

# Что такое исследование

Владислав Горбунов — Head of DS



# О чем поговорим

- Типичный подход построения ML-решений и проблемы
- Что такое исследование?
- Что такое научный метод, из чего он состоит?
- Что такое воспроизводимость?

# Типичный подход в проекте с DS

- Придумали крутую идею, которая позволит заработать миллионы долларов;
- Установим побольше GPU, поновее и подороже;
- Соберем данные из какого-либо бизнес-процесса (чаще всего плохого качества, мало данных);
- Докинем данные из Интернета (могут быть неточные и предвзятые);
- Поддаемся предвзятости подтверждения ([Склонность к подтверждению своей точки зрения](#)):
  - Изучаем собранные данные, чтобы сформировать гипотезу (т. е. какие  $X$ ,  $y$ , какие ML-алгоритмы использовать, подбираем метрики под используемый алгоритм);
  - Используем по существу те же данные, которые были получены при генерации гипотезы, для проверки нашей гипотезы;
  - Проверяем нашу гипотезу с помощью высокопроизводительного алгоритма обучения, который может хорошо соответствовать практически любому набору слабо коррелированных  $X$  и  $y$ ;
  - Изменяем нашу гипотезу до тех пор, пока наши результаты не станут «хорошими», например, в процессе подсматриваем в тестовую выборку, чтоб “подобрать” параметры;
- И, конечно же, не беспокоимся о воспроизведении; мы ведь точно всё сделали правильно и с первого раза всё получится.

При создании таких продуктов успех - ошибка выжившего.

Вы сжигаете много денег компании пока верите в идею и до последнего стараетесь сделать ее, потому что она должна же работать.

Люди увольняются, когда понимают что не могут решить задачу или решение постоянно ломается в продуктиве на новых данных.

Приходят новые люди и не понимают как поддерживать то что уже было сделано и начинают делать с самого начала...

Провал таких проектов на данных очень вероятен, т.к. DS становятся продуктом предвзятости подтверждения, а это нельзя называть наукой.

# У науки о данных есть проблемы с наукой

Большая часть науки о данных кажется похожей на науку карго-культа — фразой, придуманной Ричардом Фейнманом для описания деятельности, использующей атрибуты и внешний вид науки без понимания зачем эти атрибуты нужны.

Это практики, которые имитируют внешние признаки научной деятельности, но не следуют её фундаментальным принципам — критическому мышлению, стремлению к объективности, систематической проверке гипотез и качественному анализу данных.

## Почему так происходит?

Многие организации не зависят от DS в критически важной деятельности.

Они занимаются «наукой» о данных ради извлечения выгоды из ажиотажа вокруг этой технологии, независимо от того, работает она или нет.

Даже если зависят, то всегда могут найтись такие сотрудники, которые сведут свою деятельность к подобной, ради успеха в который верят. Или же ради инженерии....

# Monkey ML

Это неструктурированный, экспериментальный подход к разработке моделей, цель которого попробовать различные подходы, алгоритмы или наборы данных в надежде найти работающее решение, часто без тщательного планирования или глубокого понимания принципов, лежащих в основе этих методов

- **Подход проб и ошибок**, при котором разработчики экспериментируют с различными моделями, параметрами и техниками обработки данных, не полагаясь строго на теорию или предварительные исследования.
- **Использование случайного поиска** параметров или архитектур моделей вместо более систематических методов оптимизации.
- **Автоматизация выбора модели** - применение AutoML, для поиска оптимальных моделей и их параметров без необходимости глубокого понимания каждого подхода.
- **Параллельное тестирование** множества различных подходов и моделей, чтобы увидеть, что работает лучше всего для конкретной задачи.

Хотя такой подход может показаться неэффективным или расточительным из-за большого количества времени и ресурсов, потраченных на эксперименты, он может быть полезен в ситуациях, когда вы столкнулись с тупиком в генерации новых идей и гипотез. Этот метод также может способствовать обнаружению неожиданных решений или новаторских подходов к задачам машинного обучения.

