

وضعیت آغاز شود. ایده‌ال آن است که تعامل سازنده‌ای با محوریت انجمن ریاضی ایران بین مؤلفان کتاب‌های درسی و کارگروهی تخصصی از سوی انجمن برقرار گردد. در غیر این صورت، انجمن می‌تواند با همکاری دانشگاه‌های بزرگ، چاره‌اندیشی نموده و نتایج را به صورت دستورالعمل‌های واحد به دانشگاه‌ها ابلاغ نماید.

یکی از راهکارهای عملی، تعریف دروس جبرانی یا اختیاری با سرفصلی است که نقصان‌های آموزشی مورد اشاره را پوشش دهد. با توجه به تراکم بالای سرفصل‌های درس ریاضی عمومی و شکایت جدی اساتید از کمبود وقت، امکان گنجاندن مطالب جدید در این دروس میسر نیست. بنابراین، ارائه‌ی راه‌حلی خارج از چارچوب فعلی، مانند همین پیشنهاد، ضروری به نظر می‌رسد.

*دانشگاه اصفهان

دوره متوسطه به شدت کاهش یافته است، به گونه‌ای که نمی‌تواند پایه مناسبی برای دروس دانشگاهی فراهم کند. از سوی دیگر، تراکم غیرمنطقی سرفصل‌های درس ریاضی عمومی ۱ در دانشگاه، فرصت لازم برای جبران این کاستی را از استادان و دانشجویان سلب کرده است. این وضعیت، مثلثات را به حلقه‌ای مفقوده در زنجیره آموزش ریاضی تبدیل کرده که نتیجه آن، ضعف گسترده دانشجویان در درک و به کارگیری این ابزار ریاضی حیاتی است.

بی‌تردید، همکاران محقق می‌توانند نمونه‌های بیشتری از کاستی‌های محتوایی را به موارد اشاره شده در این یادداشت بیفزایند. اساتید دانشگاهی که درس ریاضی عمومی را تدریس می‌کنند، در عمل به طور مستمر با پیامدهای این نارسایی‌ها دست‌به‌گریباند. به نظر می‌رسد زمان آن فرا رسیده که اقدام عملی برای بهبود این

علوم داده، علم داده‌ها یا علم داده

علی دولتی*

بایستی به کار ببریم؟ انتخاب معادل فارسی مناسب برای یک واژه بیگانه، تنها یافتن یک ترجمه تحت‌اللفظی نیست؛ بلکه فرآیندی است زبان‌شناختی که باید ایجاز، دقت، روانی و همخوانی با ساختارهای زبان مقصد را در نظر بگیرد. با این معیارها می‌توان به تحلیل این سه گزینه پرداخت:

علوم داده

این ترکیب، واژه «علوم» را جمع به کار برده است. در زبان فارسی، جمع بستن بخش اول یک ترکیب (مضاف) معمولاً بر مجموعه‌ای از دانش‌ها دلالت دارد. برای مثال، «علوم پزشکی» خود شامل شاخه‌های متعددی آناتومی، فیزیولوژی و بیوشیمی است. از آنجا که «دیتا ساینس» خود یک شاخه علمی واحد و مستقل محسوب می‌شود (هرچند بین‌رشته‌ای)، استفاده از جمع می‌تواند این تصور نادرست را ایجاد کند که با حوزه‌ای متشکل از چندین علم مختلف روبرو هستیم.

علم داده‌ها

این ترکیب بر کثرت و حجم انبوه داده‌ها تأکید دارد. اگرچه این ویژگی ذاتی این حوزه است، اما اضافه کردن «ها» باعث می‌شود ترکیب از

عبارت «دیتا ساینس»^۱ اصطلاحی است که برای یکپارچه‌سازی آمار، تحلیل داده‌ها، و روش‌های مربوط در نظر گرفته شده است [۳]، [۴]. این علم از سه مرحله‌ی طراحی برای داده‌ها، جمع‌آوری داده‌ها، و تحلیل داده‌ها تشکیل شده است. با این تعریف، دیتا ساینس به عنوان یک حوزه علمی میان‌رشته‌ای ترکیبی از علوم کامپیوتر (چگونگی پردازش داده‌ها شامل مهارت‌های برنامه‌نویسی، توانایی کار با داده‌های بزرگ و دانش الگوریتم‌ها، ساختار داده‌ها، و مهندسی نرم‌افزار برای تولید راه‌حل‌های مقیاس‌پذیر و کارآمد)، آمار و ریاضیات (چگونگی طراحی آزمایش‌ها، ایجاد مدل‌های آماری و یادگیری ماشین برای پیش‌بینی و استنباط) و دانش حوزه تخصصی (به منظور درک عمیق از زمینه‌ای که داده از آن می‌آید و توانایی طرح سوال درست و مرتبط از داده‌ها و تفسیر نتایج مدل‌ها در چارچوب واقعی کسب‌وکار یا مسئله علمی و تبدیل خروجی‌های فنی به بینش قابل اجرا) است [۲]. با ورود این مفهوم به زبان فارسی، معادل‌های مختلفی برای آن به کار برده شده است: «علم داده»، «علوم داده» و «علم داده‌ها». پرسش اینجاست که کدام یک از این سه گزینه را

