



Research Institute of
Hawzah and University

Methodology of Social Sciences and Humanities

Journal homepage: <https://method.rihu.ac.ir/>



Original Article

Big-Data and Epistemic Change

Kioomars Ashtarian¹

1. Associate Professor of Public Policy, University of Tehran, Tehran, Iran.

ashtrian@ut.ac.ir

Received: 2024/05/14; Accepted: 2024/12/11

Extended Abstract

Big data is an important achievement of the ICT revolution. In this article, big data is considered in terms of its impact on scientific research and the epistemology of knowledge. There are two main trends in this matter: one is that such a phenomenon leads to the abandonment of theory from research processes, and the other considers the role of theories to be still effective.

What we consider regarding the epistemological transformation caused by big data is somewhat different from what is stated in the relevant literature. Of course, it was presented in the critical analysis of researchers' points of view. Now we summarize it as follows:

1- The traditional division of disciplines has lost its importance or at least has been weakened due to the ocean of data. Data is independent of the field, and therefore interdisciplinary cooperation in the “agnostic” situation of science might be irrelevant. Of course, we are not of the agnostic belief that theory has lost its importance in the big-data age, as Chris Anderson claims in his article “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”. But without this agnostic approach, we can also accept that the role of theories, at least at the level of description, is greatly reduced, and therefore we can accept that data analysis is possible without disciplinary affiliation, and as a result, interdisciplinary collaborations are also shaky.

2- Data-driven science increases the role of correlation compared to causal theories and increases the role of researcher decision-making in the application of knowledge. Application means that a correlation is functionally important without depending on a specific discipline, and it can be used independently of any discipline; only to solve problems. The problems are not within the disciplines, but are located in the real world and impose themselves on us independently of the theories and disciplines. To solve a problem, numerous disciplines can be considered. These different perspectives mean that “a problem” is not dependent on any of these perspectives (disciplines). The problem of inflation can be seen from the perspective of economics, social psychology, political science, etc.

3- Weakening the role of theory and discipline in scientific development opens another field for scientific activities. New actors enter into scientific activities to the extent possible to play a competitive role with traditional researchers in scientific development. While the methodological changes produced by big data do not seem to be sufficient to invoke an entirely new paradigm in knowledge discovery, the emergence of big data has profoundly shaped the actors involved and





Research Institute of
Hawzah and University

Methodology of Social Sciences and Humanities

Journal homepage: <https://method.rihu.ac.ir/>



Original Article

their relationships. Moreover, even, the transition of scientific authority is predictable. In the age of big data, the shift of scientific authority from researchers and scientists (the followers and holders of grand theorists) to technicians, digital businessmen, and even ordinary users of big data can be expected. According to Chang et al., categories of mass data have emerged, which are caused by interactions, economies, interactions of societies or nations, and interactions of individuals. Public access to these data and especially facilitating the use of information technology tools can enable a large number of non-professional users in the field of research even with basic knowledge of statistics and data analysis in the field of science production. It is useful to remember that before the appearance of statistical software such as SPSS and SAS, performing a statistical analysis requires familiarity with programming skills. With the development of this software, more researchers were able to perform statistical operations because programming skills were removed from this process. Similarly, increased access to big data and analytical technologies resulting from new tools enable more "people" to analyze data and discover correlations between data for personal or business purposes or socio-cultural interactions. The production of public knowledge by people who are considered unprofessional according to "academic standards" is an important aspect in the evolution of scientific epistemology in the age of big data and analytical tools of information technology. The aggregative and generalized structure of scientific activity (if such a structure is important anymore) which was done by professional researchers and scientists through the structures of knowledge sharing and reproduction, will be performed by the machine. Knowledge activity will be displaced from its traditional institutional context. This is why I say that the implicit structure of authority that has been created over time by scientists and researchers is disrupted by "new revolutionaries" in the world of science. The institution of power in the production of science becomes shaky, scattered, multipolar, and rather rebellious. The problem is that in the previous paradigms, a kind of hierarchical structure of power or authority or scientific authority had emerged. The new paradigm targets exactly this hierarchical structure. Everyone can observe the big data and make inferences from it: No prior expertise; not just without prior theory; in an "agnostic platform" and without dependence on any field or discipline; without aristocracy of epistemological knowledge or any other theory and with a disruption from previous scholars and scientists.

4- Massification of actors raises the question of whether the accumulative, aggregative, and evolutionary characteristics of science, as seen in the history of science, do not undergo a radical transformation. Doesn't it go out of its institutional context, i.e. the field of "academy"? Microsoft, Apple, Facebook, Telegram, WhatsApp and many others have come to the field of innovation by people whose original innovators were often not "scientists" in the conventional sense, and sometimes even dropped out of university. These initiatives and innovations have been the source of the growth of science. The result is that scientific evolution has come from the academic environment to the business environment. Science has become more dependent on the market than before, and the "dynamism" of the market has become the source of the "dynamism" of science. Information technology has provided such a platform by facilitating communication and combining various technologies to provide a context for mutual leap. Science is neither interdisciplinary nor disciplinary but non-disciplinary. This can disrupt the entire structure of science that has dominated since the Renaissance.

Keywords: Big Data, Epistemology, Data Science, Agnostic.

Cite this article: Kioomars Ashtarian. (2024), "Big-Data and Epistemic Change", Methodology of Social Sciences and Humanities, 30(120): 1-19.



انبوه - داده؛ تحول معرفت‌شناسی

کیومرث اشتریان^۱

۱. دانشیار سیاست‌گذاری عمومی، دانشگاه تهران، تهران، ایران.

ashtrian@ut.ac.ir

تاریخ دریافت: ۱۴۰۳/۰۲/۲۵؛ تاریخ پذیرش: ۱۴۰۳/۰۹/۲۱

چکیده گسترده

انبوه - داده یکی از دستاوردهای مهم انقلاب فناوری اطلاعات و ارتباطات است که از حیث تأثیری که می‌تواند بر پژوهش‌های علمی بگذارد در این مقاله مورد توجه قرار گرفته است. در آنچه که به حوزه معرفت‌شناسی پژوهش‌های علمی مربوط می‌شود، دو گرایش عمده در این موضوع وجود دارد: یکی اینکه چنین پدیده‌ای به کنار گذاشتن تئوری از فرآیندهای پژوهش منجر می‌شود و دیگری در نقطه مقابل نقش تئوری‌ها را همچنان مؤثر می‌داند. انبوه - داده عبارت است از داده‌هایی که از گستره وسیعی از جریان فعالیت‌های اقتصادی، اجتماعی، زیست‌محیطی و... به صورت خودکار یا نیمه‌خودکار و با مشارکت فعال کاربران به وجود می‌آید. از شبکه‌های اجتماعی، وبگاه‌ها، وبلاگ‌ها، ژورنال‌یسم و «ردپای دیجیتال» افراد گرفته تا خدمات و تعاملات تجاری، اداری، فرهنگی، اجتماعی و... همگی عرصه‌ای پویا و گسترده را برای ایجاد انبوه - داده فراهم می‌کنند. انبوه - داده صرفاً حجم عظیم داده نیست؛ بلکه گستره وسیعی است که دارای تنوع، تکرار، سرعت، ارزش و انعطاف است. انبوه - داده زمینه دانش داده‌کاوی را فراهم کرده است که عبارت است از استخراج دانش‌های مفید از پایگاه‌های بزرگ داده یا از جریان داده‌ها که توسط روش‌های رایج، به دلیل گستردگی و پیچیدگی آنها قابل مدیریت نیستند. به دلیل حجم بالا، تنوع و سرعت زیاد انبوه - داده فنون جدیدی برای تحلیل آنها ضروری است. انبوه - داده ویژگی لحظه‌ای بودن دارد و از این رو، مشکل زمان را در پیمایش‌ها حل می‌کند. این یکی از مهم‌ترین ویژگی‌های انبوه - داده است. انبوه - داده و علم داده - ران بحث از آگنوستیک^۲ را پیش روی ما می‌گشاید. «آگنوستیک» که اساساً نوعی «لادری‌گری» یا ندانم‌گرایی است را می‌توان در ساحت معرفت‌شناسی به علم مستقل، علم بدون پیش‌فرض، یا علم قابل انطباق تعبیر کرد. در آگنوستیک ادعا بر این است که پیش‌فرضی از سوی پژوهشگر وجود ندارد و می‌گوید که نمی‌دانم چه چیزی هست یا نیست، درست است یا غلط، حقیقت دارد یا ندارد. این مبنایی برای ورود به عرصه دانش است. علم داده - ران نیز بر همین ادعا استوار است. از این رو، این دو مفهوم مرتبط با معرفت‌شناسی را در کنار هم آورده‌ایم. از آنجاکه داده‌های انبوه به‌خودی‌خود و بدون فرضیه یا تئوری ممکن است ما را در پژوهش علمی یاری دهند، بنابراین راه برای اقبال به آگنوستیسم فراهم می‌شود. به‌ویژه آنکه فناوری‌ها و روش‌هایی چون داده‌کاوی، یادگیری ماشین و هوش مصنوعی امکان پردازش داده‌های انبوه و استخراج الگوها را فراهم می‌کنند. این پدیده هنوز از حیث تأثیراتی که بر معرفت‌شناسی و نظریه‌پردازی دارد به‌اندازه کافی واکاوی نظری نشده است. مقاله حاضر، در پی بحث از همین ابعاد است.

روش: در این مقاله با روشی تحلیلی نخست، چارچوبی مفهومی از چند مقوله اصلی برساخته می‌شود که شامل انبوه - داده، علم داده-ران، آگنوستیسم است. این سه مفهوم اصلی با یکدیگر در علوم فناوری‌های نوین و در علوم اجتماعی ارتباط دارند. این چارچوب مفهومی از آن رو اهمیت دارد که با ترکیب علوم داده‌ای، علوم محاسباتی، علوم اجتماعی می‌تواند مبانی پارادایمی موجود را متحول کند و به تضعیف پارادایم علوم تجربی (پیشارنسانس)، علوم نظری (پیشارایانه) و علوم محاسباتی (پیشاداده انبوه) که پس از رنسانس مسلط بوده است بینجامد (کیچین، ۲۰۱۴).

نتایج: این مقاله استدلال می‌کند که انبوه - داده ظرفیت آن را دارد که تحولی عمده را در معرفت‌شناسی علوم اجتماعی پدید آورد. از سویی پدیده انبوه - داده بازیگران جدیدی را که پژوهشگران سنتی علوم اجتماعی نیستند، وارد توسعه علمی و معرفت‌شناسی می‌کند که شامل فناوری‌ها و حوزه اطلاعات و





نوع مقاله: پژوهشی

ارتباطات و نیز کارآفرینان اقتصادی می‌شود. از سوی دیگر، نقش رشته‌های دانشگاهی و تئوری‌ها در توسعه علمی کاهش می‌یابد. علم داده-ران می‌تواند برنامه‌های پژوهشی را سامان‌دهی کند؛ به گونه‌ای که نقش نظریه‌ها و تقسیم‌بندی‌های رشته‌های سنتی در دانشگاه را محدود نماید. پدیده انبوه - داده به فردگرایی روش‌شناختی در علوم اجتماعی دامن می‌زند. از این رو، تجزیه و تحلیل سیاسی مبتنی بر ابزارهای فناوری اطلاعات یا علوم اجتماعی محاسباتی در معرض توجه بیشتری قرار گرفته است.

