Commercial Surveys Bimonthly
دوماهنامه بررسی‌های بازرگانی
Vol.23, No.132, Aug-Sep 2025, 65-86
سال بیست و سوم، شماره ۱۳۲، مرداد و شهریور ۱۴۰۴، ۸۶-۶۵

Legal Challenges of Anonymization and Re-identification of Personal Data

https://doi.org/10.22034/bs.2025.2056995.3090

Amir Mohammad Ghorban Nia, Master's Graduate in Private Law, University of Tehran, Tehran, Iran.
Zahra Shakeri*, Associate Professor, Faculty of Law and Political Science, University of Tehran, Tehran, Iran.

Accepted: 1 Apr 2025 Revised: 11 Aug 2025 Received: 21 Sep 2025

Data Anonymization / Data Re-identification / Informational Privacy / Personal Data

Anonymization of personal data is a key mechanism for privacy protection and compliance with legal requirements in data processing. However, challenges such as the technical complexity and cost of proper implementation, reduced data accuracy and usability, and the existence of vast amounts of auxiliary data about individuals, which may enable re-identification, undermine the effectiveness of anonymization. This study, employing a descriptive-analytical method, addresses the legal gap in Iranian law regarding anonymization, the discrepancies in international definitions, and the lack of detailed regulations and the associated shortcomings in international legal frameworks. The article examines issues such as the legal criteria for determining whether data is truly anonymized, how the risk of identity disclosure can be minimized, and which technical methods are most suitable for anonymization. The findings indicate that anonymization is highly vulnerable to re-identification when auxiliary data is available. Consequently, policymakers and data governance stakeholders should move beyond reliance on anonymization as a definitive solution and adopt a multi-layered approach to mitigate re-identification risks, thereby formulating a novel policy for the sharing and disclosure of personal data.

Data Availability
The data used or generated in this research are presented in the text of the article.

Conflicts of interest
The authors of this paper declared no conflict of interest regarding the authorship or publication of this article.

* Corresponding Author: zshakeri@ut.ac.ir

دوماهنامه بررسی‌های بازرگانی
Commercial Surveys Bimonthly
سال بیست و سوم، شماره ۱۳۲، مرداد و شهریور ۱۴۰۴، ۸۶-۶۵
Vol.23, No.132, Aug-Sep 2025, 65-86
مقاله پژوهشی
https://doi.org/10.22034/bs.2025.2056995.3090

چالش‌های حقوقی ناشناس‌سازی و بازشناسایی داده‌های شخصی

دریافت: ۱۲ / ۰۱ / ۱۴۰۴ بازنگری: ۲۰ / ۰۵ / ۱۴۰۴ پذیرش: ۳۰ / ۰۶ / ۱۴۰۴

امیرمحمد قربان‌نیا[۱]
زهرا شاکری[۲] (نویسنده مسئول)

چکیده

ناشناس‌سازی داده‌های شخصی، یکی از راه‌کارهای کلیدی در حفاظت از حریم خصوصی و رعایت الزامات قانونی پردازش داده‌ها است؛ لکن چالش‌هایی مثل پیچیدگی فنی و هزینه‌های اجرای صحیح این روش‌ها، کاهش دقت و کاربرد داده‌ها و وجود حجم زیادی از داده‌های مکمل راجع به اشخاص و امکان بازشناسایی ایشان، مانع تاثیرگذاری حداکثری ناشناس‌سازی می‌شوند. مقاله حاضر از نوع کاربردی و با روش توصیفی-تحلیلی، با عنایت به سکوت قوانین ایران درمورد ناشناس‌سازی و اختلاف بین تعاریف خارجی و همینطور عدم پرداخت به جزییات و کاستی‌های آن در مقررات بین‌المللی، به معرفی ناشناس‌سازی و پاسخ‌دهی به مسائلی مانند اینکه معیار تشخیص داده‌های ناشناس از منظر قانونی چیست و اینکه چگونه می‌توان ریسک افشای هویت را کاهش داد و چه روش‌های فنی برای اجرای ناشناس‌سازی مناسب‌تر است؛ می‌پردازد. یافته‌ها و نتایج نشان می‌دهد که ناشناس‌سازی در برابر داده‌های مکمل، بسیار آسیب‌پذیر است. در نتیجه سیاست‌گذاران و فعالان حوزه داده باید به جای اتکا به ناشناس‌سازی به‌عنوان راه‌حلی قطعی، با رویکردی چندلایه برای کاهش ریسک بازشناسایی به تدوین سیاستی نوین برای اشتراک‌گذاری یا انتشار داده‌های شخصی بپردازند.

طبقه‌بندی JEL: K24، K33، K42، O33، L86

بازشناسایی داده‌ها / حریم خصوصی اطلاعاتی / داده‌های شخصی / ناشناس‌سازی داده‌ها

۱. دانش آموخته کارشناسی ارشد، حقوق خصوصی، دانشگاه تهران، تهران، ایران amir.ghorbannia@ut.ac.ir
۲. دانشیار دانشکده حقوق و علوم سیاسی دانشگاه تهران، تهران، ایران zshakeri@ut.ac.ir

۱. مقدمه: طرح مسئله

با پیدایش فناوری‌های نوین و توسعه تجارت الکترونیک، داده‌های شخصی به دارایی‌هایی ارزشمند برای سازمان‌ها، شرکت‌ها و نهادهای دولتی تبدیل شده‌اند و در تحلیل‌های کلان داده‌ها، تصمیم‌گیری‌های تجاری و سیاست‌گذاری‌های عمومی نقش کلیدی ایفا می‌کنند (هدایت نظری و محمدی، ۱۴۰۳، ص ۲۸). با این حال، استفاده‌های ثانویه از این داده‌ها، به‌ویژه فراتر از مقاصدی که اشخاص موضوع داده در هنگام جمع‌آوری به آن رضایت داده‌اند، با عدم شفافیت، چالش‌های جدی در حوزه حریم خصوصی و حفاظت از داده‌ها ایجاد کرده است (ناصر و تاج، ۱۳۹۹، ص ۵۹). این مسئله نوعی تعارض میان منافع اقتصادی بهره‌برداران داده‌ها و حق افراد بر کنترل اطلاعات شخصی خود را نشان می‌دهد. یکی از مهم‌ترین راه‌کارهایی که برای کاهش این تعارض و پیشنهاد شده، ناشناس‌سازی (Anonymization) داده‌ها است.

ناشناس‌سازی فرایندی است که با حذف یا تغییر اطلاعات هویتی، امکان شناسایی مجدد افراد را کاهش می‌دهد و از این طریق، پردازش داده‌ها را با الزامات قانونی و اخلاقی سازگارتر می‌سازد. این فرایند شامل روش‌هایی همچون حذف شناسه‌ها، تعمیم، تصادفی‌سازی و رمزنگاری است. در مقابل، بازشناسایی (Re-identification) به فرایندی اطلاق می‌شود که طی آن، داده‌های ناشناس‌شده مجدداً به هویت واقعی افراد مرتبط می‌شوند. این کار عمدتاً از طریق داده‌های کمکی (Auxiliary Data) انجام می‌شود که از منابع مختلفی مانند شبکه‌های اجتماعی، پایگاه‌های داده عمومی و اطلاعات منتشرشده توسط نهادهای دیگر استخراج می‌شوند.

پیشرفت‌های محاسباتی و در دسترس بودن داده‌های کمکی موجب شده است که ناشناس‌سازی، برخلاف تصور اولیه، لزوماً به معنای غیرقابل‌شناسایی شدن داده‌ها نباشد. روش‌های جدیدتر ناشناس‌سازی مثل تعمیم و مبهم‌سازی داده‌ها نیز اگرچه باعث افزایش حریم خصوصی از طریق کاهش احتمال بازشناسایی می‌شوند، اما ممکن است دقت تحلیل‌های داده‌ای را کاهش دهند. به هرحال مسئله بازشناسایی نه‌تنها یک چالش فنی، بلکه یک دغدغه حقوقی مهم است، زیرا مشخص نیست چه سطحی از ناشناس‌سازی می‌تواند داده‌ها را از شمول مقررات حفاظت از داده خارج کند.

این پژوهش، چالش‌های ناشناس‌سازی و امکان بازشناسایی داده‌های شخصی را، به‌ویژه از طریق داده‌های کمکی، بررسی می‌کند. در این راستا، سه پرسش اساسی مطرح می‌شود: ۱. تا چه میزان روش‌های فعلی ناشناس‌سازی قادر به جلوگیری از بازشناسایی داده‌ها هستند؟ ۲. داده‌های کمکی چه نقشی در فرایند بازشناسایی ایفا می‌کنند؟ ۳. چارچوب‌های حقوقی فعلی چگونه باید با چالش‌های ناشناس‌سازی و بازشناسایی مواجه شوند؟ بر این اساس، پژوهش حاضر با روش توصیفی-تحلیلی و ضمن بررسی تطبیقی مقررات مختلف، در دو بخش اصلی (بخش اول در بررسی جزییات ناشناس‌سازی و بخش دوم متضمن چالش‌های آن) تلاش دارد تا به سوالات فوق پاسخ دهد. یافته‌های این پژوهش نشان از این امر دارند که حتی پیشرفته‌ترین روش‌های ناشناس‌سازی نیز در برابر تکنیک‌های بازشناسایی، به‌ویژه با استفاده از داده‌های کمکی، آسیب‌پذیر هستند، چارچوب‌های حقوقی موجود در تعریف دقیق سطح مطلوب ناشناس‌سازی دچار ابهام هستند و این امر اجرای مقررات را دشوار کرده است و در نهایت مقررات‌گذاری مؤثر در این حوزه مستلزم اتخاذ رویکردی پویا و مبتنی بر مدیریت ریسک است که امکان شناسایی تهدیدات نوظهور را فراهم کند.

۲. پیشینه پژوهش

در مورد روش‌های فنی حفاظت از حریم خصوصی اطلاعات در منابع فارسی مطالب بسیار کمی یافت می‌شود. با این حال

مرادی، عباس (۱۴۰۰)، از منظر مطالعات آماری و اهمیت حفظ محرمانگی هویت افراد در انتشار نتیجه آماربرداری‌ها به امر ناشناس‌سازی پرداخته است. البته این مقاله فاقد بحث حقوقی و تنها متضمن مباحث فنی و تکنیکی در طراحی نرم‌افزارها و الگوریتم‌ها برای انتشار صحیح آمار است. در پژوهش دیگری، فاطمه، امیری (۱۳۹۹)، با عنایت به تاثیر منفی ناشناس‌سازی بر سودمندی داده‌ها، مدلی ارائه کرده است که با به‌کارگیری آن می‌توان داده‌ها را به شکل ناشناس، ولی با کمترین میزان اتلاف منتشر نمود که این مقاله نیز از نظر حقوقی به امر ناشناس‌سازی نپرداخته است. فرحزادی و ناصر (۱۴۰۰)، نیز در مقاله‌ای به امر تبادل داده‌های اکتسابی توسط ابزارهای اینترنت اشیاء پرداخته‌اند که با وجود اهمیت تبادل ناشناس داده‌ها، در این نوشتار صحبتی از ناشناس‌سازی به میان نیامده است. در سایر منابع فارسی نیز به همین منوال، مطلب زیادی که به شکل تخصصی مربوط به ناشناس‌سازی باشد، جز در حد ارائه تعریف، یافت نمی‌گردد. لذا بیشتر منابع این پژوهش را منابع خارجی تشکیل می‌دهند که در متن مورد ارجاع قرار گرفته‌اند. بنابر توضیحات فوق، ویژگی این مقاله در این است که به شکل تخصصی به ناشناس‌سازی و چالش‌های اجرایی آن می‌پردازد.

