

Data Ethics: Conceptual Analysis of Dimensions and Implications in Data Policy

Mehdi Shaghghi*

Assistant Prof.; Faculty of Psychology and Education; Dept. of Information Science and Knowledge Studies Shahid Beheshti University; Tehran, Iran Email: m_shaghghi@sbu.ac.ir

Received: 16, Jun. 2025

Accepted: 15, Nov. 2025

Abstract: This article investigates the implications of the field of data ethics for data policy-making. To achieve this, we employed a structural conceptual analysis method with an implication-finding approach. First, the field is introduced, and the distinction between data ethics and data science ethics is discussed. It is argued that their relationship is one of specific commonality (*umum wa khusûs min wajh*), as data science ethics generally focuses on the issue of ethical data processing by software professionals, whereas data ethics concentrates on the production/ collection, storage, optimization, distribution, sharing, and use of data by corporations. Next, the difference between this field and big data ethics is addressed. Following a review of relevant research in data ethics, the field's implications for data policy are presented across six dimensions: Data Economy, Data Collection and Creation, Data Ownership, Open Data, Data Quality, and Data Protection. The findings show that data ethics, extending beyond common discussions in data science ethics or big data ethics, possesses a set of fundamental implications for data policy. In the absence of these implications, the groundwork is laid for structural biases, data injustice, and corporate exploitation of users. Firstly, the data economy transforms users into unpaid knowledge workers, turning their data into a source of commercial value, a tool for power consolidation, and a mechanism for directing consumer behavior. The conceptual analysis of the texts indicates that data collection and creation are never neutral; they are always influenced by ontological, epistemological, methodological assumptions, and power structures. Therefore, raw, unbiased data or "fact without interpretation" does not exist. In the domain of data ownership, it becomes clear that the distinction between use, control, and access has different consequences for individual rights, and many common ownership models practically favor data producers rather than their potential owners. Reviews show that open data only leads to transparency and public measurability if informational justice, analytical capacity, and equal access are guaranteed. Otherwise, it becomes a tool for strengthening the monopoly of technological elites. In the dimension of data quality, it was revealed that completeness, consistency, timeliness, and especially "relevance" play a crucial role in preventing confused decisions and data-driven

Iranian Journal of
Information
Processing and
Management

Iranian Research Institute
for Information Science and Technology
(IranDoc)

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 41 | No. 2 | pp. 549-594

Winter 2026

<https://doi.org/10.22034/jipm.2025.2063775.2037>



* Corresponding Author

policy errors. Furthermore, the analysis of the findings indicates that data protection - including preventing unauthorized access, safeguarding data against unforeseen secondary uses, and controlling the transfer of risk from organizations to users - is a fundamental prerequisite for public trust and reducing misuse by corporations and governments. Finally, the research shows that some organizational ethical documents and charters are mostly performative in function and can serve as a cover for the continuation of unethical behaviors.

Keywords: Data Ethics, Data Collection, Data Creation, Data Ownership, Open Data, Data Quality, Data Economy

اخلاق داده: تحلیل مفهومی ابعاد و دلالت‌ها در سیاست‌گذاری داده‌ها

مهدی شاقی

دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛
دانشگاه شهید بهشتی؛ تهران، ایران؛
پدیده‌آور رابط m_shaghghi@sbu.ac.ir



مقاله برای اصلاح به مدت ۵۰ روز نزد پدیده‌آور بوده است.

پذیوش: ۱۴۰۴/۰۸/۲۴

دریافت: ۱۴۰۴/۰۳/۲۶

نشریه علمی | رتبه بین‌المللی
پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شابا (چاپی) ۲۲۵۱-۸۲۲۳

شابا (الکترونیکی) ۲۲۵۱-۸۲۳۱

نما به در SCOPUS، ISI، LISTA و

jipm.irandoc.ac.ir

دوره ۴۱ | شماره ۲ | صص ۵۴۹-۵۹۴

زمستان ۱۴۰۴

<https://doi.org/10.22034/jipm.2025.2063775.2037>



چکیده: در این مقاله سعی شده دلالت‌های حوزه مطالعاتی اخلاق داده برای سیاست‌گذاری داده‌ها بررسی شود. در این بررسی از روش تحلیل مفهومی از نوع ساختاری با رویکرد دلالت‌پژوهی استفاده شد. برای انجام این کار ابتدا، به معرفی این حوزه و سپس، به تفاوت حوزه اخلاق داده و حوزه اخلاق علم داده پرداخته شده و استدلال شده که نسبت آن‌ها با هم عموم و خصوص من وجه است؛ زیرا اخلاق علم داده به‌طور عام به مسئله پردازش اخلاقی داده‌ها توسط متخصصان نرم‌افزار تمرکز دارد؛ در حالی که اخلاق داده به تولید یا گردآوری، ذخیره‌سازی، بهینه‌سازی، توزیع، و به‌اشتراک‌گذاری و استفاده از داده‌ها توسط شرکت‌ها متمرکز است. در ادامه، به تفاوت این حوزه با حوزه اخلاق کلان‌داده اشاره شده و پس از مرور برخی پژوهش‌های حوزه اخلاق داده، دلالت‌های حوزه اخلاق داده برای سیاست‌گذاری داده‌ها در شش بُعد اقتصاد داده، گردآوری و خلق داده، مالکیت داده، داده‌های باز، کیفیت داده، و حفاظت داده ارائه شده است. یافته‌ها نشان می‌دهد که اخلاق داده، فراتر از مباحث رایج در اخلاق علم داده یا اخلاق کلان‌داده، دارای مجموعه‌ای از دلالت‌های بنیادین برای سیاست‌گذاری داده است که در غیاب آن‌ها، زمینه برای سوگیری‌های ساختاری، بی‌عدالتی داده‌ای و بهره‌کشی شرکتی از کاربران فراهم می‌شود. از جمله این دلالت‌ها اینکه اقتصاد داده، کاربران را به کارگران دانشی بی‌مزد بدل کرده و داده‌های آنان را به منبع ارزش تجاری، ابزار تحکیم قدرت و سازوکاری برای جهت‌دهی رفتار مصرف‌کنندگان تبدیل می‌کند. تحلیل مفهومی متون نشان می‌دهد که گردآوری و خلق داده هیچ‌گاه خنثی نیست و همیشه تحت تأثیر مفروضات هستی‌شناختی، معرفت‌شناختی، روش‌شناختی و ساختارهای قدرت انجام می‌شود و از این‌رو، داده خام بی‌سوگیری یا «واقعیت بدون تفسیر» وجود ندارد. در حوزه مالکیت داده نیز روشن می‌شود که تمایز میان استفاده، کنترل، و دسترسی پیامدهای متفاوتی

برای حقوق افراد دارد و بسیاری از الگوهای رایج مالکیت، در عمل به نفع تولیدکنندگان داده و نه صاحبان بالقوه آن عمل می‌کنند. بررسی‌ها نشان می‌دهد که داده‌های باز تنها زمانی به شفافیت و امکان سنجش عمومی می‌انجامند که عدالت اطلاعاتی، توان تحلیل، و دسترسی برابر تضمین شده باشد و در غیر این صورت، به ابزاری برای تقویت انحصار نخبگان فناور تبدیل می‌شود. در بُعد کیفیت داده نیز آشکار شد که کامل بودن، سازگاری، به‌هنگام بودن و به‌ویژه «رابط»، نقش اساسی در جلوگیری از تصمیمات مغشوش و خطاهای سیاستی مبتنی بر داده دارند. افزون بر این، تحلیل یافته‌ها نشان می‌دهد که حفاظت داده اعم از جلوگیری از دسترسی نامجاز، صیانت از داده در برابر استفاده‌های ثانویه پیش‌بینی نشده، و کنترل انتقال ریسک از سازمان‌ها به کاربران، پیش‌نیازی بنیادین برای اعتماد عمومی و کاهش سوءاستفاده‌های شرکت‌ها و دولت‌هاست. سرانجام، پژوهش نشان می‌دهد که برخی اسناد و منشورهای اخلاقی سازمانی، بیشتر کارکردی نمایشی داشته و می‌توانند پوششی برای تداوم رفتارهای ضد اخلاقی باشند.

کلیدواژه‌ها: اخلاق داده، گردآوری داده، خلق داده، مالکیت داده، داده‌های باز، کیفیت داده، اقتصاد داده

۱. مقدمه

روند روبه‌رشد فناوری، قلمروهای نوی را پیش روی آورده تا بتوان طبیعت را هرچه بیشتر کنترل و همچون منبعی از آن استفاده و ذخایرش را استخراج نمود که این خود شایسته توجه اخلاقی ویژه‌ای است. یکی از علومی که به انسان‌ها در مدیریت و کنترل محیط اطراف خود کمک می‌کند، علم داده است که هدف غایی آن ساخت ابزارها و توسعه مفاهیمی برای شناخت و پیش‌بینی آینده برای انواع مداخله و سرمایه‌گذاری است. در مدرنیته متأخر که مبتنی بر سرمایه‌داری مالی است، انبوه داده‌ها از فعالیت‌های روزانه افراد گردآوری می‌شود تا مسیر سرمایه‌گذاری آینده پیش‌بینی شود. در نتیجه این تحلیل‌ها چشم‌اندازی حاصل می‌شود که روند واردات، صادرات، تجارت، خرده‌فروشی، مدیریت نیروی انسانی، روش‌های بهداشتی و درمانی، سازوکارهای سیاسی، تغذیه، انتظارات، و حتی ازدواج و تولید مثل را متأثر می‌سازد. حوزه اخلاق داده ایجاد شده تا مخاطرات تولید استفاده غیرمسئولانه از داده‌ها را مطالعه کند و مبنایی برای سیاست‌گذاری این حوزه ایجاد نماید. اخلاق داده شاخه جدیدی از علم اخلاق است که به مطالعه و ارزیابی مسائل و مشکلات اخلاقی مرتبط با داده‌ها (شامل تولید، ذخیره‌سازی، بهینه‌سازی، پردازش، توزیع، به اشتراک‌گذاری، و استفاده)، الگوریتم‌ها (شامل هوش مصنوعی، عوامل مصنوعی، یادگیری ماشین و ربات‌ها) و رویه‌های حاصل از داده‌ها (شامل نوآوری مسئولانه،

برنامه‌نویسی، هک، و دستورالعمل‌های حرفه‌ای) می‌پردازد تا با معرفی راه‌حل‌های اخلاقی، پایه‌ای برای سیاست‌گذاری ایجاد نماید. اخلاق داده یک مرتبه بالاتر از اخلاق اطلاعات و اخلاق رایانه است؛ زیرا در همان حال که بر داده تمرکز دارد، مسایل حوزه‌های یادشده را نیز کل‌نگرانه بازبینی می‌کند (Floridi & Taddeo 2016). هدف حوزه اخلاق داده، اخلاقی‌تر کردن «علم داده»، ارتقای اخلاق حرفه‌ای در علم داده، یا ارائه اصولی برای استفاده اخلاقی از روش‌های علم داده (مثل آنچه در کتاب «چارچوبی برای استفاده اخلاقی از روش‌های پیشرفته علم داده در امور بشردوستانه»^۱ (۲۰۲۰) آمده) نیست، بلکه هدفش زمینه‌سازی برای سیاست‌گذاری مطلوب در باب تولید، توزیع، پردازش، و مصرف داده‌هاست تا در آینده، سازمان‌ها و شرکت‌ها مؤظف به رعایت معیارهای خاصی شوند و با آن مورد سنجش قرار گیرند.

یکی از خلأهای سیاست‌گذاری در ایران، درباره استفاده منفی از الگوهای مستخرج از داده‌های عمومی است. برای مثال، شرکت بیمه می‌تواند با استخراج الگو از سوابق درمانی انسان‌ها، گروه‌های پرهزینه را از دایره شمول خود حذف کند (Crawford 2013). دیگری در استفاده بی‌اجازه از داده‌های فعالیت کاربران علیه آن‌ها و انتقال ریسک (Gille 2013) به آن‌ها از راه‌هایی مانند فروش داده‌ها و عادت‌دهی اعتیادگونه به برخی کالاهاست که ساده‌ترین شکل آن فروش داده‌های شخصی مردم به شرکت‌ها برای ارسال پیام تبلیغاتی است. دیگری، استفاده از داده‌های شرکتی توسط دولت برای رصد، پایش، شناسایی افراد، فیلترینگ، و سایر اقداماتی است که نوعی انتقال ریسک از دولت به ملت است. برای مثال، جام جم آنلاین افشا کرد که شرکت‌های ارائه‌دهنده خدمات اینترنتی، داده‌های کاربران را به شرکتی به نام «سحاب‌پرداز» می‌فروشنند تا در زمینه فیلترینگ هوشمند از آن‌ها استفاده شود (جام جم آنلاین ۱۳۹۶). مورد دیگر، فروش تراکنش‌های کاربران به شرکت‌های خارجی است. برای مثال، تراکنش افراد و شرکت‌ها در سایت صرافی‌های خرید و فروش ارز دیجیتال به‌سادگی می‌تواند فروخته شود و در اختیار کشورهای دیگر قرار گیرد. ایجاد اختلال در قیمت سکه، دلار و کالاهای اساسی، به‌واسطه وجود داده‌های عرضه و تقاضای کالا و خرید مردم به‌سادگی امکان‌پذیر است. و مورد آخر، سوگیری در گردآوری، خلق، و پردازش داده و از این‌رو، معوج شدن شکل‌دهی افکار عمومی یا

1. A Framework for the Ethical Use of Advanced Data Science Methods in the Humanitarian Sector

سیاست‌گذاری است. از آنجا که اصل ۴۴ قانون اساسی در ایران، ضامن حرکت شتابان به سمت اقتصاد شرکتی و رشد سرمایه‌داری شرکتی است و چون شرکت‌ها علاقه زیادی به بازاریابی اجتماعی، پیش‌بینی رفتار مراجعان، پیش‌بینی رفتار صنایع داخلی و خارجی، و نقش‌آفرینی در سیاست‌گذاری ملی و صنعتی و مشارکت در بازی قدرت دارند (Diaz 2008)، به همین دلیل، می‌توانند بیشترین استفاده را از داده‌های کاربران برای شرکت در بازی قدرت ببرند. هرچند ممکن است دولت بخواهد از داده‌های هویت‌های حقیقی و حقوقی به نفع امنیت ملی یا مدیریت مصرف کالاهای اساسی استفاده کند، ولی همین داده‌ها می‌تواند مورد سوء استفاده نظام‌های حاکمیتی نیز قرار گیرد و یا توسط شرکت‌های خصوصی به ضرر مردم و دولت بهره‌برداری شود و در واقع، شمشیری دو لبه است که نبود سیاست‌های اخلاقی ملموس می‌تواند مشکلاتی جدی ایجاد کند. بنابراین، مسئله این است که با توجه به اینکه تمرکز حوزه اخلاق داده بر مطالعه سیاست‌های داده‌ای است و با توجه به سوء استفاده‌ها و سوگیری‌هایی که بازتاب‌دهنده خلأ سیاستی است، این حوزه مطالعاتی چه دلالت‌هایی برای سیاست‌گذاری داده‌ها دارد و چه ابعادی را می‌توان بر اساس مطالعات این حوزه برای سیاست‌گذاری داده‌ها پیشنهاد کرد.

روش

مطالعه حاضر به روش «تحلیل مفهومی» انجام شده که روشی فلسفی-تحلیلی برای شفاف کردن و فهم‌پذیر ساختن مفاهیم و یافتن دلالت‌هاست (Furner 2004). در این روش، به جای تمرکز بر داده‌های تجربی، به کندوکاو معنا، ساختار، و روابط میان مفاهیم برای پیدا کردن دلالت‌ها تمرکز می‌شود. هسته اصلی روش تحلیل مفهومی روشن‌تر کردن مفاهیم جهت معلوم ساختن ابعاد ضروری یک مفهوم خاص است. «کاستریک» روش تحلیل مفهومی را شامل سه روش فرعی می‌داند: روش ساختاری، روش ردیابی، و روش تقلیل‌دهی. تحلیل مفهومی ساختاری عبارت است از روشن کردن مفهوم از طریق نگاه به کاربردهای مختلف آن و تلاش برای فرموله کردن تعریفی دقیق یا مشخص کردن ابعاد یک مفهوم؛ در تحلیل مفهومی ردیابی، شفاف کردن مفهوم یا از طریق مثال نقض اتفاق می‌افتد و یا از طریق ارائه نمونه‌هایی انجام می‌شود که با تعریف آن مفهوم منطبق است، ولی شهوداً و در بدو امر احساس می‌شود که آن نمونه‌ها نمی‌تواند نمونه‌ای مناسب برای آن مفهوم باشد؛ و در تحلیل مفهومی به روش تقلیل‌دهی، یک مفهوم پیچیده به اجزای پایه

تجزیه می‌شود و با تعریف آن اجزای پایه، مختصات مفهوم به دست می‌آید (Kosterec 2016). در مقاله حاضر از روش تحلیل ساختاری با رویکرد دلالت‌پژوهی استفاده شده و سعی گردیده با بررسی مفاهیم یک حوزه مطالعاتی، دلالت‌های آن برای سیاست‌گذاری داده مشخص شود. «جیمز مور» در مقاله معروف «اخلاق رایانه چیست؟» از همین روش برای به فهم درآوردن این حوزه موضوعی استفاده کرده و اتفاقاً، در آن مقاله یکی از مهم‌ترین مشکلات مفاهیم جدید حوزه رایانه، فناوری سایبر، و فضای مجازی را «آشفته‌گی مفهومی» خواند و آن را عامل عدم امکان پُر کردن خلأهای سیاستی دانست (Moor 1985). در اینجا نیز شاید یکی از دلایل وجود خلأ سیاستی در خصوص داده و اخلاق داده، آشفته‌گی مفهومی حول و حوش این مفهوم باشد. در مقاله حاضر (۱) پیکره‌ای از متون مرتبط با اخلاق داده گردآوری شد (۹۳ مقاله و کتاب؛ ۲) آن متون با توجه به مفاهیم مشترک‌شان ذیل شش موضوع کلی دسته‌بندی شد؛ (۳) کتب و مقالات نامرتب با مضمون اصلی حذف شد؛ (۴) تعداد ۴۴ مقاله و کتاب در شش موضوع کلی بررسی شد؛ (۵) مضامین اصلی این پیکره استخراج گردید؛ و (۶) این مضامین بنیادی به‌عنوان دلالت‌های متون حوزه اخلاق داده برای سیاست‌گذاری داده عرضه شد.

