

## Конспект к уроку «Что такое данные и как с ними работать?»

Данный урок входит в серию уроков «Культура работы с данными» и является первым в ней. Урок можно провести обособленно или в комплексе с другими уроками серии.

**Цель урока:** сформировать представление обучающихся о данных, источниках данных и их роли в современном мире.

### Задачи урока:

- познакомить обучающихся с понятием «данные», их свойствами, видами, типами и форматами;
- выделить ключевые области применения и источники данных в окружающем мире
- обсудить риски и проблемы ненадежных источников данных; определить критерии поиска и проверки надежных источников данных;
- на практике познакомить обучающихся с примерами открытых надежных источников данных и способами их поиска и извлечения;
- совместно с обучающимися сформулировать ценность данных в современном мире для себя, общества, бизнеса и государства.

### Сценарный ход урока

**Педагог (слайд №1):** Добрый день! Сегодня мы будем говорить о данных, что это такое, где данные применяют и откуда их берут. Как вы понимаете термин данные?

**Обучающиеся отвечают на вопрос педагога.**

**Педагог (слайд №2):** В современном обществе данные используются повсеместно. Важность данных подчеркивается различными законами, правилами, стандартами, регламентирующими работу с ними на уровне государств и различных сфер деятельности. Самый первый Комитет по данным для науки и техники (англ. Committee on Data for Science and Technology; CODATA) создан в 1966 году. Он работает до сих пор. Его цель - сбор, критическая оценка, хранение и предоставление данных для задач науки и техники. <https://codata.org/>

**Педагог (слайд №3):** На уровне нашего государства принят стандарт **ГОСТ Р ИСО/МЭК Р С 20546-2021**. «Информационные технологии. Большие данные», основанный на международных стандартах ИСО (Международная организация по стандартизации). В РФ, согласно этому стандарту, Данные (data) — это представление информации в формальном виде, пригодном для передачи, интерпретации или обработки. Примечание - Данные могут быть обработаны автоматически или вручную. Мировое сообщество разрабатывает различные регламенты по работе с данными, наиболее известным является Свод знаний в области управления данными — DAMA-DMBOK. Этот документ создан Международной ассоциацией управления данными (Data Management Association International, DAMA), которая объединяет профессионалов в области управления данными. Первое издание «Руководства

DAMA к своду знаний по управлению данными» (Data Management Body of Knowledge, DMBOK) вышло в 2009 году. Имеется русская версия руководства.  
<https://dataliteracy.ru/dmbok> на русском языке.

**Педагог (слайд №4):** Кроме официально определённых терминов данные, в зависимости от сферы их использования встречаются и другие понятия. Например, в сфере информационных технологий приняты следующие определения:

**Данные** – числовые и нечисловые значения характеристик кого-либо (чего-либо), с которыми выполняет операции человек или компьютер. Данные, которые можно обрабатывать средствами вычислительной техникой называют машиночитаемыми.

**Данные** — это зафиксированные и сохранённые человеком или машиной (компьютером) факты или сведения (значения, показателя, характеристики в любом формате) о событии, явлении, процессе или объекте из окружающего мира.

**Данными** считают факты или наблюдения, неорганизованные и необработанные, не передающие никакого конкретного смысла и не имеющие ценности, потому что они лишены контекста и интерпретации.

**Педагог (слайд №5):** Как вы думаете, что тогда определяет ценность данных? Почему говорят, «что данные - новая нефть», а также стремятся собирать и хранить как можно больше данных? В 2006 году бизнесмен Клайв Хамби ввел метафору «Данные — это новая нефть», а вице-президент американской Ассоциации рекламодателей (Association of National Advertisers) Майкл Палмер развил эту мысль: «Данные похожи на сырьё. Оно полезно, но в необработанном виде непригодно для использования. Его необходимо преобразовать в газ, пластмассу, химикаты и тому подобные вещи...». Для получения из данных пользы (информации, знаний) их надо обработать (проанализировать и визуализировать), верно интерпретировать результаты анализа и далее предпринять то или иное действие (например, принять решение на основе данных).

