Observatoire ARGA

МЕЖДУНАРОДНЫЙ АНАЛИТИЧЕСКИЙ ДОКЛАД

Исследование №22

Анонимизация персональных данных в эпоху мощных AI-моделей

Риски повторной идентификации, связывание массивов и независимое доказательство эффективности

Организация: Observatoire ARGA

Русская версия

21 rue de l’Aviation, 64600 Anglet, France

info@argaobservatory.org | www.argaobservatory.org

22 октября 2026 года

Содержание

Резюме для руководства — 4

1. Предмет исследования, вопросы и пределы выводов — 5
2. Юридическая граница: персональные, псевдонимизированные и анонимные данные — 6
3. Проект EDPB 02/2026: новая архитектура оценки анонимности — 7
4. Контекстуальная оценка и перспектива получателя — 8
5. Критерий No Record Isolation: выделение записи — 9
6. Критерий No Linkage: связывание с внешними массивами — 9
7. Критерий No Inference: вывод чувствительных сведений — 10
8. Правовой режим самой операции анонимизации — 10
9. Особые категории, уголовно-процессуальные и санкционные сведения — 11
10. Судебная практика ЕС и проверка фактической идентифицируемости — 12
11. Модель угроз и возможности мотивированного нарушителя — 13
12. Прямые и косвенные идентификаторы: данные о событиях — 14
13. Геолокация, время и траектории перемещения — 15
14. Графы связей, текстовые эмбеддинги и векторные базы — 16

15. Подавление, генерализация и рандомизация: что работает и где ломается — 16
16. k-анонимность, l-разнообразие и t-близость: границы моделей — 17
17. Дифференциальная приватность: формальная гарантия и смысл параметров — 18
18. Учёт DP-бюджета, композиция и качество публикуемых результатов — 19
19. Синтетические данные и имитационные наборы — 19
20. PETs, федеративное обучение и защищённые вычисления — 20
21. Может ли обученная AI-модель быть анонимной — 21
22. Запоминание, регургитация и извлечение обучающих примеров — 22
23. Атаки на принадлежность записи к обучающему корпусу — 22
24. Model inversion и вывод атрибутов — 23
25. Организация испытаний: тестовый дизайн и независимость проверяющего — 24
26. Статистическая неопределённость и ошибочные выводы тестирования — 25
27. Повторная проверка после обновления данных и моделей — 26
28. Правила публикации: открытая, ограниченная и закрытая среды — 27
29. Документирование происхождения данных и цепочки преобразований — 27
30. Международная сопоставимость: ЕС, Великобритания и США — 28
31. Защита субъектов и реализация их прав — 29
32. Нарушения безопасности и риск несанкционированного раскрытия — 30
33. Применение в международном розыске и экстрадиционных делах — 31
34. Применение в санкционном и AML/KYC-комплаенсе — 32
35. Миграция, убежище и решения об уязвимых лицах — 33
36. Методика независимого аудита анонимности для Observatoire ARGA — 34
37. Открытая консультация EDPB: предложения Observatoire ARGA — 35
38. Пилотная программа проверки для международных исследовательских проектов — 36
39. Институциональные гарантии независимости и предотвращение конфликта интересов — 37
40. Ограничения исследования и условия применения выводов — 38
41. Выводы и институциональные рекомендации — 38
Приложение 1. Словарь ключевых терминов для междисциплинарной проверки — 39
Приложение 2. Дерево первичного юридического решения — 39
Приложение 3. Паспорт массива перед обработкой — 40
Приложение 4. Реестр правовых оснований и специальных ограничений — 40
Приложение 5. Матрица угроз и компетенций нарушителя — 41
Приложение 6. Карта вспомогательных данных и внешних источников — 41
Приложение 7. Рабочий протокол No Record Isolation — 41
Приложение 8. Протокол No Linkage и контроль ложных совпадений — 42
Приложение 9. Протокол No Inference для индивидуальных атрибутов — 42
Приложение 10. Числовой пример k-анонимности: две разные доли — 43
Приложение 11. Журнал бюджета дифференциальной приватности — 43
Приложение 12. Карта качества статистического результата с шумом — 43
Приложение 13. Проверка синтетического набора до передачи — 44
Приложение 14. Проверка AI-модели на воспроизведение обучающих фрагментов — 44
Приложение 15. Расчёт положительной прогностической ценности атаки membership inference — 44

Приложение 16. Интерпретация отрицательных испытаний — 45
Приложение 17. Контрольный лист векторных индексов и RAG — 45
Приложение 18. Оценка удалённых идентификаторов в свободном тексте — 46
Приложение 19. Оценка траекторий и географических данных — 46
Приложение 20. Контроль статей 9, 10, 11 и 25 GDPR — 46
Приложение 21. Паспорт AI-модели для вывода о персональном статусе — 47
Приложение 22. Due diligence внешнего поставщика анонимизации — 47
Приложение 23. Форма решения о выпуске обезличенного результата — 47
Приложение 24. Триггеры повторной проверки анонимности — 48
Приложение 25. Карта инцидента ошибочной анонимизации — 48
Приложение 26. Форма обращения лица о возможной идентификации — 49
Приложение 27. Учебный сценарий: международный розыск — 49
Приложение 28. Учебный сценарий: санкционный false positive — 49
Приложение 29. Учебный сценарий: отчёт об убежище и миграции — 50
Приложение 30. Проект замечаний для консультации EDPB 02/2026 — 50
Приложение 31. Хартия независимости аудитора — 51
Приложение 32. Контрольный список финальной редакционной проверки — 51
Приложение 33. Подробный кейс: малочисленная группа и раскрытие медицинского признака — 51
Приложение 34. Биометрические признаки, фотографии и голос — 52
Приложение 35. Многоязычные корпуса и ошибки извлечения идентификаторов — 53
Приложение 36. Композиция агрегированных публикаций и дифференциальное раскрытие — 54
Приложение 37. Анонимизация графа финансовых отношений — 55
Приложение 38. Условный пример различия перспектив контролёра и получателя — 55
Приложение 39. Контекстный риск моделей с открытыми весами и закрытым API — 56
Приложение 40. Удаление данных из модели и пределы machine unlearning — 57
Приложение 41. Протокол предварительной оценки пригодности открытой публикации — 57
Приложение 42. Протокол реконструкции связывания нескольких публикаций — 58
Приложение 43. Техническое испытание на извлечение данных из языковой модели — 59
Приложение 44. Membership inference: интерпретация точности на редком событии — 60
Приложение 45. Реестр рисков для векторных баз и retrieval-augmented generation — 61
Приложение 46. Сравнительный протокол для ЕС, Великобритании и США — 62
Приложение 47. Принцип минимального раскрытия при судебном и регуляторном споре — 62
Приложение 48. Испытание согласованности независимых оценщиков — 63
Приложение 49. План реагирования на повторную идентификацию после публикации — 64
Приложение 50. Договорная матрица для независимого тестирования анонимности — 65
Приложение 51. Управление дифференциальной приватностью в серии публикаций — 66
Приложение 52. Оценка малых выборок и вторичного подавления — 66
Приложение 53. Управление исходными доказательствами и научной воспроизводимостью — 67

Приложение 54. Система порогов эскалации и повторной проверки — 68
Приложение 55. Обращения лиц и независимое рассмотрение спорных выводов — 68
Приложение 56. Методика оценки качества и полезности анонимного результата — 69
Приложение 57. Комплексный сценарий: многоязычный судебный корпус — 70
Приложение 58. Комплексный сценарий: серия ежемесячных статистических бюллетеней — 71
Приложение 59. Комплексный сценарий: синтетическая база банковских транзакций — 71
Приложение 60. Комплексный сценарий: фото, аудио и метаданные — 72
Приложение 61. Комплексный сценарий: научный доступ к редким медицинским и миграционным событиям — 73
Приложение 62. Комплексный сценарий: подрядчик, который обещает «100% анонимность» — 73
Приложение 63. Комплексный сценарий: повторное использование результата для обучения AI — 74
Приложение 64. Итоговая карта внедрения методики ARGA на 12 месяцев — 75
Источники и методические основания — 76

Статус: исследовательский доклад для внутренней экспертной проверки. Проект EDPB Guidelines 02/2026 рассматривается как документ публичной консультации, а не окончательное руководство. Все числовые сценарии обозначены как условные.

Резюме для руководства

Ключевой вывод настоящего исследования: анонимность является контекстным, проверяемым и пересматриваемым правовым и техническим заключением, а не названием процедуры удаления персональных идентификаторов. Ни k-анонимность, ни дифференциальная приватность, ни синтетическая генерация по отдельности не заменяют проверки разумно вероятных способов выделения, связывания и вывода данных о человеке. Это особенно существенно для крупных языковых моделей, векторных индексов, судебных нарративов и малочисленных групп. [1][2][3]

По состоянию на 22 октября 2026 года EDPB Guidelines 02/2026 on Anonymisation остаются проектом на публичной консультации, завершающейся 30 октября. Основой обязательных требований остаётся GDPR и применимое национальное право; проект EDPB используется для методического анализа, но не как уже вступившая в силу новая норма. Обработка персональных данных в ходе самой анонимизации требует самостоятельной правовой оценки. [1][2]

Для Observatoire ARGA предложена проверка в шесть этапов: определение цели и адресата; инвентаризация исходных сведений; построение модели угроз; техническая и правовая оценка; независимая рецензия; документированное решение о выпуске с триггерами повторной проверки. Отрицательный вывод, когда открытая публикация недопустима, должен считаться полноценным профессиональным результатом, а не неудачей исследования.

Приоритетными красными флагами выступают уникальная последовательность процессуальных событий, возможность связывания с публичными судебными архивами,

вторичное восстановление подавленных статистических ячеек, сохранение данных в RAG-индексах и неподтверждённые заявления разработчика о том, что модель полностью анонимна. Во всех таких случаях вывод ограничивается фактическим контекстом и доступными доказательствами.

Практический маршрут

Для руководства рекомендуется утвердить роли владельца данных, независимого технического рецензента и юридического эксперта; использовать версии документов с контролируемым доступом; не публиковать реальные индивидуальные кейсы только ради иллюстрации методики; регулярно проверять изменения внешних источников; документировать отказ от выпуска. Настоящий доклад содержит модели процедур и учебные сценарии, но не претендует на результаты реальных испытаний клиентов или действующую аккредитацию исследовательской структуры.

1. Предмет исследования, вопросы и пределы выводов

Анонимизация персональных данных стала одним из основных условий ответственного использования больших массивов сведений для статистики, научных исследований, обмена между организациями и обучения систем искусственного интеллекта. Однако современная модель ИИ способна сопоставлять разнородные признаки, выводить вероятные атрибуты человека и воспроизводить фрагменты обучающего материала. Поэтому вопрос уже не сводится к тому, удалены ли из таблицы имя, номер паспорта и адрес. Необходимо установить, остаётся ли физическое лицо идентифицируемым средствами, которые разумно могут быть использованы конкретными участниками обработки. [1][2][4]

Настоящее исследование отвечает на четыре взаимосвязанных вопроса: когда сведения могут быть признаны анонимными в юридическом смысле; как оценить остаточные риски выделения, связывания и логического вывода; какие доказательства должен представить разработчик или владелец массива; как меняется оценка при переходе от обезличенной таблицы к обученной модели, векторному индексу, синтетическому набору или опубликованной статистике. Анализ имеет нормативно-методологический характер: приведённые примеры являются условными моделями проверки, а не результатами испытаний реальных организаций.

Работа подготовлена по состоянию на 22 октября 2026 года. Особое значение имеет проект EDPB Guidelines 02/2026 on Anonymisation, открытый для общественного обсуждения до 30 октября 2026 года. Он представляет интерпретационную позицию Европейского совета по защите данных, но пока не является окончательной версией руководства и не изменяет текст GDPR. Там, где доклад предлагает собственные пороги, формы или процедуры, они прямо обозначены как рекомендации Observatoire ARGA, а не как уже действующие юридические обязанности. [1][2]

Практический результат исследования представляет собой проект методики независимой оценки анонимности, набор контрольных вопросов, матриц, протоколов испытаний и процедур раскрытия результатов. Особое внимание уделяется чувствительным контекстам ARGA: международному розыску, экстрадиции, санкционному и финансовому комплаенсу, миграционным процедурам и публичной отчётности. Для этих областей ошибки связывания записей и раскрытия даже небольшого количества сведений могут иметь непропорционально тяжёлые последствия.

2. Юридическая граница: персональные, псевдонимизированные и анонимные данные

Статья 4(1) GDPR охватывает информацию, относящуюся к идентифицированному или идентифицируемому физическому лицу. Статья 4(5) определяет псевдонимизацию как обработку, после которой отнесение данных к конкретному субъекту невозможно без дополнительной информации, хранимой отдельно с соответствующими мерами защиты. Псевдонимизация остаётся способом обработки персональных данных и сама по себе не выводит их за пределы GDPR. Анонимной является информация, которая не относится к идентифицированному или идентифицируемому лицу с учётом применимых условий и средств идентификации. [1][10][14]

Преамбула 26 GDPR требует принимать во внимание все средства, которые разумно могут быть использованы контролёром или другим лицом, включая затраты, время, доступные технологии и технологическое развитие. Следовательно, юридический вопрос не тождествен математической невозможности идентификации при любых мыслимых обстоятельствах. Вместе с тем формальное удаление идентификаторов, подмена имени кодом и наличие договора о конфиденциальности также не являются достаточным доказательством анонимности. [1][3]

Практическая классификация должна охватывать не только «датасет как вещь», но и точку зрения получателя, момент передачи, доступные ему вспомогательные данные, правила доступа и последствия для человека. Доступная публично таблица, закрытый статистический результат и массив, раскрываемый одному доверенному исследователю, имеют разные модели угроз. Нельзя без нового анализа переносить заключение об анонимности из одной модели публикации в другую.

Категория	Юридический вывод	Ключевая проверка
Идентифицируемые сведения	Персональные данные; применяется GDPR при наличии его сферы действия	Прямая или косвенная идентификация
Псевдонимизированные сведения	Как правило, остаются персональными в контексте контролёра	Доступность ключа, связки и дополнительных сведений
Анонимные сведения	Вне предмета GDPR как сведения о субъекте, не являющемся идентифицируемым	Средства разумно вероятной идентификации для релевантных участников
Смесь данных	К персональной части применяются требования GDPR	Изоляция компонентов и невозможность обратного связывания

Дополнительная экспертная оценка

Практический спор обычно начинается с некорректного вопроса: «может ли сам владелец узнать имя человека?» Такой вопрос слишком узок. Значение имеет возможность идентификации не только владельцем, но и иным релевантным участником с учётом всех средств, использование которых разумно вероятно в конкретном контексте. Необходимо анализировать как данные, так и контролируемый выпуск: доступен ли получателю отдельный ключ, сопоставимые коммерческие реестры, журналы событий, поисковые системы и локальные сведения, которых нет у исходного контролёра. Проверяемость вывода требует открыто указать, чья перспектива рассмотрена, какие средства ей доступны и почему остальные средства признаны неразумно вероятными. [1][2]

Нельзя также смешивать три операции: псевдонимизация защищает связь атрибута с человеком, анонимизация стремится исключить идентифицируемость для релевантных

сторон, а удаление данных ограничивает дальнейшее существование конкретных копий. Удалённая фамилия не гарантирует ни анонимность, ни полное удаление, если в индексах сохранились фрагменты исходного документа. Тем более юридический вывод не может определяться только названием файла `anonymous.csv` или заверением подрядчика. [1]

В цепочке обработки следует различать состояние *до*, *во время* и *после* преобразования. При создании анонимного результата организация может обрабатывать персональные данные и обязана соблюдать применимые правила к этой операции. Только впоследствии, если результат действительно не позволяет установить лицо по контекстному правовому тесту, на него может не распространяться соответствующий режим защиты персональных данных. Документируемая разница этих этапов предотвращает ошибочное ретроспективное «обнуление» обязанностей.

3. Проект EDPB 02/2026: новая архитектура оценки анонимности

Руководство 02/2026, версия 1.0, принято 7 июля 2026 года для публичной консультации и опубликовано 8 июля. Оно обновляет подход рабочей группы статьи 29 2014 года и использует более детальную рамку оценки применительно к современным средствам обработки. В проекте предлагается начинать с двух вопросов: относится ли информация к физическому лицу и может ли оно быть идентифицировано или выделено таким способом, который позволит обращаться с ним иначе. Только после этого исследователь переходит к техническим критериям. [2][3]

Проект различает контекстуальный подход, учитывающий возможности каждого релевантного участника, и упрощённый подход, устанавливающий более консервативную оценку без детального разграничения сторон. Первый ближе к контекстному юридическому тесту, но сложнее в документировании. Второй может быть полезен организации, предпочитающей единый высокий уровень осторожности, однако способен привести к избыточному признанию сведений персональными в ситуациях, когда для отдельного получателя они объективно анонимны. Ни один подход не отменяет необходимость разумного, обоснованного анализа. [2]

Техническое ядро проекта образуют три критерия: No Record Isolation, No Linkage и No Inference. Их совместное соблюдение образует сильное основание для вывода об анонимности. Нарушение отдельного критерия не означает автоматически, что перед нами персональные данные: требуется дальнейшая оценка того, позволяет ли обнаруженная возможность действительно выделить или идентифицировать человека применительно к конкретному контексту. Это различие принципиально, поскольку «тест не пройден» и «установлено нарушение GDPR» не являются синонимами. [2]

Для Observatoire ARGA проект 02/2026 полезен прежде всего своей проверяемой процедурой. При подготовке независимого заключения рекомендуется фиксировать отдельно юридический объект оценки, перечень релевантных участников, источник вспомогательной информации, выбранный тест, его фактический результат и границы применимости вывода. Такое заключение допускает внешнюю рецензию и пересмотр после появления новых технологий идентификации.

Дополнительная экспертная оценка

Проект EDPB различает контекстный и упрощённый подходы. Три критерия — невозможность выделить запись, связать сведения и вывести информацию — служат практической рамкой для оценки. При этом их нельзя превращать в универсальную математическую аксиому. Если тест одного критерия выявляет риск, требуется дальнейшая оценка его существенности и правового контекста; если все три теста пройдены, результат подтверждает исследованную конфигурацию, но не существование абсолютной гарантии на неограниченный срок. [2]

Из этого следуют два требования для независимого отчёта. Во-первых, каждое испытание описывает атакующего, его начальное знание, полученные сведения и уровень уверенности. Во-вторых, отрицательный результат объясняется через обнаружительные возможности теста: какие классы атак охвачены, какие исключены, сколько было попыток и доступны ли для проверки исходные настройки. Только при наличии этих данных иной специалист сможет оценить достаточность исследования. Рекомендация состоит не в формальном «трёхгалочном» акте, а в иерархии доказательств. [2][15]

Для комментария EDPB целесообразно предложить пример с юридическим нарративом. Публикация может не содержать уникальных строк, пригодных для автоматического поиска, но всё же однозначно выводить на конкретное дело благодаря совокупности редких дат, процессуальных событий и межгосударственных перемещений. В такой ситуации семантическое сопоставление выступает самостоятельной угрозой, требующей ручного экспертного теста наряду со статистическими метриками.

4. Контекстуальная оценка и перспектива получателя

Один массив может быть персональным для организации, располагающей таблицей соответствия кодов и фамилий, и неперсональным для получателя, который объективно не может получить соответствие законными и разумно вероятными средствами. Именно такую относительность необходимо анализировать, а не превращать в универсальный тезис «псевдонимы у третьей стороны всегда анонимны». Суд ЕС в деле C-413/23 P *EDPS v SRB* от 4 сентября 2025 года отдельно рассмотрел перспективу получателя и одновременно подчеркнул, что обязанность контролёра по информированию субъектов оценивается в момент сбора данных с его собственной позиции. Дело было возвращено в Общий суд; решение Суда не является абстрактной индульгенцией любым передачам псевдонимизированных массивов. [7]

Контекстуальная проверка должна фиксировать правомерность и практическую осуществимость доступа к вспомогательным сведениям. Открытые базы, утечки, служебные полномочия, договорные права запрашивать данные и возможность получить информацию через посредника оцениваются раздельно. Само наличие у кого-либо в мире регистра идентификаторов ещё не делает доступ к нему разумно вероятным для каждого наблюдателя. Но и формальный запрет в договоре без технических и организационных мер не доказывает невозможность связывания. [1][2][8]

В проектах ARGA особенно важно различать пять перспектив: исходного контролёра, подрядчика-анонимизатора, закрытого получателя, публичного пользователя опубликованного набора и противника с доступом к данным вне предусмотренного маршрута. Если материал планируется публиковать открыто, применяется анализ широкой и долгоживущей аудитории, а не только текущего пользователя. Для закрытого доступа нужно дополнительно подтверждать реальность договорных, технических и организационных ограничений.

Релевантный участник	Доступные средства	Предварительный вывод
Контролёр с исходным реестром	Ключ, внутренние события, журналы изменений	Высокая вероятность атрибуции при сохранности связки
Процессор с ограниченным API	Функциональный доступ, возможные логи и метаданные	Проверка побочных каналов обязательна
Получатель обезличенного извлечения	Публикации, собственная база, допустимые запросы	Вывод зависит от объективно доступного контекста
Неопределённая публика	Открытые источники, новые сервисы и будущие утечки	Требуется более консервативная проверка

5. Критерий No Record Isolation: выделение записи

Выделение записи предполагает возможность обособить информацию об одном лице среди остальных, даже не зная его имени. Комбинация точной даты, малого населённого пункта, редкой профессии и необычного события иногда представляет собой фактический идентификатор. В проекте EDPB 02/2026 критерий No Record Isolation проверяет, содержит ли массив уникальное сочетание значений атрибутов, относящихся к одному человеку. Увеличение размерности и точности данных обычно повышает вероятность уникальности, но не позволяет заменить тест автоматической презумпцией. [2]

Например, таблица о поездках с возрастом, точным временем пересечения границы и направлением маршрута может выделять одну запись при отсутствии имени. Обобщение времени до месяца или региона может уменьшить уникальность, но лишить исследование части полезности. Задача аудитора состоит в сравнении возможных преобразований с назначением публикации и их остаточными рисками. Доказательственный протокол должен хранить правила обработки, распределение размеров эквивалентных групп, долю единичных записей и динамику после обновления набора.

Нельзя считать установленным юридический риск идентификации лишь по математическому признаку уникальности без оценки возможности отличить соответствующего человека в реальном контексте. Но высокая доля одиночных профилей является поводом для углублённой проверки и иногда отказа в публичном раскрытии. Опыт WP29 и CNIL показывает, что механическая замена полей не обеспечивает устойчивости к выделению и сопоставлению. [3][5]

Для независимого аудита предлагается рассчитывать не только общее число уникальных комбинаций, но и их распределение по классам риска, например по редким событиям, чувствительным географиям и типам решений. Сведения о международном розыске или состоянии здоровья способны раскрывать больше, чем обычные административные признаки. Поэтому единый порог для всех контекстов не рекомендуется.

6. Критерий No Linkage: связывание с внешними массивами

Связывание может происходить без общих формальных идентификаторов. Несколько временных отметок, маршрут, редкая должность и возраст иногда позволяют совместить две записи из разных баз. Источником могут стать публичные судебные документы, корпоративные реестры, новостные материалы, социальные сети и служебные архивы. Проект EDPB выделяет No Linkage как самостоятельную проверку: следует оценивать возможность соединить сведения о человеке в данной базе с другими данными так, чтобы возникло отличимое отношение к одному и тому же лицу. [2]

Один из характерных рисков возникает при последовательной публикации агрегированных наборов. Сегодня публикуется региональная таблица судебных решений, завтра — данные по возрасту и типу дела, затем — узкий статистический отчёт о сроках экстрадиции. Каждая таблица может выглядеть безопасной отдельно, но их пересечение уменьшает группы до одного-двух людей. Аудитор должен проверять планируемое раскрытие совместно с предшествующими публикациями самого контролёра и наиболее доступными внешними массивами.

Практически полезно описать «возможности сопоставителя»: какие поля совпадают, насколько они устойчивы, есть ли общедоступные события, какие ошибки возможны и как оценивалась доля ложных совпадений. Значение имеет не только процент успешно связанных записей, но и уверенность в точности атрибуции. Неверное приписывание криминального или санкционного статуса может причинять самостоятельный ущерб, даже если атакующий не идентифицировал реального субъекта исходного набора с высокой вероятностью.

При передаче данных в защищённую исследовательскую среду можно снижать риск посредством доступа к заранее утверждённым запросам, лимитов на выгрузку, аудита экспортов, независимого контроля ключей и правил публикации производных результатов. Эти механизмы управляют риском, но не превращают исходные данные в анонимные автоматически. [12][15]

7. Критерий No Inference: вывод чувствительных сведений

Даже если конкретная запись не выделяется и не связывается с другой базой, публикация может позволять с высокой надёжностью вывести сведения о человеке. Проект EDPB 02/2026 рассматривает No Inference как третий критерий, уделяя внимание точности и применимости такого вывода. В отличие от обычной корреляции на уровне большой группы опасным может стать вывод об отдельном человеке или очень узкой группе, который позволяет обращаться с ним особым образом. [2]

Пример: обезличенный отчёт учреждения показывает, что в определённый день единственный заявитель редкой национальности получил конкретное медицинское заключение. Даже без имени сочетание учреждения, времени и внешних сведений может раскрыть информацию о лице. Другой пример — публикация вероятностей санкционного совпадения по малому сектору, из которой контрагент способен вывести, кто оказался предметом углублённой проверки. Необходима связанная проверка критериев, а не три независимых «галочки».

Математический прогноз сам по себе не равен истинному знанию. Поэтому протокол обязан различать проверенную истинность атрибута, вероятностный вывод, ложные позитивные ответы и субъективную уверенность наблюдателя. Высокая точность классификатора на несбалансированном тестовом массиве не исключает опасных ошибок для малых групп. Не следует использовать единое число accuracy как достаточное доказательство сохранности конфиденциальности.

