



Российский форум
«Молодёжь и наука»

15-17 мая 2021

Большие данные и общество: возможности и вызовы для молодёжи

Концепция **ЦИ**фровой **ПлатФ**ормы «**ЦИПФ**» для управления социальной и политической активностью молодёжи

Воронцов Константин Вячеславович

д.ф.-м.н., проф. РАН, заведующий лабораторией Машинного Интеллекта МФТИ

Карабулатова Ирина Советовна

д.филол.н., проф., профессор-исследователь РУДН

Предпосылки

1. Постоянно возникают новые инициативы и виды позитивной профессиональной и общественно-политической активности
2. Молодёжная активность хорошо заметна в СМИ и соцмедиа
3. Её выявление, систематизация и своевременная поддержка
 - эффективнее, чем формирование и навязывание инициатив сверху
 - эффективнее, чем сбор заявок на поддержку будущих активностей
 - является средством пассивного предупреждения негативных активностей
4. Возможные виды поддержки:
 - информационная, организационная, финансовая

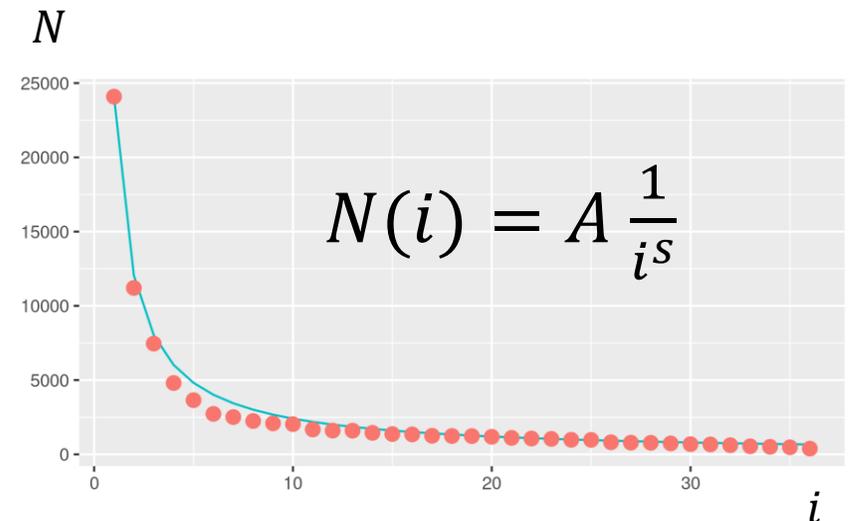
Закон Ципфа

- частота слов в текстах
- доходы компаний / людей в стране
- численность населения в городах
- численность людей, вовлечённых в различные социальные активности

Распределение с «тяжёлым хвостом»: нельзя пренебрегать большим числом по отдельности редких активностей



Джордж Ципф
(1902—1950)