سوابق مطالعاتی در حوزه اخلاق داده

در ایران دانش عمومی و سازمانی در خصوص اخلاق داده اندک است. با جست‌وجوهای اولیه در محیط وب، مشاهده می‌شود که کمتر از دو یا سه اثر درباره «اخلاق داده» یا «اخلاق مدیریت داده» وجود دارد. می‌توان متون و منابع مختلفی را به فارسی درباره اخلاق کلان‌داده یافت، ولی اخلاق کلان‌داده فقط به پردازش داده‌ها تمرکز دارد و به گردآوری، حفاظت، استفاده، و پایش داده‌ها بی‌توجه است. عموم مطالعات از جنبه کاربردی به کلان‌داده نگریسته‌اند و وقتی از چالش‌ها صحبت کرده‌اند نیز بیشتر چالش‌های عملی مطرح بوده و جنبه‌های اخلاقی آن را مغفول گذاشته‌اند. اشتیاق به استفاده ابزار از کلان‌داده‌ها در این متون به شدت مشهود است. برای مثال، «عرب‌مازار یزدی و مرادی» اظهار کرده‌اند که با تحلیل کلان‌داده‌ها می‌توان دقت را در شناسایی و پیش‌بینی درآمدهای مالیاتی افزایش داد؛ یعنی می‌توان آن‌هایی را که باید مالیات بدهند به‌خوبی شناسایی نمود، اعتبارسنجی سوابق پرداخت را برای متقاضیان وام در بانک‌ها به‌خوبی انجام داد، الگوهای مخارج سازمانی را در سازمان‌ها استخراج، و ناکارآمدی‌های عملکردی را شناسایی کرد (۱۳۹۹). «میرعرب

و نوروبی فیروز» نیز به حمل و نقل و رفتار مردم در استفاده از اتوبوس‌ها و پیش‌بینی ایستگاه‌های شلوغ و روزها و ساعات پرتردد اشاره کرده یا به کاربرد کلان‌داده‌ها در پیش‌بینی پاندمی‌ها تأکید نموده‌اند (۱۳۹۹). پژوهش‌های «میرعرب، میرطاهری و اصغری» (۱۴۰۰)، «سعادت و مهرشاد» (۱۳۹۶)، «هللی، مظلوم و هادیان» (۱۳۹۴)، «فیروزی» و همکاران (۱۴۰۰)، و «جمشیدی و دهقانی سانچ» (۱۳۹۹) نیز از این نوع‌اند. همه این تحقیقات بر دو مؤلفه کلیدی «شناسایی» و «پیش‌بینی» تأکید دارند. برخی پژوهش‌ها به مشکلات سیاست‌گذاری کلان‌داده‌ها در ایران پرداخته‌اند. برای مثال، «دادگستر» از طریق مصاحبه، این مشکلات را به ترتیب اهمیت در هفت مقوله سیاسی، فناوریانه، اقتصادی، حقوقی، اجتماعی، آموزشی، و امنیتی قرار داده ولی خبری از مقوله اخلاقی نیست (۱۴۰۲). در برخی پژوهش‌ها مثل پژوهش «رهبری»، گلایه شده که چرا نهادهایی همچون «ساترا»^۱ اجازه نمی‌دهند که از کلان‌داده‌ها در افزایش قدرت بازاری شرکت‌های ایرانی استفاده شود! (برای مثال، «ساترا» در ماده ۴ دستورالعمل خود امکان استفاده از داده‌های اشخاص ثالث توسط رسانه‌های کاربرمحور دروازه‌بان را رد کرده است) (۱۴۰۲). برخی پژوهش‌ها مثل پژوهش «ملکی، منصوری و احمدی» به بررسی جنبه‌های اخلاقی کلان‌داده‌ها از منظر قرآن کریم پرداخته‌اند، ولی معیارهایی که عرضه کرده‌اند، بسیار کلی، مبهم، و غیرعملیاتی است (۱۴۰۰). «فیروزی» و همکاران ذکر کرده‌اند که هر نوع استفاده‌ای از کلان‌داده که باعث شود قناعت، عدم تجسس، خیرخواهی، حسن ظن و امثال آن زیر سؤال رود، غیراخلاقی است، ولی مشخص نکرده‌اند که مثلاً چه استفاده‌ای از کلان‌داده باعث زیر سؤال رفتن قناعت می‌شود و چگونه تشخیص می‌دهیم که قناعت را زیر سؤال برده است (۱۴۰۰). گروهی دیگر از پژوهش‌ها نیز جنبه‌های اخلاقی کلان‌داده‌ها را از منظر فلسفه اخلاق بررسی کرده‌اند. برای مثال، «قاسمی و ثقه‌الاسلامی» بحث کرده‌اند که امکان گردآوری و تحلیل کلان‌داده‌ها، عاملیت اخلاقی انسان‌ها را تحت تأثیر قرار داده است. برای مثال، کسانی که با استفاده از کلان‌داده می‌توانند روند رشد قیمت‌ها در کالاهای اساسی (زمین، طلا، مسکن، ارز دیجیتال، و امثال آن) را پیش‌بینی کنند، زودتر از بقیه دست به خرید می‌زنند که منجر به محدود شدن (اراده آزاد و امکان انتخاب) امکان خرید سایرین می‌شود (۱۴۰۰).

هرچند مطالعه اخلاقی داده‌ها در متون فارسی‌زبان بسیار محدود است، ولی بدنه

منسجمی از مباحث مهم در خصوص اخلاق داده در متون انگلیسی قابل ردگیری است. ابتدا «فلوریدی و تادئو» پیشگامانه بر معرفی حوزه‌ای به نام اخلاق داده تمرکز کرده‌اند. طبق نظر آن‌ها در اخلاق داده بررسی می‌شود که اساساً گردآوری اخلاقی داده‌ها چگونه است. آیا ما می‌توانیم داده‌ها را به هر شکل ممکن گردآوری کنیم و به لحاظ اخلاقی مجاز به این کار هستیم؟ آیا می‌توانیم آن‌ها را به هر شکل دلخواه مورد پردازش و استفاده قرار دهیم و آیا از نظر اخلاقی مجاز هستیم هر طور که دلمان خواست، آن‌ها را ذخیره‌سازی کنیم؟ همچنین اگر بخواهیم تراکنش داده‌ها را مورد پایش قرار دهیم و در واقع، اقدام به پایش داده‌ای کنیم، باید چه معیارهایی را مورد توجه قرار دهیم. بحث اخلاق داده از آنجا اهمیت دارد که بسیاری از داده‌ها ممکن است هرگز تبدیل به اطلاعات نشوند، ولی در پیش‌بینی و تصمیم‌سازی مؤثر باشند. بحث در باب عملیات پردازش و نحوه عمل سخت‌افزار و نرم‌افزار روی داده‌ها و نحوه تغییردهی آن‌ها و چگونگی جلوگیری از سوگیری پردازشی نیز از نظر آن‌ها جزو دامنه اخلاق داده است و اخلاق داده نوعی رویکرد کلان اخلاقی است (Floridi & Taddeo 2016). گروهی دیگری از متون انگلیسی‌زبان در حوزه اخلاق داده، جنبه عملیاتی و آیین‌نامه‌ای دارند و به ذکر قواعد کلی اخلاق داده در سازمان‌ها و شرکت‌ها می‌پردازند. برای مثال، «چارچوب اخلاق داده» سندی است که دولت بریتانیا برای هدایت و سنجش عملکرد داده‌ای سازمان‌ها و شرکت‌ها ارائه کرده است. در این سند سه اصل اساسی شفافیت، پاسخگویی و انصاف و پنج اقدام خاص شامل ۱) تعریف نفع عمومی و نیاز کاربران؛ ۲) درگیر کردن تخصص‌های مختلف در کار داده‌ها، ۳) هماهنگی فعالیت‌ها با قانون، ۴) مرور کیفیت و محدودیت‌های داده‌ها، و ۵) ملاحظه و ارزیابی عواقب سیاست‌ها و تصمیمات ارائه شده است (UK Government Digital Services 2020). کتاب «اخلاق داده: خط‌مشی‌ها و اصولی برای نهادها، شرکت‌ها [و سازمان‌ها]» اثر (Tranberg et al. 2018) و موارد متعدد دیگر نیز از همین نوع است. گروه دیگری از متون به نقد کلان‌روایت‌ها و پیش‌فرض‌های اخلاق داده پرداخته و پرسیده‌اند که گفتمان اخلاق داده چگونه و چرا شکل گرفته است. برای مثال، «هوبلر» و «زفیرو» در مقالات خود به مطالعه و نقد پیش‌فرض‌هایی چون سودگرایی، حضور راه‌حل داده‌کاوانه در همه چیز، و طرف صحبت بودن طراحان، مهندسان و صاحبان شرکت و هیچ‌انگاری سایرین پرداخته‌اند (Häußler 2021; Zeffiro 2021). پژوهش‌هایی نیز به چگونگی استفاده از کلان‌داده‌ها در بحران‌ها و موقعیت‌های حساس و نحوه تعهد و التزام به قواعد اخلاق داده پرداخته‌اند. برای مثال، «یالوپ و

علی‌اصغر» به نحوه به کارگیری الگوهای حکمرانی داده در بحران‌هایی نظیر همه‌گیری کرونا پرداخته و گفته‌اند که به خاطر تماس بیشتر مردم با ابزارهای الکترونیک در دوره کرونا و افزایش حساسیت آن‌ها نسبت به حریم داده‌ای، اصول اخلاق داده چگونه باید تغییر پیدا کند (Yallop and Aliasghar 2020). سرانجام، یک گروه مهم دیگر از نوشته‌ها نیز اسناد راهبردی کشورها و سازمان‌های بین‌المللی در خصوص اخلاق داده است که وظایفی اخلاقی را به دوش کارکنان دولت یا بخش خصوصی می‌گذارند. برای مثال، در «راهبرد داده‌ای فدرال» مربوط به ایالات متحده آمریکا که مخاطب آن کارکنان دولت است، وظایف اخلاقی در هفت بخش رعایت مقررات، احترام به مردم، احترام به حریم خصوصی، صداقت عمل، پاسخگویی، شفافیت، و آگاهی از مفاهیم حوزه مدیریت داده و علم داده آورده شده است (White House, Office of Management and Budget, President's Managerial Agenda 2019). یا می‌توان به سند «اصول اعمال شایسته در خصوص اخلاق داده در بخش عمومی» اشاره کرد که توسط «سازمان همکاری و توسعه اقتصادی» تنظیم شده است. در این سند، به کارکنان بخش عمومی توصیه شده که متعهد به ۱۰ اصل اخلاقی در خصوص داده‌ها باشند: (۱) گردآوری صادقانه، (۲) به کارگیری اخلاق داده در عمل، (۳) توجه به مقررات دولتی در خصوص اخلاق داده، (۴) کنترل ورودی‌های داده‌ای، (۵) توجه به نحوه‌های استفاده از داده‌های مردم، (۶) تعیین حدود برای گردآوری داده، (۷) شفافیت، (۸) بیشتر کردن کنترل کارکنان بخش‌ها روی داده‌های حوزه کاری‌شان، (۹) انتشار داده‌های باز، و (۱۰) پاسخگو بودن هنگام مدیریت ریسک (OECD 2020). بنابراین، می‌بینیم که حجم قابل توجهی از متون توسط سازمان‌ها، نهادهای غیرانتفاعی، و ناشران بین‌المللی در خصوص اخلاق داده منتشر شده که نیاز است درونی‌سازی، هضم، بازاندیشی شده، و استفاده گزینشگرانه و هوشمندانه از آن‌ها صورت گیرد و در جهت منافع ملی استفاده شود.

مفهوم‌شناسی: تفاوت اخلاق داده با اخلاق علم داده و اخلاق کلان داده

گفته شد شاخه‌ای از اخلاق حرفه‌ای در علوم مهندسی رایانه وجود دارد که «اخلاق علم داده» نامیده می‌شود. اخلاق علم داده، زیرمجموعه اخلاق حرفه‌ای مهندسان رایانه است که در دامنه شمول اخلاق حرفه‌ای توسعه‌دهندگان سیستم‌های خبره، الگوریتم‌های یادگیری ماشین، و برنامه‌نویسان نیز می‌گنجد. مهم‌ترین دغدغه فکری مهندسان رایانه و

توسعه‌دهندگان سیستم‌های خبره، «سوگیری» و «نویز» است. آن‌ها به این تمرکز دارند که الگوریتم‌هایی که توسعه می‌دهند یا مدل‌هایی که می‌سازند، چقدر از سوگیری‌های ذهنی و تکثر ارزیابی‌ها تأثیر می‌پذیرد. اگر مدل یا الگوریتمی خطا داشته باشد، چه عواقبی برای کاربران خواهد داشت و چگونه باید در توسعه الگوریتم‌ها، از نفوذ سوگیری‌ها و تشتت ارزیابی یا پیش‌بینی جلوگیری کرد و منافع عمومی را در نظر گرفت و منافع عمومی را مقدم بر منافع سرمایه‌گذار دانست. بنابراین، متخصصان رایانه و توسعه‌دهندگان سیستم‌های خبره و متخصصان علم داده اصول و قواعدی اخلاقی را برای خود تنظیم می‌کنند تا الگوریتم‌ها را از سوگیری و نویز مصون کنند.

اگر هنگام پردازش مجموعه‌های عظیم داده‌ها، برخی اصول اخلاقی رعایت نشود، آنگاه هویت بسیاری از افراد برای پایگاه‌داران آشکار می‌شود؛ برخی گرایش‌های فکری، مصرفی، یا عادات آن‌ها که مربوط به مسئله مورد پردازش نیست، هویدا می‌شود؛ و الگوهای به‌دست می‌آید که با آن می‌توان از مردم سوء استفاده کرد و بر نقاط ضعفشان انگشت گذاشت. بنابراین، هنگام پردازش کلان‌داده‌ها باید اصولی را رعایت کرد (برای مثال، دو مجموعه داده نامرتب به هم، هنگام پردازش به هم مرتبط نشوند) تا هویت‌ها، عادات، سلیقه، الگوی مصرف، گرایش فکری و امثال آن آشکار نگردد و از ضعف‌های مردم برای تعدی به آن‌ها یا فروختن کالا و خدمات به آن‌ها استفاده نشود (Harron et al. 2017). بنابراین، اخلاق کلان‌داده نیز یکی از زیرمجموعه‌های اخلاق داده است که صرفاً به اصول اخلاقی پردازش داده تمرکز دارد و به تولید، ذخیره‌سازی، به‌اشتراک‌گذاری، استفاده، پایش و امثال آن بی‌توجه است؛ در حالی که اخلاق داده، حوزه کلانی است که همه جنبه‌های اخلاقی داده را مد نظر قرار می‌دهد و نگاه آن، اجتماعی و سیاست‌گذارانه است.

ابعاد قابل توجه در سیاست‌گذاری اخلاقی داده‌ها

۱. اقتصاد داده

حتی قبل از ظهور «وب ۲/۰» که اجازه ایجاد، انتشار و تبادل محتوا بدون نیاز به اتکا به واسطه‌ها را می‌داد، محققان موضوع «نیروی کار دیجیتال رایگان» را مطرح کرده بودند (Scholz ۲۰۱۳) که در واقع، امروز به‌عنوان کارگران دانشی شناخته می‌شوند. در اواخر دهه ۱۹۹۰، کاربران خدمات آنلاین چون تابلوهای پیام و خدمات چت شرکت آمریکا

آنلاین (AOL) در واقع، نقش «رهبران جامعه» و مدیران ملی را بازی می‌کردند و کارشان این بود که روی محتواهای نوشته‌شده در تابلوهای پیام کامنت می‌گذاشتند و آن‌ها را تحلیل می‌کردند، ولی کسی نمی‌دانست که آن‌ها مزدبگیر شرکت‌اند و در اوایل دهه ۱۹۹۰ به این کاربران تقریباً ۳/۵ دلار در روز پول پرداخت می‌شد (Margonelli 1999). بنابراین، نیروی کار کم‌حقوق ولی مؤثری تشکیل شده بود که به سود شرکت کمک می‌کردند؛ زیرا محتوا را به روشی تحلیل یا پُررنگ می‌کردند که آن را برای دیگران جذاب‌تر می‌کرد و همچنین باعث می‌شد که کاربران دیگر، برای دسترسی به AOL پول پرداخت کنند. این کاربران فعال، در تغییر ذائقه یا جهت فکری مردم نیز بسیار مؤثر بودند و به همین دلیل نقش پنهان رهبری جامعه را هم داشتند و در همان حال، به سود شرکت کمک می‌کردند. اکنون نیز اوضاع بر همین منوال است. مشتریان خدمات جدید و سایت‌های شبکه‌های اجتماعی نیز برای ایجاد ارزش تجاری بسیار مهم هستند؛ زیرا محتوایی را در بخش کامنت‌ها و گفتگوها تولید می‌کنند که دیگران را به استفاده از پلتفرم یا سایت شرکت تشویق می‌کنند. این رویکرد که طی آن کاربران به «مصرف‌کننده تولیدگر»^۱ تبدیل می‌شوند، در زمینه‌های کسب‌وکار مورد سوء استفاده فراوانی قرار گرفته و بعدها به‌عنوان شکلی از کار آزاد توصیف گردیده است (Fuchs 2006) که می‌تواند درآمدهایی برای کاربران فعال هم داشته باشد. این داده‌ها از شش جنبه برای شرکت‌ها منفعت مالی دارند: ۱) تجارت مستقیم داده (فروش مستقیم داده، پلتفرم‌های ابری داده به‌عنوان خدمت (DaaS)^۲، ساخت الگوریتم از داده و فروش آن‌ها)، ۲) کاهش هزینه‌ها و افزایش کارایی (بهینه‌سازی فرایندهای کاری، پیش‌بینی بازار، تخصیص بهینه منابع)، ۳) ارزش‌آفرینی (شخصی‌سازی، توسعه کالاهای جدید، بهبود کالاهای موجود)، ۴) راهبردسازی (تحلیل رقبا، فروش هدفمند، بهینه‌سازی قیمت)، ۵) مدیریت ریسک (کشف تقلب‌ها، کشف انحرافات شرکت از قوانین و اصلاح آن، تحلیل ریسک شرکا، بدهکاران و قسط‌دهندگان)، و ۶) خدمات به دولت (فروش داده‌های کاربران کشور میزبان به دولت توسط شرکت‌های چندملیتی جهت بهره‌گیری اقتصادی-امنیتی).

در این حالت، کاربران هدف تبلیغات و خدمات اضافی شرکت‌های فناوری و شرکای

1. prosumer

2. Data as a Service (DaaS) cloud-based software tools

تجاری‌شان واقع می‌شوند. تعاملات کاربران با یکدیگر و محتواهای حاصل از تحلیل داده‌ها که به‌عنوان خوراک اولیه به خودِ کاربران تزریق می‌شود، برای تعیین نوع محتوای تولیدشده حیاتی است. «فوکس» خلاصه می‌کند که: «هر زمان که کاربران فیس‌بوک آنلاین هستند، داده‌هایی ایجاد می‌کنند که به پروفایل‌ها و رفتار آنلاین آن‌ها اشاره دارد. این داده‌ها به مشتریان تبلیغات چی فیس‌بوک فروخته می‌شود؛ گروه‌هایی که قادر به ارائه تبلیغات هدفمند بر روی پروفایل کاربران هستند». وی یادآور می‌شود که از همین داده‌ها برای کامنت گذاشتن و جذب مشتری در وبسایت‌های دیگر شرکت استفاده می‌شود. محتواهایی که توسط خود کاربران تولید می‌شود و مشارکت کاربران را هم زیاد می‌کند، محدود به پلتفرم‌هایی که به تولید و بازتولید داده‌های کاربران می‌پردازند، نیست (Fuchs 2006). همان‌گونه که «گریتز و هلموند» (۲۰۱۳) نشان می‌دهند، بیشتر داده‌ها در گیرودار تعاملات پیچیده بین پلتفرم‌های مختلف ایجاد می‌شود. امکان ادغام شبکه‌های اجتماعی مختلف و «پروتکل گراف باز»^۱ امکان «اتصال بک‌اند»^۲ بین پلتفرم‌ها را ایجاد می‌کند. آنچه محققان از آن به‌عنوان «اقتصاد لایک» نام می‌برند، یک محیط آنلاین تعاملی را در اختیار کاربران قرار می‌دهد که جمع‌آوری داده‌های غیرمتمرکز را با تحلیل‌های متمرکز امکان‌پذیر می‌کند و افزون بر آن، کاربرانی با حقوق بسیار اندک در پشت صحنه وجود دارند که لایک می‌کنند و دیگران را به لایک کردن ترغیب می‌نمایند.

