

## Модуль 4

### Обеспечение качества и безопасности данных в системах машинного зрения

#### Формы

Решение кейсов

#### Диагностические инструменты

Внимание информационной безопасности при работе с данными — один из самых недооцененных аспектов работы инженера компьютерного зрения в сфере БАС. В этом задании Вам необходимо реализовать адекватные задаче меры безопасности при работе с данными.

**Кейс:** Вы работаете инженером машинного зрения в сфере БАС. Руководство Вашего подразделения попросило передать мини-датасет изображений и сопутствующую телеметрию субподрядчику для выполнения заказанных работ. Вы знаете что:

- для работы им нужны только некоторые данные в телеметрии — а именно данные по динамике полета (высота, углы...),
- у них система «заточена» под общий формат (стандарт телеметрии для БАС данного типа датасетов): строго определенное число столбцов и названий столбцов, при этом ячейки не должны быть пустыми (или содержать значения, интерпретируемые как пустые — пустые строки, например). Ваш файл телеметрии полностью соответствует стандарту,
- помните случай в прошлом году, когда утечка координат и времени полета из файла телеметрии БАС-оператора «Компании X» привела ее к серьезным убыткам,
- не знаете никого в организации субподрядчика и в первый раз с ними работаете...

Проявляя разумную осмотрительность на этапе передачи данных, Вы решаете (предварительно договорившись с субподрядчиком использовать защищенный протокол передачи данных, например SSH):

1. Зашифровать столбцы телеметрии с данными координат и времени полета.
2. Удалить все EXIF-метаданные изображений

**Для этого Вам надо:**

1. Прочитать датасет в pandas и зашифровать соответствующие данные, а также
2. Написать функцию, которая будет копировать изображения в другую директорию, но уже без метаданных.

#### Материал для выполнения:

Архив, содержащий:

1. Набор из нескольких изображений с метаданными «0001.jpg, 0002.jpg...»
2. Файл фрагмента телеметрии БАС «tele\_part.xlsx»

*Внимание! Не нужно загружать на Платформу изображения — они довольно объемные и могут превысить лимит 50Мб. Достаточно отправить jupyter notebook с решением и комментариями в нем по ходу работы.*

#### Критерии оценки:

1. Шифрование избранных данных в файле телеметрии (0-10 баллов):

- 10 баллов: скрипт выполняется, все необходимые данные зашифрованы криптостойким методом, который крайне трудно расшифровать. Скрипт подробно прокомментирован.
- 8-9 баллов: скрипт выполняется, все необходимые данные зашифрованы, но недостаточно криптостойким методом
- 6-7 баллов: скрипт выполняется, но все необходимые данные зашифрованы
- 4-5 баллов: скрипт выполняется, ошибки не носят принципиальный характер (опечатки в именах переменных, забытые переменные)
- < 4 баллов: скрипт не выполняется, в целом правильный — но имеет серьезные недоработки или не завершен, отсутствуют пошаговые комментарии

## 2. Удаление метаданных изображений:

- 10 баллов: скрипт функции, удаляющей метаданные работает как ожидается, задача реализуется полностью. Скрипт достаточно подробно и пошагово прокомментирован.
- 8-9 баллов: скрипт в целом работает, но не все метаданные удаляются
- 6-7 баллов: скрипт не отрабатывает как нужно, выдает ошибку, которая легко исправляется и не имеет принципиального характера (опечатки в именах переменных, забытые переменные)
- 4-5 баллов: скрипт не выполняется, хотя в целом написан, в скрипте не хватает существенных фрагментов или он не дописан. Метаданные не удаляются
- < 4 баллов: скрипт представляет разрозненные фрагменты и не может быть запущен с выполнением функции без значительных доработок