

## Increase of Data and Development of Science

*Morteza Mardiha* (Allameh Tabatabaei University),  
mardihamortaza@yahoo.com

---

### ARTICLE INFO

---

#### Article History

**Received:** 2020/03/02

**Accepted:** 2020/03/16

#### Key Words:

Science,  
scientist,  
development,  
data,  
stagnation,  
discovery

### ABSTRACT

---

Despite what it may seem, some indicators show that during the last quarter of twentieth-century science has been in the state of stagnation and there is less news about big discoveries. However, this period has seen a vast increase in time and money spent on research as well as real augmentation in the number of scientists and articles. Moreover, due to the new tools, hardware, and software, there is a huge amount of data, from microphysics and astronomy to biology and social science, in the disposition of researchers. This article considers those indicators, asks of its causes, and concentrates on relationship between increase of data and development of science. It proposes that development of science depends on testable hypotheses, but massive data and considering more variants make the very act of hypothesizing more difficult. In other words, having more experimental backing does not always contribute to scientific progress, sometimes it becomes an obstacle.

## افزایش داده‌ها و پیشرفت علم

مرتضی مردیها (دانشگاه علامه طباطبائی، mardihamortaza@yahoo.com)

### چکیده

برخلاف آنچه ممکن است به نظر برسد برخی شاخص‌ها نشان می‌دهند علم از حدود ربع چهارم قرن بیستم به بعد دچار رکود شده است و از کشف‌های بزرگ کمتر نشانی هست. این در حالی است که این دوره شاهد بیشترین تعداد دانشمند، بیشترین بودجه تحقیقاتی و بیشترین تعداد مقالات علمی در جهان بوده است. افزون‌براین، انواعی از ابزارهای سخت و نرم جدید، انبوهی از داده‌های تجربی در عموم رشته‌ها از میکروفیزیک تا نجوم و از زیست‌شناسی تا علوم اجتماعی را فراهم کرده‌اند که علی‌القاعده باید به رشد علم شتاب بیشتری بدهند.

این مقاله با بررسی این شاخص‌ها از چرایی واقعه می‌پرسد و بر این نکته متمرکز می‌شود که انبوهی داده‌ها مشکل چندانی از رشد علم حل نمی‌کند؛ زیرا رشد علم در گرو فرضیات آزمون‌پذیر است و داده‌های انبوه‌تر و گرفتن متغیرهای بیشتر امکان فرضیه‌پردازی را دشوار می‌کند. به‌دیگرسخن، پشتوانه تجربی همواره کمک‌کننده به غنای تبیینی نیست، بلکه گاه منافی آن است.

### اطلاعات مقاله

تاریخ دریافت: ۹۸/۱۲/۱۲

تاریخ پذیرش: ۹۸/۱۲/۲۶

### واژگان کلیدی:

علم،

پیشرفت،

رکود،

داده،

کشف و دانشمند

## مقدمه

مشهور است که در دوران باستان و میانه، فیلسوفان و دانشمندان، به تجربه به‌عنوان ابزار شناخت واقعیت اعتنای چندانی نداشتند. بسیاری از تلاش‌ها برای شناخت چیزها متکی به نوعی استدلال بود. از ارسطو که می‌گفت مدار سیارات دایره است، چون حرکت در آن از سر نیاز برای نیل به هدفی نیست، پس شکل کامل است و متناسب با موجودات کاملی چون اجرام آسمانی، تا ابن‌سینا که می‌گفت خورشید عنصری بسیط است؛ چون در چنان سطحی از وجود نمی‌تواند کون و فساد ناشی از ترکیب وجود داشته باشد، می‌کوشیدند چستی پدیده‌ای ملموس را نه با مشاهده که با استدلال معلوم کنند.

حتی برخی بنیانگذاران علوم نوین هم از چنین گرایشی برکنار نبودند. کپرنیک بنیاد کار خود را بیشتر بر محاسبه قرار داد تا مشاهده؛ و دکارت مدعی بود که درک درد مستلزم داشتن روح است و حیوان چون روح ندارد نمی‌تواند دردی حس کند. (این استناد تجربه‌گرایی به دوران پیشامدرن البته چندان درست و دقیق نیست. پیشینیان اغلب هنگامی که امکان عملی تجربه برایشان موجود بود از آن روگردان نبودند؛ جایی که دست تجربه به آن نمی‌رسید به فرض و حدس و نظریه متوسل می‌شدند، و این کاری است که دانشمندان اکنون هم می‌کنند؛ ولی این قدر هست که محک تجربه در قیاس با سامان فلسفی و کلامی در گذشته کمتر مهم بود).

این نیز مشهور است که ترک استدلال و رو کردن به تجربه، شامل مشاهده و آزمون، عاملی بود که حرکت به سوی علوم جدید را کلید زد. بویل و پاسکال نمونه‌های دقیق آن محسوب‌اند و گالیله و نیوتن نمونه‌های مشهور و مهم آن. علم تجربی با قهرمانانی چون دالتون، داروین، مندلیف، پاستور، فارادی، ماکسول و دیگران در قرن نوزده به اوج رسید؛ قرنی که در آن ویول با گذاشتن واژه «ساینست» به جای «نچرال فیلازفر» گویی جدایی قطعی علم تجربی و فلسفه استدلالی را اعلام کرد.

در قرن بیستم بحث بر سر ماهیت تجربه و نقش آن مکاتب فلسفی علم‌شناسی رقیبی ایجاد کرد. از جمله برخی به دلایل گوناگونی یکه‌تازی تجربه در عرصه علم را به معرض تردید گذاشتند. در نیمه دوم قرن بیستم، نازک‌بینی‌هایی در این باب

صورت گرفت و در چستی داده‌های تجربه و تعمیم تجربی و میزان خلوص و اعتبار آن و به‌ویژه نسبت آن با عقلانیت علمی مباحث مفصلی رفت. در عین حال، در اصل محوریت داده‌های تجربی برای کشف علمی انکاری نرفته، ولی ورای این هر چیز دیگری، از اثرپذیری مشاهده از نظریه تا اعتبار تعمیم تجربی، از نسبت کشف جدید مبتنی بر مشاهدات با یافته‌های جا افتاده پیشین، یا عقلانیت علمی، تا نقش امور اجتماعی در انجام تجربه و تفسیر داده‌ها، مورد نقض و ابرام بوده است.

یکی از موضوعاتی که در همین چشم‌انداز می‌توان بررسی کرد این است که چه نسبتی میان گسترش داده‌های خام و عرضه شناخت علمی دقیق‌تر و جامع‌تری از پدیده‌ها وجود دارد؛ به‌ویژه در دوران معاصر که به کمک ابزارهای جدید فراهم آمدن داده‌های انبوه ممکن شده است. به‌دیگرسخن، با فرض اینکه، فارغ از بحث اهمیت نظریه و تأثیرات آن، پیشرفت علم مدرن بر اثر توجه به تجربه و مشاهده و گردآوری داده‌های مشاهده‌ای بوده است، آیا افزایش داده‌ها در دهه‌های اخیر به پیشرفت بیشتر در علم منجر شده است؟ اگر پاسخ منفی است، علت آن چیست؟ چرا میزان مبسوط‌تر و باریک‌تری از داده‌ها، مثلاً نسبت به یک قرن پیش، روند رسیدن به نظریه‌ای درست‌تر و دقیق‌تر و کشف‌های بزرگ را افزایش نداده است؟

## پیشرفت علم

ماکس پلانک، از پیشگامان فیزیک جدید، در ابتدای قرن بیستم چنین می‌نویسد: هرچه علم بیشتر پیشرفت کند پیشرفت‌های بعدی آن سخت‌تر می‌شود. از همان زمان به بعد اختلاف نظر میان دانشمندان طراز اول و عدم پیشرفت در کشفیات بزرگ به چشم می‌خورد. از جمله، اختلاف نظرهای اساسی میان مکتب کپنهاگ (بور و هایزنبرگ) و اینیشتین و پیروان آنها جدی بوده است (kim, 2007)؛ مشکلات نظریه جاذبه عمومی نیوتن ورای منظومه شمسی مستلزم فرضیه کمکی جرم سیاه شد تا به کمک نظریه جاذبه عمومی بیاید تا در برابر ناسازی‌های تجربی به حفظ آن کمک کند (Moskvitch, 2018).