*Например, по ТВ звучит прогноз погоды (то есть информация) на текущий день. Прогноз погоды составлен на основе метеонаблюдений (то есть данных), обработанных и проанализированных на компьютере. На основе прогноза погоды мы принимаем решения: берем зонт или шляпу от солнца, одеваемся теплее или обувает резиновые сапоги.*

Или, на сайте интернет-магазина (в его базе данных) хранятся данные о покупках (сырьё): наименование товара, количество проданных штук, регион покупателя, общая сумма покупки. Но если эти данные (сырьё) обработать - провести их анализ по разным параметрам, то можно получить полезную информацию, на основе которой грамотный специалист сделает выводы и примет эффективные управленческие решения, направленные на повышение продаж товара (что является основной задачей магазина). Так, информация о востребованности определённого товара в



регионе (например, термобелье в районе Севера и Сибири) позволит создать его запас на складе региона, что сократит срок доставки до потребителя и увеличит удовлетворенность покупателя от сделки.

Таким образом, информация — форматированные данные, обработанные и интерпретированные с определенной целью, которым придан смысл посредством добавления контекста. Обработанные данные являются информацией.

Данные и информация тесно переплетены и по отдельности не существуют. Данные — это форма информации, а информация — это форма данных.

### Обучающиеся отвечают на вопрос педагога.

Педагог: Приведите примеры данных из своей жизни.

**Педагог (слайд №6):** Работа с данными требует разносторонних знаний из различных областей науки и технологий. Познакомимся с некоторыми понятиями, которые применяют при работе с данными. Это классификация данных по разным основаниям. Рассмотрим классификацию данных по их назначению:

1. **Метаданные.** Это данные о данных (о их составе, содержании, статусе, происхождении, местонахождении, качестве, форматах, объёме, условиях доступа, авторских правах и т. п.). Например, при отправке по электронной почте сообщения, кроме самого текста (содержания) сообщение содержит дату и время отправки, указание отправителя и получателя, тип вложения, его объём и прочие характеристики. Это и есть метаданные.

*Вопрос педагога: А какие примеры метаданных вы можете привести?*

Любой файл содержит кроме данных и метаданные - сведения о самом файле – дата и время его создания или изменения, объем, описание, данные автора.

Каталог библиотеки – метаданные о данных – описание о книгах – автор, издательство, цена, год выпуска, тираж и т.д.

2. **Справочные данные** — данные из справочников, международных, общероссийских и отраслевых классификаторов и т. п.;

*Вопрос педагога: А какими справочниками данных вам приходилось пользоваться или вы знаете?*

Таблицы Брадиса, периодическая система химических элементов, телефонный справочник, справочник калорийности продуктов, справочник для поступающих в вузы, справочник лекарств.

3. **Основные данные** — данные об объектах, их характеристики.



*Вопрос педагога: Приведите примеры данных.*

Записи метеонаблюдений, чек в магазине, записи в телефоне или записной книжке, оценки в журнале, базы полиции, паспортного стола, мобильных операторов.

*Вопрос педагога: Приведите примеры того, как из данных получается информация?*

Запись в справочнике номера телефона (данные, набор цифр) трансформируется в информацию (анализируем и обрабатываем данные), если мы ищем номер нужного объекта, например гостиницы, чтобы позвонить туда для бронирования номера.

Анализ данных классного журнала (данные в виде цифр и символов фамилий) позволяет получить информацию об отличниках и хорошистах, а на основе этой информации директор школы принимает решение о поощрении их туристической поездкой.

**Обучающиеся отвечают на вопрос педагога, приводят примеры.**

**Педагог (слайд №7):** Молодцы! Вы все верно поняли и привели интересные примеры. Еще данные делят на:

**Структурированные** данные — данные, имеющие строго определённую структуру, определяемую формальной моделью данных (например, таблицы);

**Неструктурированные** данные — данные, произвольные по форме, не имеющие строго определённой структуры и не организованные по определённым правилам.

