

Квантовая механика информационной приватности: Корпускулярно-волновой дуализм персональных данных и тензорная теория регулирования

Автор: **Алексей Мунтян**, генеральный директор ООО «Лекс Инжиниринг» ([Privacy Advocates](#)), соучредитель Общественного учреждения «Сообщество профессионалов в области информационной приватности» (RPPA.pro).

Актуальность материала: **02.02.2026** (с учетом стилистических правок и дополнительных пояснений на основе обратной связи читателей)

О чем вообще речь?

Представьте, что законы о персональных данных (все эти галочки «Я согласен») написаны для мира, где информация — это физический предмет, как кирпич или письмо в конверте. Если затереть на конверте имя (анонимизировать), то никто не узнает, чье оно. Это «ньютоновская механика» — старый, понятный мир.

Но мы живем в цифровом мире. Здесь данные ведут себя не как кирпичи, а как свет или волна в квантовой физике. Они могут быть в двух местах одновременно, менять свои свойства в зависимости от того, кто на них смотрит, и мгновенно связываться друг с другом.

Автор статьи описывает тензорную теорию регулирования обработки данных, которая предлагает перестать безуспешно обновлять аналоговые законы и мучить людей бесконечной дачей согласий на обработку их данных. Вместо этого должна применяться цифровая автоматика регулирования оборота данных, работающая по законам физики и математики:

1. Она взвешивает данные по «массе» (уникальность и чувствительность).
2. Смотрит на «скорость» их циркуляции и использования.
3. Сверяет различные «направления» (цели и контекст) использования данных.

Если риск (импульс) использования данных слишком велик или направление не совпадает — система просто не даст обработать данные. Это превращает защиту приватности из бюрократии в точную науку.

Заслуги автора более чем скромны с точки зрения новизны самого подхода к модернизации регулирования оборота персональных данных — большинство тезисов опирается на идеи, ранее сформулированные профессиональным и научным сообществом (как говорится, «стоя на плечах гигантов»). Вклад данной работы заключается в компиляции и структурировании разрозненных элементов в логичную, внутренне непротиворечивую концепцию. Читателям, заинтересованным в углубленном изучении темы, настоятельно рекомендую обратиться к разделу [«Первоисточники»](#). Сама идея тензорного регулирования родилась в ходе дискуссии с

коллегой из отраслевого контрольно-надзорного органа.

Автор приносит извинения тем, чьи взгляды на регулирование приватности могли войти в противоречие с положениями статьи, а также надеется на понимание новаторского характера работы. Буду признателен за конструктивную [обратную связь](#). На текущем этапе, в силу объективных причин, я не готов к развернутой дискуссии, однако данная публикация не ставит целью исчерпать тему. При наличии интереса со стороны аудитории планируется продолжение цикла статей, посвященных практической реализации тензорного регулирования.

Аннотация

Современное регулирование защиты прав субъектов персональных данных переживает фундаментальный эпистемологический кризис¹⁾. На протяжении последних сорока лет, начиная с Конвенции 108 Совета Европы, руководящие принципы и правовые рамки в области оборота персональных данных строились на концептуальном фундаменте, который можно сравнить с классической ньютоновской механикой. В этой картине мира данные рассматриваются как статические, дискретные объекты — «корпускулы» информации. Имя, идентификационный номер, адрес электронной почты воспринимаются как атомарные единицы, обладающие фиксированными свойствами «персональности» или «анонимности» независимо от контекста их наблюдения или скорости их движения в цифровой среде.

Однако аналоговая реальность XX трансформировалась в цифровую реальность XXI века: в эпоху больших данных, повсеместного машинного обучения и алгоритмического профилирования информация больше не ведет себя как инертная материя. Она демонстрирует свойства, напоминающие квантовые феномены: суперпозицию состояний, когда одни и те же данные могут быть одновременно приватными и публичными; квантовую запутанность, где разрозненные биты информации оказывают мгновенное влияние друг на друга на огромных расстояниях; и, что наиболее важно, корпускулярно-волновой дуализм, при котором природа данных меняется в зависимости от акта наблюдения (обработки).