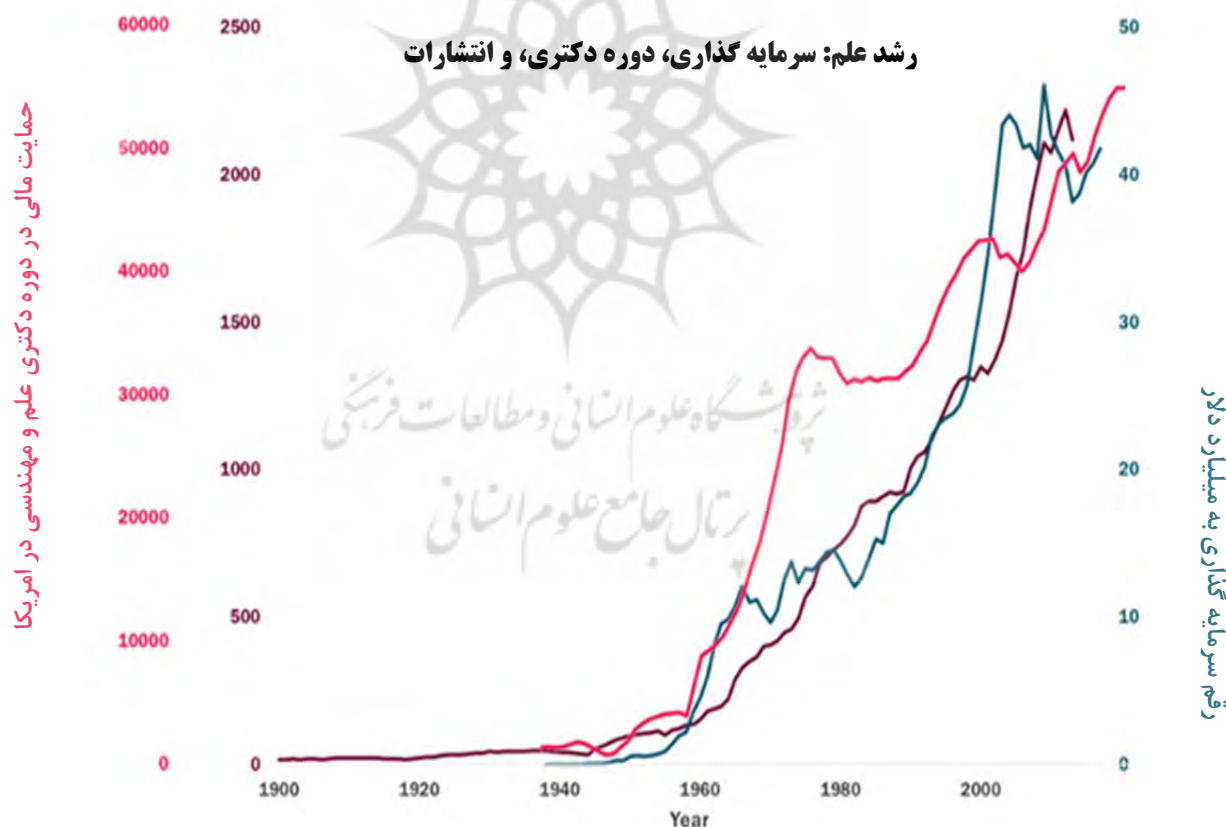
در علوم دیگر، حتی زیست‌شناسی، که پیشرفت‌های مهمی

دورکیم، و وبر، و کینز، و اینشتین و نظایر اینها هستند تا چه حد قابل قیاس با امروزیان از جمله کسانی هستند که در فهرست فارین پالیسی آمده‌اند و علت چیست؟ پژوهشگر دیگری با استناد به شواهد گوناگون می‌نویسد: «پیشرفت‌های مهم در علم به‌طور فزاینده‌ای سخت‌تر و سخت‌تر می‌شود»، و سپس این پرسش را مطرح می‌کند که «به‌رغم این حجم شگفت‌مقالات علمی که هر ساله منتشر می‌شود، چرا علم سریع‌تر از این رشد نمی‌کند؟» (Ramirez, 2018).

مجله آتلانتیک (۱۶ نوامبر ۲۰۱۸) در مقاله‌ای به منحنی به شدت روبه رشد تعداد دانشجویان دکتری و بودجه اختصاص داده شده به آنها و به انتشارات دانشگاهی (در امریکا در طول یک قرن گذشته) اشاره کرده، و این پرسش را مطرح می‌کند که آیا این حجم از افزایش به رشد علم و کشفیات مهم کمکی کرده است.

در نیمه دوم قرن بیستم داشت، اختلاف‌نظرهای اساسی، از جمله در مورد ماهیت ژن (Pearson, 2006)، و چگونگی فرایند انتقال اطلاعات وراثتی از ژن به دی ان ای و از آن به سلول‌ها (Schwanhausser, 2011; Elowitz, 2002) در کار بوده است. در مورد چستی علت یک بیماری همچون سرطان اختلاف میان دانشمندان طراز اول در بیشترین حد ممکن است (Tomasetti, 2017; Knapton, 2018). این مجموعه مشاهدات و ملاحظات را می‌توان طولانی‌تر کرد و نشان داد که طرح موضوعی به نام رکود علم نباید تعجیبی برانگیزد.

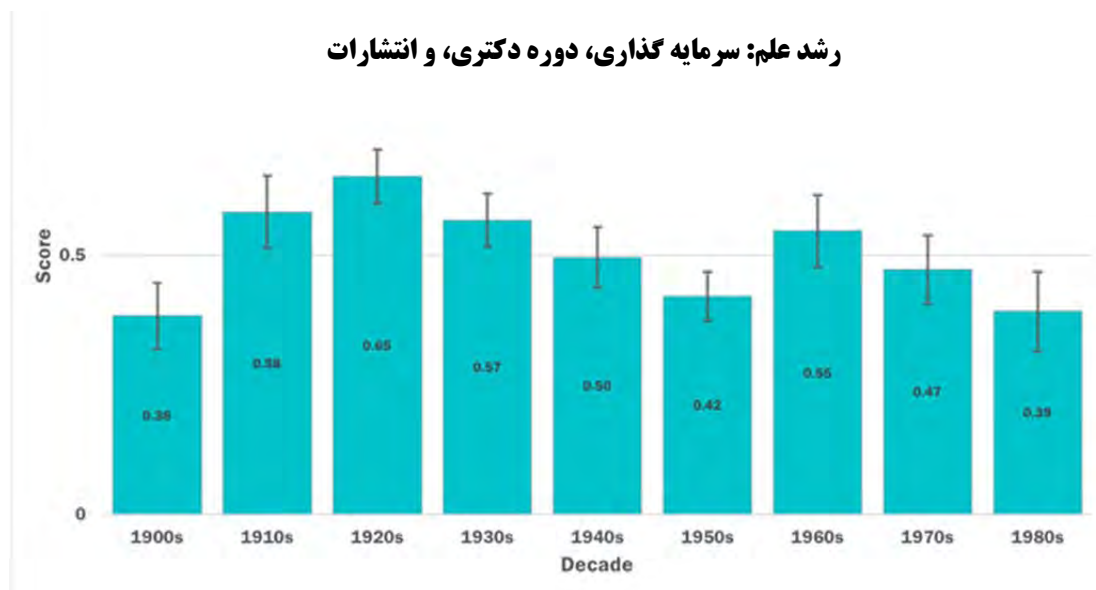
ناظری به اشاره به فهرست مجله فارین پالیسی در معرفی صد متفکر برتر قرن بیست و یکم، پرسشی مطرح کرده است مبنی بر اینکه اگر فهرستی از متفکران بزرگ نیمه دوم قرن نوزدهم یا نیمه اول قرن بیستم تهیه کرده و با این فهرست مقایسه کنیم به چه



برای پاسخ به این پرسش، نویسندگان مقاله پیمایشی در میان بیش از هزار تن از استادان فیزیک، شیمی، و فیزیولوژی برترین دانشگاه‌های دنیا اجرا کرده و از آنان پرسیدند مهم‌ترین اکتشافات