При работе с данными часто используют понятие датасет. Датасет (англ. dataset) — это обработанный и структурированный массив данных, пригодный для анализа, исследования и использования в различных вычислительных задачах. В основном, датасеты представляют собой таблицы или матрицы, где каждая строка соответствует отдельному наблюдению, а каждый столбец — признаку или атрибуту, описывающему это наблюдение. Пример датасета, таблица с записями данных о продажах интернет-магазина, где каждая покупка — это строки, а в столбцах указаны имена покупателей, наименование товара, цены и даты покупки. Датасет можно проанализировать и выявить спрос на товары, сезонный и суточный спрос на товары и другую информацию

**Педагог (слайд №8):** кроме вышеназванных данных в современном мире оперируют с понятием большие данные (Big Data). Этот термин связан с техническими аспектами сбора и обработки данных; он не предполагает конкретные виды данных. Большие данные — это разнообразные крупные и сложные наборы данных, поступающие с высокой скоростью. Размер этих наборов данных настолько велик, что традиционные программы для обработки не могут с ними справиться. Большие данные характеризуются основными свойствами:



**Большой объем (Volume)** данных представляет собой количество данных, доступных для анализа с целью извлечения полезной информации. Ключевым для развития технологий обработки больших данных стал рост объемов данных вследствие интернет-активности. Например, на YouTube загружается 300 часов видео каждую минуту, а объем мобильного трафика достиг 6,2 млрд гигабайт в месяц.

**Скорость обработки данных (Velocity)** — это скорость потока создания, хранения, анализа и визуализации данных. Быстрота изменений приводит к необходимости обработки большого количества данных за короткий промежуток времени. Так, Google обрабатывает 40 тысяч поисковых запросов в секунду — то есть 3,5 млрд запросов в день.

**Разнообразие источников и форматов данных (Variety)** - анализ разных типов структурированных и неструктурированных данных из различных предметных областей.

*Вопрос педагога: приведите примеры больших данных.*

К большим данным относятся аудио- и видеофайлы, изображения, данные постов в социальных сетях и другие текстовые форматы, данные о переходе по гиперссылкам, данные датчиков и сенсоров, умный дом, интернет вещей, умный город.

**Педагог (слайд №9):** Для хранения и обработки машиночитаемых данных используются специальные цифровые форматы.

**TSV (tab-separated values)** — это текстовый формат файла, в котором данные в столбцах разделены знаками табуляции, а строки — знаками перевода строки. Удобно работать с TSV-файлами в редакторах электронных таблиц или текстовых редакторах, таких как Блокнот или Notepad++ на Windows или TextEdit на Mac OS.

**CSV (Comma-Separated Values)** — это текстовый формат данных, в котором значения разделены запятыми. Он широко используется для обмена данными между различными приложениями и является одним из самых простых форматов для чтения и записи данных. Но могут быть проблемы с кодировкой и разделителями (могут быть использованы как точка так и запятая), пригоден только для таблиц.

**JSON (JavaScript Object Notation)** — это формат данных, основанный на синтаксисе JavaScript, он используется для представления структурированных данных. JSON является популярным форматом для передачи данных между клиентом и сервером в веб-приложениях.

**XML (eXtensible Markup Language)** — это язык разметки, используемый для описания структуры данных. Он может быть использован для хранения и обмена сложных данных и широко применяется в веб-службах и других системах.



**Педагог (слайд №10):** Также, данные делятся на открытые и закрытые. Закрытые данные недоступны широкому кругу пользователей в силу разных причин – государственная или коммерческая тайна, персональные данные. Открытые данные доступны всем. Открытые данные – это информация, размещенная в сети интернет в виде систематизированных данных, организованных в формате, обеспечивающем ее автоматическую обработку без предварительного изменения человеком, в целях неоднократного, свободного и бесплатного использования. Открытые данные свободно доступны для машиночитаемого использования и дальнейшей републикации (повторная публикация) без ограничений авторского права, патентов и других механизмов контроля.

Идея открытых данных зародилась во времена эпохи Просвещения, когда сформировалась концепция о вольном распространении научных знаний и об открытости для всех граждан публичной сферы деятельности, а закрепились в декларациях французской революции 18 века, о свободе печати и публичных бюджетов.

В современной истории открытые данные базируются на доктрине Открытого правительства, которая поддерживает право граждан на доступ к документам и действиям государства с целью эффективного общественного контроля за государственным регулированием. Россия присоединилась к доктрине Открытого правительства и в 2009 году был принят закон «Об обеспечении доступа к информации о деятельности государственных органов и органов местного самоуправления», который регламентирует порядок доступа граждан и организаций к информации о деятельности государственных органов и органов местного самоуправления.

**Педагог (слайд №11):** Мы уже знаем, что такое данные и в чём их отличие от информации. Знаем, как применять данные для получения информации. Познакомились с форматами и видами данных. А теперь давайте подумаем, где берут данные для обработки?