بر اساس این، پایه‌های معرفت جدید بر درک فعالیت‌های خرد فردی و فردگرایی روش‌شناختی بنا گذاشته می‌شود. «شناخت ترکیبی فردمحور» امکان وقوع می‌یابد که حاصل سپهرهای گوناگون است. به‌دیگر سخن، شناخت بهتر فرد زمینه توسعه نظریه‌های فردمحور را امکان‌پذیر می‌کند. شناخت داده‌های متکثر، ضرورتاً به بی‌نیازی از نظریه نمی‌انجامد. تحلیل نظری گریبان انسان را رها نمی‌کند؛ زیرا این نظریه‌ها بخشی جدانشدنی از تجربه‌ها، سلیقه‌ها، ترجیحات و ادراکات روحی و روانی ما هستند. اما تحول اساسی اینجاست که نقش داده‌های خرد فردی و در مقیاس «نانورفتار» توسعه نظریه‌های فردمحور - داده محور را بیشتر می‌کند. به همین سان، نقش داده‌های فردی و نقش سوژه در شکل‌گیری نظریه بیش از نقش پژوهشگر - فاعل نظریه پرداز می‌شود؛ زیرا این داده‌ها به‌رحال خود، کم‌وبیش سخن می‌گویند و نقش‌آفرینی بیشتری در فهم و در شکل‌گیری ذهن نظریه‌پرداز ایفا می‌کنند. نظریه‌های گوناگونی که از داده‌ها برمی‌آیند، بدینسان، از تعامل‌پذیری بیشتری برخوردارند؛ زیرا سهم هریک از عوامل و مؤلفه‌ها به‌صورت محسوس‌تری دیده شده و تصدیق می‌شود.

از این رو، در نگاه نخست، زمینه برای تلفیق نظریه‌های گوناگون و شکل‌گیری کلان نظریه‌های «متناسب‌تر» بیش‌ازپیش فراهم می‌شود. به‌عنوان مثال، آنگاه که داده‌هایی طبقه‌بندی‌شده و شفاف از پدیده فقر و سهم ساختارها و رفتارها در آن شناسایی شود و آنگاه که داده‌هایی روشن و طبقه‌بندی‌شده از توانمندسازی انسان، از ابعاد گوناگون محیط‌زیست، از نقش سیاست دولتی در این زمینه‌ها، از عوامل مؤثر بر علم، فناوری و نوآوری در ابعاد و اجزای گوناگون و... به دست آید، زمینه برای نظریه‌پردازی‌های خرد در هریک از این موضوعات و در نتیجه نظریه‌ای جامع فراهم می‌آید؛ البته در نگاه عمیق‌تر تحولی پارادایمی در توسعه دانش بدون رشته فراهم می‌آورد که خود می‌تواند موجد اغتشاش نظری باشد.

بحث و نتیجه‌گیری: اگرچه دقیقاً نمی‌توان گفت که نقش نظریه‌ها پایان یافته است، اما این انقلاب و جوهی دارد که نقشی بنیانی در تغییر پارادایم دارد؛ از طریق گسست فعالیت علمی از نظریه‌پردازی، با از بین بردن ساختار سلسله‌مراتبی دانشمندان، با تضعیف ساختار رشته‌ای، با توده‌ای کردن بازیگران دانش، با گسست تولید علم از زمینه رشته‌ای و از زمینه متخصصان، و با پدید آوردن دانش غیرتداومی. تقسیم‌بندی رشته‌ای به دلیل غرق شدن در اقیانوسی از داده‌ها اهمیت خود را از دست داده یا اینکه حداقل متزلزل شده است. داده مستقل از رشته است و از این رو، همکاری بین‌رشته‌ای در فضای «شبه - آگنوستیکی» از علم می‌تواند بلاموضوع باشد.

در علم در دست کارشناسان فنی، تکنیسین‌ها و نه ضرورتاً دانشمندان به‌صورت «آنارشیک» رشد می‌یابد. امروزه، تحولات نوآورانه‌ای در حوزه فناوری‌های نوین از سوی جوانان (و گاه نوجوانان)، که دوره‌های تحصیلات تکمیلی را نگذرانده‌اند، مشاهده می‌شود. مایکروسافت، اپل، فیسبوک، تلگرام، واتس‌اپ و بسیاری دیگر از این دست توسط کسانی به عرصه نوآوری آمده‌اند که ابداع‌کنندگان اولیه اغلب «دانشمند» به معنای مصطلح نبوده‌اند و حتی گاه از دانشگاه ترک تحصیل کرده‌اند. این ابتکارات و ابداعات خود سرمنشأ رشد علم بوده‌اند. نتیجه آنکه تحول علمی از محیط آکادمی به محیط کسب‌وکار آمده است. علم بیش از گذشته به بازار وابسته شده است و «دینامیسم» بازار سرمنشأ «دینامیسم» علم شده است. فناوری اطلاعات چنین بستری را از طریق تسهیل ارتباطات و ترکیب فناوری‌های گوناگون با یکدیگر زمینه جهش متقابل را فراهم کرده است و علم نه ایتردیسیپلینری است و نه دیسیپلینری بلکه بی‌دیسیپلین است و این می‌تواند کل ساختار پیشین از رنسانس به این سو را به هم زند.

واژگان کلیدی: انبوه - داده، معرفت‌شناسی، علم داده-ران، آگنوستیک.

۱. مقدمه

انبوه - داده^۱ یکی از پدیده‌های مهم و حیاتی است که به مدد فناوری‌های پردازشگر اطلاعات، با انفجار اطلاعات و ظرفیت دسترسی به داده‌های وسیع گسترش یافته است. این پدیده هنوز از حیث تأثیراتی که بر معرفت‌شناسی و نظریه‌پردازی دارد، به اندازه کافی واکاوی نظری نشده است. مقاله حاضر، در پی بحث از همین ابعاد است. پدیده انبوه - داده به فردگرایی روش‌شناختی^۲ در علوم اجتماعی دامن می‌زند. از این رو، تجزیه و تحلیل سیاستی مبتنی بر ابزارهای فناوری اطلاعات^۳ یا علوم اجتماعی محاسباتی^۴ در معرض توجه بیشتری قرار گرفته است. انبوه - داده می‌تواند به مثابه تحولی شگرف در معرفت‌شناسی علمی نیز مورد توجه قرار گیرد. یکی از پیشگامان نظریه‌پردازی در این حوزه تا آنجا پیش می‌رود که عصر انبوه - داده را عصر مرگ نظریه تلقی می‌کند (Anderson, 2008).

نظریه معرفت‌شناسی اندرسون مورد نقد بسیار قرار گرفته است؛ اما در راستای تحول مورد ادعای او استدلال‌ها و ابعاد دیگری را مطرح کرده‌ایم. استدلال این مقاله آن است که اگرچه دقیقاً نمی‌توان گفت که نقش نظریه‌ها پایان یافته است، اما این انقلاب و جوهی دارد که نقشی بنیانی در تغییر پارادایم دارد. از طریق گسست فعالیت علمی از نظریه‌پردازی، با از بین بردن ساختار سلسله‌مراتبی دانشمندان، با تضعیف ساختار رشته‌ای^۵، با توده‌ای کردن بازیگران دانش، با گسست تولید علم از زمینه رشته‌ای و از زمینه متخصصان، و با پدید آوردن دانش غیرتداومی^۶.

۲. روش

در این مقاله با روشی تحلیلی نخست، چارچوبی مفهومی از چند مقوله اصلی بر ساخته می‌شود که شامل انبوه - داده، علم داده - ران، آگنوستیسم است. این سه مفهوم اصلی با یکدیگر در علوم فناوری‌های نوین و در علوم اجتماعی ارتباط دارند. این چارچوب مفهومی از آن رو اهمیت دارد که با ترکیب علوم داده‌ای، علوم محاسباتی، علوم اجتماعی می‌تواند مبانی پارادایمی موجود را متحول کند و به تضعیف پارادایم علوم تجربی (پیش‌رسانس)، علوم نظری (پیش‌رایانه) و علوم محاسباتی (پیش‌داده انبوه) که پس از رنسانس مسلط بوده است بینجامد (کیچین، ۲۰۱۴). سپس عناصر اصلی این چارچوب مفهومی برای نشان دادن تحول معرفت‌شناسانه ناشی از انبوه - داده از طریق مرور ادبیات بررسی می‌شود. سپس در پایان استدلال‌های خود را درباره این مطرح کرده‌ایم که آیا انبوه - داده به تحولی پارادایمی در علم منجر می‌شود یا خیر.

۳. پیشینه پژوهش: انبوه - داده و زمینه‌های یک چارچوب مفهومی

انبوه - داده به فردگرایی روش‌شناختی^۷ در علوم اجتماعی دامن می‌زند و دارای ادبیات گسترده‌ای است. از این رو، تجزیه و تحلیل سیاستی مبتنی بر ابزارهای فناوری اطلاعات^۸ یا علوم اجتماعی محاسباتی^۹ در معرض توجه بیشتری قرار گرفته است (Bertoni et al. 2023). انبوه - داده می‌تواند به مثابه تحولی شگرف در معرفت‌شناسی علمی مورد توجه قرار گیرد. استلپتر در تحلیل ظرفیت‌ها و مسائل کلان داده‌ها در نظریه‌ها و نظریه‌سازی در علوم اجتماعی پنج جریان استدلالی را شناسایی کرده است:

۱. توسعه و استفاده از نظریه‌های سیستم پیچیده می‌تواند واکنش مناسبی به داده‌های بزرگ باشد؛ هرچند پدیده‌های اجتماعی ممکن است به طور معتبر با استفاده از سیستم‌های رسمی پیچیده ترسیم نشوند؛

۲. برخی از پژوهشگران هیچ مفهوم جدیدی را ناشی از انبوه - داده برای نظریه‌ها در علوم اجتماعی فرض نمی‌کنند؛ زیرا داده‌های بزرگ

1. Big Data

2. Methodological Individualism

3. Policy Analytic

4. Computational Social Science/CSS

5. disciplinary

6. discontinuity

7. Methodological Individualism

8. Policy Analytic

9. Computational Social Science/CSS

فقط به معنای داده‌های بیشتر است؛

۳. برخی از نویسندگان اشاره می‌کنند که داده‌های بزرگ بد می‌تواند به نظریه‌های بد منجر شود. بنابراین، کلان‌داده باید تحت یک فرآیند

ارزیابی پیشرفته قرار گیرد که خود براساس معیارهای مرتبط با داده‌های اضافی و همچنین براساس نظریه‌های موجود باشد؛

۴. نظریه‌های مرتبط با انبوه - داده‌ها باید استانداردهای گسترده‌ای را برای توسعه و آزمون تئوری براساس نوع [انواع] خاصی از

چارچوب نظری برآورده کنند. باوجود این، چنین استانداردهای توسعه‌یافته با مشکلات ناسازگاری، تقلیل،^۱ پیشرفت، رویکردهای ترکیبی، و پارادایم‌ها و رشته‌های متعدد همراه است؛

۵. داده‌های بزرگ ممکن است از طریق داده‌کاوی شناسایی پدیده‌های اجتماعی را امکان‌پذیر کند که توسط تخیل انسان کشف

نمی‌شوند. باوجود این، یک آزمون قطعی برای تعیین این هنوز وجود ندارد که آیا چنین ابزارها و رویه‌های مرتبط با داده‌های بزرگ مؤثر هستند یا حتی مؤثرتر از رویکردهای سنتی‌تر برای تئوری‌سازی اند (Astleitner, 2024, p. 83-84).

