

Анастасия Валеева: Что такое дата-журналистика?

Хотите узнать, что такое **дата-журналистика**, она же **журналистика данных**?

У каждого дата журналиста есть свой ответ на этот вопрос. А еще есть много мифов и заблуждений. Поэтому давайте сначала оттолкнемся от обратного. Дата-журналистика – это необязательно истории с большими данными или истории, наполненные цифрами; журналистика данных – это не всегда инфографика или что-то очень сложное.

Хотя, действительно, в журналистике данных есть и цифры, и инфографика, и даже могут быть использованы большие данные. А еще это действительно сложно.



Но гораздо важно другое – журналистика данных – это, в первую очередь, **журналистика, которая использует данные**.

Именно так и гласит одно из наиболее [распространенных определений](#): Дата-журналистика – это “сбор, очистка, организация, анализ, визуализация и публикация данных для создания журналистских материалов”. В этом определении перечислены основные шаги для создания дата истории. Но обо всем по порядку.

Этап №1. Сбор данных

Первый этап создания дата истории – сбор данных. Но как вы поймете, какие данные надо собирать? Тут есть два варианта: либо у вас изначально есть вопрос, ответ на который вы хотите найти в данных, либо к вам попадают данные, в которых вы хотите найти интересные моменты.

Вдохновение для первого способа может прийти откуда угодно. Например, из повестки дня. Вы открываете новостной сайт и читаете, что в вашем городе обещают построить новую школу. Вы можете превратить это в вопрос, ответ на который можно найти в данных. Например: а сколько сейчас в городе школ, сколько в них учеников и решит ли проблему переполненности одна новая школа? Значит, чтобы ответить на эти вопросы, вам нужно собрать данные по статистике среднего школьного образования.

А иногда данные попадают к вам сами. Допустим, Министерство образования и науки опубликовало данные о количестве выпускников высших учебных заведений по стране. Глядя на эти данные, вы можете задаться вопросом: каких профессий выпускают в стране больше всего? Это новый тренд или стабильная ситуация из года в год? Есть ли рынок для таких специалистов? Часть данных у вас уже есть, но возможно, вам понадобятся дополнительные данные, например, количество вакансий по специальностям.

Этап №2. Подготовка и анализ данных

Когда вы наконец собрали все данные, которые нужны для истории, вам, вероятно, не терпится приступить к их анализу и обработке. Но не стоит слишком торопиться, иначе можно сделать много ошибок, из-за которых работу потом придется начинать сначала.

Ключевые, хотя и самые трудоемкие процессы в дата журналистике – это очистка и организация данных. Как правило, эта работа делается в Microsoft Excel или других табличных процессорах. Обучиться этому можно [онлайн](#) или на тренингах по обработке данных. В результате, ваши данные должны стать “чистыми”, то

есть пригодными для анализа с помощью компьютера.

Как компьютер может анализировать данные? В первую очередь, это простые математические операции, такие, как подсчет суммы и среднего значения, вычисление доли или изменения показателя. Разница в том, что на компьютере это можно сделать гораздо быстрее, чем на калькуляторе, и произвести эти вычисления с огромным количеством показателей.

Но самое главное здесь именно журналистика; то есть то, как эти вычисления помогут вам ответить на вопросы для истории. Например, чтобы узнать, сколько в среднем в школах вашего города учеников в одном классе, надо сложить всех учеников вашего города и поделить на сумму всех классов.

А чтобы узнать, какой класс самый переполненный, надо посчитать количество учеников в каждом классе и отсортировать эти значения от максимального к минимальному.

Чем сложнее вычисления, тем более неожиданную историю вы можете создать. Например, используя методы статистического анализа, вы можете установить корреляцию или даже зависимость одного показателя от другого. Зависят ли результаты национального школьного теста от того, сколько учеников в классе?

Этап №3. Коммуникация данных

Теперь, когда вы закончили свой анализ и готовы рассказать об этом вашей аудитории, самое время подобрать правильные слова, изображения и героев для вашей истории.

О данных не нужно писать скучно, всю математику можно оставить за кадром, предоставив ссылку на ваши вычисления. А в самой истории описать, как то, что вы обнаружили в данных, проявляется в жизни и к каким последствиям может привести.

Подумайте, какие вам нужны эксперты и герои, чтобы проиллюстрировать проблему, которую вы хотите поднять в своем

материале. Может быть, вашим героем будет мальчик, который учился хуже всех в классе, пока родители не перевели его в менее переполненную школу?

Один из способов рассказать вашу историю – визуализация данных. При этом важно знать базовые принципы дизайна и уметь правильно подобрать тип графика под то, что вы хотите донести до аудитории. Классические графики, такие как столбчатая диаграмма для сравнения данных по категориям и линейный график для обозначения тренда – никогда не выходят из моды. А если вы хотите поэкспериментировать, то есть много [каталогов](#) и [онлайн ресурсов](#), которые помогут вам выбрать нужный тип графика.

Непосредственно для создания визуализаций также существует много онлайн инструментов, в том числе бесплатных. Новичкам рекомендуется начать с [Датараппера](#) (на сайте есть [блог с инструкциями](#)), [Инфограм](#) и, а продвинутым пользователям стоит познакомиться с [Флориш](#) и [Табло](#).

Советы:

- *не бойтесь цифр
- * организуйте работу по этапам, от идеи до визуализации
- * читайте новости и старайтесь превращать новостной повод в вопрос, ответ на который можно найти в данных
- * читайте аналитические отчеты и исследования и старайтесь понимать этот язык
- * изучите функции Microsoft Excel
- * если вы новичок в визуализации данных, попробуйте начать с сервиса “Datawrapper”