۳. آشنایی با مفاهیم اصلی

حریم خصوصی به معنای دور نگه داشتن زندگی شخصی از دسترسی عمومی بر مبنای حق بر گمنامی است. به‌طور کلی چهار نوع حریم خصوصی شامل حریم اطلاعاتی، بدنی، مکانی و ارتباطی وجود دارد. حریم خصوصی اطلاعاتی مربوط به محدودیت جمع‌آوری، مدیریت، تجزیه و تحلیل و انتشار داده‌های شخصی می‌باشد؛ هرچند داده‌های شخصی ممکن است برای تحقیقات دانشگاهی، وظایف داده‌کاوی دولت، اهداف تجاری و اهداف اجتماعی، منتشر گردند (Ji et al, ۲۰۱۷, p ۱۳۰۷). حال در طول انتشار یا تجزیه و تحلیل داده‌ها، سه دسته از تهدیدات برای افراد متصور است؛ اولا، افشای هویت زمانی اتفاق می‌افتد که یک شخص بتواند به‌درستی دیگری را در مجموعه‌ای از داده‌های منتشر شده تشخیص دهد. ثانیا، افشای ویژگی زمانی رخ می‌دهد که فردی با اطلاعات حساس مربوط به خود مرتبط شود و ثالثا، افشای عضویت که یعنی شخص خارجی بتواند استنباط کند که فرد خاصی در مجموعه داده‌های منتشر شده حاضر است یا غایب. برای مثال تشخیص حضور شخص الف در مجموعه اطلاعات آماری منتشر شده به شکل ناشناس در مورد تعداد مبتلایان به سرطان، می‌تواند تهدیدی جدی برای زندگی شخصی ایشان باشد. تلاش روش‌هایی مثل ناشناس‌سازی که با مفهوم آن آشنا خواهیم شد، جلوگیری از افشای تمام موارد فوق و تهدیدات مذکور می‌باشد (Abdul & Sungchang, ۲۰۲۰, p ۳).

۳-۱. تعریف ناشناس‌سازی

ناشناس‌سازی، یک روش کلیدی مدیریت داده است که مستقیماً با به اشتراک گذاری داده‌ها مرتبط است و از حریم خصوصی افراد محافظت می‌کند. این روش شامل غیرقابل شناسایی کردن افراد با حذف/تغییر اطلاعات شناسایی شخصی (Personally Identifiable Information – PII) در مجموعه داده‌ها می‌باشد (Brasher, ۲۰۱۸, p ۲۰۹). اطلاعات هویتی یا اطلاعاتی هستند که امکان شناسایی مستقیم را فراهم می‌کنند (مثلاً نام یک شخص) یا بخش‌هایی از اطلاعاتی که در ترکیب با یکدیگر ممکن است به شناسایی یک فرد منجر شوند (به‌عنوان مثال، تاریخ تولد به اضافه محل سکونت به اضافه جنسیت) (Stam & Kleiner, ۲۰۲۰, p ۳). به‌عنوان نمونه با هدف تحلیل پایگاه مشتریان و نیازهای بازار توسط شرکت ثالث، در اشتراک گذاری پرونده شخصی به نام علی، ساکن تهران، متولد سال ۱۳۷۹ که از خدمات شرکت خاصی که در زمینه ارائه فضای ابری فعال

است استفاده می‌کند، پس از ناشناس‌سازی تنها نوع و جزییات خدمات ارائه شده بدون اشاره به اسم و اطلاعات هویتی افشا می‌شود. حال در سطوح اولیه تنها به حذف نام و تاریخ تولد اکتفا می‌شود و در سطح بالاتر به‌دلیل احتمال ترکیب اطلاعات و شناسایی افراد، آدرس محل سکونت، آدرس آی پی، شماره تماس و اطلاعات کاربری ثبت شده در سایت نیز پنهان می‌گردند. به هر حال معمولاً درمورد ناشناس‌سازی داده‌ها، غیرقابل بازگشت بودن شرط است؛ لذا سطوح مبتدی و با احتمال منطقی بازشناسایی، ناشناس‌سازی محسوب نمی‌شوند، لکن درمورد نحوه احراز تحقق این شرط اختلافات بسیاری در قوانین و مقررات مختلف وجود دارد که هرکدام معیارهای متفاوتی در این خصوص ارائه می‌نمایند که در بخش‌های آتی بررسی خواهند شد.

اکثر برنامه‌های کاربردی تجزیه و تحلیل داده‌ها و مطالعات آماری نیازی به شناسایی کاربران ندارند و می‌توانند با استفاده از داده‌های ناشناس انجام شوند (مرادی، ۱۴۰۰: ص ۱۴۸). از آنجایی که موضوع تمام قوانین حمایتی داده‌های شخصی هستند، داده‌های ناشناس به‌دلیل عدم پیوند با شخص خاصی، از محدوده حمایت‌های این قوانین خارج می‌شوند. این امر بدان معنا است که داده‌های ناشناس معمولا می‌توانند بر خلاف مقررات مربوط به محدودیت زمانی نگهداری داده‌های شخصی، به‌طور نامحدود ذخیره شوند، برای هر هدفی مورد استفاده قرار گیرند و با هر شخص ثالثی به اشتراک گذاشته شوند. ضمن اینکه تمام این موارد بدون نیاز به رضایت افراد موضوع داده‌ها صورت می‌پذیرد. البته ناشناس‌سازی برای اینکه موفق باشد، باید به تعادلی میان حفظ حریم خصوصی و دقت و کاربرد داده‌ها نائل آید (Gadotti, ۲۰۲۴, p ۱).

در نهایت ناشناس‌سازی هم یک موضوع فنی و هم قانونی است. از یک طرف، دانشمندان علوم کامپیوتر تکنیک‌های ناشناس‌سازی را طراحی می‌کنند و با آزمایش روش‌های احتمالی شناسایی افراد بر روی داده‌های ظاهرا ناشناس، سطح تضمین ارائه شده برای محرمانگی هویت افراد را می‌سنجند و از طرف دیگر قوانین حفاظت از داده و حریم خصوصی تعریف می‌کنند که چه داده‌هایی ناشناس هستند و چه احکامی بر آن‌ها اعمال می‌شوند. چنین تعاریفی معمولاً بسیار کلی هستند و از این رو توسط حقوقدانان، نهادهای نظارتی و در نهایت توسط دادگاه‌ها تفسیر می‌شوند. لذا پرداختن به جزییات امر ناشناس‌سازی هم از نظر فنی و هم از نظر حقوقی حائز اهمیت بسیار می‌باشد.

۳-۲. مزایا و معایب ناشناس‌سازی داده‌ها

برای ناشناس‌سازی از جهات مختلف فنی، حقوقی و اقتصادی می‌توان مزایا و معایب متعددی را در نظر گرفت که آشنایی با آن‌ها می‌تواند به تشخیص سطح و نوع مناسب الزامات قانونی در این خصوص کمک نماید. لازم به ذکر است معایبی که در ادامه بیان می‌گردند، مواردی هستند که اغلب ناشی از نحوه اجرا و یا ذات ناشناس‌سازی بوده و با وجود برخی اشتراکات، با چالش‌های حقوقی ناشناس‌سازی که در بخش دوم ذکر می‌شوند متفاوت می‌باشند.

مزایای ناشناس‌سازی

برای مقابله با موانع قانونی پردازش داده‌ها و توانمند ساختن شرکت‌ها برای دستیابی به اهداف تجاری مشروع خود در مواقعی که اطلاع از هویت اشخاص مورد نیاز نیست ناشناس‌سازی و پردازش داده‌های ناشناس می‌تواند بهترین راه‌حل باشد. در حقیقت رعایت تعهدات حفاظت از داده‌های شخصی بر مبنای قوانین حمایتی تنها در صورتی بر عهده شرکت‌ها است که داده‌ها قابل پیوند به یک شخص حقیقی باشند، اما در مورد داده‌های ناشناس چنین امکانی وجود ندارد (Stummer, ۲۰۲۲, p ۱۸۲). بنابراین برای مثال تعهد نسبت به حذف داده‌ها پس از محقق شدن هدف

جمع آوری و پردازش که در واقع همان محدودیت مدت زمان نگهداری داده‌ها می‌باشد، در مورد داده‌های ناشناس قابل اعمال نیست و داده‌های ناشناس را می‌توان به مدت نامحدود ذخیره نمود (عباس نیا، ۱۴۰۲، ص ۱۹). یکی دیگر از تعهدات مندرج در قوانین حمایتی محدودیت اهداف پردازش و نحوه استفاده از داده‌ها به مواردی است که در حین جمع آوری اعلام شده و افراد نسبت به آن اعطای رضایت نموده‌اند. اما داده‌های ناشناس را می‌توان با اهداف ثانویه‌ای متفاوت از اهداف اولیه پردازش نمود (Boté, Vericad, & Termens, ۲۰۱۹, P ۳۳۳) همچنین درمورد داده‌های شخصی در صورت انتقال آن‌ها به پردازشگر ثالث با رعایت شرایط قانونی، وظیفه نظارت بر پردازش و اعمال پردازشگر و همچنین ارائه دستورالعمل‌های لازم همواره با کنترل‌کننده اصلی خواهد بود؛ در حالی که در مورد داده‌های ناشناس و پس از انتقال آن‌ها کنترل‌کننده هیچ مسئولیتی نسبت به نظارت بر نحوه استفاده از آن‌ها و حفظ امنیت داده‌ها نخواهد داشت. به‌طور کلی نیز داده‌های ناشناس را می‌توان بدون محدودیت با دیگران به اشتراک گذاشت یا به‌طور عمومی منتشر نمود. همینطور در مواردی که به‌دلایل مختلف از جمله به‌دلیل اعتراض یا بازپس‌گیری رضایت از جانب اشخاص موضوع داده، حذف نمودن داده‌ها لازم می‌آید کنترل‌کننده می‌تواند ناشناسازی داده‌ها را جایگزین حذف آن‌ها نماید (Stam & Kleiner, ۲۰۲۰, p ۴).

ناشناس‌سازی از نظر فنی و اقتصادی نیز با مزایای بسیاری همراه است؛ از جمله:

۱. تسهیل همکاری: به‌دلیل رفع موانع قانونی اشتراک‌گذاری داده‌ها از نظر فنی با موانع و محدودیت‌های کمتری همراه خواهد بود و با هزینه کمتر و با ملاحظات فنی متعادل صورت می‌پذیرد و از این جهت همکاری بین شرکت‌ها تسهیل می‌گردد.

۲. تسهیل مطالعات آماری و پروژه‌های تحقیقاتی: ناشناس‌سازی کمک می‌کند تا داده‌ها مجدداً برای اهداف ثانویه حتی بدون رضایت اشخاص موضوع داده مورد استفاده قرار گیرند. همچنین در برنامه‌های اجتماعی توسط نهادهای عمومی و در راستای تصمیم‌گیری و سیاست‌گذاری‌ها نیز استفاده از داده‌ها با موانع کمتری همراه خواهد بود (امیری، ۱۳۹۹، ص ۲۱۱).

۳ . امنیت داده‌های شخصی: ناشناس‌سازی داده‌ها می‌تواند ریسک‌های مرتبط با نقض حریم داده‌ها را با حذف یا پنهان کردن اطلاعات حساس و یا اطلاعات هویتی، کاهش دهد. ضمن اینکه هزینه‌های اجرای اقدامات امنیتی نیز کاهش می‌یابد چراکه نگرانی در مورد نتایج نامطلوب حاصل از افشای داده‌ها توسط مهاجمان به حداقل می‌رسد. همچنین با اجرای ناشناس‌سازی داده‌ها، جذابیت داده‌ها برای هکرها یا سارقان کاهش یافته و به‌طور بالقوه از تلاش برای دسترسی، سرقت یا فروش آن‌ها جلوگیری می‌شود.

۴. بهبود اعتماد و شهرت تجاری نزد کاربران: با ناشناس کردن داده‌ها و شفاف بودن ارتباط با کاربران و اطلاع‌رسانی در مورد چگونگی و چرایی پردازش، شرکت نشان می‌دهد که برای حریم خصوصی ارزش قائل است. این یکی از راه‌های ایجاد اعتماد است که در نهایت به افزایش مزیت رقابتی منجر خواهد شد.