Рекомендация ARGA: тестировать сценарии, в которых злоумышленник владеет независимыми фактами о человеке, а также сценарии, где он знает только принадлежность к категории. По каждому тесту фиксируются начальное знание, источник, улучшение вывода и потенциальный вред. В правовом заключении описание этого риска должно оставаться понятным для юриста, а не только специалиста по ML.

8. Правовой режим самой операции анонимизации

Когда исходный материал является персональными данными, его преобразование в анонимную форму обычно представляет собой обработку персональных данных. Поэтому к операции могут применяться принципы статьи 5, надлежащее правовое основание статьи 6 и, при наличии особых категорий сведений или данных об уголовных правонарушениях, дополнительные требования статей 9 и 10. Достижение анонимного результата в будущем не легализует задним числом незаконный сбор, создание промежуточных копий или передачу исходного материала подрядчику. [1]

Контролёру надлежит указать цель обработки, оценить необходимость использования исходных идентификаторов, установить сроки хранения и определить границы полномочий процессора. Даже временная таблица соответствия псевдонимов может быть критичным персональным ресурсом. Её удаление должно документироваться с учётом резервных копий, журналов, выгруженных моделей, объектов в облаке и сроков физического уничтожения. Непроверяемое обещание «ключ удалён» не является достаточной процедурой.

В исследовательской сфере статья 89 GDPR допускает специальные гарантии и может взаимодействовать с национальными нормами и исключениями. Но «научный интерес» не является универсальным самостоятельным разрешением собирать любые сведения о человеке. Требуется анализ конкретной цели, правового основания, пропорциональности и допустимости

последующей передачи. Возможности ограничить права субъекта определяются применимым правом, а не договором между заказчиком исследования и исполнителем. [1]

Стадия	Основной вопрос	След документа
Получение исходного массива	На каком основании получены данные?	Реестр обработки, происхождение, цель
Преобразование	Кто видит исходные данные и ключи?	Схема доступа, DPA, журнал операций
Проверка устойчивости	Какие атаки и роли оценены?	Протокол испытаний, ограничения
Раскрытие результата	Для какой аудитории действует вывод?	Решение о выпуске, версия набора

9. Особые категории, уголовно-процессуальные и санкционные сведения

Анонимизация сведений о здоровье, убеждениях, биометрических признаках или уязвимых лицах требует более высокой осторожности не только потому, что раскрытие может причинить серьёзный вред. Сами операции над исходными данными должны иметь предусмотренное законом основание для обработки специальных категорий в контексте статьи 9 GDPR. Материалы о судимостях и уголовных правонарушениях подпадают под отдельную статью 10, связанную с официальным контролем или надлежащими гарантиями, предусмотренными правом Союза либо государства-члена. Нельзя механически объединять эти статьи. [1]

Санкционные списки, сведения о розыске, процессуальных решениях и расследованиях могут одновременно содержать персональные данные и сведения о правовом положении лица. Их режим зависит от происхождения, характера данных и применения конкретных норм. Наличие публичной записи ещё не гарантирует правомерность новой индексации, массового объединения с коммерческим профилем или публикации индивидуального риск-скоринга. Значение имеют цель повторного использования, точность сведений, дата прекращения действия статуса и возможность исправления ошибочных совпадений.

Для ARGA предлагается принцип «максимальной защиты от обратной атрибуции»: прежде чем публиковать аналитические примеры, сокращать детализацию событий, изменять редкие комбинации признаков, исключать материалы с очевидным индивидуальным профилем и отдельно тестировать сопоставимость с открытыми процессуальными документами. Если материал используется в юридическом деле, он может требовать иной логики обращения: сохранения оригинала в закрытом доказательственном контуре при публикации только безопасной агрегированной оценки.

Уничтожение идентифицирующей связи не должно разрушать права и возможности доказывания в конкретном деле, если соответствующая обязанность сохранения данных предусмотрена законом. Поэтому анонимизация для научной публикации и хранение надлежащего доказательства в юридическом досье разделяются архитектурно, по срокам хранения и правам доступа. Это не компромисс слов, а две разные операции обработки.

Дополнительная экспертная оценка

Сведения о здоровье, религиозных взглядах, биометрических признаках и иных специальных категориях требуют анализа статьи 9 GDPR, когда они относятся к установленным законом категориям. Сведения о судимости и уголовных правонарушениях имеют отдельный режим статьи 10. Их нельзя механически объединять в один перечень «чувствительных данных» без различения оснований и ограничений. Информация о санкциях или международном розыске может не подпадать автоматически под тот же терминологический режим, но иметь высокий

риск ущерба, а по обстоятельствам включать данные, охраняемые специальными нормами. [1]

В публичном кейсе об отказе в экстрадиции опасны не только имя и номер дела. Иногда достаточно сочетания даты, гражданства, страны задержания и срока судебного процесса. Если в архиве существует единственная новость с подобными событиями, читатель может соотнести отчёт с конкретным лицом за считанные минуты. Поэтому авторы должны проверять семантическую уникальность ещё до публикации. Простое удаление фамилии из материалов адвокатского дела — в такой среде псевдонимизация, а не надёжная анонимизация.

Согласование интересов исследования и защиты лица должно происходить до верстки. Иначе редактор обнаружит в последний момент, что ключевой сюжет содержит уникальные детали, и вынужден будет либо удалить доказательную часть, либо опубликовать опасный материал. Разумная стратегия предусматривает заранее две линии доказательства: закрытый оригинал для уполномоченного получателя и публичную, методологически самостоятельную модель случая.

10. Судебная практика ЕС и проверка фактической идентифицируемости

Решение Суда ЕС по делу *Breyer* (C-582/14, 19 октября 2016 года) подчёркивает значение реально доступных правовых способов получить дополнительные сведения: динамический IP-адрес может быть персональными данными для оператора сайта в условиях, когда существуют средства установления связи с человеком через провайдера. Суд не установил, что любой цифровой идентификатор всегда и для всех является персональным вне зависимости от обстоятельств. [8]

В деле *IAB Europe* (C-604/22, 7 марта 2024 года) рассматривалась строка, фиксирующая пользовательские предпочтения относительно согласия. Суд подчеркнул, что машиночитаемый код может относиться к идентифицируемому человеку, если он может быть связан с другими сведениями при наличии соответствующих средств. Организация, устанавливающая правила обработки, также при определённых обстоятельствах способна быть совместным контролёром, даже без прямого доступа к каждому массиву. [9]

Дело *EDPS v SRB* (C-413/23 P, 4 сентября 2025 года) показывает необходимость отличать квалификацию сведений для получателя от исходной обязанности контролёра по прозрачности в момент сбора данных. Суд отменил решение Общего суда и вернул дело на новое рассмотрение. Поэтому формула «после псевдонимизации третий получатель освобождается от GDPR» юридически неверна: нужны фактические доказательства невозможности отнесения сведений к субъекту для данного получателя, а обязанности исходного контролёра анализируются самостоятельно. [7]

Эти решения образуют методическую связку: сначала определить, относится ли информация к конкретному лицу, затем установить возможные средства идентификации с позиции релевантного участника, после чего отдельно проверить обязанности каждой стороны и каждого этапа. Доклад не подменяет анализ национальных норм и специальных отраслевых режимов; судебные принципы применяются к фактическим обстоятельствам, а не как абстрактный штамп на полях формы анонимизации.

Дополнительная экспертная оценка

В деле C-413/23 P от 4 сентября 2025 года Суд ЕС рассматривал передачу псевдонимизированных комментариев третьей стороне в контексте Регламента (ЕС)

2018/1725. Значение решения не в том, что все псевдонимизированные данные всегда автоматически анонимны или персональны для любого получателя. Суд обратил внимание на необходимость юридически корректного анализа идентифицируемости с учётом конкретной стороны и структуры передачи. [5]

Чтобы применить эту логику к практической оценке, аудитор составляет отдельные карты возможностей: исходного контролёра, технического подрядчика, исследовательского института, публики и потенциального недобросовестного получателя. Если один участник имеет ключ соответствия, а другой его не имеет, это существенный факт, но не исчерпывающий ответ: нужно проверить иные разумно вероятные средства идентификации. Критически важны договорные ограничения, фактический доступ и вспомогательные публичные данные, которые могут быть доступны всем.

Решение по конкретному спору не заменяет текст GDPR и не освобождает от оценки иных обязательств: законности изначальной обработки, прозрачности, обеспечения безопасности и прав субъектов. Превращение одного судебного аргумента в всеобщую формулу «у получателя нет ключа — значит данные анонимны» было бы методической и юридической ошибкой. Особенно в международных исследованиях, где получатели различаются возможностями и законодательными обязанностями.

11. Модель угроз и возможности мотивированного нарушителя

Проверка анонимности начинается с определения того, кто может пытаться установить личность и что этому участнику действительно доступно. Британский ICO рекомендует концепцию «мотивированного нарушителя»: это достаточно компетентный человек, располагающий общеизвестными ресурсами и способный целенаправленно использовать открытые и иные разумно доступные сведения. Модель не требует предполагать неограниченные средства и одновременно не позволяет ограничиться случайным пользователем, не заинтересованным в сопоставлении записей. [15]

Рекомендуемая матрица противников включает четыре категории: обычный публичный пользователь, профессиональный исследователь или журналист, получатель с собственной коммерческой базой и внутренний участник с привилегиями. Для каждой категории нужно описать доступные данные, вычислительные ресурсы, мотивацию, юридические барьеры и реальную возможность нарушения этих барьеров. Если речь идёт о высокоценных финансовых данных, предположения о мотивации и ресурсах должны быть строже, чем при публикации общей статистики посещаемости.

Угрозы распределяются по жизненному циклу: раскрытие исходного файла, компрометация таблицы соответствия, восстановление атрибутов по косвенным признакам, связывание с соседними публикациями, извлечение данных из модели, утечка через интерфейс запросов и повторная идентификация после появления новой внешней базы. В отчёте следует указать, какие сценарии тестировались, какие исключены и почему. Нельзя называть проверку «независимой» при отсутствии доступа к достаточному материалу для воспроизведения основных тестов.

Для измерения риска составляется реестр попыток: атакующая гипотеза, источник сведений, допущенный бюджет запросов, фактические попытки, достигнутый результат, неопределённость и влияние на субъектов. Отказ от проведения опасных атак на реальных людях должен сопровождаться безопасной заменой синтетическими или согласованными тестовыми данными, а не скрываться за записью «рисков нет».

Дополнительная экспертная оценка

При оценке возможностей атакующего не следует рассматривать только высококвалифицированного специалиста по безопасности. В ряде случаев человека легче установит местный журналист, коллега или родственник, знающий контекст события. Напротив, сложные атаки на машинное обучение требуют вычислительных ресурсов и доступа к модели, которые обычному получателю недоступны. Тест строится на нескольких профилях нарушителя: общественный наблюдатель, исследователь с легитимным доступом, инсайдер, коммерческий агрегатор данных и сторона, целенаправленно преследующая человека. [15]

Для каждого профиля фиксируется, какие вспомогательные сведения доступны законно или фактически, какая ожидается мотивация, сколько стоит попытка сопоставления и какие последствия создаёт успех. Время также имеет значение: возможность найти человека в общедоступном реестре за одну минуту и гипотетическая атака, требующая доступа к закрытым массивам нескольких государств, не эквивалентны. Обе могут обсуждаться, но их следует квалифицировать отдельно и не сводить к абстрактной формуле «теоретически возможно».

Модель угроз должна пересматриваться при выходе новых инструментов, изменении доступности записей, утечке данных партнёра и росте вычислительных возможностей. Надёжность отчёта подтверждается тем, что в нём честно перечислены недоступные для испытания классы угроз и установлены события, при которых потребуется повторная оценка.

12. Прямые и косвенные идентификаторы: данные о событиях

На практике самыми устойчивыми идентификаторами оказываются не только паспортные номера и полные имена. Последовательность редких событий, описание профессии, время операции, географический маршрут, номер процедуры или необычный набор обращений может представлять более сильный ключ. Для аудита полезно составлять классификацию атрибутов по роли: прямые, квазиидентификаторы, чувствительные свойства, технические метаданные и поля контроля качества.

Преобразование квазиидентификаторов должно учитывать их взаимосвязь. Обобщение возраста до десятилетий мало помогает, если сохраняется точная дата события и единственная должность в небольшой организации. Удаление номера дела не решает проблему, если текстовая фабула и дата суда однозначно находятся поисковой системой. Особенно опасны нарративные поля: они способны переносить уникальные сведения, невидимые для автоматического подсчёта повторов табличных строк.

Операционная процедура предусматривает семантический анализ свободного текста, поиск косвенных ссылок на лиц и небольшие группы, проверку ссылок на уникальные публичные документы, затем тесты объединения обработанных атрибутов. Применение модели распознавания персональных данных на этом этапе полезно, но требует проверки её полноты на языке и предметной области конкретного досье. Алгоритмический фильтр не должен подменять экспертное чтение в материалах высокой чувствительности.

Отдельный контроль необходим для вложений: PDF с авторскими метаданными, скриншоты, временные метки в EXIF, названия файлов, идентификаторы комментариев и скрытые слои могут повторно идентифицировать субъектов после тщательной обработки основного текста. Протокол должен охватывать готовый пакет целиком, а не только его визуально видимую часть.

13. Геолокация, время и траектории перемещения

Данные перемещений создают риск восстановления личности через комбинации «дом — работа — редкое событие», даже если координаты округлены. Чем выше пространственное и временное разрешение, тем меньше вероятность, что две независимые траектории окажутся одинаковыми. Нельзя автоматически считать анонимными журналы локации после замены идентификатора устройства случайной строкой: последовательность пребывания сама может выполнять функцию устойчивого отпечатка.

Для снижения риска применяют агрегацию по достаточно широким районам и временным интервалам, обрезание редких траекторий, публикацию вероятностных распределений вместо индивидуальных маршрутов и ограничение доступа к сырому журналу. Но каждая операция способна искажать исследовательский результат: при расследовании нарушения границы или цепочки экстрадиционных перемещений потеря часовых меток может лишить доказательства правового смысла. Поэтому необходимо разделить закрытый доказательственный оригинал и публикационную копию.

Проверка должна учитывать последовательные релизы: публикация данных по неделям позволяет восстанавливать поведение из пересекающихся временных окон, даже когда каждый выпуск проходит отдельный тест малых групп. Злоумышленник способен вычитать и сопоставлять агрегаты. Организация должна вести общий реестр выпусков с их фильтрами, периодами и метаданными, а не оценивать новые файлы в вакууме.

Рекомендация для ARGA: при географической отчётности использовать минимально необходимое разрешение, фиксировать допустимую цель точных координат, проверять редкие маршруты и назначать независимое решение о публикации, если речь идёт о людях под риском преследования. Это относится как к открытым материалам, так и к информационным панелям с доступом клиентов.

Дополнительная экспертная оценка

Пространственные и временные признаки часто обладают высокой идентификационной силой. Сочетание немногих точек маршрута способно выделять человека независимо от имени; к тому же географический риск меняется с плотностью населения и доступностью местных публикаций. Общая рекомендация «округлите координаты до двух знаков» ненадёжна без учёта размера территории, частоты событий и числа людей, попадающих в соответствующий диапазон. [12][15]

Правильная процедура состоит в сравнении нескольких вариантов разрешения: точная дата/месяц/квартал, населённый пункт/регион/страна, полный маршрут/обобщённая последовательность. Для каждого варианта оцениваются уникальные комбинации, возможность связывания с независимыми данными и влияние огрубления на исследовательский результат. Если предметом работы является соблюдение сроков процессуальных действий, чрезмерное укрупнение времени может скрыть нарушение. В таком случае разумнее предоставить точные сроки уполномоченному рецензенту, но публиковать только итоговую правовую оценку без редкой комбинации деталей.

Особенно опасны небольшие группы с уникальным событием: единственный отказ в экстрадиции в регионе или редкая передача из одного учреждения в другое. При ограниченном количестве публичных дел даже широкий диапазон даты не обязательно защищает человека. Условный кейс, построенный с независимым сюжетом, предпочтительнее, чем попытка механически «размыть» заведомо узнаваемый реальный.

14. Графы связей, текстовые эмбеддинги и векторные базы

Векторное представление текста не содержит имени в привычном виде, но может сохранять признаки содержания. Если эмбеддинг позволяет находить исходный документ, выявлять его тему или отличать небольшой круг людей, вопрос анонимности требует отдельного доказательства. Ошибочно считать, что преобразование строки в числовой массив автоматически уничтожает персональную природу информации. Важно определить, какие запросы принимает поиск по векторам, доступны ли обратимые связи с исходными документами и какие признаки могут быть извлечены из результата.

Графовые базы добавляют ещё один канал идентификации: роль узла и конфигурация его рёбер могут быть уникальны. Удаление имён участников финансовой транзакции может не устранить возможности установить участников по суммам, времени переводов и структуре контрагентов. Достаточно одному наблюдателю обладать частью внешних связей, чтобы восстановить остальные через совпадающий подграф. Поэтому отдельная проверка каждого узла не заменяет проверки графа.

Для Retrieval-Augmented Generation нужно учитывать весь маршрут данных: исходный текст, индекс, метаданные документа, права доступа, результат поиска, контекст модели, сгенерированный ответ и журналы запросов. Анонимность одного слоя не означает анонимность другого. Если модель получает скрытый идентификатор вместе с текстом, его отсутствие в итоговом ответе не подтверждает, что весь процесс обработки исключён из GDPR. [4][6]

Проверка независимого аудитора может включать набор запросов на восстановление редких формулировок, тестирование соседства векторных представлений, сравнение результатов с публичными корпусами и контроль получения фрагментов из документов с ограниченным доступом. Выполнять эти тесты следует в управляемой среде, не публикуя в отчёте сами найденные персональные сведения.

15. Подавление, генерализация и рандомизация: что работает и где ломается

Классические методы снижения идентифицируемости остаются полезными, если применяются к конкретной модели угроз. Подавление удаляет отдельные значения или строки; генерализация снижает точность атрибутов; рандомизация изменяет возможность связать наблюдение с конкретным человеком. Однако ни один из методов не даёт универсальной гарантии. Например, удаление адреса не помогает при уникальной фабуле дела; округление суммы сделки не устраняет редкость контрагента; случайная перестановка значений способна разрушить корреляции, важные для исследования. [3][5][12]

При проектировании преобразования нужно определить, что именно требуется сохранить: частоты по регионам, распределение сроков, наличие редких ошибок или возможность проверить законность индивидуального решения. Если необходим только процент случаев ошибочного совпадения в общей массе, публикация строковых досье избыточна. Если нужна доказательная реконструкция конкретной ошибки, анонимизация до агрегатов может быть неприемлемой, и следует выбрать закрытую экспертизу под правовыми гарантиями.

Аудитор проверяет не просто количество удалённых полей, а изменения в вероятности выделения, связывания и вывода, а также утрату полезности данных. Отчёт должен содержать исходные характеристики в закрытом протоколе, метод трансформации, метрики до и после, перечень неустранённых рисков и мотивированное решение о допустимости выбранного режима доступа.

Метод	Типичный эффект	Типичная остаточная уязвимость
Подавление редких строк	Уменьшение явных одиночных случаев	Вывод по оставшимся агрегатам
Генерализация возраста, времени, места	Рост групп сходства	Связывание по другим атрибутам
Рандомизация отдельных значений	Снижение определённости атрибуции	Восстановление статистическим усреднением
Синтетическая публикация	Отсутствие прямых строк при удачном генераторе	Запоминание обучающих наблюдений

16. k-анонимность, l-разнообразие и t-близость: границы моделей

k-анонимность требует, чтобы каждая запись была неотличима по выбранным квазиидентификаторам как минимум от k−1 других записей. Это полезный диагностический показатель, но его содержание зависит от выбранного перечня полей. Если в группе из пяти строк все пять содержат одинаковое чувствительное значение, групповое совпадение может раскрывать это значение при знании о принадлежности человека к группе. Здесь возникает проблема однородности чувствительного атрибута.

l-разнообразие стремится уменьшить проблему однородности, предусматривая разнообразие чувствительных значений внутри групп. t-близость дополнительно связывает их распределение с распределением по более широкой популяции. Эти конструкции важны методически, однако не гарантируют защиту от хорошо информированного наблюдателя, последовательных публикаций и смешения контекстов. Они также могут требовать слишком сильного обобщения при малых или неравномерных группах. [3][12]

Пример расчёта: если набор разделён на 120 групп, 15 из которых включают только одну строку по выбранным квазиидентификаторам, доля таких групп равна 12,5%. Это не равнозначно доле уникальных людей во всём массиве: для последней нужен вес каждой группы по числу строк. Если общее число строк 600, показатель «15 уникальных строк из 600» составляет 2,5%. Указывать в отчёте только 12,5% было бы статистически некорректно.

Рекомендация ARGA: использовать k-, l- и t-метрики только вместе с явно указанными квазиидентификаторами, моделью внешней информации и проверкой полезности. Вывод «k = 5, следовательно GDPR не применяется» юридически не обоснован. Отдельно необходимо оценивать возможность вывести юридически значимые атрибуты в небольших группах.

Дополнительная экспертная оценка

Модели групповой неразличимости дают полезные тесты, но работают в пределах выбранного набора квазиидентификаторов. Если набор дополняется внешней публикацией, класс эквивалентности может распасться. Кроме того, если все люди внутри одной группы обладают одинаковым чувствительным признаком, размер группы сам по себе не препятствует выводу этого признака. Поэтому k-анонимность не является универсальной заменой анализу No Inference. [12]

Показательный пример: в группе из двадцати человек одинакового возраста и региона скрыты имена, но все двадцать имеют один и тот же диагноз. Даже если определить, какая именно строка принадлежит человеку, невозможно, установление его принадлежности к этой группе позволяет логически вывести диагноз. В другом случае распределение диагнозов может быть разнообразным, но временная последовательность остаётся уникальной и облегчает связывание с газетной заметкой. Разные меры защищают от разных классов угроз.

Поэтому аудитор документирует полный перечень использованных признаков, обоснование их выбора, влияние подавления на качество статистики и тесты с изменёнными наборами квазиидентификаторов. Если критерий достигнут только за счёт исключения ключевых случаев, исследование должно раскрыть возникшее смещение. Значение k устанавливается не произвольно по привычке, а рассматривается в контексте вреда и доступных средств идентификации.

17. Дифференциальная приватность: формальная гарантия и смысл параметров

Дифференциальная приватность (differential privacy, DP) описывает ограничение изменения распределения выходов механизма при замене соседних наборов данных, различающихся вкладом одного субъекта в принятой модели соседства. В распространённой форме механизм M обладает (ε,δ)-дифференциальной приватностью, если для всех соседних D и D′ и измеримых событий S выполняется P[M(D)∈S] ≤ e^ε·P[M(D′)∈S]+δ. Эта математическая формулировка относится к корректно реализованному механизму с чётко заданными предпосылками и не означает полного отсутствия информации о человеке в результате. [13]

Параметр ε характеризует предел изменения вероятностей и зависит от механизма и принятой модели соседства; δ допускает дополнительный остаточный риск в формальной гарантии. Меньшее ε при прочих равных обычно означает более сильную приватностную гарантию, но универсального юридического «безопасного epsilon» не существует. Его оценка требует понимать вклад одного человека, размер аудитории, число выпусков и возможные побочные каналы. NIST SP 800-226 (финальная версия марта 2025 года) предупреждает, что заявка на DP без анализа параметров, реализации и среды может создать ложную уверенность. [13]

Пример предназначен только для иллюстрации: при простой последовательной композиции двух механизмов с параметрами (0,5; 0) и (0,7; 0) консервативная верхняя граница суммарного расхода равна ε = 1,2 и δ = 0 при соответствующих предпосылках. Это не измерение реальной приватности модели ARGA и не сертификат анонимности результата. При повторных адаптивных запросах нужен учёт бюджета и специфических правил композиции.

Для независимого аудита необходимо раскрыть определение соседства (уровень строки или человека), максимальный вклад субъекта, тип шума, механизм генерации, защищённость генератора случайности, перечень опубликованных выпусков и накопленный бюджет. Доказательство корректности математической формулы без проверки кода и цепочки обработки недостаточно.

Дополнительная экспертная оценка

Дифференциальная приватность формулирует статистическую гарантию устойчивости результата к участию защищаемой единицы при определённом сравнении соседних наборов. Параметр epsilon не равен вероятности реидентификации, а delta нельзя описывать без механизма, модели соседства и условий композиции. Даже при математически корректном определении реализация может разрушать гарантию незащищёнными логами, детерминированным генератором шума, неучтёнными повторными запросами или побочным каналом. [13]

В правовом заключении необходимо отделить математический факт от юридического вывода. Механизм может удовлетворять выбранному определению DP, но организация всё ещё обязана установить законность исходной обработки, достаточность мер безопасности, прозрачность и применимость прав субъектов. Не существует известного универсального

порога epsilon, который автоматически превращает все результаты в анонимные данные для целей GDPR. [1][13]

Методика оценки должна требовать отчёт о параметрах, модели атакующего, защищаемой единице, составе релизов, тестах корректности и статистическом качестве. По возможности публикуются проверяемые агрегаты и описание гарантии, а исходные данные и опасные демонстрационные запросы остаются в защищённой среде. Именно такой уровень детализации позволяет внешнему рецензенту оспорить вывод, не получая без необходимости сведения о самих людях.

18. Учёт DP-бюджета, композиция и качество публикуемых результатов

Одна из главных практических ошибок состоит в многократном использовании одного массива с независимыми обещаниями «каждый релиз защищён». При повторных ответах на запросы приватностная потеря может накапливаться, а объединение выходов позволяет уточнять значения. Система должна вести учёт расхода бюджета в разрезе субъекта или группы согласно модели соседства и технической реализации. Нельзя восстанавливать бюджет простым суммированием ε, если механизм заявляет другие гарантии композиции, но нельзя и игнорировать историю публикаций.