Но если вы используете этот метод, не как просто полезный инструмент в генерации новых идей, а заменяете им полностью процесс, считая что случайный эксперимент — это всё доказательство правдоподобности и эффективности, что вам нужно, то вы скатываетесь в науку карго-культа

# К чему приводит такой подход в науке о данных

- Открытая база данных инцидентов с ИИ алгоритмами;
- Тестовый беспилотный автомобиль Uber, который сбил и убил женщину из Аризоны в 2018 году, имел недостатки в программном обеспечении, сообщил Национальный совет по безопасности транспорта, добавив, что автономные автомобили компании попали в 37 аварий за предыдущие 18 месяцев;
- Статья 2020 года о том, как недостаточная спецификация создает проблемы доверия к современному машинному обучению;
- Многие социальные события почти непредсказуемы, даже при наличии хороших данных и большого количества специалистов по обработке данных;
- ИИ по сути потерпел неудачу перед лицом COVID;

## Q: Разве все так плохо?

A: К сожалению, да. На нашей практике очень многие люди, называющие себя DS, DA или MLE подвержены таким практикам.

Q: Как можно попробовать уйти от этих проблем?

A: Изучить что такое исследования в науке и провести связи с наукой о данных.

Выделить значимые события и артефакты, с которыми нужно работать, чтобы увеличить шансы на успех проектов. Или хотя бы сжечь не так много денег для компании, пока вы будете проверять гипотезы и пытаться довести их до внедрения.

Q: А как же популярные типовые задачи, ведь NLP и CV продвинулись далеко вперед?

A: Все сильно зависит от бизнес-задачи, данных и гипотез. Не забываем что есть ошибка выжившего и даже monkey-MLщики могут выдать хороший результат, который даст бизнесу выгоду.

# Что же такое исследование?

**Научное исследование** (англ. Scientific research) — это творческая и систематическая работа, проводимая для увеличения знания, основанная на применении научного метода. Очень часто, говоря “исследование”, подразумевают именно “научное исследование”.

Научное исследование направлено на получение знаний в форме предположений о каком-то явлении, на основе которых делаются предсказания, проверяемые экспериментально, по результатам чего делается вывод: согласуются или противоречат экспериментальные данные прогнозам.

- Предположения со временем становятся более сильными, если мы продолжаем получать, согласующиеся с ними, результаты экспериментов;
- Предположения могут изменяться, когда появляются экспериментальные данные, противоречащие предположениям.
- В науке всегда допускается вероятность того, что предположение, даже если есть множество согласующихся с ним результатов экспериментов, может быть ложным, потому что, как только принимается решение, что истина найдена, со 100% уверенностью, то становится невозможным дальнейшее увеличение знания в этом вопросе. Ведь никакие новые данные не могут опровергнуть 100% истину.

Что еще можно почитать:

- [Types of scientific research. Canadian Institute for Knowledge Development;](#)
- [Scientific Inquiry in Social Work: Types of research. Pressbooks](#)

К сожалению отсутствует единая типизация научных исследований. Тем не менее, попробуем их классифицировать.



# Типы научных исследований - По назначению

- **Фундаментальное научное исследование** (англ. Basic research, pure research, fundamental research) — научное исследование, целью которого является улучшение научной теории для лучшего понимания природных или других явлений. [Basic research. Wikipedia](#);
- **Прикладное научное исследование** (англ. Applied research) — научное исследование, целью которого является разработка технологий или методов, которые можно использовать для вмешательства и изменения природных или других явлений с помощью научной теории. [Applied research. Wikipedia](#);