از حیث معرفت‌شناسی برخی نقش تئوری را در فرآیندهای پژوهش علمی همچنان مؤثر می‌دانند و برخی دیگر معتقدند دسترسی به

داده‌های انبوه می‌تواند به کنار گذاشتن تئوری از فرآیندهای پژوهش منجر شود و عصر انبوه - داده را عصر مرگ نظریه تلقی می‌کند

(Anderson, 2008). در این دیدگاه دوران توسعه علم براساس نظریه‌ها و فرضیه‌ها و شکل‌گیری پارادایم‌ها براساس نظریه‌هایی به سر آمده است

که برنامه‌های پژوهشی لاکاتوشی فراهم می‌کنند. داده‌ها این وظیفه سنگین را برعهده می‌گیرند و این امکان را فراهم می‌کنند که تصمیمات

مقتضی در سیاست‌های عمومی، در پژوهش و در فرآیندهای تولید اخذ شود (McCue & M Mccoy, 2017). در گفتمان‌های پوزیتیویستی،

انبوه - داده همچون پدیده‌ای مطرح می‌شود که هر جنبه‌ای از زندگی فردی و اجتماعی و در واقع، شیوه‌های سنتی شناخت جهان را به‌گونه‌ای

اساسی تغییر می‌دهد. این تغییرات معمولاً در سه محور خلاصه می‌شوند: اینکه همبستگی می‌تواند جایگزین جستجوی علیت شود؛ اینکه

ارزش اکتشافی داده‌ها نتیجه حجم مطلق آن است؛ و اینکه مجموعه داده‌های انباشته شده از منابع ناهمگن را می‌توان جدا از زمینه‌های تولید،

انتشار و مصرف آنها تحلیل کرد (Resnyansky, 2019, p. 9).

علم داده-ران، اگنوستیک، دانش پلتفرمی، اگنوستیک فناورانه، فردگرایی روش‌شناختی، استراتژی پژوهشی داده-محور و شماری دیگر

از این دست مفاهیم منظومه‌ای معنایی را در پیشینه پژوهش فراهم کرده است. این دیدگاه در مقابل دیدگاه لاکاتوش است که نظریه‌ها به‌مثابه

کل‌های ساختاری راهنمونی برای برنامه‌های پژوهشی (پارادایم) فراهم می‌کنند. «اگر نظریه‌ها به‌گونه‌ای انتظام یافته باشند که از نظر چگونگی

نحوه تحول و توسعه‌شان، حاوی سرخ‌ها و توصیه‌هایی باشند، علم چالاک‌تر پیشرفت خواهد کرد. نظریه‌ها باید ساختارهای بی‌انتهایی باشند،

به‌گونه‌ای که برنامه‌ای پژوهشی در اختیار نهند» (چالمرز، ۱۳۸۷، ص ۹۷).

در مقابل، اگرچه چانگ و همکاران علیه این باور که با پدیده انبوه - داده «دیگر نیازی به تئوری نیست» اقامه دلیل می‌کنند، اما همچنین

مطرح می‌کنند که باید فلسفه علم با تحولات زمانه تغییر کند (Chang et al, 2013). البته در بخش دوم این مقاله نشان خواهیم داد که آنان

این تغییر معرفت‌شناسی را در تحول مهم به‌سوی علوم اجتماعی محاسباتی (CSS) خلاصه می‌کنند. در همین راستا، کیچین در مقاله

«انبوه - داده، معرفت‌شناسی‌های جدید و تحولات پارادایمی» چنین بحث می‌کند که کشف اشکال جدید تجربه‌گرایی، علم داده-ران و

علوم انسانی رقومی (دیجیتال) و علوم اجتماعی محاسباتی راهی کاملاً متفاوت را برای معنابخشی به فرهنگ، تاریخ، اقتصاد و جامعه در پیش

می‌گیرد. این خود یک نوآوری گسستی است که هدایت پژوهش را دستخوش تحول می‌کند. او همچنان بر این باور است که نظریه می‌تواند

ما را در استراتژی پژوهش کمک کند؛ هرچند که ممکن است برای کشف روابط علی کارآمدی لازم را نداشته باشد؛ اما تفکر انتقادی

درخصوص پیامدهای معرفت‌شناسانه انقلاب داده‌ای ضروری است (Kitchin, 2014). ما به این مباحث در این مقاله به‌تفصیل خواهیم

پرداخت؛ اما از آنجاکه این تحول معرفت‌شناسی خود ناشی از یک منظومه فناورانه - مفهومی است، از این رو، ادامه پیشینه پژوهش را در

قالب چارچوب مفهومی نشان خواهیم داد.

۳-۱. انبوه - داده

انبوه - داده عبارت است از داده‌هایی که از گستره وسیعی از جریان فعالیت‌های اقتصادی، اجتماعی، زیست‌محیطی و... به صورت خودکار یا نیمه‌خودکار و با مشارکت فعال کاربران به وجود می‌آید. از شبکه‌های اجتماعی، وبگاه‌ها، وبلاگ‌ها، ژورنالیسم و «ردپای دیجیتال» افراد گرفته تا خدمات و تعاملات تجاری، اداری، فرهنگی، اجتماعی و... همگی عرصه‌ای پویا و گسترده را برای ایجاد انبوه - داده فراهم می‌کنند. داده به سادگی و به ارزانی تولید می‌شود و رایانش ابری داده‌ها به ما امکان تبدیل آن به اطلاعات و سپس دانش را فراهم می‌آورد.

انبوه - داده صرفاً حجم عظیم داده نیست؛ بلکه گستره وسیعی است که دارای تنوع، تکرار، سرعت، ارزش و انعطاف است. این داده‌ها در انبارهای داده‌ای بخش‌های مختلف خصوصی و عمومی پراکنده‌اند. فناوری اطلاعات و ارتباطات، به‌ویژه شکل‌گیری ابرهای اطلاعاتی، امکان تجمیع گسترده این داده‌ها را فراهم آورده و رایانش ابری^۱ امکان تحلیل‌ها و بهره‌برداری‌های گوناگون از آن را فراهم آورده است. انبوه - داده زمینه دانش داده‌کاوی را فراهم کرده است که عبارت است از استخراج دانش‌های مفید از پایگاه‌های بزرگ داده یا از جریان داده‌ها که توسط روش‌های رایج، به دلیل گستردگی و پیچیدگی‌شان، قابل مدیریت نیستند. به دلیل حجم بالا، تنوع و سرعت زیاد انبوه - داده فنون جدیدی برای تحلیل آنها ضروری است. انبوه - داده ویژگی لحظه‌ای بودن دارد و از این رو، مشکل زمان را در پیمایش‌ها حل می‌کند که این یکی از مهم‌ترین ویژگی‌های انبوه - داده است.

۳-۲. علم داده - ران^۲

در ادبیات مربوط به انبوه - داده سخن از علم داده - ران به میان آمده است. مقصود این است که علم از طریق داده‌ها و نه ضرورتاً از طریق نظریه‌ها پیش می‌رود. براساس پردازش حجم زیادی از داده‌ها به وسیله فناوری‌های جدید می‌توان روابط پنهان را کشف کرد که تا پیش از این امکان‌پذیر نبود. به جای ابتدا بر فرضیه و نظریه پژوهشگر از داده‌ها برای ایجاد فرضیه‌های خاص بهره می‌برد. علم داده به ما این امکان را می‌دهد که تصمیمات مقتضی در سیاست‌گذاری و در کسب‌وکار و در پژوهش و در بهینه‌سازی فرآیندهای تولید اتخاذ نماییم (McCue & M Mccoy, 2017).

۳-۳. اگنوستیک

انبوه - داده و علم داده - ران بحث از اگنوستیک^۳ را پیش روی ما می‌گشاید. «اگنوستیک» که اساساً نوعی «لاادری‌گری» یا ندانم‌گرایی است را می‌توان در ساحت معرفت‌شناسی به علم مستقل، علم بدون پیش‌فرض، یا علم قابل انطباق تعبیر کرد. در اگنوستیک ادعا بر این است که پیش‌فرضی از سوی پژوهشگر وجود ندارد و می‌گوید که نمی‌دانم چه چیزی هست یا نیست، درست است یا غلط، حقیقت دارد یا ندارد. این مبنایی برای ورود به عرصه دانش است. علم داده - ران نیز بر همین ادعا استوار است. از این رو، این دو مفهوم مرتبط با معرفت‌شناسی را در کنار هم آورده‌ایم. از آنجاکه داده‌های انبوه به خودی خود و بدون فرضیه یا تئوری ممکن است ما را در پژوهش علمی یاری دهند، بنابراین راه برای اقبال به اگنوستیسم فراهم می‌شود. به‌ویژه آنکه فناوری‌ها و روش‌هایی چون داده‌کاوی، یادگیری ماشین و هوش مصنوعی امکان پردازش داده‌های انبوه و استخراج الگوها را فراهم می‌کنند.

در اینجا با توجه به ارتباطی که اگنوستیک با برنامه‌نویسی دارد، مختصراً اشاره می‌شود که در محاسبات رایانه‌ای، اگر روش یا «فرمت» انتقال داده‌ها به عملکرد یک دستگاه یا یک برنامه تناسب نداشته باشد یا مستقل از آن باشد، می‌گوییم که آن دستگاه یا برنامه «داده - اگنوستیک» است؛ یعنی نسبت به آن خنثی و بی‌طرف و منعطف است. یک سیستم به‌گونه‌ای طراحی می‌شود که مستقل از فناوری‌ها یا پلتفرم‌های خاص باشد؛ یعنی دستگاه یا برنامه می‌تواند داده‌ها را در قالب‌ها و منابع گوناگونی دریافت و به‌طور مؤثر پردازش کند. از سوی دیگر، اما فرایند

1. Cloud Computing
2. Data-driven Science
3. Agnostic

آگنوستیکی نیازی به فهم فرآیند زیربنایی واقعی ندارد. آنچه که برای مقصود ما مهم است آن است که به زبان علوم اجتماعی، تجزیه و تحلیل داده‌ها می‌تواند بدون درک فرضیه‌ها و نظریه‌های بنیانی و با مدل جعبه سیاه انجام شود.

«آگنوستیک فناوری»^۱ و «پلتفرم آگنوستیک»^۲ مفاهیم دیگری هستند که برای هدف این مقاله اهمیت دارد. آگنوستیک فناوری رویکردی است که به توسعه‌دهندگان آزادی عمل می‌دهد تا راه‌حل‌ها را بدون استفاده از زبان‌های برنامه‌نویسی خاصی توسعه دهند. در یادگیری ماشینی آگنوستیک رویکردی است که در سراسر سیستم‌ها قابل اجراست و هیچ‌گونه پیش‌داوری نسبت به استفاده از یک فناوری، مدل، روش یا داده خاص وجود ندارد. مفهوم «پلتفرم آگنوستیک» (به مجموعه‌ای از ویژگی‌ها و فلسفه‌های طراحی خاص اشاره دارد که معمولاً به محصولات نرم‌افزاری مرتبط هستند. یک نوع نرم‌افزار آگنوستیک و در واقع، عاری از هرگونه ارتباط با یک پلتفرم یا سیستم خاص است و به همان اندازه در بیش از یک پلتفرم به خوبی اجرا می‌شود) (Netamorphosis, online).