داده‌هایی که شرکت‌های فناوری و پلتفرم‌های شبکه‌ای آن‌ها تولید می‌کنند، دارایی‌های حیاتی آن‌ها هستند، ولی شیوه‌های گردآوری این شرکت‌ها باید قانون‌مند شود. نگرانی‌هایی که وجود دارد، یکی مربوط به حریم خصوصی و دیگری درباره‌ی نیاز به چارچوب‌های قانونی برای ضربه‌نزدن مقررات حریم خصوصی به پیشرفت‌های فناوری است (Tene and Polonetsky 2013). به‌رغم انتقادهای فراوان در باب استفاده‌های شرکتی از داده‌ها، تحقیقاتی که بر اساس این داده‌ها انجام می‌شود نیز تحت تأثیر این نگرانی‌هاست. ممکن است استدلال شود که رفاه، کاربری اقتصادی فردی، و ارزش تعاملی استفاده از پلتفرم‌ها برای توجیه کنار گذاشتن رضایت آگاهانه کافی است، به‌ویژه اینکه دولت‌ها برای شهروندان «وظیفه‌ی مشارکت» در اقتصاد شرکتی را قائل‌اند تا بدین وسیله حداکثر رفاه

1. open graph protocol
2. back-end connection

مطلوب برای شهروندان و حداکثر توان کنترل نظم امور برای دولت حاصل آید. ولی نباید مشارکت آگاهانه یک فرد در برخی اقدامات عمومی را با ثبت اجتناب‌ناپذیر و ناخواسته داده‌های شخصی دیجیتال و کشاندن ناخواسته او به سمت ثبت داده‌هایش یکی دانست. «لوپتون» با بررسی فعل و انفعال بین مصرف‌کننده تولیدگر و استفاده‌کنندگان ابزاری از داده‌ها در حوزه سلامت شخصی و عمومی استدلال می‌کند که امروزه صنعت سلامت و بهداشت، یک «اقتصاد تجربه دیجیتال بیمار» را ایجاد کرده که در آن، تمایل افراد برای ارائه داده (و اطلاعات) در مورد بیماری‌ها و درمان از نظر اخلاقی ارزشمند و حتی پولی شده است (Lupton 2014). این روش با طراحی وب‌سایت‌های نظردهی (مثل نی‌نی سایت در ایران) به کار می‌رود تا بیماران تجارب درمانی خود را تشریح کنند. اخلاقی جلوه دادن عرضه داده درباره تجارب شخصی برای بهره‌مندی صاحبان نگاه‌ها، در تمایل به ایجاد انتظار اخلاقی از کاربران برای کمک از طریق عرضه داده‌ها و اطلاعات شخصی برای مصلحت عمومی انعکاس یافته است؛ غافل از اینکه بسیاری از این داده‌ها برای تخمین شاخص باروری در کشورهای مختلف (اینکه حدوداً چه تعداد زن در سن باروری هستند، امید به باروری حدوداً چقدر است و امثال آن) جهت تخمین چشم‌انداز امنیتی، نظامی و تولیدی کشورهای مختلف در سال‌های آینده (مثلاً ۲۰ سال آینده) به راحتی قابل استفاده است. بنابراین، در دو سوی وبگاه‌های شرکت‌ها، یعنی هم در سوی صاحبان و مالکان و هم در سوی کاربران و حتی در سوی مخالفان، دست‌هایی پنهان وجود دارد که یکی نقش کاربر معمولی، دیگری نقش مخالف و یکی نیز نقش صاحب شرکت را بازی می‌کند و در این بازی داده‌ها جهت تثبیت هژمونی اقتصادی، یکسان‌سازی، تسلط بر رقبا، انقیاد مصرف‌کنندگان، و تقویت «برنده همه را می‌برد» استفاده می‌شود.

۲. گردآوری و خلق داده

موضوع گردآوری و خلق داده، بیشتر به قالب (چارچوب) گردآوری داده اشاره دارد؛ یعنی چه داده‌هایی گردآوری و خلق شود، قالب گردآوری یا خلق داده چطور مفهوم‌سازی، مقوله‌بندی یا طیف‌بندی شود، و به واسطه این قالب چه داده‌هایی مرتبط دانسته شود. قالب گردآوری داده از سه جنبه واجد اهمیت است: هستی‌شناسی داده‌ها (O'Neil 2016)، معرفت‌شناسی داده‌ها (Mittelstadt et al. 2016)، و بر ساخت اجتماعی داده‌ها (Acemoglu and Johnson 2023).

از جنبه هستی‌شناختی، سه بُعد قابل بررسی است: الف) تعریف واقعیت و اینکه چه چیز دانستنی است: عمل خلق داده به‌طور تلویحی بیانگر ایستار هستی‌شناختی خالق یا گردآورنده داده‌هاست (O'Neil 2016). وقتی کسی تصمیم می‌گیرد که درباره «رفاه اجتماعی» داده خلق یا گردآوری کند، ابتدا تعریف می‌کند که «رفاه اجتماعی» چیست و سپس ابعاد قابل سنجش آن کدام است (مثلاً درآمد، سلامت، مشارکت اجتماعی و غیره) که هر دو این اعمال، هستی‌شناختی است. این تعاریف، برساخت انسانی است و واجد حقیقت ذاتی نیست. ب) قالب‌بندی مسئله‌ها: اول اینکه چه چیز اساساً مشکل دانسته و تعریف شود و چه چیز از دایره مشکلات خارج گردد. اگر موضوع رفاه اجتماعی در کشور ایران در دایره مشکلات قرار نگیرد، اساساً داده‌ای درباره آن تولید یا گردآوری نخواهد شد. اینکه چه افراد، نهادها و ارگان‌هایی مشکلات را تعریف و ابلاغ می‌کنند، چه گروه‌هایی مشکلات را تعریف و وارد گفتمان عمومی می‌کنند، چه گروه‌هایی برخی مشکلات را پنهان و برخی مشکلات دیگر را در دستور جلسات وارد می‌کنند، به‌خودی‌خود یک کنش هستی‌شناختی در گردآوری و خلق داده است (Rogers 2013). ج) قالب‌بندی مقیاس: تعریف اینکه چه چیز قابل اندازه‌گیری است و چه چیز نیست، چه چیز نامربوط است و چه چیز به‌خاطر پیچیدگی بالا واجد اهمیت در گردآوری داده نیست. اگر چیزی غیرقابل اندازه‌گیری، نامربوط و یا بسیار پیچیده تلقی شود (مثلاً عشق، خرد، حکمت، عقلانیت و امثال آن)، از دایره هستی‌داده‌های کمی خارج می‌گردد. دوم اینکه در هر یک از شاخص‌ها، چه ابعادی کیفی است و تبدیل آن به شاخص کمی زمان‌بر یا هزینه‌بر است و باید کنار گذاشته شود که عامل ایجاد خلأ داده‌ای است (Golebiewski and Boyd 2019). مثلاً در موضوع آموزش، تکیه بر ابعادی که به‌آسانی قابل سنجش است، باعث غفلت از جنبه‌های پیچیده و کیفی آموزشی می‌شود و خلأ داده‌ای، جهان آموزش را معوج نشان می‌دهد؛ همان‌گونه که داده‌های علم‌سنجی امروزه ارزش علمی مقالات را معوج نشان می‌دهد.

کلیدی‌ترین نکته در معرفت‌شناسی داده‌ها آن است که داده گرانبار از نظریه است. این نکته از چند جنبه قابل بررسی است (Frické 2015): الف) چیزی به نام «داده خام» وجود ندارد و «اصطلاح داده خام افسانه‌ای بیش نیست» (Rosenberg 2013). داده‌ها به‌طور مشخص یا تلویحی با یک عدسی نظری گردآوری یا خلق می‌شوند و نظریه‌ها، پیش‌فرض‌ها و فرضیه‌ها هدایت می‌کند که دنبال چه بگردیم، چطور به آن

بنگریم، آن را اندازه‌گیری کنیم، و آن را معنادار تلقی کنیم. مثلاً درآمد زمانی متغیر معنادار برای رفاه اجتماعی است که یک جامعه مبتنی بر مناسبات کارمزدی پیش‌فرض باشد و در اقوام، اجتماعات، و گروه‌های خاصی از مردم کره زمین که چنین مناسباتی وجود ندارد، رفاه اجتماعی را نمی‌توان با درآمد سنجید. ب) طبقه‌بندی‌های پیشینی و پیش‌فرض انگاشته‌شده، بستر قرارگیری داده‌هاست و گردآوری و خلق داده اساساً با این طبقه‌بندی‌های پنهان و پیشینی معنا پیدا می‌کند. طبقه‌بندی‌های گیاهان، طبقه‌بندی امراض پزشکی، طبقه‌بندی‌های جمعیت‌شناختی و امثال آن تعیین می‌کنند که چه داده‌ای باید گردآوری یا خلق شود و پس از گردآوری و خلق، کجا بنشیند (Bowker and Star 2000). این طبقه‌بندی‌ها بر پایه نظریات پزشکی، زیست‌شناختی، انسان‌شناختی و غیره ایجاد شده‌اند و ذاتی نیستند. ج) عملیاتی‌سازی شاخص‌ها (تعریف عملیاتی متغیرها) عملی نظری است که مفاهیمی مثل عدالت، خوش‌حالی، و هوش را اندازه‌پذیر می‌کند. روش عملیاتی‌سازی این مفاهیم تعیین می‌کند که چطور آن‌ها را فهم کنیم و نحوه تعریف عملیاتی آن‌ها وابسته به نظریه است و تعیین می‌کند که واقعیت چیست (Hartlapp & Falkner 2009). د) موضوع جبر روش‌شناختی و اینکه داده با چه روشی (خودارزیابی، مشاهده، نگرش‌سنجی، ردپای دیجیتال، و غیره) گردآوری شود، بر دانش حاصل از آن اثر دارد. برای مثال، در سنجش سوگیری ضمنی به روش «آی‌آی‌تی»^۱ پاسخ درست یا غلط مطرح نیست، بلکه سرعت کلیک بر روی گزینه‌ها سنجیده می‌شود. بنابراین، عده‌ای بر این باورند که حقیقت، کارکرد روش است (Rogers 2013). ه) واسطه‌های ابزاری و اینکه داده با چه ابزاری و چگونه گردآوری شود (حسگر، الگوریتم، رابط کاربر، برگه کاغذ، و غیره) بر اینکه چه داده‌ای گردآوری شود، اثر دارد (Wiggins and Jones 2023). و) نگاه خیره‌گردآورنده یا خالق داده‌ها که در گردآوری کمی دال‌بر تأکیدات گردآورنده است. محقق کمی در پرسشنامه رتوریک خاصی به کار می‌برد که خواه‌ناخواه بر چیز خاصی تأکید بیشتری دارد و این تأکید با هجا، لحن و مدالیت کلمات، تعداد سؤال‌هایی که به یک بُعد خاص از متغیر تخصیص می‌یابد، و با طول جملات بازنمایی می‌شود (ibid) (ز) زمان، مکان، فضای اجتماعی و رویدادهایی که هنگام گردآوری داده مردم را تحت تأثیر قرار داده‌اند بر داده‌ها اثر اجتناب‌ناپذیر دارد (O'Neil 2016). ح) موضوع آگاهی و

1. implicit association test (IAT)

اینکه فرد آگاهانه داده‌ای ارائه دهد (مثلاً پرسشنامه‌ای را پُر کند) یا ناآگاهانه از او داده گردآوری شود یا فرد، آگاهی ضمنی به ابزار گردآوری داده داشته باشد یا نداشته باشد و پژوهشگر با حضور میدانی‌اش مداخله‌ای در آگاهی پاسخ‌دهنده ایجاد کرده باشد یا خیر، بر داده‌های گردآوری‌شده اثر دارد (Gamez 2016). ط) مواضع نظری پاسخ‌دهنده، تعلق پارادایمی او و شبکه مفاهیم ذهنی‌اش در بازخوردی که می‌دهد و داده‌ای که بر جای می‌گذارد، مؤثر است. هر فردی یک «قاب معرفتی» است؛ به‌طوری که هر اثری که بر جای می‌گذارد، از آن قاب عبور می‌کند (Lupton 2016). ظ) برخی داده‌ها (رسانه‌ای، خبری، و امثال آن) حاصل تعامل و مذاکره‌اند و از این‌رو، برخی معتقدند که داده دارای هویت مذاکره‌ای است (Wiggins and Jones 2023). لایک‌ها، کامنت‌ها، و اشتراک‌گذاری فایل‌ها و مطالب برساخت اجتماعی‌اند و حقایق عینی نیستند، بلکه از کنش متقابل انسان‌ها در موقعیت ناشی می‌شوند.

موضوع قدرت، اقتدار و صلاحیت مهم‌ترین بحث در برساخت اجتماعی داده‌هاست که خود دارای چند بُعد است (Espeland and Stevens 2008): الف) چه افراد، نهادها، سازمان‌ها یا الگوریتم‌هایی قدرت و صلاحیت این را پیدا کرده‌اند که شکل گردآوری داده را تعریف و گردآوری و تولید داده کنند. زاویه دیدها، ارزش‌ها و علایق عوامل نامبرده به درون داده نفوذ می‌کند (ibid). برای مثال، اینکه شاخص امنیت غذایی با چه طیفی (پیوسته یا گسسته)، با چه مقیاس‌ها و خرده‌مقیاس‌هایی و در چه ابعادی سنجیده شود و چه نمره‌ای کمینه مطلوب در این شاخص تلقی گردد، وابسته به افراد، گروه‌ها و نهادهایی است که اقتدار و صلاحیت گردآوری و خلق داده را پیدا کرده‌اند. ب) در تعیین شاخص‌ها، بی‌عدالتی شناختی رخ می‌دهد و دیدگاه اقلیت‌ها در چستی مفاهیمی همچون عدالت، سلامت، و اخلاق و امثال آن دیده نمی‌شود (Benjamin 2019). ج) مسئله جعبه سیاه گردآوری داده و اینکه وقتی الگوریتمی خودیادگیرنده به گردآوری خودکار داده می‌پردازد، ارزیابی فرایند گردآوری داده برای همگان حتی تولیدکنندگان الگوریتم ناممکن است (Diakopoulos 2014). داده‌ها توسط بات‌ها گردآوری می‌شود، ولی منطق و پیشفرض‌های گنجانده‌شده در الگوریتم و شکل خودیادگیری بات برای کسی آشکار نیست و بنابراین، تأیید روایی معرفت‌شناختی داده‌ها ناممکن است.

۳. مالکیت داده

در باره مالکیت داده‌های موجود در وبگاه‌ها و پایگاه‌های اطلاعاتی شرکت‌ها یا سازمان‌ها، چهار دیدگاه وجود دارد (Mittelstadt & Floridi 2016): (۱) استفاده-دسترسی، (۲) استفاده-عدم دسترسی، (۳) کنترل-دسترسی، و (۴) کنترل-عدم دسترسی. مالکیت در معنای «استفاده-دسترسی» یعنی فرد بتواند از خروجی تحلیل داده‌های شرکت یا سازمان رایگان استفاده کند و نیز هرگاه اراده کرد بتواند خود یا توسط غیر تحلیل کند و پایگاه‌ها مکلف به ارائه اصل داده‌ها و تحلیل‌ها به افراد هستند. مالکیت در معنای «استفاده-عدم دسترسی» یعنی «امکان استفاده» صرفاً شامل استفاده از خروجی تحلیل‌هاست؛ زیرا از نظر طرفداران این رویکرد، خود داده‌ها (مثل داده‌های بیولوژیک) قابل تفسیر توسط عموم افراد نیست. مالکیت در معنای «کنترل-دسترسی» یعنی فرد افزون بر اینکه باید بتواند درباره انواع استفاده‌های دیگران (انواع تحلیل‌های ثانویه، انواع الگوییابی، فروش و تبادل) تصمیم بگیرد، در صورت تشخیص به اصل داده‌های خود یا نتایج تحلیل آن‌ها هم دسترسی داشته باشد. و مالکیت از نوع «کنترل-عدم دسترسی» نیز به این معناست که فرد دسترسی به اصل داده‌ها یا دسترسی رایگان به خروجی تحلیل داده‌ها نخواهد داشت و فقط می‌تواند درباره استفاده از داده‌ها در تحلیل‌های ثانویه، الگوییابی‌ها، تبادل و فروش تصمیم بگیرد. به چند دلیل، بسیاری از پژوهشگران از گزینه کنترل-عدم دسترسی حمایت می‌کنند (Watson, Kay & Smith 2010): الف) فرد منبع داده است، نه خالق داده، ب) خلق داده نیازمند ابزار (بستر نرم‌افزاری، وسایل آزمایشگاهی) و دانش (آزمایش، استخراج، مقوله‌بندی (فرا داده‌ها)، تحلیل) است که شامل مالکیت معنوی می‌شود، ج) بدون وجود زمینه خلق داده، خلق رخ نمی‌دهد، د) حرکت نخست برای خلق داده، خالق داده بوده نه منبع داده، ه) منبع داده حق دارد آسیب ندیدن از داده‌هایی را که او منبعش بوده، کنترل کند (حق کنترل آسیب بر خود)، و) خلق داده در بدو امر برای رفع نیاز منبع داده (در مورد داده‌های آزمایشگاهی) اتفاق افتاده و داوطلبانه نبوده، و زمانی که داوطلبانه بوده، شرط استفاده شخصی توسط منبع داده ذکر نشده است، و ز) هر خلق داده از منبع داده با پردازش اولیه و تغییر فرمت همراه است که آن را از حالت خام درمی‌آورد و ایجاد مالکیت می‌کند. بنابراین، هنگامی که کاربر قرار است داوطلبانه نقش منبع داده را ایفا کند (مثل مصاحبه‌شونده‌ها یا داوطلبان اهدای تخمک یا اسپرم به پایگاه داده ژنتیک)، خالق داده (پژوهشگر) موظف است شروط استفاده از داده‌ها را بپرسد و در آن باره تصمیم‌گیری کند (McDonough & Brandenburg

2019). اگر «شرط استفاده شخصی از داده‌ها» توسط فرد قید شود و پژوهشگر یا مرکز بپذیرد، آنگاه مالکیت از نوع کنترل-دسترسی خواهد بود، وگرنه از نوع کنترل-عدم دسترسی است، و اگر که داده‌ها برای رفع نیاز کاربر گردآوری شده باشد، مالکیت کاربر از نوع مالکیت-عدم دسترسی است.

