



# Речевая аналитика для служб безопасности

## Автоматическое определение тематик телефонных переговоров

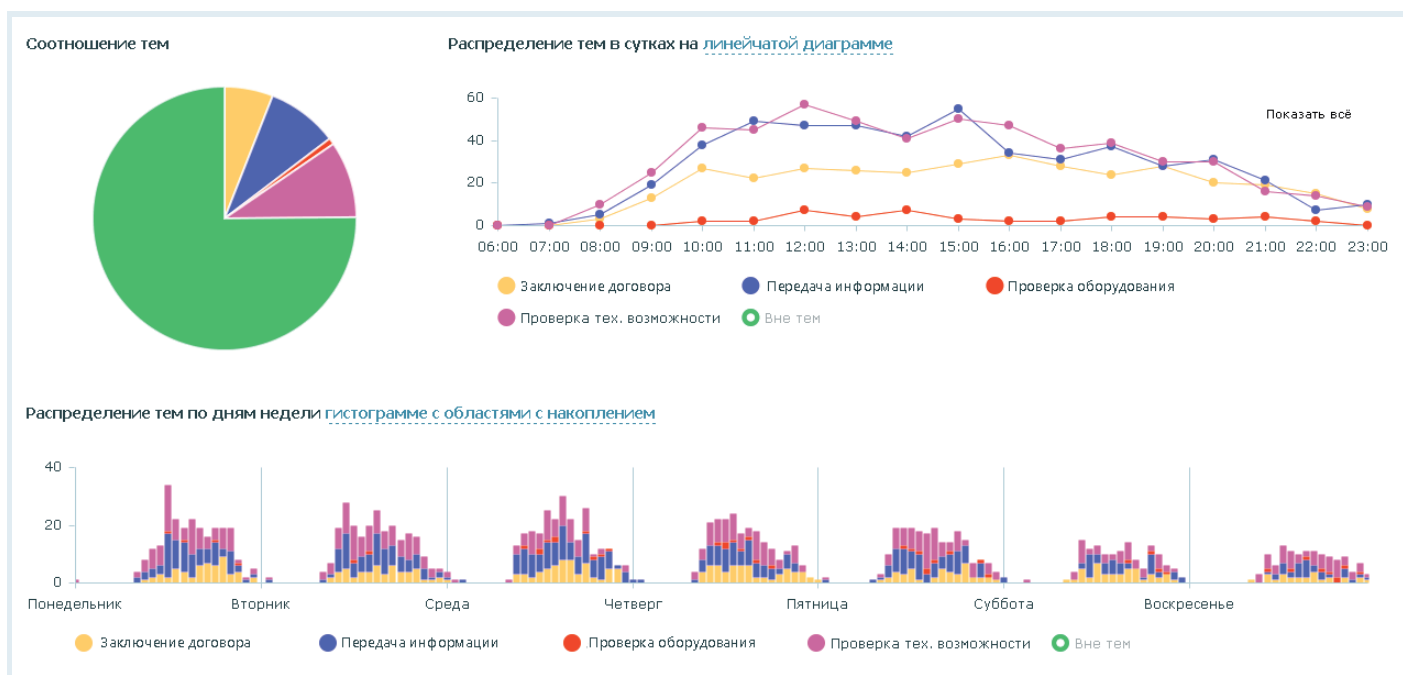
Благодаря высокому уровню развития технологии распознавания спонтанной слитной русской речи, клиентам ООО «ЦРТ» стали доступны инновационные и высокоэффективные средства мониторинга и анализа речевой информации, передаваемой по голосовым каналам связи.

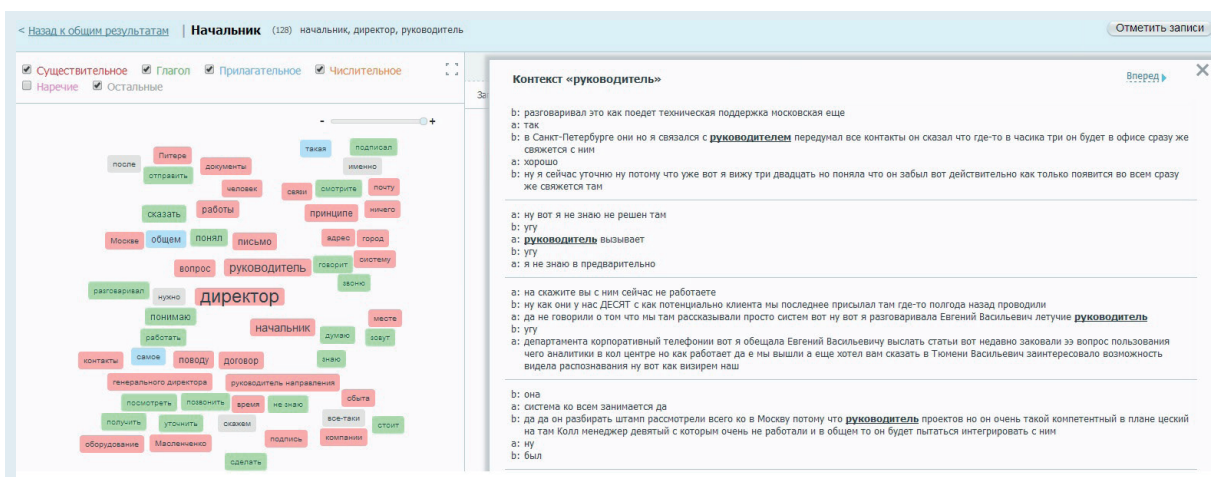
Автоматический перевод фонограмм телефонных переговоров в текст позволяет осуществлять поиск целевой информации и проводить автоматический анализ содержания разговоров, выявлять тематические кластеры на огромных массивах речевой неструктурированной информации в считанные секунды, что еще совсем недавно казалось невозможным.