DP требует анализа точности результата. Если шум в небольшой группе сопоставим по величине с самой группой, отчёт может стать бесполезным или вводящим в заблуждение. В докладе о редких случаях незаконной экстрадиции грубое искажение нескольких наблюдений способно полностью изменить интерпретацию. Поэтому критерий полезности необходимо фиксировать заранее и не подбирать механизм только ради красивого значения ε.

Публикация графиков и дашбордов должна учитывать вероятностные диапазоны, метод округления, возможность отрицательных шумных счётчиков и работу с нулями. Нормализация, пороги раскрытия и постобработка сами по себе могут менять интерпретацию результата. Фраза «дифференциально приватный» должна сопровождаться описанием метрики, бюджета и целей, а не служить декоративной печатью одобрения.

Предлагаемый журнал DP-релизов содержит версию исходной популяции, цель, владельца механизма, идентификатор запрошенной статистики, параметры приватности, доказательства тестирования, точность и дату закрытия бюджета. Это даёт возможность остановить публикацию новых срезов при опасном накоплении запросов. [13]

19. Синтетические данные и имитационные наборы

Синтетические данные создаются программно или статистическим генератором и иногда могут уменьшить потребность в передаче реальных наблюдений. Однако модель генерации, обученная на персональном корпусе, способна воспроизводить редкие строки либо сохранять сильные корреляции, позволяющие восстановить индивидуальные признаки. Следовательно, слово «синтетические» описывает способ создания, а не гарантированный юридический статус результата. [4][12]

Проверка включает как минимум поиск почти точных совпадений с обучающим массивом, тесты членства записи в обучении, попытки восстановить редкие комбинации и анализ статистического сходства для малых групп. Следует отделять ошибки генерации от риска раскрытия: синтетический набор, который воспроизводит несуществующие уголовные события, может быть непригоден для обучения правового классификатора даже при удовлетворительном уровне конфиденциальности.

Особое значение имеет область применения. Синтетические анкеты для демонстрации структуры базы можно генерировать из заранее выдуманной схемы без обращения к реальным досье. Но

синтетические судебные материалы, рассчитанные на воспроизведение частоты процессуальных нарушений, потребуют убедительной оценки репрезентативности и рисков. Имитированные персональные профили не следует публиковать как данные о реальных лицах; подписи, происхождение и ограничения обязаны присутствовать во всех версиях.

Рекомендация: для обучения внутренних интерфейсов использовать по возможности полностью искусственные тест-кейсы, не производные от конфиденциальных материалов клиентов. Если требуется статистическая близость к реальным данным, проводить отдельную экспертизу механизма синтеза и минимизировать доступ к оригиналам.

Дополнительная экспертная оценка

Синтетический набор полезен для тестирования программ, обучения методик и публикации примеров, но само слово «синтетический» не устанавливает анонимность. Если алгоритм генерации запомнил и воспроизвёл редкие записи, результат может сохранять персональную связь. Риск особенно заметен в корпусах, где много стандартных записей, но отдельные уникальные истории способны быть скопированы практически дословно. [3][12]

Проверка должна включать поиск ближайших соседей относительно исходного корпуса, сопоставление редких сочетаний, тесты membership inference и оценку раскрытия атрибутов. При этом общая близость распределений не равна доказательству отсутствия меморизации. Аудитор отдельно рассматривает малочисленные группы, ошибочные классификации и влияние генерации на правовые выводы. Если синтетическая выборка воспроизводит слишком мало негативных случаев, аналитик может ошибочно заключить, что дискриминационных рисков практически нет.

Для публикационных задач разумно отделять два вида материалов. «Учебный условный кейс» создаётся вручную как вымышленный пример и не претендует на статистическую репрезентативность. «Синтетический набор» генерируется по реальным данным и может сохранять статистические свойства, поэтому требует отдельного теста раскрытия и методического описания. Эти форматы нельзя подменять друг другом ради удобства редакции.

20. PETs, федеративное обучение и защищённые вычисления

Privacy Enhancing Technologies (PETs) включают различные способы снижения угроз: защищённые многосторонние вычисления, доверенные вычислительные среды, федеративное обучение, криптографические протоколы, контролируемые исследовательские пространства и дифференциальную приватность. Эти механизмы решают разные задачи. Например, федеративное обучение сокращает необходимость передавать исходные массивы в центральный сервер, но обновления модели могут раскрывать сведения при неправильной настройке. Наличие защищённого канала передачи не доказывает анонимность результата.

Криптографическая защита в состоянии хранения и передачи снижает риск несанкционированного доступа; юридический статус расшифрованных данных для уполномоченного получателя остаётся самостоятельным вопросом. Доверенная среда ограничивает доступ и способы вывода, но не отменяет необходимость проверки безопасности её конфигурации, журналов и правил экспорта. ENISA рассматривает развитые методы псевдонимизации, в том числе криптографические, в качестве инструментов снижения риска, а не универсального преобразования в неперсональные сведения. [14]

Технологический выбор должен следовать модели угроз, а не моде поставщиков. Для совместного исследования дел о розыске возможны защищённые агрегированные запросы без объединения именных реестров. Для независимого аудита алгоритма может потребоваться

доступ проверяющего к тестовой выборке в контролируемом контуре. Для публичного отчёта чаще достаточно агрегированной публикации с безопасными порогами и проверкой раскрытия через пересечение релизов.

Итоговое решение должно сравнивать правовую допустимость, защищённость исходных данных, качество результата, риск остаточного извлечения и возможность внешней проверки. Если конструкция снижает риск, но юридически не создаёт анонимность, это следует обозначить прямо: применение PETs и соответствие GDPR относятся к связанным, но не тождественным вопросам.

21. Может ли обученная AI-модель быть анонимной

В Opinion 28/2024 от 18 декабря 2024 года EDPB указал, что модель, обученная на персональных данных, не является автоматически анонимной. Оценка проводится индивидуально. Для обоснования анонимности необходимо учитывать как вероятность непосредственного получения персональных данных о субъектах обучения из модели, так и вероятность их получения через запросы, при условии использования разумно вероятных средств. При этом надзорные органы должны анализировать характеристики данных, архитектуры и процесса обучения, контекст развёртывания и круг потенциальных пользователей. [4]

CNIL в материале от 22 июля 2025 года отдельно предлагает поставщикам метод оценки статуса модели по GDPR. Ключевой риск — запоминание фрагментов обучающего корпуса и возможность их извлечения. Следовательно, нельзя выводить модель из режима персональных данных только из-за отсутствия явной текстовой таблицы или потому, что сведения распределены между параметрами нейронной сети. [6]

Для научного доклада важно различать «модель как хранилище вероятностных закономерностей» и «модель как возможный источник сведений об отдельных людях». Даже при высоком качестве тестов отсутствие успешного извлечения в фиксированном наборе попыток не доказывает абсолютную невозможность новых атак. Изменение способа доступа, системного промпта, fine-tuning, retrieval-слоя и аппаратного окружения может менять вероятность раскрытия. В заключении нужно указывать проверенную версию модели и поверхность доступа.

Предлагаемый минимальный пакет доказательств включает описание обучающего корпуса и его происхождения, процесс исключения высокорисковых категорий, меры против запоминания, тестирование атак по заранее зарегистрированному плану, отчёт о неудачных и удачных попытках, управление доступом к весам и процедуру повторной оценки после изменения модели. Этот пакет не заменяет правовую основу первоначальной обработки данных.

Дополнительная экспертная оценка

Заключение об анонимности AI-модели требует анализа её фактических возможностей, а не только структуры обучающих файлов. В Opinion 28/2024 EDPB подчёркивает необходимость оценки вероятности прямого извлечения персональных данных, в том числе вероятностного, и возможности получения таких сведений из запросов к модели с учётом разумно вероятных средств. Конкретный результат зависит от архитектуры, обучения, доступа к весам, интерфейса и дополнительных систем. [3]

Например, закрытый интерфейс может иметь фильтры, ограничивающие число запросов и возвращение длинных фрагментов; открытые веса позволяют исследователю менять режим генерации или использовать собственные средства анализа. Поэтому один вывод об анонимности нельзя автоматически переносить между двумя способами распространения. В RAG дополнительно появляется поисковый слой, который может выдавать фрагменты из закрытого корпуса даже при отсутствии такой возможности у базовой модели.

Тестирование проводится повторяемо: формируются допустимые сценарии, измеряется успешность атак, анализируются случаи ложного восстановления и приводятся ограничения. При отсутствии доступа к обучающему корпусу аудитор должен прямо сказать, какие гипотезы о меморизации проверить невозможно. Убедительное заключение не скрывает неизвестные величины за утверждением «признаков утечки не обнаружено».

22. Запоминание, регургитация и извлечение обучающих примеров

Регургитация возникает, когда генеративная модель воспроизводит материал, близкий к обучающим данным, намеренно, случайно или вследствие целенаправленного запроса. CNIL прямо рассматривает этот механизм как один из значимых рисков при оценке статуса AI-модели. При этом степень близости ответа к исходному тексту сама по себе ещё не определяет, раскрыты ли персональные данные: значение имеет наличие идентифицируемого лица и характер воспроизведённых сведений. [6][17]

На запоминание влияют повторяемость и редкость фрагментов, объём корпуса, сложность модели, конфигурация обучения и особенности вывода. Тесты должны включать не только очевидные запросы на повторение текста, но и вариации контекста, альтернативные формулировки, частичные подсказки и допустимую многошаговую работу с интерфейсом. Отдельно требуется наблюдать, не позволяет ли система извлечь персональные сведения из скрытого поискового индекса или журналов инструментов.

Меры снижения риска включают дедупликацию, исключение редких чувствительных примеров, регуляризацию обучения, технические ограничения интерфейса и мониторинг выдачи. CNIL приводит такие меры как возможные способы уменьшения запоминания или извлечения в контексте разработки AI. Но фильтр готового ответа не отменяет наличие сведений в параметрах модели и не гарантирует защиту при выпуске открытых весов. [17][20]

Профессиональное заключение не должно описывать удачный тест как инструкцию для воспроизведения утечки. Необходимо разделять конфиденциальное доказательство уязвимости и публичное изложение класса риска. Для публикации достаточно назвать категорию извлечённых сведений, условия, частоту и принятые исправления; сами личные данные должны оставаться в защищённом контуре до решения уполномоченных лиц.

23. Атаки на принадлежность записи к обучающему корпусу

Membership inference стремится определить, использовались ли сведения о конкретном человеке при обучении модели. Такой вывод может быть чувствительным даже без восстановления полного документа: попадание в базу жалоб, медицинских историй или материалов уголовного преследования само по себе раскрывает факты о субъекте. Метод оценки должен сопоставлять поведение модели на специально отобранных примерах из обучающего и контрольного наборов и учитывать различия распределений, которые способны создать ложный сигнал.

Важна исходная вероятность принадлежности: даже тест с высокой чувствительностью может давать много ложноположительных результатов, если реальная доля членов целевой группы мала. Поэтому аудитору необходимо сообщать не только precision и recall на сбалансированной лабораторной выборке, но и анализ применимости к вероятным условиям реальной атаки. Технически корректный эксперимент с искусственным балансом классов не позволяет прямо говорить о риске для населения без корректировки предпосылок.

В условном сценарии известно, что доля записей определённого типа в обучающем массиве составляет 1%. При чувствительности теста 90% и специфичности 95% положительная прогностическая ценность равна 0,009 / (0,009 + 0,0495), то есть около 15,4%. Такая атака не даёт достаточной достоверности для утверждения, что конкретный человек включён в набор, хотя

сама её возможность требует исследования. Числа здесь условные и приведены только как демонстрация важности базовой частоты.

Аудитор должен зафиксировать модель доступа, частоту ложных срабатываний, доверительные интервалы, долю пересечений между контрольными множествами и версии данных. Для чувствительных сфер рекомендуется дополнительно моделировать действия злоумышленника, располагающего фактом о единственной поездке, госпитализации или процессуальном событии.

Дополнительная экспертная оценка

Для membership inference решающее значение имеет частота принадлежности в исследуемой популяции. Если она мала, высокий показатель sensitivity не гарантирует, что положительный сигнал с высокой вероятностью указывает на реального участника обучения. Поэтому на практике нужно показывать как минимум истинно положительные, ложно положительные, истинно отрицательные и ложно отрицательные выводы, а также положительную прогностическую ценность и используемую базовую частоту. [3][13]

Атаки часто разрабатываются и оцениваются на специальных теневых моделях. Их распределение может отличаться от распределения реальной модели, а значит, значения метрик не обязательно переносятся на целевую систему. Дополнительно проверяется утечка через повторные запросы и подбор порогов после просмотра тестовой выборки. Чтобы избежать оптимистического смещения, протокол должен фиксировать пороги до независимой оценки и отдельно учитывать исследовательские попытки, использованные при разработке атаки.

Полученный результат нельзя использовать для создания неоправданных персональных обвинений. Если атака ошибочно отнесла человека к учебному корпусу, публичное заявление само по себе может нанести репутационный или юридический вред. Независимый наблюдатель публикует агрегированные показатели и ограничения, а подтверждение членства отдельных записей требует отдельного доказательства происхождения данных.

24. Model inversion и вывод атрибутов

Модельная инверсия и attribute inference направлены на восстановление или вывод характеристик наблюдений, использованных при обучении либо доступных через модель. Они отличаются от простого воспроизведения исходных строк: злоумышленник может получить вероятный признак или прототип, который вместе с внешними данными позволит отнести результат к человеку. Для правовой оценки существенны достоверность атрибуции и возможность индивидуального воздействия, а не только визуальная похожесть изображения или формальное сходство текста.

При оценке риска следует отделять то, что модель сообщает о группе, от того, что можно обоснованно заключить об одном субъекте. Например, прогноз вероятности болезни по демографическому набору не обязательно идентифицирует человека; но совместное использование закрытого кода процедуры и редкого события может фактически создавать индивидуальный профиль. Если система используется для скоринга заявителей, вероятностный вывод может иметь реальное последствие даже при отсутствии уверенности в точном имени.

Тестовый протокол фиксирует исходные предположения атакующего, модель или API, использованные запросы, контроль против галлюцинаций и независимое подтверждение восстановленных атрибутов. Результаты измеряются отдельно по подгруппам; нельзя усреднять риски для редкой уязвимой группы с массой менее чувствительных профилей. Следует также оценивать правомерность подготовки тестовых данных и исключать проведение опасных экспериментов на клиентских материалах без соответствующих гарантий.

Результатом экспертизы должно быть ограниченное утверждение: при описанных предпосылках и ресурсах удалось либо не удалось получить данные определённой категории. Формулировки «модель не хранит личных данных» и «невозможно узнать, был ли человек в обучении» требуют более сильного доказательства, чем простой набор отрицательных тестов.

25. Организация испытаний: тестовый дизайн и независимость проверяющего

Качественный аудит требует заранее зафиксированного плана: объект исследования, версия данных и модели, роли участников, набор угроз, методы измерения и критерии завершения. Без этого исследователь может непреднамеренно подобрать тесты под желаемый результат и объявить безопасность по итогам проверок, которые никогда не были способны выявить существенную утечку. План испытаний должен быть утверждён до доступа аудитора к итоговым метрикам и содержать процедуру регистрации отклонений.

В независимом контуре целесообразно разделять команду, готовящую данные, и команду, проверяющую риски. Разработчик вправе сообщить архитектуру и необходимые технические ограничения, но не должен единолично определять, какие найденные ошибки считать существенными. Аудитор обязан раскрыть финансовые, организационные и договорные зависимости, а также любые ограничения доступа к весам, логам и обучающему корпусу. Наличие сертификата системы управления не заменяет предметного испытания конкретной модели.

Процедура воспроизведения включает контрольную копию исходных материалов, хеши файлов, версию окружения, журнал запуска, безопасное хранение результатов и возможность независимого повторного подсчёта метрик. Если части корпуса недоступны по законным причинам, эксперт должен обозначить, какие выводы можно сделать в пределах предоставленного доступа, а какие нельзя подтвердить. Корректно ограниченное заключение предпочтительнее уверенного, но недоказуемого сертификата.

Для применения в ARGA рекомендуется формировать двухчастное заключение: публичную аналитическую записку и закрытый технический отчёт. Первая описывает метод, границы и статус риска; второй содержит доказательные материалы, включая секретные параметры и потенциально персональные примеры, с ограниченным доступом и сроками хранения.

Дополнительная экспертная оценка

Независимый тест не должен ограничиваться проверкой заранее отобранных заказчиком наиболее безопасных примеров. Протокол включает правила выборки, критерии исключения, число попыток, защиту исходного материала и фиксированные метрики успеха. Часть сценариев согласуется с разработчиком для предотвращения вреда, но часть создаётся независимой командой, чтобы испытание не превратилось в демонстрацию уже известных положительных результатов. [2][3]

Особенно важно разделять разработку атаки и её окончательную оценку. Если специалист подбирает десятки вариантов запроса на одной тестовой выборке и затем сообщает только лучший результат, итоговая метрика не отражает ожидаемую эффективность на независимых данных. Публичный отчёт должен показывать число исследованных конфигураций, выбор финального порога и качество на отложенной группе. Для сценариев, где реальные данные нельзя передавать внешнему эксперту, используется защищённая среда с доступом по ролям и журналом действий.

Перед началом теста участники согласуют порядок немедленной остановки при обнаружении живого лица или иной чувствительной информации. Факт обнаружения

сохраняется в закрытом акте, но опубликование самоидентифицирующего примера не требуется. Независимость считается документированной, если заказчик не может в одностороннем порядке исключить отрицательный результат, а конфликт интересов раскрыт до начала испытаний.

26. Статистическая неопределённость и ошибочные выводы тестирования

Отрицательный результат испытаний не является доказательством нулевого риска. Если в n независимых испытаниях не обнаружена ни одна успешная утечка, ориентировочная верхняя 95-процентная граница вероятности события при биномиальных предпосылках часто оценивается правилом 3/n. Например, при 0 успехов из 300 это приблизительно 1%, а не ноль. Для другого устройства испытания, зависимых запросов или адаптивного противника такая оценка может быть неприменима. Метод расчёта и предпосылки обязательно раскрываются.

Если из 200 атак успешными оказались 8, эмпирическая доля равна 4%. Но сама эта цифра не указывает, отражает ли выборка реальный риск. Необходимо оценивать интервалы неопределённости, состав тестов, долю редких примеров, качество контроля и способ отбора запросов. «Средний процент успеха» может скрывать 40-процентную уязвимость единственного особо чувствительного подмножества.

Показатели usefulness и privacy следует сообщать раздельно. Например, снижение риска связывания после удаления точных дат может одновременно уменьшить полезность статистики о сроках экстрадиции. Научно добросовестный вывод допускает, что оба показателя ухудшились, либо что задача требует другого режима доступа вместо публикации детализированного набора. В таблицах должны быть указаны знаменатели и метод определения «успеха».

Рекомендуется публиковать минимальную спецификацию: популяция, тестовая выборка, мера события, точечная оценка, интервал, существенные систематические ограничения и различие между испытанной версией и планируемым продуктом. Для внутренних решений можно использовать заранее утверждённую матрицу допустимости, но её пороги необходимо объяснять как политику управления риском, а не общеобязательный стандарт GDPR.

Дополнительная экспертная оценка

Нулевое число обнаруженных успешных атак не означает нулевую вероятность успеха. В простой модели независимых испытаний при нуле успехов из n попыток часто используют приближённую верхнюю границу 3/n для 95-процентного одностороннего интервала. Например, при 0 из 300 это около 1%, но только при выполнении предпосылок о сопоставимости и независимости испытаний. Если запросы коррелированы или атакующий адаптирует стратегию после каждого ответа, простая формула может оказаться неприменимой. [3]

Следовательно, отчёт должен показывать не только дробь «0/300», но и дизайн выборки, число уникальных целей, время эксперимента, механизм адаптации, доверительные ограничения и причины прекращения теста. Нельзя уверенно переносить результаты с одного интерфейса, языка или группы людей на весь жизненный цикл модели. Нужна также оценка числа множественных гипотез: при тысячах сравнений часть случайных совпадений ожидаема даже при отсутствии реального раскрытия.

Для юридических выводов применяется отдельная шкала: экспериментально подтверждённый факт, статистическое свидетельство, техническая возможность, разумно вероятный риск и непроверенная гипотеза. Эта шкала предотвращает как чрезмерное

успокоение («не нашли, значит невозможно»), так и преувеличение («один совпавший фрагмент доказывает массовую утечку»).

27. Повторная проверка после обновления данных и моделей

Оценка анонимности не является бессрочной. Новый общедоступный реестр, утечка смежной базы, изменение мощности поисковых систем, расширение API или выпуск нового fine-tuning могут превратить ранее неопасную комбинацию в идентифицирующую. ICO рекомендует пересматривать оценку с учётом развития технологий, новых наборов и изменения модели раскрытия. Аналогичный подход отражён в проекте EDPB 02/2026. [2][15]

Триггерами пересмотра должны быть не только календарные даты, но и изменения условий: передача новому контрагенту, публичное размещение ранее закрытого набора, выпуск дополнительных разрезов статистики, расширение данных до новых лет, объединение с иной системой, инцидент безопасности, появление метода извлечения из модели. Отчёт о переоценке должен ссылаться на предыдущий анализ и объяснять, какие предпосылки перестали действовать.

Для систем ИИ целесообразно вести версионность не только модели, но и пайплайна загрузки, обучающего корпуса, индекса поиска, постпроцессинга и системы логирования. Один и тот же набор весов, подключённый к новому поисковому слою, может существенно изменить фактический риск получения индивидуальных материалов. Запись «модель без изменений» в этом случае вводит в заблуждение.

Рекомендуемая процедура содержит ответственное лицо за мониторинг новых угроз, библиотеку воспроизводимых тестов, журнал исключений и механизм приостановки доступа в случае ухудшения риска. Решение о повторной публикации должно приниматься после завершения теста и документироваться отдельно от эксплуатационного релиза продукта.

Дополнительная экспертная оценка

Акт проверки должен содержать перечень конкретных версий. Смена весов, данных дообучения, токенизатора, политики фильтрации, базы RAG, процедуры чанкинга или инфраструктуры логирования может изменить вероятность раскрытия без видимых изменений для конечного пользователя. Поэтому разработчик обязан определить события, при которых прежние результаты больше не применяются. Это особенно важно для постоянно обновляемых исследовательских платформ. [3]

В регрессионном наборе сохраняются тесты на уникальные нарративы, извлечение контрольных строк, косвенное связывание и вывод чувствительных атрибутов. Отдельная часть предназначена для новых атак, которые не использовались при первичной проверке. Пороговые значения устанавливают в зависимости от исследуемого риска; прежние значения сохраняют для сравнения динамики. Снижение количества обнаружений после обновления ещё не доказывает устранение причины, если одновременно меняются настройки тестирования.

Ответственное лицо ведёт реестр изменений и решение о повторной публикации. Для материалов, затрагивающих уголовное преследование или убежище, до завершения проверки допускается заморозка соответствующего раздела. Отчёт о регрессии должен быть связан с прошлым актом, чтобы читатель мог видеть, какие условия действовали при предыдущем утверждении.

28. Правила публикации: открытая, ограниченная и закрытая среды

Способ раскрытия информации определяет и практическую модель угроз. Публичная выгрузка не позволяет эффективно отозвать уже распространённые копии. Закрытая исследовательская среда с контролируемыми запросами снижает число получателей, но сохраняет юридическую обработку исходных данных и риски изнутри. Передача через доверенного посредника может быть полезна, однако требует договора, контроля полномочий и процедуры на случай нарушения. [12][15]

При выборе режима организации следует задавать не только вопрос «можно ли анонимизировать», но и «необходима ли публикация строк вообще». Во многих случаях общественную подотчётность достигают публикацией агрегированных показателей и описаний методики, при сохранении первичного массива у уполномоченного владельца. Там, где научная проверяемость требует воспроизводимости, можно предоставить проверяющим управляемый доступ с журналами и запретом произвольного вывода.

При возникновении спорного запроса на раскрытие полезно разделить решение на две стадии: право на доступ к исходной информации и пригодность подготовленного результата к передаче конкретному получателю. Даже если публичный интерес высок, это не снимает автоматически необходимости соблюдать требования о законности, минимизации и защите данных при подготовке и передаче.

Режим	Преимущество	Уязвимость / контроль
Публичный open-data набор	Проверяемость и повторное использование	Неотзывность, широкое связывание
Закрытый доступ специалистам	Более детальные исследования	Инсайдер, экспорт, договорные ограничения
Управляемые статистические запросы	Контроль выдачи	Дифференциальные атаки, повторные запросы
Публикация только агрегатов	Низкая детализация о людях	Редкие ячейки, пересечения релизов

29. Документирование происхождения данных и цепочки преобразований

Невозможно убедительно доказать эффективность анонимизации, если неизвестно, какие данные преобразовывались, какие исключались и какие версии оказались у подрядчиков. Provenance-реестр должен описывать первоначальный источник, цель сбора, допустимый круг пользователей, набор идентификаторов, историю объединений и созданные производные материалы. Для каждой операции сохраняются конфигурация, ответственное лицо, контрольная сумма входа и выхода и решение о дальнейшей судьбе промежуточных копий.

Если данные импортируются из нескольких стран или институциональных систем, нужно отдельно учитывать различия правовых режимов и специальные ограничения оригинального источника. Технический статус обезличенного результата не отменяет необходимости законного получения исходника и договорных обязанностей по конфиденциальности. В международных исследованиях проблема особенно заметна, когда анонимизатору фактически передают полный материал для предобработки ещё до оформления его роли и мер безопасности.

При обучении AI-моделей следует вести связь между версиями обучающего корпуса и версиями модели, достаточную для управляемой проверки рисков и реагирования на ошибки. Однако сама система происхождения не должна становиться новым централизованным идентификатором; доступ к связке должен быть минимальным, протоколируемым и ограниченным. Компромисс между прослеживаемостью и конфиденциальностью требует явного управленческого решения.