معایب ناشناس‌سازی

ناشناس‌سازی گاه منجر به کاهش دقت و اتلاف داده‌ها می‌شود. استفاده از روش‌های ناشناس‌سازی داده‌ها اغلب به معنای از دست دادن اطلاعات ارزشمند است که می‌تواند به دست آوردن جزییات مفید برای تجزیه و تحلیل و تحقیق را دشوار کند. این امر اثربخشی سیاست‌گذاری و تصمیم‌گیری مبتنی بر داده را محدود می‌کند. همچنین ممکن است به‌دلیل کاهش کیفیت داده‌ها، برخی اشخاص ثالث با انگیزه‌های خاص، از سرمایه‌گذاری یا

مشارکت با کنترل‌کننده در امر پردازش داده‌ها منصرف شوند. در حقیقت سودمندی داده‌ها و میزان حفاظت از حریم خصوصی، دو هدف متعارض می‌باشند و لذا داده‌های مفید، اغلب ناشناس نیستند (Chawla, ۲۰۰۵, p ۵). بنابراین ناشناس‌سازی می‌تواند تحقق اهداف پردازش داده‌ها را با مانع مواجه کند. به‌عنوان مثال در زمینه بازاریابی و تبلیغات، با عدم امکان تعیین شخص موضوع داده، دیگر داده‌ها برای تبلیغات هدفمند بر اساس علایق فرد کاربرد ندارند (Chunchun et al, ۲۰۲۲, p ۳۸۵). همچنین ناشناس‌سازی داده‌ها برای تجزیه و تحلیل داده‌های با حجم بالا، پایگاه‌های داده و کلان داده‌ها مفید است اما در سطح فردی می‌تواند مانعی برای انجام تحقیقات و آزمایشات مربوط به اشخاص معین باشد. ضمن اینکه حتی با ناشناس‌سازی داده‌ها، این خطر وجود دارد که بتوان افراد را با عنایت به رشد فناوری و همچنین در دسترس بودن اطلاعات مرتبط به آن‌ها از طریق منابع دیگر و ترکیب آن با داده‌های ناشناس، دوباره شناسایی کرد. بنابراین، ناشناس‌سازی همیشه به معنای حفظ حریم خصوصی کامل نیست و ابزارهای معکوس کردن ناشناس‌سازی روز به روز قدرتمندتر و در دسترس تر می‌شوند. این مسئله زمانی تبدیل به یک چالش مهم می‌گردد که بدانیم با ناشناس‌سازی، همان‌طور که به‌عنوان یک مزیت برای شرکت‌ها گفته شد، داده‌ها از دامنه حمایتی قوانین مربوط به حفاظت از داده‌های شخصی خارج می‌شوند و لذا در صورت بازشناسایی یا رخداد هر خطر دیگری از طریق داده‌های ناشناس برای حریم خصوصی افراد، قانون حمایتی از ایشان به عمل نخواهد آورد. لذا ناشناس‌سازی از این منظر به مثابه یک شمشیر دو لبه عمل می‌کند.

۴. چالش‌های حقوقی ناشناس‌سازی و روش‌های رفع آن

در میان معایب مذکور، فارغ از چالش‌های فنی و اقتصادی ناشناس‌سازی ناظر به نحوه اجرای تکنیک‌ها و فرایندها و همینطور اتلاف داده‌ها و کاهش کارکرد تجاری آن‌ها، آنچه از منظر حقوقی به‌عنوان چالشی جدی برای ناشناس‌سازی در نظر گرفته می‌شود، ریسک بازگشت‌پذیری ناشناس‌سازی و احتمال افشای هویت افراد و همچنین اختلاف در تعیین معیار ناشناس‌سازی صحیح در قوانین می‌باشد. در این زمینه انتخاب تعریفی صحیح از داده‌های ناشناس و سپس بکارگیری برخی روش‌های نوین ناشناس‌سازی می‌تواند از بسیاری از چالش‌های مطروحه جلوگیری نماید.

۴-۱. بازشناسایی اشخاص موضوع داده‌ها

همان‌طور که بیان شد از مهم‌ترین چالش‌های قانونی در امر ناشناس‌سازی داده‌های شخصی احتمال بازگشت این فرایندها و بازشناسایی اشخاص موضوع داده‌ها است. بازشناسایی به معنای هر نوع عملیات نسبت به داده‌های ناشناس شده برای شناخت مجدد هویت افراد می‌باشد. برای این امر، روش‌های متعددی متصور است که تحولات خاصی در دنیای مدرن امروز با افزایش حجم داده‌های شخصی منتشر شده، این فرایندها را تسهیل می‌کنند.

برای مثال، کلان داده‌ها و پایگاه‌های داده، شامل حجم بزرگ یا پیچیده‌ای از داده‌ها هستند که قرارگیری داده‌های شخصی در قالب این مجموعه‌ها، واجد ریسک می‌باشد. اول اینکه، هر چه مقدار داده بیشتر باشد، احتمال شناسایی مجدد افراد حتی در مجموعه داده‌هایی که به نظر می‌رسد در آن اطلاعات پیوند مشخصی ندارند، بیشتر می‌شود و به‌دلیل حجم و پیچیدگی، حفظ امنیت حریم ایشان نیز دشوارتر می‌گردد؛ دوم، تجزیه و تحلیل کلان داده‌ها می‌تواند از ترکیب داده‌های شخصی کم اهمیت، اطلاعات جدید و حساسی را استنباط کند که بسیار مهم‌تر است و توقع نمیرفته که فاش شوند (Gruschka et al, ۲۰۱۸, p ۵۰۲۷). لذا به‌طور کلی نگهداری داده‌ها در قالب کلان داده‌ها و پایگاه‌های داده، به‌دلیل حجم زیاد و احتمال یافتن پیوندهای معنادار میان داده‌های مختلف مربوط به یک

شخص واحد، به بازشناسایی افراد کمک شایانی می‌کند (VicenA & Navarro-Arribas, ۲۰۱۶, p ۱۹). همچنین افزایش میزان داده‌های کمکی (به این داده‌ها مکمل یا بیرونی و یا دانش پس زمینه نیز گفته می‌شود) به دنبال پیدایش یا رشد برخی فناوری‌ها نیز به ترکیب داده‌های منتشر شده با داده‌های ناشناس و بازشناسایی هرچه راحت‌تر افراد منجر می‌گردد (Abdul & Sungchang, ۲۰۲۱, p ۱۵). پیشرفت هوش مصنوعی و ماهر شدن مدل‌های یادگیری ماشین (Machine Learning) در تجزیه و تحلیل الگوها و الگوریتم‌ها در مجموعه داده‌های ناشناس نیز می‌تواند با نگرانی مهندسی معکوس فرایندهای ناشناس‌سازی و کشف هویت افراد همراه باشد (Richman, ۲۰۲۳). ضمن اینکه فناوری‌هایی مثل اینترنت اشیاء و خانه‌های هوشمند، وسایل پوشیدنی دیجیتال و اتومبیل‌های خودران، امکان جمع‌آوری انواع جدیدی از داده‌های مصرف‌کننده را فراهم می‌کنند (Brasher, ۲۰۱۸, p ۲۱۳). مهم‌تر از تمام این‌ها، با رشد پیام‌رسان‌ها و شبکه‌های اجتماعی، حجم انتشار داوطلبانه اطلاعات نیز بسیار بالا رفته است. روش‌های غیرقانونی مثل هک یا افشای اطلاعات نیز به دنبال لو رفتن داده‌های شرکت‌های بزرگ و به‌ویژه سرویس‌های موبایل و مخابرات، نرم‌افزارهای بانک‌ها و تاکسی‌ها یا فروشگاه‌های اینترنتی، به افزایش فراوانی داده‌های شخصی در منظر عموم کمک می‌کنند. لذا بنابر این مطالب، تقریبا در حوزه اطلاعات و داده‌های شخصی، حریم خصوصی به معنای واقعی وجود نداشته و اطلاعات تمام افراد، حداقل در کشورهایی با خلاء مقرارت یا عدم توجه به حریم افراد، به شکل عمومی در دسترس می‌باشند.

لذا اصولا نمی‌توان نوع و مقدار اطلاعات خارجی را که اشخاص ثالث می‌توانند به آن دسترسی داشته باشد، پیش‌بینی کرد. با وضعیتی که گفته شد، اصلا نباید گمان کرد که مهاجم قادر به یافتن داده‌های مورد نیاز برای باز کردن قفل داده‌های ناشناس نیست. به همین دلیل نگرشی که در گذشته وجود داشت، مبنی بر «امنیت از طریق ابهام» (Security through obscurity) (به این معنا که در دنیای فناوری چون رسیدن به قطعیت دشوار است، صرف ابهام و عدم اطمینان از وجود تهدید، برای فرض وجود امنیت کافی است) امروزه دیگر قابل پذیرش نیست (Dwork, ۲۰۰۶, pp ۱-۲) و لذا باید اقدامات امنیتی را با بدبینی برنامه‌ریزی نمود. حتی این بدبینی و ایجاد احتمالات جدید برای بازشناسایی تا حدی پیش رفته که برخی معتقدند استفاده از اصطلاح داده‌های ناشناس اشتباه است و به جای آن باید از اصطلاح «داده‌های شبه ناشناس» (Pseudo-anonymized data) استفاده نمود (Vokinger, ۲۰۲۰, p ۲۳۱).

حال با آشنایی با مفهوم بازشناسایی و زمینه‌های آن، مطرح نمودن سه مثال بسیار معروف در این زمینه برای آشنایی با ابعاد عملی بحث خالی از فایده نیست. البته با اینکه پرونده‌ها و موارد جدیدتر نیز در دسترس هستند اما این سه مورد به‌طور خاص به شکل‌گیری ادبیات حقوقی بحث ناشناس‌سازی منجر شده‌اند.

به‌عنوان نمونه اول، پرونده دکتر سوییینی (Latanya Sweeney - Massachusetts Governor Re-identification) مهم‌ترین و قدیمی‌ترین نمونه بازشناسایی داده‌های ظاهراً ناشناس شده است. در دهه ۱۹۹۰، دولت ایالت ماساچوست مجموعه‌ای از داده‌های پزشکی بیماران را برای مقاصد پژوهشی منتشر کرد. این داده‌ها شامل اطلاعاتی مانند تاریخ‌های بستری، تشخیص بیماری و روش‌های درمانی بود. اما نام، آدرس و سایر شناسه‌های مستقیم بیماران حذف شده بود. هدف از این انتشار، فراهم کردن بستری برای تحلیل‌های علمی بدون افشای هویت بیماران بود. لاتانیا سوییینی، پژوهشگر علوم کامپیوتر و حریم خصوصی، نشان داد که این داده‌ها همچنان قابل بازشناسایی هستند. او با استفاده از فهرست رأی‌دهندگان عمومی، که شامل نام، تاریخ تولد و کدپستی افراد بود، موفق شد هویت فرماندار وقت ماساچوست، ویلیام ولد را از مجموعه داده‌های پزشکی

استخراج کند. سوینی با تطبیق تاریخ تولد، جنسیت و کدپستی بیماران با داده‌های رأی‌دهندگان، سوابق پزشکی فرماندار را شناسایی کرد و این یافته را برای او ارسال کرد تا آسیب‌پذیری روش‌های سنتی ناشناس‌سازی را نشان دهد. این پرونده تأثیر عمیقی بر سیاست‌گذاری‌های حفاظت از داده‌های شخصی گذاشت و به توسعه استانداردهای سخت‌گیرانه‌تری در حوزه ناشناس‌سازی، منجر شد (Ohm, ۲۰۰۹, p ۱۷۱۹).