# Типы научных исследований - По типу данных

- **Количественное научное исследование (англ. Quantitative research)** — это стратегия научного исследования, которая фокусируется на сборе и анализе количественных данных. Количественные исследования применяются как в рамках индуктивного подхода для выявления причинно-следственных связей и статистических закономерностей, так и в рамках дедуктивного подхода для проверки гипотез. [Quantitative research](#);
- **Качественное научное исследование (англ. Qualitative research)** — это стратегия научного исследования, которая фокусируется на сборе и анализе качественных (нечисловых) данных. Качественные исследования ориентированные на исследование индивидов, понимание и объяснение субъективных аспектов их взаимодействия. [Qualitative research. Wikipedia](#);
- **Смешанное научное исследование (англ. Mixed research)** — это стратегия научного исследования, которая объединяет в себе методы из качественного научного исследования и количественного научного исследования. [Multimethodology. Wikipedia](#);

# Типы научных исследований - По цели

- **Разведочное научное исследование** (англ. Exploratory research, formulative research) — это предварительное научное исследование для выяснения точного характера проблемы, которую предстоит решить. Оно используется для учета существующих результатов и методов во время выполнения новых исследований, а также для определения приоритетов исследований, сбора данных и изучения предметной области. [Exploratory research. Wikipedia](#); [Difference Between Exploratory and Descriptive Research. Surbhi S](#);
- **Описательное научное исследование** (англ. Descriptive research) — это научное исследование для описания характеристик изучаемой популяции или явления. Оно не отвечает на вопросы о том, как/когда/почему возникли эти характеристики. Скорее, он затрагивает вопросы «что» или «каков» (каковы характеристики изучаемой популяции или явления?). Характеристики, используемые для описания явления или популяции, обычно представляют собой своего рода категориальную схему, также известную как описательные категории. [Descriptive research. Wikipedia](#); [Difference Between Exploratory and Descriptive Research. Surbhi S](#);
- **Причинное научное исследование** (англ. Causal research, explanatory research) — это научное исследование для выявления степени и характера причинно-следственных связей. [Causal research. Wikipedia](#); [Causal Research \(Explanatory research\). research-methodology.net](#);

# Научный метод

**Научный метод** — это итеративный процесс исследования, в котором элементы постоянно сменяют друг друга при появлении новой информации.

- Предмет исследования и его характеристики
- Гипотезы и теории
- Прогнозы
- Эксперименты

# Научный метод - Характеристики предмета исследования

Научный метод зависит от того, насколько тщательно и точно зафиксированы характеристики предмета исследования и собрана информация о нем. Характеристики содержат:

- **Наблюдения** - Подробная информация о наблюдаемом явлении, с максимальной фиксацией контекста наблюдения, строго сформулированный вопрос на основе зафиксированных результатов наблюдений и предыдущий опыт наблюдений за предметом исследования;
  - Если вы предполагаете, что наблюдаете некоторую зависимость, стоит также пронаблюдать, а не случается ли явление случайным образом, само по себе, без тех факторов, от которых вы думаете, что оно зависит. Для этого важен размер выборки, ибо чем она больше, тем меньше вероятность случайного совпадения наблюдений.
  - Часто можно наблюдать ошибки в работе связанные с тем, что недостаточно точно был сформулирован вопрос на основе наблюдений;
- **Определения** - Терминология, меры и метрики.
  - Научное определение термина иногда существенно отличается от его использования в естественном языке, поэтому для исключения разночтения стоит фиксировать используемую терминологию
  - Научные величины часто характеризуются своими единицами измерения и методиками оценки, которые позже можно описать в общепринятых терминах при передаче работы другим исследователям
- **Измерения**
  - Систематический, тщательный сбор измерений или подсчетов соответствующих величин часто является решающим различием между псевдонаукой и наукой;
  - Измерения в научной работе также обычно сопровождаются оценками их неопределенности. Неопределенность часто оценивается путем повторных измерений желаемой величины. Неопределенности также могут быть рассчитаны путем рассмотрения неопределенности отдельных используемых базовых величин.
  - По возможности стоит проводить измерения в контролируемых условиях. Если такая возможность отсутствует, стоит зафиксировать неконтролируемые параметры;

# Научный метод - Гипотезы

**Гипотеза** — это предполагаемое объяснение явления (наблюдения и измерения предмета) или аргументированное предложение о возможной корреляции между явлениями.