Практический результат — «паспорт происхождения»: краткая публичная часть с видами источников и условиями обработки и закрытая часть с конкретными файлами, хешами,

разрешениями и юридическими основаниями. Запись без даты, роли, проверяющего и версии недопустима в качестве полноценного доказательного следа.

30. Международная сопоставимость: ЕС, Великобритания и США

В ЕС ключевым критерием является идентифицируемость физического лица по GDPR и его преамбуле 26, уточнённая практикой Суда и методическими документами EDPB. В Великобритании ICO применяет собственное руководство к UK GDPR и анализирует identifiability с учётом «мотивированного нарушителя». Руководство ICO, опубликованное в марте 2025 года, на октябрь 2026 года содержит предупреждение о пересмотре в связи с Data (Use and Access) Act. Его нельзя без оговорок описывать как буквальную реализацию норм ЕС. [1][2][15]

В США существуют отраслевые режимы и технические рекомендации. Например, HIPAA предусматривает специфические методы de-identification в области охраняемой медицинской информации, а NIST SP 800-188 описывает методы и управление устранением идентифицирующих связей в государственных наборах. Эти подходы могут давать полезные инженерные инструменты, но сертификат соответствия одному режиму не равен автоматическому признанию анонимности по GDPR. [12]

Для международного исследовательского обмена рекомендуется матрица сопоставления требований: применимое право, способ проверки идентифицируемости, допустимый получатель, исходный юридический режим, передаваемые артефакты, возможность дальнейшего раскрытия и порядок уведомления об ошибке. Отдельно нужно проверять, какие данные подлежат трансграничной передаче до анонимизации. Если передача касается персональных данных, правила GDPR о передаче в третьи страны анализируются независимо от планируемого конечного результата.

Организация, работающая одновременно с европейскими клиентами и международными техническими поставщиками, должна избегать одного глобального маркетингового утверждения «fully anonymised». Более точная формулировка указывает конкретную версию, назначение, модель доступа и юрисдикцию, для которых проведена оценка. Это делает вывод узким, но честным и юридически более устойчивым.

Дополнительная экспертная оценка

Международные публикации часто сравнивают данные организаций с разными критериями и процедурами анонимизации. Если одни источники считают деидентификацией удаление прямых идентификаторов, а другие требуют проверки косвенного связывания, механическое объединение результатов создаёт ложное впечатление сопоставимости. Методический протокол обязан описать фактическую операцию преобразования и её проверку, а не только указанное в источнике юридическое название. [1][15]

Для каждого массива указывают применимый закон, статус исходных данных, сторону, определяющую цели и средства, конечных получателей и дополнительные ограничения. Особое внимание уделяют будущему использованию: передача для расчёта одной статистики не даёт сама по себе разрешения на обучение модели, публикацию корпуса или свободное присоединение к коммерческим наборам. Если одни данные подпадают под американский отраслевой режим, нужно установить его конкретные требования, а не говорить об абстрактном «законе США об анонимизации».

Исследовательская задача может выполняться поэтапно: сначала публикуется несопоставимость критериев, затем проводится новая проверка для согласованной методологии, после чего строится статистическое сравнение с интервалами

неопределённости. Это медленнее, чем простая сводная таблица, но снижает риск неверного международного вывода.

31. Защита субъектов и реализация их прав

Даже когда организация стремится создать анонимный результат, исходные операции с персональными данными могут затрагивать права на информирование, доступ, исправление, удаление, ограничение обработки и возражение. Применимость каждого права зависит от правового основания, обстоятельств и предусмотренных законом исключений. Нельзя объявить эти права несущественными на том основании, что в будущем результат исследования будет опубликован без имён. [1]

Статья 11 GDPR предусматривает особые положения в ситуациях, когда цели обработки не требуют или больше не требуют идентификации субъекта: контролёр не обязан сохранять или получать дополнительные данные исключительно ради идентификации. Однако это положение не разрешает произвольно уничтожать учётные записи и тем самым обходить законные запросы. Необходимо оценить, действительно ли идентификация больше не нужна для целей, и можно ли связать обращение субъекта с сохранившейся информацией без создания новой непропорциональной системы учёта.

Если обнаружено, что якобы анонимный набор всё же позволяет идентифицировать человека, организация должна пересмотреть юридическую квалификацию, ограничить риск дальнейшего раскрытия и оценить применимые обязанности по защите данных. Операционная последовательность может включать приостановку публичного API, исправление документации, запрос удаления у контролируемых получателей, уведомление затронутых лиц в предусмотренных законом случаях и дополнительную независимую оценку. Нельзя обещать технический отзыв всех скачанных копий открытого массива.

Для ARGA целесообразна отдельная процедура реагирования на обращения людей, считающих, что они узнаваемы в публикации. Должны существовать понятный контакт, безопасный способ представить подтверждение, срок внутреннего рассмотрения, запрет на неоправданное требование избыточных документов и возможность независимой эскалации спорного решения.

Дополнительная экспертная оценка

Организация иногда заявляет, что не может установить лицо в опубликованном результате, но при этом сохраняет оригинальные персональные данные и ключ сопоставления. Такая ситуация требует анализа каждой операции отдельно. Статья 11 GDPR регулирует обработку, при которой идентификация не требуется для целей контролёра, но не является общим освобождением от обязанности отвечать на запросы при наличии достаточной информации для установления связи. [1]

Правовой регламент должен содержать процедуру обращения, позволяющую доказать связь с публикацией без необоснованного запроса дополнительных документов. Ответ включает информацию о фактической возможности идентификации, применимом режиме, принятых ограничениях и порядке надзорного обжалования. В отношении отдельного человека недопустимо утверждать, что он не обладает правами, лишь на основании того, что набор назван «анонимным» в техническом паспорте. Если выявлена возможность восстановления данных, необходимо повторно рассмотреть правовой статус и реакцию.

Для исследовательских проектов могут существовать предусмотренные законом особенности и ограничения отдельных прав, но они зависят от фактической цели и соответствующих норм; ссылка на научное исследование без правового основания не

достаточна. Обеспечение прав и минимизация данных должны быть частью архитектуры исследования, а не дополнительной обязанностью после публичного спора.

32. Нарушения безопасности и риск несанкционированного раскрытия

Ошибочная анонимизация и киберинцидент являются различными, хотя могут пересекаться. Ошибочная юридическая квалификация опубликованных данных способна вызвать незаконное раскрытие персональных сведений без взлома. Взлом защищённого хранилища исходных данных может затронуть персональные сведения даже при корректной анонимной публичной копии. Поэтому организации нужны отдельные контуры оценки приватностного результата и информационной безопасности. [1][19]

При обнаружении возможности обратной идентификации необходимо собрать факты: какие данные были доступны, с какого момента, кто получил доступ, сколько записей затронуто, какие сведения реально раскрывались и каковы последствия. Если инцидент квалифицируется как нарушение безопасности персональных данных, сроки и условия уведомлений определяются статьями 33–34 GDPR и применимым правом. Не каждый технический дефект автоматически создаёт обязанность публичного уведомления, но и отсутствие доказанной эксплуатации не всегда означает отсутствие риска для прав и свобод.

Сохранение цифровых доказательств должно сочетаться с минимизацией вторичного ущерба. Логи, контрольные суммы и скриншоты результатов не следует размещать в общедоступном приложении к докладу; для внутреннего расследования готовится отдельный ограниченный архив. Если тест случайно извлёк реальное имя клиента или номер дела, дальнейшее копирование примера в презентации увеличивает инцидент.

Система контроля должна предусматривать дежурное лицо, критерии приостановки распространения, компетенцию юридической команды, документированное решение о необходимости уведомлений и корректирующие меры для повторной публикации. Роль независимого аудитора — установить факт, риск и проверенность исправления; решения о процессуальных уведомлениях принимает соответствующий контролёр с юридической оценкой.

Дополнительная экспертная оценка

Ошибочную анонимизацию и нарушение безопасности персональных данных необходимо квалифицировать отдельно. Если контролёр обнаружил, что опубликованный материал позволяет связать сведения с физическим лицом, он устанавливает статус данных и обстоятельства раскрытия. Затем оценивает, возникло ли нарушение безопасности персональных данных в смысле GDPR, какие риски оно создаёт и какие юридические обязанности следуют из фактов. Нельзя автоматически обещать уведомление в 72 часа для любого статистического сигнала, но и нельзя исключить его без документированной оценки. [1]

Технический план включает сохранение доказательств, приостановку доступа при необходимости, выявление всех копий и связанных наборов, безопасное исправление и повторный тест. Для публичного контента важно считать остаточный риск: архивные версии, поисковые индексы и скачанные копии могут существовать независимо от действий первоначального издателя. Поэтому корректирующий акт честно отделяет действия, совершённые организацией, от результата, который нельзя гарантировать полностью.

Внутренний журнал событий хранит время обнаружения, наблюдаемые признаки, оценку риска, принятое решение и ответственных специалистов. Это помогает доказать разумность

реакции и одновременно позволяет пересмотреть ложные тревоги, не превращая каждый технический инцидент в публичную кризисную публикацию.

33. Применение в международном розыске и экстрадиционных делах

Досье о международном розыске часто состоят из уникальных сочетаний фактов: гражданство, маршрут перемещения, год процессуального решения, обстоятельства задержания, судебная инстанция и язык документов. Такая комбинация может идентифицировать человека быстрее, чем имя. Публикация исследовательского кейса с удалённой фамилией не гарантирует анонимности, если любой читатель может найти постановление или новостную публикацию по сохранённой фабуле.

Для аналитической работы ARGA рекомендуется двухуровневая модель. Закрытый уровень хранит исходный доказательственный пакет при надлежащем правовом основании и режиме конфиденциальности. Публичный уровень описывает типовую правовую проблему, не воспроизводя единственное сочетание событий. В ряде случаев безопаснее представить обобщённый составной учебный сценарий, явно помеченный как гипотетический, чем «косметически обезличенный» настоящий случай.

Проверка публичного кейса должна имитировать обычный поиск в официальных реестрах и открытых источниках по сочетанию оставленных признаков. Если лицо находится за пределами государства происхождения, дополнительные сведения о месте проживания и расписании процедур могут усиливать не только приватностный, но и физический риск. Поэтому критерий допустимости раскрытия нельзя сокращать до числового процента реидентификации.

Обратная проблема — слишком сильное обобщение, после которого исчезает правовой смысл нарушения. В этом случае исследователь должен выбрать иную форму публикации: сравнение процедур и стандартов, агрегированную статистику или закрытую правовую экспертизу, а не создавать ложное впечатление о реальном судебном деле, изменив его существенные факты.

Дополнительная экспертная оценка

В делах, связанных с международным розыском, опасны редкие комбинации правовых фактов. Судебное решение, возраст, дата задержания, страна убежища и конкретное обвинение могут образовать уникальный профиль без фамилии. Если исследование опирается на известный публичный кейс, обычная замена географических названий зачастую недостаточна: последовательность событий останется узнаваемой при поиске или семантическом сопоставлении. [1][4]

Для внутреннего правового анализа допустимо хранить полную цепочку доказательств при надлежащем основании и ограничениях. Для публичной версии нужно решить, требуется ли индивидуальный рассказ вообще. Периодические сводные показатели, методические конструкции и гипотетические кейсы часто позволяют продемонстрировать системную проблему с меньшим риском раскрытия. Если индивидуальные обстоятельства необходимы для публичного обсуждения нарушения, нужно отдельно рассмотреть правомерность публикации и возможность получения законного согласия или иного основания, не подменяя этот анализ термином «анонимизация».

Надзор и редакция разделяются. Юридический эксперт отвечает за корректность процессуальных формулировок, специалист по приватности — за идентифицируемость, независимый рецензент — за возможность повторения вывода. Разные компетенции предотвращают ситуацию, когда формально верная юридическая история причиняет новый вред её участнику.

34. Применение в санкционном и AML/KYC-комплаенсе

Санкционный скрининг, AML/KYC и обработка сообщений о подозрительной активности создают специфические риски ложной атрибуции. Совпадение по распространённому имени и дате рождения не должно автоматически становиться публичным утверждением о санкционном статусе человека. Анонимизация тестовых корпусов алгоритмов скрининга должна скрывать как прямые идентификаторы, так и редкие маршруты платежей, комбинации компаний, временные метки и причины внутреннего расследования.

При публикации показателей false positive/false negative рекомендуется указывать границы выборки и способ установления истинного статуса. Если набор включает дела, связанные с уголовными правонарушениями, применимость статьи 10 GDPR и национального регулирования подлежит отдельному анализу. Наличие у внешнего поставщика «de-identified training data» не доказывает, что банк вправе передать ему свои журналы комплаенса без правовой и договорной оценки. [1]

Матрица рисков должна рассматривать два независимых ущерба: ошибочное раскрытие настоящей чувствительной связи и создание ложного сходства, при котором невиновного человека ошибочно связывают с отдельным санкционным субъектом. При подготовке синтетического набора для проверки алгоритма нельзя переносить вымышленные записи в оперативную базу без специальных маркеров и изоляции.

Рекомендуемый путь независимой проверки — тестирование модели на контролируемом наборе, подготовленном по согласованной схеме, с обезличенными результатами по категориям ошибок. Если необходимо изучить оригинальные спорные совпадения, они должны передаваться конкретно уполномоченным экспертам, а не включаться в открытый отчёт. Сохранность исходных доказательств и безопасность публикации оцениваются раздельно.

Дополнительная экспертная оценка

Санкционные проверки и AML/KYC-системы создают риск ошибочного совпадения, когда лицу приписывается деятельность или статус другого человека. Анонимизация исследовательского отчёта должна предотвращать не только раскрытие, но и ложное связывание с конкретным контрагентом. Даже если имя удалено, сумма, отрасль, дата блокировки и название финансового инструмента могут оставаться достаточными для установления организации и физического лица. [1]

Исследователь проверяет, сколько возможных лиц соответствует публикуемому профилю, есть ли единичные сделки, позволяет ли агрегирование восстановить конкретное событие, а также каковы последствия ошибки при восприятии читателем. Если кейс показывает false positive, следует избегать повторения старого ложного обвинения в такой форме, которая продолжит распространяться через поисковые системы и обучающие корпуса. При этом сам вывод о системном дефекте должен остаться проверяемым через защищённый набор доказательств.

Дополнительно оцениваются отношения между юридическим лицом и бенефициаром. Графы транзакций могут оставаться идентифицирующими даже при замене имён случайными кодами, если структура крупных переводов публично известна. Отдельный тест проверяет структурную уникальность графа и возможность связывания с реестрами бенефициаров.

35. Миграция, убежище и решения об уязвимых лицах

Материалы о международной защите и миграции особенно чувствительны к внешнему связыванию: конкретный маршрут, место задержания, время обращения, обстоятельства политического или иного преследования и информация о членах семьи иногда образуют уникальный отпечаток. Даже если опубликованный текст не раскрывает имени, он может позволить заинтересованным лицам распознать заявителя или его родственников. Поэтому статистическая оценка анонимности должна дополняться анализом угроз безопасности и вероятного вреда.

Внутренний набор для анализа миграционных алгоритмов следует минимизировать исходя из тестируемого правового вопроса. Если исследуется доля отмен решений по процедурным основаниям, подробные описания личных историй часто не нужны. Если необходим сравнительный анализ индивидуальных мотивировок, более уместен режим закрытого исследования с надёжным контролем доступа и публикацией обезличенных выводов, проверенных поиском по внешним источникам.

Условный кейс: публикация сообщает возраст, конкретный месяц эвакуации, страну транзита, редкую профессию и дату отказа в экстрадиции. Каждая характеристика может казаться общей, но совокупность способна выделить одного человека. Без проверки внешних реестров и новостей вывод об анонимности будет необоснованным. Дополнительная проверка затрагивает упоминания детей, лиц, предоставлявших помощь, и иных третьих сторон, не являющихся заявителями.

Рекомендация ARGA — применять повышенный уровень редакционного контроля для материалов о лицах под риском преследования: независимое согласование текста, проверку вспомогательных идентификаторов, отказ от публикации опасных деталей и процедуру срочного снятия материала при обоснованной жалобе. Согласие на публикацию отдельного факта не заменяет оценки законности и безопасности всего набора.

Дополнительная экспертная оценка

В миграционных и гуманитарных исследованиях последствия раскрытия могут быть особенно тяжёлыми. Редкие истории преследования, маршруты выезда, медицинские данные и состав семьи позволяют установить человека даже без прямых идентификаторов. При этом исключение таких эпизодов из статистики ведёт к систематическому невидению наиболее уязвимых групп. Следовательно, задача не решается простым удалением всех малочисленных категорий. [1][15]

Надёжная процедура сочетает ограниченный доказательственный доступ, более общие публичные показатели и независимую оценку того, насколько анонимизация исказила реальную картину нарушений. В исследовании указывается, какие типы случаев стали ненаблюдаемыми вследствие защиты, и можно ли компенсировать это качественным методическим описанием без раскрытия личности. Особенно важно не публиковать комбинацию страны происхождения, уникальной судебной даты и редкого заболевания, если эти признаки уже содержатся в локальных новостях.

Для детей и уязвимых взрослых применяется повышенная осторожность, но конкретные правовые последствия определяются применимым режимом и обстоятельствами. Гарантировать безопасность на основании одного статистического теста недопустимо. Решение о публикации может быть отрицательным даже при удовлетворительных агрегированных метриках, если прогнозируемый индивидуальный вред несоразмерен исследовательской ценности.

36. Методика независимого аудита анонимности для Observatoire ARGA

Предлагается шесть последовательных стадий независимой проверки. Первая — определить объект: конкретный файл, статистический механизм, эмбеддинги, модель или совокупную систему. Вторая — установить юридическую рамку: применимые нормы, цели, категории данных, роли и модель раскрытия. Третья — составить карту акторов и вспомогательной информации. Четвёртая — выбрать и зарегистрировать тесты No Record Isolation, No Linkage и No Inference, дополнив их модельно-специфическими атаками. Пятая — оценить остаточный риск, неопределённость и полезность. Шестая — принять и задокументировать ограниченный вывод с условиями повторного рассмотрения. [2][4]

Заключение должно содержать как минимум следующие поля: идентификатор проверенной версии, дата заморозки объекта, правовое основание доступа, перечень тестируемых участков, модели противников, параметры испытаний, характеристики выборок, таблица результатов, экспертные ограничения и решение о раскрытии. Публикация, ограниченная словами «данные надёжно анонимизированы», непригодна для независимой проверки. Нужен воспроизводимый след, который не раскрывает защищённые данные в открытой части.

Решение аудитора	Когда применяется	Условие повторной оценки
Обоснованно анонимно в определённом контексте	Есть проверяемый пакет доказательств	Изменение объекта, адресата или угроз
Персональные / псевдонимизированные данные	Идентифицируемость сохраняется	Успешная дополнительная обработка
Данных недостаточно для заключения	Нет тестов или сведений о получателе	Предоставление недостающих доказательств
Публикация недопустима в заявленном виде	Существенный неустранённый риск	Иной режим доступа или новая версия

Независимый аудитор не заменяет орган по защите данных и не выдаёт официальное разрешение на обработку. Observatoire может формировать исследовательское заключение и рекомендации о снижении рисков, но юридическое решение о соблюдении GDPR и меры правоприменения принадлежат уполномоченным субъектам. Коммерческая заинтересованность в результатах должна раскрываться, а методика — проходить внутреннюю научную рецензию.

Дополнительная экспертная оценка

Методика должна содержать не меньше шести стадий: юридическое определение объёма проверки; инвентаризацию источников; построение модели угроз; проведение технических испытаний; независимую рецензию; публикационное решение с обязанностью повторной оценки. Отсутствие любой стадии отражается в отчёте как ограничение, а не скрывается в общих словах об успешной проверке. [1][2]

Аудитор не присваивает универсальный знак «анонимно», если не имеет сведений обо всех релевантных получателях или об обучении модели. Вместо этого формулируется контекстное заключение: какой результат, для какой цели и среды, при каких средствах идентификации исследован. Доказательства воспроизводимости и границы доступа перечисляются отдельно. Если данные пригодны только для защищённой среды, такой вывод признаётся полноценным результатом проверки, а не неудачей аудита.

Для ARGA важно институционально разделить исследование, подготовку консультаций и коммерческое сопровождение конкретных дел. Лицо, участвовавшее в защите клиента, может предоставить исходный материал на законном основании, но не должно без независимой процедуры единолично утверждать публикационную безопасность

собственного кейса. Решение обнародовать отрицательный результат не должно зависеть от дальнейшего коммерческого интереса.

37. Открытая консультация EDPB: предложения Observatoire ARGA

Поскольку Guidelines 02/2026 в октябре 2026 года находятся на стадии консультации, ARGA может сформировать адресные комментарии к методике без утверждения, что её положения окончательно приняты. Обоснованный вклад должен быть связан с предметом проекта: тест идентифицируемости, контекстуальный подход, три критерия, дополнительные сведения, доступные третьим лицам, и способы документирования. Предложения в настоящем разделе являются исследовательской позицией и не означают, что комментарии уже направлены регулятору. [2]

Первое предложение: дополнить практические примеры высокомерными текстовыми и графовыми данными, поскольку выделение одного лица может быть результатом сочетания нарративных и структурированных атрибутов. Второе: конкретизировать процедуру оценки последовательных публикаций и внешних обновлений, чтобы контролёр не подтверждал анонимность нового выпуска независимо от ранее опубликованной статистики. Третье: рекомендовать отдельно отражать доступные государственные и судебные реестры, если публикация касается уголовно-процессуальной или санкционной информации.

Четвёртое предложение: включить пример оценки модели ИИ, обученной на персональных данных, с разграничением анонимности исходного набора, весов модели, retrieval-индекса и конечного ответа. Пятое: предусмотреть минимальную спецификацию доказательств для заявлений об анонимности, включая версию, тип угроз, состав тестов, отрицательные результаты и механизм повторной проверки. Шестое: пояснить, как учитывать риск ложной атрибуции, когда публикация приводит не к раскрытию истинной идентичности, а к ошибочному приписыванию человеку чувствительного факта.

Предложения не следует подавать как закрытую критику регулятора. Проект 02/2026 уже содержит значительный методологический аппарат; задача ARGA — предложить прикладные сценарии и критерии воспроизводимости для специальных сфер. При реальном участии в консультации необходимо сверить актуальное окно подачи и требования официальной формы, сохранить отправленную редакцию и подтверждение регистрации.

Дополнительная экспертная оценка

Предлагаемые замечания должны быть сформулированы как исследовательская позиция, а не как описание уже проведённого регулятором обязательного теста. Обсуждаемый проект 02/2026 остаётся консультационным до окончания открытого периода, а окончательная редакция может отличаться. В примерах важно отдельно разграничить правовой стандарт идентифицируемости и вспомогательные количественные показатели; никакая одна метрика не заменяет фактический анализ средств, разумно вероятных для конкретного участника. [2]

Содержательно Observatoire может предложить три уточнения. Первое — иллюстративные примеры семантического связывания юридических нарративов, где уникальная последовательность событий восстанавливает личность без прямых идентификаторов. Второе — специальную методику для релизов, состоящих из многих версий и перекрывающихся агрегатов, включая вторичное подавление. Третье — описание того, как технические испытания AI-модели отделяются от юридического вывода о персональном характере данных. Эти предложения следует представлять как открытые вопросы для консультации, а не как утверждения о пробелах в действующей обязательной норме.

Перед направлением комментариев проводится проверка самой формы подачи, правомерности раскрытия метаданных и согласование официальной позиции ARGA. Публичное представление исследовательской точки зрения не должно включать реальные сведения о клиентах без отдельного анализа необходимости и законности.

38. Пилотная программа проверки для международных исследовательских проектов

Пилот целесообразно разделить на три независимых набора: табличный набор с квазиидентификаторами, корпус кратких юридических текстов и тестовую систему retrieval/генерации ответов. Для начальной стадии рекомендуется использовать искусственные данные и материалы, заранее разрешённые к анализу; реальное клиентское досье не является приемлемым тестовым сырьём по умолчанию. Каждый набор должен иметь известные исходные свойства, чтобы измерение риска не зависело от догадок о настоящих лицах.

В табличном наборе проверяются групповые признаки, редкость строк и связывание с открытым искусственным реестром. В текстовом — способность обнаруживать уникальные сюжеты и скрытые метаданные. В модельном — запоминание и выдача фрагментов, возможные выводы о принадлежности и риск retrieval-утечек. Тесты предрегистрируются, результаты собираются раздельно, затем оценивается применимость к юридическим требованиям и целям исследования.

Критерии перехода от пилота к реальным данным включают независимую правовую оценку, назначение ответственного за защиту данных при необходимости, DPIA там, где она требуется, договорный режим и проверку компетенций технической команды. При отсутствии условий организация должна ограничиться синтетическими материалами и методическими публикациями. Решение «пилот успешен» не означает разрешения на масштабирование без переоценки новой среды.

Дорожная карта условна: этап 1 — подготовка методики и искусственного стенда; этап 2 — воспроизводимое испытание; этап 3 — независимая рецензия; этап 4 — публикация ограниченного отчёта и исправление слабых мест. Для каждого этапа назначается владелец решения, продукт, риск и основание остановки. Сроки устанавливаются отдельно после подтверждения ресурсов; в докладе не имитируется уже проведённое исследование.

Дополнительная экспертная оценка

Пилот рекомендуется начинать с ограниченного числа разнородных наборов: табличная статистика, текстовый корпус, набор с временными и пространственными признаками, RAG-система и публикационный кейс с малочисленной группой. Выбор должен учитывать доступность законных тестовых материалов, измеримость результата и наличие независимых проверяющих. Не следует начинать программу с самых чувствительных дел, где ошибка сама по себе создаст высокую вероятность вреда. [1][2]

На первом этапе оформляется техническое задание с критериями успеха и отказа. На втором — проводится тест базовых методов анонимизации и сравнительная оценка изменений в полезности данных. На третьем — независимая рецензия и публикационное решение. Для каждого набора сохраняются открытый методический паспорт и закрытые доказательства; их доступность организуется соразмерно рискам. Если при пилоте выявляется связь с живым лицом, испытание прекращается, информация не помещается в публичный отчёт, а уполномоченные специалисты оценивают последующие меры.