پرونده آمریکا آنلاین (AOL Search Data Leak) یکی از دیگر نمونه‌های شاخص در زمینه افشای داده‌های ظاهراً ناشناس‌شده و امکان بازشناسایی کاربران است. در سال ۲۰۰۶، شرکت مذکور به‌عنوان بخشی از یک پروژه تحقیقاتی، مجموعه‌ای از حدود ۲۰ میلیون جستجو را که توسط ۶۵۰ هزار کاربر در بازه سه‌ماهه وارد سایت ایشان شده بودند، منتشر کرد. این داده‌ها شامل تاریخچه جستجوهای کاربران بود اما برای حفظ حریم خصوصی، نام‌های کاربران حذف شده و هر فرد با یک شناسه عددی منحصربه‌فرد جایگزین شده بود. هدف شرکت از انتشار این داده‌ها، کمک به پژوهشگران برای بهبود الگوریتم‌های جستجو بود. بااین‌حال، مدت کوتاهی پس از انتشار، محققان و خبرنگاران دریافتند که حتی بدون وجود نام یا اطلاعات شناسایی مستقیم، برخی کاربران قابل بازشناسایی هستند. الگوهای جستجو شامل اطلاعاتی مانند آدرس، نام اعضای خانواده، سوابق پزشکی و حتی برنامه‌های سفر بودند که امکان تطبیق آن‌ها با داده‌های عمومی را فراهم می‌کرد. در یکی از موارد مشهور، روزنامه‌نگاران نیویورک تایمز موفق شدند یک زن ۶۲ ساله از ایالت جورجیا را تنها بر اساس جستجوهای او در شرکت شناسایی کنند. این افشاگری باعث شد که شرکت مجبور به حذف داده‌ها شود، اما این اطلاعات پیش از آن در اینترنت منتشر شده بود. در نتیجه، این رخداد منجر به کناره‌گیری تعدادی از مدیران شرکت مذکور، از جمله مدیر ارشد فناوری شرکت شد و بحث‌های گسترده‌ای را در مورد حریم خصوصی در داده‌های بزرگ و روش‌های صحیح ناشناس‌سازی برانگیخت (Ohm, ۲۰۰۹, pp ۱۷۱۷-۱۷۲۰).

در آخر پرونده شرکت نتفلیکس (Netflix Prize Data) زمانی آغاز شد که این شرکت در سال ۲۰۰۶ یک مجموعه داده شامل ۱۰۰ میلیون امتیاز کاربران به فیلم‌ها را منتشر کرد تا پژوهشگران بتوانند الگوریتم‌های بهینه‌تری برای پیشنهاد فیلم توسعه دهند. این مجموعه داده شامل اطلاعاتی درباره فیلم‌های تماشا شده، تاریخ امتیازدهی، و امتیازهای ثبت شده و فاقد نام کاربران یا سایر اطلاعات شناسایی مستقیم بود. هدف نتفلیکس از انتشار این داده‌ها، ایجاد رقابتی برای بهبود سیستم پیشنهاد فیلم بود، به این صورت که برنده مسابقه کسی بود که بتواند دقت پیش‌بینی‌های الگوریتم را نسبت به مدل موجود شرکت بهبود ببخشد. بااین‌حال، در سال ۲۰۰۷، پژوهشگرانی از دانشگاه تگزاس نشان دادند که می‌توان با استفاده از داده‌های عمومی سایت دیگری (IMDb)، برخی از کاربران را بازشناسایی کرد. آن‌ها دریافتند که با تطبیق الگوهای امتیازدهی و تاریخ‌های ثبت امتیاز، می‌توان هویت کاربران را کشف کرد. این کشف باعث شد که کاربران بر علیه نتفلیکس اقامه دعوا کنند. در نتیجه، نتفلیکس تصمیم گرفت که انتشار مجموعه داده دوم را لغو کند و پرونده در نهایت با توافق خارج از دادگاه خاتمه یافت (Ohm, ۲۰۰۹, p ۱۷۲۰). در مجموع سه مورد مذکور مهم‌ترین و شناخته‌شده‌ترین موارد ناظر به ناشناس‌سازی و بازشناسایی می‌باشند.

۴-۲. اختلاف در تعیین معیار ناشناس بودن داده‌ها

بنابر توضیحات ارائه شده و احتمال بالای بازشناسایی اشخاص، ناشناس‌سازی در دنیای امروز به دو معنا قابل طرح می‌باشد. در معنای اول با ناشناس‌سازی یا گمنامی مطلق مواجه هستیم، که به موجب آن هیچکس قادر به شناسایی اشخاص موضوع داده نیست که البته

چنین فرضی بسیار ایده‌آل و نادر می‌باشد (Rubinstein & Hartzog, ۲۰۱۵, p ۷۰۴). در معنای دوم که به واقعیت نزدیک‌تر است، ناشناس‌سازی قانونی یک امر نسبی و فرضی است که به‌موجب آن احتمال بازشناسایی به‌طور کامل منتفی نمی‌گردد، بلکه طبق معیارهای قانونی، حالتی است که ضمن آن داده‌های هویتی با تلاش معقول قابل بازگشت نباشند. در حقیقت به‌دلیل ریسک بازشناسایی و عدم امکان نیل به ناشناس‌سازی مطلق و واقعی، در قوانین به‌دلیل لزوم تسهیل فعالیت‌ها و اطمینان‌بخشی نسبی به افراد، انجام اقدامات خاصی پیش‌بینی شده تا با تامین برخی ویژگی‌ها برای ناشناس‌سازی، عدم امکان شناسایی افراد فرض گردد؛ هرچند در عمل خلاف این فرض قابل اثبات باشد. البته در تعیین همین معیارها و معنای عدم امکان بازشناسایی در قوانین مختلف نیز اختلاف است که باعث ایجاد چالش‌های قانونی بسیاری می‌گردد (GFDP, ۲۰۲۲, p ۱۶).

در نظام فقهی اسلامی، حریم خصوصی دارای مبانی مستحکمی است که در قالب مفاهیمی همچون حق مالکیت، اصل برائت، مصونیت از تجسس و تفتیش، ممنوعیت افشای اسرار و استراق سمع، حرمت غیبت، اصل احترام به حریم مؤمن و ضرورت حفظ حرمت آبرو، حیثیت و کرامت افراد تبلور یافته است (حسینی و برزویی، ۱۳۹۶: ص ۱۲۹). از مهم‌ترین قواعد فقهی در این زمینه، قاعده تسلیط است که تصرف مالکانه در اموال را تضمین می‌کند و جنبه سلبی آن، منع تصرف غیرمجاز دیگران را بر حوزه حریم خصوصی نیز تعمیم می‌دهد (اوسط و دیگران، ۱۴۰۲: ص ۵۲۶). این اصول فقهی مبنای حمایتی قدرتمندی برای حریم خصوصی به شمار می‌روند. با توجه به این مبانی، در نظام حقوقی ایران نیز اصولی از قانون اساسی به صیانت از حریم خصوصی اختصاص یافته است. اصل ۲۲ قانون اساسی تضمین می‌کند که حیثیت، جان، مال و حقوق افراد مصون است مگر در موارد مقرر قانونی و اصل ۲۵ ممنوعیت استراق سمع، ضبط و افشای مکالمات و نامه‌ها را بدون حکم قانون پیش‌بینی کرده است. منشور حقوق شهروندی نیز، با وجود فقدان اعتبار قانونی، حق حفاظت از داده‌های شخصی در فضای مجازی را به رسمیت شناخته است. قوانین عادی از جمله قانون تجارت الکترونیکی مصوب ۱۳۸۲ به تعریف داده پیام شخصی، شرایط ذخیره و پردازش داده‌ها با رضایت افراد و حمایت از داده‌های حساس پرداخته‌اند. در مواد ۵۸ و ۵۹ این قانون به لزوم اخذ رضایت برای پردازش داده‌های حساس و اصول کلی پردازش داده‌ها اشاره شده است (فرحزادی و ناصر، ۱۴۰۰: ص ۱۱۸). همچنین قانون انتشار و دسترسی آزاد به اطلاعات، قانون جرایم رایانه‌ای و آیین دادرسی کیفری، مقررات حمایتی در حوزه داده‌ها و جرایم مرتبط را وضع کرده‌اند. قانون احترام به آزادی‌های مشروع و حفظ حقوق شهروندی مصوب ۱۳۸۲، با رویکرد کیفری بر حفظ حریم خصوصی متهمان تأکید می‌کند. علاوه بر این، مصوبات شورای عالی فضای مجازی نیز بر حفاظت از داده‌ها و حریم خصوصی تأکید دارند. حتی در قوانین غیرمرتبط نظیر قانون مالیات‌های مستقیم، قانون ثبت احوال، قانون مرکز آمار و قانون پایانه‌های فروشگاهی سامانه مودیان نیز حمایت‌هایی از داده‌های جمع‌آوری شده و عدم افشای آن دیده می‌شود. سند چشم‌انداز جمهوری اسلامی ایران در افق ۱۴۰۴ نیز به‌صورت کلی بر جایگاه حریم خصوصی، در چارچوب مفاهیمی چون آزادی‌های مشروع، حفظ کرامت انسانی و امنیت اجتماعی، تأکید دارد (اکبری و فلاحیان، ۱۴۰۰: ص ۳۷۲). بنابراین احترام به کلیت حق بر حریم خصوصی که یکی از جنبه‌های آن حق بر حریم خصوصی اطلاعاتی است و یکی از مهم‌ترین حقوق منشعب از آن، حق بر گمنامی می‌باشد، از طریق این موارد قابل اثبات است؛ هرچند متاسفانه قانون خاصی در زمینه حریم خصوصی و حمایت از داده‌های شخصی به تصویب نرسیده است.

به‌دلیل فقدان قانون جامع حمایت از داده‌های

شخصی در ایران، تعاریف و ضوابط دقیق درباره مفاهیمی همچون ناشناس‌سازی و بازشناسایی در قوانین موجود پیش‌بینی نشده و تنها در برخی دستورالعمل‌ها، طرح‌ها و پیش‌نویس‌ها به‌صورت کلی و ضمنی مطرح شده‌اند. لذا در سال‌های اخیر، برای جبران این خلأ، اسناد متعددی تدوین شده است؛ از جمله لایحه حمایت از داده و حریم خصوصی در فضای مجازی (۱۳۹۶)، پیش‌نویس لایحه حمایت از داده‌ها و حریم خصوصی (۱۳۹۷) تهیه‌شده توسط سازمان فناوری اطلاعات، و طرح حمایت و حفاظت از داده‌ها و اطلاعات شخصی که در مجلس در حال بررسی است. این اسناد هرچند هنوز به نتیجه نرسیده‌اند، اما گام‌های مثبتی در جهت حفاظت از داده‌ها محسوب می‌شوند. همچنین شورای عالی انقلاب فرهنگی مقرراتی درباره شبکه‌های اطلاع‌رسانی رایانه‌ای تصویب کرده و پلیس فتا آیین‌نامه دفاتر خدمات حضوری اینترنت را منتشر نموده که حفاظت از داده‌های مراجعان را الزامی کرده است. در سطح اجرایی، «دستورالعمل اجرایی بهبود حفاظت از حریم خصوصی کاربران و شیوه جمع‌آوری، پردازش و نگهداری اطلاعات کاربران» مصوب ۱۴۰۲ کمیسیون عالی تنظیم مقررات فضای مجازی و تأییدشده توسط شورای عالی فضای مجازی، الزامات تدوین سیاست‌های حریم خصوصی داخلی شرکت‌ها را تعیین کرده و «استاندارد ملی امنیت اطلاعات، امنیت سایبری و حفاظت از حریم خصوصی» مصوب ۱۴۰۳ کمیته ملی فناوری اطلاعات نیز برای کلیه سازمان‌ها لازم‌الاجرا شده است. گرچه این اسناد به رمزگذاری و روش‌های کلی حفاظت اشاره دارند، اما مقرره‌ای اختصاصی درباره ناشناس‌سازی ارائه نکرده‌اند (نخجوانی، ۱۴۰۰: ص ۹۶-۱۱۰).

در میان متون فوق، تنها برخی مواد به مفهوم ناشناس‌سازی نزدیک شده‌اند؛ چنان‌که در طرح حفاظت از داده‌های شخصی ۱۴۰۳، ماده ۹ مقرر می‌دارد رضایت به پردازش داده‌ها به معنای اجازه افشای هویت شخص موضوع داده نیست و حق ناشناس ماندن باید حفظ شود و تبصره آن افشای هویت را هر فعل یا ترک فعلی می‌داند که منجر به آگاهی غیرمجاز اشخاص از نام، نام خانوادگی یا شناسه‌های اختصاصی فرد شود. همچنین در پیش‌نویس لایحه حمایت از داده‌ها، ماده ۱۳ استفاده تجاری از داده‌های شخصی را به شرط ناشناس بودن هویت اشخاص مجاز می‌شمارد و ماده ۱۱ به حق گمنامی در پردازش تصریح می‌کند که رضایت به پردازش به معنای آشکاری هویت نیست و شخص حق دارد گمنامی او در محدوده رضایت رعایت شود؛ تبصره این ماده نیز تعریف مشابهی از آشکاری هویت ارائه کرده است. این احکام هرچند رویکردی مثبت دارند، اما فاقد جزییات فنی، معیارهای حقوقی و ضوابط اجرایی لازم برای تضمین ناشناس‌سازی و پیشگیری از بازشناسایی هستند.