۴. داده‌های باز

مفهوم شفافیت در عصر تفوق حقوق بر تکالیف، در حال تبدیل شدن به گفتمان است. ادعای برخی کنشگران اجتماعی، سیاست‌مداران، اندیشمندان و رهبران احزاب آن است که بسیاری از دعاوی سیاسی و اجتماعی که به‌عنوان شاهدی جهت توجیه اتخاذ برخی تصمیمات ارائه می‌شود، ناشی از تحلیل نادرست داده‌هاست. آن‌ها مدعی‌اند که اگر داده‌ها در معرض استفاده و تحلیل جمع‌گسترده‌تری از گروه‌ها قرار گیرد، شاهد تصمیمات سیاسی بهتری خواهیم بود (Levy & Johns 2016). در جهان علم نیز، تکرار برخی آزمایش‌های مطرح‌شده در مقالات علمی به معضل بزرگی تبدیل شده است. یک شیمیدان در مقاله‌ای علمی از ساخت سلول یا ترکیب شیمیایی جدیدی سخن می‌گوید که افراد دیگر در تکرار آن ناکام می‌مانند. نویسنده مقاله به عمد اطلاعات گمراه‌کننده‌ای در مقاله می‌آورد که فرد دیگر نتواند محصول جدید را بازتولید کند و یا اساساً آزمایش وی مبتنی بر داده‌های واقعی نیست یا اصلاً آزمایشی در کار نیست. امروزه، شرکت‌ها از جمله شرکت‌های بیمه و خدمات، تصمیماتی را علیه برخی گروه‌های آسیب‌پذیر (سالمندان، بیماران، افراد کم‌سواد و امثال آن) می‌گیرند که بنا به ادعای آن‌ها متکی به داده‌هاست، ولی خود کلان‌داده‌ها برای آزمایش اعتبار تصمیمات در دسترس نیست (Gurstein 2011). بنابراین، امروزه مطالبه شفافیت تا حدودی به مطالبه داده‌های واقعی متمرکز شده است. مطالبه داده‌های باز، نوعی مطالبه برای رفع شکاف اطلاعاتی بین افراد یا حتی کشورها نیز هست. گرایش معاصر به شعار «ذات اطلاعات آن است که به اشتراک گذاشته شود» مبتنی بر ریشه کلمه اطلاعات (in-form-ation) به معنای در فرم کردن یا در قالب کردن (شکل‌دهی دیگران) است که نمی‌تواند بدون اشتراک اطلاعات ممکن شود (Capurro 2009). داده‌ها که منشأ ظهور اطلاعات‌اند نیز از این قاعده مستثنا نیستند؛ زیرا اطلاعات حاصل دسته‌بندی داده‌هاست. داده از نظر برخی منتقدان، هستی‌مند طبیعی است و بنابراین، نه جنبه خلقی، که جنبه کشفی دارد. حتی با پذیرش الگوی کنترل-عدم دسترسی نیز باید

اذعان کرد که افراد از گردآوری نامحسوس مطلع نمی‌شوند که بخواهند آن را کنترل کنند. در جهانی که داده‌گرایان معتقد به پایان نظریه شده‌اند و مدعی‌اند که امروزه نظریه‌ها از کلان‌داده‌ها قابل استخراج است و نیازی به نظریه‌پردازی پیشینی نیست، مطالبه داده‌های باز راهی است برای تعدیل سیاست‌ها و تصمیماتی که در مقیاس کلان برای انسان‌ها گرفته می‌شود. البته باید در نظر داشت که این گفتمان ممکن است سلاحی علیه جوامع دیگر باشد؛ یعنی مطالبه داده برای اعمال قدرت و تغییر در مناسبات سیاسی؛ و هشدارهای (Iliadis & Russo 2016) و (Levy & Johns 2016) از این‌رو است. در دسترس بودن داده برای همگان به معنای ایجاد عدالت نیست، زیرا نفس قابل دسترس بودن تفاوتی ایجاد نمی‌کند، بلکه توان استفاده از داده است که ایجاد تفاوت می‌کند. به همین دلیل، تکنوکرات‌ها و نخبگان فناوری هستند که از داده‌های باز سودهای کلانی به دست خواهند آورد و عامه مردم از این سیاست سودی نخواهند برد. اگر داده باز ولی دانش استفاده از آن انحصاری باشد یا به نحو پدیدارشناسانه به قول «گوروویچ» (۱۳۷۱) طوری فوت‌وفنی شود که زبان استفاده از آن از زبان عامه فاصله چشمگیری بیابد، آنگاه باز بودن داده جز برای گروه‌های خاص فایده‌ای نخواهد داشت. این است که مطالبه داده از مشتریان که مبنای تصمیم‌گیری نهادها برای افزایش یا کاهش تولید، افزایش یا کاهش ورودی‌ها و خروجی‌ها، و واردات و صادرات داخلی یا خارجی است، نیاز به شفافیت داده‌ای دارد و این شفافیت جز از طریق نظارت مردم یا نمایندگان مشروع آن‌ها بر فرایند گردآوری داده‌ها و بازتفسیر آن‌ها ممکن نیست. البته، برخی محققان اظهار کرده‌اند که معنای باز بودن داده آن نیست که در اختیار همگان قرار گیرد، بلکه بدین معناست که گردآوری با فرایندی باز، شفاف و قابل ردگیری رخ دهد. در پژوهش‌ها، این به معنای ذکر گام‌به‌گام مراحل و ابزارهای گردآوری داده، روش‌های دسته‌بندی آن‌ها، نحوه برخورد با داده‌های پرت، نحوه برخورد با داده‌های گمشده و فرایندهای تبدیل است. برخی از منتقدان نیز گفتمان داده‌های باز را واجد نوعی ساختار اسطوره‌ای می‌دانند؛ زیرا قبول داده‌های باز به معنای حکمرانی باز است (Janssen, Charalabidis & Zuiderwijk 2012). همچنین، سیاست داده‌های باز راه را برای ترویج اطلاعات کاذب عمدی و غیرعمدی^۱ باز می‌کند؛ زیرا می‌توان با قدرت بیشتری به دستکاری محیط‌ها و انسان‌ها، قبول‌اندن باورها، تنفیذ سیاست‌ها و تصمیم‌ها و

1. disinformation and misinformation

افزایش قدرت و مشروعیت اقدام کرد (Johnson 2014) و از کنترل رسانه‌ای جهان کم‌کم به سمت کنترل داده‌ای آن حرکت نمود.

۵. کیفیت داده

کیفیت داده که شاید مهم‌ترین بُعد در بین ابعاد شش‌گانه این مطالعه باشد، برای سازمان‌ها و شرکت‌هایی که از منبع مشترکی داده دریافت می‌کنند یا از داده‌های مشترکی استفاده می‌کنند، ولی هر یک به نحو خاصی به تحلیل می‌پردازند، دغدغه مهمی است. در مراحل مختلف تحلیل داده نیز افراد مختلفی روی آن کار می‌کنند و به همین دلیل داده‌ها ممکن است گاهی دچار تغییراتی شود که نتایج را دچار سوگیری نماید. مسئله گردآوری و خلق داده که قبلاً ذکر شد نیز به اهمیت کیفیت داده می‌افزاید. بنابراین، معیارهای مختلفی برای ارزیابی کیفیت داده در نظر گرفته شده است. «سازمان بهداشت جهانی (دبلیو‌اچ‌ا)» معیارهای خوبی در چهار بُعد ارائه داده که سازمان‌های بین‌المللی دیگر مثل سازمان همکاری و توسعه اقتصادی نیز از آن اقتباس کرده‌اند: (۱) کامل بودن، (۲) به‌هنگام بودن، (۳) سازگاری درونی، و (۴) سازگاری بیرونی با سایر مجموعه داده‌ها (World Health Organization 2019). کامل بودن داده‌ها مربوط به چند چیز است: اول اینکه داده‌ها تا چند رقم اعشار دقت دارند و تحلیل‌ها چقدر تحت تأثیر اندازه دقت داده‌ها قرار می‌گیرد. دوم اینکه چقدر داده گمشده وجود دارد. اگر میزان داده‌های گمشده بیش از ده درصد باشد، کامل بودن داده‌ها دچار تردید است. اگر مجموعه داده‌ای فاقد داده گمشده باشد نیز نشان‌دهنده احتمال جعل در گردآوری داده‌هاست؛ زیرا چنین نیست که بتوان داده را به‌طور کامل گردآوری کرد (حتی با سیستم اطلاعاتی) و همواره مقداری داده گمشده وجود دارد. نکته دیگر اینکه آیا داده‌های گمشده از الگوی خاصی تبعیت می‌کند یا نه. اگر داده‌های گمشده در یک متغیر یا یک مقوله به‌طور معناداری بیش از مقولات دیگر باشد، نشان‌دهنده الگودار بودن داده‌های گمشده است و این، خدشه‌ای بر مفهوم کامل بودن به حساب می‌آید. از نظر مقیاس نیز هر چه داده‌های گردآوری شده بیشتر در مقیاس نسبتی باشند و کمتر در مقیاس اسمی یا رتبه‌ای، کیفیت داده‌ها بهتر است؛ زیرا داده‌های مرتبه بالا را می‌توان به داده‌های مرتبه پایین تقلیل داد، ولی برعکسش ممکن نیست. از زاویه دید معیار به‌هنگام بودن، در هر مجموعه داده باید متغیری به نام زمان (در ستون) وجود داشته باشد که زمان گردآوری هر ردیف داده را تعیین کند. نبود چنین مقوله‌ای، هم نقصانی در کامل بودن است و هم

نقصانی در به‌هنگام بودن. برای انواع مختلف داده‌ها، معیارهای معینی برای به‌هنگام بودن وجود دارد. برای مثال، استاندارد به‌هنگام بودن داده‌های قند خون افراد ممکن است ۶ ماه، استاندارد به‌هنگام بودن داده‌های پاپ‌اسمیر زنان ۲ سال و استاندارد به‌هنگام بودن اطلاعات ژنتیک، ۵ سال باشد. همچنین، استاندارد به‌هنگام بودن داده‌های سکونت افراد ۶ ماه، داده‌های تغییر شغل ۱ سال و داده‌های عدالت غذایی، ۳ سال باشد. تحلیل‌هایی که از داده‌های نابهنگام حاصل شود، وجاهت خاصی برای وضع موجود نخواهد داشت و افراد تحت تأثیر تصمیمات حاصل را دچار نوسانات ناعادلانه‌ای خواهد نمود. برای مثال، فرض کنید که ارائه‌ی یارانه به افراد بر اساس داده‌های یکسال گذشته از دارایی‌ها و شغل افراد باشد (خانه، اتومبیل، و شغل (درآمد) به‌عنوان سه متغیر اصلی). اگر برای متغیر به‌هنگام بودن، از استاندارد درستی تبعیت نشود، آنگاه بسیاری از افراد ممکن است تحت تأثیر تصمیماتی قرار گیرند که مناسب آن‌ها نیست. معیار سوم از نظر (World Health Organization 2019)، سازگاری درونی است. منظور از سازگاری درونی آن است که تغییرات و نوسانات موجود در داده‌ها (پراکندگی داده‌ها) از الگوی استاندارد تبعیت کنند. همچنین، انتظار می‌رود که داده‌های مربوط به متغیرهای طبیعی (مثل سن، طول عمر، یا ساعات خواب) از توزیع نرمال تبعیت کند و متغیرهای غیرطبیعی (مثل درآمد) خیر. معیار داده‌های پرت در مقوله‌ی انسجام درونی یعنی این داده‌ها در متغیرهای مختلف تابع استاندارد خاصی باشد و از حد معینی بیشتر نباشد. نبود داده‌ی پرت نیز خود، نشان‌دهنده‌ی نبود انسجام درونی است؛ زیرا در مجموعه داده‌هایی که در مقیاس کلان و به‌درستی گردآوری شده باشند، به احتمال زیاد دارای داده‌های پرت خواهیم بود. سازگاری درونی باید در طول زمان نیز تأیید شود؛ یعنی اگر مجموعه داده‌ها در نقطه‌ی زمانی t دارای n داده‌ی پرت بود، در نقطه‌ی زمانی $t+1$ نیز حدوداً همان مقدار داده‌ی پرت در مجموعه داده‌ها حضور داشته باشد. اندازه‌های گرایش به مرکز و پراکندگی نیز در مجموعه داده‌ها برای متغیرهای یکسان نباید در طول زمان نوسان زیادی نشان دهد. یکی از معانی انسجام درونی آن است که مقادیر داده‌ها به یک میزان افزایش یا کاهش از خود نشان دهند. برای مثال، آزمایش فشار و قند خون به‌طور معمول، در دستگاه واحد و یکجا گرفته می‌شود. اگر تعداد داده‌های قند خون افزایش، و داده‌های فشار خون کاهش از خود نشان دهد، نشان‌دهنده‌ی نبود انسجام درونی است. برای مثال، همیشه انتظار می‌رود که داده‌های متغیرهای سن و فشار خون با هم افزایش پیدا کنند؛ یعنی با افزایش سن انتظار می‌رود که فشار خون نیز بیشتر شود. تحلیل‌گر باید دقت

کند که این دو متغیر بایستی همبستگی مثبتی را با هم نشان دهند. ولی اگر این همبستگی منفی شد، نشانه وجود نوعی عدم انسجام درونی است که نیاز به بررسی دقیق‌تر را ایجاب می‌کند. معیار آخر، سازگاری بیرونی با سایر مجموعه داده‌هاست. داده‌های نمونه‌گیری‌های متعدد موجود در مجموعه بزرگ کلان داده بایستی با داده‌های دیگری که به‌طور مستقیم از جامعه برگرفته شده، سازگار باشد و متغیرها، اندازه‌های گرایش به مرکز و پراکنندگی تقریباً یکسانی را از خود نمایش دهند. فرض کنید داده‌های حاصل از نمونه‌گیری‌های متعدد از متغیر مصرف گوشت در مطالعه‌ای مربوط به عدالت غذایی در شهر تهران، نشان‌دهنده میانگین مصرف ۱/۵ کیلوگرم در ماه باشد، ولی داده‌های حاصل از ترازوهای فروشگاه‌های گوشت، نشان‌دهنده مصرف ۷۵۰ گرم در ماه است. در این صورت، هر دو مجموعه داده بحث‌برانگیز و غیرقابل اعتماد خواهند بود.

«پیبی‌نو، لی و وانگ» نیز معیارهای جامعی در ۱۶ بُعد درباب کیفیت داده ارائه کرده‌اند: (۱) دسترس‌پذیری، (۲) حجم مناسب داده‌ها، (۳) باورپذیری، (۴) کامل بودن، (۵) بازنمایی فشرده (یا خلاصه)، (۶) بازنمایی سازگار (یا منسجم)، (۷) سهولت دستکاری، (۸) عاری از خطا بودن، (۹) قابل تفسیر بودن، (۱۰) عینیت، (۱۱) ربط، (۱۲) حسن ظن، (۱۳) امنیت، (۱۴) به‌هنگام بودن، (۱۵) قابل فهم بودن، و (۱۶) ارزش افزوده داشتن (Pipino, Lee and Wang 2002). در دسترس بودن یعنی داده‌ها تا چه اندازه موجود هستند و آیا هرگاه بخواهیم، در اختیار ما هست یا اینکه اگر به آسانی در اختیار ما نیست، چقدر به راحتی قابل استخراج است. گاهی داده در سیستم‌هایی ذخیره شده است که دستیابی به آن نیازمند ابزارها و نرم‌افزارهای دیگری است که خود آن‌ها نیاز به روزآمدسازی دارند و بنابراین از دسترسی و استفاده خارج‌اند. هرچه واسطه‌های استخراج و استفاده از داده بیشتر باشد، کیفیت آن پایین‌تر است. ممکن است برای دسترسی به داده فقط به نام کاربری و کلمه عبور نیاز باشد یا اینکه نیاز باشد تا قفل نرم‌افزاری را فعال کنیم، نرم‌افزار خاصی را نصب کنیم، نام کاربری و کلمه عبور چند سیستم مختلف را بدانیم، و پس از استخراج داده‌ها به نرم‌افزار دیگری برای تغییر فرمت یا یکپارچه‌سازی نیاز داشته باشیم. این است که کیفیت داده‌های مورد اول، بالا و مورد دوم، پایین است.

منظور از بُعد «حجم مناسب داده‌ها» در بحث کیفیت آن است که اندازه داده‌ها باید تا حد ممکن به اندازه جامعه‌ای که قرار است برایش تصمیم‌گیری شود، نزدیک باشد. برای مثال، اگر داده‌های درآمد ۳ میلیون تهرانی برای جمعیت ۱۲ میلیونی در دسترس

باشد، آنگاه این مقدار داده دارای کیفیتی پایین‌تر از فایل داده‌هایی است که شامل ۶ میلیون مورد است. حجم مناسب داده‌ها نیز اشاره دارد به داشتن داده‌هایی از یک موضوع در تاریخ‌های مختلف برای انجام مقایسه. برای مثال، اگر داده‌های درآمدی تهرانی‌ها ۶ میلیون مورد در سه سال اخیر باشد، بهتر از ۶ میلیون مورد برای یکسال است.

باورپذیری اشاره دارد به صحت و اعتبار داده‌ها. از نظر «اسپانگ هانسن» اعتبار داده‌ها (و اطلاعات) را نمی‌توان بدون مفهوم سند درک و یا اثبات کرد (Spang-Hanssen 2001). فرض کنید فایل اکسلی از اینترنت دانلود کرده‌اید که نام و نشانی ندارد و معلوم نیست مربوط به کدام سازمان یا شرکت است. این داده‌ها اگر درست نیز گردآوری شده باشند، معتبر تلقی نمی‌شود؛ زیرا مستند به هیچ سندی که نشان‌دهنده منشأ آن‌ها باشد، نیست. حال اگر همان فایل داده‌ها از سایت سازمانی دانلود شود، نسبت به داده‌هایی که از وبلاگ با عنوان «مربوط به سازمان X» دانلود شده باشد، معتبرتر است. بنابراین، اعتبار داده به مفهوم سند و سندیت وابسته است و هرچه سندیت یک مجموعه داده به منشأ آن نزدیک‌تر و ربط آن به منشأ با مدارک بیشتری قابل اثبات باشد، همان‌قدر معتبرتر است. این اعتبار می‌تواند از طریق شهادت افراد واجد صلاحیت، از طریق تناسب داده‌ها با وظایف و فعالیت‌ها، و استفاده‌های قبلی از مجموعه داده‌ها و تناسب داده‌های جدید با داده‌های قبلی اثبات شود.

پس از بُعد «کامل بودن» (که قبلاً به تفصیل بحث شد)، بُعد کیفیت است که از نظر Pipino, Lee and Wang (2002)، دال بر بازنمایی فشرده بوده و اینکه چه اندازه فشرده و خلاصه گردآوری شده است. فرض کنید در یک صفحه گسترده، سن زنان و مردان جدا آمده و چون باید یک متغیر سن داشته باشیم، مجبوریم آن دو را ترکیب کنیم. نیاز به دستکاری و احتمال بروز خطا یا سوگیری در ترکیب باعث کاسته شدن از کیفیت آن می‌شود. چون ترکیب داده‌ها و ساخت متغیرها یا مقولات جدید ممکن است باعث بروز خطا یا ایجاد سوگیری گردد، بنابراین، داده‌های باکیفیت از همان ابتدا فشرده و خلاصه هستند و ذیل مقولات درستی جمع آمده‌اند.

بُعد ششم یعنی «بازنمایی سازگار» یعنی داده باکیفیت آنی است که در فرمت یا قالب ارائه شود. قالب را به دو صورت می‌توان فهم کرد: قالبی که داده‌ها در آن ذخیره شده‌اند، و قالبی که داده‌ها بر اساس آن نگاهشته شده‌اند. فرض کنید داده‌هایی گردآوری کرده‌اید که یکی در قالب xlsx، دیگری در قالب csv و دیگری در قالب sav یا xml است. این

داده‌ها باکیفیت نیست؛ زیرا هر یک در یک قالب خاصی عرضه شده است. نیز فرض کنید که مجموعه داده، ترکیبی است از داده‌های عددی، متنی، متنی-عددی، تاریخ و ارز. این داده‌ها کیفیت مناسبی ندارند؛ زیرا بازنمایی یکسانی از نوع نگارش داده حاصل نشده و برای یکدست‌سازی و ایجاد بازنمایی مناسب و مقوله‌بندی، تغییرات زیادی نیاز است که خود ممکن است باعث بروز خطا گردد.