Существующие регуляторные подходы, пытающиеся зафиксировать статус данных в момент их сбора через определение цели обработки, напоминают попытки измерить положение электрона без учета его импульса. Это приводит к системным сбоям: «анонимизированные» наборы данных внезапно повторно сопоставляются с субъектами данных (коллапс волновой функции), безобидные метаданные превращаются в инструменты слежки (эффект наблюдателя), а разделенные базы данных сливаются в единый профиль человека (мозаичный эффект).

В данной статье мы предпринимаем попытку преодолеть ограничения бинарного подхода, предлагая в качестве теоретической модели для обновления законодательства о персональных данных использовать концепт тензорного регулирования. Опираясь на гипотезу о корпускулярно-волновом дуализме и физической концепции импульса ($P=mv$)²⁾, мы постулируем необходимость введения понятия «тензор данных» (Т). Этот математический объект объединяет в себе «массу» данных (их состав и чувствительность) и «вектор» их движения (цель и скорость обработки). Мы утверждаем, что регулированию должна подлежать не статичная обработка персональных данных, а динамическое применение тензоров операторами, что позволит гармонично объединить корпускулярный (фокус на характеристиках и свойствах данных) и волновой (контекстуальная целостность обработки данных для защиты прав, свобод и интересов субъектов данных) подходы.

Корпускулярная теория и иллюзия фиксированной идентичности

Традиционный подход к регулированию персональных данных, доминирующий в законодательстве США и, отчасти, ЕС, базируется на материалистическом, или корпускулярном, представлении. В этой модели информация рассматривается как набор дискретных частиц — идентификаторов (атрибутов).

В США этот подход проявляется наиболее ярко в секторальном законодательстве (например, HIPAA в здравоохранении или VPPA в видеопрокате), которое применяется только при наличии строго определенных элементов («частиц») информации в процессе обработки данных. Например, HIPAA перечисляет 18 конкретных идентификаторов (имена, даты, номера социального страхования), удаление которых превращает создающие риски для субъектов персональные данные в относительно безопасные «обезличенные» данные³⁾. Это сродни утверждению, что если из атома убрать протон, то атом перестает быть атомом.

Европейский подход, закрепленный в Общем регламенте по защите данных (GDPR), хотя и является более гибким, вводя понятие идентифицируемости, все же остается привязанным к субстанции данных. Регуляторные механизмы — право на доступ, право на забвение, переносимость данных — оперируют данными как физическими объектами, которые можно найти, изъять, переместить или уничтожить. Этот подход предполагает, что свойство быть персональным присуще самой единице (биту) информации. Рассмотрим актуальную проблематику определения персональных данных в ЕС на двух примерах:

1. **Механизм вероятностного определения сведений в качестве персональных данных:** «дело Брейера»⁴⁾ — это судебное разбирательство между Патриком Брейером и властями ФРГ по поводу регистрации и хранения последней интернет-адреса («IP-адреса»), присвоенного г-ну Брейеру при доступе к нескольким интернет-сайтам, принадлежащим федеральным учреждениям Германии. В определении персональных данных отсутствует указание на связь между указанной в нем информацией и прямой или косвенной определенности или «определяемости» физического лица. Соответственно, отсутствует однозначное понимание того, в каких случаях данные будут относиться к персональным, а в каких — нет. Суд подтвердил, что должны существовать средства, которые «с разумной вероятностью могут быть использованы» для объединения элементов. В деле «Брейера» это был правовой механизм, с помощью которого правительство могло обязать интернет-провайдеров раскрывать информацию.
2. **Механизм предположительного определения сведений в качестве персональных данных:** «дело Главной комиссии по служебной этике в Литве»⁵⁾ — по мнению суда, обработка данных, которые косвенно могут раскрыть чувствительную информацию («специальные категории персональных данных») о физическом лице, не исключается из режима усиленной защиты, т.к. необходимо принять во внимание обработку не только изначально чувствительных данных, но и данных, раскрывающих информацию такого характера косвенно, после интеллектуальной операции, включающей дедукцию (аналитику) или перекрестные ссылки. Например, публикация имени супруга или партнера будет равнозначна обработке специальной категории данных, поскольку она может раскрыть сексуальную ориентацию.