**Обучающиеся отвечают на вопрос педагога, приводят примеры.**

Специалисты многих профессий собирают и используют различные наборы данных для их анализа и принятия решений. Наборы статистических и научных данных могут понадобиться школьникам и студентам для проведения исследований по интересующему предмету. Частные лица, то есть мы с вами, можем использовать данные для своих потребностей, например искать книги, журналы, статьи в электронных библиотеках или пользоваться электронными справочниками, анализировать свои предпочтения, состояния, определять интересы, следить за здоровьем и физическими показателями, выбирать товары, планировать путешествия и многое другое.



## Основными надежными Источниками данных могут быть:

1. Я сам (человек). Каждый может фиксировать, накапливать достоверные данные о себе, например, вести личный дневник/блог, фотографии, видео, медицинские анализы, передвижения в пространстве, физические показатели (рост, вес, сила), успеваемость, денежные траты, эмоции и многое другое).
2. Государственные учреждения. Государство является основным сборщиком и агрегатором данных о жизни в стране и конкретной местности. Это и данные переписи населения, выборов, статистические и прочие данные министерств и ведомств.
3. Коммерческие компании и некоммерческие организации. Например, социальная сеть может предоставить данные о предпочтениях участников (какую музыку слушают, какие темы просматривают, на кого подписаны), их возрасте, регионе проживания.
4. Научные учреждения. Университеты и НИИ проводят исследования по различным направлениям науки, техники, технологий. Полученные данные публикуют для ознакомления и использования всеми заинтересованными лицами. Часто медицинские исследования становятся популярными источниками информации.
5. Независимые профессиональные сообщества и международные организации. ООН, ЮНЕСКО, ВОЗ хранят данные о мировых исследованиях в разных областях.

**Педагог (слайд №12):** Поскольку данные нужны для получения информации и принятия верных решений, то к качеству данных предъявляются особые требования. Качество данных (data quality) — обобщенное понятие, отражающее «степень пригодности информации для решения соответствующей задачи» (по ГОСТ Р ИСО 8000-2-2019). Чтобы понять какие данные качественные, то есть им можно доверять и использовать для анализа, а какие нет - познакомимся с критериями оценивания источников открытых данных.

1. **Владелец данных.** Самые достоверные данные размещаются на официальных порталах и сайтах. Эти информационные ресурсы принадлежат государственным органам и организациям, промышленным и технологическим корпорациям, компаниям, оказывающим услуги. Осторожно относимся к данным, автор которых не установлен или не является признанным экспертом в данной профессиональной области.
2. **Реквизиты данных (метаданные).** База данных снабжена метаданными, где указаны сведения об авторе или ответственном за размещение данных на портале, дата публикации данных или обновления, формат и объем файла с данными.
3. **Надежность места хранения данных и наличие точки входа.** Ссылка на доступ к данным (точка входа) размещается на портале или сайте. Если это неофициальный сайт источника данных, то оцените его на надежность: отсутствие рекламы на государственных сайтах или минимальная реклама на сайтах предприятий; отсутствие грамматических, синтаксических ошибок в текстах; указание на первоисточники данных.





4. **Доступность данных.** Любой желающий имеет возможность просмотреть и скачать файл с открытыми данными. Все действия с открытыми данными бесплатны для пользователя. Если за скачивание открытых данных просят провести оплату, то сайт не является официальным источником этих данных и им лучше не пользоваться.
5. **Наличие информации о методологии и способах сбора данных.** Сведения о сборе и обработке данных позволяют составить представление о качестве опубликованной информации. На официальных порталах кроме самых данных обязательно публикуют методологические пояснения о методах сбора и обработки данных: наименования методов сбора данных; охват обследуемой аудитории (респондентов); методы расчета и корректировки данных и многое другое. Например, на сайте Росстата опубликованы методология и нормативно-справочная информация о сборе и обработке всех представленных материалов, а также методологические пояснения по каждому виду данных.

**Педагог (слайд №13):** Сбор данных, или DAQ (Data Acquisition) представляет процесс оцифровки данных из окружающего нас аналогового мира, чтобы их можно было отображать, анализировать и хранить на ПК. Сбор первичных, необработанных (сырых) данных организуется различными инструментами. Важно помнить, что данные должны быть машиночитаемыми.