بعد هفتم یعنی «سهولت دستکاری» یعنی چقدر داده‌ها قابل دستکاری‌اند که واجد دو جنبه است: الف) اینکه داده‌ها از نوع داده‌های مرتبه بالا باشند و بتوان آن‌ها را به داده‌های مرتبه پایین تبدیل کرد؛ و ب) اینکه داده‌ها دارای جنبه‌های همپوشان باشند و بتوان بر اساس آن‌ها مقولات جدیدی تعریف کرد. درباره مورد اول پیشتر ذکر گردید که داده‌های نسبتی باکیفیت‌اند؛ زیرا هم پیوسته‌اند و هم قابل تقلیل به داده‌های اسمی و رتبه‌ای. درباره مورد دوم، در نظر بگیرید که فایلی دارای داده‌هایی راجع به نوع اتومبیل افراد، میزان درآمد ماهیانه، و متراژ و محل مسکن آن‌هاست. با تجمیع داده‌های این چهار متغیر می‌توان متغیر جدیدی به نام «سرمایه اقتصادی» را شکل داد. هر چه یک مجموعه داده، دارای مقولات همگرایی بیشتری باشد، قدرت دستکاری بیشتر و بنابراین، کیفیت بیشتر است. بعد هشتم کیفیت داده‌ها از نظر Pipino, Lee and Wang (2002) یعنی «عاری از خطا بودن» اشاره به «دقت» دارد و با «وثاقت» یا «اعتبار» متفاوت است. دقت یعنی داده‌ها تا چه حد دقیق گردآوری شده است؛ بدین معنا که داده‌ها تا چند رقم اعشار دقت دارند و اینکه ابزارهای دقیقی برای گردآوری داده استفاده شده باشد و افزون بر آن، اگر با ابزار مشابهی همان داده‌ها گردآوری شود، داده‌های مشابهی جمع شود و نتایج حاصل از گردآوری تکرارپذیر باشد.

بعد نهم یعنی «تفسیرپذیری» به شفافیت تعاریف مقولات و متغیرها و ربط بین داده‌ها و تعاریف مقولات معطوف است؛ بدین معنا که تعریف عملیاتی دقیقی انجام شده باشد تا معلوم شود چه نوع داده‌ای دال بر چه مفهومی است. مثلاً اگر داده‌های حضور در محل کار را دال بر فعالیت کاری بگیریم، در تعریف خود دقت کافی به خرج نداده‌ایم.

بعد دهم یعنی «عینیت» از دو جنبه قابل بررسی است: اول، از نظر مفهوم سطح انتزاع

1. numerical

2. string

و دوم، از نظر مفهوم سوگیری و به‌ویژه «سوگیری ضمنی»^۱، اول اینکه داده عینی داده‌ای است که ذیل متغیرها یا مقولاتی با حداقل سطح انتزاع گرد آمده باشد؛ یعنی مقوله یا متغیری که داده ذیل آن گرد آمده، به‌طور مستقیم، سنجش‌پذیر باشد و به یک شیء بیرونی ملموس مثل زمان و مکان و اندازه و قیمت و امثال آن مربوط باشد. دوم اینکه داده عینی داده‌ای است که از سیستم یا ابزاری با حداقل سوگیری ضمنی استخراج شده باشد. نرم‌افزاری را در نظر بگیرید که در آن، ساعت کاری کارکنان بر اساس ساعت خروج منهای ساعت ورود ثبت می‌شود و حاصل ذیل مقوله‌ای با همین نام ثبت می‌گردد. چنین سیستمی از اساس واجد سوگیری ضمنی است؛ زیرا سطح مرجع یا پایه را در نظر نمی‌گیرد. اگر فردی ساعت ۱۱ وارد محل کار شده و ساعت ۱۴ خارج شود، سه ساعت کار برای وی ثبت می‌گردد، و فردی دیگر، ساعت ۹ صبح وارد شود و ساعت ۱۲ خارج گردد نیز سه ساعت کار برای او ثبت می‌شود. از سه ساعت کار اول، یک ساعت استراحت ظهر است، ولی ساعات ۹ الی ۱۲ ظهر ساعات پرتراфик اجرایی است.

«ربط» به‌عنوان بُعد یازدهم یعنی داده با کار و تصمیم هدف، ربط و تناسب داشته باشند. از داده‌های به‌کارگیری نرم‌افزارها، سامانه‌ها و خودپردازهای بانکی نمی‌توان رضایت افراد از خدمات بانک را نتیجه گرفت و از داده‌های حقوق و مزایای افراد نمی‌توان «رضایت از کار» را استنتاج کرد. در رضایت شغلی «فاصله تا محل کار»، روابط با فرادستان و فرودستان، ربط بین رشته تحصیلی و شغل، و میزان ریسک فعالیت فرد نیز مؤثر است. «حسن ظن» که بُعد دوازدهم است، به شهادت خبرگان و حسن ظن آن‌ها نسبت به منبع گردآوری داده‌ها (سیستمی که داده‌ها از آن استخراج شده یا ابزاری که داده‌ها با آن گردآوری شده) برمی‌گردد. «حسن ظن» یعنی خبرگان قبول داشته باشند که سازمان، افراد کارآمدی برای گردآوری داده دارد که به صحت داده‌ها حساسیت دارند و خود سازمان نیز در توجه به دقت و صحت داده‌ها دارای حسن شهرت باشد. این بُعد به شهرت نیک نرم‌افزارها، سیستم‌ها و سامانه‌هایی که داده‌ها از آن استخراج می‌شود نیز دلالت دارد. بُعد سیزدهم یعنی «امنیت» از منظر Pipino, Lee and Wang (2002) یعنی هرچه سطوح دسترسی پیچیده‌تر، تعداد افراد دارای امکان ورود به سیستم و استخراج داده کمتر و امنیت سیستم در برابر ورود و دستکاری داده‌ها بیشتر باشد، کیفیت داده‌ها بیشتر است.

1. implicit bias

از بُعد چهاردهم یعنی به‌هنگام بودن که بگذریم (قبلاً بحث شد)، بُعد پانزدهم یعنی «قابل فهم بودن» دال بر آن است که آیا داده‌ها در قالب اعداد یا کلمات قابل فهم وارد شده یا در قالب علائم و رموزی درج گردیده که برای همه قابل فهم نیست. گاهی برای جلوگیری از سوء استفاده از علائم و رموز خاصی برای ثبت داده استفاده می‌شود، ولی این کار اساساً از کیفیت داده‌ها می‌کاهد، چون: الف) علائم، قراردادی‌اند و ممکن است فراموش شوند و در چند سال آتی نتوان از آن داده‌ها به‌نحو صحیحی استفاده کرد (ب) فردمحور هستند و ممکن است فرد یا افرادی که رموز داده‌ها را می‌دانند در آن شرکت یا سازمان مشغول کار نباشند. فرض کنید که در یک مجموعه داده، علامت * برای نشان دادن مرد و علامت # برای نشان دادن زن به کار رفته باشد. یا حرف «م» برای نشان دادن مکان و حرف «ز» برای نشان دادن زمان استفاده شده باشد و پس از چند سال ممکن است تحلیلگران این حروف را با «زن» و «مرد» اشتباه بگیرند.

Pipino, Lee and Wang (2002) مفهوم «ارزش افزوده» را شانزدهمین و آخرین بُعد کیفیت داده‌ها دانسته‌اند که به جنبه‌های اجتماعی و کلان استفاده از داده‌ها اشاره دارد. ممکن است داده‌هایی داشته باشیم که هر ۱۵ بُعد کیفیت را دارد و درباره تعداد دفعات مراجعه افراد به سرویس‌های بهداشتی است و بر اساس آن بخواهیم تصمیماتی درباره نحوه استفاده، ساعات استفاده و دفعات استفاده از سرویس‌های بهداشتی بگیریم یا داده‌های موثقی در سازمان داشته باشیم درباره تعداد دفعات ورود افراد به سامانه ایمیل و بخواهیم با آن تصمیماتی درباره استفاده از ایمیل بگیریم. شاید داده‌ها از هر لحاظ با کیفیت باشد، ولی در نگاهی اجتماعی‌تر و سازمانی‌تر مفید نباشد که بر اساس آن‌ها تصمیمی در خصوص یکی از جنبه‌های انسانی محیط گرفته شود. به همین دلیل، گفته می‌شود که مجموعه داده برای اتخاذ برخی تصمیمات خاص فاقد کیفیت مناسب است. شاید بتوان مفهوم «ربط» را کلیدی‌ترین مفهوم در کیفیت داده‌ها دانست و همه ابعاد را با مفهوم ربط توضیح داد: ربط داده‌ها با تصمیمی که قرار است اتخاذ شود، ربط داده‌ها با یکدیگر (سازگاری درونی)، ربط داده‌ها به مجموعه داده‌های دیگر در زمان‌های مختلف (سازگاری بیرونی)، ربط داده‌ها به مقولات، ربط داده‌ها به رموز یا علائم و قابل فهم بودن آن‌ها، ربط داده‌ها به ابزارها و سیستم‌ها، ربط داده‌ها به منابع (سازمان‌ها) و امثال آن، در معنای همان تفاسیری است که در باب ابعاد شانزده گانه ذکر گردید و به همین دلیل شاید بتوان همه مقولات کیفیت را با مفهوم «ربط» توضیح داد و تبیین کرد.

۶. حفاظت داده

گردآوری داده‌های کاربران توسط وبگاه‌های تجاری و سازمانی نگرانی‌های اخلاقی عدیده‌ای در باب عدالت و انصاف ایجاد می‌کند و پژوهشگران، این نگرانی‌ها را در قالب مقاله‌ها و پژوهش‌های متعدد عرضه کرده‌اند. در این بخش برخی از این نگرانی‌ها فهرست می‌شوند و در انتها الگویی برای ارزیابی اخلاقی بودن برخورد وبسایت شرکت یا سازمان با داده‌های مخاطبان عرضه می‌شود. ابتدا به صورت گزاره‌ای و فهرست‌وار به نگرانی‌ها می‌پردازیم:

برخی شرکت‌ها، اطلاعات پروفایل کاربران و داده‌های آن‌ها را به شرکت‌های تبلیغاتی یا شرکت‌های شریک می‌فروشند و در برخی کشورها قانونی درباره این رفتارها وجود ندارد (Numerico 2019). برای مثال، می‌توان به نرم‌افزار بیکن^۱ اشاره کرد که در سال ۲۰۰۸، روی سامانه فیس‌بوک سوار شده بود و کارش شناسایی و دسته‌بندی علائق کاربران و فروش علائق به شرکت‌های تبلیغاتی بود. با تحلیل کلان‌داده‌ها حتی مسایل جنسی کاربران هم تحلیل شده بود و حتی به برخی از آن‌ها که دو-جنسیتی بودند، ابزارهای جنسی خاصی نیز پیشنهاد داده شده بود که برایشان بسیار تعجب‌آور بود (McGuire 2010).

برخی مؤسسات یا سازمان‌های حوزه پزشکی و سلامت از بدن انسان‌ها و حتی بدن اجساد سوء استفاده می‌کنند تا از آن داده استخراج کنند. برای مثال، گزارش موثقی در متون علمی و رسانه‌ها منتشر شده مبنی بر اینکه اجزای بدن اجساد را استخراج و بدون اجازه نگهداری می‌کردند تا داده‌هایی را با توجیه استفاده برای منفعت عمومی استخراج کنند و این در حالی است که در این نوع اعمال، از انسان به‌مثابه ابزار استفاده شده است (Lipworth et al. 2017).

در گذشته، انسان‌ها در برابر حوادث آینده شانس مساوی داشتند، ولی الگوییابی از داده‌ها برخی افراد را نسبت به وقایع آینده در موقعیت برتر قرار می‌دهد. در ورود به آینده، دیگر افراد شانس برابری نخواهند داشت و پیش‌بینی‌پذیر شدن آینده برای برخی و پیش‌بینی‌ناپذیر بودن آن برای برخی دیگر باعث می‌شود که آن‌ها این امکان را بیابند تا آینده را به ضرر برخی دیگر و به نفع خود رقم بزنند (Hasselbalch 2019). برای

مثال، تحلیل رفتار خرید خانه از طریق الگوییابی از بانک اطلاعات مسکن باعث می‌شود برخی از شرکت‌ها بتوانند گوی سبقت را در خرید خانه و مسکن برابند و الگوییابی از داده‌های خرید خواروبار یا لوازم خانگی توسط برخی شرکت‌ها باعث می‌شود در واردات برخی کالاها یا وارد نکردن برخی دیگر سبقت بگیرند و دیگران را متضرر کنند. در واقع، رفتار خرید و کشیدن کارت اعتباری در کارتخوان‌ها که بر اساس شغل و کسب‌وکار خریدار قابل الگوییابی است، می‌تواند نشان دهد که کدام کسب‌وکارها یا کدام نوع از افراد بیشترین خرید از کالایی خاص را از طریق کارتخوان انجام داده‌اند و این نیز به‌نوبه خود می‌تواند به‌راحتی برای واردکنندگان کالاها که امتیاز انحصاری این کار را از دولت دریافت کرده‌اند، برای به‌هم زدن بازار داخلی و کسب سود بیشتر استفاده شود.

برخی شرکت‌ها، به‌ویژه شرکت‌های ارائه‌دهنده خدمات اینترنتی یا شرکت‌های مخابراتی از الگوی مصرف اینترنت یا مصرف مکالمه مثل توزیع کنندگان مواد مخدر استفاده می‌کنند. آن‌ها الگوییابی می‌کنند که چه کسانی تازه به شبکه مصرف اینترنت یا مخابرات پیوسته‌اند و چه کسانی قدیمی‌اند. سپس به کسانی که تازه پیوسته‌اند، اینترنت و بسته‌های مکالمه مجانی ارائه می‌دهند تا مثل فروشندگان مواد مخدر، مشتریان را به مصرف اینترنت یا مکالمه معتاد کنند. این در حالی است که به مشتریان قدیمی بسته هدیه مکالمه یا اینترنت ارائه نمی‌دهند؛ زیرا آن‌ها پیشتر به شبکه قفل شده‌اند و تغییر شرکت خدمات‌دهنده برایشان هزینه‌زاست. این نوع استفاده از داده کاوی نیز نوعی استفاده ابزاری از انسان‌هاست که شوربختانه تداول دارد (McGuire 2010).

لینک شدن دو یا چند پایگاه داده مربوط به دو شرکت مختلف برای الگوییابی. مهم‌ترین مثال آن، همان نرم‌افزار «بیکن» و تحلیل علایق کاربران فیس‌بوک بود که با اطلاعات پروفایل کاربران موجود در پایگاه اطلاعاتی فروشندگان لوازم آرایشی و بهداشتی لینک شده بود (Jurkiewicz 2018).

چه داده‌هایی از کاربران، واقعی است و چه داده‌هایی آمیخته با کذب؟ بسیاری از داده‌هایی که از کارکنان دولت یا شرکت‌ها در پروفایل کاری آن‌ها وجود دارد، واقعیت آن‌ها را نشان نمی‌دهد؛ در حالی که هیچ انسانی به موتورهای جست‌وجو از جمله گوگل دروغ نمی‌گوید؛ زیرا خواهان راه حلی برای مشکل واقعی خویش است. مردم ممکن است بیماری یا انحراف جنسی یا سایر امور خصوصی خود را پنهان کنند، ولی این موارد را از گوگل و سایر موتورهای جست‌وجو پنهان نمی‌کنند و این یعنی داده‌های ذخیره‌شده

در موتورهای جست‌وجویی مثل گوگل واقعی‌ترین داده‌ها دربارهٔ انسان‌هاست که در اختیار یک شرکت قرار می‌گیرد (Stephens-Davidowitz 2017).

برخی از شرکت‌ها، داده‌های گروه‌ها یا کانال‌های موجود در شبکه‌های اجتماعی را به صورت دسته‌بندی شده استخراج می‌کنند، با این استدلال که «چون داده‌های این گروه‌ها عمومی است، پس استخراج آن‌ها برای تحلیل مجاز است». آن‌ها این نوع داده‌ها را به نوعی جزو «داده‌های باز» تلقی می‌کنند؛ در حالی که آن گروه‌ها شامل اشخاص خاصی هستند که به واسطهٔ داشتن یک علاقهٔ مشترک آنجا گردآمده‌اند و یک ویژگی مشترک دارند و یک برچسب مشترک خورده‌اند (مثلاً گروه نشانه‌شناسی رسانه در تلگرام یا ایتا) و هر تحلیل از داده‌های آن گروه، به ویژگی‌های آن گروه بار می‌شود و این یکی از مباحث کلیدی در موضوع «محرمانگی گروهی» است (Mittelstadt 2017).

یکی از اصول مهمی که بر اساس نوع مالکیت «کنترل-عدم دسترسی» تعریف شده، «رضایت آگاهانه» است. بر این اساس، وبسایت‌هایی که دست به ثبت نام کاربران می‌زنند و اطلاعات تراکنش یا تعامل آن‌ها را ذخیره می‌کنند باید طی یک اعلامیهٔ آشکار در سایت خود اعلام و تصدیق کنند که از داده‌های کاربران برای چه مواردی استفاده می‌کنند. چنین چیزی در بسیاری از وبگاه‌های داخلی و خارجی وجود ندارد و این به یکی از مطالبه‌های حوزهٔ جدید «اخلاق داده» تبدیل شده است. شایستهٔ ذکر است که بر اساس نظر برخی اندیشمندان، نبود تحمیل پیشفرض پنهان رضایت آگاهانه است و به همین دلیل، آن‌ها رضایت گرفتن از خانوادهٔ بیمار برای جراحی را از نوع رضایت آگاهانه نمی‌دانند؛ زیرا اگر رضایت ندهند، عضو خانواده‌شان جراحی نمی‌شود. فرم رضایت آگاهانه باید شامل مواردی باشد که فرد بتواند به آن رضایت ندهد.