Гипотезы стоит рассматривать как интерпретации наблюдаемых явлений, то есть предположения, которые подвержены личному опыту и взглядам выдвигающего их человека.

Важно, формулировать гипотезу таким образом, чтоб её можно было не только экспериментально подтвердить, но и опровергнуть (принцип Фальсифицируемости).

**Фальсифицируемость** - принцип, который предложил философ Карл Поппер, и он гласит, что: любая научная гипотеза или теория в принципе должны быть опровержимы (если мы ошибаемся, должен быть способ эту ошибку уловить)

Критерий Поппера не требует, чтобы гипотеза была правдоподобной, она может быть очень фантастической. Всё, что требует критерий Поппера, это чтобы ваша гипотеза, в принципе, могла бы быть хоть как-то опровергнута.

Обычно гипотезы имеют форму математической модели. Также их можно сформулировать и как логические рассуждения.

- **Логическое рассуждение** — это умственная деятельность, направленная на строгое заключение. Оно происходит в форме умозаключений или аргументов, начиная с набора предпосылок и доводя до вывода, поддерживаемого этими посылками. Посылки и заключение являются утверждениями, то есть истинными или ложными утверждениями о предмете рассуждения. Вместе они образуют аргумент. Типы логических рассуждений:
  - **Дедуктивное рассуждение** строится от общих положений к частным случаям. Предлагает самую сильную поддержку: посылки обеспечивают вывод, а это означает, что вывод не может быть ложным, если все посылки верны. (Для недедуктивного логического рассуждения посылки делают вывод рационально убедительным, не гарантируя его истинность. Это часто понимается с точки зрения вероятности: посылки делают вывод более вероятным)
  - **Индуктивное рассуждение** — это форма обобщения, которая выводит универсальный закон из закономерности, обнаруженной во многих отдельных случаях.
  - **Абдуктивное рассуждение** обычно понимается как вывод от наблюдения к факту, объясняющему это наблюдение. Часто выражение «максимально правдоподобная интерпретация» используется как синоним.
  - **Аналогическое рассуждение** предполагает сравнение двух систем на предмет их сходства. Оно начинается с информации об одной системе и выводит информацию о другой системе на основе сходства между двумя системами.

# Научный метод - Работа с гипотезами

О чем важно знать во время работы с гипотезами:

- **Бритва Оккама** - эмпирическое правило для выбора среди группы равнозначных гипотез наиболее желательной.
  - Если существует несколько логически непротиворечивых объяснений какого-либо явления, объясняющих его одинаково хорошо, то следует, предпочитать самое простое из них при прочих равных условиях. Содержание принципа можно свести к следующему: не надо без необходимости вводить новые сущности, чтобы объяснить какое-то новое явление, если это явление можно исчерпывающе объяснить без них.
  - Бритва Оккама требует предпочесть простое объяснение только в том случае, если оно объясняет явление не менее точно, чем сложное, учитывая весь известный на текущий момент массив наблюдений, то есть если отсутствуют объективные основания для того, чтобы предпочесть более сложное объяснение простому.
- **Альтернативные гипотезы** ([Strong inference](#)) - Важно проверять и поддерживать несколько гипотез, чтобы свести к минимуму предвзятость подтверждения, которая может возникнуть в результате проверки и принятия лишь одной гипотезы.
  - Предвзятости подтверждения — тенденция человека искать и интерпретировать такую информацию или отдавать предпочтение такой информации, которая согласуется с его точкой зрения, убеждением или гипотезой. Эффект проявляется сильнее в отношении значимых вопросов и глубоко укоренившихся убеждений. Люди также склонны интерпретировать неоднозначные свидетельства таким образом, чтобы поддерживалась их точка зрения.
- **“Обращение к новизне”** (люди склонны предпочитать новые, свежие утверждения) - это может привести к поиску доказательств того, что новое истинно, даже если сейчас ещё мало доказательной базы в пользу этой гипотезы.

# Научный метод - Теории

Проверенные гипотезы становятся теориями или интегрируются в существующую теорию.

После проверки гипотезы не становятся истиной.