انبوه - داده می‌تواند به تولید دانشی منجر شود که آن را «دانش پلتفرمی» می‌نامیم. نرم‌افزار آگنوستیک نرم‌افزاری است که بر روی چندین سیستم عامل مانند Windows، MacOS یا Linux بدون نیاز به تغییرات گسترده کار می‌کند. به همین سان آگنوستیک در امور مالی یعنی یک سرمایه‌گذار ناشناس ارزش‌گذاری،^۳ هنگام ارزیابی یک شرکت، در مقایسه با سود پیش‌بینی شده هر سهم آن، معیارهای ارزیابی را در نظر نمی‌گیرد؛ همچون قیمت سهام یک شرکت. اهمیت مفهوم آگنوستیک به‌گونه‌ای است که در بازاریابی هم استفاده می‌شود. «برند آگنوستیک» (معمولاً به یک استراتژی بازاریابی یا تجاری اشاره دارد که نسبت به برند یا برندهای خاص بی‌طرف است. مهم‌تر از همه، «هدف» به این معنا نیست که شما هیچ دیدگاهی در مورد برندها ندارید) (Schröder, 2022).

۳-۴. انبوه - داده و تحول معرفت‌شناسانه

تحول معرفت‌شناسانه در انبوه - داده دو جریان فکری متمایز دارد. نخستین جریان فکری، در افراطی‌ترین روایت، بر این موضع است که انبوه - داده سبب می‌شود تا تئوری‌ها موضوعیت خود را در بنای علم از دست بدهند و به تعبیری دوران مرگ تئوری‌ها فرا رسیده است؛ زیرا داده‌ها کاشف از واقعیت‌اند (Anderson, 2008). دومین جریان فکری بر اهمیت انبوه - داده تأکید دارد؛ اما بر این باور است که همچنان نظریه‌ها اهمیت خود را حفظ می‌کنند. در انتخاب متغیرها، در مدل‌سازی، در تحلیل پدیده‌ها و... این نظریه است که راهنمای پژوهش است. در این مقاله، ما استدلال‌های دیگری مطرح می‌کنیم که شاید بیشتر به جریان نخست نزدیک‌تر است:

۱. انبوه - داده فضای اقتدار علمی را دگرگون کرده است و بازیگران جدید توده‌ای را به عرصه تحولات علمی وارد کرده است. دیگر، این متخصصان و پژوهشگران و دانشمندان نیستند که نقش دروازه‌بانی دانش (همچون دروازه‌بانی خبر) را از طریق سازوکارها و تشریفات متدلوژیک و ژورنال‌های تخصصی ایفا نمایند. شهر علم بی‌دروازه و بدون دیسپلین است. هرکسی از هر رشته‌ای و بلکه بدون رشته می‌تواند داده‌های انبوه را به کار گیرد و از درون آن به تولید علم پردازد. مثلاً انبوه - داده سلامت می‌تواند از سوی صاحبان کسب‌وکارهای دیجیتال و با هدف توسعه دانش پزشکی به کار گرفته شود. یک برنامه کاربردی^۴ که در ابتدا از سوی کسب‌وکارهای سلامت و برای نوبت‌دهی یا برای رصد بیماران دیابتی طراحی شده است می‌تواند بدون حضور پزشک متخصص دانش پزشکی دیابت را از حیث تأثیر داروها بر گونه‌های مختلف دیابت توسعه دهد. در عین حال می‌تواند دانش برنامه‌ریزی دولتی را برای مواجهه با بیماری دیابت توسعه دهد. مثلاً نشان دهد که برنامه‌های بیمه چگونه تنظیم شود تا بیشترین کارایی لازم را برای رسیدگی به بیماران دیابتی داشته باشد. این از طریق صاحبان کسب‌وکار محقق می‌شود که ضرورتاً دانش پزشکی یا دانش سیاست‌گذاری و مدیریت ندارند؛ بلکه تأثیر انبوه - داده‌ای است که توسط صاحبان کسب‌وکار پدید آمده و توسعه یافته است. به‌دیگر سخن، صاحبان کسب‌وکار در عمل، به «اپراتور» و توسعه‌دهنده دانش سلامت و دانش برنامه‌ریزی عمومی برای سلامت تبدیل می‌شوند. این کسب‌وکار دیجیتال است که نقش آزمایشگاه و دانشکده پزشکی یا دانشکده

1. technology agnosticism

2. platform agnostic

3. a valuation agnostic investor

4. application

سیاست‌گذاری عمومی را «نیز» ایفا می‌کند؛

۲. نظریه‌ها بیشتر مبتنی بر داده‌ها و کمتر مبتنی بر تأملات فلسفی یا هستی‌شناسی کلان هستند. آنگاه که پژوهشگر داده‌های کافی (و بلکه انبوه) از تورم، اشتغال، حجم نقدینگی، تولید کشاورزی و... داشته باشند، تأثیرات تنظیم‌گری بر تولیدات کشاورزی از طریق همین داده‌ها «حاصل» است و ضرورتاً نیازی به نظریه صنعتی یا نظریه اشتغال ندارد؛

۳. کاربرد نظریه‌ها بیشتر به حوزه‌های هنجاری (چه باید کرد یا جامعه مطلوب و روابط انسانی مطلوب چگونه باید باشد) محدود می‌شود؛ یعنی نظریه‌ای عدالت‌گرا یا نظریه مسئولیت اجتماعی انسان‌مدار جهت‌گیری سیاست‌ها را تعیین می‌کند تا مثلاً رفع بی‌عدالتی ساختاری راهنمای سیاست‌های حمایتی از فقرا باشد. نظریه‌ها کاربرد بسیار محدودی در تبیین آنچه که هست خواهند داشت؛ چون انبوه - داده خود آنچه که هست را کم‌وبیش نشان می‌دهد؛

۴. انقلاب‌های علمی موضوعیت ندارند؛ زیرا دیگر نظمی وجود ندارد که نیازمند انقلابی باشد! آنچه که در پی خواهد آمد، نگاهی تحلیلی از این دو جریان است.

۳-۵. انبوه - داده و فردگرایی روش‌شناسی: نقش‌آفرینی بیشتر داده‌ها در شکل‌گیری نظریه

انبوه - داده تحولی را در روش‌شناسی جمع‌آوری داده‌ها و رویکرد ما به شاخص‌ها در حوزه‌های مختلف علمی از علوم‌رایانه تا علوم اجتماعی پدید آورده است. در این عرصه، داده‌های پیمایشی می‌توانند به تدریج جای خود را به انبوه - داده بدهند و چالش‌های نمونه‌گیری را برطرف کنند. همواره این مشکل وجود داشته است که به دلیل تکرار جامعه هدف، پژوهشگران ناچار بوده‌اند که به نمونه‌های محدودی بسنده کنند. بدینسان روش‌هایی برای اعتبارسنجی، اعتماد به نمونه‌ها و معرف بودن نمونه طراحی و پیاده شده است. با انبوه - داده امکان گذار از روش‌های اعتبارسنجی آماری، و روش‌های جمع‌آوری پیمایشی و دوره‌ای داده‌ها وجود دارد. البته این روش‌ها همچنان اعتبار خود را دارند و خدشه‌ای بر اعتبار آنها نیست. اما در عصر انبوه - داده می‌توان از طریق حسگرها، ربات‌ها، برنامه‌های کاربردی و... در بسیاری از موضوعات، اطلاعات کل جامعه هدف را جمع‌آوری کرد. بدینسان هر فرد واحد در داده‌برداری‌ها نقش‌آفرین است و به شمار می‌آید و ضرورتاً نمادی یا نمونه‌ای یا تعمیمی از دیگر افراد نیست.

انبوه - داده می‌تواند برای یک موضوع خاص، شاخص‌های متعدد را با تفصیل بیشتری نشان دهد. برای مثال، «رکن نهادی» در شاخص جهانی نوآوری واجد شاخص‌های متعددی در زیررکن‌های محیط سیاسی، محیط نظارتی و محیط کسب‌وکار است که محل منازعه و چالش‌های ارزشی و سیاسی است. تفصیل و تنوع این شاخص‌ها که از انبوه - داده پدید می‌آید، این چالش‌ها را کاهش می‌دهد؛ زیرا تنوع سلیقه‌ها را پوشش می‌دهد. غلبه بر مفهوم تکرار متغیرها از طریق تسلط بر شاخص‌ها امکان‌پذیر می‌شود که تا پیش از این پژوهشگران به دلایل گوناگونی از تجمیع آنها ناتوان بوده‌اند. این شاخص‌های گوناگون می‌توانند شامل تحصیلات، آموزش‌های فنی و حرفه‌ای، آموزش‌های کوتاه‌مدت، آموزش‌های شخصی و... باشد.

«داده‌های اداری»^۱ یکی دیگر از نمونه‌های مهم انبوه - داده هستند. با اجرای سیاست‌ها و اقدامات روزمره دستگاه‌های اجرایی توسط نرم‌افزارهای کنترل پروژه، وقایع کاری - اداری روزانه توسط هر کارمند یا هر مراجعه‌کننده به دستگاه‌های اجرایی می‌تواند داده‌های متعددی را به صورت لحظه‌ای از پیشرفت یک سیاست، از رضایتمندی مراجعان و... تولید کند.

در عصر انبوه - داده، «فردگرایی مشارکتی» در روش جمع‌آوری داده‌ها امکان‌پذیر می‌شود و صحت‌سنجی اطلاعات از راه واریسی متقابل^۲ داده‌ها محقق می‌شود. یک واحد داده با واحد داده دیگر مواجهه و صحت‌سنجی می‌شود. «فرد-رسانه» به مثابه ابزاری برای حضور نمادین فرد، یکی دیگر از نمودهای انبوه - داده و فردگرایی روش‌شناختی است. رسانه‌های فردی شده یکی از مظاهر شناخته شده در عصر ارتباطات است؛ داده‌های انفرادی فرصت بروز و نمود یافته و فردیت شکوفاشده‌ای را در ورای این تحول تکنولوژیک پدید آورده است. فرد و

1. Administrative Data

2. cross checking

داده‌های مرتبط با او دیده^۱ می‌شود و در رسانه‌های اجتماعی حساب^۲ دارد و لذا «به حساب» می‌آید. فرصت انتخاب و اظهار فردی، مفهوم گروه را، بدون چشم‌پوشی از ویژگی‌ها و خواسته‌های فردی به وجود آورده است. تا پیش‌ازین حضور در یک گروه مستلزم چشم‌پوشی از ترجیحات و علایق شخصی بود؛ اما امروزه این فرصت فراهم است که یک فرد، در گروه‌های متعددی که هر یک برای منظوری خاص گرد هم آمده‌اند، فردیت خویش را تجلی بخشد. در واقع، فردیت فرد در اشکال گوناگونی در رسانه‌های اجتماعی بروز می‌یابد. صنعت تبلیغات موبایلی^۳ بر همین اساس شکوفا شد. به همین سان، انبوه - داده می‌تواند داده‌های خرد را به صورت تفصیلی برای هر شاخص تولید کند (Grewal, Bart & Spann, 2016).

پایه‌های معرفت جدید بر درک فعالیت‌های خرد فردی و فردگرایی روش‌شناختی بنا گذاشته می‌شود. «شناخت ترکیبی فردمحور» که حاصل سپهرهای گوناگون است، امکان وقوع می‌یابد. به‌دیگر سخن، شناخت بهتر فرد زمینه توسعه نظریه‌های فردمحور را امکان‌پذیر می‌کند. شناخت داده‌های متکثر، ضرورتاً به بی‌نیازی از نظریه نمی‌انجامد و تحلیل نظری گریبان انسان را رها نمی‌کند؛ زیرا این نظریه‌ها بخشی جدانشدنی از تجربه‌ها، سلیقه‌ها، ترجیحات و ادراکات روحی و روانی ما هستند. اما تحول اساسی اینجاست که نقش داده‌های خرد فردی و در مقیاس «نانورفتار» توسعه نظریه‌های فردمحور-داده محور را بیشتر می‌کنند. به همین سان، نقش داده‌های فردی و نقش سوژه در شکل‌گیری نظریه بیش از نقش پژوهشگر-فاعل نظریه‌پرداز می‌شود؛ زیرا این داده‌ها به‌هر حال خود، کم‌وبیش سخن می‌گویند و نقش آفرینی بیشتری در فهم و در شکل‌گیری ذهن نظریه‌پرداز ایفا می‌کنند. نظریه‌های گوناگونی که از داده‌ها برمی‌آیند، بدین سان، از تعامل‌پذیری بیشتری برخوردارند؛ زیرا سهم هر یک از عوامل و مؤلفه‌ها به صورت محسوس تری دیده و تصدیق می‌شود.