نتیجه‌ای می‌رسیم (Leherer, 2011). اگر در این مدعا متفکران را به دانشمندان محدود کنیم چه پاسخی به این پرسش می‌توان داد؟ فهرستی که در آن کسانی چون داروین و پلانک و فروید و

## رشد علم: سرمایه گذاری، دوره دکتری، و انتشارات



بنتلی گلاس، در سال ۱۹۷۱ در مقاله‌ای در مجله ساینس می‌نویسد: دوران پرشکوه علم سپری شده است. او همچنین می‌نویسد: «در خلال یکی دو نسل، پیشرفت علم به احتمال متوقف خواهد شد... ما شبیه کاوشگران یک قاره بزرگیم که عمده کوه‌ها و رودخانه‌ها و دشت‌های آن را شناسایی کرده‌ایم. هنوز بی‌شمار نکات ریز هست که باید کشف شود، ولی دیگر افق ناپیدایی در کار نیست» (Martin, 2005). جان هورگان، نویسنده امریکن ساینتیفیک، در کتاب خود، پایان علم، می‌نویسد: «ما در آینده به کشف چیزی شبیه جاذبه عمومی، نسبیّت، مکانیک کوانتوم، یا تکامل داروینی نخواهیم رسید» (Horgan, 1996: 16).

هرچند بحث حاضر ناظر به علم است، ولی، شاید حتی نرخ رشد تکنولوژی هم برخلاف ظاهر رو به کاهش داشته باشد. از ۱۹۷۰ به بعد عمده کارها در حوزه رایانه و اینترنت و تلفن همراه بوده است. این در قیاس با اختراع موتور درون‌سوز، موتور الکتریکی، اتومبیل، هواپیما، کشتی‌های بزرگ، وسایل راه‌سازی، دست‌کم در حوزه اقتصاد وزن فراوانی ندارد (Gordon, 2016)؛ هرچند مهندسان مخترعی هستند که از این ایده دفاع می‌کنند که پیشرفت تکنولوژی نه فقط شتاب خطی که شتابی به شکل هندسی افزاینده دارد، و به فراوانی از تعبیر انقلاب برای توصیف انواع پیشرفت‌های نانو و نظیر آن استفاده می‌کنند (Kurzweil, 2005)؛ هرچند در اینجا هم به نظر می‌رسد نسبت میان

قرن بیستم کدام‌ها بوده‌اند و سپس پاسخ‌ها را براساس تاریخ کشفیات به نمودار تبدیل کردند. حاصل آن اینکه دهه‌های اخیر از نظر اهمیت اکتشافات رو به افول جدی داشته است.

چیزی که گراف نشان می‌دهد فقط روند رو به کاهش کشفیات دارای اهمیت (در حد برنده شدن جایزه نوبل) نیست؛ چنان‌که مشهود است گراف تا دهه ۱۹۸۰ بیشتر نیست. علت آن است که کمیته نوبل در سال‌های اخیر ترجیح داده است جایزه هایش را بیشتر به کشفیات دهه‌های هفتاد و هشتاد اختصاص بدهد. دهه‌های ۱۹۹۰ و ۲۰۰۰ دهه‌هایی است که در آن کمیته نوبل کمترین تعداد کشف‌های همان دهه را شایسته جایزه تلقی کرده و جایزه را به کارهای دهه‌های قبل‌تر داده است. این یعنی واقعیت ستون‌های هفتاد و هشتاد در گراف پیشین از آنی که دیده می‌شود کمتر است و در این صورت، سیر نزولی آشکار است.

لئو کادانوف، فیزیکدان برجسته، در مورد میزان پیشرفت‌های اخیر در علم می‌نویسد: «حقیقت این است که هیچ چیز، هیچ چیز، در سطح کشف مغناطیس، مکانیک کوانتوم، نسبیّت، یا ملکول دی‌ان‌ای در کار نیست. هیچ کشفی در این حد و اندازه در دهه‌های اخیر صورت نگرفته است» (Collision & Nielsen, 2018). این ادعای او نتیجه مجموعه تحقیقاتی است که به‌عنوان یک فرد آشنا به حوزه‌های مختلف دانش انجام داده است. زیست‌شناس ممتاز و رئیس انجمن امریکایی پیشرفت علوم،



پژوهشی معتبر شاهد کشفیات بزرگ می‌بودیم که نسبت به یک یا دو قرن پیش کمتر هستیم.

### افزایش داده‌ها و تأیید فرضیه

از زمان‌های قدیم که اهل منطق راجع به استقرا و مقدمات قیاس بحث می‌کردند، این پرسش مطرح بود که چه میزان استشهاد تجربی برای تقویم یک مدعا و احياناً تعمیم آن لازم یا کافی است. در عصر جدید که استناد به سامان به داده‌های تجربی مسیر اصلی کشف علمی دانسته شده، بحث قدیمی استقرا بازپردازی‌های متناسب با علم جدید یافت، و پرسش از میزانی از شواهد تجربی که می‌تواند یک مدعا علمی را از مرز تز و فرضیه و حدس و گمان فراتر برد و جامعه یک نظریه به آن بپوشاند همچنان موقعیت محوری خود را حفظ کرد.

کسانی چون کارنپ بر یک منطق کمی تأیید تأکید کردند و کسان دیگری چون همپل بر منطق غیرکمی تأیید. و البته کسانی مثل کوهن هم وجه اجتماعی تأیید علمی را مورد توجه قرار دادند. امثال پوپر و گودمن استدلال کردند که استقرا، حتی استقرای تام (به فرض ممکن بودن) به یقین منجر نخواهد شد. در تمامی این مسیر، فارغ از پیچیدگی‌های فنی این مباحث، کسی در این اصل بدیهی تردید نکرده است که «اگر داده‌های خام نباشد، علمی هم در کار نیست» (Miyakawa, 2020). باری، به موازات این، به نظر می‌رسد یک چیز دیگری هم مسلم انگاشته شده است: جمع‌آوری میزان بیشتری از داده‌های خام می‌تواند احتمال بیشتری برای کشف علمی فراهم می‌کند.

فارغ از بحث مبانی منطقی یا روان‌شناختی استقرا و عدم امکان حصول یقین در باب یافته‌های آن، در ظاهر در این تردیدی نیست که در گزاره «همه الف‌ها ب است»، هرچه تعداد بیشتری الف ب و تعداد کمتری الف غیر ب یافته شود، اعتماد به صدق گزاره بیشتر می‌شود؛ اعم از اعتماد منطقی یا روان‌شناختی - جامعه‌شناختی. باری، بحث میزان داده‌ها فقط منحصر به تعداد شواهدی نبوده است که یک گزاره مشاهده‌تی (مثلاً هرچه ارتفاع بالاتر باشد، فشار هوا کمتر است) را تقویت می‌کند. گاه یک نظریه به درجاتی از پیچیدگی می‌رسد که به

تکنولوژی‌های نیمه اول قرن بیستم و نیمه اول قرن بیست و یکم به نسبت میان کشف‌های بزرگ نیمه دوم قرن نوزدهم و نیمه دوم قرن بیستم یکسان باشد. کارکرد تکنولوژی‌های متأخر هم بیشتر متوجه سخت‌افزارهای بسیار ریز و نرم‌افزارهاست. باری، به هر حال، سخن ما در اینجا در اصل متوجه علم است نه تکنیک.

اگر این ادعا در کلیت خود قابل تأمل باشد، پرسش بعدی چرای آن است. پاسخ‌های گوناگونی به این پرسش داده شده است؛ از جمله در یک نظرسنجی از ۲۷۰ نفر از دانشمندان، به مسئله رقابت بر سر گرفتن اشاره شده که معیار آن تعداد مقالات است (Bellus, 2016). این رقابت بر سر چاپ تعداد بیشتری مقاله، باعث سوگیری مقالات و گزارش‌های علمی به سویی می‌شود که انتظار کشف علمی مهم از آن کمتر می‌رود؛ از جمله اینکه مقالات به موضوعاتی رو می‌آورند که به نتیجه مشخصی برسد؛ چون پذیرفته نیست کسی بگوید روی این فرضیه مدت زیادی کار کردم و سرانجام داده‌ها آن را تأیید نکرد.