*Вопрос педагога: Как вы думаете, как собираются данные, какие инструменты используют?*

**Обучающиеся отвечают на вопрос педагога, приводят примеры.**

Методы сбора данных:

1. **Ручной сбор.** Данные (числовые, текстовые) вводятся с клавиатуры или сенсорного экрана сотрудником, ответственным за внесение значений. Это может быть менеджер банка, учреждения, организации, который вводит данные со слов клиента или переносит данные с бумажных носителей в цифровой формат. Кроме того, каждый из нас, заполняя поля данных при покупке товаров, билетов, бронировании отелей, регистрации на сайтах и интернет-сервисах, также вручную вносит свои данные. Метод подвержен ошибкам (человеческий фактор) из-за неверного ввода, а также трудоемок и затратен по времени.
2. **Автоматизированный или автоматический сбор.** К автоматизированному сбору относится оптическое распознавание символов (рукописных и печатных) и преобразование их в текст машинного формата, например, сканирование и распознавание бумажных документов, бланков, текстов. Автоматический сбор данных происходит без участия человека средствами программного обеспечения. Примеры таких систем - централизованные системы бронирования, интернет-магазины, платформы социальных сетей. Здесь для обмена данными используется API (Application programming interface, программный интерфейс приложения) — это





слой ПО, позволяющий программам взаимодействовать друг с другом. Для прямого доступа к своим данным большинство современных платформ раскрывают публичные или приватные API. Благодаря API системы могут автоматически собирать любой контент.

**Сбор данных с помощью сенсоров и датчиков.** Датчики используют в любом электромеханическом устройстве, которым пользуются люди – от автомобиля до медицинского оборудования. Благодаря датчикам возможно удаленное управление домами, транспортными средствами, рабочими пространствами и целыми городами (пока только отдельными городскими службами) с помощью обычного смартфона. Датчики не только собирают данные внешнего аналогового физического мира: температуру, напряжение, силу тока, деформацию и давление, расстояние и скорость, звук, массу, положение, изображение и т.п., но и преобразуют их в цифровой формат, пригодный для обработки, анализа, принятия решений. Собранные и накопленные большие массивы данных используют для анализа, моделирования, прогнозирования. Простым примером является процесс измерения температуры в помещении в виде цифрового значения с использованием аналогового датчика, такого, например, как термометр.

3. **Опросы и интервью.** Для сбора социологических данных (данных о совместной жизни групп и сообществ людей) применяют опросы (заполнение анкет) и интервью (личное общение) — методы сбора первичных данных, предусматривающие устное или письменное обращение с опрашиваемыми людьми. С их помощью можно получить данные о мотивах, интересах, планах, настроениях, предпочтениях людей, структуре отношений, семейной жизни. Проанализировав социологические данные, можно выявить взгляды и настроения граждан, составить полную картину общественного мнения по тому или иному вопросу. Сферы, в которых применяются соцопросы: политика, бизнес, образование, воспитание, благоустройство города и т.д. На основе опросов составляются рейтинги и принимаются управленческие решения, например, по благоустройству города или района.

*Ресурсы социологических данных:*

- ☐ Единый открытый архив экономических и социологических данных от Высшей школы экономики содержит 1700 социологических исследований и 200 временных рядов основных показателей российской экономики.  
<http://sophist.hse.ru>.
- ☐ Портал социологических данных РАНХиГС публикует открытую базу данных актуальных социологических и экономических исследований.  
<https://social.ranepa.ru/>
- ☐ Рейтинговое агентство RAEX («РАЭК-Аналитика») занимается подготовкой рейтингов в сфере образования, некоммерческого сектора, устойчивого развития, инфраструктуры развития бизнеса. <https://raex-rr.com>



**4. Статистические данные.** Статистические данные отражают массовые (количественные или качественные) данные по разным наименованиям и критериям, что позволяет анализировать количественные стороны массовых общественных явлений в числовой форме. Например, в каждом образовательном учреждении ведется статистика успеваемости обучающихся, на основе которой собирается статистика района, города, области, региона и страны. На основе государственных данных собираются данные в мировом масштабе.

Например, в образовании большинства стран популярна Международная программа по оценке образовательных достижений обучающихся (PISA (Programme for International Student Assessment), в рамках которой оцениваются знания и навыки обучающихся школ в возрасте 15 лет. В Российской Федерации национальным центром проведения исследования PISA является ФГБУ «Федеральный институт оценки качества образования». <https://fioco.ru/pisa>.

