NLP В ЗАДАЧАХ КИБЕРБЕЗОПАСНОСТИ

Чернышов Юрий
ychernyshov@ussc.ru

Язык людей и машин

NLP одно из направлений на пути движения к сильному искусственному интеллекту (AGI)

Рей Курцвейл – неизбежен симбиоз людей и машин
“Transcend: nine steps to living well forever”

Некоторые примеры задач NLP

Анализ языка:

language modelling,

sentiment analysis,

text classification,

named-entity recognition,

natural language inference,

relation extraction,

semantic parsing,

co-reference resolution,

entity linking,

relational reasoning,

semantic composition,

language identification and translation,

entity and information extraction,

intent detection and classification,

stance and fake news detection, rumor detection, hate speech detection, clickbait detection, abuse detection.

Генерация языка:

question-answering systems,

text and dialogues generation,

text summarization,

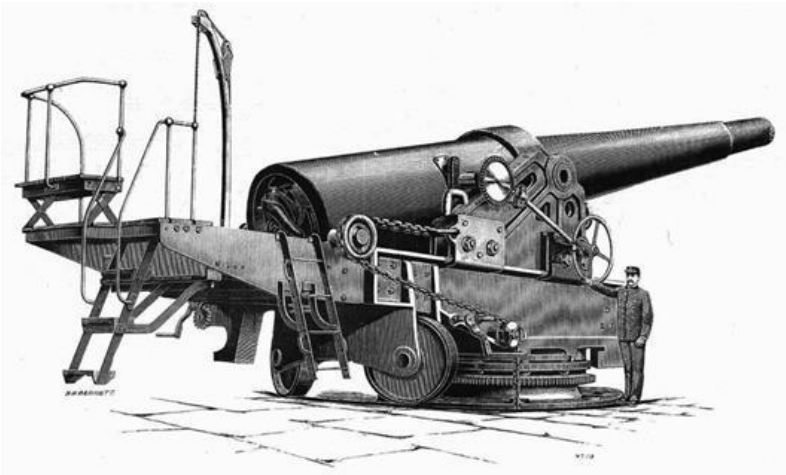
slot filling for knowledge base population tasks,

scripts and programming code generation.

Применение ИИ в кибербезопасности

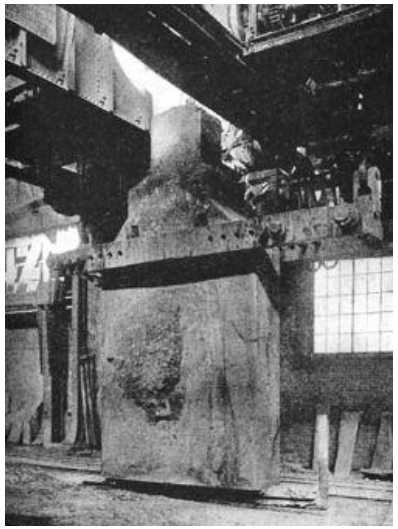
Атака

- маскировка, адаптация зловреда
- интеллектуальный подбор данных
- фишинг
- DeepFake
- RL для определения тактики атаки



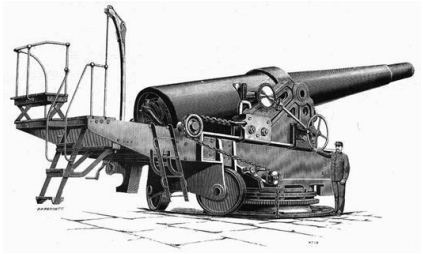
Оборона

- продвинутый мониторинг и анализ
- распознавание 0-day атак
- форензика
- реагирование
- RL для определения защитных мер

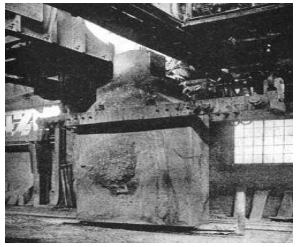


Некоторые задачи кибербезопасности, в которых используются методы NLP

Сбор информации (crawlers, scrappers)



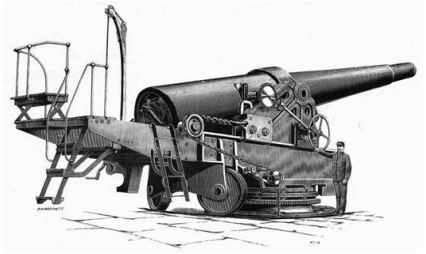
- OSINT (социальные сети, новости, сайты, форумы)
- поиск информации об уязвимостях ПО, системы, объекта
- Сбор данных для DeepFake (текст, видео, аудио)



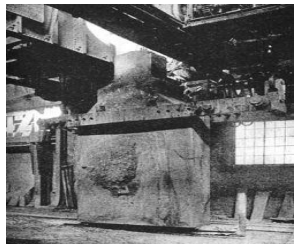
- анти-OSINT
- Поиск сведений об уязвимости (ПО, системы, объекта)
- Выявление опасного для предприятия или человека контента
- Распознавание информационной атаки на предприятие
- Распознавание злонамеренной активности (подготовка атаки APT группировки)