برای همین کمتر کسی به راه‌های نرفته و به سوی سرزمین‌های دوردست قدم می‌گذارد. نیز گاه مقالات با توجه به علایق حرفه‌ای اسپانسرهای محتمل به سمتی می‌رود که در بازار قابل استفاده باشد. دسته دیگری از پاسخ‌ها اشاره دارد به مشکل هم عصری و ضرورت فاصله گرفتن از یک دوره برای درک اهمیت اتفاقات آن، و نیز وفور دانشمندان، و از آن مهم‌تر، جانمایی تیم‌های تحقیقاتی به جای دانشمند منفرد. و موارد دیگری از جمله اینکه سن رسیدن به خلاقیت علمی اکنون حدود چهار سال است؛ در حالی که در اوایل قرن حدود سی سال و در قدیم مثلاً زمان لئوناردو داوینچی بیست سال بود. چون زمان بسیار بیشتری صرف آموختن میراث علم می‌شود تا نوبت به مرزهای علم و امکان کشف برسد (Leherer, 2011).

به‌دیگرسخن، بسا که دانشمندان امروز بیش از گذشتگان خود قوه ابتکار دارند و نیرو به خرج می‌دهند؛ ولی کار دشوارتر شده است. این توضیحات و جوهری از علل مشکل را توضیح می‌دهند؛ اما کافی به نظر نمی‌رسند. چون در این صورت لازم بود دست‌کم از سوی اقلیتی که وارد بازی گرنت نشده‌اند کارهایی صورت گیرد، و سرانجام در انتهای عمر دانشمندان یا از گروه‌های

Do we (preference) یا حتی طرح (establish) یک فرضیه. Do we have enough data to confirm; When is enough data enough; و مواردی از این دست. بنابراین تقاضا برای داده بیشتر وجود دارد و این تقاضا به بازار داغی برای داده‌های علمی منجر می‌شود.

بسیاری از دانشگاه‌ها مراکز جمع‌آوری داده به راه انداخته‌اند. نیز پیشرفت‌های نرم‌افزاری زمینه‌ساز کار با داده‌های انبوه شده است. تصور بر این است که صنعتی شدن علم به انقلابی در دسترسی به پایگاه‌های داده انجامیده است که عمده فعالیت‌های دنیای علم را در چنگ دارد. نیز تکنولوژی‌هایی فراهم شده به این منظور که سیستم‌های رایانه‌ای را به این توانا کند که کار روی داده‌های انبوه فراهم شده از شواهد تجربی را ساماندهی کرده و کارآمدی آن را افزایش دهند (Hidary, 2019).

نرم‌افزارهایی طراحی شده است تا مبالغی از داده را به‌طور اتوماتیک از حافظه فعال خارج به اصطلاح پنهان یا ذخیره کند تا زمانی که در طول فرایند پردازش به میزان بیشتری داده نیاز پیدا شود، یا به تعبیر دقیق‌تر فراوانی بیش از حد داده‌ها نسبت به امکانات پردازش مشکل ایجاد نکند و امکان پردازش دسته‌های محدود داده به صورت طبقه‌بندی شده فراهم باشد (vander, 1997). کاربست وسیعی از این شیوه به سبب رشد اطلاعات و داده‌ها گریزناپذیر به نظر می‌رسد؛ درحالی‌که از نگاه برخی، «این‌گونه فهرست کردن داده‌ها از منظر هزینه-فایده بررسی نشده است» (Murdock, 1980). ممکن است کسی امید برد که «این عصر طلایی داده‌ها بتواند به تصمیم‌گیری‌ها در حوزه علوم کمک کند. ولی انبوهی داده‌ها می‌تواند به شکلی پارادوکسیکال کار را دشوارتر کند» (tufekci, 2019).

ممکن است چنین به نظر بیاید که داده‌های انبوه وعده فهم بهتری از موضوع به کاربران می‌دهد؛ ولی «بیشتر همیشه بهتر نیست. گاه چنین می‌شود که به موازات درک بهتر، داده‌های انبوه مقدار زیادی هم‌بستگی‌های آماری مشکوک را القا می‌کند» (Machlis, 2013). پژوهشگران بسیاری مدل‌های آماری گوناگونی را برای تحلیل رشد داده‌ها به کار بسته‌اند، «ولی هیچ فرایند یا قاعده قابل تعمیمی از سوی آنان عرضه نشده است؛ لذا

آسانی نمی‌توان مجموعه‌ای از الف‌های ب را تعریف کرد که مشاهده تعداد بیشتری از آن موارد گزاره را تقویت و مشاهده موارد الف‌های غیر ب آن را تکذیب یا تضعیف کند.

بسیاری از فرضیات و نظریات علمی بر مجموعه‌ای از الف‌های ب، جیم‌های دال، ه‌های واو... متکی است؛ یعنی چنین نیست که فقط مشاهده همراهی دو پدیده برای تقویت فرضیه کافی باشد، تا بحث به گستره داده‌های خام و احتمال افزایش تعداد موارد مؤید محدود شود. مسئله مهم در اینجا این است که اگر شواهد بیشتر به تقویت بعضی و تضعیف بعضی دیگر از مجموعه گزاره‌های یک فرضیه انجامید تکلیف چیست؟ فرض کنیم مجموعه‌ای گسترده از مشاهدات نشان دهد که تعداد بیشتری از الف‌ها ب و جیم‌ها دال هستند؛ ولی هم‌زمان نشان دهد که ه‌ها و زهای جدیداً مشاهده شده واو و ح نیستند، در این صورت تئوری مورد بحث تضعیف شده است یا تقویت؟ فرض کنیم بر میزان مشاهدات و داده‌ها علی‌الدوام بیفزاییم و نتایج مشابهی به دست آوریم، به نظر می‌رسد افزایش داده‌ها گرهی از کار نمی‌گشاید و به تعیین تکلیف قطعی یا حتی احتمالی فرضیه کمکی نمی‌کند.

افزون‌براین، حالت سومی هم متصور است: حالتی که افزایش داده‌ها نه فقط کمکی به پیشرفت تبیین نمی‌کند؛ بلکه برای آن مشکل می‌آفریند. با افزایش مشاهدات و انباشت داده‌ها، ممکن است نه فقط دسته‌ای از گزاره‌های یک فرضیه تقویت و دسته‌ای دیگر تضعیف شود، بلکه معلوم شود گزاره‌های دیگری هم باید در نظر گرفته می‌شده که نشده است. افزودن آن مشاهدات و این گزاره‌های متناظر به مجموعه گزاره‌ها و مشاهدات مربوط به فرضیه، ممکن است آن را مبهم‌تر و بیشتر غیرقابل تصمیم کند. به‌دیگر سخن، داده بیشتر به دور شدن بیشتر از اثبات یا تکذیب (تأیید یا تضعیف) فرضیه منجر می‌شود.

در بیشتر رشته‌ها و در بسیاری از موضوعات پژوهشی، حجم قابل ملاحظه‌ای از مقالات یافت می‌شود که حتی در عنوان آنها تعبیر enough data به صورت انکاری یا استفهامی به کار رفته است حاکی از انکار یا تردید در کفایت داده برای اثبات (determination) یا ترجیح احتمالی (probabilistic)

این سؤال مطرح می‌شود که دو نظریه را که با داده‌های انبوه سازگار درمی‌آیند چطور می‌شود مقایسه کرد» (Saxena, 2007).  
 فقره ذیل به عنوان موضوع یک تحقیق گویا است: «چگونه تکنولوژی‌های جدید تعمیم و توسعه داده‌های مربوط به محیط زیست از روی رفتار شهروندان به مشکلات جدیدی در تحلیل و معنادار کردن داده‌ها میدان می‌دهد؟» (gabrys, 2016). گفته شده است که گاه داده‌های محدود پیمایش‌ها به دلیل محدودیت و حساب‌شدگی قابل اعتمادتر از داده انبوه رسمی و کمی است (Cunningham, 1997). نیز بر این تأکید شده است که «الگوهای هوش مصنوعی موجود، از جمله BACON، یک الگو برای کشف علمی، به این جهت ناکارآمدند که سهم بالای وجه مفهومی یا غیرداده‌ای تحقیق را نادیده می‌انگارند» (Chalmers, 1992).