Также можно отметить попытки философско-логического обоснования определения свойств персональных данных как уникальных и неизменных через совокупность тождеств **Idem**

(идентичность структурного порядка — **неизменность**) и **Ipse** (идентичность качественного порядка — **уникальность**):

- Тожество **Idem** является логической структурой, в рамках которой можно утверждать, что данный человек или предмет — тот же самый, что был год назад, вчера, тот, кто совершил некое действие и т.п. Этот тип идентичности (Дерек Парфит называет его numerical identity, «нумерической идентичностью») подразумевается неизменность во времени, пространственную идентификацию и, как следствие, количественную идентичность в смысле «тот же самый, а не другой» в рамках пары «тождественный — различный». Данного типа тождественность, описанная логическим языком с учетом временного оператора, выглядит следующим образом: предмет P1 в момент времени t1 является тем же самым, что и предмет P2 в момент времени t2 если и только если $K(P1(t1), P2(t2))$, где K — некоторый критерий диахронического тождества предметов P1 и P2 во времени.
- Тожество **Ipse** — это самость, в терминологии Парфита qualitative identity, «качественная идентичность». В этом аспекте два предмета являются тождественными, если все свойства и отношения, характеризующие один из них, характеризуют и другой, и наоборот. В ракурсе времени Ipse подразумевает, что у предмета или человека есть набор неизменных качеств, по которым можно его идентифицировать. Для образования этой идентичности принципиально, что в нее диалектически включается иное, то есть тождественность возникает относительно инаковости.

Корпускулярная теория терпит крах, когда сталкивается с реальностью высокой размерности данных. Основной ошибкой здесь является вера в то, что удаление явных (неизменных и уникальных) идентификаторов устраняет риски нарушения конфиденциальности/безопасности персональных данных. Это явление известно как ошибка статической анонимизации. История изобилует примерами, доказывающими несостоятельность этого подхода. Приведем два таких примера, демонстрирующих, что «персональность» данных — это не статическое свойство частицы, а динамическая вероятность, зависящая от контекста:

1. В 2006 году компания AOL опубликовала логи поисковых запросов 650 000 пользователей, заменив их имена на случайные числовые идентификаторы. С точки зрения корпускулярной теории, обработка данных была безопасна: идентифицирующие единицы информации (имена) были удалены. Однако содержание запросов — семантическое поле намерений, страхов и интересов — осталось. Журналисты The New York Times смогли идентифицировать пользователя № 4417749 как Тельму Арнольд, просто анализируя спектр ее запросов: «ландшафтные дизайнеры в Лилберне, Джорджия», «люди с фамилией Арнольд», «онемение пальцев». Совокупность этих запросов создала уникальный интерференционный узор (цифровой отпечаток), который мог принадлежать только одному человеку в мире.
2. В 2009 году Netflix опубликовал анонимизированные рейтинги фильмов для конкурса алгоритмов. Исследователи Нараянан и Шматиков продемонстрировали, что зная всего несколько оценок фильмов и приблизительные даты их просмотра (информация, доступная в публичных профилях IMDb), можно с высокой точностью идентифицировать пользователя в «анонимном» наборе. Здесь проявилась высокая размерность пространства данных: разреженность матрицы предпочтений делает каждый индивидуальный вектор поведения уникальным, как отпечаток пальца. Индивидуальные оценки фильмов сами по себе не были персональными данными, но их конфигурация (волна) однозначно указывала на личность.

В действующей парадигме право пытается наделить конкретный набор байтов (например, IP-адрес или cookie-файл) фиксированным «правовым зарядом» — является ли он персональным

сам по себе. Однако свойство персональности не имманентно (т.е. внутренне) присуще данным как физическому объекту, а носит вероятностный характер, зависящий от дополнительных знаний оператора. Попытка закрепить за «корпускулой» данных статичный статус создает правовую иллюзию: данные считаются анонимными (нейтральными частицами), пока не попадают в контекст, где они мгновенно коллапсируют⁶⁾ в состояние идентифицируемости (прямой или косвенной определяемости физического лица).