در شرایط فعلی که مقررات جامع و الزام‌آور در زمینه حمایت از داده‌های شخصی وجود ندارد و ابعاد حقوقی ناشناس‌سازی تعریف نشده است، با توجه به مبانی فقهی و حقوقی موجود و ضرورت رعایت حریم خصوصی اطلاعات، می‌توان بر مبنای کلیات حقوقی کشور و قوانین مسئولیت مدنی اقدام کرد. ماده ۱ قانون مسئولیت مدنی مقرر می‌دارد: هر کس بدون مجوز قانونی، به‌صورت عمدی یا ناشی از بی‌احتیاطی به جان، مال، حیثیت یا هر حق قانونی دیگر افراد لطمه وارد کند که موجب ضرر مادی یا معنوی شود، مسئول جبران خسارت است. همچنین مواد ۸ و ۹ این قانون، حق جبران خسارت مادی و معنوی ناشی از آسیب به حیثیت و اعتبار شخصی یا خانوادگی را به رسمیت شناخته‌اند. بنابراین در صورت بازشناسایی غیرمجاز و نقض حق گمنامی، خسارت قابل جبران خواهد بود. علاوه بر این، ماده ۲۲۵ قانون مدنی، با توجه به عرف و اراده ضمنی طرفین قرارداد، مبنایی برای استنباط حقوق حریم خصوصی و ناشناس‌سازی در قراردادهای مرتبط با پردازش داده‌ها فراهم می‌آورد (لطیف زاده و دیگران، ۱۴۰۲: ص ۹۹۹).

در نهایت، اگرچه مبانی فقهی و حقوقی حریم خصوصی و حق ناشناس ماندن در فضای مجازی در حقوق ایران پذیرفته شده است، فقدان تعاریف و معیارهای عملی ناشناس‌سازی و بازشناسایی ضرورت تدوین چارچوب‌های حقوقی جامع و دقیق را نمایان می‌سازد. این چارچوب‌ها باید هم تعاریف فنی و حقوقی ناشناس‌سازی را مشخص کنند، هم تضمین‌های کافی برای پیشگیری از بازشناسایی غیرمجاز فراهم آورند و هم با بهره‌گیری از اصول مسئولیت مدنی، حقوق افراد را در برابر نقض حریم محفوظ نگه دارند.

در نظام حقوقی کشور آمریکا، با وجود اینکه مقرره جامعی در سطح فدرال برای حمایت از داده‌های شخصی به‌طور کلی وجود ندارد، اما هم در حوزه‌های مختلف آموزشی، بهداشتی، کیفری، مالی و ... به‌طور جداگانه و در قوانین مختلف از داده‌های شخصی حمایت شده و هم در برخی ایالات قوانین جامعی برای این منظور تصویب شده است. برای شناخت رویکرد آمریکا در مورد ناشناس‌سازی ذکر دو مقدمه ضروری است. اول اینکه در اکثر قوانین حفاظت از داده‌های ایالات متحده، مفهومی به نام «بندرگاه‌های امن»(Safe Harbors) برای داده‌های ناشناس وجود دارد که به شرکت‌ها اجازه می‌دهند در صورت انجام فرایندهای صحیح ناشناس‌سازی از برخی الزامات سخت‌گیرانه قوانین حریم خصوصی معاف شوند (Brasher, ۲۰۱۸, p ۲۲۲). دوم اینکه به‌طور کلی برای توصیف اثرات پردازش داده‌ها بر حقوق اساسی افراد، اصطلاحات «تاثیر»، «ریسک» و «ضرر» رایج هستند. تاثیر در مرتبه اول قرار می‌گیرد که به معنی هر اقدامی است که منجر به ایجاد تغییر در وضعیت حریم خصوصی افراد می‌گردد؛ ریسک به مرحله‌ای گفته می‌شود که تاثیرات در آن به حدی برسند که با توجه به شرایط واقعی، احتمال کافی برای ایراد آسیب به حریم اشخاص وجود داشته باشد و مرحله ضرر نیز در انتها ناظر به موردی است که ریسک احتمالی به واقعیت پیوسته و خسارتی متوجه افراد گردد (Rupp & Max von, ۲۰۲۴, p ۴). معیار قوانین آمریکا برای ورود داده‌ها به بندرگاه‌های امن، مبتنی بر مدیریت ریسک می‌باشد، یعنی اگر شرکتی با اقدامات خود ریسک بازشناسایی را کاهش دهد و به حداقل برساند، داده‌های در اختیار آن ناشناس تلقی و از دایره قانون مستثنی می‌شوند.

حال مهم‌ترین بندرگاه‌های امن برای داده‌های ناشناس در ایالات متحده شامل این موارد هستند: ۱. روش‌های حذف اطلاعات هویتی شخصی مطابق با قانون قابلیت انتقال و پاسخگویی بیمه سلامتی مصوب ۱۹۹۶ که دو روش را برای ناشناس‌سازی داده‌ها مشخص کرده است. اول روش تشخیص متخصص (Expert Determination) که ذیل آن یک متخصص آماری یا ریاضیاتی ارزیابی می‌کند که آیا اطلاعات باقی‌مانده هنوز می‌توانند به یک فرد خاص ارتباط داده شوند یا خیر. متخصص باید به این نتیجه برسد که احتمال بازشناسایی فرد بر اساس اطلاعات باقی‌مانده بسیار کم است. این نتیجه باید بر اساس اصول علمی و روش‌های آماری معتبر باشد. بنابراین رویکرد این قانون یک رویکرد مبتنی بر ارزیابی ریسک است. دوم روش حذف مشخصات هویتی (Safe Harbor Method) که به معنای حذف ۱۸ دسته از اطلاعات قابل‌شناسایی (مانند نام، آدرس، شماره‌های شناسایی) برای کاهش احتمال بازشناسایی می‌باشد. ۲. استثنای داده‌های ناشناس در قانون حریم خصوصی مصرف‌کنندگان ایالت کالیفرنیا مصوب ۲۰۱۸ که داده‌های ناشناس را به‌عنوان اطلاعاتی تعریف می‌کند که به‌طور منطقی نمی‌توان از آن‌ها برای شناسایی فردی خاص استفاده کرد، مشروط بر اینکه شرکت اقداماتی را انجام دهد. این اقدامات شامل تدابیر فنی (حذف یا تغییر داده‌ها به‌گونه‌ای که شناسایی مجدد افراد غیرممکن باشد)، تعهدات قراردادی (ممنوعیت تلاش برای شناسایی مجدد داده‌های ناشناس‌شده توسط اشخاص ثالث) و حفاظت از داده‌ها (حفظ امنیت داده‌ها و جلوگیری از هرگونه دسترسی غیرمجاز) می‌باشند. این

حکم نشان‌دهنده رویکرد مبتنی بر ریسک در قانون مذکور است؛ به‌طوری که اگر ریسک شناسایی مجدد داده‌ها به‌طور معقولی پایین باشد و تدابیر مناسب اتخاذ شود، داده‌ها را می‌توان ناشناس شده تلقی نمود.

کمیسیون تجارت فدرال (Federal Trade Commission - FTC) نیز در سال ۲۰۱۲ با ارائه گزارشی، رهنمودهایی را برای شرکت‌هایی که از داده‌های مصرف‌کنندگان استفاده‌های تجاری می‌کنند منتشر نمود. طبق این گزارش داده‌های ناشناس از برخی الزامات مربوط به حریم خصوصی معاف هستند. در اینجا معیار اینطور بیان شده که اگر داده‌ها به‌طور منطقی قابل پیوند به هویت یک شخص حقیقی باشند ناشناس نیستند. طبق این مقرره تا جایی که یک شرکت اقدامات منطقی را برای اطمینان از عدم شناسایی داده‌ها انجام دهد، متعهد شود که برای شناسایی مجدد داده‌ها تلاش نخواهد کرد و به‌طور قراردادی گیرندگان پایین دستی را از تلاش برای شناسایی مجدد داده‌ها منع کند، از الزامات قانونی معاف است (Brasher, ۲۰۱۸, p ۲۲۵). تمام این موارد منجر به مدیریت ریسک بازشناسایی می‌شود، پس این رویکرد هم مبتنی بر ریسک می‌باشد و در حقوق آمریکا رویکرد غالب همین است.

ظاهراً در اتحادیه اروپا برخلاف آمریکا دو رویکرد اصلی در خصوص معیار ناشناس بودن داده‌ها وجود دارد که گاه به حد تعارض می‌رسند. رویکرد اول مبتنی بر ریسک معقول و سهل‌گیرانه است و رویکرد دوم در ناشناس‌سازی داده‌ها بسیار سختگیرانه عمل می‌کند (Bourdillon & Burt, ۲۰۲۳).

در ابتدا دو مقرره حفاظت از داده‌های اتحادیه اروپا شامل دستورالعمل حفاظت از داده‌های اتحادیه اروپا ۱۹۹۵ (EC - DPD/۴۶/۹۵ Data Protection Directive) و مقرره عمومی حفاظت از داده‌های شخصی سال ۲۰۱۶ (General Data Protection Regulation (GDPR))که جایگزین آن شده، در خصوص معیار ناشناس بودن از رویکرد مبتنی بر ریسک حمایت می‌کنند. این رویکرد دامنه داده‌های ناشناس را بسیار محدود می‌کند چرا که باید تمام احتمالات منطقی برای بازشناسایی داده‌ها در نظر گرفته شوند و درصورتی‌که هیچ ریسکی وجود نداشته باشد داده‌ها را ناشناس تلقی کنیم. البته در مقرره عمومی نسبت به مقرره پیشین دامنه بحث کمی گسترده‌تر و سختگیری‌ها بیشتر شده است. در مقرره سابق اگر داده‌ها بتوانند با وسایلی که به‌طور معقول در دسترس کنترل‌کننده داده یا شخص ثالث است دوباره به فرد مربوطه متصل شوند، آن داده هنوز شخصی محسوب می‌شود. بنابراین مقرره مذکور اجازه می‌داد که ناشناس‌سازی نسبی (با احتمال پایین بازشناسایی) قابل قبول باشد. اما در مقرره عمومی، و در ماده ۴، اگر هرگونه احتمال شناسایی مجدد وجود داشته باشد (حتی در صورت ترکیب داده‌ها با اطلاعات اضافی)، داده همچنان شخصی محسوب می‌شود. لذا محتوای این سند سخت‌گیرانه‌تر است و ناشناس‌سازی طبق آن باید قطعی باشد، نه نسبی. البته یادآوری شماره ۲۶ همین قانون در مورد ناشناس‌سازی اینطور بیان می‌کند: «برای تعیین اینکه آیا یک شخص حقیقی قابل شناسایی است یا خیر، باید تمام ابزارهایی که به‌طور منطقی احتمالاً مورد استفاده قرار می‌گیرند، در نظر گرفته شود، چه توسط کنترل‌کننده یا توسط شخص دیگری برای شناسایی مستقیم یا غیرمستقیم شخص حقیقی. برای اطمینان از اینکه آیا احتمال معقولی برای شناسایی شخص حقیقی وجود دارد یا خیر، باید تمام عوامل عینی، مانند هزینه‌ها و مدت زمان مورد نیاز برای شناسایی، با در نظر گرفتن فناوری موجود در زمان پردازش و پیشرفت‌های فناوری در نظر گرفته شود.» بنابراین یادآوری مذکور، با عدول از سختگیری ماده ۴، کمی به رویکرد مقرره سابق نزدیک‌تر می‌شود.