**5. Веб-скрейпинг или парсинг.** Веб-скрейпинг или парсинг – автоматический или ручной сбор данных с веб-страниц. Веб-скрейпинг (от англ. web scraping — соскабливание веба) — это автоматизированное извлечение данных с веб-страниц: цен, промоакций, авторских статей, описаний товаров, информации о клиентах и другого ценного контента, влияющего на коммерческий успех бизнеса. Парсинг (от глагола to parse — «разбирать») — процесс, где из скачанных данных извлекается нужная информация и превращается в читаемый формат.

**Педагог (слайд №14):** Данные являются основным сырьем для большинства процессов в экономике, науке, работе государства, особенно если эти процессы связаны с автоматической обработкой данных средствами вычислительных устройств. Статус данных подкрепляется и тем фактом, что для их сбора и изучения была выделена отдельная дисциплина - **наука о данных** (англ. data science; иногда даталогия — datalogy) — раздел информатики, изучающий проблемы анализа, обработки и представления данных в цифровой форме. Наука объединяет методы по обработке данных, в том числе больших данных, статистические методы, методы интеллектуального анализа данных и приложения искусственного интеллекта для работы с данными, методы проектирования и разработки баз и хранилищ данных.

### **Сферы применения Науки о данных:**

*Прогнозирование спроса.* В учреждениях и компаниях собираются большие массивы данных о товарах, услугах, клиентах за некоторый период времени. Анализ этих данных, особенно с помощью искусственного интеллекта, поможет найти закономерности, спрогнозировать будущий спрос и перестроить бизнес-процессы под нужное количество товаров и услуг.

*Система рекомендаций.* Интернет-сервисы хранят данные о просмотрах пользователями контента: видео, фильмов, музыки, статей или страниц товаров и

услуг. Проанализировав предпочтения клиентов, можно разработать предложение по контенту, наиболее соответствующее каждой целевой аудитории.

*Динамическое ценообразование.* В интернет-торговле и системах бронирования собираются и хранятся данные о продажах товаров и услуг разным категориям покупателей. Data Science анализирует эти данные и предлагает покупателям оптимальные цены на товары и услуги, что помогает увеличить выручку.

*Поиск аномалий.* Data Science помогает компаниям находить ошибки в данных бизнес-процессов и отчетности, вовремя заметить неточности и аномальные изменения данных, понять их причину и изменить работу компании.

Каждый из нас должен владеть культурой работы с данными, а именно - уметь собирать и обрабатывать данные; знать источники открытых данных и скачивать данные из них; анализировать данные при помощи специальных программ и математических функций; уметь визуализировать данные в виде динамических графиков и аналитического сторителлинга. Этим компетенциям мы и научимся на уроках серии «Культура работы с данными».

А сейчас предлагаю на практике поработать с данными и закрепить пройденный на уроке теоретический материал.

### **Практическая работа к уроку «Данные – новая нефть».**

#### **Практическая работа 1. Поиск данных в открытых источниках**

Задание 1.1. Поиск на портале открытых данных Министерства культуры РФ.

Задание 1.2. Поиск на портале открытых данных Министерство просвещения Российской Федерации.

Задание 1.3 Поиск открытых данных на портале Федеральной службы государственной статистики.

Задание 1.4 Поиск на портале открытых данных коммерческой компании Домклик.

#### **Практическая работа 2. Работы с данными в концепции Data Science (сбор, подготовка, обработка, анализ, коммуникация) на примере организации социологического опроса.**

Задание 2.1. Сбор первичных данных при помощи Яндекс или Гугл формы. Ручной ввод данных.

Задание 2.2. Проверка данных и их очистка от нерелевантных сведений.

Задание 2.3 Обработка данных. Классификация (разделение на выборки и группы), применение математических моделей.

Задание 2.4 Проведение качественного и количественного анализа, подтверждение или опровержение поставленных изначально гипотез.

Задание 2.5. Визуализация данных, создание аналитических отчетов и предложений для последующего принятия решений.



**Педагог:** Наше занятие подходит к концу. Надеюсь, вы смогли по-новому взглянуть на данные. Сегодня мы познакомились с понятиями данные и информация, на практике познакомились с источниками открытых данных и концепцией Data Science .