Научное знание построено на том, что мы выдвигаем предположения и проводим эксперименты, которые поддерживают существующую гипотезу или противоречат ей.

Даже если гипотеза стала теорией после сотни удачных проверок, достаточно одного примера, где предсказания теории будут расходиться с новыми результатами, чтобы модифицировать эту теорию, иногда полностью её изменив.

- Новые теории могут быть более всеобъемлющими, чем те, что им предшествовали, и, таким образом, быть в состоянии объяснить больше, чем предыдущие (большой объем наблюдений).
- Теории могут быть включены в состав других теорий. Таким образом, в определенных случаях независимые, несвязанные между собой научные теории могут быть связаны, увеличивая таким образом свою объяснительную и предсказательную силу.



# Научный метод - Прогнозы

- Любая полезная гипотеза позволит делать прогнозы посредством дедуктивных рассуждений. Она может предсказать результат эксперимента в лабораторных условиях или наблюдения за явлением в природе.
- Важно, чтобы результат проверки такого прогноза в настоящее время был неизвестен. Только в этом случае успешный результат увеличивает вероятность того, что гипотеза верна. Если результат уже известен, он называется следствием и уже должен был быть учтен при формулировании гипотезы.
- Если предсказания недоступны наблюдению или экспериментированию, это значит что гипотеза не поддается проверке и, следовательно, в строгом смысле слова остается ненаучной. Новая технология или теория могут сделать необходимые эксперименты возможными. Но пока стоит перейти к проверяемым гипотезам.

# Научный метод - Эксперименты

Как только предсказания сделаны, их можно проверять с помощью экспериментов.

- Если результаты испытаний противоречат предсказаниям, лежащие в их основе гипотезы подвергаются сомнению и становятся менее обоснованными.
- И наоборот, успешные предсказания не гарантируют истинности гипотезы, они лишь её поддерживают, делают её более вероятной, до появления новых экспериментальных данных, противоречащих гипотезе.

# Научный метод - Воспроизводимость экспериментов

Изначально термин “воспроизводимые исследования” применялся для обозначения прозрачности исследований в вычислительных науках. Таким образом характеризовалось исследование, выпущенное с использованием ПО, которое позволяло читателю статьи видеть весь процесс обработки данных, от сырых до различных рисунков и таблиц.

Важность повторяемых независимых исследований, которые подтверждают один и тот же результат стала признаваться во всех науках и термин “воспроизводимые исследования” стал расширяться.

По итогу началась путаница, так как под воспроизведением стали понимать два различающихся понятия:

- Возможность исследователя взять предыдущее исследование и с помощью предоставленных материалов (методов и данных) повторить результаты первоначального исследования;
  - Такая воспроизводимость не добавляет веса полученным результатам, а лишь показывает, что методы были описаны корректно и были предоставлены все необходимые материалы.
  - Обычно именно это обозначается термином *reproducibility*, но не всегда.
- Возможность исследователя повторить результаты предыдущего исследования, используя те же методы, но другие данные;
  - Таким образом появляются новые доказательства в пользу результатов предыдущего исследования, подкрепляя его как научное знание.
  - Это же стало обозначаться термином *replicability*, но тоже не всегда, иногда подменяя его на *reproducibility*.

Давайте посмотрим на разные варианты определения воспроизводимости, созданные научным сообществом

# Научный метод - Воспроизводимость экспериментов

Определения терминов от Ассоциации вычислительной техники ([Artifact Review and Badging. Association for Computing Machinery](#))

На данный момент это актуальная версия определения, но всего пару лет назад, авторы меняли местами воспроизводимость и реплицируемость.:

- **Повторяемость (англ. Repeatability)** - Same team, same experimental setup
  - Измерение может быть выполнено с заявленной точностью одной и той же командой, с использованием одной и той же процедуры измерения, одной и той же измерительной системы, в одних и тех же условиях эксплуатации, в одном и том же месте в ходе нескольких испытаний. Для вычислительных экспериментов это означает, что исследователь может достоверно повторить свои собственные вычисления.
- **Воспроизводимость (англ. Reproducibility)** - Different team, same experimental setup
  - Измерение может быть получено с заявленной точностью другой командой, используя ту же процедуру измерения, ту же измерительную систему, в тех же условиях эксплуатации, в том же или другом месте в ходе нескольких испытаний. Для вычислительных экспериментов это означает, что независимая группа может получить тот же результат, используя первоначальные артефакты.
- **Реплицируемость (англ. Replicability)** - Different team, different experimental setup
  - Измерение может быть получено с заявленной точностью другой командой, другой измерительной системой, в другом месте в ходе нескольких испытаний. Для вычислительных экспериментов это означает, что независимая группа может получить тот же результат, используя артефакты, которые они разрабатывают совершенно независимо.

# Научный метод - Воспроизводимость экспериментов

В работе [Steven N. Goodman. 2016. What does research reproducibility mean?](#) поднимается вопрос путаницы терминов и потому вводится альтернативная терминология, обозначающая 3 уровня воспроизводимости:

- **Воспроизводимость методов (англ. Methods reproducibility)**

- Воспроизводимость методов призвана уловить первоначальный смысл “воспроизводимости”, то есть способность максимально точно реализовать экспериментальные и вычислительные процедуры с использованием одних и тех же данных и инструментов для получения одинаковых результатов.

- **Воспроизводимость результатов (англ. Results reproducibility)**

- Воспроизводимость результатов относится к тому, что ранее описывалось как «репликация», то есть получение подтверждающих результатов в новом исследовании с использованием тех же экспериментальных методов.

- **Воспроизводимость выводов (англ. Inferential reproducibility)**

- Воспроизводимость выводов относится к получению качественно аналогичных выводов либо в результате независимого повторения исследования, либо в результате повторного анализа исходного исследования.
- Воспроизводимость выводов не идентична воспроизводимости результатов или воспроизводимости методов, поскольку ученые могут делать одни и те же выводы из разных наборов исследований и данных или могут делать разные выводы из одних и тех же исходных данных, иногда даже если они согласны в аналитических результатах.
- Этот аспект воспроизводимости, хотя и недооценен, но может быть самым важным. Несмотря на то, что задаются куда более понятные определения, разграничивающие явно 3 уровня воспроизводимости, в статье обращается внимание на то, что необходимо в рамках каждой области проработать методики обеспечения воспроизводимости методов, а также для каждого исследования задавать погрешность при воспроизведении результатов.

Воспроизводимость выводов же считается наивысшей формой поддержки результатов научного исследования, но не отменяет первые два уровня, ведь без уверенности в вычислениях, без уверенности в результатах, ещё меньше будет уверенности в интерпретации этих результатов, для формирования вывода.

# Научный метод - Полезные практики для экспериментов

- **Проведение “решающих экспериментов”** - в своей основе они противопоставляют несколько альтернативных гипотез, делая по итогу одну из них более вероятной, чем другие.
- **Метод контрольной группы** - с его помощью можно бороться с ошибкой наблюдения, сравнивая разницу между несколькими выборками, наблюдениями или популяциями в разных условиях, чтобы увидеть, что меняется, а что остается неизменным.
  - **Методы исследования причинных связей** могут помочь нам выяснить, какие факторы являются важными.
  - **Факторный анализ** в свою очередь позволит сократить количество переменных, отобрав наиболее важные.
- **Слепой метод** — процедура проведения исследования реакции людей на какое-либо воздействие, при котором испытуемые не посвящаются в важные детали проводимого исследования. Метод применяется для исключения субъективных факторов, которые могут повлиять на результат эксперимента.
  - **Двойной слепой метод** заключается в том, что не только испытуемые, но и экспериментаторы остаются в неведении о важных деталях эксперимента до его окончания. Двойной слепой метод исключает неосознанное влияние экспериментатора на испытуемого, а также субъективизм при оценке экспериментатором результатов эксперимента. При планировании эксперимента важно учитывать соотношение затрат и выгоды. Часто важной и более сложной задачей является не проведение экспериментов, а дизайн их таким образом, чтобы увеличить выгоду и сократить затраты на проведение.