بر مبنای پیش‌گفته بیان می‌کنیم افزون‌بر عواملی که پیش‌تر به اشاره نقل شد، عامل مهم دیگری در رکود علم مؤثر است، و آن اینکه با کشف شدن آن کشفیاتی که به نسبت آسان‌تر و دم‌دست‌تر بوده، به جاها و مسائلی رسیده‌ایم که دشواریاب‌تر است؛ اما این دشواریابی از تمثیل کشف و تکمیل جغرافیای یک قاره فراتر می‌رود، و تا حدود زیادی ناشی از حجم فزاینده و انبوه پرسش‌های مطرح، و حجم فزاینده و انبوه داده‌هاست که یافتن ربط علی یا هم‌بستگی آماری را، حتی در مقام حدس اولیه یک فرضیه، دیریاب‌تر می‌کند؛ چون مسئله‌های علمی را بیشتر به چند معادله و چند مجهول تبدیل می‌کند؛ و نیز آزمون کردن فرضیات، با توجه به تعدد بیشتر عوامل دخیل، دچار وضعیتی می‌شود که تکرار یک تجربه در شرایط مشابه و فرض عدم دخالت عوامل دیگر دشواری و دیریابی فزاینده‌ای پیدا می‌کند.

هنگامی که با مجموعه محدودی از داده‌ها به ساخت یک فرضیه تبیینی پرداخته می‌شود، ممکن است فرضیه ما به دلیل در نظر نگرفتن بسیاری از ریزه‌کاری‌ها و عامل حاشیه‌ای، خام و ساده‌انگار باشد. حال اگر بر میزان و دقت مشاهدات افزودیم و حجم بزرگی از داده‌های نازک‌بینانه فراهم شد؛ البته از آن ساده‌سازی و خام‌اندیشی فراتر رفته‌ایم؛ ولی لزوماً به فرضیه یا نظریه بهتری نمی‌رسیم؛ چون گاه داده‌ها به قدری متنوع می‌شوند

که قرار دادن همه آنها در یک نقشه تبیینی دشوار می‌شود. به‌دیگرسخن، دقت بیشتر و داده بیشتر، بسا که به ضعف قوه تبیینی منجر شود. گویی میان پشتوانه تجربی و غنای تبیینی رابطه مستقیمی وجود ندارد. به همین دلیل فرضیه‌سازی در علم گاه از درجاتی از ساده‌سازی ناگزیر است. معمولاً بنا بر این است که برخی مؤلفه‌هایی که به نظر می‌رسد تأثیر جدی و تعیین‌کننده ندارند در نظر گرفته نشوند، تا فرضیه بیش از حد شلوغ و امکان‌سنجش تجربی آن دشوار نشود.

به نظر می‌رسد این بیش از آنکه استثنا باشد قاعده است. بسیاری از نظریه‌های علمی مشهور مثل قانون جاذبه عمومی، تکامل انواع زیستی، مدارهای الکترونی عناصر و مانند آن، که هم از اهمیت برخوردارند هم اثبات شده تلقی می‌شوند، قطعیت و دقت آنها اغلب مشروط به در نظر نگرفتن برخی پارامترهاست که، بنا به فرض و در چارچوب میزان خاصی از حساسیت، تأثیر اساسی ندارند.

از نگاه برخی نظریه‌پردازان روش‌شناسی علم، حتی آن دسته از نظریه‌های بزرگ تاریخ علم مدرن که از حیث دقت و پیچیدگی و نوآوری اعجاب‌آور تلقی شده‌اند، آن‌قدرها هم دقیق نبوده‌اند. در مشاهدات و محاسبات، اعدادی تا حدودی نزدیک به مقتضای فرضیه اولیه به دست آمده و چون عوامل اصلی تأثیرگذار محدود بوده‌اند، یا در تقریب خاصی محدود تلقی شده‌اند، کشف بزرگ صورت گرفته است.

برای نمونه می‌توان به کشف قوانین سه گانه حرکت و مشخصاً جاذبه عمومی نیوتن استناد کرد. چنان‌که از سوی برخی فیزیکدانان بحث شده است اعداد و ارقام نیوتن هرگز به‌طور دقیق در قالب نظریه عکس مجذور قرار نمی‌گرفت (Rothman, 1970: 8; Duhem, 1989: 303)؛ ولی از آنجاکه عوامل اصلی دخیل در جاذبه اجرام، در حدی که عوامل مؤثر متعدد محتمل دیگر، با اثرگذاری‌هایی در اندازه کوچک، نتیجه را به کلی تغییر ندهند محدود بود، این کشف صورت گرفت. همین اثرات ظاهراً کوچک ولی در هر حال نهایتاً ناهم‌خوان بودند که بعدتر باعث شد توجه اینشتین به وجود عوامل مؤثر دیگر جلب شود و در قالب نظریه‌های نسبیّت، نظریه جاذبه عمومی را تعمیر و تکمیل کند؛ اما



پنهان‌تر و پیچیده‌تر دور از دسترس‌تر واقعیت است (چون وجوه نزدیک‌تر و آسان‌تر پیش‌تر کشف شده)، و از سویی انبوهی داده‌هایی که به یمن ابزارهای جدید در دسترس است.

در بسیاری از موارد، نظر به وجوه پیچیده‌تر امر واقع، افزایش گسترده داده‌ها همواره یا اغلب به‌طور هماهنگ به تأیید یا تضعیف گزاره‌های یک فرضیه یا نظریه منجر نمی‌شود؛ بلکه عوامل جدیدی را وارد عرصه کرده و تأثیرات احتمالی آنها را مطرح می‌کند که کار تبیین را از قبل مشکل‌تر می‌کند. رامیرز موارد متعددی را همچون دلیل رکود علم برمی‌شمرد؛ ولی موردی که به بحث ما در اینجا مربوط است اینکه: تکنیک‌های آماری و پردازش داده‌ها دائم پیچیده‌تر می‌شود؛ زیرا محققان به دشواری تحقیق شرایط کاملاً مشابه برای تکرار آزمایش‌ها آگاه شده‌اند. «نتیجه این شده است که آنچه به‌عنوان پژوهش عرضه می‌شود، به‌طور فزاینده‌ای، شامل نمونه‌های فراگیرتر و نتایج کوچک‌تر است؛ و به نظر می‌رسد این مشکل در چشم‌انداز آینده بدتر هم بشود» (Ramirez, 2018).

بر این مبنای، آنچه در دهه‌های اخیر ذیل عنوان علم پیچیدگی یا علوم پیچیده مطرح شده است، تا حدودی قابل تسری به بسیاری از حوزه‌های دیگر علم است. علوم پیچیده علمی است همچون آب و هوا، بیولوژی بدن انسان، امور مالی بازار، که فرایندهای مورد بررسی آن مشتمل بر شمار بالایی از اجزا و عواملی است که در کار هم‌کنش با یکدیگرند و از رهگذر این مجموعه تعاملات «پدیده‌های غیرقابل نادیده‌گرفتنی به وجود می‌آید که با تحلیل فقط اثر مستقیم خود اجزا و عوامل قابل تبیین نیست» (Zeng, 2017). نکته مهم در اینجا قید «غیرقابل نادیده‌گرفتن» یا صفت «نه کم‌اهمیت» است. معنای این سخن آن است که در بسیاری جاهای دیگر هم که به معنای خاص کلمه علوم پیچیده نیست، همین شبکه اجزا بسیار و هم‌کنش‌های زنجیره‌ای در کار است؛ هرچند برخی آثار آنها ممکن است قابل‌نظر باشد. به‌دیگر سخن، یک اتفاق به‌ظاهر ساده فیزیکی یا یک ترکیب شیمیایی هم در اصل می‌تواند شامل اجزای زیاد و هم‌کنش‌های زنجیره‌ای باشد؛ باری، نتایجی که این اجزا و هم‌کنش‌ها ایجاد

ادامه همین کار تکمیلی، یعنی نظریه میدان واحد (یکی کردن دو نظریه جاذبه و الکترومغناطیس که بخش عمده تلاش اینشتین در مدت سی سال آخر عمر و نیز تلاش محققان دیگری را در طول یک قرن گذشته به خود اختصاص داده است) به نتیجه مورد نظر یعنی یک فرمول واحد برای توضیح مجموعه پدیده‌هایی نرسید که در این دو نظریه وجود دارد (Tretkoff, 2005).