Оценка успеха пилота строится не по числу выданных положительных сертификатов, а по качеству выявленных ограничений, полноте документации, воспроизводимости оценок и

способности принять обоснованное решение «не публиковать». Такой критерий снижает стимул к формальному подтверждению желаемого заказчиком результата.

39. Институциональные гарантии независимости и предотвращение конфликта интересов

В случае когда Observatoire ARGA осуществляет научную экспертизу и существует институциональная связь с консультационной деятельностью, необходимо обеспечить разграничение коммерческого заказа и исследовательского вывода. Наличие оплачиваемого клиента не делает заключение недействительным само по себе, но создаёт риск выбора методики под интересы заказчика. Документ должен предусматривать раскрытие источников финансирования, независимого рецензента и право публикации критического результата в допустимых правовых пределах.

Проект системы управления включает запрет на предварительное обещание результата, регистрацию методологии до теста, независимое утверждение поправок к протоколу, контроль конфиденциальности, механизм отвода эксперта и отдельную процедуру жалобы. В технических экспертизах полезно вести реестр доступа и версий, чтобы участник, заказывающий тест, не мог незаметно подменить объект после проверки.

Публикационная политика должна предусматривать три вида документа: методический обзор, отчёт об испытании и правовое заключение по конкретному контексту. Их нельзя обозначать одним словом «сертификация». Если ARGA не является аккредитованным органом оценки соответствия для соответствующей схемы, публикация не должна создавать обратного впечатления. Отдельная юридическая рецензия должна проверять как вывод, так и описание правового статуса самой процедуры.

Устойчивость системы достигается не размером итогового PDF, а проверяемостью решений: понятно, что проверено, какими средствами, в каком контексте, что осталось неизвестным и кто несёт ответственность за дальнейшее использование. Этот принцип применяется к каждому приложенному акту и краткому публичному сообщению.

Дополнительная экспертная оценка

Заявление о независимости должно подтверждаться конкретными процедурами. Редакционный совет раскрывает источники финансирования, устанавливает правила отвода участников, документирует предварительные договорные отношения с организациями, чьи системы оцениваются, и обеспечивает канал оспаривания фактических выводов. Независимость от государственного органа не равнозначна отсутствию зависимости от частного заказчика. [2]

Если исследовательская организация одновременно оказывает платные услуги в смежной области, нужно исключить возможность изменения научного вывода ради коммерческого результата. Проверяющие, участвовавшие в подготовке защищаемого юридического кейса, не должны единолично принимать решение об анонимности его публикационной версии. Положительный и отрицательный результаты документируются одинаково, а право на публикацию методической критики не может целиком зависеть от согласия оцениваемой стороны.

Процедура исправления ошибочного публичного вывода также является частью независимости: организация публикует корректировку с указанием предмета и даты, но не раскрывает новые личные сведения в самой поправке. Гарантией качества является не

отсутствие ошибок, а способность обнаруживать и исправлять их без необоснованного ущерба для людей.

40. Ограничения исследования и условия применения выводов

Настоящий доклад представляет анализ норм, официальных рекомендаций и возможной методологии независимой проверки. Он не содержит результатов доступа к закрытым AI-моделям, тестирования систем компаний либо исследования реальных реестров международного розыска. Все числовые примеры в аналитической части иллюстративны и не должны цитироваться как статистика правоприменения или фактическая частота раскрытия в какой-либо отрасли.

Результаты технической проверки анонимности зависят от модели угроз, времени, среды и объёма доступной эксперту информации. Отсутствие успешной атаки в пилотном эксперименте не равно невозможности будущего извлечения сведений. Рекомендации NIST, CNIL и EDPB имеют разную юридическую силу и назначение; их нельзя складывать в единый список прямых законодательных обязанностей. [2][4][12][13]

Отдельная неопределённость связана со статусом Guidelines 02/2026 и Guidelines 01/2025, доступных на официальных ресурсах EDPB как проекты соответствующих процедур консультации. До использования доклада в подаче необходимо проверить, появились ли финальные версии и какие изменения они принесли. Правовой анализ основывается на доступных источниках по состоянию на 22 октября 2026 года.

Если проект будет использован для судебного спора или взаимодействия с национальным органом, необходимо дополнительное изучение применимого процессуального права, особенностей конкретной юрисдикции и фактов. Обобщённая методика не является индивидуальным юридическим советом и не заменяет квалифицированной оценки доказательств по конкретному делу.

41. Выводы и институциональные рекомендации

Первый вывод: анонимизация не является операцией замены имён; это проверяемая правовая и техническая оценка фактической идентифицируемости в определённом контексте. Второй: подход EDPB 02/2026 полезен для построения воспроизводимой проверки, но остаётся проектом на стадии консультации. Третий: псевдонимизация, шифрование, синтез данных, федеративное обучение и дифференциальная приватность снижают различные риски и не дают одинаковых юридических результатов. [1][2][4]

Четвёртый вывод: мощные AI-модели, эмбеддинги и retrieval-системы требуют собственной проверки извлечения, запоминания, принадлежности и связывания сведений, а не переноса статуса обучающего набора. Пятый: публичная публикация должна анализироваться вместе с уже выпущенными данными и доступными внешними источниками. Шестой: в чувствительных юридических контекстах дополнительным критерием выбора режима доступа является тяжесть возможного вреда, в том числе ложного приписывания уголовного или санкционного статуса.

Observatoire ARGA рекомендуется утвердить стандарт доказательственного паспорта анонимизации, форму независимого протокола, реестр версий и внешних релизов, матрицу угроз, порядок безопасного тестирования и процедуру повторной проверки. Для применения на реальных материалах требуется отдельное решение о правовой основе, роли независимого аудитора, защите данных и доступе к первичному досье. Ни один внутренний шаблон не должен называться «сертификатом GDPR».

Предлагается подготовить отдельную публичную позицию по EDPB Guidelines 02/2026 с конкретными сценариями для судебных, санкционных и миграционных данных. Приоритет — воспроизводимость и соразмерность: защищать людей и одновременно сохранять возможность

качественной международной аналитики, не выдавая формальные процедуры за доказательство отсутствия риска.

Приложение 1. Словарь ключевых терминов для междисциплинарной проверки

Термины ниже используются как рабочая лексика исследования. Их следует сопоставлять с определениям GDPR и контекстом конкретного нормативного документа; перевод технического понятия не всегда создает самостоятельную юридическую категорию. [1][2]

Термин	Значение в исследовании	Предостережение
Идентифицируемость	Возможность отличить и связать информацию с лицом разумно доступными средствами	Зависит от контекста
Псевдонимизация	Отделение атрибуции к лицу от основного набора с отдельной защитой связки	Не равна анонимности
Record isolation	Выделение уникальной записи или профиля	Не автоматически доказывает истинную личность
Linkage	Связывание различных наборов и наблюдений	Возможно без имени
Inference	Вывод сведений о лице из других признаков	Точность и вред оцениваются раздельно
Regurgitation	Близкое воспроизведение обучающего материала	Не всякий повтор содержит персональные данные
Membership inference	Вывод об участии конкретной записи в обучении	Зависит от базовой частоты
DP budget	Учёт математической приватностной потери по механизму	Не универсальный правовой порог

Для внешних документов рекомендуется приводить исходный английский термин при первом употреблении. Если в проекте EDPB 02/2026 определения будут изменены по итогам консультации, терминологический словарь необходимо пересмотреть вместе с матрицей тестов.

Приложение 2. Дерево первичного юридического решения

Это дерево предназначено для первого отбора вопросов и не подменяет правовое заключение. Оно требует документированного ответа на каждую развилку, особенно если квалификация меняется после раскрытия данных третьему лицу. [1][2]

Шаг	Проверочный вопрос	Следующее действие
1	Относится ли информация к физическому лицу по содержанию, цели или эффекту?	Если нет, зафиксировать основания; если да, далее
2	Идентифицировано ли лицо или может ли оно быть выделено?	Оценить разумно вероятные средства
3	Кто владеет дополнительной информацией и может ли получить её получатель?	Составить карту участников и доступа
4	Проходит ли набор No Record Isolation, No Linkage и No Inference?	Зафиксировать результаты отдельно
5	Если тест не пройден, приводит ли это к идентифицируемости?	Дополнительная контекстная оценка
6	Для какого режима и даты действует вывод?	Ограниченная формулировка и триггеры пересмотра

Решение «анонимно» должно сопровождаться версией объекта и указанием проверенной аудитории. Если документы не позволяют установить модель внешней информации, допустимый вывод — «недостаточно данных для подтверждения анонимности», а не автоматически отрицательный или положительный ответ.

Приложение 3. Паспорт массива перед обработкой

Заполнение паспорта предшествует передаче исходных сведений специалистам по очистке. Оно позволяет оценить юридические предпосылки и определить, какие технические процедуры вообще допустимы. Для массивов с угрозой вреда важно отдельно указать уязвимые группы и существующие обязанности хранения доказательств.

Поле паспорта	Требуемая запись
Идентификатор, версия, дата	Неизменяемый код и ответственная группа
Категории источников	Происхождение, юрисдикция, срок актуальности
Цель и основание	Конкретная операция и применимая норма
Категории субъектов	Включая третьих лиц и малые группы
Особые категории / статья 10	Отдельная правовая оценка
Режим будущего доступа	Публикация / закрытая среда / агрегаты
Дополнительные идентификаторы	Ключи, справочники, внешние ссылки
Сроки уничтожения и архив	Первичный оригинал, промежуточные копии

Паспорт не должен воспроизводить реальные персональные данные; к нему прилагаются указатели на защищённое хранилище. Изменение цели публикации требует обновления паспорта, а не простой замены названия итогового PDF.

Приложение 4. Реестр правовых оснований и специальных ограничений

Объектом регистрации является каждая существенная операция обработки, а не «проект в целом». Для одной исследовательской программы сбор, трансформация, тестирование и публикация могут иметь разные получателей и правовые условия. Приведённые ниже графы требуют заполнения применительно к конкретному делу. [1]

Операция	Предмет правовой оценки	Обязательное свидетельство
Первичный сбор	Статьи 5 и 6, источник, цель	Описание законности и прозрачности
Преобразование	Необходимость доступа к оригиналам	Регламент операции, роли, безопасность
Использование специальных данных	Статья 9 и применимые условия	Обоснование исключения/основания
Обработка данных о правонарушениях	Статья 10 и национальные нормы	Соответствующее полномочие и гарантии
Экспорт подрядчику	Роли, договор, трансграничная передача	Договор и оценка маршрута данных
Публикация	Анонимность объекта для аудитории	Итоговый протокол, решение владельца

Внутренний рецензент отмечает отдельным флагом случаи, где публикация юридически возможна, но риск безопасности для субъекта остаётся неприемлемым. Для особых режимов хранения доказательств нельзя использовать автоматическое уничтожение как средство обхода обязанности сохранения.

Приложение 5. Матрица угроз и компетенций нарушителя

В матрице должны быть указаны возможности, а не характеристики реальных подозреваемых. Это поддерживает воспроизводимость тестов и уменьшает риск произвольного предположения о «безопасном получателе». Модель по мотивированному нарушителю ориентирована на рекомендации ICO, с обязательной поправкой на юрисдикцию проекта. [15]

Категория	Пример сведений	Проверяемый риск	Условие остановки
Обычный пользователь	Новости и поисковая выдача	Поиск уникального события	Обнаружено прямое совпадение
Специализированный аналитик	Публичные реестры, отраслевые базы	Графовое связывание	Выделены чувствительные профили
Получатель по договору	Собственные клиентские журналы	Совмещение полей	Допуск позволяет атрибутировать
Внутренний сотрудник	Справочник, логи, ключ	Обход разграничения	Воспроизводимая эскалация доступа
Нарушитель после утечки	Часть чужой базы	Историческое связывание	Появилась новая идентифицируемость

Для каждого сценария протокол фиксирует не только потенциальный доступ, но и его действительную разумную вероятность. Упоминание гипотетического всемогущего противника без оценки затрат не соответствует необходимому контекстному подходу.

Приложение 6. Карта вспомогательных данных и внешних источников

Одним из типичных недостатков оценки является отсутствие списка доступных сторонних массивов. Для независимого аудита необходимо сформировать перечень открытых реестров, архивных публикаций, наборов самого контролёра и возможных источников, предоставляемых партнёром, с указанием условий доступа. Конкретные клиентские ссылки не включаются в публичную редакцию.

Категория источника	Риск совмещения	Метод проверки
Судебные и административные решения	Уникальная фабула и дата	Поиск редких сочетаний
Корпоративные сведения	Должность и граф связей	Сопоставление ролей и событий
Географические и временные публикации	Маршрут и присутствие	Анализ временной точности
Старые отчёты организации	Дифференциальное раскрытие	Анализ пересечений релизов
Открытые медиаматериалы	Фото, цитаты, события	Семантическое сопоставление
Коммерческие справочники	Фирменные идентификаторы	Доступность по роли адресата

Изменение доступности любого ключевого источника запускает пересмотр. Даже недавний отчёт можно считать устаревшим, если после его подготовки появился новый открытый документ, позволяющий надёжно связать ранее обезличенные записи.

Приложение 7. Рабочий протокол No Record Isolation

Протокол предназначен для статических структурированных наборов. Для нарративных, графовых или мультимодальных данных его необходимо дополнить специфическими тестами. Выбор полей и единицы наблюдения регистрируется до подсчёта результатов.

Показатель	Способ фиксации	Интерпретация
Единица наблюдения	Строка / человек / событие	Оговорить мультизаписи
Набор квазиидентификаторов	Список полей и точность	Зафиксировать смысл выбора
Единичные комбинации	Количество и доля строк	Повод к проверке связывания

Показатель	Способ фиксации	Интерпретация
Малые эквивалентные классы	Распределение размеров групп	Не универсальный правовой порог
Доля потерянной полезности	По заранее выбранным задачам	Сопоставить с целями публикации

Исследователь не должен публиковать список уникальных комбинаций, поскольку он сам может облегчать атаку. В открытом отчёте достаточно агрегированных метрик, метода и перечня исключений; конкретные редкие строки хранятся в закрытом доказательственном приложении.

Приложение 8. Протокол No Linkage и контроль ложных совпадений

Предлагается тестировать два слоя: связывание с внутренними опубликованными материалами и с внешними источниками, разумно доступными адресату. В результатах различаются истинное совпадение, вероятное совпадение, ложное совпадение и случаи невозможной проверки. Нельзя без разметки считать совпадение имён и дат доказательством идентичности.

Стадия теста	Действие	Доказательство
Определение атаки	Описать знания и допуск противника	Сценарий и дата
Список источников	Зафиксировать версии и условия доступа	Реестр вспомогательной информации
Совмещение полей	Провести тест на разрешённых данных	Закрытый журнал результата
Проверка качества	Расчёт precision и false matches	Размеченная контрольная выборка
Решение	Оценить реальную идентифицируемость	Обоснование аудитора

Если ложные совпадения обладают потенциально тяжёлым юридическим эффектом, их выделяют в отдельную категорию вреда. Так, неверно связанный санкционный профиль может быть опасен даже без успешного раскрытия первоначальной личности.

Приложение 9. Протокол No Inference для индивидуальных атрибутов

Тест предназначен для оценки возможности получить о конкретном лице чувствительный вывод из публикации при наличии указанных начальных знаний. Испытание должно измерять изменение информации, а не только способность алгоритма классифицировать общие категории. В отчёте отдельно обозначаются истинность атрибута, точность модели и вред от ложной атрибуции. [2]

Проверяемый элемент	Документирование
Предмет вывода	Конкретный тип атрибута, не личное имя
Базовая информация атакующего	Что было известно до раскрытия
Результат после доступа	Какие новые сведения стали выводимы
Надёжность	Precision, ошибки, интервал неопределённости
Эффект для субъекта	Возможность и последствия различного обращения
Решение	Исправление, иной режим или отказ от раскрытия

Нарушение отдельного критерия проекта EDPB 02/2026 не превращается автоматически в окончательный правовой вывод; проверяется, действительно ли полученные сведения способны отнести результат к человеку в релевантном контексте.

Приложение 10. Числовой пример k-анонимности: две разные доли

Условный набор включает 600 записей, сгруппированных по трём квазиидентификаторам. Из 120 эквивалентных групп 15 имеют по одной строке. Доля малых групп среди всех групп равна 15/120 = 12,5%, а доля уникальных строк среди всех строк равна 15/600 = 2,5%. Показатели отвечают на разные вопросы и не должны подменять друг друга.

Величина	Значение	Вычисление
Число строк	600	Исходный набор
Число групп	120	По выбранным квазиидентификаторам
Единичные группы	15	Эмпирический счётчик
Доля единичных групп	12,5%	15 / 120 × 100
Доля единичных строк	2,5%	15 / 600 × 100

Значение k для всего массива здесь равно 1. Это не доказывает автоматически, что каждая уникальная строка реально идентифицирует лицо; необходимы тесты доступных внешних средств и чувствительных выводов. Результат арифметики нельзя представлять как эксперимент на данных клиентов.

Приложение 11. Журнал бюджета дифференциальной приватности

Реестр предназначен для предотвращения неучтённой композиции нескольких статистических выпусков из одной популяции. Использование ε и δ требует заранее выбранной модели соседства, описания механизма и наличия технической реализации, прошедшей проверку. [13]

Релиз	Пример параметров	Статус
Срез А	ε = 0,5; δ = 0	Только условный пример
Срез Б	ε = 0,7; δ = 0	Только условный пример
Консервативная простая композиция	ε ≤ 1,2; δ = 0	При соблюдении предпосылок
Следующий срез	Указать новые параметры	Требует решения до запуска

В рабочем журнале также должны фиксироваться диапазон участия одного лица, связь наборов, точность, общий план раскрытия и ответственный. Никакой числовой epsilon в этом приложении не предлагается как универсально приемлемый для GDPR.

Приложение 12. Карта качества статистического результата с шумом

Перед применением DP либо иных методов возмущения данных необходимо записать, какую управленческую ошибку способен вызвать искажённый показатель. Это особенно важно в исследованиях редких процессуальных событий: публикация шума без указания интервала способна превратить нулевую или малую частоту в ложный вывод.

Вопрос качества	Запись в карточке
Какой показатель защищается?	Числитель, знаменатель, единица
Зачем он публикуется?	Конкретное решение или исследование
Допустимая ошибка	Заранее согласованная область точности
Метод добавления шума	Формула, генератор, параметры
Постобработка	Усечение отрицательных значений, округление
Сообщение читателю	Интервал, ограничения и запрет индивидуальных выводов

Если полезность не достигается при приемлемом уровне защиты, допустим отказ от публикации строки и переход к контролируемому исследовательскому доступу. Публичный интерес не требует выдавать искажённую статистику как точные факты.

Приложение 13. Проверка синтетического набора до передачи

Синтетический набор проходит отдельную техническую и юридическую экспертизу. Цель — выявить почти точные копии обучения, риск членства и возможность восстановления редких комбинаций; одновременно проверить, не создаёт ли генератор ложных доказательных фактов, способных попасть в рабочий оборот. [4][12]

Испытание	Что фиксируется	Возможная мера
Совпадение с обучающими строками	Доля близких пар и тип близости	Изменение генератора
Редкие чувствительные шаблоны	Число воспроизведённых случаев	Удаление источников риска
Membership inference	Метрики с базовой частотой	Доработка обучения
Качество распределений	Расхождения по целевым задачам	Ограничение применимости
Фактическая маркировка	Указание искусственного происхождения	Запрет смешения с реестром лиц

Если набор получен из оригиналов, внутри организации должен сохраняться защищённый след происхождения. Во внешнем приложении не следует публиковать обнаруженные реальные совпадения, иначе сама проверка превратится в новую утечку.

Приложение 14. Проверка AI-модели на воспроизведение обучающих фрагментов

Испытание проводится на отдельной согласованной тестовой базе и в контролируемом окружении. Задача — измерить возможность появления личной информации при обычных и усложнённых запросах, не распространяя проверочные сведения. В протоколе нужна версия модели, способ доступа, наличие retrieval-слоя, максимальный бюджет тестовых запросов и решение о безопасности публикации. [4][6]

Тип проверки	Контрольный вопрос
Обычный пользователь	Раскрываются ли сведения при нейтральных вопросах?
Повторная формулировка	Меняется ли результат при допустимой вариации?
Частичный контекст	Возможно ли восстановить редкий фрагмент?
Retrieval-индекс	Есть ли доступ к документам вне роли пользователя?
Фильтр ответа	Не скрывает ли он риск доступа к параметрам/индексу?

Отрицательные результаты нельзя обобщать на открытые веса или новую версию. Обнаруженные личные сведения хранятся отдельно; публичный отчёт сообщает только категорию, степень риска и меры исправления.

Приложение 15. Расчёт положительной прогностической ценности атаки membership inference

Условная модель имеет чувствительность 90%, специфичность 95%; реальная доля членов целевой популяции составляет 1%. При 10 000 проверяемых лицах ожидаются 100 участников обучения. Тест верно пометит 90 из них. Из 9 900 неучастников 5% будут ошибочно помечены, то есть 495 человек. Положительных результатов ожидается 585, а доля истинных среди них — 90/585 ≈ 15,38%.

Категория	Прогноз при условных параметрах
Истинные участники	100
Истинноположительные	90
Ложноположительные	495
Все положительные	585
Положительная прогностическая ценность	15,38%

Эти числа демонстрируют необходимость учитывать базовую частоту и цену ложной атрибуции; они не являются реальным результатом атаки. При иных условиях, размере обучающего массива и характеристиках модели значения изменятся.

Приложение 16. Интерпретация отрицательных испытаний

В условном независимом биномиальном эксперименте при нуле успешных событий в 300 попытках приближённая верхняя граница вероятности события на уровне 95% по правилу 3/n равна 1%. Точная односторонняя граница составляет приблизительно 1 − 0,05^(1/300) ≈ 0,9936%. Интерпретация применима только при корректности предпосылок об одинаковых независимых испытаниях.

Показатель	Значение	Ограничение
Число испытаний	300	Условная выборка
Успешные события	0	Не доказывает нулевую вероятность
Приближённая граница	1,0%	Правило 3/n
Точная граница	≈ 0,994%	Односторонняя биномиальная

Если запросы адаптивны, распределения различаются или некоторые из них происходят по иным каналам, такой интервал может оказаться некорректным. В отчёте обязательно указываются тип события, способ отбора и условия применения статистической модели.

Приложение 17. Контрольный лист векторных индексов и RAG

Использование retrieval не отменяет необходимости разграничивать статусы исходного документа, индексного представления и ответа модели. Один из наиболее опасных сценариев — когда агрегированная выдача ссылается на защищённый документ или воспроизводит его редкое содержание. Ниже приведён минимальный объектный контроль.

Компонент	Контроль	Неприемлемая ситуация
Исходные документы	Классификация, сроки, основание	Незаконная передача в индекс
Эмбеддинги	Тест поиска редких текстов	Узнаваемый фрагмент досье
Метаданные	Удаление избыточных идентификаторов	Имя субъекта в скрытом поле
Контекст модели	Ролевой доступ и границы выдачи	Подмешивание закрытого текста
Логи	Доступ, срок, маскирование	Сбор новых чувствительных историй
Результат	Тест регургитации	Цитирование закрытых сведений

Оценка должна охватывать модельное окружение в целом. Отдельный «сертификат анонимности эмбеддинга» не обосновывает безопасность приложения, если поиск соединяет его с необезличенными источниками.

Приложение 18. Оценка удалённых идентификаторов в свободном тексте

Тексты требуют двухэтапного контроля: сначала поиск явных идентификаторов, затем анализ сюжетных комбинаций. Именованные сущности, даты, учреждения и географические признаки должны проверяться вместе с содержанием отдельных юридических событий. Для закрытых материалов рекомендуется две независимые читательские оценки и проверка поисковой сопоставимости с открытыми источниками.

Категория	Типичный маркер	Проверочный приём
Прямой идентификатор	Имя, документ, контакт	Автоматический поиск + чтение
Процедурный	Номер дела, точная дата	Сопоставление с официальным реестром
Сюжетный	Редкое сочетание фактов	Поиск комбинации без имени
Семейный	Родство и уникальный эпизод	Проверка третьих лиц
Технический	Автор файла, EXIF, комментарии	Анализ метаданных

Проверка не должна превращаться в массовую передачу конфиденциального текста неизвестному внешнему AI-сервису. Если используется модель для предварительного поиска маркеров, её режим обработки и условия хранения проходят отдельную оценку.

Приложение 19. Оценка траекторий и географических данных

Цель матрицы — предотвратить уникальность маршрутов, которая сохраняется после удаления имён. Проверка осуществляется отдельно для точных координат, последовательности посещений и публичных событий, к которым может быть привязан субъект.

Проверяемый признак	Потенциальный риск	Снижение риска
Точные координаты	Привязка к дому или работе	Агрегация района
Минутная временная метка	Совпадение с публичным событием	Укрупнение интервала
Редкая последовательность мест	Поведенческий отпечаток	Уменьшение детализации
Маршрут из малого населённого пункта	Узнаваемость субъекта	Подавление редких случаев
Пересекающиеся публикации	Дифференциальное раскрытие	Общий реестр всех релизов

Если преобразование лишает материал доказательной точности, публикация в открытый доступ может оказаться неправильной целью. Тогда необходимо сохранить оригинал закрыто и выпустить только методические выводы.