# Научный метод - Научный метод - Рецензирование

Каждый элемент научного метода подлежит рецензированию на предмет возможных ошибок.

- Наука – это социальное занятие, и научная работа, как правило, принимается научным сообществом, если она подтверждена. Крайне важно, чтобы экспериментальные и теоретические результаты могли быть воспроизведены другими членами научного сообщества.
  - Если эксперимент не может быть повторен для получения тех же результатов, это означает, что первоначальные результаты могли быть ошибочными.  
В результате один эксперимент обычно проводится несколько раз, особенно когда имеются неконтролируемые переменные или другие признаки возможной экспериментальной ошибки.
- В процессе рецензирования эксперты, иногда анонимно, высказывают свою оценку определенным аспектам научного исследования с точки зрения элементов научного метода.
  - Желательно заранее иметь список доступных рецензентов, к которым можно обратиться за оценкой, особенно если область узкоспециализирована.
  - Рецензирование не подтверждает правильность результатов, а лишь то, что, по мнению рецензента, само исследование проведено корректно с точки зрения научного метода и современных практик на основании описания, предоставленного исследователем.
- Важно максимально фиксировать каждый элемент научного метода, каждый шаг пройденного исследования.  
Кроме того, что это важно для рецензирования, это позволит еще воспроизвести ваши результаты другим людям, что ещё больше поддержит ваши выводы в научном сообществе, а также позволит применить на практике ваши результаты.
  - Поскольку ученый не может записать все, что происходило в ходе эксперимента, сообщаются факты, отобранные по их очевидной значимости. Это неизбежно может привести к проблемам в дальнейшем, если вдруг что-то, что исследователь посчитал не важным, будет подвергнуто сомнению рецензентом или другими исследователями, которые попытаются воспроизвести результаты.
  - Поэтому хорошей практикой является избыточное описание и сохранение методов исследования, данных, инструментов. Всё равно будут ситуации, когда необходимо будет повторить исследование с учетом обратной связи от других исследователей, однако, таким образом можно очень снизить количество таких случаев.
  - Чтобы не случилось ситуации, когда вам показывают один удачный эксперимент, который просто отобрали из 100 проведенных экспериментов, где лишь один был удачным, есть принцип открытости исходных данных. То есть ученый должен описать не только один эксперимент и как он его провел, но и предоставить данные, а также указать а были ли другие попытки и эксперименты.

# Научный метод - Порождение новых знаний

Часто исследователи возвращаются к уже пройденным элементам, чтобы их обновить, когда появляется новая информация, чтобы улучшить качество исследования.

- Сложности с формулированием гипотез, могут привести к тому, что исследователь переосмыслит предмет исследования.
- Сложности с проверкой предсказаний полученных гипотез могут привести к пересмотру гипотезы или определения предмет исследования.
- Незначительные и неинтересные результаты эксперимента могут заставить исследователя пересмотреть экспериментальный метод, гипотезу или предмет исследования.

Другие исследователи могут начать собственные исследования и включиться в процесс на любом этапе.

- Они могут принять предмет исследования и сформулировать свою собственную гипотезу.
- Могут принять гипотезу и сделать свои собственные предсказания.
- Могут даже создать свой дизайн эксперимента для проверки выдвинутых гипотез.



# Научный метод - Научный метод - Кризис воспроизводимости

Усилия по оценке воспроизводимости научных публикаций дали тревожные результаты. Например, в 2015 году группа исследователей психологии, получившая название «Открытое научное сотрудничество», рассмотрела 100 экспериментов, опубликованных в высокорейтинговых рецензируемых журналах. Из этих 100 исследований только 68 репродукций дали статистически значимые результаты, совпадающие с исходными данными.

Согласно данным другого анализа, до 85 % всех проведённых в мире исследовательских работ в области биомедицины не привели к значимым результатам.