به دلیل وجود مسائلی که از نظر گذشت، می‌توان با تکیه بر کتاب DAMA-DMBOK: Data Management Body of Knowledge که به سرویراستاری «لراسباستین کولمن»^۱ و همکاران تنظیم شده، اصولی را پیشنهاد داد که بر اساس آن بتوان تعامل اخلاقی تری را با داده‌ها در جامعه پایه‌ریزی کرد. این اصول در ذیل چهار بُعد قابل دسته‌بندی است (۱) بُعد چارچوب اخلاقی، (۲) بُعد کاربری اخلاقی، (۳) بُعد پردازش اخلاقی، و (۴) بُعد حفاظت اخلاقی. بُعد چارچوب اخلاقی اشاره به ساختار و زیربنای اخلاقی یک چیز دارد. ساختارها

دلالت بر قواعد و اصول کلی حاکم بر یک چیز را دارند. منظور از چارچوب اخلاقی آن است که بررسی شود پایگاه اطلاعاتی از چه قواعد کلی اخلاق برای هدایت رفتارها و رفع چالش‌ها استفاده می‌کند. در بعد این چارچوب، این پرسش‌ها مطرح است: الف) آیا پایگاه اطلاعاتی دارای اصول ارزشیابی اخلاقی است؟ یعنی آیا پایگاه اطلاعاتی اعلام کرده که متکی به کدام نظریه اخلاقی است؟ (مثلاً اخلاق نتیجه‌گرا، اخلاق عدالت‌گرا و امثال آن، ب) آیا پایگاه اطلاعاتی در استفاده از اصول اخلاقی، دارای وحدت رویه است؟ آیا قواعد حاصل از آن اصول، دارای وحدت رویه است؟ ج) آیا قواعد اخلاقی پایگاه اطلاعاتی با شهودهای اخلاقی یک انسان معمولی همخوانی دارد؟ بُعد کاربری اخلاقی اشاره دارد به کاربرد اخلاقی داده‌ها در پایگاه اطلاعاتی. منظور از کاربری اخلاقی آن است که آیا هنگام ثبت، انتشار، استفاده، ترکیب و دست‌ورزی با داده‌ها اصول اخلاقی رعایت می‌شود یا خیر. در بعد کاربری، پرسش‌های زیر مطرح است: الف) آیا پایگاه اطلاعاتی قواعد اخلاقی ثبت و ذخیره‌سازی داده‌ها را به کاربران اعلام می‌کند؟ ب) آیا کاربران پایگاه اطلاعاتی در مورد اینکه چه داده‌هایی ارائه می‌شود و چه کاری با آن انجام می‌شود، می‌توانند آگاهی کسب کنند؟ ج) آیا پایگاه اطلاعاتی اجازه داده است که کاربران داده‌های خود را کنترل کنند؟ د) آیا پایگاه اطلاعاتی، کاربران (نویسندگان، کاربران عادی و غیره) را از ریسک‌ها و خطراتی که ممکن است برای داده‌هایشان ایجاد شود، آگاه می‌کند؟ ه) آیا پایگاه اطلاعاتی الگوریتمی جهت نظارت بر داده‌ها (مانند کنترل دسترسی کاربران داخلی و اشخاص ثالث) دارد؟ و) آیا امکان رصد پایدار داده‌های شخصی در سراسر پایگاه اطلاعاتی تعبیه شده و پاسخگویی را در کوتاه‌مدت، میان‌مدت و بلندمدت تضمین می‌کند؟ ز) آیا پایگاه اطلاعاتی روش‌هایی برای پنهان کردن اطلاعات حساس کاربران، در هنگام نقل و انتقال یا تحلیل داده‌ها دارد؟ ح) آیا دسته‌بندی‌هایی که پایگاه اطلاعاتی از موجودیت‌ها (منابع، آثار، کاربران و غیره) ارائه کرده، سوگیرانه است؟ ط) آیا بخش‌های مختلف پایگاه اطلاعاتی به صورت سوگیرانه تقسیم‌بندی شده است؟ ی) آیا پایگاه اطلاعاتی کاربران را آگاه کرده است که چه نوع تحلیل‌هایی روی داده‌های کاربران انجام می‌شود؟ ط) آیا پایگاه اطلاعاتی منابع احتمالی سوگیری داده‌ها را بررسی و اقداماتی برای کاهش یا رسیدگی به این سوگیری‌ها انجام می‌دهد؟ ک) آیا در پایگاه اطلاعاتی سازوکاری وجود دارد که رضایت کاربر برای انواع استفاده‌ها از داده‌هایش اخذ شود و بتواند به مواردی رضایت ندهد؟ گ) آیا پایگاه اطلاعاتی سازوکاری برای

اخذ رضایت از کاربران، برای هر نوع استفاده ثانویه از داده‌های کاربران را دارد؟ (ل) آیا در پایگاه اطلاعاتی برای حفظ حریم خصوصی کاربران، سطوح دسترسی تعریف شده است؟ (م) آیا در پایگاه اطلاعاتی سازوکاری وجود دارد که از طریق آن فرد بتواند اطلاعات شخصی خود را درخواست کند؟ (ن) آیا در پایگاه اطلاعاتی غیر از اشخاص، گروه‌های اجتماعی نیز به‌عنوان صاحبان حق حریم خصوصی در نظر گرفته شده‌اند؟ (س) آیا حق فراموش شدن داده‌ها برای کاربران در نظر گرفته شده است؟ یعنی سازوکاری وجود دارد که هرگاه کاربران خواستند، داده‌های مرتبط با آن‌ها از پایگاه حذف شود؟ بُعد پردازش اشاره دارد به استفاده از ابزارهای تحلیل و پردازش و الگوییابی از داده‌ها و اخلاقی بودن آن. منظور از پردازش اخلاقی آن است که آیا ابزارها (مثلاً ربات‌ها یا نرم‌افزارها) و وسایل مورد استفاده برای پردازش داده‌ها و نحوه به‌کارگیری آن‌ها از اصول اخلاقی تبعیت می‌کند یا خیر. در بُعد پردازش این پرسش‌ها مطرح است: الف) آیا پایگاه اطلاعاتی درباره تحلیل و الگوییابی و استفاده‌های دیگر از داده‌ها، اطلاعاتی به کاربران داده است؟ آیا پایگاه اطلاعاتی سازوکاری برای اخذ رضایت درباره نوع داده‌های و الگوییابی از داده‌ها دارد؟ (ب) آیا پایگاه اطلاعاتی قابلیت‌هایی جهت رهگیری و مشاهده جست‌وجوها و انتخاب‌های کاربران ایجاد نموده است؟ و آیا آن را به کاربران اطلاع می‌دهد؟ (ج) آیا پایگاه اطلاعاتی هر نوع استفاده همزمان از داده‌های این پایگاه با داده‌های پایگاه دیگر برای الگوییابی را اطلاع‌رسانی می‌کند یا سازوکاری برای رضایت گرفتن برای آن دارد؟ (د) آیا پایگاه اطلاعاتی در برابر اطلاعات غلط، مسئولیت اجتماعی را پذیرفته است یا مکانیسمی برای رفع اطلاعات غلط دارد؟ (ه) آیا پایگاه اطلاعاتی دارای برنامه زمان‌بندی امحای داده‌های کاربران است؟ یعنی معلوم کرده است که اطلاعات فعالیت کاربران را تا چه زمانی نگهداری می‌کند؟ (و) آیا پایگاه اطلاعاتی داده‌هایی را از کاربران گردآوری می‌کند که هویت آن‌ها با آن داده‌ها قابل کشف است؟ (ز) آیا پایگاه اطلاعاتی اعلام کرده است که اطلاعات کاربران به کسی فروخته نمی‌شود یا الگوییابی از تراکنش کاربران در اختیار کسی قرار نمی‌گیرد؟ (ح) آیا پایگاه اطلاعاتی کاربران را مجبور می‌کند که نصب شدن کوکی‌ها روی سیستم‌شان را قبول کنند؟ آیا اگر قبول نکنند، نمی‌توانند از پایگاه استفاده کنند؟ (ی) آیا پایگاه اطلاعاتی از الگوریتم‌های شخصی‌سازی پاسخ جست‌وجو استفاده می‌کند؟ آیا آن را به کاربران اطلاع می‌دهد؟ و در نهایت بُعد امنیت دلالت دارد بر نگهداری و توسعه آینده پایگاه اطلاعاتی و داده‌های آن، و اخلاقی بودن این نگهداری و

توسعه آینده. منظور از حفاظت اخلاقی آن است که حفاظت از داده‌ها و توسعه فناورانه آتی آن با اصول اخلاقی منطبق است یا خیر. در این بُعد این موارد قابل پرسش است: الف) آیا سازوکار تخصیص نام کاربری و کلمه عبور در پایگاه اطلاعاتی به‌اندازه کافی پیچیده است؟ ب) آیا پایگاه اطلاعاتی سازوکاری برای تأیید و اطلاع‌رسانی ایمیلی یا پیامکی هر نوع ورود یا تغییر نام کاربری و کلمه عبور دارد؟ ج) آیا پایگاه اطلاعاتی به‌طور رسمی برنامه‌ای را برای حفاظت از داده‌های کاربران اعلام نموده است؟ د) آیا اگر کاربران، از سیستم دیگری برای ورود به پایگاه اطلاعاتی استفاده کنند، نیازمند تأیید هویت مجدد می‌شوند؟ ه) آیا هر کاربر، همزمان می‌تواند چندین حساب کاربری در پایگاه اطلاعاتی داشته باشد؟ و) آیا امکان بازیابی رمز عبور در پایگاه اطلاعاتی وجود دارد؟ و آیا به‌اندازه کافی ایمن است؟ ز) آیا پایگاه اطلاعاتی به‌صورت دوره‌ای داده‌های امنیتی کاربران را روزآمد می‌کند و برای مثال، تغییر رمز عبور را هر شش ماه یکبار الزامی می‌نماید؟ جدول ۱، می‌تواند الگویی شش‌بُعدی از سیاست‌گذاری اخلاقی داده‌ها باشد:

جدول ۱. الگوی سیاست‌گذاری بر اساس مضامین استخراج‌شده

مضامین	دلالت‌های اخلاقی	دلالت‌های ساختاری و اجتماعی	رویکردهای سیاست‌گذارانه	کنشگران کلیدی
اقتصاد داده	یکسان‌سازی، برنده همه را می‌برد، کالایی‌سازی تجربه زیسته کاربران، استخراج منفعت اقتصادی، سیاسی و اجتماعی از سوژه دیجیتال	تعمیق نابرابری اطلاعاتی، تکوین طبقات داده‌محور، بی‌قدرتی اپراتورهای خرد	طراحی نظام‌های بازتوزیعی داده، محدودسازی تراکم اطلاعاتی شرکت‌های کلان، مشروط‌سازی اقتصاد داده به حقوق جمعی	دولت، شرکت‌های پلتفرمی، نهادهای مدنی
تمرکز ساختارهای انباشت اطلاعاتی در دست کنشگران سرمایه‌داری پلتفرمی	به‌حاشیه‌رفتن فاعلان غیربازاری، غیاب تکثر اطلاعاتی	سیاست‌های ضدانحصار پلتفرمی، تأسیس نهادهای واسط داده برای مشارکت همگانی در اقتصاد داده	نهادهای ضدانحصار، اتحادیه‌های دیجیتال	
گردآوری نادیده‌انگاشته شدن مفهوم و خلق داده «رضایت آگاهانه» و تقلیل سوژه انسانی به منبع خام داده	کالاسازی بدن و ذهن افراد، تخریب حریم خصوصی، مصادره زیست‌جهان افراد	بازتعریف رضایت به‌مثابه رابطه اخلاقی و نه صرفاً فرمال، اعمال نظارت اخلاقی بینارشته‌ای بر فرایند خلق داده	کمیته‌های اخلاق پژوهش، جوامع محلی تولید، ذخیره و پردازش داده	

مضامین	داللت‌های اخلاقی	داللت‌های ساختاری و اجتماعی	رویکردهای سیاست‌گذارانه	کنشگران کلیدی
	بهره‌برداری بدون تحلیل زمینه‌ای از داده‌های عمومی و نادیده گرفتن محرمانگی جمعی	برجسب‌زنی جمعی، باز تولید طبقه‌بندی‌های تبعیض‌آمیز	بازاندیشی در تعریف «عمومی بودن»، تنظیم چارچوب‌های اخلاقی برای گردآوری از گروه‌های خاص	دانشگاه‌ها، جوامع تخصصی دیجیتال
مالکیت داده	نبودن نظر به اجتماعی معتبر برای مالکیت داده در زیست‌جهان دیجیتال	حذف عاملیت کاربران از تصمیم‌گیری بر داده‌های خود، تقلیل «حق» به «دسترسی»	پیشنهاد نظام «مالکیت رابطه‌ای داده» با تأکید بر عاملیت جمعی و حق حاکمیت کاربران بر ردپاهای دیجیتال خود	حقوق دانان داده، نهادهای حقوق بشری
	برداشت مالکانه شرکت‌ها از داده‌هایی که در تعامل اجتماعی تولید شده‌اند	تغییر توازن قدرت در تعاملات دیجیتال، انقیاد کاربران	وضع مقررات حقوق داده اجتماعی، شفاف‌سازی لایه‌های «مالکیت» و «دسترسی» و «استفاده»	دادگاه‌ها، نهادهای مقررات‌گذار
کیفیت داده	سوگیری ساختاری در داده‌های خام و الگوریتم‌ها به مثابه بازتابی از نابرابری اجتماعی	باز تولید تبعیض از طریق تحلیل‌های ظاهراً بی‌طرفانه، تکوین «نابرابری تحلیلی»	ایجاد سازوکارهای بازبینی سوگیری، طراحی الگوهای داده‌محور حساس به زمینه و طبقه	پژوهشگران داده، نهادهای ارزیاب کیفیت
	استفاده از داده بدون تحلیل زمینه‌مند از آن	کژفهمی واقعیت‌های اجتماعی، حذف تنوع معنایی زیست‌جهان انسانی	آموزش اخلاق داده‌زمینه‌محور، الزام به روایت‌مندی در تحلیل داده	جامعه علمی، پلتفرم‌های داده
داده‌های باز	تلقی داده به مثابه «منبع عمومی مطلق» بدون رعایت لایه‌بندی حقوق و تعلقات اجتماعی	درز اطلاعات حساس، تعرض به حریم گروهی، بهره‌برداری ناعادلانه	طراحی سطوح دسترسی چندلایه، تعریف حقوق جمعی در حوزه داده‌های باز، ایجاد توازن میان شفافیت و حریم	دولت، گروه‌های اجتماعی، پژوهشگران داده
	دسترسی نابرابر به زیرساخت تحلیل داده ¹	توانمندسازی شرکت‌های بزرگ و حاشیه‌نشینی بازیگران خرد	تدارک زیرساخت‌های عمومی تحلیل داده، توانمندسازی کنشگران محلی برای بهره‌مندی عادلانه از داده‌های باز	جامعه مدنی، دانشگاه‌ها، دولت محلی
حفاظت داده	فقدان نظام زیرساخت‌های امنیتی در نگهداری داده و غیبت رویکردهای پیشگیرانه	آسیب‌پذیری کاربران، سوء استفاده سازمان‌یافته از داده‌های حساس	نهادینه‌سازی سیاست‌های حفاظت فعال، رمزنگاری، مدیریت بحران داده، طراحی حق «امحای دیجیتال» سیستم	نهادهای امنیتی دیجیتال، طراحان سیستم
	ناممکن بودن فراموشی در معماری دیجیتال	تراکم سوابق دیجیتال، کنترل غیرقابل بازگشت بر سوژه	قانون‌گذاری برای حق فراموش شدن، طراحی داده‌های منقضی‌شونده و کنترل‌پذیر	قانون‌گذاران، طراحان پلتفرم

ورای نتیجه‌گیری: مواضع انتقادی

از نقدهای قابل توجهی که درباره‌ی ظهور مضامین جدید در حوزه‌ی اخلاق مطرح شده، آن است که ادعا می‌شود بین مسئله‌ی استخراج، مصرف و بهره‌کشی از منابع سرزمینی یا نیروی انسانی و ظهور مفاهیم جدید اخلاقی مانند اخلاق ماشین، اخلاق رایانه، اخلاق اطلاعات، اخلاق داده و امثال آن رابطه برقرار است. برای مثال، «ماسون» اشاره می‌کند که بین عصر انقلاب صنعتی و عصر حاضر همخوانی و هماهنگی وجود دارد؛ زیرا در هر دو این دوره‌ها، با انتشار بیش از حد متون و آثار در باب اخلاق مواجه بوده‌ایم. از نظر وی در دوره‌ی انقلاب صنعتی ساعات زیاد کار در کارخانه‌ها و تخصص‌زدایی از کارگران به‌واسطه‌ی تقسیم کار باعث بیگانگی کارگر از کار شده بود و احتمال صدمه زدن کارگران به دستگاه‌ها زیاد بود و حتی گروهی به نام لادایت‌ها (و بعدها نئولادایت‌ها)^۱ ظهور کرده بودند که به هر نوع ماشین صنعتی آسیب می‌زدند. و اتفاقاً در همین دوره بود که حجم عظیمی از ادبیات درباره‌ی اخلاق توسط اندیشمندانی چون «امانوئل کانت»، «جرمی بنتام»، «جان استوارت میل» و امثال آن تولید شد که از زمان افلاطون و ارسطو تا آن زمان نظیر نداشت. «ماسون» ادعا می‌کند که هر جا بهره‌کشی از بدن‌ها، زمین‌ها و منابع سرزمینی آغاز و موانعی از سوی مردم یا دولت ایجاد می‌شود، متون اخلاقی نیز شروع به تولید می‌کند (Mason 2007). برای مثال، درست در همان زمان که گروهی از انگلیسی‌ها وارد قاره‌ی آمریکا شده و شروع به حصارکشی و تصاحب زمین‌های بومیان می‌کنند، «لاک»^۲ رساله‌ی حقوق خود را می‌نویسد تا مالکیت انسان بر بدن خویش و بنابراین، بر کار خویش و سرانجام بر هر چیزی را که با کار انسان آمیخته شده، اثبات کند تا به این طریق بتواند از زمینی که در آمریکا تصاحب کرده و پادشاه انگلستان چشم طمع بر آن داشته، دفاع نماید (Santal 2008). بنابراین، منتقدان امروزه، این ادعای «لاک» را که «حق مالکیت حقی طبیعی است»، به پرسش می‌کشند. پرسش آن‌ها این است که اصرار «لاک» بر اینکه «حق فرد برای مالکیت، حق طبیعی است و حق ساختگی (یا ساخته‌ی انسان) نیست»، مبتنی بر چه شواهدی است. همچنین، نظریه‌ی مالکیت «لاک» پیش‌فرض می‌گیرد که شخصی که مدعی مالکیت است، «مالک بدنش است» و به همین دلیل مالک کار بدنی‌اش است. اگر حق مالکیت حقی

1. NeoLuddites

2. John Locke

طبیعی است، پس همه اشخاص باید از آن برخوردار باشند. ولی نکته این است که بردگی در دوره «لاک» موجود بوده و رواج داشته و گویا توجه به بدن بردگان و بومیان و کار آنان در میان نبوده است. بردگان طبق قانون، مالک بدن خود نبودند و بر اساس منطق استدلال «لاک»، آن‌ها نمی‌توانند ادعایی نسبت به ثمره کارشان داشته باشند و این یعنی آن‌ها حق مالکیت ندارند. بنابراین، طبق نظریه کار «لاک»، حق مالکیت قابل ارائه به همه انسان‌ها نیست و اگر بود، بومی‌های آمریکا که کارشان را با زمین مخلوط می‌کردند نیز باید دارای حق مالکیت در زمین‌های شمال و جنوب آمریکا می‌شدند. منتقدان ادعا می‌کنند که هر جا استثمار، بهره‌کشی، استخراج و سوء استفاده و امثال آن به شکل گسترده آغاز می‌شود، تولید متون اخلاقی نیز فزونی می‌یابد که البته این متون نقش سرپوش را بازی می‌کنند و چون ضمانت اجرایی ندارند، نقش مُسکّن و آرامش‌بخش را برای مردم عادی دارند تا آن‌ها را مطمئن کنند که شرکت‌ها، صنایع و صاحبان سرمایه در چارچوب‌های اخلاقی مقید شده‌اند و استفاده غیراخلاقی از منابع (مثلاً داده‌ها) تحت کنترل است. نبود ضمانت اجرایی و به‌ویژه اثرگذاری، در پژوهش «هگندورف» نمایان است. وی طی یک مطالعه کنترل‌شده روی نمونه‌ای از شرکت‌های نرم‌افزاری دارای سند اخلاق داده نشان داد که عملاً هیچ‌یک از آن‌ها اثرگذار نبوده و به‌طور تلویحی اشاره می‌کند که صرفاً دارای جنبه نمایشی بوده‌اند (Hagendorf 2020, 108). از نظر منتقدانی مثل «فن‌مانن»، اکنون که در دوره داده و بهره‌کشی از داده‌ها هستیم نیز شاهد وفور متون در خصوص اخلاق داده و اطلاعات هستیم و شرکت‌ها و بنگاه‌های تجاری و مراکز صنعتی شروع به تولید انبوه چارچوب‌های اخلاقی و اصول اخلاق سازمانی و حرفه‌ای درباره داده‌ها کرده‌اند تا اولاً از اخلاق داده به‌عنوان سرپوشی برای بهره‌کشی از داده‌ها و نفوذ به حریم خصوصی استفاده شود و ثانیاً نگرانی‌ها و حرف‌وحديث‌هایی را که درباره سیاست‌های ناسالم برخی شرکت‌ها و سازمان‌ها وجود دارد، بشوید و ببرد (Van Maanen 2022). پیش‌فرض این نوع اصول اخلاقی آن است که قوانین موجود در باب داده و اطلاعات خوب است و بقیه موارد باید با اصول اخلاقی تنظیم شود. پیش‌فرض دوم این است که فرد دارای یک آگویی خود مستقل و ممتازی است که تحت تأثیر ساختارها (سیاسی، اجتماعی و سازمانی) نیست و در وهله بعد جو سازمان نیز تأثیری روی وی ندارد و می‌تواند به‌طور مستقل تصمیم بگیرد. غیر از این، نگارش اصول اخلاقی برای سازمان‌ها تبدیل به یک کسب‌وکار پردرآمد برای عده‌ای شده است. این یک پدیده عمومی است و ایران نیز از این قضیه مستثنا نیست.