می‌توان حدس زد که (با فرض وجود میدان واحد) بخش مهمی از دشواری کار ناشی از تعدد عوامل دخیل (جاذبه و الکترومغناطیس) در فرایندهاست که باعث می‌شود پرداختن فرضیه‌ای که تمامی این عوامل را در خود دهد دشوار باشد، و سنجش تجربی آن دشوارتر؛ زیرا حتی با فرض امکان تشخیص عوامل و اندازه‌گیری آنها، بنا به نتایج اغلب غیردقیق مشاهدات دقیق (Evet, 1996)، بسا که تقریب‌های فراوانی که از مشاهدات متعدد ناشی می‌شود در هم ضرب شده و نتیجه را از چیزی دور کند که، همچون در مورد نیوتن، تقریبی قابل‌انکار باشد.

یک پژوهشگر حوزه دارو می‌نویسد مهم‌ترین مشکل کشف علمی ساختن فرضیات آزمون‌پذیر است؛ و در ابتدا پیدا کردن سؤالی مهم از میان انبوه سؤالاتی که می‌شود در هر حوزه علمی مطرح کرد. وقتی فرضیه‌ای دقیق و ربط مفروض میان متغیرها درست طراحی نشود، حال هر چقدر هم داده‌ها انبوه باشند، خروجی مطالعه به درد نخواهد خورد. مسئله‌های مهم محدودیت‌ها و ترکیب‌های پیچیده‌تری از عوامل را پیش می‌نهد و همین کار علم را سخت‌تر می‌کند. «این را می‌توان در موردی مثل کشف دارو ملاحظه کرد. دانشمند افزون‌براینکه باید دارویی کشف کند که فلان بیماری را برطرف کند (که خود با توجه به پیچیدگی‌های بیولوژیک فرضیه‌پردازی دشواری است)، چندین ملاحظه دیگر را هم، از قبیل قیمت، محیط زیست، امکان تهیه، عوارض جانبی، دشواری آزمون روی انسان، باید با استانداردهای سختگیرانه امروز در نظر بگیرد که هر کدام کار را دشوارتر می‌کند» (Van Lang, 2017).

براساس این، می‌توان گفت یکی از علل اینکه علم تا اوایل قرن بیستم رشد سریعی از خود نشان داد ولی پس از آن از سرعت پیشرفت آن کاسته شد، از یک سو موضوعیت فزاینده وجوه

است. به‌عنوان مثال، در بررسی پدیده‌ای مثل حرکت در سطح شیب‌دار در کارهای گالیه، یا ترکیب شیمیایی سوختن در کارهای لاوزیه، آنچه نتیجه کلی مورد نظر دانشمندان در آن زمان بود، کمتر از انبوهی از عوامل ریز و فرعی اثر می‌پذیرفت؛ درحالی‌که در پدیده‌هایی مثل تغییر آب و هوا و امواجی که از فضا دریافت می‌شود، عوامل فرعی یا کمتر پیدا و هم‌کنش‌های آنها می‌توانند نتیجه را به کلی برهم زنند. حال اگر موضوعات باریک‌تر جدید علمی چنان است که «ممکن است میلیون‌ها نقطه تقاطع و فرایندهای دینامیک در آنها اتفاق بیفتد» (همان)، عجیب نیست اگر داده‌های انبوه هم به سختی بتواند به تدوین فرضیه و به‌ویژه به آزمون آن میدان بدهد.

### کلان داده

منظور ما در اینجا نفی ارزش چیزی نیست که تحت نام کلان داده مطرح است. کلان داده‌ها از قضا در برخی موارد، بیشتر در حوزه علوم اجتماعی، کارایی‌هایی دارند که از داده‌های محدود بر نمی‌آید. به‌عنوان مثال، جستجو درباره جستجوهای مردم در گوگل از وجود خصلت‌ها یا تمایلات یا دغدغه‌هایی پرده برمی‌دارد که در داده‌های محدود موجود در نظرسنجی‌ها به دشواری کشف می‌شود (Davidowitz, 2017). همین‌طور عکس‌های فراوانی که از فضای عمومی گرفته می‌شود و به کمک آن مثلاً رفتار ترافیکی مردم یک شهر یا منطقه در اندازه‌هایی بررسی می‌شود که نمی‌تواند تصادفی باشد (Herrera, 2010).

ولی نکته قابل توجه این است که، برخلاف نظر داویدوویتز، از این نظر انقلابی رخ نداده و دستاوردها بیشتر اکتشافی است تا تبیینی. در اینجا هم داده‌های انبوه چنان است که گاه ممکن است فرضیاتی متعارض را تأیید کند، یا اساساً دشوار بتوان از میان آن همه داده به فرضیه اولیه‌ای رسید. فرض کنید از روی انبوه عکس‌های موجود در اینستاگرام بتوان گرایشی عمومی را در کاربران تأیید کرد مبنی بر اینکه آنها تمایلی دارند به عکس گذاشتن از احوالات خوش خود و احوالات بد دیگران. یا با بررسی داده‌های انبوه پست‌های فیس‌بوک و مقایسه آن با جستجوها در آن به این نتیجه رسید که کاربران تمایل دارند تصویری فرهیخته‌تر و اخلاقی‌تر از آنی که

می‌کنند در بسیاری از موارد، به ویژه از منظر درجه حساسیت آزمایشگر، اهمیت چندانی ندارد.

به‌عنوان مثال، درصدهای پایینی از ناخالصی در ترکیب‌های شیمیایی با درصدهای پایینی از اصطکاک در پدیده‌های مکانیکی، که حاصل برخی عوامل فرعی و هم‌کنش‌های آنهاست، که فاقد اهمیت تلقی می‌شوند. بنابراین علوم غریب‌پسیده علمی یا بخش‌هایی از علوم است که در آنها می‌توان مواردی از عناصر و نیروهای دخیل را کم اهمیت انگاشت، وگرنه پیچیدگی و ذواجزا بودن فرایندها و اشتغال آنها بر شبکه‌ای از هم‌کنش‌ها بیشتر امری عمومی است.

معمولاً زنگ زدن یک تکه آهن را امری ساده تلقی می‌کنیم (که ضمن آن آهن و اکسیژن ترکیب می‌شوند)، و آتش گرفتن یک جنگل را امری پیچیده (که ضمن همان ترکیب با اکسیژن، انبوهی از عوامل همچون باد، باران، میزان خشکی جنگل، میزان انبوهی آن، درجه حرارت هوا، درجه رطوبت، جنس چوب درختان، و... در زمان‌ها و مکان‌های مختلف فرایند آتش‌سوزی، وارد معادله می‌شوند). می‌توان حدس زد که شبکه‌ای مشابه این مورد دوم در مورد نخست هم در جریان است؛ ولی از آنجاکه از منظر حساسیت بشری نسبت به این دو، تفاوت بسیاری وجود دارد، انبوه عوامل فعال در فرایند زنگ زدن یک قطعه آهن و هم‌کنش‌های آنها موضوع مهمی برای علم، به ویژه علم کاربردی، محسوب نمی‌شود.

دانشمندان همان‌قدر از فعل و انفعالات دقیق انبوه مولکول‌های یک رشته کوه ناآگاهند که از فعل و انفعالات انبوه مولکول‌های یک طوفان دریایی. باری، این دومی به دلیل سرعت و تحرک و حساسیت، حجم بیشتری از تحقیقات علمی را به خود جلب می‌کند و پرسش‌های ریز بیشتری را باید پاسخ بدهد. این پرسش‌ها در مورد نخست هم قابل طرح است؛ ولی اهمیتی ندارد.