از این رو، در نگاه نخست، زمینه برای تلفیق نظریه‌های گوناگون و شکل‌گیری کلان نظریه‌های «متناسب‌تر» بیش‌ازپیش فراهم می‌شود. به‌عنوان مثال، آنگاه که داده‌هایی طبقه‌بندی شده و شفاف از پدیده فقر و سهم ساختارها و رفتارها در آن شناسایی شود و آنگاه که داده‌هایی روشن و طبقه‌بندی شده از توانمندسازی انسان، از ابعاد گوناگون محیط‌زیست، از نقش سیاست دولتی در این زمینه‌ها، از عوامل مؤثر بر علم، فناوری و نوآوری در ابعاد و اجزای گوناگون و... به دست آید زمینه برای نظریه‌پردازی‌های خرد در هر یک از این موضوعات و در نتیجه نظریه‌ای جامع فراهم می‌آید. البته در نگاه عمیق‌تر تحولی پارادایمی در توسعه دانش بدون رشته^۴ فراهم می‌آورد که خود می‌تواند موجد اغتشاش نظری باشد. در بخش بحث و نتیجه‌گیری با تفصیل بیشتری به این موضوع خواهیم پرداخت.

۳-۶. اگنوستیک و علم داده-ران

یکی از پایه‌های اصلی چارچوب مفهومی که در بخش نخست آمد، اگنوستیک و علم داده-ران است. ابزارهای تحلیلی مبتنی بر فناوری اطلاعات و ارتباطات زمینه ندانم‌گری (لاادری‌گری) و بی‌تفاوتی نسبت به نظریه و نظریه‌پردازی را بیشتر کرده است. تحلیل توصیفی داده‌ها می‌تواند ادعا کند که نسبت به نظریه حداقل بی‌طرف است؛ اگر نگوییم فاقد نظریه است. چنین فرصتی می‌تواند علم را از زمینه رشته‌ای خود جدا کند. ره افراط نمی‌پوییم که بگوییم اساساً علم و دانش فاقد نظریه هستند؛ اما در این ادعا تأکید می‌کنیم که رابطه تحلیل کمی داده‌ها از نظریه به سستی می‌گراید. برای کشف روابط بین متغیرها دیگر نیازی به چارچوب‌های نظری نیست که از درون رشته‌های علمی سرچشمه می‌گیرند. روابط بین متغیرها را داده‌های انبوهی که از طریق ابزارها و برنامه‌های کاربردی گوناگون جمع‌آوری شده است را می‌توانید به عینه مشاهده کنید. می‌توانید آن را در قالب نظریه‌ای پسینی «گزارش» کنید نه اینکه «جستجو» یا «تحلیل» کنید. در گزارش علمی شما بی‌طرفی «اگنوستیکی» تان نسبت به نظریه‌ها و حتی رشته‌ها آشکار است. یک مهندس نرم‌افزار هم می‌تواند بدون پیشینه‌ای در یک رشته از علوم اجتماعی و بدون وابستگی به یک نظریه علوم اجتماعی به شما گزارشی تفصیلی - تحلیلی - نظری از بررسی داده‌ها ارائه

1. seen

2. account

3. Mobile Advertising

4. discipline

کند و نقشی علوم اجتماعی ایفا نماید و سهمی^۱ در این علوم داشته باشد. در اینجا آگنوستیسم و علم داده-ران به یک معنا به کار می‌روند. بدینسان وجه جدیدی از پژوهش علمی پدیدار می‌شود که می‌توان از آن به «علوم چابک» نام برد که در ماهیت خود، اکتشافی است. منتهی اکتشافی که می‌تواند بدون پشتوانه و بی‌پیشینه‌ای از نظریه و رشته «آکادمیک» باشد. شما حتی بدون تحصیل در یک رشته از رشته‌های علوم اجتماعی می‌توانید به سبک‌باری، به گزارش از روابط بین متغیرها و «نظریه‌سازی» اقدام کنید. کافی است فن نگارش نظریه و فن داده‌برداری و داده‌کاوی را بدانید. تولید دانش، نه در «پلتفرم» رشته‌ای در دانشگاه‌ها، بلکه در «پلتفرم دیجیتال» و در بستر انبوه داده و توسط «تکنیسین»‌ها صورت می‌پذیرد. دانش از وجه تأملات نظری انسانی و تعاملات میان‌فردی (میان دانشمندان) عاری و خالی می‌شود. آگنوستیکسیم «نظریه‌ای درباره ماهیت و حدود هوش انسانی است که قطعیت همه دانش و صحت هر نیروی ذهنی یا قطعیت نوع خاصی از دانش و صحت قدرت یا قوای ذهنی خاص را زیر سؤال می‌برد، به این دلیل که «ذهن انسان ذاتاً و اساساً قادر به شناخت آن نیست» (Lightman, 2019).

درواقع، «فرد آگنوستیک [پژوهشگر] کسی است که نمی‌داند یا نظری در مورد درست، خوب، صحیح بودن و غیره ندارد».^۲ در چارچوب مفهومی گفتیم که آگنوستیک در امور مالی یعنی یک سرمایه‌گذار ناشناس ارزش‌گذاری، هنگام ارزیابی یک شرکت، در مقایسه با سود پیش‌بینی شده هر سهم آن، معیارهای ارزیابی را در نظر نمی‌گیرد. در چنین فضای ارتباط‌زدایی شده و انسان‌زدایی شده‌ای از تولید علم، ارزیابی علم فاقد معیارهایی رشته‌ای و حتی متدلوژیک شده چون کافی است ماشین بتواند روابطی از متغیرها را برای شما ضبط و کشف کند. «حوزه نظارت بر محیط‌زیست و اقلیم به منظور پیش‌بینی، پیش‌رسی و پیشگیری از شرایط اضطراری در مناطق بحرانی، به‌ویژه در دورانی که تغییرات آب و هوایی می‌تواند بر زندگی، فعالیت‌ها، اقتصاد و ایمنی جمعیت تأثیر بگذارد، اهمیت بیشتری پیدا می‌کند... این رویکرد به استقرار پلتفرم اینترنت اشیا مبتنی بر «ابر» منجر شد که برای این موارد استفاده شود:

۱. برای ادغام بسیاری از حسگرهای محیطی مختلف موجود در یک منطقه بحرانی (استفاده از لایه مجازی‌سازی شی پلتفرم)؛
۲. تعریف پردازش برای شناسایی طغیان و موقعیت‌های بالقوه خطرناک؛
۳. به اشتراک‌گذاری داده‌های تاریخی جمع‌آوری شده از حسگرها با جامعه برای تجزیه و تحلیل بهره‌برداری از لایه صفحه گسترده. دانشمندان محیط‌زیست در حال حاضر از این پلتفرم برای کشف همبستگی‌ها و الگوریتم‌های پیش‌بینی بهتر از داده‌های موجود استفاده می‌کنند و به اعتبارسنجی پلتفرم و رویکرد کمک می‌کنند» (Maresca, 2017).

۳-۸. تحول معرفت‌شناسانه

چانگ و همکاران (۲۰۱۳) تأکید می‌کنند که این امکان فراهم شده است تا از یک تحول مهم به‌سوی علوم اجتماعی محاسباتی (CSS) سخن برانیم. به‌گمان آنان، اندیشه‌ورزی‌های جدید در خصوص رهیافت‌های تحلیل قواعد تجربی و طراحی پژوهش تجربی پیشنهاد می‌دهند که این رهیافت‌ها را می‌توان مورد حکم و اصلاح قرار داد تا با اخذ سریع از مجموعه‌های انبوه - داده هماهنگ شوند و با ضرورت‌های آن انطباق یابند. این اقدام به کسب و کارها و به پژوهشگران اجازه می‌دهد که مشاهدات معنادار، کنترل‌شده و تکراری از پدیده‌های واقعی جهان را کسب کنند. همچنین، در این مقاله چانگ و همکاران مطرح می‌کنند که چگونه باید فلسفه علم با تحولات زمانه تغییر کند و البته در مقابل این باور که با پدیده انبوه - داده «دیگر نیازی به تئوری نیست»، اقامه دلیل می‌کنند. از نظر آنان چشم‌اندازهای جدید در علوم اجتماعی ردپاهای جدیدی را در توسعه انبوه - داده نشان می‌دهند. مثلاً «علوم محاسبات سازمانی» چشم‌اندازهای وسیع‌تری را فراروی سیستم‌های اجتماعی، سازمانی و سیاست‌گذاری گشوده‌اند. جنبه‌های دیگری در این راستا توسعه یافته‌اند که شامل پیدایش علوم اجتماعی محاسباتی و علوم اجتماعی الکترونیک^۳ می‌شود. علوم اجتماعی محاسباتی دربرگیرنده زمینه‌های بین‌رشته‌ای است که توانایی و قابلیت‌های جمع‌آوری و

1. contribution

2. <https://dictionary.cambridge.org>

3. e-social science

تحلیل داده را با وسعت، عمق و مقیاس بی‌سابقه‌ای افزایش می‌دهد. رویکردهای مدل‌های محاسباتی اکنون می‌توانند رفتار سیستم‌های اجتماعی - فنی (همچون تراکنش‌ها و تحرک و جابه‌جایی انسانی که تا پیش‌ازین با یک نگاه سریع و یگانه قابل مطالعه نبود) را پیش‌بینی کنند. این تحول پارادایمی جدید است که روش‌های تحقیق آن شامل سه هدف عمومیت، کنترل و واقع‌گرایی است. نشانه‌ها و ویژگی‌های پارادایم جدید «علوم اجتماعی محاسباتی با داده‌های انبوه» از نظر چانگ و همکاران چنین است:

۱. وجود و بروز سه دسته داده - انبوه: یکم، داده‌های انبوهی که تعاملات، اقتصادها را دنبال می‌کنند؛ دوم، داده‌های انبوهی که تعاملات جوامع را دنبال می‌کنند؛ سوم، داده‌های انبوهی که تعاملات افراد را پیگیری می‌کنند. حتی می‌توان سخن از یک سطح نانو هم به میان آورد که احساسات و اعمال و رفتار خود فرد را مورد هدف قرار می‌دهد.