کارگروه ماده ۲۹ (Working Party - ۲۹ Article WP۲۹) نیز که تا پیش از جایگزینی آن با هیات حفاظت از داده‌های اروپا EDPB) (تحت مقرره عمومی حفاظت

از داده‌ها فعالیت می‌کرد، در سال ۲۰۱۴ نظریه‌ای درباره ناشناس‌سازی داده‌ها منتشر کرده است. این نظریه با عنوان «نظر شماره ۲۰۱۴/۰۵ درباره تکنیک‌های ناشناس‌سازی» به بررسی روش‌های ناشناس‌سازی داده‌های شخصی و میزان کارایی آن‌ها در حفظ حریم خصوصی پرداخت. طبق این نظر ارزیابی ناشناس بودن داده‌ها باید بر اساس سه نوع ریسک مختلف انجام شود:

۱. ریسک افشای ویژگی‌ها (Attribute: Disclosure) اگر برخی ویژگی‌های فرد همچنان در داده باقی بمانند، ممکن است بازشناسایی صورت گیرد.
۲. ریسک افشای ارتباط داده‌ها (Record Linkage): حتی اگر نام حذف شود، امکان پیوند دادن رکوردهای مختلف همچنان می‌تواند شناسایی فرد را ممکن کند.
۳. ریسک استنتاج Inference Attack) (: گاهی حتی بدون نیاز به داده‌های اضافی، با تحلیل داده‌های موجود می‌توان اطلاعات جدیدی درباره فرد به دست آورد.

لذا داده‌ها تنها زمانی ناشناس هستند که هیچ‌یک از این ریسک‌ها وجود نداشته باشد. لذا این نظریه با رویکرد مبتنی بر ریسک موافق است، با این تفاوت که تامین استانداردهای سه گانه مذکور نیاز به احتمال نزدیک به صفر برای بازشناسایی دارد و این یک استاندارد ایده آلیستی و غیرعملی است که نمی‌توان آن را در عصر کلان داده‌ها تضمین نمود. حتی در این نظر از صفت غیرقابل بازگشت نیز برای داده‌های ناشناس استفاده شده است. ضمن اینکه در این سند الزام جداگانه‌ای نیز برای حذف داده‌های خام مربوط به اشخاص که داده‌های ناشناس از تغییر آن‌ها به دست آمده‌اند پیش‌بینی شده است که نشان از رویکردی بسیار سختگیرانه به نسبت مقررات دیگر دارد (Bourdillon et al, ۲۰۱۷, pp ۱۳-۱۶).

همان‌طور که بیان شد، پس از تصویب مقرره عمومی حفاظت از داده‌ها در سال ۲۰۱۶، گروه کاری ماده ۲۹ منحل شد و جای خود را به هیات حفاظت از داده‌های اروپا داد. اما هیات همچنان تا حد زیادی رویکرد سخت‌گیرانه کارگروه را حفظ نمود. طبق رویکرد این نهاد در راهنمای سال ۲۰۲۰، ناشناس‌سازی باید به‌گونه‌ای صورت پذیرد که هیچ راه عملی برای بازشناسایی داده‌ها وجود نداشته باشد. در دستورالعمل‌های جدید هیات نیز بیان شده که روش‌های بازشناسایی داده‌ها همیشه در حال بهبود هستند و شرکت‌ها باید همواره ناشناس‌سازی را تقویت کنند.

آژانس امنیت سایبری اتحادیه اروپا (European Union Agency for Cybersecurity – ENISA) نیز نهادی است که وظیفه تقویت امنیت سایبری در سراسر کشورهای عضو را بر عهده دارد. این آژانس در یک گزارش در سال ۲۰۱۵ در مورد امنیت داده‌های شخصی اینطور بیان نمود که ناشناس‌سازی فرایند تغییر داده‌های شخصی به نحوی است که افراد نتوانند مجدداً شناسایی شوند و هیچ اطلاعاتی را نتوان در مورد آن‌ها به دست آورد. در حالی که ناشناس‌سازی در سایر مقررات تنها ناظر به حذف داده‌های هویتی است که به‌طور مستقیم باعث شناسایی افراد می‌شوند، در این سند تمام داده‌های مربوط به افراد حتی آن‌هایی که به شکل غیر مستقیم به هویت ایشان پیوند می‌خورند نیز باید از مجموعه داده‌ها حذف شوند تا داده‌ها ناشناس تلقی گردند.

در مورد معیار ناشناس بودن در اتحادیه اروپا می‌توان به دو پرونده در این زمینه نیز توجه نمود. پرونده برید (reyer v Germany 2016) یکی از مهم‌ترین پرونده‌های دیوان دادگستری اتحادیه اروپا در مورد تعریف داده‌های شخصی و معیار ناشناس‌سازی است. این پرونده به‌طور خاص به آدرس‌های آی پی و این سؤال پرداخت که آیا این آدرس‌ها داده شخصی محسوب می‌شوند یا خیر؟ آقای پاتریک برید یک فعال حقوق دیجیتال، از دولت آلمان شکایت کرد. دلیل شکایت او این بود که سایت‌های دولتی آلمان آدرس‌های آی پی بازدیدکنندگان را ذخیره می‌کردند، حتی اگر بازدیدکنندگان ثبت‌نام نکرده بودند.

دولت آلمان استدلال کرد که آدرس‌های آی پی به‌تنهایی داده شخصی نیستند، زیرا بدون اطلاعات اضافی (از سوی ارائه‌دهنده خدمات اینترنتی) نمی‌توان آن‌ها را به یک فرد خاص مرتبط کرد. دیوان دادگستری اتحادیه اروپا در تاریخ ۱۹ اکتبر ۲۰۱۶ رأی خود را اینطور صادر کرد: آدرس‌های آی پی داده شخصی محسوب می‌شوند، حتی اگر دولت آلمان به‌تنهایی نتواند فرد را شناسایی کند. معیار شناسایی نباید فقط بر اساس توانایی کنترل‌کننده داده (مثلاً دولت آلمان) باشد، بلکه باید در نظر گرفت که آیا شخص ثالثی می‌تواند داده را بازشناسایی کند یا نه. اگر اطلاعات اضافی کمک کننده‌ای در دست ثالث وجود داشته باشد، داده‌ها همچنان شخصی هستند. این یک معیار سخت‌گیرانه‌تر برای ناشناس‌سازی است، زیرا نشان می‌دهد که برای ناشناس‌سازی کامل، نباید هیچ شخص ثالثی بتواند داده را بازشناسایی کند (Groos & Veen, ۲۰۲۰, p ۵۰۱).

در پرونده دوم (The SRB v EDPS case 2022) راجع به هیات واحد تصمیم‌گیری علیه ناظر حفاظت از داده‌های اروپا (هیات یک نهاد اروپایی است که مسئول اجرای مکانیزم حل و فصل اختلافات بانکی در اتحادیه اروپا است)، هیات مذکور در چارچوب وظایف خود اطلاعات مربوط به کارمندان یک بانک را در قالبی پردازش کرد که شامل حقوق و شرایط کاری آن‌ها بود، اما بدون درج نام آن‌ها. ناظر حفاظت از داده‌های اروپا، که مسئول نظارت بر رعایت مقرره عمومی حفاظت از داده‌ها در نهادهای اتحادیه اروپا است، اعلام کرد که این داده‌ها هنوز هم داده‌های شخصی محسوب می‌شوند و باید تحت قوانین حفاظت از داده‌ها پردازش شوند. هیات با این تصمیم مخالف بود و استدلال کرد که اطلاعات مذکور ناشناس هستند، زیرا نام و سایر مشخصات شناسایی از آن‌ها حذف شده است. این پرونده به این سؤال کلیدی پرداخت که آیا داده‌های فاقد اطلاعات شناسایی مستقیم (مانند نام)، اما همچنان شامل اطلاعات حقوقی و شغلی، می‌توانند ناشناس محسوب شوند؟ در این پرونده، داده‌های مربوط به حقوق و شرایط کاری کارمندان با توجه به تعداد محدود افراد در هر موقعیت شغلی در بانک، می‌توانستند به‌طور غیرمستقیم منجر به شناسایی افراد شوند. دادگاه اتحادیه اروپا حکم داد که معیار تعیین ناشناس‌سازی باید سخت‌گیرانه باشد و هرگونه احتمال بازشناسایی، حتی با اطلاعات اضافی، نشان می‌دهد که داده همچنان شخصی است. حذف نام به‌تنهایی برای ناشناس‌سازی کافی نیست؛ بلکه باید بررسی کرد که آیا داده‌ها می‌توانند به‌طور غیرمستقیم بازشناسایی شوند یا خیر. اگر حتی یک شخص ثالث بتواند با اطلاعات اضافی (که ممکن است عمومی نباشد، اما در دسترس باشد) داده‌ها را بازشناسایی کند، داده‌ها ناشناس محسوب نمی‌شوند. پس باید ریسک بازشناسایی به‌عنوان معیار اصلی تعیین داده‌های ناشناس در نظر گرفته شود.

در این خصوص توجه به نکته‌ای در مورد جبران خسارات ضروری می‌باشد. درست است که داده‌ها پس از ناشناس‌سازی از حیطه حمایت قوانین خارج می‌گردند، اما اگر اثبات شود که عمل ناشناس‌سازی با تقصیر یا عدم رعایت مقررات و معیارهای مذکور به شکل نادرستی انجام شده و به این دلیل کنترل‌کننده در باشناسایی هویت افراد موثر است، در این صورت مسئولیت هرگونه ضرر وارده به اشخاص موضوع داده‌ها (مادی یا معنوی) بر عهده ایشان خواهد بود. در همین راستا ماده ۸۲ مقرره عمومی حفاظت از داده‌های اتحادیه اروپا، هر شخصی که در نتیجه عدم رعایت این قانون متحمل هر نوع خسارتی شده باشد را مستحق دریافت غرامت از کنترل‌کننده یا پردازشگر دانسته و حتی حقوق ایران نیز با وجود عدم وضع قوانین حمایتی در مورد داده‌های شخصی، بر مبنای کلیات حقوق مسئولیت مدنی و بر اساس تقصیر، چنین خساراتی را قابل مطالبه می‌داند.

۴-۳. روش‌های ناشناس‌سازی داده‌ها

از نظر فنی چندین تکنیک و روش مختلف برای ناشناس کردن داده‌های شخصی وجود دارد که با افزایش سطح حفاظت از حریم خصوصی افراد و کاهش احتمال بازشناسایی ایشان همراه خواهد بود:

۱. پوشش داده‌ها (Data masking)

پوشش داده‌ها، شاید شناخته‌شده‌ترین روش ناشناس‌سازی داده‌ها باشد. این فرایند به معنای پنهان کردن یا حذف مقادیر در یک مجموعه داده است. برای مثال در انتشار یک مجموعه داده، با نمایش چهار رقم اول شماره تماس، می‌توان به جای باقی اعداد از ستاره استفاده کرد. این روش در عمل مشابه با روش حذف (Suppression) است. هرچند می‌توان گفت حذف کلی‌تر بوده و در معنای دقیق شامل حذف کامل اطلاعات حساس از مجموعه داده‌ها می‌شود. برای مثال در یک پایگاه داده پزشکی، برای حفظ حریم خصوصی بیماران، نام و شماره شناسایی آن‌ها حذف می‌گردد (Richman, ۲۰۲۴).

۲. آشفتگی داده‌ها (Data perturbation)

در این روش کنترل‌کننده جزییات داده‌ها را به شکل تصادفی تغییر می‌دهد تا ابهام را به یک مجموعه داده به روشی قابل پیش‌بینی و قابل بازیابی اضافه کند، بدون اینکه بر دقت تجزیه و تحلیل تأثیر بگذارد. این روش را می‌توان با وارد کردن نویز (Noise Addition) (تغییر کنترل شده مقادیر) به مقادیر عددی حساس، یا با تغییر تصادفی متغیرها انجام داد. در این تکنیک، مقادیر تصادفی کوچک به داده‌های عددی اضافه یا از آن‌ها کم می‌شود تا داده‌های واقعی مخفی شوند و شناسایی افراد سخت‌تر شود. برای مثال در یک پایگاه داده حقوق و دستمزد، مبلغ حقوق هر فرد با افزودن یا کاستن درصد کمی تغییر داده می‌شود.