Давайте подведем итоги урока в форме синквейна. Выразите свои знания, эмоции, успехи от урока в виде синквейна.

Синквейн — это пятистрочные строфы.

1-я строка – одно ключевое слово, определяющее содержание синквейна;

2-я строка – два прилагательных, характеризующих данное понятие;

3-я строка – три глагола, обозначающих действие в рамках заданной темы;

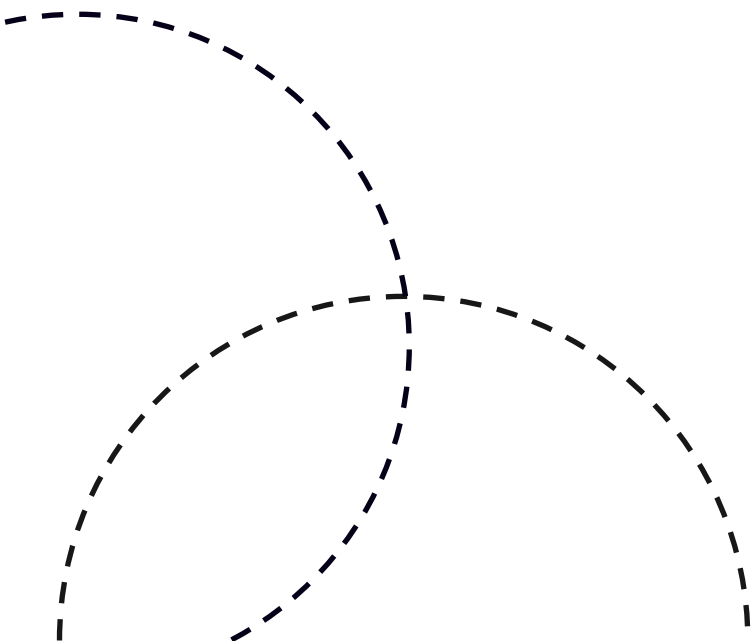
4-я строка – короткое предложение, раскрывающее суть темы или отношение к ней;

5-я строка – синоним ключевого слова (существительное).

Например:

- Данные
- Большие, разнообразные
- Ищем, обрабатываем, используем
- Данные – вторая нефть
- Сведения

**Всем спасибо за занятие!**





Вклад  
в будущее  
СБЕР



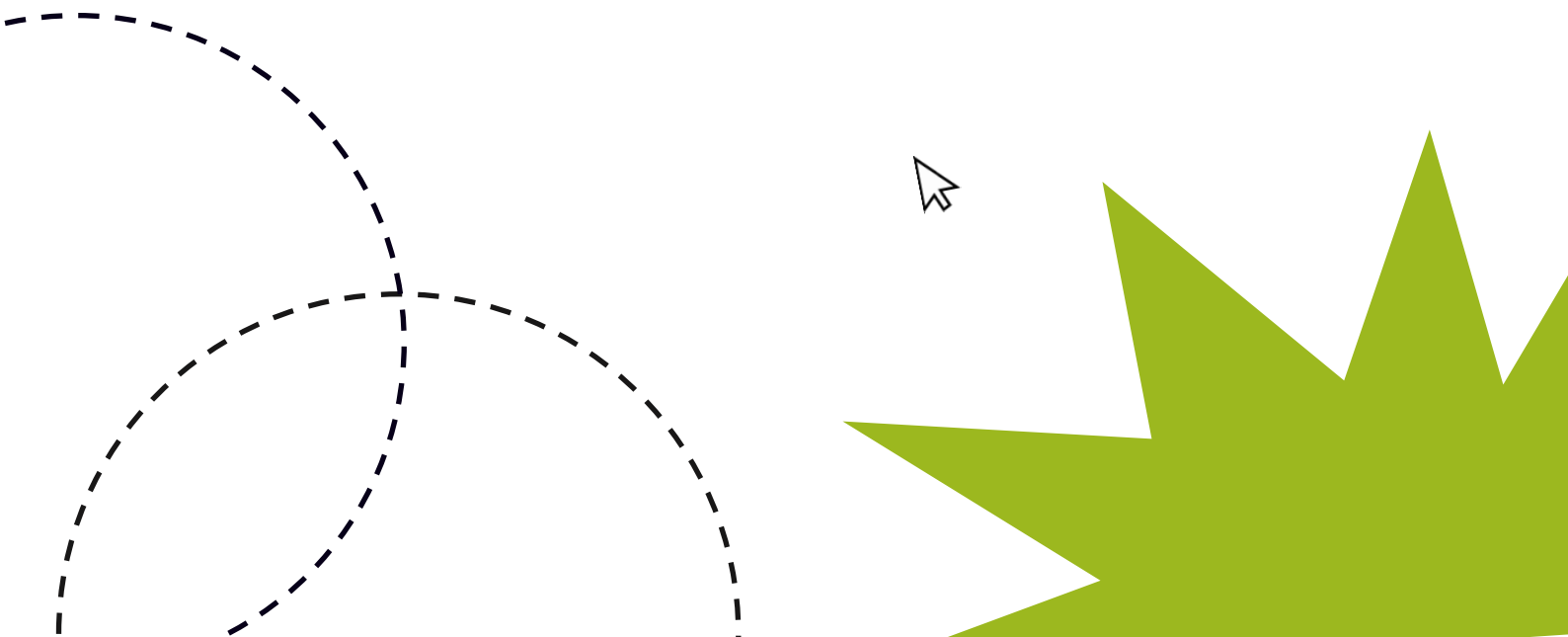
АКАДЕМИЯ  
искусственного интеллекта  
для школьников