به نظر می‌رسد یکی از دلایل سختی پیشرفت علم در دهه‌های اخیر این است که با مواردی بیشتر روبه‌رو شده است که از انبوه عوامل دخیل کمتر چیزهای کم اهمیت وجود دارد. به‌دیگر سخن، کشف‌های بزرگ پیشین همان‌هایی بوده است که این نادیده گرفتن برخی عوامل کم اهمیت در آنها مقدرتر بوده

دانشمندان بر این نکته تأکید کرده‌اند که در بسیاری موارد، داده‌های تجربی به‌خودی‌خود به فرضیه‌ای منتهی نمی‌شوند. این حدس و فرض پژوهشگر است که کلامی بر زبان آنها می‌گذارد که البته باید پس از آن حتی‌المقدور به محک تجربه نهاده شود. نیز آشنایان حوزه مطالعات کمی آگاهند که سنجش همبستگی میان متغیرها به صورت اکتشافی شانس زیادی برای به نتیجه رسیدن ندارد. باید نقشه‌ای (همان فرضیه) در دست باشد، وگرنه پژوهشگر در میان انبوه داده‌ها گم می‌شود. می‌توان پرسید دقیقاً چرا چنین است؟ چرا به فرضیه یا حدس نیاز داریم؟ پاسخ این است: به دلیل انبوهی داده‌ها.

فرض کنید چند ماده شیمیایی به ما بدهند و بخواهند که تأثیر آنها را بر یکدیگر بررسی کنیم. این نیازی به حدس و فرض ندارد. کافی است هر کدام را در معرض اثر دیگری قرار دهیم و نتیجه را ببینیم. نهایت اینکه باید هر آزمایش را با اندکی تغییر شرایط تکرار کنیم تا از نتیجه اطمینان بیشتری حاصل کنیم. حال اگر پرسش این باشد که چه اشیا یا بر هم تأثیر شدید می‌گذارند یا در قالب چه فرآیندی این تأثیر را می‌گذارند، مجبور به نوعی گزینش و برش هستیم؛ چون نه می‌توانیم همه اشیا عالم را آزمایش کنیم، و نه می‌توانیم در بررسی تأثیر دوشی بر هم، تمامی عوامل و اجزای موجود در هر کدام آنها و تمامی عوامل محیطی را در نظر بگیریم. همین برش و گزینش است که نام آن را فرضیه و حدس و نقشه می‌گذاریم.

فرضیه گزینشی از داده‌هاست که به هر دلیلی احتمال می‌دهیم ممکن است مربوط‌تر باشد؛ هرچند بسا که نادرست از آب درآید، یا حتی نادرست و درست بودن آن معلوم نشود. حال می‌گوییم اگر چنین باشد، افزودن داده‌ها در شکل اولیه آن بسا که چشم‌انداز کشف را مبهم‌تر می‌کند؛ چون این برش در داده‌هاست که راهی به رفتن می‌گشاید. در غیر این صورت، پژوهشگر زبان حالش این بیت حافظ می‌شود:

از هر طرف که رفتم جز وحشت نیفزود

زینهار زین بیابان وین راه بی‌نهایت

و یک فرضیه و حدس عالمانه مصداق این بیت:

هستند از خود نشان دهند. اینها البته ارزشمند و چشمگیر است، ولی تا حدودی شبیه همان کوه‌ها و دشت‌های بزرگ است که در آغاز هر علمی کشف می‌شود. باری، تا نوبت به موارد دورتر و ریزتر و پیچیده‌تر برسد، انبوهی داده‌ها چندان مشکل‌گشا نیست، و به سبب دشواری بیشتر طراحی نقشه تحقیق، حتی ممکن است گاه مشکل‌آفرین هم باشد.

اگر بخواهیم پس از دانستن مواردی مانند آنچه آمد، مثلاً اینکه مردم در نسبت با آن چیزی که آشکارا می‌گویند در واقع بیشتر نژادپرستند، گام‌هایی جلوتر بگذاریم، و مثلاً از این آگاه شویم که چه نسبتی میان متغیرهایی نه از قبیل سن و جنس و طبقه و درآمد و گرایش حزبی بلکه از قبیل گرایش دینی، قومیت، رادیکالیسم، خشونت، رشته علمی، شغل، وضعیت تأهل، نوع و میزان تربیت خانوادگی، تحولات زندگی، و مانند اینها با نژادگرایی وجود دارد، و چگونه میزان نژادگرایی یک فرد در طول زمان ممکن است تغییر کند، این تا حدی شبیه همان پرسش نجومی است که انواع امواج کیهانی که به منظومه شمسی می‌رسند از چه و کجاها منشأ گرفته‌اند، یا این پرسش زیست‌شناختی که عواملی که باعث می‌شوند گلبول‌های سفید گاه بافت‌های طبیعی خود بدن را به جای دشمن گرفته و مورد حمله قرار دهند چه هستند، انبوهی داده‌ها مشکل‌چندانی نمی‌گشاید، مادامی که تنوع آنها چندان است که به نظر می‌رسد یک شبکه با اجزا و اعضای زیاد ( $A1 - A_n$ ) و دینامیک ( $T1 - T_n, P1 - P_n$ ) در کنار آن کارسازی می‌کند، که شناسایی هر عامل، و تشخیص نوع و اندازه اثر آن، و برآورد مجموعه هم‌کنش‌ها و به اصطلاح بردار برآیند آن، به اضافه تغییرات ناشی از دینامیسم زمانی مکانی، چندان پیچیده و درهم است که حتی پیش نهادن فرضیه هم دشوار است؛ چه رسد به آزمایش آن. برای همین داده انبوه راه‌حل نیست. طنز داستانی که دیوید و ویس در کتاب کلان داده، داده‌های جدید (۲۰۱۷) آورده است به همین اشاره دارد. در یک کنفرانس بین‌المللی یکی از سخنرانان می‌گوید من چندین گیگابایت داده دارم، و سخنران بعدی می‌گوید اگر شما چندین گیگابایت داده دارید من چندین ترابایت دارم.

چنان‌که مشهور است بزرگان فلسفه علم و نیز برخی از

در این شب سیاهم گم گشت راه مقصود

از گوشه‌ای برون آی ای کوكب هدایت

### نتیجه

شواهد نشان از این دارد که رشد علم به‌ویژه وجوه نظری آن و آنچه به کشفیات بزرگ ناظر است در دهه‌های اخیر، با وجود افزایش علی‌الدوام ورودی بودجه و خروجی مقاله، دچار افت و رکود بوده است. مطابق برآوردهایی که کارشناسان، از جمله گزیده‌ای از خود دانشمندان کرده‌اند، دلایلی با سنخ‌های متفاوت در این کار سازسازی کرده است. یک سنخ از دلایل به سازوکارهای فعالیت‌های پژوهشی علمی در دانشگاه‌ها و مراکز پژوهشی به‌ویژه از منظر قواعد تشویقی راجع است. تشخیص برخی مشعر بر این است که این شیوه اساساً هر راهی به چشم‌انداز جدیدی را سد می‌کند؛ زیرا برای افزایش احتمال کشف‌های بزرگ این مهم است که راه برای نوعی ماجراجویی در پژوهش علمی باز باشد، و لازمه این آن است که کسان زیادی مقالاتی بنویسند در شرح فرضیات و حدس‌های علمی که داده‌ها سرانجام آن را تأیید نکرده است، یا پژوهش‌هایی تعریف و طراحی کنند که زمانی طولانی نیاز دارد، و در هر صورت، پژوهشگران مأجور و مشکور هم باشند. درحالی‌که نظام تشویقی تولید علمی در عمل چنین چیزی را نمی‌پذیرد. قواعد پذیرش مقالات هم تا حدودی به همین سمت و سواها هدایت می‌کند؛ به‌ویژه از این نظر که داورها دست‌کم تا حدودی در داخل پارادایم (های) رایج صورت می‌گیرد و نیز از این نظر که کمتر پیش می‌آید استادان طراز اول چنین داورهایی را به عهده گیرند.

