

کلان داده‌ها

جامعه اطلاعاتی

ترجمه و گردآوری

مستر سید علی اصغر رضوی

زبا محمدزاده

محمد کردان ششابی



عنوان و نام پدیدآور	: کلان داده‌ها: جامعه اطلاعاتی / ترجمه و گردآوری سیدعلی اصغر رضوی، زیبا محمدزاده، محمد کاردان نشاطی.
مشخصات نشر	: تهران: اساطیر پارس؛ چاپار، ۱۳۹۵.
مشخصات ظاهری	: ۳۳۸ ص: مصور، جدول، نمودار.
شابک	: 978-600-97299-0-6
وضعیت فهرست نویسی	: فیبا
یادداشت	: کتاب حاضر از دو کتاب "We are Big Data The Future of the Information Society, 2016" اثر سائدر کلسوس، نارت ویلارد و "Big data applications and use cases, c2016" اثر پاتریک سی. کی. هانگ به فارسی ترجمه شده است.
یادداشت	: واژه‌نامه.
یادداشت	: کتابنامه.
یادداشت	: نمایه.
عنوان دیگر	: جامعه اطلاعاتی.
موضوع	: داده‌های کلان
موضوع	: Big data
موضوع	: Information society
موضوع	: جامعه اطلاعاتی
شناسه افزوده	: شناسه سیدعلی اصغر رضوی، ۱۳۵۳ - مترجم، گردآورنده
شناسه افزوده	: شناسه زیبا محمدزاده، ۱۳۵۳ - مترجم، گردآورنده
شناسه افزوده	: کاردان نشاطی، محمد - ۱۳۲۸ - مترجم، گردآورنده
رده بندی کنگره	: ۱۳۹۵ ک۸ ۷۶/۹۷۳۵ - مترجم، گردآورنده
رده بندی دیویی	: ۰۰۵/۷
شماره کتابشناسی ملی	: ۴۶۱۲۹۱۲



تهران: ۸۸۸۹۹۶۸۰ - ۰۹۱۲۱۹۶۰۲۱۴

تهران - خیابان ولی عصر (عج) - بالاتر از زرنشت

کوچه نوریخش - پلاک ۴۰ - نشر چاپار

کلان داده‌ها، جامعه اطلاعاتی

ترجمه و گردآوری: دکتر سید علی اصغر رضوی

زیبا محمدزاده، محمد کاردان نشاطی

صفحه‌آرا و طراح جلد: رحیم کبیرصابر

ناظر فنی چاپ: مریم تفضلی

چاپ اول ۱۳۹۵

شمارگان: ۵۰۰ نسخه

شابک: ۹۷۸-۶۰۰-۹۷۲۹۹-۰-۶ ISBN: 978-600-97299-0-6

حق هرگونه چاپ و انتشار محفوظ است.

قیمت: ۱۷۰۰۰ تومان

پیشگفتار مترجمان

سپاس بزرگان خداوند عالم و همه عزیزانی که در تهیه این اثر به ما کمک کردند. امیدواریم در این راه گامی هرچند کوچک ولی مؤثر برداشته باشیم. کتاب حاضر، ترجمه و گراوری است از بخش‌های مهمی از مجموعه کتاب‌های:

الف) *We are Big Data: The Future of the Information Society* که به وسیله Sande Klous, Nart Wielaard در سال ۲۰۱۶ توسط انتشارات Atlantis Press و ب) *Big Data Applications and Use Cases* که توسط Patrick C. K. Hung در سال ۲۰۱۶ در *International Series on Computer Entertainment ...* که توسط انتشارات Springer منتشر شده‌اند.

در کتاب حاضر، موضوع‌ها و بخش‌های گوناگون و برای دانشجویان محترم تحصیلات تکمیلی - به‌خصوص دانشجویان علوم اطلاعات و دانش‌شناسی -، علاقه‌مندان و پژوهشگران به این حوزه تهیه و معرفی می‌شود. همچنین به بحث‌ها و موضوع‌های مهمی از جمله: تعریفی از کلان داده‌ها، ویژگی‌ها، کاربردها، نحوه استفاده و به‌کارگیری کلان داده‌ها در جامعه اطلاعاتی و نحوه استفاده آن در جامعه اطلاعاتی نوین و به راهکارهایی نیز اشاره شده است.

