уведомлений; защищённые API, хранение истории ручных решений, ограничение использования непроверенных внешних источников в автоматическом выводе.                                                                                                                |
| 3.1.5  | Эргономика и техническая эстетика      | Демонстрационный экран или API-клиент: отсортированная лента риск-событий, объяснения, основные публикации, график роста, подтверждение/отклонение аналитиком, фильтры и настройка уведомлений. Полный промышленный UI не обязателен.                                                                             |
| 3.1.6  | Эксплуатация и обслуживание            | Воспроизводимый запуск, журнал сообщений и оценок, мониторинг задержек, версий моделей, ложных тревог и изменений базового уровня; возможная автономная установка на собственной инфраструктуре.                                                                                                                  |
| 3.1.7  | Защита от несанкционированного доступа | Изоляция профилей клиентов, RBAC, защита webhooks, аудит всех изменений порогов и пользовательских решений, минимизация персональных данных в отчётности.                                                                                                                                                         |
| 3.1.8  | Сохранность при авариях                | Оценки и история уведомлений должны сохраняться при перезапуске; при недоступности несущественных признаков формируется статус частичной уверенности, а не безусловный нулевой риск. Дублирующая доставка не порождает повторную тревогу.                                                                         |
| 3.1.9  | Защита от внешних воздействий          | Пики сообщений, отсутствие категории/тональности, пропуски метрик, короткие тексты, задержанные публикации, сбои базы/сети/AI и временное отсутствие событийного адаптера — с управляемой деградацией и явной отметкой ограничений.                                                                               |
| 3.1.10 | Патентная чистота                      | Показать перечень моделей, библиотек и лицензий, допустимость коммерческого использования и локального развёртывания; ограничить передачу закрытых данных третьим сторонам.                                                                                                                                       |
| 3.1.11 | Стандартизация / унификация            | Форматы событий и оценок, API и схемы, версии оценочных моделей, UTC time/event-time, стабильные идентификаторы, экспорт наблюдений и                                                                                                                                                                             |

| №      | Элемент шаблона                   | Содержание и требования                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--------|-----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|        |                                   | событий для независимой оценки.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 3.1.12 | Дополнительные требования         | Передать шаблоны пользовательских профилей, документацию, тестовые последовательности и рекомендации по настройке порогов для разных информационных тематик.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 3.2.1  | Функции и критерии отказов        | Получение сигналов из временных окон и статистических отклонений, учёт независимости источников и вторичных пересказов, базовых категорий, изменения тональности, динамики вовлечённости (при доступности) и персональных интересов клиента. Сравнить пороговый метод, change-point/seasonal baseline и ML. Возвращать event_id, project_id, score, urgency, confidence, key_features, evidence_ids, timestamp и версию модели; предупреждать об ограниченных данных. Подавлять дубли тревог, обновлять риск и журнал решений. Не пытаться заново категоризировать весь поток или строить собственную полноценную событийную кластеризацию. |
| 3.3.1  | Информационное обеспечение        | Открытые, синтетические и правомерно обезличенные временные хронологии; заранее определённые периоды «до» и «после» резонанса, контрпримеры с плановыми кампаниями, новостными циклами и слабым шумом; независимая экспертная разметка полезности сигналов.                                                                                                                                                                                                                                                                                                                                                                                 |
| 3.3.2  | Лингвистическое обеспечение       | Вход — мультязычные тексты, готовые категории, языки и тональность, при необходимости нормализованные географии; полнота признаков по редким языкам оценивается отдельно, но повторное обучение категоризатора не является предметом.                                                                                                                                                                                                                                                                                                                                                                                                       |
| 3.3.3  | Программное обеспечение           | Выбор методов открыт: статистические пороги, модели временных рядов, компактные ML и выборочные LLM для объяснения. Архитектура API независима от конкретного поискового движка и базы; массовая тяжёлая LLM не является обязательной.                                                                                                                                                                                                                                                                                                                                                                                                      |
| 3.3.4  | Техническое обеспечение           | Представить ресурсы CPU/GPU/RAM, задержку оценивания, потребление памяти исторических окон и стоимость обработки 1 млн публикаций или 100 тыс. событий при согласованном распределении.                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 3.3.5  | Организационное обеспечение       | Согласовать описание событий, профили заказчиков, протокол разметки, политику неиспользования будущей информации, версии моделей, ручную проверку лучших и ошибочных тревог.                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 3.4    | Требования к программному изделию | Рабочий потоковый движок или развёртываемый сервис, API оценок и уведомлений, отдельные клиентские профили, демонстрационная лента риска, аналитический журнал, модуль обратной связи и комплект реплей-тестов.                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 3.5    | Информационная безопасность       | Применение открытых и согласованных обезличенных данных; разграничение доступа к профилям клиента, источникам и истории оценок; запрет автоматической рассылки чувствительных данных во внешние LLM без разрешения.                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 3.6    | Промышленная безопасность / ESG   | Прямых требований промышленной безопасности нет. Отдельно оценить рациональность CPU/GPU и долю тяжёлых моделей, ограничения автоматических рекомендаций и риск неверной эскалации.                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