وقتی به سازمان‌های مختلف و اصول اخلاقی آن‌ها مراجعه می‌کنیم، یک کلیشه یا شکل و قالب خاصی را می‌بینیم که نشان می‌دهد همه آن‌ها از کارخانه تولید اصول اخلاقی یک فرد خاص بیرون آمده‌اند (برای مثال، به اصول اخلاقی شرکت گاز ایران از بخش صنعت و نهاد کتابخانه‌های عمومی از بخش فرهنگ نگاه کنید که هر دو مثل هم‌اند. در خارج از کشور، حتی مطالعه‌ای توسط Floridi & Cowls (2019) انجام شده و قالب یکپارچه‌ای برای نگارش اصول اخلاقی عرضه شده است!).

انتقاد دیگری که به جریان تدوین اصول اخلاقی از جمله اخلاق داده در شرکت‌ها، سازمان‌ها و دولت‌ها وارد است، آن است که این اصول در رویه‌های جاری کسب و کارها برای تولید ثروت و حفظ مزیت رقابتی یک وصله ناجور است. شرکت‌ها برای اینکه از سهم بازار بهره بیشتری ببرند، نیازمند پنهان‌کاری (برخلاف شفافیت)، گردآوری داده از مردم برای کشف الگوهای خرید و رفتار، اجیر کردن جاسوس‌های صنعتی، و تخریب رقیب هستند. برای مثال، تبلیغات آشکارا مخرب شرکت‌هایی که در یک حوزه فعالیت می‌کنند (مثل پیسی و کوکاکولا) نشان از رقابتی بودن فعالیت‌های شرکتی دارد و افراد و گروه‌ها با نوعی تنازع برای بقا درگیرند. پرسش این است که در اقتصاد سرمایه‌داری که اساساً بر پایه رقابت، حفظ مزیت رقابتی، بالا رفتن از نردبان و سپس پایین انداختن آن (برای اینکه کس دیگری از آن بالا نرود)، کاهش حداکثری هزینه‌های نیروی کار (برای کم کردن هزینه‌های غیرسرمایه‌ای)، مدیریت دانش (برای تبدیل کردن اطلاعات، داده و دانش کارکنان به سرمایه شرکتی و کاهش قدرت چانه‌زنی آن‌ها)، انتقال ریسک (از طبقات قدرتمند به طبقات ضعیف) و امثال این مفاهیم مواجه هستیم (چنگ ۲۰۰۸)، چطور می‌توان اصول اخلاقی را با این مفاهیم جفت‌وجور کرد، به گونه‌ای که هم بتوان مزیت رقابتی را حفظ کرد، هم بتوان شناخت بیشتری از مشتریان به دست آورد، هم هزینه‌های نیروی کار را کاهش داد، هم ریسک‌های شرکت را به حداقل رساند، هم در بازی تنازع برای بقا برنده بود و هم پای‌بند به اصول اخلاقی حوزه اطلاعات و داده بود؟ در خصوص دولت‌ها نیز همین مسئله اصول اخلاقی به مثابه یک وصله ناجور وجود دارد. «تاوانی» توضیح می‌دهد که پلیس آمریکا چطور برای گیرانداختن گروه‌های ضد هنجار مانند بچه‌بازها، در فضای مجازی حساب کاربری کاذب ایجاد می‌کنند، از موبایل دختران و پسران زندانی، عکس‌های شخصی‌شان را برمی‌دارند و در پروفایل می‌گذارند تا افرادی را که در رابطه با آن‌ها هستند، گیر بیندازند و دستگیر کنند (Tavani 2015). در واقع، سؤال

این است که وقتی برای مثال، دستور کار یک سازمان حفظ امنیت، آسایش و وضع نرمال است، این اصول اخلاقی از جمله اصول اخلاق داده چطور می‌تواند در آن چارچوب مفصل‌بندی شود؟ به گونه‌ای که هم دستور کار اصلی حفظ شود، هم امنیت مردم تضمین گردد، و هم رعایت به اصول اخلاقی برقرار بماند؟ مگر جز این است که شرح مبسوطی از استثنائات یا اولویت‌بندی‌ها نیاز است تا نشان دهد در چه مواقعی به لحاظ اخلاقی می‌توان از این اصول عدول کرد؟ آیا گاهی اوقات خود این عدول کردن اخلاقی نیست؟ (برای مثال، در داستان هاگلبری فین می‌خوانیم که چطور «هاگ» بر خلاف قوانین، اصول اخلاقی، و نظم موجود سند مالکیت صاحب «جیم» را نادیده می‌گیرد و او را نجات می‌دهد و اکنون، ما این عمل وی یا این نوع اعمال را اخلاقی‌تر از پای‌بندی به اصول نوشته‌شده می‌دانیم). آیا چنین اولویت‌ها، استثنائات، و اقتضائات، در عمل، اصول اخلاقی مکتوب و قاب‌شده بر دیوارها یا قرار داده‌شده در سایت‌ها را بی‌معنا نمی‌کند؟

در یک نمونه دیگر می‌توانیم به مفهوم «جامعه ریسک» که نخستین بار توسط «اولریش بک» ارائه شده، اشاره کنیم. در جامعه ریسک، مسئله اصلی توزیع نامتناسب ثروت نیست، بلکه توزیع نامتناسب ریسک است. «بک» در کتاب «جامعه ریسک» توضیح می‌دهد که پس از برقرار شدن شیوه تولید سرمایه‌دارانه، چطور دولت‌های حامی این شیوه تولید و همچنین شرکت‌های ملی و بین‌المللی در برهه‌های مختلف تلاش کرده‌اند تا ریسک را از گروه‌های قدرتمند به گروه‌های ضعیف‌تر منتقل کنند (Beck 1992). برای مثال، نخستین نمونه بارز انتقال ریسک را در دهه ۱۸۲۰ انگلستان شاهدیم؛ جایی که «چارلز تاونزند»^۱ در مجلس اعیان انگلستان تصریح می‌کند که ارائه دادن اعانه یا به گفته‌ای یارانه اسپیناملند^۲ به مردم عادی، نجات دادن آن‌ها از گرسنگی است و تا جایی که مردم ریسک گرسنگی را احساس نکنند، در شرکت‌ها مشغول به کار نخواهند شد و بنابراین باید این یارانه قطع شود تا فشار گرسنگی و احساس خطر مرگ، آن‌ها را به کار بکشاند (پولانی ۱۹۴۴). در دوره پس از ۱۹۶۰، در آمریکا و کشورهای اروپایی و سپس در کشورهای در حال توسعه، دقیقاً پس از برافتادن سبک اقتصاد کینزی، شاهد رونق کار موقت، برداشته شدن رابطه استخدامی قطعی، و کوتاه شدن قراردادهای هستیم که ریسک کار برای شاغلان

1. Charles Townshend

2. Speenhamland subsidiary

افزایش یابد. در حوزه صنعت و کشاورزی برای مثال، در حوزه میوه و یا غلات شاهد رشد روزافزون استانداردهای کیفیت هستیم که باعث می‌شود بسیاری از میوه‌ها یا غلاتی که اندکی کوچک‌تر هستند یا لکه دارند، روی دست کشاورزان غیرصنعتی خرده‌پا باقی بماند تا ریسک فروش میوه برای شرکت‌های توزیع، کاهش یابد و در عوض کشاورزان خرده‌پا ضرر کنند و مجبور باشند یا محصول را دور بریزند و یا در بازارهای محلی به قیمت پایین عرضه کنند. همچنین، شاهد یارانه‌های دولتی کشورهای توسعه‌یافته به کشورهای فقیر (مثلاً یارانه گندم ایالات متحده به ایتالیایی) هستیم که از یک سو، ریسک ضرر را برای کشاورزی صنعتی آمریکا کاهش می‌دهد و از سوی دیگر، صرفه اقتصادی تولید گندم را برای کشاورزان ایتالیایی از بین می‌برد؛ زیرا نمی‌توانند گندمی تولید کنند که قیمت آن کمتر از گندم آمریکایی و کیفیت آن بیشتر باشد. همچنین، شاهد روند روزافزون استفاده از نگهدارنده‌های مواد غذایی هستیم که ریسک خرابی محصول را برای شرکت تولیدکننده کاهش، و ریسک ابتلا به بیماری را برای مصرف‌کنندگان افزایش می‌دهد. یا شاهد افزایش روزافزون استفاده از آفت‌کش‌ها و مواد شیمیایی هستیم که ریسک خرابی محصول را برای کشاورزی صنعتی کاهش، و ریسک ابتلا به بیماری یا آلودگی هوا را برای مصرف‌کنندگان عادی افزایش می‌دهد. این در حالی است که بعضی از دارندگان همین صنایع، در مناطق مرتفع شهر زندگی می‌کنند تا از خطر آلاینده‌ها در امان باشند و از غذاهای ارگانیک استفاده می‌کنند تا خطر ابتلا به بیماری برایشان کاهش یابد. ذکر این مصادیق از این جهت بود که نشان داده شود در جامعه ریسک، ما شاهد روند روزافزون انتقال ریسک از سمت نیروی سرمایه و حامیان آن به سمت مصرف‌کنندگان و نیروی کار هستیم. ولی نکته اینجاست که اصول اخلاق داده (از جمله آن‌هایی که در سند راهبرد داده‌ای اتحادیه اروپا که یکی از بندهای آن عبارت از «دادن حق، ابزار و مهارت به کاربران جهت تضمین کنترل کامل آن‌ها روی داده‌هایشان» است)، این جریان انتقال ریسک را کاهش می‌دهد یا با کندی مواجه می‌سازد و یا در عمل، ریسک شرکت‌ها و دولت‌ها را افزایش می‌دهد و این برخلاف رویه انتقال ریسک از سوی قدرتمندان و ثروتمندان به مردم عادی و مصرف‌کنندگان است. برای مثال، فرض کنید شرکت یا دولتی بخواهد به مردم آموزش دهد که چطور و با چه ابزارهایی از داده‌های خود محافظت کنند که دست دولت‌ها و شرکت‌ها به آن‌ها نرسد و بتوانند محرمانه و بدون زیر نظر بودن زندگی کنند. به‌رغم مثبت بودن نظر این شرکت یا دولت، خواه‌ناخواه سئوالاتی مطرح می‌شود که

پاسخ دادن به آن‌ها چندان ساده و آسان نیست. برای مثال، آیا چنین آموزش‌هایی خلاف امنیت اجتماعی یا حفظ نظم اقتصادی جاری نیست؟ آیا به‌جای اینکه ریسک دولت‌ها و شرکت‌ها را کاهش دهد، آن را افزایش نمی‌دهد؟ آیا این بر خلاف روح زمانه نیست؟ به همین دلیل، استدلال می‌شود که اصول اخلاقی از نقطه‌نظر انتقادی یک نوع وصله ناجور تلقی می‌شود و برای عمل کردن نیست. به همین دلیل، بسیاری از پژوهشگران معتقدند بایستی فهرست‌هایی مبتنی بر مطالعات دوره‌ای عرضه شود تا میزان تبعیت از اصول اخلاقی خودنوشته سازمان‌ها و شرکت‌ها را بر اساس عملکردها و تخلفاتشان نشان دهد تا معلوم شود بین مدعا و عمل چقدر فاصله است. همچنین، اصول اخلاقی باید معیار بررسی سوء رفتارهای سازمانی باشد و سایر آیین‌نامه‌ها (مثل آیین‌نامه تعارض منافع، تخلفات اداری و امثال آن) باید از دل منشورهای اخلاقی به‌علاوه قوانین جاری درآید و به‌عنوان ابزار کنترل استفاده شود، نه اینکه زینت‌المجالس باشد و هیچ ربط و ضبطی به کنش‌ها و رویه‌های فعلی سازمان‌ها و شرکت‌ها نداشته باشد.

از دیگر نقدهای مهم به اصول اخلاق داده، جامع نبودن، ابهام، کلی‌گویی، و غیربافت‌مند بودن آن‌هاست (Tavani 2015). نکته اول اینکه مطالعات اخلاقی و خروجی‌های آن از جمله اصول اخلاقی، اساساً «نامتناهی، تأملی، و انتقادی» هستند. چون اخلاق و از جمله اخلاق داده حوزه‌ای مطالعاتی است که مسائل در آن باید ارزیابی، تفحص، بحث، تأمل، و استدلال شوند و بنابراین، فرایندی تأملی و استدلالی را ایجاد می‌کند. دستورات فهرست‌شده در اصول اخلاق داده ممکن است کارکنان را به این اشتباه بیندازد که همه چیز که لازم دارند این است که حسب موقعیت، یک اصل را پیدا و کورکورانه از آن تبعیت کنند. چون اصول اخلاقی موقعیت‌مند نیستند، و بنابراین به کارکنان گفته نمی‌شود که در موقعیت‌هایی که دو یا چند قاعده یا دستور در تعارض با هم قرار می‌گیرند، چه باید کرد. فرد در این‌گونه موقعیت‌ها نیاز به تأمل دارد، ولی اصول حرفه‌ای نوعاً سلسله‌مراتبی از قواعد یا سازوکاری برای انتخاب یک دستور به جای دستور دیگر در اختیار قرار نمی‌دهند. همچنین، با وجود این اصول، ابهاماتی در سطح اخلاق خُرد و اخلاق کلان به‌وجود می‌آید؛ زیرا کارکنان شرکت‌ها و سازمان‌ها به‌طور معمول، یک متخصص در یک رشته خاص هستند و به اصول اخلاق حرفه‌ای خاص خود متعهدند و به همین دلیل، این ابهام در اصول اخلاقی از جمله اخلاق داده است که کدام مسئولیت‌ها متوجه افراد حرفه‌مند است و کدام مسئولیت‌ها متوجه خود حرفه است و کدام‌ها مربوط

به سازمان است. مسائل اخلاقی خُرد دربارهٔ روابط شخصی بین حرفه‌مندان و سایر افراد نظیر ارباب رجوع است، ولی مسائل اخلاقی کلان دربارهٔ مسائل اجتماعی است که اعضای یک حرفه در مجموع، به صورت گروهی با آن مواجه می‌شوند. بنابراین، مسائل اخلاقی خُرد عبارت از به‌کارگیری مفاهیم اخلاقی معمول مثل صداقت و نزاکت است که در مواقع دیگر غیر از موقعیت‌های حرفه‌ای نیز لازم است. ولی مسائل اخلاقی کلان بسیار پیچیده‌اند؛ زیرا مستلزم تدوین سیاست‌هایی برای نظم اجتماعی‌اند. بنابراین، نیاز است بین پرسش‌هایی نظیر «من به‌عنوان یک متخصص حرفه‌ای چه مسئولیت‌هایی در فلان و فلان موقعیت دارم» و پرسش‌هایی مثل «حرفهٔ ایکس به‌عنوان یک حرفه چه مسئولیت‌هایی در فلان و فلان موقعیت دارد»، تمایز قایل شویم. ولی نکته اینجاست که همهٔ اصول اخلاقی از جمله اخلاق داده در ایجاد چنین تمایزی به ما کمک نمی‌کنند و فرد نمی‌تواند تمایزی بین اصول اخلاقی حرفهٔ خودش، اصول اخلاقی حرفه به صورت یک کل، و اصول اخلاق داده تمایزی قائل شود و آن‌ها را از هم تفکیک کند یا اولویت‌بندی نماید.

نکتهٔ آخر اینکه وصل کردن مجازات‌ها و شیوه‌های انضباطی به اصول اخلاقی برای ایجاد ضمانت اجرایی نیز اساساً آن‌ها را تبدیل به قواعد حقوقی یا «قواعد قهری» می‌کند، نه قواعد اخلاقی. نقش اخلاق عموماً ارزیابی، نقد، و حتی دفاع از قوانین، قواعد و مقررات است و نقش مجازات و تحمیل را ندارد. همچنین وقتی افراد مجبور می‌شوند از دستورات اطاعت کنند، از استقلال و قدرت انتخاب که عاملی ضروری برای اخلاقی اندیشیدن است، محروم می‌شوند. بنابراین اصول اخلاقی اگر بخواهد ضمانت اجرایی پیدا کند، آنگاه استوار بر فهم غلطی از ماهیت و هدف اخلاق می‌شود که سرانجام به مجموعه‌ای از آشفتگی‌های فکری و اخلاقی منتهی می‌شود. ولی اگر چنین نیز نشود، با توجه به جریان جاری عمل و روح زمانه، تبدیل به یک نسخه برای عمل نکردن می‌شود. در نتیجه، این مسئله همیشه هست که اصول اخلاقی، به‌ویژه در حوزه‌های جدیدی چون اخلاق داده، بیشتر اوقات به ابزاری نمایشی و بی‌اثر تبدیل می‌شوند که بیشتر به‌عنوان سرپوشی برای بهره‌برداری‌های غیر اخلاقی از منابع و داده‌ها استفاده می‌شوند. اصول اخلاقی این حوزه‌ها به‌طور معمول، ابهام‌آمیز، کلی، و بدون ضمانت اجرایی‌اند و در عمل، سازمان‌ها و شرکت‌ها را از رویه‌های غیر اخلاقی خود دور نمی‌کنند. در نظام اقتصادی و اجتماعی ما که رقابت و حفظ مزیت‌های اقتصادی اولویت دارد، پیاده‌سازی اصول اخلاقی در عمل دشوار و حتی گاهی متناقض با اهداف سازمان‌هاست. این تناقض‌ها و عدم تطابق اصول اخلاقی با واقعیت‌های

روزمره، به انفعال و بی‌تأثیری آن‌ها منجر می‌شود و اصول اخلاقی، به جای ایجاد تحول واقعی، در بسیاری موارد به ابزارهایی برای پوشاندن عیوب و جلوگیری از ایجاد تغییرات عمیق و مؤثر تبدیل می‌شوند.