سنخ دیگری از دلیل این است که در قدیم علم به این سبب پیشرفت می‌کرد که حرفه جمع‌کنندگی نبود. اگر کسی که از نبوغ برخوردار بود، با چشم دیگری به طبیعت نگاه می‌کرد، و ممکن بود بتواند رازی از آن را بگشاید، و اگر حوصله و استعداد کافی به همراه مقادیری شانس داشت، ممکن بود به اثبات یا تأیید تجربی و ریاضی آن موفق شود. در قرن نوزدهم این شیوه به اوج خود رسید و کشف‌های بزرگی صورت گرفت. از اوایل قرن بیستم،

چیزی به نام دانشمند یا پژوهشگر حرفه‌ای شیوع یافت. از این پس تعداد در حال افزایشی به پاییدن طبیعت و استنتاج از آن مشغول شدند. ولی اینها همچون جویندگان طلائی بودند که به صورت سازمان‌یافته و انبوه ولی دیر به معادن می‌رسیدند؛ رگه‌های روئین و قطعه‌های بزرگ، از سوی ماجراجویان منفردی استخراج شده بود که پیش از آن راه افتاده بودند که خبرها پخش شود. امکان پیشروی متأخران کندوکاو بیشتر و یافتن طلاهای خردتر بود. هرچند البته این ادعا که دیگر نباید انتظار کشف‌های بزرگ داشت، سخن محکمی نیست. کاملاً ممکن است از پس یک دوره فترت، دنیای علم به راه دیگری برود. باری، در مقام توضیح وضع فعلی می‌توان با مقادیری احتیاط چنین فرضی را پذیرفت.

یک توضیح دیگر هم در این میان هست، که این مقاله در مقام طرح و تقویت آن بود: افزایش توانایی‌های فنی جدید، اعم از افزار سخت همچون تلسکوپ و میکروسکوپ‌های پیشرفته و رایانه‌ها و ماهواره‌ها، و افزار نرم همچون برنامه‌ها و مدل‌های تنظیم و پردازش داده، و سرانجام در اختیار داشتن میزان بسیار زیادی از داده، تأثیر محدودی بر پیشرفت علم به ویژه کشف‌های بزرگ داشته است؛ مثلاً در قیاس با تلسکوپ بازیچه گونه‌ای که گالیله به کمک آن توانست اهله مشتری را ببیند و دلیل تجربی متقنی بر نظام خورشید مرکزی بیاورد. علت آن هم این است که هرچه داده‌ها بیشتر باشد، طراحی فرضیه‌ای که در میان آن داده‌ها دست به انتخاب بزند دشوارتر، و آزمون تجربی آن باز هم دشوارتر می‌شود. میان غنای تبیینی و پشتوانه تجربی بسا که یک رابطه وارونه وجود دارد. هرچه پارامترهای بیشتری در نظر گرفته شود تا از لحاظ تجربی متقن‌تر باشد، به احتمال بیشتر به روابط ضعیف‌تر و کمتر تأیید شده میان متغیرها می‌انجامد.



## منابع

- Belluz, Julia, et. all, (2016), "The 7 Biggest problems facing science, according to 270 scientists", Vox com, 7 September.  
<https://www.vox.com/2016/7/14/12016710/science-challenges-research-funding-peer-review-process>.
- Chalmers, David (1992), "High level perception, representation, and analogy: A critique of artificial intelligence methodology", *Journal of Experimental & theoretical Artificial Intelligence*, 4 (3), pp. 185-211.
- Collison Patrick, Nielsen, Michael (2018), "Science is getting less bang for its buck", *The Atlantic*, 16 November.
- Cunningham, Alastair (1997), "Quantifying survey data", Bank of England, *Quarterly Bulletin*, 37 (3), pp. 292-300.
- Davidowitz, Stephene (2017), *Every body lies: Big data, New data, and what the internet can tell us ...*, Dey Street Books.
- Duhem, piere (1989), *La Theorie physique, Son object, sa structure*, Paris, Vrin.
- Elowitz, Michael (2002), "Stochastic Gene Expression in a Single Cell", *Science*, 16 August, 297,( 5584), pp. 1183-1186, DOI: 10.1126/science.1070919
- Evet, IW (1996), "Expert Evidence and forensic misconceptions of the nature of exact science", *Science and Justice*, 36 (2), pp. 118-122.
- Gabrys, Jennifer, et. all (2016), "Just good enough data", *Big data & Society*, 1 december, Doi.org/10.1177/2053951716679677
- Gordon, Robert (2016), *The rise and fall of American growth*, Princeton university press.
- Herrera Juan (2010), "Evaluation of traffic data obtained via GPS", *Transportation Research: Emerging Technologies*, 18 (4), pp. 568-583.
- Hidary, Jack (2019), *Quantom computing: An applied Approach*, Springer.
- Horgan, John (1996), *The End of science*, Basik Books.
- Kim, Y.S. (2007), "Can you do quantum mechanics without Einstein?", *AIP Conference Proceedings* 889, 152, 15 March, <https://doi.org/10.1063/1.2713454>.  
<https://aip.scitation.org/doi/abs/10.1063/1.2713454>.
- Knaption, Sarah (2018), "Nine in Ten Cancers Caused by Lifestyle", <http://www.telegraph.co.uk/news/health/news/12055206/html>.
- Kurzweil, Ray (2005), The accelerating power of technology, *Ted Talk*.  
[https://www.ted.com/talks/ray\\_kurzweil\\_the\\_accelerating\\_power\\_of\\_technology?language=en](https://www.ted.com/talks/ray_kurzweil_the_accelerating_power_of_technology?language=en)
- Lehrer, Jonah (2011), The difficulty of discovery (where have all geniuses gone?), *Wired*, 26 January, <https://www.wired.com/2011/01/the-difficulty-of-discovery/>
- Machlis, Sharon (2013), "More data isn't always better", *Computerworld*, 8/ may/ 2013.
- Martin, Douglas (2005), *H. Bentley Glass, Provocative Science Theorist*, New York Times, 20 January
- Miyakawa, Tsuyoshi (2020), "no raw data, no science", *Molecular brain*, 13 (24), Editorial, 21 february.
- Moskvitch, Katia (2018), "What if everything we know about dark matter is totally wrong?", *Wired*, 28 September. <https://www.wired.co.uk/article/dark-matter-worth-searching-for-null-results>
- Murdock, John (1980), "Numerical data indexing", *Journal of chemical information and computer sciences* 20 (3), pp. 132-136.
- Pearson, Helen (2006), "What is gene?", *Nature*, 441, pp. 398-401, 24 may. <https://www.nature.com/articles/441398a>.
- Ramirez, Israel (2018), *Why isn't science advancing faster?*, *Quora*, 27 September.  
<https://www.forbes.com/sites/quora/2018/09/27/why-isnt-science-advancing-faster/#451d7e463c14>
- Rothman, Milton (1970), *Discovering the natural laws: The experimental basis of physics*, Dover publication, New York.



- Saxena, Anurag (2007), "Exploring Models for the growth of literature data", *Journal of library & information Technology*, 27 (3), pp. 45-67.
- Schwanhauser, Bjorn, et. all (2011). "Global quantification of mammalian gene expression control", *Nature*, 473, pp. 337-342.
- Tomasetti, Christian, Lu Li, Bert Vogelstein, (2017) "Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention", *Science*, Vol. 355, Issue 6331, pp. 1330-1334.
- Trethoff, Emie (2005), "Einstein's quest for a unified theory", *American Physical Society*, 14 (11). <https://www.aps.org/publications/apsnews/200512/history.cfm>
- Tufceki, Zeynep (2019), "Big data and small decisions", *Scientific American*, 320, pp. 3-73, doi: !0.1038/scientificamerican0319-73.
- Van Lang, Christopher (2017), *What is the most difficult part of drug discovery?*, *Quora*, 10 November <https://www.forbes.com/sites/quora/2017/11/10/what-is-the-most-difficult-part-of-drug-discovery?/#>
- Vander, steven (1997), "When caches aren't enough: Data prefetching techniques", *Computer* 30 (7), pp. 23-30.
- Zeng, An & Stanley, Eugene (2017), "The science of science: from the perspective of complex systems", *Physics Reports*, V. 714-715, 16 November, Doi. Org/10.1016/j.physicdp.2017.10001.