Кроме того, классическое право ошибочно пытается регулировать данные как отчуждаемые объекты, существующие отдельно от личности. В цифровой среде этот разрыв исчезает: информация становится не просто описанием, а цифровым продолжением субъекта. Необходимо отметить главенство контекста над текстом — одна и та же единица информации в разных социотехнических условиях (контекстах) меняет свою природу. Таким образом, механистическое представление о данных как о фиксированных, изолированных объектах регулирования игнорирует их динамическую связь с субъектом, превращая защиту информационной приватности в защиту пустых оболочек, а не самой личности.

Волновая теория и квантовая суперпозиция данных

Чтобы преодолеть ограничения корпускулярного подхода, необходимо рассмотреть волновую природу информации. Как и в физике, где свет демонстрирует свойства волны (дифракцию, интерференцию), персональные данные проявляют свойства, зависящие от среды распространения и взаимодействия с другими потоками данных.

Один из фундаментальных постулатов квантовой теории — принцип неопределенности Гейзенберга — гласит, что невозможно одновременно точно измерить координату и импульс частицы. Как только вы фиксируете, где именно в моменте находится объект наблюдения, вы перестаёте точно знать, куда и с какой скоростью он движется.

Квантовая теория это один из возможных инструментов для понимания процессов, происходящих в информационном обществе. Только теперь в социальной системе в роли квантов выступает информация о человеке — его персональные данные, — а эффект наблюдателя начинает работать лишь тогда, когда субъект понимает, что его персональные данные собирают, анализируют и используют.

Квантовый «эффект наблюдателя» в эксперименте Юнга: если наблюдателя нет, то электроны, проходя сразу через две щели, ведут себя как волны. Когда наблюдатель возникает и пытается определить, через какую именно из щелей пролетели электроны, то они начинают вести себя как частицы. Аналогично (хотя любая аналогия априори ложна) информация определяется в качестве персональных данных не абсолютно, а относительно наличия наблюдателя/оператора — как заинтересованного в таком определении лица.

Рис. 1. Эффект наблюдателя и эффект оператора [персональных данных]

В квантовой механике частица может находиться в суперпозиции состояний до момента измерения. Аналогично, элемент данных может находиться в суперпозиции состояний «приватное» и «публичное» (или «персональное» и «неперсональное») до тех пор, пока он не будет помещен в определенный контекст или не будет обработан определенным алгоритмом.

Рассмотрим пример с IP-адресом. Для интернет-провайдера, обладающего журналом привязки IP к абонентам, этот набор цифр является строго персональными данными. Для администратора веб-сайта, видящего лишь статистику посещений, это может быть анонимным техническим параметром. Статус данных не определен жестко; он коллапсирует в то или иное состояние в зависимости от наблюдателя.

Другой пример — GPS-координаты. Одиочная координата в пустыне — это просто геодезические данные. Но когда миллионы таких координат агрегируются в «тепловую карту» активности, они начинают интерферировать. В контексте городского парка эта интерференция создает «информационный шум» (популярный маршрут для бега). В контексте секретной военной базы, где бегают только солдаты, эта же интерференция создает четкий сигнал, раскрывающий расположение патрульных путей и инфраструктуры базы. Данные переходят из состояния «обезличенная статистика» в состояние «военная тайна» исключительно благодаря контекстуальной суперпозиции.

Для описания информационного поля, в котором распространяются информационные волны, идеально подходит теория контекстуальной целостности Хелен Ниссенбаум⁷⁾. Ниссенбаум утверждает, что информационная приватность — это не просто конфиденциальность данных о человеке, а соблюдение норм информационного потока, специфичных для конкретного контекста (здравоохранение, образование, семья). Нарушение инфоприватности происходит не тогда, когда данные раскрываются/компрометируются, а когда они перемещаются из одного контекста в другой с нарушением устоявшихся норм. Это аналогично переходу электрона на другую орбиталь с излучением энергии. Если врач передает диагноз другому врачу — это нормальный внутриорбитальный переход. Если врач передает диагноз маркетологу — это нарушение, вызывающее «излучение» риска. Волновой подход требует регулирования именно этих переходов, а не самих данных.