### Система автоматического определения тематик способна:

- без участия эксперта сформировать выборки фонограмм по темам, требующим постоянного и пристального контроля со стороны специалистов служб безопасности
- в автоматическом режиме выявить кластеры телефонных разговоров, в которых происходит, например, обсуждение конфиденциальных вопросов и коммерческих интересов сторон, передача персональных данных или важных технологических особенностей
- из многотысячного массива телефонных звонков крупного предприятия выявить телефонные переговоры сотрудников на бытовые темы и отбросить их при рассмотрении инцидентов производственного характера.

Для работы системы тематической кластеризации не требуется какого-либо обучения движков распознавания речи, подбора словарей контролируемой лексики или других технологических процедур. Система приступит к анализу речевой информации сразу, как только получит доступ к фонограммам звонков.





В основе технологического базиса системы лежат технологии распознавания слитной речи на большом словаре (LVCSR – Large Vocabulary Continous Speech Recognition) и извлечения информации с помощью кластерного анализа данных (Data Mining clustering). В результате автоматического распознавания речь дикторов преобразуется в текстовый индексируемый файл, пригодный для автоматического лексико-семантического анализа.

Решение о принадлежности звонка к абстрактному тематическому кластеру проводится с учетом частотности и связности слов и словосочетаний, употребляемых дикторами в ходе телефонной беседы.

Система проста в работе, имеет продуманный эргономичный интерфейс, основная задача которого дать корректное, полное и обобщенное представление результатов семантического анализа огромных массивов в тысячи и десятки тысяч звонков. Для решения такой нетривиальной задачи используется представление в виде «семантического облака» или «облака тэгов». Легенда «облака» содержит информацию о мощности тематического кластера, частоте использования основных лексических единиц, определяющих тему.

Возможно управление «облаком» на уровне переопределения смыслового центра темы, вынесения отдельных слов в стоп-лист, отключения неинформативных для конкретного кейса отдельных частей речи (существительных, глаголов, числительных и др.). Также доступна контекстная сводка, представляющая дословное содержание ключевых фрагментов разговоров по выявленной теме. Возможен поиск тематического кластера в экспертном режиме, предусматривающем создание пользователем собственной темы с целевым набором ключевых слов, характеризующих, по мнению, эксперта тему разговора.

## ОБЛАСТИ ПРИМЕНЕНИЯ СИСТЕМЫ

- корпоративные службы безопасности: профилактика утечек информации, контроль эффективности взаимодействия компании с клиентом, выявление нарушений корпоративной этики
- государственные силовые структуры: интеллектуальный мониторинг и автоматический анализ телефонных разговоров в круглосуточном необслуживаемом режиме

## ЭФФЕКТЫ ВНЕДРЕНИЯ СИСТЕМЫ

- многократное сокращение времени поиска целевого контента в массивах неструктурированной речевой информации
- повышение эффективности мероприятий за счет тотального автоматического мониторинга разговоров по целевым темам
- профилактика и предупреждение инцидентов за счет выявления разговоров по новым и нехарактерным темам
- повышение производительности обработки речевой информации за счет выявления кластеров разговоров по непродуктивным темам



### Санкт-Петербург

194044, Выборгская наб., 45, лит. Е  
Тел.: +7 812 325 8848  
Факс: +7 812 327 9297  
stc-spb@speechpro.com

### Москва

109004, Земляной Вал, 59, стр. 2  
Тел.: +7 495 669 7440  
Факс: +7 495 669 7444  
stc-msk@speechpro.com

### Партнеры в Беларуси

220037, Минск, пер. Уральский, 15, оф. 403  
Тел.: +375 17 266 9312  
Факс: +375 17 266 9312  
info@speetech.by  
www.speetech.by



WWW.SPEECHPRO.RU