۳. جابه‌جایی داده‌ها (Data swapping)

این روش متضمن تنظیم مجدد داده‌ها در یک مجموعه داده به گونه‌ای است که مقادیر مشخصه، دیگر با داده‌های اصلی مطابقت نداشته باشند. این تکنیک ناشناس‌سازی داده‌ها که به آن درهم ریختگی داده یا جایگشت داده نیز گفته می‌شود، ممکن است یک ویژگی را از ردیف ۱ گرفته و آن را با یک ویژگی از ردیف ۷۸ همان ستون تعویض کند (Devane, ۲۰۲۴).

۴. تعمیم (Generalization)

در این تکنیک، جزییات داده‌ها با مقادیر کلی‌تر جایگزین می‌شوند تا شناسایی دقیق فرد دشوار گردد. برای مثال به جای ذکر آدرس دقیق یک فرد (مانند خیابان آزادی، پلاک ۱۲۳)، فقط نام شهر (تهران) ذکر می‌شود یا به جای ذکر کدپستی، تنها دو شماره اول در جدول نمایش داده می‌شود و یا به جای تاریخ تولد دقیق، تنها سال تولد درج می‌گردد (Gruschka et al, ۲۰۱۸, p ۵۰۲۹).

۵. داده‌های مصنوعی (Synthetic data)

داده‌های مصنوعی، داده‌هایی هستند که به‌صورت الگوریتمی و بدون استفاده مستقیم از داده‌های واقعی تولید می‌شوند. این داده‌ها معمولاً برای آموزش مدل‌های یادگیری ماشین، تست نرم‌افزارها، یا انجام تحقیقات در شرایطی که دسترسی به داده‌های واقعی محدود یا ممنوع است، استفاده می‌شوند (Usercentrics, ۲۰۲۴). در یک تکنیک خاص ناشناس‌سازی، می‌توان به جای انتشار داده‌های اصلی، داده‌های مصنوعی تولید شده با همان ترتیب‌ها و با همان مقیاس‌ها را منتشر نمود.

۶. تجمیع (Aggregation)

این روش شامل گروه‌بندی داده‌های فردی به دسته‌های کلی‌تر است. برای مثال به‌جای نمایش نمرات هر دانش‌آموز،

میانگین نمرات کل کلاس ارائه می‌شود (Brasher, ۲۰۱۸, pp ۲۱۶-۲۲۰). در یک فرایند ناشناس‌سازی صحیح و قانونی، یکی یا مجموعه‌ای از این تکنیک‌ها با توجه به نوع داده و میزان حساسیت آن انتخاب می‌شوند.

۵. نتیجه‌گیری و توصیه‌های سیاستی

برای سال‌ها، این باور رواج داشت که ناشناس‌سازی داده‌ها به‌طور کامل می‌تواند خطر شناسایی مجدد را از بین ببرد و بنابراین، بسیاری از سیاست‌های انتشار داده بر این فرض استوار شدند. اما پیشرفت‌های فنی در پردازش کلان‌داده‌ها، گسترش منابع اطلاعاتی و توسعه الگوریتم‌های بازشناسایی نشان داده که این فرض دیگر معتبر نیست. امروزه حتی داده‌های ظاهراً ناشناس، در بسیاری از موارد، با ترکیب با سایر پایگاه‌ها یا تحلیل پیشرفته، قابلیت بازشناسایی دارند. این واقعیت به‌ویژه در کشورهایی که فاقد مقررات جامع و ضمانت اجراهای مؤثر هستند، مخاطرات جدی برای حریم خصوصی و آزادی‌های فردی ایجاد می‌کند.

سیاست‌های سنتی در سطح بین‌المللی، یا به انتشار آزادانه داده‌ها پس از ناشناس‌سازی بدون کنترل پسینی (انتشار و فراموشی) (Release-and-forget anonymization) متکی بوده‌اند، یا با رویکردی سخت‌گیرانه، انتشار داده‌های پرریسک را به‌طور کامل مسدود کرده‌اند. هر دو روش در عمل ناکارآمد بوده‌اند؛ نخستین رویکرد به‌دلیل نادیده گرفتن احتمال بازشناسایی، و دومین رویکرد به‌دلیل ایجاد مانع در برابر استفاده مشروع و مفید از داده‌ها.

بر این اساس، پیشنهاد می‌شود سیاست انتشار داده‌ها بر مدیریت ریسک بازشناسایی استوار گردد. چارچوب پیشنهادی عبارت است از:

۱. تعریف قانونی و فنی سطوح ناشناس‌سازی و شبه‌ناشناس‌سازی، همراه با شاخص‌های قابل سنجش برای هر سطح (مانند میزان تغییر داده، سطح تجمیع، و احتمال بازشناسایی در بازه زمانی مشخص).

۲. الزام به ارزیابی ریسک پیش از انتشار و همچنین الزامی شدن ارائه گزارش ارزیابی ریسک برای همه نهادهای دولتی و خصوصی پیش از انتشار داده، شامل: ماهیت داده، حساسیت، گروه هدف، دریافت‌کننده و سطح مهارت‌های فنی ایشان، روش‌های ناشناس‌سازی، احتمال ترکیب با منابع خارجی، شرایط دسترسی و نتیجه آزمون بازشناسایی.

۳. تعیین سطوح دسترسی مبتنی بر ریسک، به نحوی که داده‌های پرریسک فقط در بسترهای کنترل‌شده و با محدودیت تحلیل منتشر شوند.

۴. تفکیک کانال‌های انتشار به سه دسته:

انتشار عمومی (فقط داده‌های کم‌ریسک)، انتشار محدود (در بسترهای امن با کنترل دسترسی) و انتشار محرمانه (فقط برای پژوهشگران مجاز در محیط‌های امن)

۵. استفاده از ترکیب اقدامات فنی و حقوقی، از جمله مستعارسازی، تجمیع داده‌ها، تعهدات قراردادی عدم بازشناسایی، و پایش دوره‌ای و الزام به بازبینی سطح ریسک داده‌ها حداقل هر ۱۲ ماه پس از انتشار.

۶. ایجاد نظام پاسخ‌گویی و نظارت مستقل برای پایش فرایند انتشار داده، دریافت گزارش تخلفات و اعمال ضمانت اجراهای بازدارنده.

در حقوق ایران، هرچند قانون خاصی در زمینه ناشناس‌سازی و بازشناسایی داده‌ها وجود ندارد، اما مقرراتی چون قانون تجارت الکترونیکی، قانون جرایم رایانه‌ای و قانون انتشار و دسترسی آزاد به اطلاعات، ظرفیت‌هایی برای ساماندهی این موضوع دارند. همچنین، طرح حفاظت از داده‌های شخصی ۱۴۰۳، گرچه هنوز به تصویب نرسیده، اما پایه‌هایی برای تعریف مفاهیم کلیدی و تعیین مسئولیت‌ها ارائه کرده است. با این حال، نبود معیارهای حقوقی برای سنجش سطح ناشناس‌بودن، فقدان الزام قانونی به ارزیابی ریسک و نبود ضمانت اجراهای خاص برای بازشناسایی، از کاستی‌های مهم به شمار می‌آیند.

برای رفع این خلأها، علاوه بر چارچوب فوق، پیشنهاد می‌شود فراتر از اقدامات انجام شده در قالب ارائه طرح‌ها و پیش‌نویس‌های متعدد بدون سرانجام، مقرره‌ای خاص همچون مقررات عمومی حمایت از داده‌های اتحادیه اروپا برای حمایت کلی از داده‌ها، متضمن ماده‌ای مشابه با ماده ۲۶ مقرره اتحادیه اروپا درباره ناشناس‌سازی، در نظام حقوقی ایران پیش‌بینی شود. متن پیشنهادی به شرح زیر است:

ماده پیشنهادی - ناشناس‌سازی و بازشناسایی داده‌های شخصی

۱. تعاریف:

الف) «ناشناس‌سازی» عبارت است از پردازش داده‌های شخصی به نحوی که شناسایی مستقیم یا غیرمستقیم شخص موضوع داده، با استفاده از هر ابزار معقول و متعارف، غیرممکن گردد.

ب) «شبه‌ناشناس‌سازی» عبارت است از پردازش داده‌ها به گونه‌ای که شناسایی شخص، تنها با استفاده از اطلاعات تکمیلی که به‌طور جداگانه و تحت تدابیر امنیتی خاص نگهداری می‌شود، ممکن باشد.

۲. الزامات:

الف) هرگونه انتشار یا اشتراک‌گذاری داده‌های ناشناس یا شبه‌ناشناس باید مسبوق به ارزیابی ریسک بازشناسایی باشد.

ب) ناشران داده مکلف هستند متناسب با سطح ریسک، اقدامات فنی و سازمانی لازم را برای کاهش احتمال بازشناسایی اتخاذ کنند.

ج) انتشار داده‌های پرریسک صرفاً در بسترهای کنترل‌شده و با سطح دسترسی محدود مجاز است.

۳. مسئولیت:

الف) بازشناسایی عمدی داده‌های ناشناس یا شبه‌ناشناس بدون مجوز قانونی، ممنوع و مشمول مسئولیت مدنی و کیفری است.

ب) نگهداری، افشا یا استفاده از داده‌های بازشناسایی‌شده برخلاف مقررات، مشمول جریمه‌های مقرر در این قانون خواهد بود.

۴. نظارت:

الف) نهاد ناظر حفاظت از داده‌های شخصی (که ساختار، وظایف و اختیارات آن در همین قانون مشخص شده) مکلف است بر فرایند ناشناس‌سازی و انتشار داده‌ها نظارت کند و گزارش سالانه‌ای از وضعیت رعایت الزامات این ماده منتشر نماید.

ب) نهاد ناظر مجاز است در صورت احراز ریسک غیرقابل‌قبول بازشناسایی، دستور توقف انتشار یا حذف داده‌ها را صادر کند.

تبصره ۱: معیارهای سنجش سطح ناشناس‌سازی و ارزیابی ریسک بازشناسایی، ظرف شش ماه از تصویب این قانون، توسط نهاد ناظر و با مشورت متخصصان فنی و حقوقی تدوین و ابلاغ می‌گردد.

تبصره ۲: در صورت بروز اختلاف در تشخیص سطح ناشناس‌سازی یا میزان ریسک بازشناسایی، نظر نهاد ناظر ملاک عمل خواهد بود.

به‌منظور عملیاتی‌سازی چارچوب پیشنهادی، می‌توان یک نقشه راه سه‌مرحله‌ای طراحی نمود. در فاز کوتاه‌مدت (۶ تا ۱۲ ماه)، لازم است دستورالعمل‌های موقت انتشار داده‌ها توسط نهادهای ذی‌ربط همچون مرکز ملی فضای مجازی یا وزارت ارتباطات تدوین و ابلاغ گردد، کلیه سازمان‌ها و شرکت‌های خصوصی ملزم به انجام و ثبت ارزیابی ریسک پیش از انتشار شوند و فهرستی از داده‌های حساس که انتشار آن‌ها بدون ارزیابی دقیق ممنوع است، ایجاد شود.

در فاز میان‌مدت (۱ تا ۳ سال)، تشکیل «مرکز ارزیابی ریسک داده» به‌عنوان نهاد ملی مستقل، ایجاد دوره‌های آموزشی و صدور گواهی‌نامه تخصصی برای مسئولان انتشار داده یا نمایندگان شرکت‌ها و راه‌اندازی سامانه ملی انتشار داده به‌صورت تعاملی که امکان تحلیل را بدون دسترسی به داده خام فراهم کند، ضروری خواهد بود. در فاز بلندمدت (۳ تا ۵ سال)، تصویب «قانون جامع حمایت از داده‌های شخصی» با فصل ویژه ناشناس‌سازی و بازشناسایی، ایجاد نظام نظارت قضایی و اعمال جریمه‌های بازدارنده برای بازشناسایی غیرمجاز و همچنین پیوستن به شبکه‌های بین‌المللی تبادل داده امن به‌منظور بهره‌برداری علمی و اقتصادی از داده‌ها، از جمله اقدامات اساسی است.