Импульс данных как оценка риска

Физика преодолела разрыв между корпускулярной и волновой теориями через введение понятий, работающих в обоих доменах, таких как импульс ($P=m \cdot v$). В квантовой механике импульс связан как с массой частицы, так и с длиной волны ($P=h/\lambda$)⁸⁾. Для преодоления неопределенности в регулировании той или иной обработки персональных данных с точки зрения сопутствующей меры риска нам необходимо ввести аналогичную величину — импульс данных.

В нашей модели «масса» данных (m) отражает их потенциальную энергию риска и социальную значимость:

1. **Уникальность**⁹⁾ — мера информационной насыщенности (энтропии или «неожиданности» информации), отвечающая на вопрос «Насколько легко найти вас в толпе?». Высокая энтропия (отпечаток пальца) означает высокую массу, так как уникально идентифицирует субъекта. Низкая энтропия (пол, город проживания) имеет малую массу.
2. **Чувствительность** — аналог плотности вещества. Медицинские данные, биометрия, политические взгляды обладают высокой плотностью («тяжелые» данные). Технические логи, метаданные обладают низкой плотностью («легкие» данные).
3. **Связанность** — оценка риска мозаичного эффекта, то есть способности разрозненных, казалось бы, независимых наборов данных складываться в единую картину, раскрывающую информацию, которой не было ни в одном из исходных наборов. Например, лайки в соцсетях могут быть глубоко связаны с интимными характеристиками личности: сексуальной ориентацией, политическими взглядами, уровнем интеллекта и употреблением психоактивных веществ. Алгоритм, анализирующий паттерн лайков, не просто угадывает эти характеристики; он измеряет скрытую переменную, которая детерминирована запутанностью поведения пользователя.

«Скорость» данных (v) в данном контексте — это интенсивность и направленность обработки.

1. **Долгосрочное (например, «холодное» или «ледяное»¹⁰⁾ хранение** — низкая скорость обработки данных. Данные лежат в архиве, доступ ограничен. Риск (импульс) минимален, даже если масса велика.
2. **Регулярное использование и аналитика** — средняя скорость обработки данных. Данные используются для внутренних процессов и отчетов.
3. **Обработка в режиме реального времени** — высокая скорость обработки данных. Данные о пользователе мгновенно транслируются сотням рекламных сетей за миллисекунды (Real-Time Bidding — RTB). Здесь даже данные малой массы (cookie ID) приобретают колоссальный разрушительный импульс из-за гигантской скорости распространения.

Представляется, что наиболее эффективным способом управления рисками обработки персональных данных является контроль не факта наличия данных (массу), а их импульс. Формула расчета риска обработки персональных данных: ($\text{Риск} = m_{\text{данные}} \cdot v_{\text{обработка}}$).

Далее рассмотрим несколько примеров применения формулы для лучшего понимания:

- *Хранение медицинских архивов на изолированном сервере* — большая масса (m), околонулевая скорость (v), что дает небольшой импульс, то есть значение (Риск) невелико.
- *Трансляция геопозиции в реальном времени* — средняя масса (m), большая скорость (v), что дает значительный импульс и значение (Риск) существенно.
- *Утечка биометрических данных* — большая масса (m), большая скорость (v), что дает огромный импульс и катастрофическое значение (Риск).

Рис. 2. Матрица расчета импульса данных

Тензорная теория регулирования

В статье 5 Федерального закона от 27.07.2006 № 152-ФЗ «О персональных данных» принцип ограничения целью сформулирован текстом: обработка персональных данных должна ограничиваться достижением конкретных, заранее определенных и законных целей; не допускается обработка персональных данных, несовместимая с целями сбора персональных данных. На практике это приводит к расплывчатым формулировкам в политиках обработки персональных данных и согласиях на обработку персональных данных («для улучшения качества услуг» и т.п.).

Мы предлагаем перейти от скалярного регулирования обработки данных к тензорному регулированию. Тензор — это геометрический объект, описывающий линейные отношения между векторами, скалярами и другими тензорами. Он инвариантен относительно смены системы координат, что идеально подходит для современного цифрового информационного пространства с разными особенностями и спецификами обработки данных.