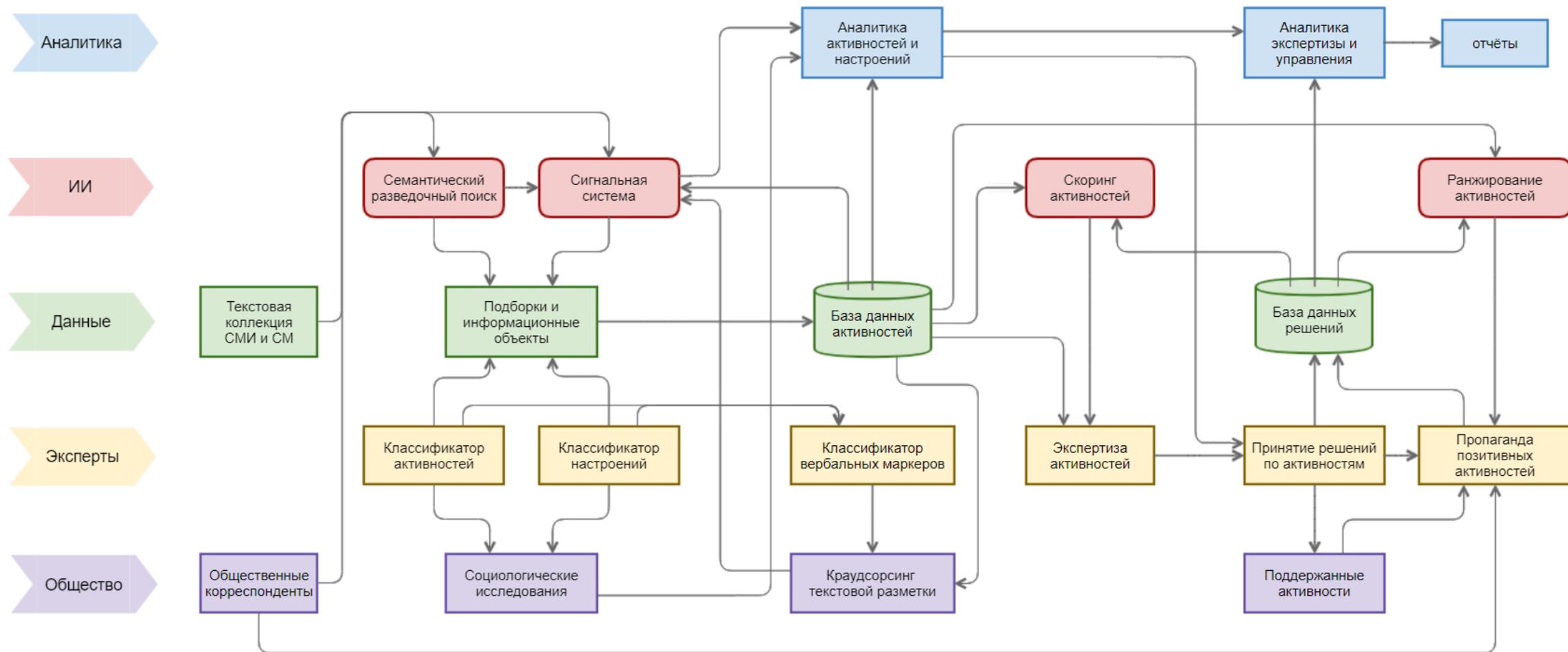
Активность — основной объект анализа

Активность — конкретная инициатива, мероприятие или серия регулярных мероприятий

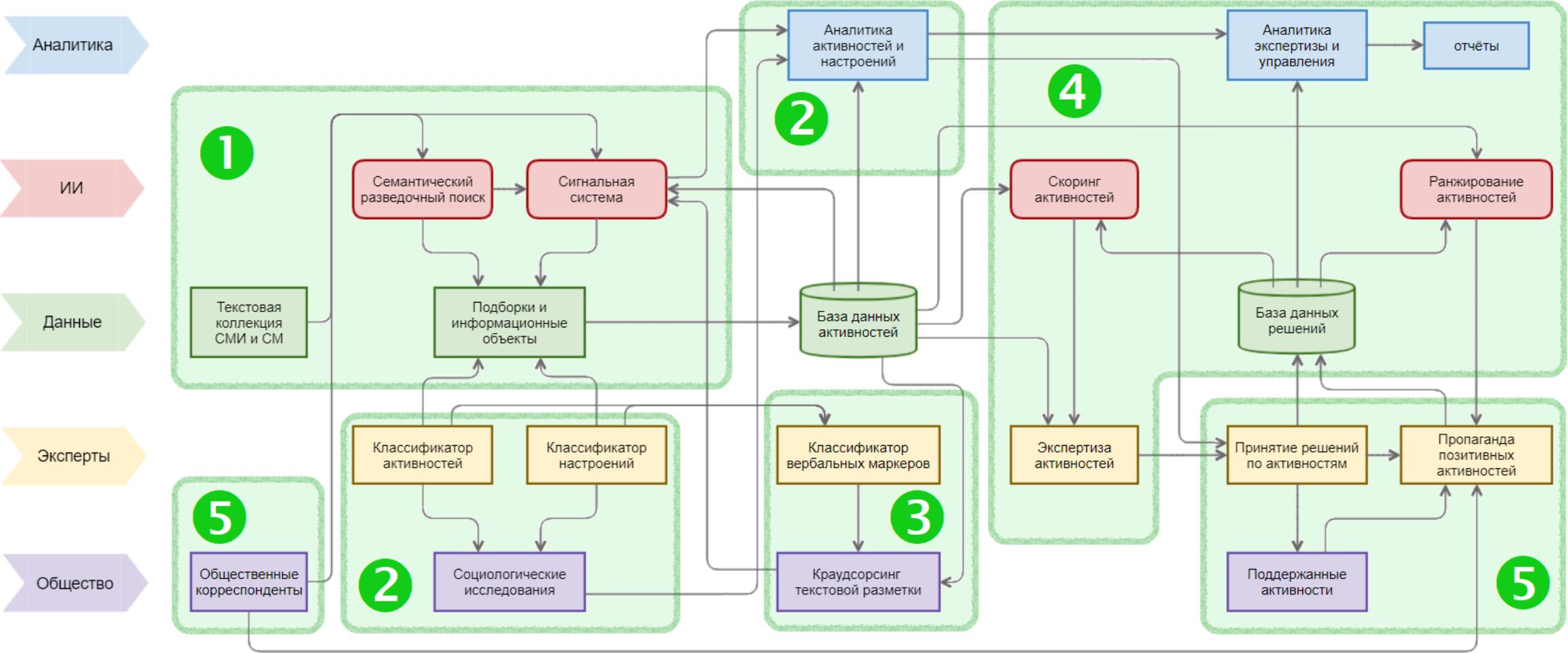
Паспорт активности:

- подборка текстовых материалов СМИ и соцмедиа
- атрибутивное описание:
организатор, дата-время, локация, формат, число участников,...
- полярность: позитивная / негативная / нейтральная
- информация о поддержке

Бизнес-логика процессов и потоки данных

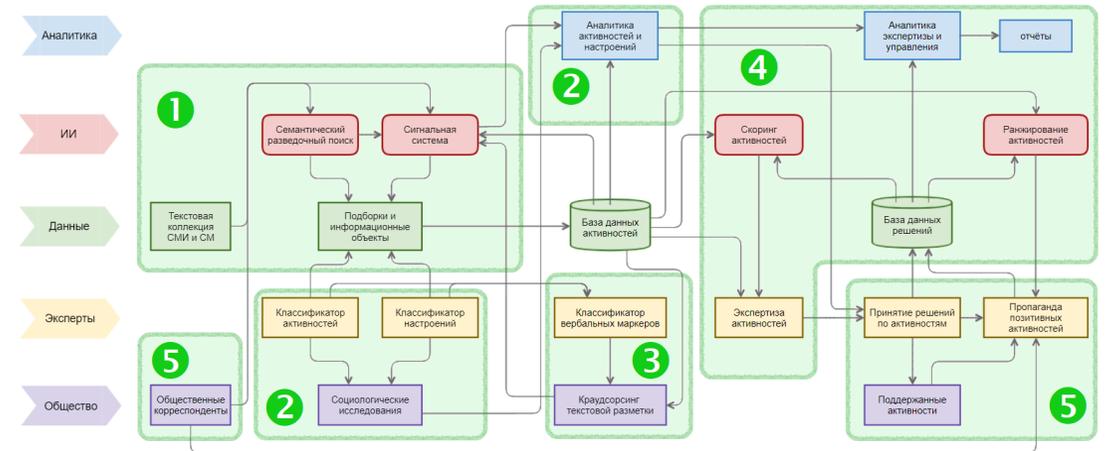


Бизнес-логика процессов и потоки данных



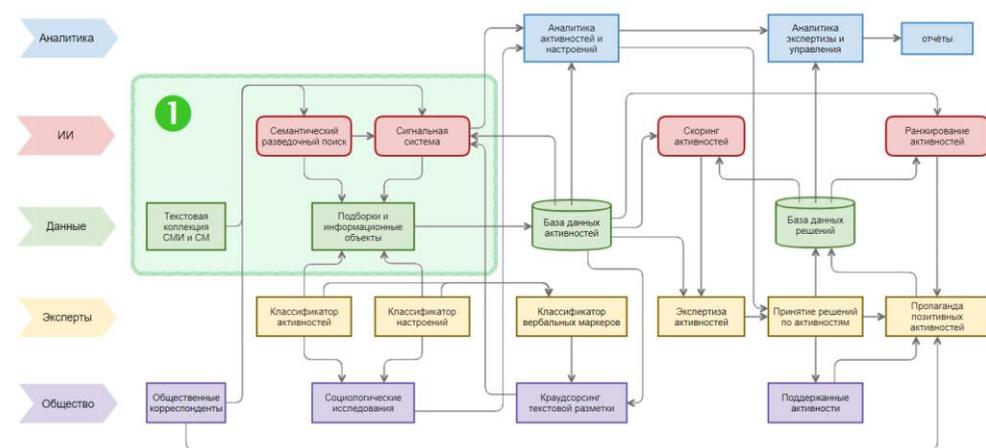
Пять взаимосвязанных контуров

- 1 Мониторинг СМИ и соцмедиа**
— формирование базы активностей
- 2 Социогуманитарная экспертиза**
— классификация активностей
и общественных настроений
- 3 Лингвистическая экспертиза**
— выделение вербальных маркеров
- 4 Принятие решений**
— экспертиза и скоринг активностей
- 5 Исполнение решений и обратная связь**
— поддержка активностей, масштабирование и PR



