

Цель

Создать утилиту, позволяющее мониторить и категорировать утечки баз данных в сети Интернет.

Задачи

1. Отобрать список наиболее актуальных источников, в которых появляются сливы баз данных.
2. Написать сборщик информации об утечках из нескольких мест, позволяющий периодически обходить и собирать новые данные об утечках.
3. Дополнить сборщик классификатор утечек.

Требования

☐ Обязательные категории источников (минимум по 1 источнику в каждой):

1. Telegram-каналы
2. Форумы в даркнете

☐ Классификатор утечек должно происходить по двум критериям:

1. Тип утечки. Например, социальная сеть, форум, корпорация и тд. Чем больше категорий, тем лучше.
2. Стоимость утечки: платно + цена, бесплатно.

☐ Сохранение в БД данных об утечках:

1. ссылки на источник утечки (страница на форуме, сообщение в Telegram-канале)
2. ссылки на БД утечки либо ссылка на torrent-файл (при наличии)
3. torrent-файл/magnet-ссылка (при наличии)
4. размер утечки
5. класс утечки по источнику
6. стоимость утечки (при наличии)

☐ Интерфейс утилиты: не имеет значение (веб UI приветствуется)

☐ БД: PostgreSQL

☐ Добавить файл README.md с описанием:

1. алгоритма(ов) работы проекта
2. алгоритма запуска проекта
3. структуры коды проекта