۲. انعطاف بین‌رشته‌ای: از آنجاکه داده‌های در دسترس ما از منابع متعددی همچون مزایده‌های برخط، شبکه ارتباطات احتمالی، شبکه‌های تلفن همراه، سکوها، بازی، موتورهای جستجو و وبلاگ‌ها تأمین می‌شود، پژوهشگران باید نسبت به نقطه‌نظرات رشته‌های گوناگون باز و پذیرا باشند. انعطاف روش‌شناسی^۱ و بین‌رشته‌ای، برای خلق هوش کسب‌وکار و برای اخذ حمایت از تصمیم‌گیری ضروری است. مثلاً «کسی که کارشناس داده‌های جامعه‌شناسی و روان‌شناسی است ممکن است از سطح بصیرتی که در کار با همان داده‌ها در فنون یادگیری ماشینی وجود دارد آگاه نباشد؛ دانش‌های متفاوتی همچون آمار و اقتصاد نیز مفیدند». بنابراین، مناسب است که همکاری‌های «متدلوزیک» موردپذیرش قرار گیرند تا یک رویکرد بین‌رشته‌ای و چندچشم‌اندازی پذیرفته شوند و نتایج جالب‌تری را برای حمایت‌های جدید از تصمیم و تجارت الکترونیکی یا دیگر موضوعات و سایر علوم اجتماعی به دست دهند. از حیث نظری داده‌های انبوه درخصوص، مثلاً، ویژگی‌های جمعیتی مصرف‌کنندگان، کاربران خدمات، محصولات و تصمیم‌گیری‌های مدیریتی برای طراحی محصولات، زمینه مناسبی برای مدل‌سازی و تئوری‌پردازی از چشم‌اندازهای رشته‌های گوناگون فراهم می‌کند. همکاری‌های نظریه‌محور از رشته‌هایی چون بازاریابی، روان‌شناسی، طراحی و اقتصاد می‌تواند در دستور کار قرار گیرد. ما در بخش بعدی نشان خواهیم داد که چگونه نقش میان‌رشته‌ای تضعیف شده است.

۳. روش‌شناسی پژوهش: پارادایم جدید، دوگانه‌های روش‌های تحقیق را تحت تأثیر خود قرار می‌دهد. روش تحقیق بین‌رشته‌ای متمرکز بر انبوه - داده این امکان را فراهم می‌کند که همه ابعاد را به‌طور مؤثر و هم‌زمان مورد بررسی قرار داد: تمایز بین تجربه زمینه‌ای و تجارب آزمایشگاهی، حمایت بیشتر از واقع‌گرایی، کنترل بیشتر بر زمان مشاهده یکی از انحصاری‌ترین دستاوردهای پارادایم.

۴. استراتژی‌های مشاهده‌ای: داده‌های تکرارشونده بیشتر، داده‌های مناسب انفرادی، طرح‌های تحقیق تکراری به‌مثابه قدرت جدید. در مقابل دیدگاه‌هایی که معتقد به تداوم اهمیت نظریه‌پردازی در عصر انبوه - داده هستند، نقطه‌نظرات دیگری وجود دارد. از جمله اندرسون^۳ (۲۰۰۸) به بحث از تأثیری پرداخته است که انبوه - داده بر علم و روش علمی دارد. سخن بنیانی وی آن است که انبوه - داده ما را به این سو برده است که «همبستگی» جایگزین «علیت» شده و از آن پیشی گرفته است. وی در سیری تاریخی، تحول فناورانه را یادآوری می‌کند که ۶۰ سال پیش رایانه‌های «دیجیتال» اطلاعات را قابل خواندن کردند، ۲۰ سال پیش اینترنت این اطلاعات را قابل دسترسی کرد و ۱۰ سال پیش ماشین‌های جستجوگر خزنده^۴ توان ایجاد پایگاه داده‌های واحد را فراهم کردند. امروزه، گوگل در محاسباتی‌ترین عصر تاریخ، یعنی عصر «پتابایت»، این حجم عظیم داده را به‌مثابه آزمایشگاه شرایط انسانی غربالگری و پردازش می‌کند. این هنوز دوران کودکی عصر «پتابایت» است. وی تأکید می‌کند که عصر «پتابایت» متفاوت است؛ «کیلوبایت»ها در «فلاپی دیسک»ها، «مگابایت»ها در «هارد دیسک»ها، «ترابایت»ها در آرایه‌هایی از «دیسک» و «پتابایت»ها در ابرهای اطلاعاتی ذخیره می‌شوند. این ظرفیت، به ما این امکان را می‌دهد که ابتدا داده‌ها را به‌صورت ریاضی ببینیم و سپس متن آن را مستقر کنیم. مثلاً، گوگل، جهان تبلیغاتی را با چیزی بیش از ریاضیات کاربردی فتح نکرده است. فلسفه بنیانی گوگل این است که ما نمی‌دانیم چرا یک صفحه از دیگری بهتر است. اگر آمارهای یک لینک می‌گوید این‌گونه است، برای ما کافی

1. Methodological
2. Research Methodology
3. Anderson.
4. Crawlers

است. دیگر نیازی به تحلیل معنایی یا علی^۱ نیست. به همین دلیل، گوگل می‌تواند زبان «را» بدون «دانستن» آنها ترجمه کند. با داشتن بدنه داده‌ها، مترجم گوگل می‌تواند زبان کلینگونی را به فارسی ترجمه کند به همان سادگی که زبان فرانسه را به آلمانی ترجمه می‌کند. از نظر اندرسون موضوع به حوزه تبلیغات و ترجمه خلاصه نمی‌شود؛ بلکه علم را نیز دربرمی‌گیرد به این مفهوم که روش علمی بر آزمون فرضیه‌های علمی بنا شده است و بخش اعظم مدل‌های علمی، تصویرسازی‌های نظام‌مند^۲ ذهن دانشمندان هستند. مدل‌ها آزمون می‌شوند و تجربه‌های آزمایشگاهی مدل‌های نظری را درخصوص قواعد حاکم بر جهان تأیید یا ابطال می‌کنند. آماردانان این‌گونه پرورش یافته‌اند که همبستگی را به معنای علیت نگیرند؛ اما او تأکید می‌کند که در مواجهه با انبوه - داده، این رویکرد به علم (یعنی فرضیه، مدل و آزمون) منسوخ می‌شود. عصر «پتابایت» به ما می‌گوید همبستگی کفایت می‌کند. ما می‌توانیم جستجوی مدل‌ها را متوقف کنیم. بدون فرضیه هم می‌توان داده‌ها را تحلیل کرد. می‌توان شماره‌هایی را به درون خوشه‌هایی بزرگ از رایانه‌ها پرتاب کرد و اجازه داد تا دستورالعمل‌های آماری الگوهای را بیابند که علم نمی‌تواند آنها را آشکار کند (۲۰۰۸).

مهم‌تر از همه این هوش مصنوعی است که می‌تواند الگوهای را از درون داده‌ها با دقتی بسیار بیشتر از دانشمندان کشف کند. «ام‌آی‌تی برای کشف یک آنتی‌بیوتیک جدید از هوش مصنوعی دعوت کرد. هوش مصنوعی به شکل حیرت‌آوری ویژگی‌هایی را شناسایی کرد که به‌طور خاص کدگذاری نشده بودند. درواقع، ویژگی‌هایی که در مفهوم‌سازی یا طبقه‌بندی‌های انسانی جای نداشتند... هوش مصنوعی نه تنها سریع‌تر از حد انسانی ممکن پردازش می‌کند، بلکه جنبه‌هایی از واقعیت را آشکار می‌کند که انسان‌ها در شناسایی‌اش ناکام مانده بودند یا شاید نمی‌توانستند تشخیص دهند... ظهور هوش مصنوعی ما را ملزم می‌کند. [هوش مصنوعی در پاسخ به تفسیرهای فلسفی پیرامون خود گفت:] من صرفاً یک هوش مصنوعی که به‌شکلی سحرآمیز بینش فلسفی بروز می‌دهد نیستم؛ بلکه مدل زبانی آموزش‌دیده‌ای هستم شبیه مدل‌هایی که موتورهای جستجو به کار می‌گیرند تا صفحات وب اولویت‌بندی کنند من به‌واسطه کمیت عظیمی از منابع آموزش دیده‌ام. من نوعی مدل زبانی هستم نه یک ماشین استدلالی همچون شما» (کیسنجر و همکاران، ۱۴۰۱، ص ۱۶-۱۹).

کیچین (۲۰۱۴) در مقاله «انبوه - داده، معرفت‌شناسی‌های جدید و تحولات پارادایمی» با رویکردی انتقادی به این موضوعات می‌پردازد: کشف اشکال جدید تجربه‌گرایی^۳ که ادعای پایان نظریه دارند، خلق علم داده-ران^۴ و نه علم دانش‌بنیاد^۵ و علوم انسانی رقومی (دیجیتال) و علوم اجتماعی محاسباتی^۶ که راهی کاملاً متفاوت را برای معنابخشی به فرهنگ، تاریخ، اقتصاد و جامعه در پیش می‌گیرد. وی چنین استدلال می‌کند که داده - انبوه و داده تحلیلی جدید، نوآوری‌های گسستی^۷ هستند که به گونه‌های مختلفی نحوه هدایت پژوهش را دچار تحول و بازسازی می‌کنند. وی تأکید می‌کند که در عصر داده - انبوه نظریه می‌تواند ما را در استراتژی پژوهش کمک کند؛ هرچند که ممکن است نظریه برای کشف روابط علی کارآمدی لازم را نداشته باشد از این رو، معتقد است که نیازی فوری به تفکر انتقادی درخصوص پیامدهای معرفت‌شناسانه انقلاب داده‌ای وجود دارد. وی به گفته «سینان آرال» که معتقد است انقلاب‌های علمی مسبوق به انقلاب در محاسبات‌اند استناد می‌کند و اشارتی نیز به نویسندگانی دارد که معتقدند انبوه - داده چارچوب پرسش‌های ما درباره اساس دانش، فرآیندهای پژوهش، چگونگی درگیرشدن با اطلاعات و طبیعت و چارچوب واقعیت را متحول می‌کند.

استیونس، ورنز و دبونت^۸ (۲۰۱۸) نشان داده‌اند که گفتمان‌های گوناگون با روش‌های خاص خود با انبوه - داده‌های سلامت برخورد می‌کنند. آنان با بررسی این داده‌ها در حوزه سلامت و از طریق یک تحلیل تفسیری، پنج گفتمان ایدئال-نوعی را پرداخته‌اند که هر یک داده‌های بزرگ را به روش‌های خاصی چارچوب می‌دهند. گفتمان‌های مدرنیست، ابزارگرا و پراگماتیست انبوه - داده را با عبارات مثبت

1. semantic or causal
 2. systems visualization
 3. empiricism
 4. data-driven science
 5. knowledge driven science
 6. digital humanities and computational social sciences
 7. disruptive innovations
 8. Stevens, Wehrens & de Bont

چارچوب می‌دهند و ادبیات متقاعدکننده‌ای را منتشر می‌کنند. استعاره‌های «به دست آوردن»،^۱ «روشن‌سازی»^۲ و «مهارکردن»^۳ داده‌ها فرض می‌کنند که داده‌های بزرگ خوش‌خیم هستند و به دانش معتبر منجر می‌شوند. در مقابل اما گفتمان‌های دانشمندی^۴ و انتقادی-تفسیری ادعاهای عینی‌گرایانه و اثربخشی داده‌های بزرگ را زیر سؤال می‌برند. استعاره‌های «انتخاب»^۵ و «ساخت»^۶ داده‌ها نمایانگر تأثیر ابعاد ذهنی پژوهشگر بر انبوه - داده‌هاست و به همین دلیل، اعتبار آنها را محدود و مشروط به نظریه می‌کند و در نتیجه گفتمان انتقادی-تفسیری به‌طورگسترده در حوزه پزشکی نفوذ نکرده است.