Мы определяем тензор данных (Т) как внешнее произведение двух векторов — вектора цели обработки данных ($P \rightarrow$) и вектора состава данных ($D \rightarrow$). Иначе говоря, мы предлагаем регулировать не сами данные, а тензор данных — векторную пару, описывающую «куда» и «с чем» движется оператор: ($T = P \rightarrow \otimes D \rightarrow$).

В тензорной модели цель — это вектор в многомерном семантическом пространстве, а оси этого пространства могут определяться, например, социальными контекстами (по Ниссенбаум): здравоохранение, финансы, социальное взаимодействие, безопасность. Магнитуда вектора отражает важность цели (например, жизненно важные интересы субъекта имеют большую длину вектора, чем законный интерес оператора).

Статья 5 Федерального закона «О персональных данных» разрешает дальнейшую обработку персональных данных, если новая цель совместима с исходной, что можно проверить через тест на совместимость. В тензорной алгебре совместимость вычисляется математически как косинус угла¹¹ между вектором исходной цели ($P \rightarrow_{\text{исх}}$) и вектором новой цели ($P \rightarrow_{\text{нов}}$).

$$\cos(\theta) = \|P \rightarrow_{\text{исх}}\| \cdot \|P \rightarrow_{\text{нов}}\|$$

Вместо бинарного деления на персональные и «неперсональные» данные вектор данных ($D \rightarrow$) описывает свойства данных по нескольким измерениям, ранее уже упомянутым — идентифицируемость ($D \rightarrow_{\text{ид}}$), чувствительность ($D \rightarrow_{\text{чувст}}$), связанность ($D \rightarrow_{\text{связь}}$), — и рассчитывается как их совокупность.

Рассмотрим пример с базой данных пациентов. Если эти данные запрашивает университет для научного исследования, его вектор цели ($P \rightarrow_{\text{наука}}$) имеет высокую проекцию на ось медицинского контекста ($\cos \approx 0.8$). При этом данные агрегируются (снижается масса вектора ($D \rightarrow$)), и результирующая «энергия» тензора остается в зеленой зоне — доступ разрешен. В случае запроса тех же данных брокером для рекламы медицинской страховки, вектор цели ($P \rightarrow_{\text{реклама}}$) оказывается ортогонален медицинскому контексту ($\cos \approx 0$). В этой ситуации, даже если «масса» данных ($D \rightarrow$) невелика, несовпадение направлений создает критический «вращательный момент» (нарушение контекста), и алгоритм ограничивает операцию как рискованную для субъекта данных, либо в принципе блокирует такую операцию как публикацию данных пациентов ($\cos 0$).

Применение тензорного регулирования смещает фокус с контроля «данных на диске» на контроль применения тензора. Например, при попытке оператора использовать данные для новой цели (отличной от изначальной), автоматизированная система вычисляет новый тензор. Если новый вектор ($P \rightarrow \text{нов}$) выходит за пределы разрешенного сегмента, операция блокируется смарт-контрактом или архитектурой безопасности автоматизированной системы.

Рис. 3. Косинусное сходство целей

Заключение: от статики к динамике

Предложение ввести правовые субинституты¹²⁾ импульса данных и тензора данных в регулирование персональных данных — это не просто метафора, а необходимый шаг эволюции правовой мысли. Представляется нецелесообразным далее полагаться на статические списки идентификаторов и атрибутов в мире, где идентификация (определение лица) — это вероятностный процесс, а контекст определяет содержание.

Тензорное регулирование позволяет:

1. **Легализовать дуализм**, то есть признать, что персональные данные могут быть одновременно и информационным активом, и проекцией личности, в зависимости от приложенного тензора.
2. **Измерять неопределенность**, то есть использовать математические метрики

(энтропия, косинусное сходство) вместо субъективных суждений экспертов.

3. **Регулировать энергию, а не материю**, то есть сосредоточиться на риске обработки (импульсе данных), позволяя свободное движение данных там, где импульс мал, и создавая защитные барьеры там, где он разрушителен.