Моделирование (генерация) информации

Существенные технологические успехи, например, GPT-3 от OpenAI

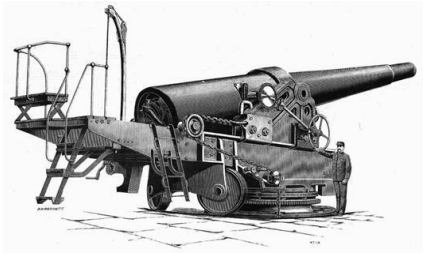


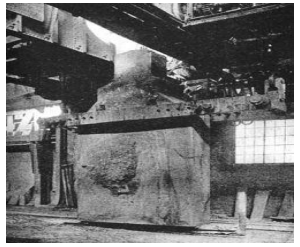
- Фейки (фишинг, сайты, DGA)
- Фейковые чат-боты
- Обфускация
- Создание «искажающего» контента
- Интеллектуальный подбор (пароль, URL)



- Honey-pots (ловушки)
- Обфускация
- Создание сложных паролей

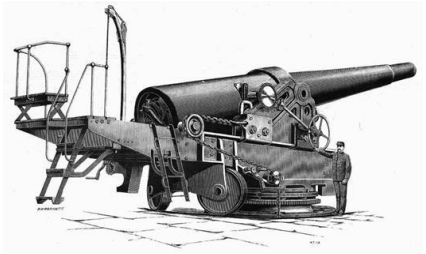
Анализ трафика



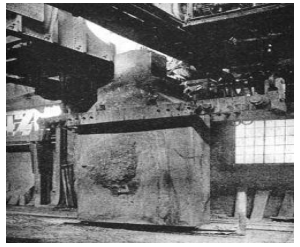


- OSINT
  - Распознавание структуры сети
  - Определение состава и версий программного обеспечения (подбор уязвимостей)
  - сотрудники, отделы, отношения, …
- Распознавание зловредов в трафике
- Реверс-инжениринг неизвестных протоколов (структура взаимодействия, C&C язык)
- Выявление автоматически сгенерированных доменных имен для “domain fluxing”, используемых ботнетами

Анализ и защита кода

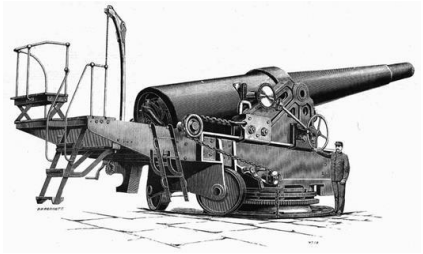


- Поиск информации об уязвимостях в открытых источниках
- Интеллектуальный анализ открытого кода (и, если доступен, закрытого)

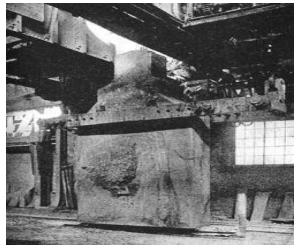


- Анализ кода, тестирование
- Рекомендации по повышению защищенности кода
- Автоматизация написания защищенного кода

Поведенческий анализ



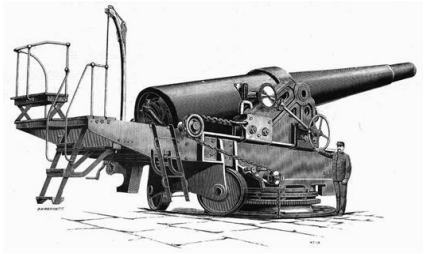
- Определение пользователей, склонных поддаться фишинговой атаке, формирование правил эффективной атаки
- Повышение эффективности фейкового чат-бота

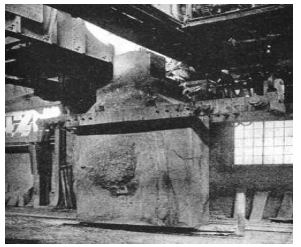


- Выявление злонамеренного или опасного поведения
- Дополнительная идентификация
- Анализ эмоций

Базы знаний, Threat Intelligence

- MITRE, CVE, CWE, CAPEC, БДУ ФСТЭК, …





- определение новых, неизвестных способов атаки (0-day),
- автоматизация атаки

- Парсинг источников (сайты, форумы, каналы в мессенджерах)
- Поиск информации об уязвимостях
- Мониторинг активности APT групп
- Автоматизация защиты (SOAR)
- Обогащение информации об инцидентах
- Ретроспективная корреляция (форензика)

Ссылки

- Хобсон Лейн, Ханнес Хапке, Коул Ховард «Обработка естественного языка в действии»
- https://medium.com/@ursachi/role-and-applications-of-nlp-in-cybersecurity-333d9280c737
- Переведенная на русский язык матрица MITRE от Positive Technologies https://mitre.ptsecurity.com/ru-RU/techniques
- Сравнение лучших практик по анализу рисков ИБ с методикой моделирования ФСТЭК https://sborisov.blogspot.com/2021/11/blog-post.html
- Introducing D3fend https://www.youtube.com/watch?v=7_DY1ULVapQ
- https://insights.sei.cmu.edu/blog/artificial-intelligence-in-practice-securing-your-code-using-natural-language-processing/