<https://www.science.org/doi/10.1126/science.aad7243>

# Научный метод - Научный метод - Кризис воспроизводимости

- **Выборочная публикация результатов (Publication Bias)**

Исследователи и журналы часто предпочитают публиковать только те исследования, которые показывают значимые или положительные результаты, игнорируя те, которые не обнаруживают значимых эффектов или подтверждают нулевую гипотезу. Это приводит к искажению общей картины в определенной области исследований.

- **P-hacking и манипуляции со статистикой**

Исследователи могут проводить множество статистических анализов или выбирать подходящие методы анализа данных до тех пор, пока не получат значимые результаты, что увеличивает вероятность ложноположительных выводов.

- **Недостатки в дизайне исследований**

Малые размеры выборок, отсутствие контрольных групп, неправильное использование статистических методов и другие методологические недостатки могут привести к ненадежным результатам.

- **Отсутствие доступа к данным**

Недостаток открытого доступа к полным наборам данных и к методологии исследований затрудняет проверку и повторение результатов другими учеными.

# Научный метод - Научный метод - Кризис воспроизводимости

- **Давление на публикацию**

“Publish or perish” — культура в академических кругах создает давление на исследователей, чтобы те публиковали статьи.

Это может привести к снижению качества исследований и увеличению количества спорных результатов.

- **Конфликт интересов**

Финансирование исследований коммерческими организациями может привести к смещению в отборе, интерпретации и публикации результатов в пользу финансирующей стороны. Многие коммерчески успешные результаты не публикуются.

- **Проблемы с рецензированием**

Система рецензирования может не всегда эффективно выявлять методологические недостатки и другие проблемы в представленных на публикацию исследованиях.

- **Отсутствие стандартизации**

В разных областях исследований могут отсутствовать единые стандарты проведения экспериментов и анализа данных, что затрудняет воспроизведение и сравнение результатов.

# Q: Разве все так плохо?

Уважаемое издание, тысячи цитирований статьи

Data Science Methodologies:  
Current Challenges and Future Approaches

Iñigo Martinez<sup>a,b,\*</sup>, Elisabeth Viles<sup>b,c</sup>, Igor G Olaizola<sup>a</sup>

<sup>a</sup>Basque Research and Technology Alliance (BRTA), Donostia-San Sebastián 20009, Spain  
<sup>b</sup>School of Engineering, University of Navarra, Donostia-San Sebastián 20018, Spain  
<sup>c</sup>Data Science and Artificial Intelligence, University of Navarra, Pamplona 31009, Spain

## 1. Introduction

In recent years, the field of data science has received increased attention and has employed great research efforts in developing advanced analytics, improving data models and cultivating new algorithms. The latest achievements in the field are a good reflection of such undertaking [1]. In fact, the data science research community is growing day by day, exploring new domains, creating new specialized and expert roles, and sprouting more and more branches of a flourishing tree called data science. Besides, the data science tree is not alone, as it is nourished by the neighboring fields of mathematics, statistics and computer science.

However, these recent technical achievements do not go hand in hand with their application to real data science projects. In 2019, VentureBeat [2] revealed that 87% of data science projects never make it into production and a NewVantage survey [3] reported that for 77% of businesses the adoption of big data and artificial intelligence (AI) initiatives continue to represent a big challenge. Also [4] reported that 80% of analytics insights will not deliver business outcomes through 2022 and

Цитирование в нескольких статьях



James Roberts · 3-й  
VP - Manufacturing Intelligence & Ad



CIO DIVE Deep Dive Opinion Library Events Press Releases  
IT Strategy Cloud Security

OPINION

## 4 reasons why most data science projects fail

Published May 2, 2017

By James Roberts in

[Publish](#)

*The following is a guest article from James Roberts, chief data officer at Quisitive.*

Experts have called 2017 the year of data literacy and digital transformation. While data is a key component that drives transformation, too often companies approach data and analytics projects the wrong way. In fact, a mere 13% of data and analytics projects reach completion, and of those that do, only 8% of companies report being completely satisfied with the outcome.