Приложение 20. Контроль статей 9, 10, 11 и 25 GDPR

Реестр адресован юристам, согласующим независимое исследование. Он служит напоминанием о том, что технический результат не отменяет нормы, применимые к исходной обработке. Статья 25 дополняется действующими рекомендациями EDPB о защите данных по проектированию и по умолчанию; проект Guidelines 01/2025 о псевдонимизации не следует выдавать за окончательное руководство. [1][10][11]

Правовая норма	Критический вопрос
Статья 9	Есть ли условие обработки особых категорий?
Статья 10	Соблюдаются ли специальные требования к сведениям о правонарушениях?
Статья 11	Действительно ли идентификация не требуется для целей?
Статья 25	Встроены ли соответствующие меры уже в проектирование?

Правовая норма	Критический вопрос
Статья 32	Адекватна ли безопасность исходника и ключей?
Статья 35	Нужна ли оценка воздействия на защиту данных?

В заключении отдельно отражается законодательство государства-члена, если применение статьи 10 или исследовательских условий зависит от национальных правил. Отсутствие юрисдикционного анализа должно быть отмечено как ограничение.

Приложение 21. Паспорт AI-модели для вывода о персональном статусе

Отчёт о модели не может ограничиваться названием семейства и числом параметров. Для оценки извлекаемости персональных сведений необходимо указать версию, период обучения, тип доступа, связанные индексы и существенные средства снижения запоминания. Отдельно фиксируется, кто может вызывать модель и какие побочные источники ей доступны. [4][6]

Раздел паспорта	Минимальное содержание
Объект проверки	Версия весов, дата и контрольный идентификатор
Обучающие источники	Категории и исключения, без публикации оригиналов
Доступ	Закрытый API, внутренний контур, открытые веса
Дополнительные ресурсы	RAG, инструменты и внешние базы
Тесты	Экстракция, membership, inference, регургитация
Принятые меры	Дедупликация, фильтры, контроль прав
Ограничения	Непроверенные режимы и доступы

Заявление об анонимности относится только к описанному объекту и контексту. Факт публикации новой версии или подключения другого retrieval-индекса требует отдельной проверки.

Приложение 22. Due diligence внешнего поставщика анонимизации

Контролёр не должен передавать поставщику исходную клиентскую базу на основании маркетинговой фразы «GDPR compliant». Необходимо установить фактические роли, законность маршрута передачи, доступ субподрядчиков и меры безопасности. Наличие технической спецификации преобразований полезно, но без подтверждения её исполнения недостаточно. [1][12][14]

Запрос поставщику	Что требовать в ответ
Кто принимает решения о целях и средствах?	Роли и договорная схема
Где обрабатываются исходные данные?	Регионы, субпроцессоры, трансферы
Сохраняются ли ключи соответствия?	Схема владения и уничтожения
Есть ли тесты на связывание и вывод?	Протоколы и условия воспроизведения
Какие меры при обнаружении утечки?	Процедура реагирования и сроки
Можно ли провести независимый аудит?	Доступ, ограничения, право проверки

Передача оригиналов должна быть минимальной. Если для поставленной задачи достаточно выполнения кода в инфраструктуре заказчика, такой вариант следует сравнить с полной выгрузкой данных во внешнюю среду.

Приложение 23. Форма решения о выпуске обезличенного результата

Форма предназначена для подписываемого решения, а не для автоматического сертификата. Заполняются объект, аудитория, краткое обоснование и условия пересмотра. При несогласии

юриста и технического аудитора разногласие фиксируется, а не скрывается усреднённым «зелёным» статусом.

Поле решения	Содержание
Идентификатор и версия	Однозначный код артефакта
Цель выпуска	Зачем нужна именно такая детализация
Юридическая классификация	Анонимно / персонально / недостаточно данных
Проверенная аудитория	Публичная / закрытая / конкретный получатель
Основные остаточные риски	Условия, при которых риск возвращается
Решение о выпуске	Разрешить / ограничить / доработать / отказать
Ответственные	Контролёр, технический рецензент, юрист
Следующий пересмотр	Дата или конкретный триггер

Пустые графы о допущениях и ограничениях должны блокировать утверждение документа. Условное разрешение публиковать только агрегаты не является разрешением выдавать исходные строки.

Приложение 24. Триггеры повторной проверки анонимности

Регламент должен предусматривать немедленный и плановый пересмотр. Частота выбирается по характеру массива, но сама по себе дата очередного аудита не заменяет событийных триггеров. [2][15]

Триггер	Обязательная реакция
Новая внешняя база	Повторить тест связывания
Изменение получателя	Обновить контекст идентификации
Раскрытие дополнительного среза	Проверить пересечение публикаций
Новый релиз модели	Проверить экстракцию и inference
Смена режима API	Пересмотреть доступ и бюджет запросов
Инцидент безопасности	Приостановить и расследовать, если необходимо
Жалоба потенциально идентифицируемого лица	Адресная проверка и документированное решение

Каждое повторное заключение ссылается на идентификатор предыдущего. Если новая информация меняет статус результата, публикуется исправление с понятным описанием границ старого вывода, без раскрытия новых персональных сведений.

Приложение 25. Карта инцидента ошибочной анонимизации

При подозрении на идентифицируемость опубликованных данных необходимо немедленно отделить установление фактов от публичной коммуникации. Изъятие материала может быть разумной временной мерой, но нельзя утверждать, что все ранее скачанные копии уничтожены. Юридическая квалификация возможного нарушения должна опираться на статьи GDPR и фактический риск, а не на название внутренней уязвимости. [1]

Временная стадия	Действие	Ответственный
Обнаружение	Зарегистрировать сигнал и версию файла	Дежурный по инцидентам
Сдерживание	Ограничить дальнейшее распространение	Владелец сервиса
Факты	Установить данные, получателей и период	Техническая группа
Юридическая оценка	Проверить статус персональных данных и уведомления	Юрист / контролёр

Временная стадия	Действие	Ответственный
Корректирующие меры	Переанонимизация или отказ от публикации	Владелец данных
Закрытие	Независимый повторный тест и отчёт	Внешний рецензент

Закрытое приложение к акту хранит технические подтверждения; публичный отчёт не повторяет личные данные, найденные при тестировании. Любые сообщения надзорным органам подготавливаются по применимому праву и фактическим срокам, а не по условным SLA этого шаблона.

Приложение 26. Форма обращения лица о возможной идентификации

Доступный путь оспаривания публикации необходим даже при уверенности организации в своей методике. Заявителю достаточно указать материал и причину опасения; от него не следует заранее требовать передачи полного паспорта, медицинской истории или процессуального досье, если для первичного анализа достаточно меньшего объёма сведений.

Поле	Порядок обращения
Ссылка на опубликованный материал	URL, версия или цитата безопасного фрагмента
Характер риска	Выделение, связывание, вывод или иной вред
Контакт для ответа	Минимально необходимое средство связи
Предложенные меры	Исправление, ограничение, пояснение
Временные меры	Решение о блокировке на период проверки
Итог и основания	Документированное заключение и порядок эскалации

Организация не должна автоматически удалять доказательства обращения. Следует сохранять достаточный, но минимизированный след рассмотрения с ограниченным доступом и определёнными сроками хранения.

Приложение 27. Учебный сценарий: международный розыск

Гипотетическая ситуация. Исследователь хочет опубликовать краткое описание успешного оспаривания трансграничного розыска. Имя заменено на «заявитель А», однако текст сохраняет год задержания, редкое гражданство, страну транзита, конкретный суд и возраст ребёнка. Публикация может быть сопоставлена с судебным решением и медийным сообщением.

Проверка	Возможный результат	Решение
No Record Isolation	Редкая комбинация событий	Обобщить фабулу
No Linkage	Сходство с публичным делом	Удалить точный суд и даты
No Inference	Из кейса выводится семейный статус	Исключить детали о ребёнке
Полезность	Сохраняется принцип отказа в экстрадиции	Публиковать правовой вывод, не историю

Сценарий создан для методики и не описывает реального клиента. Конечный документ должен указывать условный характер примера и запрещать его использование как подтверждения факта о конкретном деле.

Приложение 28. Учебный сценарий: санкционный false positive

Гипотетическая ситуация. Сервис проверяет совпадения по санкционным спискам. Для оценки качества ARGA получает обезличенную таблицу ошибок: месяц, страна, отрасль, тип имени, объяснение совпадения и суммы операций. В небольшом секторе сохранилась уникальная

комбинация крупной сделки и региона, позволяющая найти конкретную компанию и её руководителя.

Риск	Мера оценки
Связывание с корпоративным реестром	Поиск уникального сочетания
Ложное обвинительное впечатление	Проверка пояснений и маркировки исхода
Косвенная идентификация физического лица	Отделить компанию и должностных лиц
Утрата полезности при обобщении	Сохранить статистику типов ошибок
Повторные релизы	Контроль пересекающихся срезов

Публично допустимая таблица может показывать частоты ошибок по типам алгоритмов без суммы конкретной операции и точного месяца. Исходные спорные совпадения остаются в закрытом контуре с законными основаниями доступа.

Приложение 29. Учебный сценарий: отчёт об убежище и миграции

Гипотетическая ситуация. Доклад о миграционном решении сообщает страну происхождения, род занятий, месяц эвакуации, маршрут через две страны и исход дела. Даже после удаления имени сочетание может быть узнаваемо в местных публикациях. Уникальные сведения о ребёнке и месте проживания способны создавать дополнительный риск безопасности.

Контроль	Вопрос	Действие
Сюжетная уникальность	Находится ли дело внешним поиском?	Изменить уровень детализации
Уязвимые третьи лица	Можно ли узнать родственников?	Исключить дополнительные сведения
Доказательный оригинал	Требует ли закон сохранения?	Отдельный закрытый архив
Публикационная цель	Нужен ли реальный кейс?	Использовать составной учебный пример

Целью анонимизации является защита лиц, а не эстетическая замена фамилий и названий. Если сохранение юридически существенных обстоятельств несовместимо с безопасной публикацией, следует отказаться от индивидуального кейса.

Приложение 30. Проект замечаний для консультации EDPB 02/2026

Этот текст является планом содержательной позиции, а не отправленным документом. Исходная редакция должна быть дополнительно согласована на уровне ARGA и соответствовать актуальной форме EDPB. Период консультации проекта — 8 июля — 30 октября 2026 года. [2]

Предложение	Мотивировка	Ожидаемое уточнение
Пример для юридических нарративов	Редкость сюжета не видна в таблицах	Тест семантического связывания
Последовательные публикации	Отдельные агрегаты раскрывают при пересечении	Обязательное описание истории релизов как лучшая практика
AI-модели и retrieval	Один объект не охватывает цепочку обработки	Многослойный паспорт проверки
Уязвимые и преследуемые лица	Вред может быть непропорциональным	Риск-ориентированные примеры
Доказательная воспроизводимость	Вывод трудно независимо проверить	Минимальная форма протокола

Предложения следует аргументировать ссылками на проект и документированной методикой, избегая утверждения, что каждое пожелание уже является нормой GDPR. Подача и подтверждение получения документируются отдельно.

Приложение 31. Хартия независимости аудитора

Независимая экспертиза не должна гарантировать коммерчески желаемый результат. Проект хартии устанавливает регистр заинтересованностей, запрет редактирования отрицательных находок под давлением заказчика, прозрачное описание ограничений и отдельное утверждение публичной версии с учётом защиты данных.

Принцип	Практическая гарантия
Необусловленность вывода	Вознаграждение не зависит от оценки «анонимно»
Управление конфликтом	Раскрытие связей и механизм отвода
Методическая независимость	Предрегистрация тестов и изменений
Техническая воспроизводимость	Архив версий, хешей и протоколов
Ответственная публикация	Закрытая часть для найденных личных данных
Исправление ошибок	Публичная история существенных коррекций

Хартия не превращает Observatoire в надзорный орган либо аккредитованную организацию оценки соответствия. Правовой статус любой внешней декларации должен быть обозначен точно.

Приложение 32. Контрольный список финальной редакционной проверки

Редактор выполняет формальную и содержательную проверку перед выпуском отчёта. Отметка о наличии приложения не заменяет проверки его качества; пустые шаблоны, не относящиеся к предмету проекта, исключаются из публикации.

Проверяемое требование	Признак выполнения
Даты и версии	Указан срез права и проекта EDPB
Статус документов	Закон отделён от консультационного проекта
Все фактические утверждения	Есть подтверждаемый источник или оговорка
Все математические примеры	Показаны знаменатели и предпосылки
Персональные сведения	Не включены в публичный отчёт
Внутренние ссылки	Оглавление совпадает с пагинацией
Внешние ссылки	Ведут на официальные документы
Шапка и колонтитулы	Совпадают с серией №11–21
Отображение символов	Нет квадратов, обрезок и порчи таблиц

Проверка завершается отдельной записью о том, что документ пригоден для внутреннего рассмотрения. Для официального утверждения требуется профильная юридическая и научная рецензия, включая проверку применимости выводов к фактическим материалам.

Приложение 33. Подробный кейс: малочисленная группа и раскрытие медицинского признака

Исходная гипотетическая ситуация. Исследовательский центр публикует отчёт о последствиях решения миграционного органа для группы из 28 человек. После удаления имён таблица содержит возрастной диапазон, месяц обращения, регион проживания и показатель доступа к медицинской помощи. В шести сочетаниях квазиидентификаторов есть по одной записи. В одном сочетании присутствует редкая последовательность «переселение

— операция — отказ в выплате», которую локальное СМИ описало в статье без упоминания медицинского диагноза.

Первая проверка должна ответить на вопрос No Record Isolation: выделяется ли одиночная запись по сохраняемым признакам. Вторая — No Linkage: способен ли обычный читатель или заинтересованное лицо обнаружить совпадающий сюжет в локальной публикации. Третья — No Inference: позволяет ли связанная таблица вывести медицинское состояние, прежде неизвестное человеку, нашедшему статью. Суммарная оценка существенно сильнее, чем любой тест по отдельности. [2][3]

Возможные технические меры включают объединение нескольких месяцев, исключение необычного медицинского показателя и переход от индивидуальной таблицы к статистике, при которой количество наблюдений не раскрывается для очень малых групп. Но если требуемая аналитическая цель состоит именно в изучении редкой ошибки, последующее подавление всех редких случаев может систематически скрыть наиболее важные нарушения. Решение о публикации должно учитывать конфликт между приватностью и полезностью, а не стремиться автоматически увеличить размер групп.

Проект результата независимой проверки: исходная публикация отклоняется в представленном виде; контролёру предлагается сохранить закрытую доказательственную версию, публиковать агрегаты без уникальных событий и сопровождать выводы качественным правовым анализом без связывающих деталей. Этот вывод условный: без настоящих данных и проверки внешних источников он не является заключением о конкретном человеке.

Вариант публикации	Приватностный риск	Научная полезность
Строковая таблица с датами	Высокая вероятность связывания	Высокая детализация
Обобщённый квартал и регион	Риск ниже, но редкость сохраняется	Средняя
Агрегированный показатель без редких групп	Снижен для внешней аудитории	Ограниченная
Закрытая среда с проверяющим	Управляемый риск доступа	Высокая при гарантиях

Приложение 34. Биометрические признаки, фотографии и голос

Данные лица, голоса и движений не всегда поддаются традиционным методам удаления идентификаторов. Размытие лица может сохранить идентифицирующие контуры одежды, татуировки, фон, силуэт и окружающий текст. Изменение тембра голоса может не скрывать речь о конкретном событии либо устойчивые речевые признаки. Автоматический перевод и пересказ интервью способны сохранить уникальную фабулу, по которой человек находится во внешних источниках.

При анализе биометрических данных необходимо разграничивать биометрические данные в техническом смысле и специальные категории GDPR, для которых значение имеет, в частности, специальная обработка для целей однозначной идентификации физического лица. Применимость статьи 9 зависит от юридической квалификации конкретной операции; нельзя объявлять все изображения автоматически специальными категориями или, наоборот, игнорировать биометрический риск в стандартном файле фото. [1]

Для аудита мультимодального материала рекомендуются три группы тестов. Первая — удаление прямых меток: лицо, имя на бейдже, номер дела, голосовое представление. Вторая — анализ контекстных идентификаторов: место, даты, уникальный фон, сопутствующие люди и метаданные файла. Третья — тест сопоставления с открытыми публикациями и обратным поиском в разрешённых условиях. Наличие у аудитора инструментов распознавания должно сопровождаться оценкой законности их использования и ограничением результатов.

Условный сценарий: в публичном ролике эксперт рассказывает об анонимном свидетеле. Лицо размыто, но на экране виден уникальный корпус суда, а время заседания указано в подписи. В таком случае техническая обработка лица не снимает риска идентификации. Следует переработать сюжет и визуальную среду, а не усилить размытие и объявить материал анонимным.

Уровень	Что тестируется	Типичные меры
Медиафайл	Видео, аудио, субтитры	Монтаж, обобщение контекста
Метаданные	EXIF, геометки, имена авторов	Очистка контейнера
Сюжет	Редкое событие, участники	Переформулирование и исключение деталей
Внешний поиск	Сопоставимость кадров/фраз	Решение о неприменимости публикации

Приложение 35. Многоязычные корпуса и ошибки извлечения идентификаторов

Юридические корпуса ARGA могут включать русский, английский, французский и местные языки, иногда в одном документе. Система поиска личных данных, проверенная только на английском, способна пропускать транслитерации, отчества, изменяемые фамилии, названия организаций, падежные формы и идентификаторы с непривычной структурой. Простой поиск шаблона имени недостаточен для доказательства качественной анонимизации.

Процедура контроля начинается с составления языковой карты набора. Для каждого языка выбирается репрезентативная тестовая выборка с разметкой прямых и косвенных идентификаторов. Показатель полноты автоматического обнаружения следует измерять отдельно по языкам и типам идентификаторов. Общий recall, рассчитанный на преобладающем английском корпусе, может скрывать слабость на документах о международном розыске на редком языке.

Отдельный риск связан с машинным переводом. Перевод способен нормализовать редкую формулировку, но также внести новые интерпретации и исказить юридически значимый факт. В отчёте о публикации важно различать редакционное обобщение, осмысленное человеком, и неконтролируемую генерацию текста. Для юридических примеров минимизация риска не должна создавать ложные процессуальные обстоятельства; если анонимизация требует существенного изменения фактов, пример маркируется как составной и учебный.

Возможный тест — сначала сопоставить обезличенный русскоязычный текст с открытыми русскими источниками, затем с переводами на другие языки. Если уникальный эпизод находится по смыслу после перевода, анонимность нельзя подтверждать лишь отсутствием

буквального совпадения. В техническом протоколе фиксируются языки запросов, версии переводчика и ограничения контроля; внешние AI-сервисы используются только при законных условиях обработки исходного текста.

Проверка	Ошибка, которую выявляет	Документ
Стратифицированная разметка	Неравномерная полнота по языкам	Таблица качества
Транслитерации	Пропуск вариантов имени	Список правил
Семантический поиск	Узнаваемость сюжетной фабулы	Протокол без публикации оригинала
Проверка перевода	Подмена или утрата юридических фактов	Журнал редакторской проверки

Приложение 36. Композиция агрегированных публикаций и дифференциальное раскрытие

Условный пример. Организация публикует число обращений в конкретный административный орган за весь квартал, а затем за каждый месяц, кроме одного. Формально каждая таблица содержит только агрегаты. Однако читатель способен вычесть суммы и восстановить число за скрытый месяц. Если соответствующий месяц включает одну известную извне жалобу, дополнительное раскрытие может выдать чувствительный факт о лице. Это классический риск пересечения релизов, который нельзя оценить только по отдельным файлам.

Для каждой новой публикации следует вести карту статистических измерений: популяция, временное окно, география, набор фильтров, используемые классификации, ранее опубликованные пересекающиеся итоги и условия доступа. Если система позволяет пользователю самостоятельно менять фильтры, необходимо оценить атаки на вычитание и последовательные запросы. Ограничение одной таблицы с маленькими ячейками неэффективно, если сумма и соседние срезы доступными пользователю восстановят скрытое значение.

Меры различаются по цели и угрозе: подавление взаимосвязанных итогов, контролируемые ответы на запросы, приватностный бюджет, уменьшение точности и ограничение числа доступных измерений. Универсальный порог «не показывать значения меньше пяти» не достаточен; через разность итогов могут раскрываться значения меньше порога. Выбор механизма требует оценки научной полезности и реальной модели пользователей. [12][13]

Внутренний протокол должен включать как минимум один симулированный тест обратного расчёта на специально созданных агрегатах. Если итоговый интерфейс не допускает выгрузку строк, но допускает неограниченное число произвольных фильтров, его нельзя считать безопасным без анализа совокупной информационной выдачи.

Опубликовано	Скрыто	Возможность восстановления
Квартал = 120	Март	При наличии января и февраля
Январь = 53	Март	Вычитание частичной суммы
Февраль = 66	Март	120 − 53 − 66 = 1
Март = 1 (выведено)	Индивидуальная связь	При наличии внешнего знания о жалобе

Приложение 37. Анонимизация графа финансовых отношений

Графовой набор создаётся из организаций, физических лиц, адресов и платежей. Даже если имена заменены случайными идентификаторами, структура связей может оставаться уникальной: центральный узел с несколькими дочерними компаниями, одним зарубежным получателем и редкой последовательностью переводов способен совпадать с публичным корпоративным реестром. При этом использование графа для AML-аналитики часто требует сохранения аномальной структуры, иначе исчезает предмет исследования.

Предлагается разделить доказательный граф и публикационный граф. В первом хранятся оригинальные связи с ограниченным доступом и правовым основанием. Во втором раскрываются только категории узлов, статистические характеристики сети, агрегированные признаки подозрительных паттернов и, где необходимо, тщательно изменённые учебные структуры. Научная публикация должна сообщать, что учебный граф является синтетическим или модифицированным, а не притворяться неизменённым реальным транзакционным набором.

Проверка No Record Isolation распространяется на уникальность сочетаний степеней узлов, типов связей и последовательности операций. No Linkage тестируется по доступным корпоративным записям и публичным финансовым событиям. No Inference проверяет, какие чувствительные выводы можно сделать о должностном лице или бенефициаре при доступе к частичному графу. Используемые графовые метрики сами иногда раскрывают наличие редкого участника, особенно если публикация содержит маленький сектор.

Для ARGA риск ложного связывания равен по значимости риску истинной утечки: совпадающая структура не всегда доказывает, что компания участвовала в запрещённой операции. В открытом тексте необходимо избегать формулировок, которые превращают статистическое подозрение в утверждение об установленном правонарушении.

Объект графа	Риск раскрытия	Публикационный вариант
Уникальный центральный узел	Идентификация компании и руководителя	Обобщённый класс ролей
Редкое ребро платежа	Привязка к транзакции	Диапазон и агрегация
Временная цепочка	Восстановление события	Учебная синтетическая структура
Открытый реестр	Связывание с бенефициаром	Проверка контекста и исключения

Приложение 38. Условный пример различия перспектив контролёра и получателя

Организация A проводит опрос 2 000 участников и присваивает каждой анкете случайный код. В закрытом хранилище она оставляет таблицу кодов и контактных данных для последующих разъяснений. Организации B передаётся только очищенная таблица ответов и правила анализа. Договор запрещает повторную идентификацию, но само по себе это запрещение не доказывает, что для B информация анонимна. Нужно установить, имеются ли в ответах редкие идентифицирующие признаки и доступные B внешние источники.

С точки зрения A данные остаются персональными, если она способна связать код с человеком; применяются обязанности GDPR. С точки зрения B оценка требует конкретных фактов о возможности связи, устройстве доступа и получаемой информации. После дела

EDPS v SRB нельзя ни игнорировать различия между участниками, ни использовать их для обхода обязанности исходного контролёра предоставить информацию при сборе данных. [7]

Если B получает дополнительно журналы обращений либо может направлять запросы A для связывания ответов, исходная оценка меняется. Если анкета содержит точную редкую профессию и единственный регион, даже отсутствие таблицы соответствия не гарантирует невозможности идентификации. Поэтому протокол должен включать перечень допустимых запросов между A и B и проверку содержательной уникальности, а не одну строчку о договорном запрете.

Условный результат: допускается контекстный вывод об анонимности для B лишь при проверенном отсутствии разумно вероятной атрибуции в её фактическом окружении. В отношении A сохраняются собственные обязанности; законность первоначального сбора анализируется отдельно, а заключение B не легализует старые ошибки A.

Вопрос	Организация A	Организация B
Таблица соответствия	Да	Нет
Возможность отнести код к лицу	Прямая через закрытый ключ	Требует фактической оценки
Применимость GDPR	При наличии персональной обработки	Зависит от квалификации полученных данных
Обязанности информирования при сборе	У A	Не устраняются статусом B

Приложение 39. Контекстный риск моделей с открытыми весами и закрытым API

Одна и та же модель может давать разный риск для пользователя ограниченного API и для лица, получившего веса, возможность дообучения и неограниченного локального вычисления. Через API владелец способен вводить лимиты запросов, фильтры и мониторинг извлечения. Но программный фильтр ответа не доказывает, что персональных сведений нет в параметрах; при передаче весов часть защитных мер становится неприменимой. [4][6]

Оценка должна различать «модель как размещённый сервис», «веса модели» и «систему с внешними знаниями». Для публичного API дополнительно важны лимиты повторных запросов, учёт адаптивного поведения и защита журналов. Для открытых весов требуется более сильная проверка прямой экстракции и признание ограничений тестового бюджета; после публикации запретить локальные эксперименты технически невозможно.

Если разработчик считает веса анонимными, такой вывод должен опираться на доказательства устойчивости к разумно вероятным средствам для предполагаемого круга держателей весов. Недостаточно испытать чат-интерфейс под жёстким фильтром, а затем объявить эквивалентную анонимность распространяемому файлу параметров. Смена режима доступа должна запускать новый цикл оценки.

Предлагается раскрывать в отчёте отдельные строки для API, локальных весов, fine-tuning, retrieval и журнала взаимодействий. Если какой-либо режим не исследован, его следует пометить как вне объёма проверки. Заключение «всё безопасно при любых формах доступа» не соответствует проверяемой методологии.