С практической точки зрения тензорная модель создает естественный барьер для утечек данных и минимизирует их последствия. Требование четкой векторизации целей исключает накопление избыточных данных «на всякий случай»: если вектор цели отсутствует или ортогонален вектору данных, тензор обнуляется, делая хранение информации нелегитимным и технически невозможным в архитектуре системы. Это обеспечивает математически гарантированную минимизацию поверхности атаки. Кроме того, применение формулы импульса данных принудительно снижает ущерб от потенциальных инцидентов. Поскольку система запрещает придавать высокую скорость обработки данным с большой массой, наиболее критичная информация (например, биометрия или медицинские данные) автоматически вытесняется в защищенные сегменты с низкой кинетической энергией, становясь недоступной для массового мгновенного хищения через высокоскоростные интерфейсы.

В мире квантовой неопределенности данных и цифровых технологий обработки информации, тензор данных — это инструмент, который возвращает наблюдателю (и правоприменителю) способность понимать и контролировать цифровую реальность, не разрушая ее. Будущее инфоприватности — это не бесконечные согласия на обработку персональных данных, а умная автоматическая система, которая измеряет риски по законам физики и математики, не позволяя нашим данным стать оружием против нас.

Что дальше?

Введение тензорного регулирования и проектирование алгоритмов автоматизированного регулирования оборота данных требует пересмотра не только законов, но разработку соответствующей методологии и стандартов. Как минимум, нам потребуется:

- **Стандарты векторизации целей** — отраслевые онтологии, сопоставляющие бизнес-процессы с векторами контекста ($P \rightarrow$).
- **Метрология данных** — сертифицированные инструменты для измерения свойств данных ($D \rightarrow$), своего рода аналог весов и линеек в торговле.
- **Контейнеризация правил** — данные должны быть упакованы в контейнер вместе с правилами их использования, «прилипающим» к данным намертво.

В последующих статьях мы постараемся более подробно рассмотреть и проанализировать каждый из вышеупомянутых элементов, а также описать возможные пути по их реализации. И в помощь нам будет разработанная в Privacy Advocates методология аудита и учета обработки персональных данных, включающая в себя [каталог из почти 170 групп и почти полутора тысяч атрибутов персональных данных](#).

В качестве наиболее амбициозной цели, автор статьи надеется пройти долгий путь:

1. теоретическая основа□
2. онтология□
3. методология□
4. дизайн алгоритма□

5. проектирование автоматизированной системы

Если есть желающие присоединиться к автору на этом пути, [напишите](#) — он будет очень рад единомышленникам.

Первоисточники

1. Solove, D. (2004). The Digital Person: Technology and Privacy in the Information Age.
2. Barbaro, M., & Zeller, T. (2006). A Face Is Exposed for AOL Searcher No. 4417749.
3. Narayanan, A., & Shmatikov, V. (2008). Robust De-anonymization of Large Sparse Datasets.
4. Nissenbaum, H. (2009). Privacy in Context: Technology, Policy, and the Integrity of Social Life.
5. Ohm, P. (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization.
6. Mc.Crory, D. (2010). Data Gravity — in the Clouds.
7. Kosinski, M., et al. (2013). Private traits and attributes are predictable from digital records.
8. Cosson, C. (2015). “Tool Without a Handle”: 21st Century Privacy — A Quantum Puzzle.
9. Purtova, N. (2018). The law of everything. Broad concept of personal data and future of EU data protection law.
10. Hern, A. (2018). Strava suggests military users “opt out” of heatmap.
11. Nissenbaum, H. (2019). Contextual Integrity: Up and Down the Data Food Chain.
12. Савельев, А. (2021). Научно-практический постатейный комментарий к Федеральному закону «О персональных данных».
13. Alexin, Z. (2022). Entropy based approach to personal data.
14. Ntelis, P. (2025). Advancing Tensor Theories.
15. Krantz, T. & Jonker, A. (2025). What is cosine similarity?
16. Дмитрик, Н. (2025). История и теория приватности.

1)

Эпистемологический кризис — это поломка механизмов познания и смыслообразования, разрушение способности превращать знания в понимание. Речь идёт не о нехватке знаний, а о разрушении способности превращать знания в понимание.

2)

В классической ньютоновской механике импульс электрона (скорость) определяется как $P=mv$, где (m) — масса электрона, (v) — скорость движения.

