

Требования к данным (DATASET)

Данные, передаваемые органами государственной власти субъектов Российской Федерации, должны соответствовать требованиям и рекомендациям, предъявляемым к открытым государственным данным следующими нормативно-правовыми актами и организационно-распорядительными документами. Технические требования сформулированы в Методических рекомендациях по публикации открытых данных версии 3.0 (<https://data.gov.ru/metodicheskie-rekomendacii-po-publikacii-otkrytyh-dannyh-versiya-30>).

Передавая данные для использования в рамках конкурса, участник предоставляет неограниченному кругу лиц право использования данных в некоммерческих целях без ограничения по времени и территории действия.

Обработка данных осуществляется в статистических или иных исследовательских целях.

Координатор Конкурса имеет право использовать данные в дальнейшем по соглашению сторон.

Данные не должны содержать сведений, распространение которых ограничено законодательством Российской Федерации, в том числе:

- персональные данные граждан Российской Федерации или иных лиц;
- сведения, составляющие государственную тайну;
- сведения, составляющие коммерческую тайну;
- сведения, распространение которых может привести к нарушению прав физических или юридических лиц.

Набор данных может быть либо загружен в форму приема заявки на сайте asi.ru, либо опубликован органом власти в формате открытых данных с указанием гиперссылки (url) на паспорт набора данных.

Ответственность за нарушение прав третьих лиц вызванное публикацией данных, содержащих сведения, распространение которых ограничено законодательством Российской Федерации несет лицо, предоставляющее набор данных.

При загрузке набора данных через форму на сайте asi.ru набор данных должен представлять собой zip-архив, содержащий один или несколько файлов, содержащих данные. Рекомендуемым форматом описания набора данных является data-package (<https://frictionlessdata.io/specs>), однако к участию в конкурсе будут приняты и zip-архивы, не соответствующие спецификации data-package, но соответствующие рекомендациям, изложенным далее.

В качестве формата записи набора данных могут применяться: CSV, TSV, JSON, XML.

CSV «Comma-Separated List». Архив набора данных должен содержать:

– CSV файл машиночитаемой структуры данных, который должен представлять описание каждого информационного поля данных в следующих столбцах:

1. №
2. Наименование поля
3. Описание содержания поля
4. Формат значения поля

– CSV файл данных, соответствующий предоставленной структуре,
– (опционально) CSV файл с контрольными результатами. Содержит данные, на основе которых будет проверяться работа решающего задачу алгоритма.

TSV «Tab Separated Values». Архив набора данных должен содержать:

– TSV файл машиночитаемой структуры данных, который должен представлять описание каждого информационного поля данных в следующих столбцах:

1. №
2. Наименование поля
3. Описание содержания поля
4. Формат значения поля

– TSV файл данных, соответствующий предоставленной структуре.
– (опционально) TSV файл с контрольными результатами. Содержит данные, на основе которых будет проверяться работа решающего задачу алгоритма.

JSON «JavaScript Object Notation». Архив набора данных должен содержать:

– JSON Schema (<https://json-schema.org/>) файл машиночитаемой схемы данных, который должен представлять описание каждого информационного поля данных в формате JSON.

– JSON файл данных, соответствующий предоставленной схеме.
– (опционально) JSON файл с контрольными результатами. Содержит данные, на основе которых будет проверяться работа решающего задачу алгоритма.

XML «eXtensible Markup Language». Архив набора данных должен содержать:

– XSD (<https://www.w3.org/TR/xmlschema11-1/>) файл машиночитаемой схемы данных, который должен представлять описание каждого информационного поля данных в формате XML.

– XML файл данных, соответствующий предоставленной схеме.
– (опционально) XML файл с контрольными результатами. Содержит данные, на основе которых будет проверяться работа решающего задачу алгоритма.

Координатор Конкурса имеет право использовать данные в дальнейшем по соглашению сторон.