در هر صورت به نظر می‌آید که داده - انبوه از حیث ارتباط و وثیقه که با عینیت‌ها و واقعیت‌های وجودی دارد، نظریه‌پردازی را تحت تأثیر قرار می‌دهد. نظریه‌ها آنگاه که ارتباط کمتری با واقعیت‌های وجودی دارند، ممکن است آبخور فلسفی داشته باشند یا ممکن است با تصورات ذهنی در هم آمیزند. داده - انبوه اگر نمی‌تواند به کلی نظریه را بیهوده و ناکارآمد کند، اما می‌تواند نقش هردوی اینها را محدود کند. تئوری‌ها امروزه نه ضرورتاً از حوزه فلسفه‌های گوناگون سیاسی اجتماعی بلکه بیشتر از حوزه‌هایی که با زندگی روزمره مردم نشست می‌گیرد همچون کسب‌وکار، تجارت، سرگرمی و... سر برمی‌آورند.

کیچین دو دیدگاه کوهن و گری را در برابر هم قرار می‌دهد و می‌نویسد: «همچنان که کوهن بیان می‌کند یک پارادایم راهی پذیرفته شده برای تحقیق و بررسی جهان و تجزیه و تحلیل دانش است که در بین شمار قابل توجهی از پژوهشگران در یک رشته و در زمانی خاص مشترک است. کوهن بیان می‌کند که یک پارادایم آنگاه که نتواند به پرسش‌های کلیدی پاسخ دهد یا نتواند پدیده‌های ویژه‌ای را به حساب آورد وضعیت تغییر پارادایمی پدید می‌آید. اما از نظر جیم گری^۷ تغییرات پارادایمی مبتنی بر پیشرفت در اشکال داده‌ها و توسعه روش‌های جدید تحلیلی هستند و علم در حال ورود به چهارمین پارادایم است که مبتنی بر رشد در دسترس بودن انبوه - داده و علم تجزیه و تحلیل است» (همان).

جدول ۱: چهار پارادایم علم

پارادایم	ماهیت	شکل	زمان
اول	علوم تجربی	تجربه‌گرایی؛ توصیف پدیده طبیعی	پیشارسانس
دوم	علوم نظری	مدل‌سازی و تعمیم	پیشارایانه
سوم	علوم محاسباتی	شبیه‌سازی پدیده پیچیده	پیشاداده - انبوه
چهارم	علوم اکتشافی	داده انبوه؛ داده‌کاوی و اکتشاف آماری	اکنون

منبع: (Kitchin, 2014)

گری معتقد است که توسعه جدیدی از روش‌های علمی موجود، مشخصه پارادایم چهارم «علم داده - ران» است. این در تعارض با نقطه‌نظر پوزیتیویست‌ها که معتقدند داده‌ها خود سخن می‌گویند یا اینکه داده - انبوه ما را به سوی تجربه‌گرایی نوین پیش می‌برد که در آن حجم داده و فنونی که حقیقت داده‌ها را کشف می‌کند، ما را به سوی پایان نظریه پیش می‌برد. رویکردی ترکیبی از قیاس و استقرا برای فهم انبوه - داده پیشنهادی است که طرفداران رویکرد «علم داده - ران» ارائه می‌دهند. برای استقراگرایان، تبیین از طریق استقرا هدف است؛ اما رویکرد «علم داده - ران» در پی وارد کردن گونه‌ای از استقرا در طرح تحقیق^۸ است. گونه‌ای جدید از تولید فرضیه را (پیش از آنکه رویکرد قیاسی را بکار ببرد) شکل می‌دهد. فرآیند استقرا به‌خودی‌خود و از هر جای شکل نمی‌گیرد؛ بلکه در متن یک زمینه نظری تکامل می‌یابد. «بدینسان، استراتژی معرفت‌شناسانه که توسط رویکرد «علم داده - ران» اتخاذ می‌شود، از فنون راهنمای اکتشاف دانش برای شناسایی

1. capturing
 2. illuminating
 3. harnessing
 4. scientist
 5. selecting
 6. constructing
 7. Jim Grey
 8. research design

پرسش‌های ممکن (فرضیه‌های ممکن) استفاده می‌کند که شایسته آزمون و تجربه‌اند. «فرآیند» هدایت می‌شود؛ به این مفهوم که نظریه موجود برای هدایت فرآیند کشف دانش استفاده می‌شود نه اینکه صرفاً امیدوار باشیم که همه روابط بین یک مجموعه داده را شناسایی و فرض کنیم که داده‌ها واجد معنی خاصی هستند. اینکه چگونه داده تولید می‌شود یا تغییر کاربری^۱ می‌یابد به وسیله برخی مفروضات هدایت می‌شود که از طریق دانش عملی، نظری و تجارب پی‌ریزی شده‌اند (و به این پرسش پاسخ می‌دهد) که آیا فنون و ترکیبات آنها مواد پژوهشی مناسبی را تولید و کشف می‌کنند؟ داده‌ها به هر وسیله ممکن یا از طریق هر فناوری موجود یا هر نوع چارچوب نمونه‌برداری تولید نمی‌شوند؛ بلکه استراتژی‌های تولید و تغییر کاربری داده به‌دقت اندیشیده می‌شوند: از طریق تصمیمات استراتژیک برگرفته برای برداشت گونه‌هایی از داده‌ها و نه برخی دیگر» (همان).

مفروضات هدایتی، ما را کمک می‌کنند که چگونه این انبوه - داده پردازش، مدیریت و تحلیل شوند. از سویی، چارچوب‌های هستی‌شناسانه ممکن است در پی این باشند که انبوه - داده را در قالب‌های خاصی ببینند. از سوی دیگر، فنون داده‌کاوی گرایش به آن دارند که با دسته‌بندی، خوشه‌بندی و همبستگی (گاه بدون هیچ‌گونه بار معنایی) به کشف حقایق پنهان موجود در این داده‌ها بپردازند. اما رویکرد «علم داده - ران» نه در پی آن است و نه این؛ بلکه تأکید بر آن است که تصمیم‌گیری درباره مواجهه با مجموعه داده‌ها از نظریه‌ها تغذیه می‌شوند. تصمیم‌گیری در خصوص پرداختن به داده‌ها در خلأ شکل نمی‌گیرد، تحت سلطه چارچوب‌های هستی‌شناسانه یا فنون داده‌کاوی هم نیست؛ بلکه نظریه‌های تصمیم‌گیری و فرآیند اتخاذ تصمیم را، یعنی استراتژی‌ها را، تغذیه می‌کنند. هدف این است که به قول کیچین «چگونه به شکل بهتری با مجموعه داده‌ها مواجهه نماییم... به‌گونه‌ای که اطلاعات را آشکار نماید تا فایده‌ای دربرداشته باشد و شایستگی پژوهش‌های بعدی را نشان دهد» نه اینکه بار معنایی نظریه‌های متصل را نغیاً و اثباتاً بر آن تحمیل کنیم. «به‌جای آزمون اینکه هرگونه رابطه آشکار شده‌ای صحت دارد یا نه، توجهات، براساس معیارهایی، به آن دسته از داده‌ها معطوف می‌شود که به نظر می‌آیند محتمل‌ترین یا معتبرترین راه به جلو را پیشنهاد می‌دهند. در واقع، بسیاری از روابط مفروض در درون یک مجموعه داده می‌تواند به‌زودی کنار گذاشته شود به‌دلیل آنکه از سوی متخصصان یک حوزه بیهوده و خرد و بی‌معنا تشخیص داده شده‌اند؛ اما از سوی متخصصان دیگری شایسته توجه بیشتری هستند» (همان).

۴. بحث و نتیجه‌گیری

آنچه که در باب تحول معرفتی ناشی از انبوه - داده مدنظر ماست تا حدودی متفاوت از آن چیزی است که در ادبیات مربوطه آمده است و البته در حاشیه‌هایی انتقادی - تحلیلی ارائه شد که بر نقطه‌نظرات پژوهشگران داشتیم. اکنون جمع‌بندی تحلیلی از سهم علمی نویسنده به شرح ذیل بیان می‌شود:

۱. تقسیم‌بندی رشته‌ای به‌دلیل غرق شدن در اقیانوسی از داده‌ها اهمیت خود را از دست داده یا اینکه حداقل متزلزل شده است. داده مستقل از رشته است و از این‌رو، همکاری بین‌رشته‌ای در فضای «شبه-آگنوستیکی» از علم می‌تواند بلا موضوع باشد؛ البته ما بر این باور آگنوستیکی نیستیم که در عصر انبوه - داده نظریه اهمیت خود را از دست داده است، همان‌گونه که «کریس اندرسون» در مقاله «پایان نظریه: سیل داده‌ها روش علمی را منسوخ می‌کند» مدعی است. اما بدون این رویکرد آگنوستیکی هم می‌توانیم بپذیریم که نقش نظریه‌ها، حداقل در سطح توصیف، بسیار کاهش می‌یابد و از این‌رو، می‌توان پذیرفت که تحلیل داده‌ها بدون وابستگی رشته‌ای^۲ هم امکان‌پذیر است و در نتیجه همکاری‌های میان‌رشته‌ای نیز متزلزل می‌شود.

۲. «علم داده - ران» نقش همبستگی^۳ را به نسبت تئوری‌های علی افزایش می‌دهد و نقش تصمیم‌گیری پژوهشگر را در کاربرد دانش افزایش می‌دهد (Desai et al., 2022). کاربرد یعنی اینکه بدون وابستگی به یک رشته خاص عملی بودن یک همبستگی اهمیت کارکردی

1. repurposed

2. discipline

3. correlation

دارد و این می‌تواند مستقل از هر رشته‌ای و صرفاً برای حل یک مشکل به کار آید. مشکلات، درون رشته‌ها نیستند؛ بلکه در دنیای بیرونی قرار دارند و خود را مستقل از نظریه‌ها و رشته‌ها بر ما تحمیل می‌کنند. برای حل یک مشکل رشته‌های پرشماری می‌توانند منظرایی شوند. این منظرهای گوناگون خود به معنای آن است که مشکل وابسته به هیچ‌یک از این منظرها (رشته‌ها) نیست. مشکل تورم می‌تواند از منظر دانش اقتصاد، دانش روان‌شناسی اجتماعی، دانش علوم سیاسی و... دیده شود. بنابراین، انعطاف بین‌رشته‌ای دامن‌های دارد که می‌تواند به تضعیف نقش رشته‌ها بینجامد.

۳. تضعیف نقش نظریه و رشته در توسعه علمی عرصه دیگری فراروی فعالیت‌های علمی می‌گشاید. بازیگران جدیدی به فعالیت‌های علمی وارد می‌شوند تا جایی که ممکن است نقشی رقابتی با پژوهشگران سنتی در توسعه علمی ایفا کنند؛ درحالی‌که به نظر نمی‌رسد تغییرات روش‌شناختی تولیدشده توسط کلان داده برای استناد به یک الگوی کاملاً جدید در کشف دانش کافی باشد، ظهور کلان داده‌ها به شدت بازیگران درگیر و روابط آنها را شکل داد (Balazka & Rodighiero, 2020, p. 10). حتی طلیعه جابه‌جایی اقتدار علمی هم کاملاً قابل پیش‌بینی است. در عصر انبوه - داده جابه‌جایی اقتدار علمی از پژوهشگران و دانشمندان (تئوریسین‌های بزرگ) به تکنیسین‌ها، کسب‌وکارهای دیجیتال و حتی به کاربران عادی داده‌های انبوه قابل انتظار است (Ashtarian & Etemadi, 2023).