## 4. Работы, сроки, приёмка

| №     | Элемент шаблона                | Содержание и требования                                                                                                                                                                                                                                        |
|-------|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4.1.1 | Стадии, исполнители, документы | Этап 1 (недели 1–3): критерии полезного раннего сигнала, replay-корпус, размеченные примеры и baselines; этап 2 (4–9): модели/алгоритмы событийного scoring; этап 3 (10–13): потоковый модуль, API и алерты; этап 4 (14–16): независимые тесты качества/ложных |

| №     | Элемент шаблона                  | Содержание и требования                                                                                                                                                                                                                                                                                   |
|-------|----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|       |                                  | тревог/опережения/нагрузок; этап 5 (17–18): пилотное подключение на тестовом event-потоке, передача.                                                                                                                                                                                                      |
| 4.2.1 | НИР → ОКР → ОПЭ / внедрение      | НИР — проверяемая гипотеза опережения baseline. ОКР — производительный сервис с объяснимостью и настройками. Опытная эксплуатация — воспроизведение прошлых хронологий с имитацией поступления сообщений в реальном времени, затем пилот при согласованных ограничениях.                                  |
| 4.3.1 | Контроль и приёмка               | Комиссия заказчика проверяет эталонные хронологии, сравнивает методы на одинаковом потоке, оценивает P@10, количество ложных тревог, lead time, журнал и API, работу с задержанными событиями, поведение при сбоях и воспроизводимость запуска. Не допускается утечка будущей информации в ранние оценки. |
| 4.4.1 | Технико-экономические показатели | Таблица с P@5/P@10, false alarms на проект/сутки, медианным и P90 опережением базового сигнала, P95 latency, пропущенными резонансными кейсами, CPU/GPU стоимостью на 1 млн входов, качеством по языкам и типам событий. Пороги из 2.1.2 — ориентиры к согласованию, не гарантированные результаты.       |
| 4.5.1 | Порядок оплаты                   | По договорённости с исполнителем; возможна этапная оплата MVP и пилота. Бюджет не утверждён; стоимость определяется по итогам предложений. Условия оплаты и интеллектуальной собственности фиксируются договором.                                                                                         |

## 5. Подготовка объекта и документирование

| №     | Элемент шаблона               | Содержание и требования                                                                                                                                                                                                                                                          |
|-------|-------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 5.1   | Подготовка к вводу в действие | Подготовить тестовые event_id, многодневные хронологии публикаций, набор готовых категорий, клиентские профили и независимую экспертную разметку. На первом этапе не нужен доступ к production.                                                                                  |
| 5.2.1 | Комплект документации         | Схема сигналов и клиентских профилей; спецификация API; воспроизводимый evaluation harness; набор эталонных сценариев и негативных примеров; отчёт о метриках и ложных тревогах; инструкция запуска, наблюдаемость и ограничения применимости; перечень зависимостей и лицензий. |
| 5.3.1 | Источники разработки          | Публичное описание задачи, настоящее ТЗ, легально доступные корпуса и обезличенные данные, общедоступные методики оценки обнаружения изменений и информационной динамики.                                                                                                        |

## 6. Квалификационные требования

| №     | Элемент шаблона              | Содержание и требования                                                                                                                                                                       |
|-------|------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 6.1.1 | Материально-техническая база | Опыт потоковой аналитики, временных рядов и ML/NLP, написания API и создания работающих прототипов, организации независимой разметки и предотвращения утечки будущих событий в экспериментах. |
| 6.2   | Разрешительные документы     | Специальные разрешения для разработки ПО не предусматриваются; необходимо законное использование исходных данных, моделей и лицензий.                                                         |
| 6.3   | Гарантийные обязательства    | Исполнитель исправляет дефекты согласованного MVP, обнаруженные при приёмке, и передаёт воспроизводимый комплект тестов и инструкцию запуска; сроки гарантии согласуются договором.           |