



О создании информационной системы раннего выявления угроз в информационной сфере и прогнозирования рисков их возникновения (ИС ВЕПРЬ)

ГЛАВНЫЙ РАДИОЧАСТОТНЫЙ ЦЕНТР

КОРОСТАШОВ
Роман Николаевич

Начальник отдела анализа и прогнозирования ИТЦ ФГУП «ГРЧЦ»

2022 г.



Рост аудитории социальных сетей, форумов, сервисов мгновенного обмена сообщениями

Значительное увеличение объемов информации в СМИ и сети Интернет

Широкое распространение материалов, оказывающих на человека негативное информационно-психологическое воздействие

дестабилизация социально-политической обстановки,
манипуляция общественным мнением,
дезинформация,
дискредитация традиционных ценностей.

Невозможность выявления и прогнозирования таких материалов в «ручном» режиме.

Научно-техническая задача:

изучить возможность применения технологий ИИ для мониторинга информационного пространства, выявления угроз и прогнозирования рисков их возникновения.



Достигнутые результаты НИР ОСНОВА

Целесообразна разработка информационной системы раннего выявления угроз в информационной сфере и прогнозирования рисков их возникновения (ИС ВЕПРЬ) на базе технологий ИИ.

В ходе НИР подготовлены:

1. Методологическое обоснование работ по созданию ИС ВЕПРЬ

- Исследование состояния предметной области
- Анализ и разработка методов детекции и прогнозирования угроз с помощью ИИ

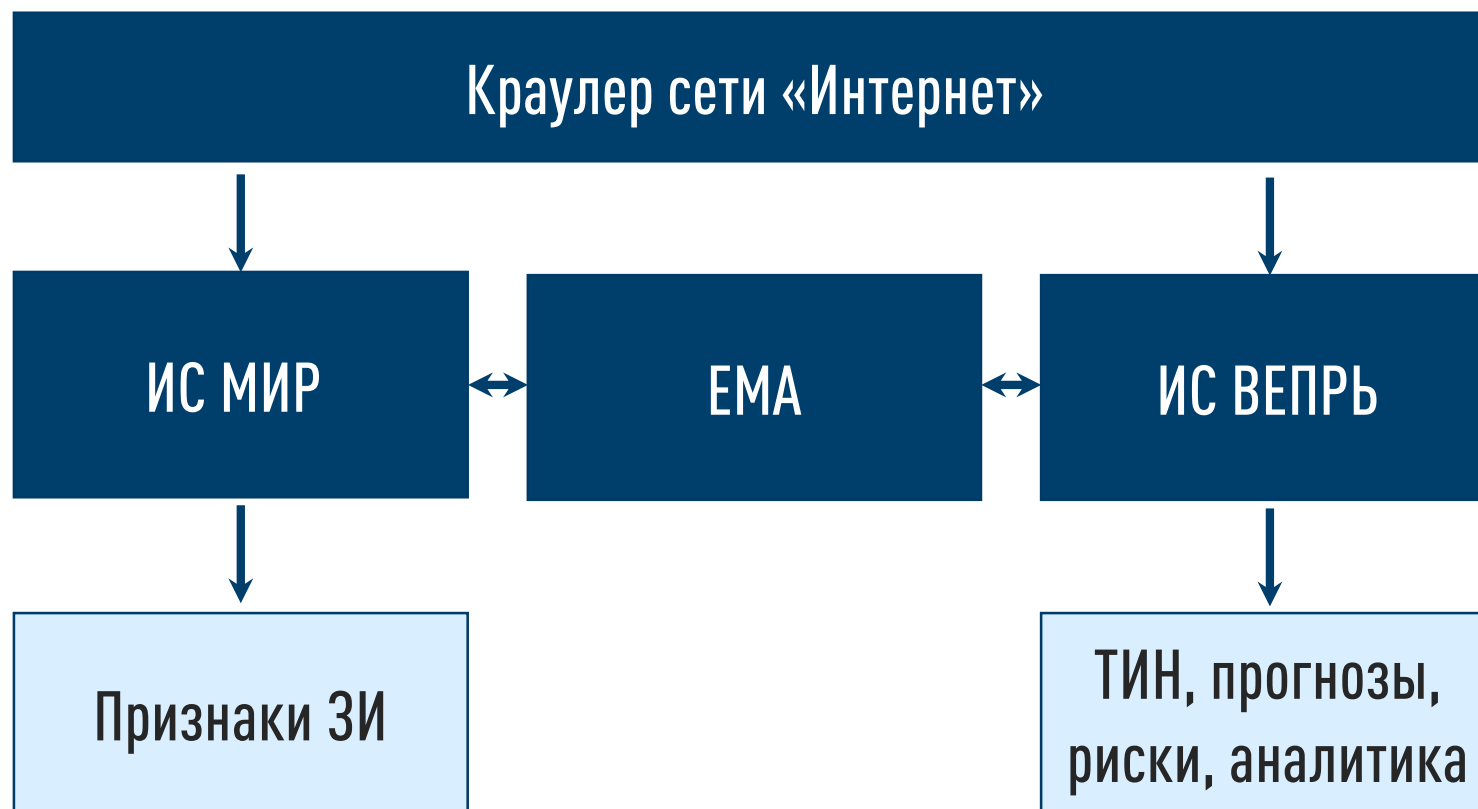
2. Определение принципов работы ИС ВЕПРЬ

- Функциональное проектирование
- Описание алгоритмов работы технических компонентов ИС

3. Комплект документации на разработку ИС ВЕПРЬ

- Концептуальная архитектура
- Техническое задание
- ТЭО
- Дорожная карта
- Требования к защите информации