فهرست منابع

- پولانی، کارل. ۱۹۴۴. *دگرگونی بزرگ: خاستگاه‌های سیاسی و اقتصادی روزگار ما*. ترجمه محمد مالجو. ۱۳۹۹. تهران: شیرازه.
- جام‌جم آنلاین. ۱۳۹۶. *جزئیات فروش اطلاعات خصوصی کاربران اینترنت در ایران*. جام‌جم آنلاین، شنبه ۹ دی‌ماه ۱۳۹۶. بازیابی شده از: <http://jjo.ir/cmzvvvca> (دسترسی در ۱۳۹۸/۶/۳۰)
- جمشیدی، بهاره، و حسین دهقانی سانج، حسین. ۱۳۹۹. *کلان داده‌های مبتنی بر اینترنت اشیاء از چشم‌انداز کشاورزی هوشمند. فصلنامه رشد فناوری ۶۳: ۱۲-۲۲*.
- چنگ، ها-جون. ۲۰۰۸. *نیکوکاران نابکار: افسانه تجارت خارجی آزاد و تاریخچه پنهان سرمایه‌داری*. ترجمه مهرداد (خلیل) شهابی و میرحمود نبوی. ۱۳۹۲. تهران: کتاب آمه.
- دادگستر، محمدرضا. ۱۴۰۲. *مسئله‌شناسی سیاست‌گذاری کلان‌داده در ایران. فصلنامه سیاست‌گذاری عمومی ۹ (۲): ۱۱۵-۱۲۷*.
- رهبری، ابراهیم. ۱۴۰۲. *تأملی بر رابطه کلان‌داده‌ها و قدرت بازاری در بازارهای پلتفرمی از منظر تحولات نوین حقوق رقابت. مطالعات حقوق خصوصی ۵۳ (۲): ۲۰۵-۲۲۸*.
- سعادت، زینب، و بتول مهرشاد. ۱۳۹۶. *اینترنت اشیاء و برنامه‌های کاربردی کلان‌داده‌ها در شهرهای هوشمند پایدار. سیاست‌نامه علم و فناوری ۷ (۳): ۱۷-۳۲*.
- عرب‌مازار یزدی، محمد، و امیر مرادی. ۱۳۹۹. *فرصت‌ها و چالش‌های کاربرد کلان‌داده‌ها در سیستم اطلاعاتی حسابداری بخش عمومی از منظر مسئولیت پاسخگویی. دانش‌حسابرسی ۲۰ (۲): ۹۵-۱۲۲*.
- فیروزی، محمدحسین، محمدرضا موحدی‌صفت، حامد حاجی‌ملا میرزایی، و سید مرتضی موسویان. ۱۴۰۰. *مدل مفهومی مدیریت کلان‌داده‌ها در رسانه‌های اجتماعی جمهوری اسلامی ایران براساس دیدگاه‌های مقام معظم رهبری پیرامون فضای مجازی. نشریه امنیت ملی ۳۹: ۳۵-۶۶*.
- قاسمی، حسین، و علیرضا ثقه‌الاسلامی. ۱۴۰۰. *تأثیر کلان‌داده‌ها بر تحول مفهومی عمل و عاملیت اخلاقی. اخلاق در علوم و فناوری ۱۶ (۲): ۱۸-۲۴*.
- گوروویچ، آرون. ۱۳۷۱. *پدیدارشناسی و فلسفه علم*. ترجمه حسین معصومی همدانی. *فصلنامه فرهنگ ۱۱: ۶۹-۹۰*.
- ملکی، داود، علیرضا منصوری، و محمدرضا احمدی. ۱۴۰۰. *نگرشی نو به آموزه‌های اخلاقی قرآن در مواجهه با چالش‌های اخلاقی حوزه کلان‌داده‌ها. قرآن و علم ۱۵ (۲۸): ۱۱۲-۱۴۶*.

میرعرب، علی، و رسول نوروزی فیروز. ۱۳۹۹. کلان‌داده و سیاست‌گذاری عمومی. *فصلنامه علوم سیاسی* ۲۴ (۱): ۱۵۱-۱۷۲.

میرعرب، علی، سیده لیلی میرطاهری، و سید امیر اصغری. ۱۴۰۰. ارائه مدلی برای خلق ارزش سازمانی با استفاده از تجزیه و تحلیل کلان‌داده. *آینده‌پژوهی مدیریت* ۳۲ (۳): ۷۰-۸۳.

هلبلی، خداداد، جلیل مظلوم، و بهرنگ هادیان. ۱۳۹۴. بررسی کاربردهای نظامی فناوری کلان‌داده و نقش آن در مدیریت صحنه نبرد. *فصلنامه علوم و فنون نظامی* ۳۳: ۴۷-۶۲.

References

- Arab-Mazar Yazdi, M. and A. Moradi. 2020. *Opportunities and Challenges of Applying Big Data in Public Sector Accounting Information Systems from an Accountability Perspective*. *Accounting Knowledge* 20 (2): 95-122. [in Persian]
- Acemoglu, D. and S. Johnson. 2023. *Power and Progress: Our 1000-year struggle over Technology and Prosperity*. New York: Hachette Book Group, Inc.
- Beck, U. 1992. *Risk Society: Towards a New Modernity*. London: Sage.
- Benjamin, R. 2019. *Race after Technology: Abolitionist Tools for the New Jim Code*. Cambridge: Polity books.
- Bowker, G.C., and S. L. Star. 2000. *Sorting things out: classification and its consequences*. Cambridge, Massachusetts: MIT press.
- Capurro, R. 2009. Past, present, and future of the concept of information. *Triple C* 7 (2): 125-141.
- Chang, H. J. 2008. *Bad Samaritans: the guilty secrets of rich nations and the threat to global prosperity*. First edition, New York: Manhattan: Random House Business. [in Persian]
- Crawford, K. 2013. The hidden biases in big data. *Harvard Business Review*. Retrieved from: <http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/> (accessed March 22, 2020)
- Diakopoulos, N. 2014. Algorithmic-accountability: The investigation of black boxes. Tow Center for Digital Journalism, Columbia University, Retrieved from: <https://academiccommons.columbia.edu/doi/10.7916/D8ZK5TW2> (accessed Feb. 25, 2025)
- Diaz, A. M. 2008. Through the Google Goggles: Sociopolitical Bias in Search Engine Design. In: *Web Search: Multidisciplinary Perspectives*. A. Spink and M. Zimmer (eds.), Berlin: Springer-Verlag, 11-34.
- Dodgson, K., P. Hirani, R. Trigwell, and G. Bueermann. 2020. *A Framework for the Ethical Use of Advanced Data Science Methods in the Humanitarian Sector*. Hague: Humanitarian Data Science and Ethics Group (DSEG).
- Espeland, W.N., and M. L. Stevens. 2008. A sociology of quantification. *European Journal of Sociology* 49 (3) 401- 436.
- Firoozi, M. H., M. R. Movahedisefat, H. Haji Mola Mirzaie and S. M. Mosavian. 2021. Conceptual model of management of clans in the social media of the Islamic Republic of Iran according to the Supreme Leader's views on cyberspace. *National Security Journal* 11 (39): 35-66. [in Persian]
- Floridi, L. & A. Taddeo. 2016. What is data ethics? *Philosophical Transactions of R. Soc, A* 374: 20160360. <http://dx.doi.org/10.1098/rsta.2016.0360>
- Floridi, L., & J. Cows. 2021. A unified framework of five principles for AI in society. In: *Ethics, Governance, and Policies in Artificial Intelligence*, Edited by Lociano Floridi, New York: Springer Verlag, pp. 5-17.

- Frické, M. 2015. Big data and its epistemology. *Journal of the Association of Information Science and Technology* 66 (4): 651-661.
- Fuches, C. 2006. Towards a critical theory of information. *TripleC* 7 (2): 243-292.
- Furner, J. 2004. Conceptual Analysis: A Method for Understanding Information as Evidence and Evidence as Information. *Archival Science* 4 (3-4): 233-265.
- Gamez, D. 2016). Are Information or Data Patterns Correlated with Consciousness? *Topoi* 35 (1): 225-239.
- Gawande, A. 2009. *The Checklist Manifesto; How to Get Things Right*. New York: Metropolitan Books.
- Gerlitz, C. and A. Helmond. 2013. The Like Economy: Social Buttons and the Data-Intensive Web. *New Media & Society* 15 (8): 1348-1365.
- Ghasemi, H. and A. Seghatoleslami. 2021. The Impact of Big data on the Conceptual Evolution of Moral Action and Agency. *Ethics in Science and Technology* 16 (2): 18-24. [in Persian]
- Gille Z. 2013. From risk to waste: global food waste regimes. *Sociological Review* 60 (S2): 27-46.
- Golebiewski, M., and D. Boyd. 2019. Data Voids: Where Missing Data Can Easily Be Exploited. Data & Society, Technical Report. Retrieved from: datasociety.net. (accessed Jan.19, 2025)
- Gurwitsch, A. 1978. Galilean Physics in the Light of Husserl's Phenomenology. In: *Phenomenology and Sociology*, ed. T. Luckmann, Penguin Books, London, 71-90. [in Persian]
- Gurstein, M. B. 2011. Open Data: Empowering the Empowered or Effective Data Use for Everyone? *First Monday*, 16 (2), Retrieved from: <https://firstmonday.org/article/view/3316/2764> (accessed Dec. 26, 2022)
- Halili, K. 2015. Investigate military applications of Big Data technology and its role in the management of battlefield. *Military Science and Tactics*, 11 (33): 47-62. [in Persian]
- Hagendorf, T. 2020. The ethics of AI ethics: An evaluation of guidelines. *Minds and Machine*, 30 (1): 99-120. <https://doi.org/10.1007/s11023-020-09517-8>.
- Harron K., C. Dibben, J. Boyd, A. Hjern, M. Azimae, M. L. Barreto, and H. Goldstein. 2017. Challenges in administrative data linkage for research. *Big Data & Society* 4 (2): 2053951717745678.
- Hartlapp, M., & G. Falkner. 2009. Problems of Operationalization and Data in EU Compliance Research. *European Union Politics* 10 (2): 281-304.
- Hasselbalch, G. 2019. Making sense of data ethics. The powers behind the data ethics debate in European policymaking, *Internet Policy Review* 8 (2): 1-19.
- Häußler, H. 2021. The Underlying Values of Data Ethics Frameworks: A Critical Analysis of Discourses and Power Structures. *Libri* 71 (4): 307-319. <https://doi.org/10.1515/libri-2021-0095>
- Iliadis, A. & F. Russo. 2016. Critical Data Studies, an Introduction. *Big Data and Society* 3 (2): 1-7.
- Jamejamonline News Agency. 2017. Details of selling personal information of internet users in Iran. Jamejamonline news agency, Saturday, December 30, 2017. Retrieved from: <http://jjo.ir/cmzvvvca> in September 21, 2019. [in Persian]
- Jamshidi, B. and H. Dehghanisani. 2020. Big IoT Data from the Perspective of Smart Agriculture. *Roshd -e- Fanavari* 16 (2): 12-22. [in Persian]
- Janssen, M., Y. Charalabidis, & A. Zuiderwijk. 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management (ISM)* 29 (4): 258-268.
- Johnson, J. A. 2014. From open data to information justice. *Ethics and Information Technology* 16: 263-274.
- Jurkiewicz, C. L. 2018. Big Data, Big Concerns: Ethics in the Digital Age. *Public Integrity* 20 (SUP1): 46-52.
- Kostrec, M. 2016. Methods of conceptual analysis. *Filozofia* 71 (3): 220-231.

- Levy, K. & D. M. Johns. 2016. When Open Data is Trojan Horse: The weaponization of Open Data in Science and Governance. *Big Data and Society* 3 (1): 1-6.
- Lipworth, W., P. H. Mason, I. Kerridge, & J. P. A. Ioannidis. 2017. Ethics and Epistemology in Big Data Research. *Bioethical Inquiry* 14 (4): 489-500.
- Loukides, M., H. Mason, & D. J. Patil. 2018. *Ethics and Data Science*. Sebastopol, CA: O'Reilly Media, Inc.
- Lupton, D. 2014. The commodification of patient opinion: the digital patient experience economy in the age of big data. *Sociology of Health and Illness* 36 (6): 856-869.
- _____. 2016. *The Quantified Self*. Cambridge: Polity books.
- McDonough, S. & R. Brandenburg. 2019. Who owns this data? Using dialogic reflection to examine an ethically important moment. *Reflective Practice* 20 (3): 355-366.
- Maleki, D., E. Arianyan, and M. R. Ahmadi. 2021. Categorizing and Presenting a New Perspective on the Moral Teachings of the Holy Quran in the Face of the Ethical Challenges of the Big Data. *The Quran and Science* 15 (28): 112-146. [in Persian]
- Margonelli, L. 1999. Inside AOL's "Cyber-Sweatshop". *Wired*. Retrieved from: <https://www.wired.com/1999/10/volunteers/> (accessed Dec. 27, 2022)
- Mason, Richard O. 2007. Four Ethical Issues of the Information Age. In J. Weckert, ed. *Computer Ethics*. Aldershot, UK: Ashgate, pp. 31-40. Reprinted from *MIS Quarterly* 10: 5-12.
- McGuire, M. 2010. Online surveillance and personal liberty. In: *Handbook of internet crime*. Yvonne Jewkes & Majid Yar (Eds.), New York: Routledge, 492-512.
- Mirarab, A. and R. Nowrozi Firooz. 2021. Big Data and Public Policy. *Political Science* 24 (1): 151-172. [in Persian]
- Mirarab, A., S. L. Mirtaheri and S. A. Asghari. 2021. A Model for Creating Organizational Value from Big Data Analytics. *Future study Management Journal* 32 (3): 70-83. [in Persian]
- Mittelstadt, B. 2017. From Individual to Group Privacy in Big Data Analytics. *Philosophy and Technology* 30 (4): 475-494. 10.1007/s13347-017-0253-7.
- _____. & L. Floridi. 2016. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics* 22 (2): 303-341.
- Mittelstadt, B. D., P. Allo, M. Taddeo, S. Wachter, & L. Floridi. 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society* 3 (2): 1-21.
- Moor, J. 1985. What is computer ethics? *Metaphilosophy* 16 (4): 266-275.
- Numerico, T. 2019. *Politics and epistemology of big data: A critical assessment*. In: *On the Cognitive, ethical and scientific dimensions of artificial intelligence*. D. Berkich and N. V. d'Alfonso (eds.). Second edition. Berlin: Springer.
- OECD 2020. Good Practices for Data Ethics in the Public Sector. Retrieved from: <https://www.oecd.org/gov/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.pdf> (accessed June 29, 2023)
- O'Neil, C. 2016. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown Publishers.
- Pipino, L. L., Y. W. Lee, and R. Y. Wang. 2002. Data Quality Assessment. *Communications of the ACM* 45 (4): 211-218.
- Polanyi, K. 1944. *The Great Transformation, the political and economic origins of our time*. Foreword by Robert M. MacIver. New York: Farrar & Rinehart. [in Persian]

- Rahbari, E. 2023. A Reflection on the Relationship between Big Data and Market Power in Platform Markets from the Perspective of the Latest Developments in Competition Law. *Law Quarterly* 53 (2): 205-228. [in Persian]
- Rogers, R. 2013. *Digital methods*. Cambridge, Massachusetts: MIT press.
- Rosenberg, D. 2013. Data before the Fact. In: *Row data is oxymoron*. Edited by Geoff rey Bowker and Paul N. Edwards. Cambridge, Massachusetts: MIT press. 25-40.
- Saadati, Z. and B. mehrshad. 2017. The Internet of Things and big data Applications in Sustainable Smart Cities. *Science and Technology Policy Letters* 07 (3): 17-30. [in Persian]
- Sandal, M. 2008. Justice: What is right to Do. Courses in Harvard University, Lecture 8. retrieved from: <https://scholar.harvard.edu/sandel/justice> (accessed Augst 8, 2020)
- Scholz, T. 2013. *Digital Labor: The Internet as Playground and Factory*. Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK: Routledge.
- Sebastian-Coleman, L., S. Earley, D. Henderson, E. Sykora, & E. Smith. 2017. *DAMA-DMBOK: Data Management Body of Knowledge*. Second edition. New Jersey: Basking Ridge.
- Spang-Hanssen, H. 2001. How to Teach about Information as Related to Documentation. *Human IT* 5 (1): 125-143.
- Stephens-Davidowitz, S. 2017. *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*. New York: Dey Street Books; Illustrated edition.
- Tavani, H. T. 2015. *Ethics and Technology; Controversies, Questions and and Strategies for Ethical Computing*. 5th edition, Danvers: Wiley.
- Tene, O., and J. Polonetsky. 2013. Big Data for All: Privacy and User Control in the Age of Analytics. *North Western Journal of Technology and Intercultural Property* 11 (5): 237-273.
- Tranberg, P., G. Hasselbalch, B. K. Olsen, & C. S. Byrne. 2018. *Data Ethics; Principles and Guidelines for Companies, Authorities & Organization*. UK, London: Spintype.com.
- UK Government Digital Services. 2020. Data ethics framework. Retrieved from: www.gov.uk/government/publications/data-ethics-framework (accessed July 22, 2022)
- Van Maanen, G. 2022. AI Ethics, Ethics Washing, and the Need to Politicize Data Ethics. *Digital society*, 1 (9), <https://doi.org/10.1007/s44206-022-00013-3>
- Watson, R. W. G., E. W. Kay, & D. Smith. 2010. Integrating biobanks: addressing the practical and ethical issues to deliver a valuable tool for cancer research. *Nature Reviews Cancer* 10 (9): 646-651.
- White House, Office of Management and Budget, President's Managerial Agenda. 2019. Federal Data Strategy. Retrieved from: <https://resources.data.gov/assets/documents/fds-data-ethics-framework.pdf> (accessed June 29, 2023)
- Wiggins, C., and M. L. Jones. 2023. *How Data Happened: A History from the Age of Reason to the Age of Algorithms*. New York: W. W. Norton & Company
- World Health Organization (2019). Analysis and Use of Health Facility Data. Retrieved December 2, 2022, from: <https://www.jstor.org/stable/resrep27953.8>
- Yallop, A. C. and Aliasghar, O. (2020). No business as usual: a case for data ethics and data governance in the age of coronavirus. *Online information Review*, 44 (6), 1217-1222.
- Zeffiro, A. (2021). From data ethics to data justice in/as pedagogy. *Studies in Social Justice*, 15 (3), 450-457.

مهدی شقاقی

متولد سال ۱۳۶۱، دارای مدرک تحصیلی دکتری در رشته علم اطلاعات و دانش‌شناسی از دانشگاه تهران است. ایشان هم‌اکنون استادیار گروه علم اطلاعات و دانش‌شناسی دانشگاه شهید بهشتی است.

اخلاق اطلاعات، اخلاق داده، جامعه‌شناسی اطلاعات، مطالعات نظری در علم اطلاعات، و روش‌شناسی تحقیق از جمله علایق پژوهشی وی است.



پژوهش نامه
پردازش و
مدیریت
اطلاعات