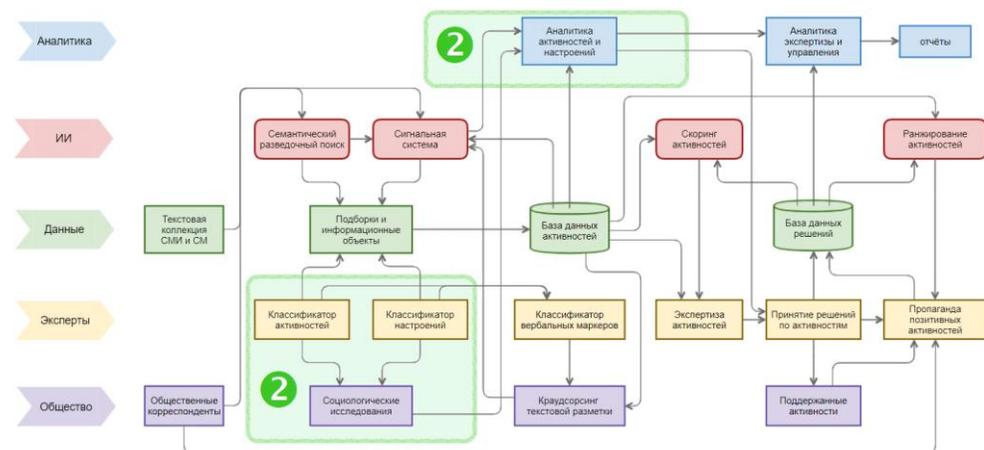
1 Мониторинг СМИ и соцмедиа

- Исходные данные:
 - новостные потоки СМИ
 - социальные медиа
 - общественные корреспонденты
- Семантический разведочный информационный поиск
 - поиск по смыслу, а не по ключевым словам
 - автоматизация рутинных операций поиска
 - формирование тематических подборок по активностям
- Сигнальная система
 - обнаружение новых активностей и общественных настроений
 - накопление данных о распространении активностей



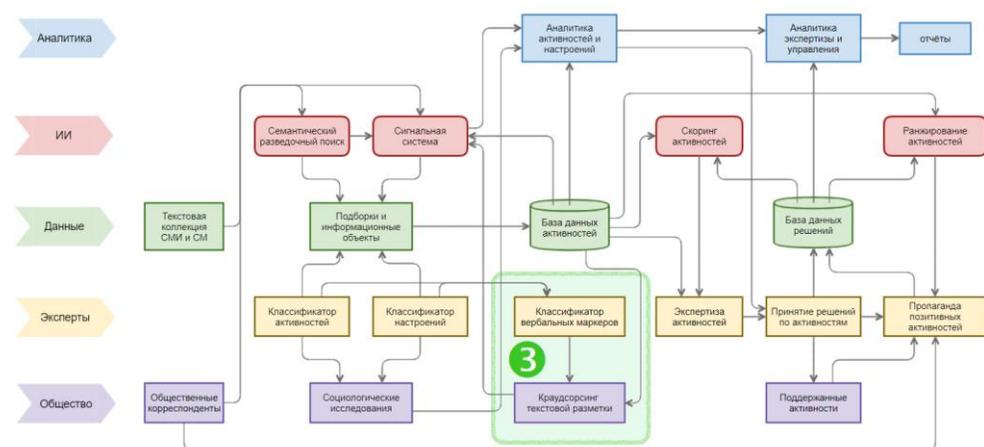
2 Социогуманитарная экспертиза

- **Классификатор активностей**
 - систематизация типов активностей
 - формирование «паспорта активности»
- **Классификатор общественных настроений и поляризации**
 - выявление ценностных ориентиров,
 - виды позитивных и негативных настроений
 - психологическая и психолингвистическая экспертиза
 - политологическая экспертиза
 - количественный и качественный анализ
- **Социологические исследования**
 - калибровка количественных оценок активностей и настроений
 - оценивание размеров соответствующих молодежных аудиторий и когортных групп