دسترسی به داده‌ها

داده‌های استفاده شده یا تولید شده در متن مقاله ارائه شده است.

تضاد منافع نویسندگان

نویسندگان این مقاله اعلام می‌دارند که هیچ گونه تضاد منافعی در رابطه با نویسندگی و یا انتشار این مقاله ندارند.

منابع

اکبری، علی و فلاحیان، مهدی، (۱۴۰۰)، حریم خصوصی در نظام حقوقی ایران و اسلام، *فصلنامه تمدن حقوقی*، ۴(۹)، صص ۳۶۱-۳۸۲. doi: ۲۰۲۲,۱۳۷۵۰۹.lc/۱۰,۲۲۰۳۴.

اوسط، امیرعلی و امام، سید محمد رضا و وزیری، مجید، (۱۴۰۲)، تحلیل و بررسی مفهوم حریم خصوصی و تاثیر آن بر سلامت اجتماعی از منظر فقه امامیه، *سبک زندگی اسلامی با محوریت سلامت*، ۷(۲)، صص ۵۴۵-۵۵۲.

امیری، فاطمه، (۱۳۹۹)، حفظ حریم خصوصی در برون‌سپاری داده‌های سامانه‌های اطلاعاتی با تکیه بر سودمندی داده، *پژوهشنامه پردازش و مدیریت اطلاعات*، ۳۶(۱)، صص ۲۱۱-۲۴۲، doi: ۱۰,۳۵۰۵۰/ JIPM۰۱۰,۲۰۲۰,۰۱۹.

حسینی، مهدی و برزویی، محمدرضا، (۱۳۹۶)، مبانی و مؤلفه‌های فقهی حمایت از حریم خصوصی افراد در فضای مجازی، *مطالعات حقوق بشر اسلامی*، ۶(۱۳)، صص ۱۱۵-۱۳۷.

عباس نیا، حامد، (۱۴۰۲)، حق پاک کردن داده‌های شخصی از فضای مجازی، پایان نامه کارشناسی ارشد، دانشکده حقوق و علوم سیاسی دانشگاه تهران.

فرحزادی، علی اکبر و ناصر، مهدی، (۱۴۰۰)، حق بر تبادل داده‌های خصوصی و راه‌کارهای رفع چالش‌های آن در سازوکار عملکرد ابزارهای اینترنت اشیا، *بررسی‌های بازرگانی*، ۱۹(۱۰۹)، صص ۱۱۵-۱۲۹، doi: ۲۰۲۱,۲۴۷۰۵۰.bs/۱۰,۲۲۰۳۴.

لطیف زاده، مهدیه و قبولی درافشان، سید محمد مهدی و محسنی، سعید و عابدی، محمد، (۱۴۰۲)، حمایت از دادهٔ شخصی در حقوق اتحادیهٔ اروپا و امکان‌سنجی آن در نظام حقوقی ایران، *فصلنامه مطالعات حقوق عمومی دانشگاه تهران*، ۵۳(۲)، صص ۹۸۱-۱۰۰۵. doi: ۱۰,۲۲۰۵۹/ ۲۰۲۱,۳۲۴۶۹۴,۲۷۸۶.jplsq.

مرادی، عباس، (۱۴۰۰)، تعادل بین دسترسی کاربران به ریزداده‌ها و حفظ حریم خصوصی واحدهای آماری در آمار رسمی، بررسی‌های آمار رسمی ایران، ۳۲(۱)، صص ۱۲۵ -۱۵۰.

ناصر، مهدی و تاج، سید امیر علی، (۱۳۹۹)، ارائه مدلی فناورانه برای حل چالش‌های حاصل از عدم تقارن اطلاعات در بیمه مالکیت صنعتی، *بررسی‌های بازرگانی*، ۱۸(۱۰۱)، صص ۴۵-۶۲.

نخجوانی، نرگس، (۱۴۰۰)، حقوق حمایت از داده‌های شخصی، چاپ اول، کتاب طه، قم.

هدایت نظری، فائزه و محمدی، اکبر، (۱۴۰۳)، موردپژوهی توسعه تجارت الکترونیک کشورهای پیشرو در عصر اقتصاد دیجیتال، *بررسی‌های بازرگانی*، ۲۲(۱۲۸)، صص ۲۷-۵۰، doi: ۱۰,۲۲۰۳۴/ ۲۰۲۴,۲۰۲۷۴۲۷,۲۹۵۹.bs.

Abbasnia, Hamed, (1402), The Right to Delete Personal Data from Cyberspace, Master's Thesis, Faculty of Law and Political Science, University of Tehran, [In Persian].

Abdul, Majeed & Sungchang, Lee, (2020), Anonymization Techniques for Privacy Preserving Data Publishing: A Comprehensive Survey, IEEE Access, 9, pp 8512-8545, doi: 10.1109/ ACCESS.2020.3045700.

Akbari, Ali & Fallahian, Mehdi, (1400), Privacy in the Iranian and Islamic Legal System, Journal of Legal Civilization, 4(9), pp 361-382. doi: 10.22034/ lc.2022.137509, [In Persian].

overcome its challenges in the functioning mechanism of Internet of Things devices, Commercial Survays, 19(109), pp. 115-129, doi: 10.22034/bs.2021.247050, [In Persian].

Gadotti, Andrea et al, (2024), Anonymization: The imperfect science of using data while preserving privacy, Sci. Adv, 10(29), pp 1-22, doi: 10.1126/sciadv.adn7053.

Germany Foundation for Data Protection, (2022), Practice Guide to Anonymising Personal Data, pp 1-59.

Groos, D. & Veen, E, (2020), Anonymised Data and the Rule of Law, European Data Protection Law Review, 6, pp 498-508, Doi: 10.21552/edpl/2020/4/6.

Gruschka, Nils & Mavroeidis, Vasileios & Vishi, Kamer & Jensen, Meiko, (2018), Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR, IEEE International Conference on Big Data pp 5027-5033, Doi: 10.1109/BigData.2018.8622621.

Hedayat Nazari, Faezeh and Mohammadi, Akbar, (1403), Case study of e-commerce development in leading countries in the era of digital economy, Commercial Survays, 22(128), pp. 27-50, doi: 10.22034/bs.2024.2027427.2959, [In Persian].

Hosseini, Mehdi & Borzoei, Mohammad Reza, (2017), Jurisprudential Principles and Components of Protecting Individuals' Privacy in Cyberspace, Islamic Human Rights Studies, 6(13), pp 115-137, [In Persian].

Ji, S & Mittal, P & Beyah, R, (2017), Graph Data Anonymization, De-Anonymization Attacks, and De-Anonymizability Quantification: A Survey, IEEE Communications Surveys & Tutorials, 19(2), pp 1305-1326, doi: 10.1109/COMST.2016.2633620.

Latifzadeh, Mahdieh & Qabouli Dorafshan, Seyed Mohammad Mahdi & Mohseni, Saeed & Abedi, Mohammad, (1402), Protection of Personal Data in European Union Law and Its Feasibility in the Iranian Legal System,Journal of Public Law Studies, University of Tehran, 53(2), pp 981-1005. doi: 10.22059/jplsq.2021.324694.2786, [In Persian].

Moradi, Abbas, (1400), Balance between users' access to microdata and maintaining the privacy of statistical units in official statistics, Iranian Official Statistics Reviews, 32(1), pp. 125-150, [In Persian].

Nasser, Mehdi and Taj, Seyed Amir Ali, (1399), Presenting a Technological Model to Solve the

Amiri, Fatemeh, (1399), Privacy protection in outsourcing information systems data based on data usefulness, Journal of Information Processing and Management, 36(1), pp. 211-242, doi: 10.35050/JIPM010.2020.019, [In Persian].

Bourdillon, Stalla & Alison, Sophie and Knight, (2017), Anonymous Data v. Personal Data, A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data, Wisconsin International Law Journal, 34(2), pp 1-38.

Bourdillon, Stalla & Burt, Andrew, (2023), The definition of 'anonymization' is changing in the EU: Here's what that means, Available at: (https://iapp.org/news/a/the-definition-of-anonymization-is-changing-in-the-eu-heres-what-that-means) Visited 2025/01/29.

Boté Vericad, Juan José & Termens, Miquel, (2019), Reusing Data: Technical and Ethical Challenges, DESIDOC Journal of Library & Information Technology. 39(6), pp 329-337, doi: 10.14429/djlit.39.6.14807.

Brasher, E, (2018), Addressing the Failure of Anonymization: Guidance from the European Union's General Data Protection Regulation, Columbia Business Law Review, (1), pp 209-253, Doi: 10.7916/cblr.v2018i1.1217.

Chawla, Shuchi et al, (2005), Toward Privacy in Public Databases, In: Kilian, J. (eds) Theory of Cryptography. TCC 2005. Lecture Notes in Computer Science, 3378. Springer, Berlin, Heidelberg, Doi: 10.1007/978-3-540-30576-7_20.

Chunchun, Ni & Li, Shan Cang & Prosanta, Gope & Geyong, Min, (2022), Data anonymization evaluation for big data and IoT environment, Information Sciences, 605, pp 381-392, doi: 10.1016/j.ins.2022.05.040.

Devane, Heather, (2024), What Are the Top Data Anonymization Techniques? Available at: (https://www.immuta.com/blog/data-anonymization-techniques/) Visited 2025/02/13.

Dwork, C, (2006), Differential Privacy, In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds) Automata, Languages and Programming. ICALP 2006, Lecture Notes in Computer Science, 4052, Springer, Berlin, Heidelberg, doi: o/10.1007/11787006_1.

Farahzadi, Ali Akbar and Nasser, Mehdi, (1400), The right to exchange private data and solutions to

Stam, A & Kleiner, B, (2020), Data anonymization: legal, ethical, and strategic considerations, FORS Guide, 11(1.0.), pp1-15, doi:10.24449/FG-2020-00011.

Stummer, Sarah, (2022), Issues of Verifying Anonymity: An Overview, Gesellschaft für Informatik (Jahrestagung) Conference, pp 179-194, doi: 10.18420/inf2022_17.

Usercentrics, (2024), Everything you need to know about data anonymization, Available at: (https://usercentrics.com/knowledge-hub/data-anonymization/) Visited 2025/01/20.

VicenÃ, Torra, & Guillermo, Navarro-Arribas, (2016), Big Data Privacy and Anonymization, pp 15-26, DOI: 10.1007/978-3-319-55783-0_2.

Vokinger, K & Stekhoven, D & Krauthammer, M, (2020), Lost in Anonymization - A Data Anonymization Reference Classification Merging Legal and Technical Considerations, J Law Med Ethics, 48(1), pp 228-231, Doi: 10.1177/1073110520917025.

Challenges Resulting from Information Asymmetry in Industrial Property Insurance, Commercial Survays, 18(101), pp. 45-62, [In Persian].

Ohm, Paul, (2009), Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, UCLA Law Review, 57, pp 1701-2010.

Owsat, Amir Ali & Imam, Seyyed Mohammad Reza & Waziri, Majid, (1402), Analysis and investigation of the concept of privacy and its impact on social health from the perspective of Imamiyyah jurisprudence, Islamic Lifestyle with a Health Focus, 7(2), pp 545-552, [In Persian].

Richman, Amitai, (2024), Data Anonymization vs Data Masking: Definitions and Use Cases, Available at (https://www.k2view.com/blog/data-anonymization-vs-data-masking/#Data-anonymization-defined) Visited 2025/01/19.

Richman, Amitai, (2023), Re-Identification of Anonymized Data: What You Need to Know, Available at: (https://www.k2view.com/blog/re-identification-of-anonymized-data) Visited 2025/02/15.

Rubinstein, Ira & Hartzog, Woodrow, (2015), Anonymization and Risk, Public Law Research Paper, 15(36), pp 1-59.

Rupp, Valentin & Grafenstein, Max von, (2024), Clarifying personal data and the role of anonymisation in data protection law: Including and excluding data from the scope of the GDPR (more clearly) through refining the concept of data protection, Computer Law & Security Review, 52, pp 1-25, Doi: 10.2139/ssrn.4409587.