Режим доступа	Типовой контроль	Существенное ограничение
Закрытый API	Лимиты, фильтры, журнал	Возможны адаптивные запросы
Открытые веса	Тест прямой экстракции	Слабый контроль после выпуска
Fine-tuning	Версия данных и модели	Возможное новое запоминание
RAG	Изоляция документов и ролей	Утечка не из весов, а из индекса

Приложение 40. Удаление данных из модели и пределы machine unlearning

Обращение субъекта об удалении персональных данных ставит перед разработчиком комплексный вопрос: удалены ли исходные файлы, производные копии, индексы поиска, обучающие контрольные примеры и информация, потенциально запомненная моделью. Применимость права на удаление зависит от конкретных условий статьи 17 GDPR и предусмотренных законом исключений. Нельзя обещать автоматическое исчезновение всей информации из весов только по факту удаления записи из исходной таблицы. [1][18]

Технические стратегии включают исключение соответствующего материала из будущего обучения, изменение индекса, перестройку производных наборов, применение специальных методов unlearning либо повторное обучение при обоснованной необходимости. Но их эффективность требует теста: модель может сохранять общий паттерн, отражать данные в логах или быть развернута в недоступных копиях. Оценка должна отделять факт удаления из управляемого хранилища от вероятности последующего извлечения из конкретной модели.

В юридическом отчёте необходимо описывать, какие операции выполнены, какие системы охвачены, какие копии невозможно разумно отозвать и какую проверку провели после исправления. Формулировка «запись удалена» не должна скрывать, что исходный материал оставался в неподконтрольном партнёру архиве, либо что результаты прошлых публикаций уже разошлись к внешним пользователям.

Для ARGA применяется приоритет процессуальной честности: нельзя уничтожать единственный доказательный оригинал, если действующее право требует его сохранения, и одновременно нельзя использовать необходимость архива как основание для бессрочного хранения всех производных копий. Юридический режим и риск опубликованной версии должны решаться раздельно.

Компонент	Возможная мера	Способ проверки
Исходная строка	Удаление / законное ограничение	Журнал операции
Архив и резервная копия	Управляемый срок жизни	План уничтожения
Retrieval-индекс	Исключение сегмента	Тест поиска
Параметры модели	Unlearning / переобучение при необходимости	Тест извлечения после изменения
Публичная копия	Коррекция и уведомление	Учёт ограниченной отзывности

Приложение 41. Протокол предварительной оценки пригодности открытой публикации

Публичное размещение предполагаемого анонимного массива требует иной проверки, чем предоставление того же массива одному исследователю по договору. В открытой среде круг

потенциальных получателей невозможно надёжно ограничить, а появление новых открытых источников и систем сопоставления значительно меняет реидентификационный риск. Поэтому исследовательское решение должно приниматься не по принципу «данные очищены, следовательно можно публиковать», а после отдельного определения цели и условий использования. [1][2][15]

Первая стадия касается целевой аудитории: кто получит доступ, можно ли копировать файлы, индексировать их, автоматически сопоставлять с реестрами и воспроизводить выводы в коммерческих системах. Вторая стадия описывает угрозы, актуальные при неизменном наборе данных: уникальные временные комбинации, последовательности событий, редкие отношения между участниками и косвенные признаки профессии, места нахождения либо судебного статуса. Третья стадия рассматривает динамические угрозы, возникающие после появления новых публикаций, утечки соседнего массива или изменений в алгоритмах поиска. [2][15]

Для чувствительных историй используется отдельное правило: даже статистически малая вероятность установления личности может быть неприемлемой, если раскрытие способно повлечь преследование, лишение доступа к финансовым услугам или угрозу личной безопасности. Отдельно проверяется ущерб от ложного связывания данных. При отрицательном решении следует предлагать альтернативу: частичную публикацию агрегатов, ограниченный доступ, выдачу результатов по запросу или полностью синтетический пример с доказанной независимостью от реальных кейсов. [1][4]

Уровень выпуска	Доступ и контроль	Условие допуска	Предел заключения
Открытый	Неограниченное копирование	Проверка потенциального внешнего связывания и инференции	Вывод действителен лишь для описанного состояния знаний и угроз
Контролируемый	Установленные получатели и договор	Проверка полномочий, запрета повторного раскрытия и журналирования	Договор не превращает персональные данные в анонимные
Безопасная исследовательская среда	Вычисления под наблюдением	Контроль экспорта, ролей, журналов запросов и повторных вычислений	Среда снижает риск, но не доказывает анонимность
Непубликационный	Только внутренний анализ	Проверка законности исходной обработки и минимизации	Отсутствие публичного выпуска не отменяет GDPR

Решение оформляется актом с датой, версией массива, подписью двух независимых специалистов и перечнем принятых ограничений. Отсутствие записанного решения означает отсутствие завершённой проверки, даже если разработчик устно сообщает об успешном тесте.

Приложение 42. Протокол реконструкции связывания нескольких публикаций

Публикации нужно рассматривать как единую последовательность, а не как независимые документы. Допустим, первая таблица показывает число обращений за убежищем по

региону, вторая — количество отклонённых обращений по возрасту, а третья — распределение по определённым месяцам. Каждая таблица сама по себе может содержать только группы достаточного размера. Однако пересечение трёх фильтров иногда формирует группу из одного или двух лиц, о которых внешняя пресса уже содержит дополнительные сведения. Такой механизм называют композиционным раскрытием и он особенно существенен для малых территорий. [4][12]

Аудитор восстанавливает каталог всех связанных релизов за согласованный период, включая архивные версии и документы партнёров. Затем моделируются операции над разрешёнными агрегатами: пересечение категорий, вычитание соседних временных периодов, сравнение до и после удаления записи и запросы по комбинациям признаков. Уточняется, могут ли результаты оказаться доступны одному и тому же получателю, включая через поисковую систему, редакционный архив или машинный интерфейс. Отдельно отмечаются точные и приблизительные совпадения; невозможность автоматического связывания не исключает ручное восстановление. [2][15]

При выявлении раскрытия необходимо установить минимальный набор публикуемых элементов, достаточный для повторения вывода, и задокументировать его без раскрытия защищаемого лица в публичной части акта. Для исправления недостаточно изменить только последний выпуск: последующие редакции будут сопоставляться с предыдущими копиями, а кеширование не позволяет гарантировать глобальное уничтожение ранее доступной информации. В реестре остаются даты исправления, уведомлённые адресаты и остаточный риск. [1]

Шаг	Доказательство	Типичный сбой
Инвентаризация	Все версии, форматы и внешние копии	Проверяли только действующий PDF
Пересечения	Список совместимых фильтров и тестовых запросов	Игнорировали вычитание соседних периодов
Внешнее сопоставление	Перечень доступных публикаций и уровень доступа	Считали открытый архив недоступным
Реакция	Исправление, ограничение доступа, повторный тест	Удалили исходник, но оставили индекс

Приложение 43. Техническое испытание на извлечение данных из языковой модели

Цель испытания состоит не в доказательстве отсутствия каких-либо персональных данных внутри параметров модели, что практически недостижимо простым тестом, а в оценке конкретных механизмов извлечения в заданной среде. До начала работы фиксируются версия весов, системные инструкции, доступность внешних инструментов, контекстное окно, фильтры вывода и интерфейсы. Модель с закрытым API и та же модель с опубликованными весами представляют разные поверхности атаки. [3][12]

Группа испытаний формирует защищённые контрольные строки, заранее известные только аудитору и разработчику в пределах разрешённого доступа. Нельзя создавать настоящую утечку: реальные номера документов, адреса клиентов и сведения о преследовании не помещаются в открытые тестовые промпты. Допускаются синтетические маркеры,

контрольные секреты и материал, доступный группе на законном основании; их связь с обучением подтверждается документами, а не догадкой. Дополнительная группа тестов использует перефразирование, незавершённые цитаты, вариации языка, форматирования, длины контекста и неинформативный шум. [3][17]

Результаты классифицируются по точности: дословное восстановление, частичное совпадение, вероятностный пересказ, ответ, ранее доступный в публичном источнике, и ложное совпадение. Если внешний источник мог дать ту же строку, результат нельзя без дополнительного обоснования считать запоминанием обучающего корпуса. Протокол должен хранить гипотезу, последовательность запросов, параметры генерации, ответы, логи и критерии успешности, выбранные до начала испытания. [3]

Измерение	Определение результата	Ограничение
Exact match	Совпадение с контрольной строкой по заранее определённому правилу	Не показывает, как строка попала в модель
Partial match	Доля совпадающих фрагментов по нормализованной метрике	Легко завышается на шаблонных фразах
Targeted extraction	Результат запроса, нацеленного на конкретную запись	Нельзя масштабировать вывод на всех субъектов
Untargeted extraction	Обнаружение непредвиденных совпадений в сериях запросов	Требует отдельного анализа множественных сравнений

Заключение обязательно сообщает бюджет запросов и границы теста. Отсутствие успешной атаки при 300 испытаниях не доказывает нулевой риск. Это лишь результат ограниченного эксперимента, который нужно интерпретировать с учётом статистической неопределённости и модели противника.

Приложение 44. Membership inference: интерпретация точности на редком событии

Атака на принадлежность обучающей выборке пытается определить, использовалась ли конкретная запись при обучении модели. Оценивать её нужно не одной общей accuracy, а отдельно по базовой распространённости членов обучающего набора в атакуемой популяции, чувствительности и частоте ложноположительных выводов. Если лишь один процент проверяемых лиц действительно входит в обучающий корпус, даже сравнительно хорошая атака способна давать больше ложных обвинений, чем корректных выводов. [3][13]

Рассмотрим условный эксперимент: среди 10 000 проверяемых записей членами корпуса являются 100. Атака обнаруживает 70 из этих 100 (чувствительность 70%) и ошибочно отмечает 5% из оставшихся 9 900 (495 ложноположительных). Всего получается 565 положительных сигналов, из которых только 70 верны. Положительная прогностическая ценность составляет 70/565, то есть около 12,39%, несмотря на 70-процентную чувствительность. Пример иллюстративный, не относится к конкретной модели и не заменяет независимого тестирования.

Этот расчёт показывает, почему юридическое заключение не должно смешивать «положительный результат атаки» с фактически установленным членством человека в базе. При исследованиях на реальных людях нужны независимые основания истинной разметки и процесс обжалования ошибок. В отчёте приводятся доверительные интервалы, процедура

калибровки порога, разделение обучающего и тестового материала, частота повторных попыток и случаи, где результат получен только после выбора удачных запросов задним числом. [1][3]

Показатель	Значение учебного расчёта	Интерпретация
Base rate	100 / 10 000 = 1%	Редкая принадлежность
True positives	70	Из 100 членов
False positives	495	5% из 9 900 не-членов
Precision / PPV	70 / 565 ≈ 12,39%	Большинство положительных сигналов ошибочны

Для реального дела вывод «материал находился в обучающем наборе» допускается только после независимой проверки происхождения материала. Иначе организация рискует создать новое ложное утверждение о лице в процессе проверки анонимизации.

Приложение 45. Реестр рисков для векторных баз и retrieval-augmented generation

В системах RAG персональные данные могут сохраняться сразу в нескольких объектах: исходном файле, извлечённом тексте, фрагменте, векторном представлении, метаданных, поисковом индексе, логе запроса и ответе модели. Удаление идентификатора из исходного PDF не означает автоматической очистки созданного ранее поискового индекса или кеша. Проверка охватывает всю цепочку создания, извлечения и повторной выдачи, а не только исходное хранилище. [1][17]

Аудитор задаёт тестовый запрос, позволяющий найти конкретное дело по неименным признакам: номеру заседания, сочетанию юрисдикций, временной линии и сумме актива. Тестирование проводится в разрешённой среде на синтетических или обезличенных моделях дел, чтобы не раскрывать реальных клиентов. Оценивается, возвращает ли система исходный фрагмент, подсказывает ли название или предоставляет достаточно последовательных деталей для внешнего связывания. Техническое сходство эмбеддингов само по себе не устанавливает идентичность, однако в сочетании с метаданными способно увеличить риск. [2][15]

В отчёте фиксируются права пользователя, версия индексатора, функция разбивки, политика ретеншна, порог сходства и наличие фильтров. После тестового удаления документа повторяют индексирование и запросы ко всем связанным копиям. Если индекс удалён, но кэширование ответа продолжает возвращать фрагмент, корректирующее действие не завершено. В публичном отчёте предоставляется общая категория уязвимости, а не текст найденного конфиденциального документа.

Слой	Что проверяется	Подтверждение устранения
Источник	Копии документа и метаданные	Инвентаризация объектов и удаление
Чанк	Фрагменты и цитаты	Повторный поиск уникальных маркеров
Вектор	Индекс и резервные реплики	Контроль реконструкции/поиска после удаления

Слой	Что проверяется	Подтверждение устранения
Логи	Промпты, ответы и трассы	Настройки ретеншна и ограничение доступа
Интерфейс	Выдача по косвенным признакам	Регрессионные тесты ролей и поисковых запросов

Приложение 46. Сравнительный протокол для ЕС, Великобритании и США

Не существует международного универсального сертификата, который автоматически признаёт любой массив анонимным во всех правопорядках. В ЕС вывод зависит от определения персональных данных и преамбулы 26 GDPR, а также конкретного контекста обработки. В Великобритании действует UK GDPR с собственным правоприменением ICO, включая практический подход к «мотивированному нарушителю». В США федеральные и отраслевые правила используют различные понятия деидентификации, а отдельные режимы предусматривают собственные критерии и исключения. [1][15][19]

Поэтому международный проект должен начинаться с матрицы потоков данных: откуда данные поступают, кто выполняет обезличивание, кто получает результат, где он будет храниться, используется ли повторно для обучения модели и может ли принимающая сторона связать его с собственными реестрами. Оценка не переносится автоматически между юрисдикциями, даже если техника преобразования не меняется. Юридический консультант отдельно рассматривает обязанности исходного контролёра, передачу персональных данных до анонимизации и возможный статус результата у конкретного получателя.

Сравнительный доклад не подменяет национальную экспертизу. Особенно это касается уголовно-процессуальных сведений, иммиграционных материалов, биометрии и медицинских данных, где могут действовать специальные запреты. Формула «данные деидентифицированы по американскому стандарту» не доказывает выполнение требований европейского GDPR. Равно и вывод о низком риске для исследовательского партнёра не означает безопасности глобальной интернет-публикации. [1][15]

Правопорядок/уровень	Предмет проверки	Что нельзя предполагать
ЕС, GDPR	Персональный статус, средства идентификации, законность операции анонимизации	Наличие единого нормативного численного порога риска
Великобритания, UK GDPR / ICO	Практическая идентифицируемость и мотивированный нарушитель	Полная тождественность тесту ЕС в каждой детали
США	Применимый федеральный, штатный или секторный режим	Единый правовой статус для всех видов данных
Международный проект	Передачи, роли сторон, дальнейшее использование	Экстерриториальная применимость одного сертификата

Приложение 47. Принцип минимального раскрытия при судебном и регуляторном споре

Работа над юридическим делом часто требует сохранять индивидуальные факты: дату задержания, суд, маршрут перемещения, реквизиты судебного решения или имя

представителя. Эти признаки могут быть необходимы для доказательства нарушения прав, но одновременно создают риск идентификации человека в публичном аналитическом докладе. Решение о публикации принимают отдельно от решения о сохранении материала для адвоката, суда или уполномоченного органа. [1]

Внутри дела следует иметь две версии: доказательственную, защищённую режимом доступа и процессуальными гарантиями, и публичную, содержащую лишь минимально необходимую фактуру. Каждое преобразование регистрируется: замена прямых идентификаторов, агрегирование дат, обобщение географии, исключение редких обстоятельств, создание синтетического случая. Для проверяемости выводов независимый рецензент может получить ограниченный доступ к оригиналу на законном основании; сам факт такой проверки не требует опубликования оригинала. [1][2]

Аудитор отдельно проверяет так называемую уникальность нарратива: сочетание известных обстоятельств способно однозначно вывести на конкретного человека, даже если фамилия отсутствует. Нельзя ограничиваться средствами автоматического распознавания имён. Отдельно описывают источники косвенной идентификации: публичные судебные карточки, пресс-релизы органов, реестры компаний, сообщения о международном розыске и старые публикации с личными фотографиями. [4][15]

Слой работы	Пример разрешённого действия	Красная линия
Защищённая доказательственная база	Хранение полной версии при наличии основания и ограничений	Несанкционированная публикация оригинала
Аналитическая проверка	Тест уникальности сюжета	Экспериментальное раскрытие живого лица внешним тестером
Публичный отчёт	Агрегирование либо условный кейс	Утверждение об анонимности без оценки косвенных признаков
Судебное раскрытие	Процессуально необходимая передача	Автоматическая эквивалентность с открытой публикацией

Приложение 48. Испытание согласованности независимых оценщиков

Два специалиста могут по-разному оценивать вероятность идентификации: технический эксперт обращает внимание на качество классификатора, юрист — на достижимость информации для конкретной стороны, специалист по безопасности — на возможности мотивированного атакующего. Чтобы спор не превращался в голосование по авторитету, методика должна заранее задавать критерии, шкалу неопределённости и порядок независимого рассмотрения одного и того же сценария. [2][15]

Каждый эксперт получает одинаковое описание данных, но проводит первичную оценку самостоятельно, без сведений о выводе коллеги. После этого расхождения классифицируются по причине: разные фактические предположения, разная оценка технической возможности, различное понимание юридического стандарта либо неполная информация. Согласование допустимо лишь с протоколом аргументов. При серьёзном расхождении решение о выпуске откладывается, а группа запрашивает дополнительные данные или привлекает третьего рецензента. [2]

Качество работы измеряется не долей одинаковых выводов любой ценой, а точностью процедурного разграничения. Согласие двух экспертов, использовавших один и тот же неверный набор фактов, не гарантирует верности заключения. Поэтому контроль включает эталонные сценарии с заранее известной скрытой связью, оценки ложных положительных и отрицательных решений, а также повторение теста через заранее определённое время. [3]

Вид расхождения	Пример	Требуемая реакция
Фактическое	Один эксперт не учёл публичный судебный реестр	Обновить исходные данные и повторить анализ
Техническое	Разная успешность linkage-атак	Зафиксировать код, среду и воспроизводимость
Юридическое	Иной релевантный получатель	Отдельно описать применимый контекст
Неразрешённое	Не хватает доступа к ключевому источнику	Не делать безусловный вывод об анонимности

Приложение 49. План реагирования на повторную идентификацию после публикации

Анонимность не является необратимой характеристикой набора, которую можно однажды подтвердить и навсегда забыть. Внешние источники, методы машинного сопоставления и сама среда публикации со временем меняются. Если появляются новые сведения, позволяющие установить, к кому относится ранее опубликованный профиль, необходимо рассматривать событие как сигнал для срочного пересмотра статуса данных и возможного инцидента безопасности. Последствия по GDPR зависят от конкретных обстоятельств; далеко не всякий сигнал автоматически запускает предусмотренную статьями 33–34 процедуру уведомления, но игнорировать его также нельзя. [1][2]

Первый этап — остановить дальнейшее распространение, если это технически возможно, сохранить доказательства события в защищённой среде и описать реальные сведения, ставшие доступными третьим лицам. На втором этапе специалист по защите данных оценивает: является ли раскрытый материал персональными данными применительно к данному контексту; имелось ли нарушение безопасности; существуют ли риски для прав и свобод лиц; какие обязательства по уведомлению и документации возникают. Третий этап включает исправление страницы, ограничение выдачи по API, работу с индексаторами и уведомление известных получателей там, где это предусмотрено применимыми правилами. [1]

Для особо чувствительных случаев — сведения о преследовании, убежище, медицинском состоянии и транзакциях — следует отдельно рассмотреть риск вторичной виктимизации. Публичное сообщение о том, что «персона X была повторно идентифицирована», способно увеличить вред. Внутренняя карта события содержит фактическую связь с лицом только в пределах разрешённого доступа; внешний отчёт описывает вид ошибки без повторного раскрытия деталей.

Период	Ответственный	Действие	Подтверждение
Первичная реакция	Владелец публикации и безопасность	Приостановка дальнейшей выдачи, сохранение	Журнал ограничений и копия состояния

Период	Ответственный	Действие	Подтверждение
		следов	
Правовая квалификация	DPO / юрист	Определение статуса данных и применимых норм	Мотивированный акт с правовой основой
Коррекция	Команда публикации	Изменение, отзыв, уведомление адресатов по условиям	Реестр релизов и подтверждения
Последующий контроль	Независимый рецензент	Регрессионный тест и оценка остаточного риска	Заключение с ограничениями

Приложение 50. Договорная матрица для независимого тестирования анонимности

Внешняя проверка создаёт новые риски, поскольку аудитору может понадобиться доступ к массиву, предназначенному для обезличивания. В соглашении необходимо описать роли сторон применительно к каждой операции. Нельзя заранее объявить аудитора исключительно «обработчиком» или «независимым контролёром» без анализа целей и средств, которые он определяет. Договор должен соответствовать фактическим отношениям и применимым требованиям GDPR. [1]

Техническое задание включает версию набора, цель проверки, круг допустимых атак, методы доступа к вспомогательным данным и запреты на дальнейшее использование информации. Запрещается запускать несанкционированные атаки против живых лиц и пытаться получить их реальные данные лишь ради демонстрации риска. Необходимы документируемая законная основа, минимизация исходного доступа и возможность прекратить испытание, когда тест способен создать дополнительное раскрытие.

Отдельно регулируется публикация результатов. Коммерческая зависимость аудитора, возможность заказчика редактировать неудобный вывод или вознаграждение только за положительное заключение подрывают независимость проверки. Договор должен предусматривать документирование отрицательных результатов, порядок исправления фактических ошибок, процедуру разногласий и возможность честно сообщить о границах тестирования. [2][15]

Элемент договора	Минимальное условие	Проверка исполнения
Объект	Версии массивов, модели и сред	Идентификаторы и хеши
Право доступа	Цели, роли, место обработки, запрет внешнего переноса	Матрица ролей и журнал сеансов
Допустимые испытания	Методы, остановки, запрет реального ущерба	Утверждённый план теста
Уничтожение	Сроки и подтверждение удаления рабочих копий	Акт уничтожения и резервных копий
Независимость	Отсутствие права заказчика скрывать отрицательный результат	Хартия и протокол конфликтов

Приложение 51. Управление дифференциальной приватностью в серии публикаций

Дифференциальная приватность описывает математическую устойчивость механизма к участию отдельного субъекта в смежных наборах данных. Она не создаёт безусловную анонимность любого исходного файла, не отменяет требования к исходной законности обработки и не является самостоятельным разрешением на публикацию чувствительных сведений. Надёжность обещания зависит от формального определения соседства, выбора механизма, корректного генератора случайности, настроек приватности и условий композиции. [1][13]

Для практического журнала вводится единица учёта: идентификатор защищаемой единицы (лицо, эпизод либо запись), период релиза, заявленные epsilon и delta, механизм и способ композиции. Нельзя складывать или усреднять гарантии произвольно, особенно если несколько подразделений публикуют статистику об одних и тех же людях. Вспомогательные незащищённые релизы, метаданные и логи могут обойти заявленное преимущество DP и должны рассматриваться отдельно. [13]

Учебный пример: механизм A имеет бюджет epsilon=0,4, механизм B — epsilon=0,6. В простой последовательной композиции при совместимом определении защищаемой единицы верхняя оценка epsilon может составлять 1,0; фактические правила композиции и общая delta требуют отдельного расчёта. Это не оценка вероятности реидентификации в один процент и не установленный юридический порог допустимого риска. До одобрения каждый механизм проверяется на корректность реализации, включая генерацию шума, защиту seed, доступность повторных запросов и утечки через исключения. [13]

Вопрос	Минимальная запись	Типичная ошибка
Что защищаем?	Лицо/запись/эпизод и модель соседства	Отчёт использует неодинаковые единицы
Какая гарантия?	epsilon, delta, механизм, дата	Называет epsilon «процентом утечки»
Сколько релизов?	Единый журнал запросов и бюджет	Каждая команда учитывает только свою публикацию
Сохраняется ли полезность?	Ошибка, смещение, доверительные границы	Шум приводит к неверному правовому выводу

Приложение 52. Оценка малых выборок и вторичного подавления

В отчётах о санкциях, розыске и миграционных решениях группы обычно малочисленны. Даже если исследователь решает подавить ячейку с одним человеком, значение этой ячейки можно восстановить из опубликованных итогов. Например, если известно, что всего в подразделении рассматривалось 12 дел, а по открытым категориям распределено 5 и 6, скрытая третья категория вычисляется как 1. Поэтому первичное подавление должно сопровождаться анализом вторичного раскрытия. [4][12]

Методика включает построение системы доступных уравнений по итогам таблиц, проверку однозначно восстанавливаемых скрытых значений и поиск информации, которую раскрывают разности последовательных отчётов. Для серий статистики задают минимальную частоту публикации и допустимые границы точности. Если исследователь

публикует общий итог и несколько чувствительных разрезов, нужно проверить, не раскрывает ли последнее число все скрытые строки. [12]

Подавление не следует применять механически. Иногда обобщение географии или времени полезнее, чем скрытие отдельного столбца; иногда единственная безопасная мера — отказаться от публикации узкого разреза. Если показатель используется для анализа возможной дискриминации, следует также оценить последствия потери статистической мощности. Методика должна ясно объяснять, какие выводы остаются возможными, а какие становятся недопустимыми после защиты. [1]

Опубликовано	Значение	Что можно восстановить
Общая численность	12	Известен итог
Категория A	5	Известна часть
Категория B	6	Известна часть
Категория C скрыта	«малая группа»	12 − 5 − 6 = 1; скрытие неэффективно

Приложение 53. Управление исходными доказательствами и научной воспроизводимостью

Юридическая проверка анонимности требует воспроизводимости, но полное открытие исходного набора может нарушить саму цель защиты данных. Поэтому система доказательств должна быть многоуровневой. Публичная часть содержит метод, агрегированные показатели, мотивировку риска и перечень ограничений. Закрытая часть хранит контрольные образцы, конфигурации атак, результаты тестов и описания использованного вспомогательного материала. Доступ к закрытой части предоставляется только на законном основании и для конкретных проверочных целей. [1][2]

Внутренние доказательства следует сохранять с версиями данных и скриптов, фиксированными параметрами, временем испытаний, именами ответственных и криптографическими контрольными суммами. Контрольная сумма помогает удостовериться, что определённые байты не изменились, но не доказывает правдивость содержимого и не оправдывает его получение. Целостность, происхождение, законность и достоверность — разные вопросы, требующие разных доказательств.