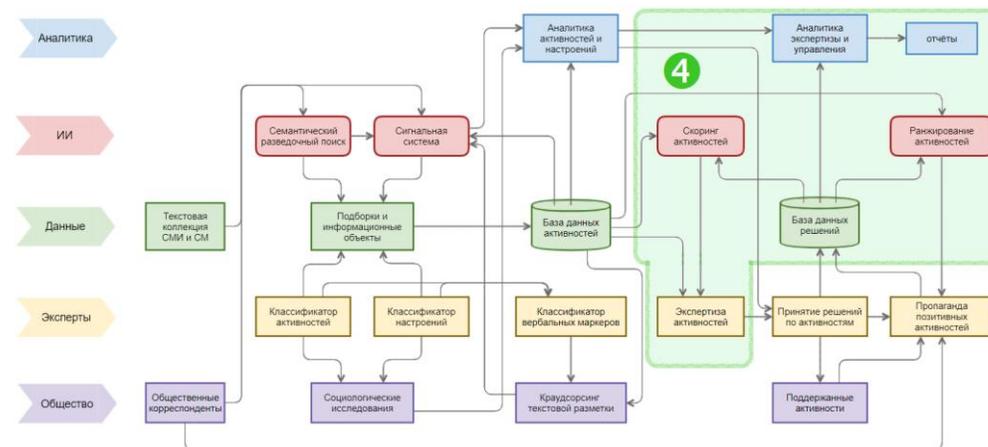
3 Лингвистическая экспертиза

- **Вербальные маркеры**
 - признаки для обучаемых моделей классификации активностей и настроений по текстовым материалам
 - привлечение экстралингвистических данных
- **Краудсорсинг**
 - методология разметки текстовых данных
 - связи между вербальными маркерами и классами активностей и настроений
- **Сигнальная система**
 - Размеченная выборка используется в контуре мониторинга для обучения моделей ИИ



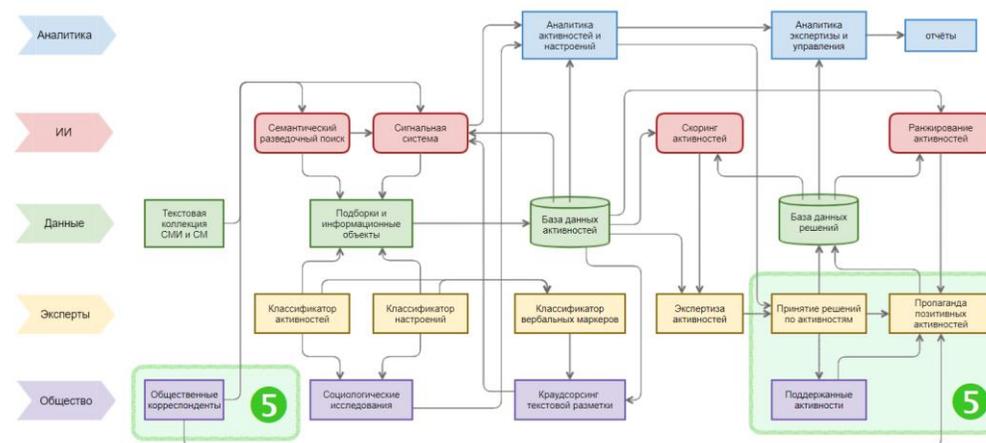
4 Принятие решений

- **Экспертиза активностей**
 - поддержка работы экспертов
- **Скоринг активностей**
 - автоматизация работы экспертов
 - генерация рекомендаций экспертам
 - ранжирование выявленных активностей
- **Принятие решений**
 - поддержка позитивных активностей
 - решения о видах и форме поддержки
 - пресечение негативных активностей
 - решения о каналах взаимодействия
 - формирование базы данных решений



5 Исполнение решений и обратная связь

- **Пиар позитивных активностей**
 - создание контента
 - организация обмена позитивным опытом
 - привлечение общественных корреспондентов
- **Исполнение решений**
 - информационная поддержка процессов исполнения решений
- **Обратная связь**
 - мониторинг случаев поддержки
 - количественные оценки эффективности управления



Резюме

- Целенаправленная поддержка позитива не оставляет негативу шансов завладеть умами
- ЦИПФ — системный подход к организации диалога поколений, способ предлагать молодёжи привлекательные альтернативы
- Цифровые технологии и технологии Искусственного Интеллекта созрели для автоматизации решения данной задачи
- Для реализации проекта уже собрана сильная междисциплинарная экспертиза: ИИ + социология + лингвистика + молодёжная политика