Сферы применения информационных систем



Первая очередь разработки ИС

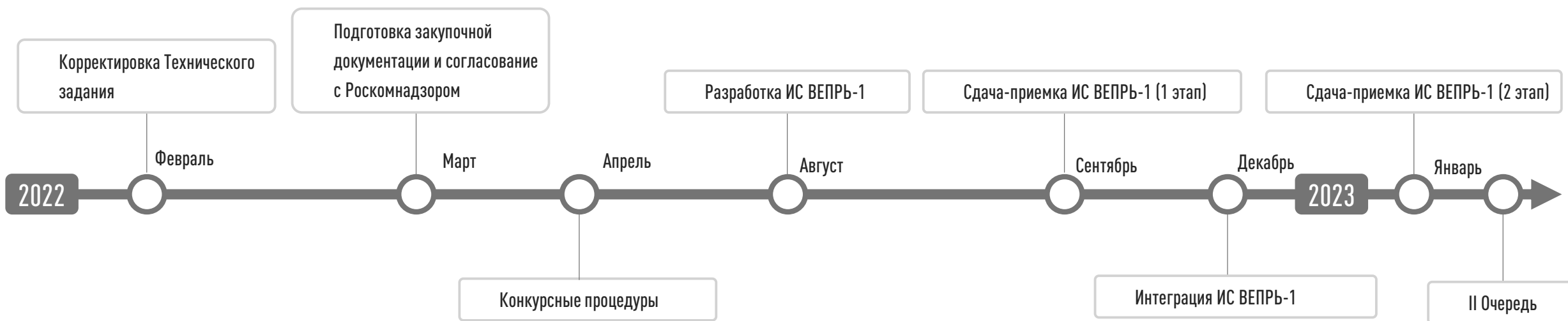


Функциональность первой очереди ИС:

- прогнозирование рисков возникновения и распространения ТИН федерального и локального уровня;
- выявление протестных настроений и фактов дестабилизации российского общества (в т.ч. по темам территориальной целостности, межнациональной розни, миграционной политики и др.);
- выявление негатива в отношении первых лиц государства, государственных структур, межгосударственных организаций;
- выявление (по представленному образцу) фейков про первых лиц государства, про государство и страну в целом;
- выявление манипуляций общественным мнением, поляризации общественного мнения (в т.ч. по темам вакцинации, внесистемной оппозиции, санкционного давления и др.);
- подмену и дискредитацию традиционных мировоззренческих, духовно-нравственных, культурных и семейных ценностей (в т.ч. пропаганды сексуальных «свобод», легализации наркотиков, прозападных ценностей и др.).



Первая очередь создания ИС



Целевые показатели:

- объем анализируемых материалов – до 20 тыс. в сутки;
- точность детекции – не ниже 85%;
- срок выявления угрозы*:
 - высокорисковые СМИ / вебсайты – 20 мин.
 - соцсети – 12 часов
 - Telegram – 20 мин.

* Скорость выявления угроз и объемы анализа материалов определяются сервисом краулера

Вторая очередь разработки ИС



01 февраля 2023 – 31 января 2024

Цели:

- выявление ТИН в изображениях, аудио и видео
- **применение риск-ориентированного подхода**
Определение на основании уже выявленных ТИН, а также вручную приоритетных для мониторинга:
 - тематик (например, «протесты»),
 - источников (напр., «оппозиционные каналы Telegram»),
 - авторов (напр., «иноагенты»).
- повышение объема анализируемых материалов с 20 до 100 тысяч в сутки
- сокращение срока выявления угрозы*:
 - высокорисковые СМИ / вебсайты – 10 мин.
 - соцсети – 6 часов
 - Telegram – 10 мин.

* Скорость выявления угроз и объемы анализа материалов определяются сервисом краулера

Третья очередь разработки ИС



01 февраля 2024 – 30 сентября 2024

Цели:

- проверка фактов и выявление дезинформации, обмана, недостоверных медиаобъектов (дипфейков), фейковых новостей
искажение истории,
фейки в отношении первых лиц государства, государственных структур и страны в целом,
намеренное распространение дезинформации с целью введения в заблуждение.
- Увеличение объёмов, анализируемых **видео** за счёт увеличения мощности серверного оборудования
- повышение объема анализируемых материалов со 100 до 200 тысяч в сутки
- повышение точности детекции не ниже 90%
- сокращение срока выявления угрозы*:
 - высокорисковые СМИ / вебсайты – 5 мин.
 - соцсети – 1 час
 - Telegram – 5 мин.

* Скорость выявления угроз и объемы анализа материалов определяются сервисом краулера

Запланированные затраты на создание ИС



Затраты, тыс. руб.	1-я очередь (2022-2023 г.)	2-я очередь (2023-2024 г.)	3-я очередь (2024 г.)
Разработка ИС ВЕПРЬ	60 000	Будет уточнено по результатам предыдущих периодов*	
Приобретение оборудования **	180 000		

* Наиболее затратным функционалом является:

- выявление дипфейков;
- проверка фактов;
- сокращение срока выявления угрозы с нескольких часов до нескольких минут.

** Общая инфраструктура для нескольких проектов (МИР, ОКУЛУС, ВЕПРЬ, ЕМА)

Ожидаемый эффект от создания ИС ВЕПРЬ



- Максимальная автоматизация процесса мониторинга и выявления точек информационной напряженности:
 - Сокращение времени выявления ТИН в следствии автоматизации процесса сбора и выявления ТИН до нескольких минут
 - Повышение качества выявления ТИН
 - Прогнозирование распространения ТИН, оценка рисков их распространения
 - Анализ всех видов контента (текст, изображения, аудио, видео)
- Предотвращение митинговой активности
- Выявление манипуляций в информационной сфере
- Оценка поляризации отношений комментаторов по интересующей тематике
- Выявление недостоверной информации

Спасибо за внимание!