**Урок подготовлен экспертами Академии искусственного интеллекта для школьников Благотворительного фонда Сбербанка «Вклад в будущее».**

Академия ИИ для школьников – Всероссийский образовательный проект, реализуемый Благотворительным фондом «Вклад в будущее» при поддержке Сбера с 2018 года. Цель проекта – формирование интереса и развитие прикладных навыков у школьников в сфере искусственного интеллекта. В рамках Академии ИИ проводятся просветительские активности, организуются соревнования, создаются образовательные курсы и вводные уроки, а также ведется работа с сообществом педагогов и амбассадоров Академии ИИ.

Все материалы опубликованы в открытом доступе на сайте Академии искусственного интеллекта для школьников и могут быть использованы для образовательных некоммерческих целей.

**Ссылка: <https://ai-academy.ru>**





## Используемые источники:

1. ПрофГид. ИТ-профессии. Режим доступа <https://www.profguide.io/professions/category/it/>
2. Росстат. Федеральная служба государственной статистики. <https://rosstat.gov.ru/statistic>
3. Доклад Центра подготовки руководителей и команд цифровой трансформации ВШГУ РАНХиГС <https://cdto.ranepa.ru/reports/data/9-1-bazovye-operacii-izmereniya>
4. «Руководство по управлению данными исследований. <http://lib-os.ru/issledovatelyam/rukovodstvo-po-upravleniyu-dannymi-issledovaniy/vvedenie/>
5. 40 источников для поиска статистики. <https://vc.ru/media/121834-bolshe-40-istochnikov-dlya-poiska-statistiki>
6. Классификация источников данных по типу данных (способу сбора информации). [https://www.hse.ru/org/hse/4432173/mathbase/db\\_datatype](https://www.hse.ru/org/hse/4432173/mathbase/db_datatype)
7. Каталог открытых данных <https://dorozhnij.com/opendata>
8. Рекомендации Совета по стратегиям цифрового правительства ОЭСР от 2014 года. [https://www.oecd-ilibrary.org/sites/059814a7-en/1/2/1/index.html?itemId=/content/publication/059814a7-en&\\_csp\\_=a527a7269ce984d41531dc563087a73c&itemI GO=oecd&itemContentType=book](https://www.oecd-ilibrary.org/sites/059814a7-en/1/2/1/index.html?itemId=/content/publication/059814a7-en&_csp_=a527a7269ce984d41531dc563087a73c&itemI GO=oecd&itemContentType=book)
9. Тихонов С. Типология данных. <https://education.yandex.ru/handbook/data-analysis/article/tipologiya-dannyh.-istochniki-dannyh>.
10. Руководства DAMA к своду знаний по управлению данными (Data Management Body of Knowledge, DMBOK), 2009 год. Версия на русском языке. <https://dataliteracy.ru/dmbok>
11. [Базы данных международных организаций](#)  
[Зарубежные базы данных](#)