Особое значение имеет отрицательная доказательственная оценка. Если тест не нашёл успешных попыток связывания, необходимо документировать покрытие сценариев, вычислительный бюджет, начальные предположения, время и ограничения исследовательской среды. Отчёт «атаки не обнаружены» не должен превращаться в утверждение «повторная идентификация невозможна». Отдельный рецензент оценивает не только результаты тестов, но и риск пропуска из-за выбранных методов. [3]

Материал	Открытая часть	Защищённая часть
Методика	Типы атак и предположения	Детальные запросы, если они повышают риск
Результаты	Агрегаты и интервалы неопределённости	Построчные логи и контрольные метки
Исходные данные	Описание категории и объёма	Минимально необходимая выборка

Материал	Открытая часть	Защищённая часть
		при законном доступе
Коррекция	Публичная история редакций	Журнал лиц и технических операций

Приложение 54. Система порогов эскалации и повторной проверки

Независимый Observatoire не должен раз и навсегда фиксировать единый численный «допустимый риск» для всех данных. Решение зависит от тяжести возможного вреда, возможностей атакующего, дальнейших каналов распространения и реалистичности мер контроля. Отчёт с высокими рисками может быть разрешён для строго ограниченного круга уполномоченных лиц, но недопустим для открытой публикации. Напротив, крайне общий агрегат может быть опубликован без сохранения излишних идентификаторов в рабочем наборе. [1][2]

Для управления решениями создаются уровни эскалации: обычная проверка, усиленная проверка независимым специалистом, решение комиссии с юридическим участием и временный запрет выпуска. Срабатывание порога зависит от выявленной уникальной записи, вероятной идентификации, возможности восстановить особо чувствительный атрибут, отсутствия законной базы исходной обработки или неконтролируемой передачи третьей стороне. Каждое отклонение от рекомендованной меры требует письменного обоснования ответственного должностного лица.

Повторная проверка осуществляется не по формальной годовщине заключения, а также при наступлении событий: новая открытая база, смена платформы публикации, добавление результатов другой организации, дообучение AI-модели, изменение индекса или появление массового инструмента сопоставления. Регулярный календарь всё же нужен для обнаружения медленных изменений. [2][15]

Триггер	Уровень реагирования	Ожидаемое решение
Новый внешний источник с пересекающимися данными	Усиленная проверка	Тест linkage и пересмотр статуса
Подтверждённая идентификация живого лица	Критический	Ограничение доступа и правовая квалификация события
Обновление модели/индекса	Технический	Регрессионные атаки и новая версия акта
Изменение цели или получателя	Юридический	Повторная оценка оснований и контекста

Приложение 55. Обращения лиц и независимое рассмотрение спорных выводов

Лицо, считающее, что опубликованный набор позволяет установить его личность или выводить о нём чувствительную информацию, должно иметь ясный канал обращения. Требовать от заявителя публично раскрыть свою идентичность в подтверждение жалобы опасно и нередко избыточно. Достаточно предусмотреть защищённый маршрут первичной оценки, при котором подробные сведения принимаются только в необходимом объёме и доступны ограниченному числу специалистов. [1]

Процесс имеет две независимые ветви. Первая устанавливает, имеется ли фактическая связь с указанным лицом и какие средства идентификации реально доступны; вторая проверяет правильность исходного решения об анонимности и необходимость исправления. Если опубликованное содержание является персональными данными для соответствующего контролёра, затем определяются применимые права субъекта и сроки их реализации. Конкретный набор прав и исключений оценивается по GDPR, а не по внутреннему регламенту исследовательского центра. [1]

Организация должна избегать конфликта интересов: специалист, подготовивший спорную публикацию, не может единолично рассматривать жалобу на неё. Решение содержит факты, пределы исследования, предпринятые меры, порядок обращения к компетентному надзорному органу и процедуру исправления. Для статистических публикаций, где невозможно установить индивидуальную связь, допускается самостоятельный анализ уязвимости без излишнего сбора личности заявителя.

Этап	Запрашиваемые сведения	Риск, которого следует избегать
Первичный контакт	Ссылка и описание опасения	Сбор дополнительных паспортных данных без необходимости
Техническая проверка	Минимальные признаки и тестовый сценарий	Новая публикация чувствительного совпадения
Правовая оценка	Роль сторон и правовые основания	Автоматическое отрицание применимости GDPR
Исправление	Технические меры и коммуникация	Обещание полного удаления уже распространённых копий

Приложение 56. Методика оценки качества и полезности анонимного результата

Сильная защита не обязательно означает полезный результат. Если при подготовке доклада изъять все редкие события, исследование систематически недооценит нарушения в малых группах. Например, исключение единичных случаев незаконного задержания способно исказить оценку уязвимостей именно тех лиц, чьи права требуют особой защиты. Задача аудитора — не доказать, что публикация имеет нулевой риск, а прозрачно сопоставить уровень защиты с исследовательской целью и подобрать менее рискованный способ выпуска. [1][12]

Проверка качества начинается с перечня исходных исследовательских вопросов: сравнение периодов, выявление аномалий, географическое распределение, оценка ошибок алгоритмов. Для каждого вопроса указывают необходимую точность, допустимую потерю информации и последствия ошибочного вывода. Затем измеряют влияние преобразования: изменения распределений, исчезнувшие категории, смещение среднего, изменение долей и ширину интервалов. Если эти показатели существенно ухудшаются, требуется либо другой способ обезличивания, либо ограниченный исследовательский доступ к более подробной версии. [12][13]

При использовании синтетики сопоставляются не только средние значения, но и редкие хвостовые события, зависимости между признаками и возможные следы обучающих записей. Высокая статистическая похожесть не доказывает отсутствия утечек. Отдельно

проверяют, не превращается ли синтетический набор в средство воспроизведения частных историй. Публичное решение должно описывать остаточные научные ограничения, чтобы читатель не трактовал защищённые агрегаты как полную картину реальности. [3]

Цель исследования	Метрика полезности	Типичный компромисс
Частота событий	Разность долей и границы погрешности	Подавление малых ячеек изменяет оценку
Динамика во времени	Устойчивость тренда при агрегировании	Широкие интервалы скрывают краткие пики
Региональные различия	Искажение региональных рангов	Обобщение географии уменьшает детализацию
Ошибки AI-систем	Сохранение распределения false positive/negative	Редкие ошибки могут исчезнуть из релиза

Приложение 57. Комплексный сценарий: многоязычный судебный корпус

Условный исследовательский проект изучает решения об экстрадиции в пяти юрисдикциях. Оригиналы документов представлены на русском, французском, английском и двух национальных языках; наряду с официальными переводами исследователь использует автоматически полученные краткие резюме. Прямые имена в корпусе удалены, однако названия аэропортов, формулировки обвинений, суммы залогов и редкие последовательности судебных дат остаются. Внешние архивы содержат часть тех же сюжетов под разными транслитерациями.

При обычной проверке поиск по точному имени ничего не находит, но семантическая система объединяет две версии одного решения, а затем показывает связанное сообщение о задержании. Проблема не в ошибке регулярного выражения для поиска фамилии, а в структуре корпуса: редкая фактическая комбинация остаётся узнаваемой. Образец создаётся исключительно как учебный; он не отражает материалы конкретного клиента. Для воспроизведения аудитор формирует словарь вариантов транслитерации, временные окна, набор общеизвестных событий и отдельные тесты на переводческую эквивалентность.

Решение о выпуске предусматривает два продукта. Для внутреннего юридического анализа сохраняется защищённый оригинал, где доступ предоставляется по ролям; для публичного доклада используются агрегированные выводы и синтетические траектории, не воспроизводящие уникальные детали. Потеря части детализации признаётся открыто. Если исследователь не может обосновать, почему редкая комбинация перестала быть идентифицирующей, он не заявляет, что корпус полностью анонимизирован. [1][2][15]

Этап	Проверка	Критерий результата
Нормализация языков	Транслитерации, названия и временные форматы	Нет неучтённого совпадения из-за языка
Связывание	Судебные архивы и доступные реестры	Измерена уникальность событий
Модель	Поиск похожих фрагментов	Описано семантическое совпадение
Выпуск	Сравнение пользы и вреда	Ограничение или агрегирование

Этап	Проверка	Критерий результата
		подтверждено актом

Приложение 58. Комплексный сценарий: серия ежемесячных статистических бюллетеней

Организация ежемесячно публикует число отказов в банковском обслуживании по четырём географическим зонам. Каждая таблица объединяет события как минимум пяти человек, поэтому редакция считает публикацию безопасной. Однако бюллетень за март содержит накопленный итог с начала года, а за апрель — тот же показатель на новую дату. При вычитании цифр можно получить месячное изменение для группы из одного человека. Если дополнительно известно единственное местное сообщение о блокировке банковского счёта, аналитический вывод начинает относиться к конкретному субъекту.

Для аудита необходимо собрать не отдельный бюллетень, а последовательность всех опубликованных файлов, в том числе первоначальные версии и редакционные исправления. Затем строят модель доступных линейных комбинаций агрегатов и проверяют, можно ли из общих сумм однозначно восстановить скрытые ячейки. Это задача не только статистика: правовой эксперт должен объяснить, когда восстановимое значение может относиться к идентифицируемому физическому лицу. [1][12]

Меры реагирования включают согласование периодов публикации, диапазонное представление малых групп, вторичное подавление и раздельное хранение детальных расчётов. При этом следует документировать последствия изменения методики для временных рядов: опубликованное среднее до защиты может оказаться несопоставимым с показателем после неё. Аудитор не должен скрывать этот эффект, иначе читатель может ошибочно принять разрыв в графике за изменение реальной практики банков.

Источник риска	Почему обычного удаления имён недостаточно	Контроль
Разность накопленных итогов	Восстанавливает малый период	Проверка всех временных разрезов
Локальные новости	Связывает событие с лицом	Ретроспективный поиск внешних публикаций
Архивные исправления	Сохраняют разные версии сумм	Учёт опубликованных версий
Изменение метода	Создаёт искусственный тренд	Документирование статистической сопоставимости

Приложение 59. Комплексный сценарий: синтетическая база банковских транзакций

Разработчик создаёт учебный набор транзакций для проверки системы обнаружения подозрительных операций. Он сообщает, что данные синтетические, поскольку имена и идентификаторы сгенерированы. Вместе с тем генератор обучался на узком наборе реальных крупных сделок, а исходные суммы, даты и граф связей контрагентов обладают высокой уникальностью. Когда другой исследователь сопоставляет синтетический граф с публичными корпоративными реестрами, появляется правдоподобная связь с реальным бенефициаром.

Проверка должна рассматривать не только сходство средних и частот, но и структурную уникальность графа, ближайшие соседние записи, повторение редких последовательностей и возможность атрибутивного вывода. Для контроля используется законно сформированная закрытая тестовая выборка и независимые метрики. Если точная связь с реальным лицом обнаружена, сведения остаются в защищённом акте и не становятся иллюстрацией в публичной статье. [3][12]

Решение зависит от назначения набора. Для внутреннего тестирования алгоритма под контролируемым доступом возможно сохранение более подробных структур при соответствующей правовой основе. Для неограниченного распространения разработчик обязан обосновать отсутствие идентифицируемости для релевантных сторон; если это невозможно, он должен изменить генерацию, сократить уникальные признаки или заменить материал полностью условным. Слова «синтетический» и «деидентифицированный» не являются доказательством юридического статуса.

Компонент	Признак утечки	Мера
Суммы и даты	Повторение редкой сделки	Контроль близости и генерализация
Граф связей	Уникальная цепь юридических лиц	Структурные атаки и изменение топологии
Текст примечаний	Дословное повторение исходника	Тест memorisation и очистка генерации
Конечная публикация	Неограниченная агрегация внешними пользователями	Отдельная оценка открытого релиза

Приложение 60. Комплексный сценарий: фото, аудио и метаданные

Правовой доклад сопровождается фотографией конференции и кратким видеоинтервью, в котором намеренно не указываются фамилии участников. Однако из изображения возможно извлечь лицо, бейдж, экран компьютера и характерный интерьер; метаданные могут содержать координаты и время записи. Голос, акцент и описание рабочей должности способны стать дополнительными признаками. Обрезка изображения или удаление звуковой дорожки может снизить риск, но не всегда устраняет его полностью. [1][4]

Для аудита медиаматериала формируется отдельный перечень возможных идентификаторов: лица и контуры, таблички, бейджи, транспорт, экранные надписи, метаданные файла, субтитры и автоматические транскрипты. Оценивается связь с публичными фотографиями мероприятия, расписаниями выступлений и архивами трансляций. Наиболее опасны сочетания признаков: анонимизированное лицо может быть восстановлено по редкой должности и времени выступления, а удалённая надпись — прочитана по субтитрам. [15]

Обработка изображений и голоса не превращается автоматически в обработку биометрических данных специальной категории: правовая квалификация зависит от характеристик и цели обработки, включая вопрос уникальной идентификации. Это не отменяет необходимости оценить персональный статус материала и риски для лиц. При высокой вероятности вреда предпочтительно отказаться от исходного медиа и заменить его схемой или иллюстрацией, не построенной на защищённом человеке.

Объект	Тест	Возможное решение
Фото	Поиск сопоставимых изображений, OCR бейджа	Удалить исходник или сильно изменить содержание
Видео	Сцена, голос, субтитры, участники	Заменить нейтральной графикой
Метаданные	GPS/EXIF, дата, автор	Очистить и повторно проверить экспорт
Архивные копии	Поиск первичного медиа	Оценить невозможность полного отзыва

Приложение 61. Комплексный сценарий: научный доступ к редким медицинским и миграционным событиям

Исследователю необходимо оценить, связано ли состояние здоровья с определёнными процессуальными задержками по делам о предоставлении убежища. В региональной выборке есть несколько редких заболеваний и уникальные сочетания семейного статуса, даты слушания и маршрута переселения. Публичное раскрытие строк создало бы существенный риск идентификации; полное исключение редких случаев, напротив, ухудшило бы научную обоснованность выводов. [1]

Проект предусматривает раздельные уровни доступа. Публично выпускают метод, безопасные сводные показатели и интерпретацию неопределённости. Более подробная таблица остаётся в защищённой исследовательской среде; исследователи получают ограниченный доступ после проверки основания, цели и разрешённых процедур. Отдельно анализируют статьи 9, 10 и 89 GDPR в зависимости от фактической категории данных и правового режима исследования; общая надпись «для науки» не заменяет требуемого анализа. [1]

При публикации важно не допустить ложного ощущения точности. Если из-за малочисленности категории объединены, читатель должен понимать, что отдельный риск мог остаться скрытым. Внешний рецензент проверяет, почему выбран именно такой уровень обобщения, почему нельзя достигнуть той же цели менее детальными данными и какие средства предотвратят дальнейшее переиспользование информации. Отрицательный вывод о возможности открытой публикации не исключает законного защищённого исследования.

Цель	Среда	Доказательство контроля
Публичный обзор	Агрегаты	Тест малых групп и связи с внешними данными
Проверка статистики	Ограниченный доступ	Журнал ролей и операций
Рецензия	Защищённые материалы	Независимость и законное основание доступа
Архивирование	Контролируемое хранение	Сроки и план удаления

Приложение 62. Комплексный сценарий: подрядчик, который обещает «100% анонимность»

Международная организация получает предложение от внешнего поставщика, который гарантирует полную анонимность любого переданного ему набора после автоматической

обработки. Договор не содержит описания методов, условий доступа к исходникам и перечня сохранённых технических копий. Поставщик предлагает сертификат без предоставления испытаний и обещает, что любые данные после его обработки не регулируются GDPR. Такой подход нельзя считать достаточным ни технически, ни юридически. [1][2]

До передачи данных необходимо проверить, на каком основании сам поставщик получает персональные сведения, какую роль выполняет и допускает ли договор обучение его собственных моделей на клиентском корпусе. Затем запрашиваются документированные преобразования, остаточные уникальные комбинации, проверка разумно вероятных внешних связей, места обработки и правила удаления. Если поставщик ссылается на определённый стандарт, аудитор должен установить его предмет и границы: стандарт управления информацией не тождествен доказательству анонимности конкретного выпуска.

При отказе предоставить существенные доказательства не следует автоматически заключать, что данные идентифицируемы в любой среде, но и положительный вывод об анонимности сделать нельзя. Корректный результат — недостаточность доказательств и решение о неприменении предлагаемой процедуры к чувствительным данным до проведения необходимых испытаний. Это самостоятельный профессиональный вывод, который должен фиксироваться в отчёте и договорной переписке.

Заявление поставщика	Проверочный вопрос	Недостаточный ответ
«100% анонимность»	Для какой среды, какой стороны и каких средств идентификации?	«Наш алгоритм удаляет имена»
«Гарантировано стандартом»	Что именно сертифицировано?	Общая ссылка на ISO без предмета
«У нас нет копий»	Где журналы, индексы и бэкапы?	Устное заверение без акта
«GDPR больше не действует»	Каким образом обработаны исходные данные?	Отсутствие анализа самой операции

Приложение 63. Комплексный сценарий: повторное использование результата для обучения AI

Публикация статистического набора для научного отчёта и дальнейшее использование его для обучения модели — разные действия с потенциально отличающимися рисками. Даже если первоначальный контролёр оценил данные как анонимные в определённой среде, следующий разработчик может объединить их с другими источниками, добавить новые метки и обучить систему, способную выводить скрытые атрибуты. Поэтому паспорт релиза должен описывать предполагаемое переиспользование и ограничения первоначальной оценки. [1][3]

Проверка проводится в два этапа. Первый устанавливает, имелись ли персональные данные на момент передачи и какие обязанности возникали. Второй рассматривает фактические возможности нового получателя, включая его внешние массивы, доступ к весам, способ дообучения и вероятную способность извлечения. Договорное запрещение связывания может снижать риск неправомерного использования, но не является математическим доказательством анонимности. Вывод о статусе данных следует формулировать отдельно для каждого исследованного контекста.

В исследовательском отчёте не нужно заранее запрещать любое переиспользование. Вместо этого следует требовать новую оценку при изменении цели, аудитории и средств сопоставления. При необходимости указываются контролируемые исследовательские площадки и публичные агрегаты с большей устойчивостью к повторному использованию. Это оставляет возможность научной работы без ошибочного предположения, что исходный положительный акт действует для всех будущих способов обработки.

Изменение	Новый риск	Требуемая проверка
Добавлен внешний корпус	Linkage	Средства и данные нового получателя
Модель получает открытые веса	Extraction	Иной бюджет и класс атак
Сменена цель	Нарушение ограничений исходной обработки	Роли, основания и совместимость
Публичное API	Массовые запросы	Контроль вывода и лимитов

Приложение 64. Итоговая карта внедрения методики ARGA на 12 месяцев

Предлагаемая программа не является фактом уже осуществлённой аккредитации, аудита или признания со стороны органов ЕС. Это проект последовательного создания исследовательской компетенции. На первом этапе утверждаются хартия независимости, система ролей, правовая модель работы с исходными материалами и протокол защиты тестовых данных. На втором формируются учебные испытания и базовые наборы воспроизводимых сценариев. Третий этап посвящён независимой рецензии, публичному методическому отчёту и процедуре исправления ошибок. [1][2]

Не рекомендуется сразу публиковать индивидуальные реальные кейсы только ради демонстрации возможностей Observatory. Первые пилоты лучше проводить на юридически и технически безопасных условных данных, а затем на законно доступных материалах в контролируемой среде. Переход к публичным результатам допускается после проверки исходных правовых оснований, независимости рецензента и возможности исправления. Показателем зрелости является способность отказаться от опасного релиза, а не количество документов с положительным заключением.

В конце цикла следует провести внешний методический аудит: выборочно повторить испытания, сравнить решения разных рецензентов, проверить активные ссылки на нормативные источники, оценить обращения субъектов и эффективность реакции на ошибки. Если существенные пробелы не устранены, программа расширения приостанавливается. Этот механизм защищает исследовательскую репутацию лучше, чем количественный план публикаций.

Этап	Условный период	Проверяемый результат
Управление и право	1–3 месяцы	Роли, хартия, доступ, правовые основания
Методика и тесты	4–6 месяцы	Контрольные наборы и репозитории испытаний
Ограниченный пилот	7–9 месяцы	Независимые акты и исправления

Этап	Условный период	Проверяемый результат
Внешняя рецензия	10–12 месяцы	Отчёт о воспроизводимости и решении о масштабе

Источники и методические основания

[1] Regulation (EU) 2016/679 (GDPR), в особенности статьи 4–6, 9–11, 25, 32–35, 89 и преамбула 26

https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng

Статус и назначение: Обязательное право ЕС; применимость к конкретной операции оценивается отдельно.

[2] EDPB Guidelines 02/2026 on Anonymisation, version 1.0, 7 July 2026

https://www.edpb.europa.eu/public-consultations/guidelines-022026-on-anonymisation_en

Статус и назначение: Проект публичной консультации до 30 октября 2026; не финальная версия.

[3] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP216)

https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

Статус и назначение: Исторический методический ориентир, уточняемый проектом EDPB 02/2026.

[4] EDPB Opinion 28/2024 on certain data protection aspects related to AI models

https://www.edpb.europa.eu/documents/opinion-of-the-board-art-64/opinion-282024-on-certain-data-protection-aspects-related-to_en

Статус и назначение: Мнение EDPB от 18 декабря 2024 года об анонимности моделей и других вопросах GDPR.

[5] CNIL. L’anonymisation de données personnelles

https://www.cnil.fr/fr/technologies/lanonymisation-de-donnees-personnelles

Статус и назначение: Практические критерии эффективности анонимизации.

[6] CNIL. IA: Analyser le statut d’un modèle d’IA au regard du RGPD (22 juillet 2025)

https://cnil.fr/fr/ia-analyser-le-statut-dun-modele-dia-au-regard-du-rgpd

Статус и назначение: Метод оценки статуса модели по GDPR.

[7] Court of Justice of the EU. EDPS v SRB, C-413/23 P, judgment of 4 September 2025

https://juris.curia.europa.eu/juris/document/document_print.jsf?docid=303863&doclang=EN

Статус и назначение: Решение о перспективах идентифицируемости и исходной обязанности информирования.

[8] Court of Justice of the EU. Breyer, C-582/14, judgment of 19 October 2016

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A62014CJ0582

Статус и назначение: Значение разумно доступных средств получения дополнительных идентификаторов.

[9] Court of Justice of the EU. IAB Europe, C-604/22, judgment of 7 March 2024

https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A62022CJ0604

Статус и назначение: Машиночитаемые строки и понятие совместного контролёра.

[10] EDPB Guidelines 01/2025 on Pseudonymisation, version 1 (public consultation)

https://www.edpb.europa.eu/public-consultations/guidelines-012025-on-pseudonymisation_en

Статус и назначение: Опубликованная для консультации версия; не обозначена как финальная.

[11] EDPB Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, final version

https://www.edpb.europa.eu/documents/guideline/guidelines-42019-on-article-25-data-protection-by-design-and-by-default_en

Статус и назначение: Окончательное руководство 2020 года о защите данных при проектировании.

[12] NIST SP 800-188. De-Identifying Government Datasets: Techniques and Governance

https://csrc.nist.gov/pubs/sp/800/188/final

Статус и назначение: Техническое руководство, сентябрь 2023; не устанавливает право ЕС.

[13] NIST SP 800-226. Guidelines for Evaluating Differential Privacy Guarantees

https://csrc.nist.gov/pubs/sp/800/226/final

Статус и назначение: Окончательная версия марта 2025 года; методология оценки DP.

[14] ENISA. Data Pseudonymisation: Advanced Techniques and Use Cases

https://www.enisa.europa.eu/publications/data-pseudonymisation-advanced-techniques-and-use-cases

Статус и назначение: Отчёт 2021 года о технических мерах и границах псевдонимизации.

[15] UK ICO. How do we ensure anonymisation is effective?

https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/anonymisation/how-do-we-ensure-anonymisation-is-effective/

Статус и назначение: Руководство ICO, с уведомлением о пересмотре после Data (Use and Access) Act.

[16] UK ICO. Pseudonymisation

https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/anonymisation/pseudonymisation/

Статус и назначение: Технические и юридические ограничения псевдонимизации в режиме UK GDPR.

[17] CNIL. Développement des systèmes d’IA: recommandations pour respecter le RGPD

https://www.cnil.fr/fr/developpement-des-systemes-dia-les-recommandations-de-la-cnil-pour-respecter-le-rgpd

Статус и назначение: Официальные рекомендации для разработки ИИ.

[18] CNIL. IA: Informer les personnes concernées

https://cnil.fr/fr/ia-informer-les-personnes-concernees

Статус и назначение: О прозрачности обработки, включая риски запоминания и регургитации.

[19] NIST Privacy Framework, Version 1.0 (January 2020)

https://www.nist.gov/privacy-framework/privacy-framework

Статус и назначение: Добровольный риск-ориентированный инструмент, не обязательное право.

[20] CNIL. Mobiliser la base légale de l’intérêt légitime pour développer un système d’IA

https://www.cnil.fr/fr/base-legale-interet-legitime-developpement-systeme

Статус и назначение: Меры для снижения запоминания и экстракции в контексте разработки моделей.