^۱Data science

ساخت واژه فارسی است که هم از اصالت زبانی برخوردار است و هم از ظرفیت لازم برای نمایندگی یک حوزه علمی نوپدید. بنابراین، به منظور جلوگیری از چندپارگی زبانی و دستیابی به یکدستی در اصطلاح‌شناسی، توصیه می‌شود در تولید محتوای رسمی و تخصصی، مطبوعات و نیز گفتمان روزمره، معادل یکپارچه و پذیرفته‌شده «علم داده» برای «دیتاساینس» به کار رود تا علاوه بر تسهیل ارتباط، به معیارسازی و غنای هرچه بیشتر فارسی علمی بینجامد.

تشکر و قدردانی

از آقای دکتر محمد قاسم وحیدی اصل و آقای دکتر سیامک کاظمی که پیش نویس این یادداشت را مطالعه کردند و نظرات ارزشمند خود را در اختیار بنده گذاشتند، تشکر و قدردانی می‌کنم.

[1] Tukey, J. W. (1962). The future of data analysis. In Breakthroughs in Statistics: Methodology and Distribution (pp. 408-452). Springer, New York.

[2] Provost, F. and Fawcett, T. (2013). Data Science for Business: What you need to know about data mining and data-analytic thinking. O'Reilly Media, Inc.

[۳] اوریت، بی. اس و اسکروندال، ای. (۲۰۱۰). کتاب فرهنگ آمار کمبریج، ترجمه وحیدی اصل، م. ق. انتشارات مبتکران، ۱۴۰۱.

[۴] دونوهو، دیوید (۲۰۱۷). پنجاه سال علم داده، ترجمه وحیدی اصل، م. ق. فرهنگ و اندیشه ریاضی، سال ۴۳ شماره ۱، صص ۲۷۳-۳۲۷.

* گروه آمار، دانشکده علوم ریاضی، دانشگاه یزد

نظر آوایی کمی طولانی‌تر و در کاربردهای روزمره کم‌روان‌تر شود. از نظر دستوری، اگرچه این ترکیب اشتباه نیست، اما ضرورتی برای استفاده از آن احساس نمی‌شود.

علم داده

این ترکیب، متداول‌ترین گزینه معادل «دیتاساینس» در بین متخصصان این حوزه است. دلایل این اقبال روشن است:

- مختصر و گویا است: در کوتاه‌ترین شکل ممکن، مفهوم «حوزه علمی که موضوع آن داده است» را می‌رساند. جان توکی [۸] بنیانگذار این رشته نیز از «دیتاساینس» به عنوان «یک علم»، نام می‌برد.

- این نامگذاری در رشته‌های علمی مشابه آن نیز سابقه دارد: به عنوان مثال حوزه‌ی میان رشته‌ای «علم اطلاعات و دانش‌شناسی» که بر نحوه جمع‌آوری، ذخیره، مدیریت اطلاعات تمرکز دارد و از علوم کامپیوتر، ریاضیات، آمار، روانشناسی، جامعه‌شناسی و سایر رشته‌ها استفاده می‌کند.

- مقبولیت یافته است: زبان پدیده‌ای پویا است و در نهایت، انتخاب و کاربرد گسترده گویشوران یک زبان است که به یک واژه مشروعیت می‌بخشد. جامعه‌ی فارسی‌زبان به طور گسترده‌ای «علم داده» را پذیرفته و به کار می‌برد.

اگرچه هر سه معادل از حیث انتقال مفهوم اصلی تا حدودی موفق عمل می‌کنند، اما ارزیابی بر اساس معیارهای زبان‌شناسی رایج‌سازی واژه، شامل دقت معنایی، ایجاز، انسجام دستوری و میزان رواج، معادل «علم داده» مناسب‌تر است. این گزینش صرفاً یک ترجمه تحت‌اللفظی نیست، بلکه یک واژه‌پردازی آگاهانه و منطبق با قواعد