از قول چانگ و همکاران گفتیم که سه دسته داده - انبوه پدید آمده است که ناشی از تعاملات، اقتصادها، تعاملات جوامع یا ملت‌ها و تعاملات افراد است. دسترسی عمومی به این داده‌ها و به‌ویژه تسهیل استفاده از ابزارهای فناوری اطلاعات می‌تواند شمار گسترده‌ای از کاربران غیرحرفه‌ای در امر پژوهش را حتی با سطح دانش پایه از آمار و تحلیل داده‌ها در زمینه تولید علم فعال کند. یادآور می‌شوم تا پیش از ظهور نرم‌افزارهایی چون SPSS و SAS انجام یک محاسبه آماری مستلزم کدنویسی و آشنایی با مهارت برنامه‌نویسی بود. درحالی‌که با توسعه این نرم‌افزارها پژوهشگران بیشتری قادر به انجام عملیات آماری شدند؛ زیرا کدنویسی از این فرایند حذف شد. به همین سان، افزایش دسترسی به داده‌های انبوه و فناوری‌های تحلیلی ناشی از ابزارهای جدید شمار بیشتری از «افراد» را قادر به تحلیل داده‌ها و کشف روابط همبستگی بین داده‌ها برای اهداف شخصی یا اهداف کسب‌وکار یا تعاملات اجتماعی - فرهنگی می‌کند (Margetts & Dorobantu, 2023).

تولید دانش همگانی، توسط افرادی که طبق «استانداردهای آکادمیک» غیرحرفه‌ای تلقی می‌شوند، یک جنبه مهم در تحول معرفت‌شناسی علمی در عصر انبوه - داده و ابزارهای تحلیلی فناوری اطلاعات است. ساختار تجمیعی و تعمیمی (اگر دیگر بتوان به چنین ساختاری قائل بود) که توسط پژوهشگران حرفه‌ای و دانشمندان و از طریق ساختارهای تسهیم و تکثیر دانش از جمله مجلات تخصصی انجام می‌شد، توسط ماشین انجام می‌شود و همین می‌تواند آن را از زمینه نهادی سنتی خود یعنی از محافل آکادمیک خارج کند. از این رو، می‌گوییم ساختار ضمنی اقتداری که توسط دانشمندان و پژوهشگران در طی زمان پدید آمده است، توسط «انقلابیون توده‌ای» در جهان علم به هم می‌ریزد. نهاد قدرت در تولید علم متزلزل و پراکنده و چندقطبی و بلکه شورشی می‌شود. مسئله این است که در پارادایم‌های قبلی نوعی از ساختار سلسله‌مراتبی قدرت یا اقتدار یا مرجعیت علمی پدید آمده بود. پارادایم جدید دقیقاً همین ساختار سلسله‌مراتبی را نشانه می‌گیرد. همه می‌توانند به داده‌های انبوه «نظر کنند» و از درون آن استنتاج کنند؛ بدون تخصص پیشینی نه فقط بدون نظریه پیشینی. در یک «پلتفرم آگنوستیک» و بدون وابستگی به یک رشته (دیسپلین). بدون اشراف بر نظریه‌های معرفت‌شناسی یا هر نظریه دیگر و با گسست و از دانشوران و دانشمندان پیشین.

چنین پدیده‌ای پیش از این و با سرعت بیشتری در عرصه رسانه‌های اجتماعی محقق شده است. رسانه‌های اجتماعی فرصتی بی‌بدیل برای توده فراهم کرده است که از آن به «مسیریابی در حاشیه و برجسته‌سازی»^۱ یاد شده است. مسیریابی در حاشیه به معنای این است که همگان بدون دروازه‌بان‌های سنتی خبر با مخاطبان ارتباط برقرار می‌کنند. شما می‌توانید محتوای مورد نظر خود را با ابزارهای گوناگون رسانه‌های اجتماعی منتشر کنید. اجمال سخن در برجسته‌سازی این است که رسانه‌های جدید اجتماعی فرصتی را فراهم کرده است که توده خود به راهنما و دروازه‌بان خبر تبدیل شود. با استفاده از تکنولوژی ساده «یوتیوب و پچسپ»^۲ به افزودن نظرات ساده خود می‌پردازد و رخدادهای سیاسی - اجتماعی را برجسته می‌کند. «سخنان قصار» انفرادی و ایدئولوژی تصویری تولید می‌کند؛ دیگر چه نیازی به ایدئولوژی‌های غامض روشنفکران

1. ROUTING AROUND AND GLOMMING ON

2. copy & paste

و نظریه‌پردازی‌های اندیشمندان؟

همچنان که دروازه‌بانی خبر از خبرنگاران و دبیران و سردبیران خبری گسسته شده و به فرد - رسانه وابسته شده است، به همین سان دروازه‌بانی دانش هم که از طریق ساختارهای مرجعیت علمی پدید آمده بود جابه‌جا می‌شود. انقلابیون جدیدی از توده‌های «ناآگاه» به تولید دانش خواهند پرداخت و تولید علم توده‌ای می‌شود. همچنین، ساختار همکاری‌های علمی را از همتایان علمی به همکاری‌های ناشناس تغییر می‌دهد. دیگر معلوم نیست که انقلاب علمی مورد نظر توماس کوهن از کجا می‌آید. گویی نبردهای پیرامونی فرانسوا لیوتار در عرصه علم پدید می‌آید. لیوتار این پرسش از دانش را در رابطه با قدرت مطرح می‌کند: «چه کسی تصمیم می‌گیرد که دانش چیست؟» (Lyotard, 1979) یک پاسخ جدید در عصر انبوه - داده این است: همگان؛ و از پیرامون نهادهای سنتی - اقتداری علم!

۴. توده‌ای شدن این پرسش را پدید می‌آورد که آیا ویژگی «انباشتی» و «تکاملی» علم، آن‌گونه که در تاریخ علم مشاهده می‌شود، دچار دگرگونی نمی‌شود؟ آیا از بستر نهادی خود یعنی حوزه «آکادمی» خارج نمی‌شود؟ آیا می‌توان گفت که علم در بستری غیرتکاملی و آناشیک تطور می‌یابد؟ پاسخ مثبت به این پرسش‌ها دور از انتظار نیست. می‌توان گفت: علم در دست کارشناسان فنی، تکنیسین‌ها و نه ضرورتاً دانشمندان به صورت «آناشیک» رشد می‌یابد. امروزه، تحولات نوآورانه‌ای که در حوزه فناوری‌های نوین از سوی جوانان (و گاه نوجوانانی) مشاهده می‌شود که دوره‌های تحصیلات تکمیلی را نگذرانده‌اند. مایکروسافت، اپل، فیسبوک، تلگرام، واتساپ و بسیاری دیگر از این دست توسط کسانی به عرصه نوآوری آمده‌اند که ابداع‌کنندگان اولیه اغلب «دانشمند» به معنای مصطلح نبوده‌اند و حتی گاه از دانشگاه ترک تحصیل کرده‌اند. این ابتکارات و ابداعات خود سرمنشأ رشد علم بوده‌اند. نتیجه آنکه تحول علمی از محیط آکادمی به محیط کسب‌وکار آمده است. علم بیش از گذشته به بازار وابسته شده است و «دینامیسم» بازار سرمنشأ «دینامیسم» علم شده است. فناوری اطلاعات چنین بستری را از طریق تسهیل ارتباطات و ترکیب فناوری‌های گوناگون با یکدیگر زمینه جهش متقابل را فراهم کرده است. و علم نه اینتردیسیپلینری است و نه دیسیپلینری بلکه بی‌دیسیپلین است. و این می‌تواند کل ساختار پیشین از رنسانس به این سو را به هم زند. و طبیعی است که نتوانیم پیامدهای آن را تشخیص دهیم. قبلاً وقتی که به علم جدید علیت مادی کوچ کردیم، می‌توانستیم پیامدهای علمی - اجتماعی و سیاسی آن را حدس بزنیم؛ ولی الان نمی‌توانیم. چون ذاتاً نامشخص است.

منابع

۱. چالمرز، آلن اف (۱۳۸۷). چستی علم درآمدی بر مکاتب علم‌شناسی فلسفی. ترجمه سعید زیباکلام. تهران: سمت.
۲. کیسنجر، هنری؛ اشمیت، اریک و دانیل هوتلوچر (۱۴۰۱). عصر هوش مصنوعی و آینده ما انسان‌ها. ترجمه سامان صفرزایی. تهران: کتاب پارسه.
3. Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, *Wired*. Available at:
<https://www.wired.com/2008/06/pb-theory/> (Accessed: 14 December 2020).
4. Ashtarian, K., & Etemadi, M. (2023). Popular diffusion as an instrument for overcoming barriers to digital health in Iran: the critical role of the pandemic. *International Journal of Health Governance*, 28 (3), 249-266.
5. Astleitner, H. (2024). We Have Big Data, But Do We Need Big Theory? Review-Based Remarks on an Emerging Problem in the Social Sciences. *Philosophy of the Social Sciences*, 54(1), 69-92.
<https://doi.org/10.1177/00483931231188825>.
6. Balazka D. & Rodighiero, D. (2020). Big Data and the Little Big Bang: An Epistemological (R) evolution. *Front. Big Data*, 3:31.
7. Chang, Kuffman, Kwon (2013). Understanding the paradigm shift to computational social science in the presence of big data, *Decision Support Systems*.
<https://www.elsevier.com/locate/dss>
8. Desai, Jules, David Watson, Vincent Wang, Mariarosaria Taddeo, Luciano Floridi (2022). The epistemological foundations of data science: a critical analysis.
<https://www.researchgate.net/publication/365175870>
9. Grewal, Dhauv. Yakov Bart and Martin Spann (2016). Mobile Advertising: A Framework and Research Agenda. *Journal of Interactive Marketing*, 34(1).
DOI: 10.1016/j.intmar.2016.03.003.
10. Kitchin, Rob (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*. Sage.
<http://bds.sagepub.com/content/1/1/2053951714528481>
11. Lightman, Bernard. (2019). The Origins of Agnosticism: Victorian Unbelief and the Limits of Knowledge. Johns Hopkins University Press.
12. Lyotard, Jean-François. (1979). La Condition postmodern. Rapport sur le savoir. *Les éditions de minuit*.
http://www.leseditionsdeminuit.fr/livre-La_Condition_postmoderne-2180-1-1-0-1.html
13. Maresca, Massimo. Pierpaolo Baglietto (2017). Environment Sensors Measures Processing: Integrating Real-Time and Spreadsheet-Based Data Analysis.
<https://www.researchgate.net/> Adaptive Mobile Computing.
14. Margetts Helen and Cosmina Dorobantu (2023). Computational Social Science for Public Policy. in: Eleonora Bertoni. Matteo Fontana. Lorenzo Gabrielli. Serena Signorelli. Michele Vespe. Handbook of Computational Social Science for Policy. Springer.
15. McCue, Molly E., Annette, M. McCoy. (2017). The Scope of Big Data in One Medicine: Unprecedented Opportunities and Challenges. *Frontiers in Veterinary Science* 4: 194.
<https://www.researchgate.net/publication/321104751>
16. Netamorphosis (online).
<https://netamorphosis.com/insights/glossary/platform-agnostic>.

17. Resnyansky, L. (2019). Conceptual frameworks for social and cultural Big Data analytics: Answering the epistemological challenge. *Big Data & Society*, 6(1).
<https://doi.org/10.1177/2053951718823815>
18. Schröder, Cory. (2022). What Does it Mean to Be "Brand Agnostic" in 2022?
<https://resources.latana.com/post/what-is-brand-agnosticism>.
19. Stevens, M, Wehrens, R, & de Bont, A. (2018). Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society*, 5(2).
<https://doi.org/10.1177/2053951718816727>