3)

Важно учитывать, что перечень из 18 идентификаторов HIPAA не является закрытым, так как последний, 18-й пункт, представляет собой открытую категорию, обязывающую удалять любые уникальные характеристики (коды, редкие признаки), которые могут прямо или косвенно указать на личность. Однако на практике эффективность такого подхода часто подвергается критике по причине формализма — организации часто ограничиваются удалением первых 17 конкретных пунктов, игнорируя 18-й, если связь данных с личностью не очевидна на первый взгляд.

4)

См. решение [CJEU, 19.10.2016, C-582/14](#) (Breyer).

5)

См. решение [CJEU, 01.08.2022, C-184/20](#) (Vyriausioji tarnybinės etikos komisija).

6)

Метафора, заимствованная из квантовой физики — *коллапс волновой функции* — мгновенное изменение описания квантового состояния (волновой функции) объекта, происходящее при измерении. Суть метафоры заключается в том, что в современном цифровом мире анонимность — это не постоянное свойство самих данных, а состояние их изоляции.

7)

Ниссенбаум выделяет 5 параметров потока данных:

1. Субъект (о ком данные).
2. Отправитель (кто передает).
3. Получатель (кто принимает).
4. Тип информации (атрибуты).
5. Принцип передачи (по согласию, по принуждению, за плату).

8)

Формула де Бройля ($P=h/\lambda$) связывает импульс частицы и длину её волны. В этой формуле (h) — постоянная Планка, (λ) — длина волны.

9)

Понятие уникальности (информационной энтропии) выражается через метрики конфиденциальности *k-anonymity*, *l-diversity*, *t-closeness*, позволяющими квантифицировать размытие данных: (1) *k-анонимность* оперирует эффектом толпы и гласит, что ваши данные должны выглядеть так же, как данные минимум еще ($k-1$) человек. Если ($k=5$), то в базе данных должно быть еще 4 человека с таким же полом, возрастом и индексом, как у вас. Злоумышленник видит запись, но не знает, кто из пятерых — вы. Вы спрятаны в группе; (2) *l-разнообразие* говорит о разнообразии секретов, так как одной толпы мало. Представьте, что вы спрятались в группе из 5 человек ($k=5$), но у всех пятерых в графе диагноз написано «грипп». Хакеру не нужно знать, кто именно вы. Он уже знает ваш диагноз, просто зная, что вы в этой группе. Поэтому внутри вашей группы секретные значения (диагнозы, зарплаты) должны быть разными (разнообразными). Если у всех одно и то же — защита не работает; (3) *t-близость* подразумевает защиту от догадок, ведь даже если диагнозы разные, но в вашей группе 90% людей болеют диабетом, а в среднем по больнице — только 5%, хакер может с высокой вероятностью угадать, что у вас диабет. Статистика внутри вашей маленькой группы должна быть близка к статистике во всем мире. Ваша группа не должна выглядеть подозрительно уникальной.

10)

Про горячее, холодное и ледяное хранение данных см. cloud.vk.com.

11)

Суть формулы: косинус угла между двумя векторами равен скалярному произведению этих векторов, делённому на произведение их длин. Это стандартная формула из линейной алгебры, используемая, в частности, для измерения сходства векторов (например, в векторной модели информации, машинном обучении). Пояснение символов: $\cos(\theta)$ — косинус угла θ между векторами; $P \rightarrow \text{исх}$ — исходный вектор; $P \rightarrow \text{нов}$ — новый вектор; $P \rightarrow \text{исх} \cdot P \rightarrow \text{нов}$ — скалярное произведение векторов; $\|P \rightarrow \text{исх}\|$ — длина (норма) вектора ($P \rightarrow \text{исх}$); $\|P \rightarrow \text{нов}\|$ — длина (норма) вектора ($P \rightarrow \text{нов}$).

12)

Субинститут в системе права — это упорядоченная совокупность правовых норм, регулирующих определённые особенности, специфику видовых общественных отношений в рамках крупного правового института.

From:

<https://www.wiki.privacy-advocates.ru/> - Privacy Advocates | Альманах

Permanent link:

https://www.wiki.privacy-advocates.ru/kvantovaja_infoprivatnost_teorija

Last update: 2026/10/07 11:12