از آنجایی که در جامعه امروز با انبوهی از اطلاعات روبه‌رو هستیم، صرفاً با ذهنی متفکر و تحلیلی امکان‌پذیر است که بتوانیم کارها را به‌درستی و راحتی به جلو پیش ببریم. در واقع از مختصات جدید دنیای امروزی شفافیت است که باعث می‌شود شرایط بهتری را برای خود طراحی کنیم و بحث شفافیت به عنوان ابزاری برای شهروندان نیز در این کتاب به آن پرداخته شده است.

در خاتمه از کلیه عزیزان و دوستانی که در تهیه این اثر به ما کمک کردند و به‌خصوص انتشارات چاپار تقدیر و تشکر به عمل می‌آید. همچنین از همه عزیزانی که این کتاب را مطالعه خواهند کرد، در صورت مجال، کمبودها و نقص‌هایی را که احتمالاً مشاهده می‌کنند به بزرگواری خودشان بخشیده و راهنمایی‌های لازم را در جهت تکمیل این اثر مبذول فرمایند.

مسلم است که هیچ مقاله، کتاب یا نوشته‌ای عاری از کاستی نیست. از طرفی رفع کاستی‌ها جز با مساعدت و همکاری و نقد و بررسی پژوهشگران و علاقه‌مندان به این موضوع امکان‌پذیر نیست.

از اینرو هر ارائه‌کننده و پیشنهاد از جانب دوستان و پژوهشگران گران‌قدر که مطرح شود در تکمیل این اثر راهگشا خواهد بود.

دکتر سید علی اصغر رضوی

زیبا محمدزاده

محمد کاردان نشاطی

اسفند ۹۵

aa_razavi@yahoo.com

پیشگفتار نویسنده

کلان داده‌ها بسیار بااهمیت هستند. بازاریاب‌های شرکت‌های مشاوره و فناوری، سرعت در حال افزایش فضایل و خوبی‌های کلان داده‌ها را می‌ستایند. شعار آن‌ها این است که به کمک فناوری نوین اطلاعات هر شرکتی می‌تواند خود را به یک ابرقدرت در صحنه کلان داده‌ها تبدیل کنند.

به شکلی مشابه، سران و کارشناسان خود خوانده سایبری و تصاویری وحشتناک از آینده کلان داده‌ها به تصور می‌کنند که در آن اراده انسانی حذف شده است. با گوش سپردن به سخنان آن‌ها، چنانچه نظر می‌رسد که با کلان داده‌ها، ما کامپیوترهای خود را به سلاح‌های گشتاور جمعی تبدیل خواهیم کرد. چنین پیش‌بینی‌هایی از روز رستاخیز ممکن است جذاب برای به نگارش در آمده در این زمینه را به فروش برساند اما پوچ‌گرایی درباره کلان داده‌ها را راستای بهبود درک ما از این واقعیت کلان داده‌ها و نیز اینکه کلان داده‌ها اقتصاد را چنانچه ما را تا چه میزان به شکلی بنیادین مورد تأثیر قرار خواهند داد، کار زیادی انجام نخواهد داد.

کلان داده‌ها اساساً چیزی بیش از یک فرآیند نیست، یک مکانیسم منحصر به فرد از اینکه ما انسان‌ها چگونه واقعیت پیرامون خود را درک می‌کنیم و بر اساس این درک درباره آینده خود پیش‌بینی‌هایی بکنیم.

در تمام تاریخ بشریت، ما انسان‌ها با تماشای جهانی که پیرامون ما را احاطه کرده است و اندیشیدن درباره آن و به واسطه نظریه‌پردازی درباره چگونگی شکل‌گیری جهان به آن معنی بخشیده‌ایم. و ما برای اثبات (یا رد) نظریات خود از

جهان پیرامونمان اطلاعات را جمع‌آوری کرده‌ایم. با گذشت زمان، ما متوجه شدیم که نیاز داریم جهان را از منظری عقلانی نظاره کنیم و به شکلی روشمند در فهم آن شرکت کنیم. تا اینکه، سرانجام این امر منجر به ایجاد «روش علمی» شد و در قرن ۱۹ و ۲۰ در فهم و استفاده از جهان، پیشرفت بزرگی را به وجود آورد. در نتیجه، مردمان بسیار بیشتری، زندگی طولانی‌تر و بهتری را پشت سر گذاشتند.

اما تا همین اواخر، گردآوری، تجزیه و تحلیل و ذخیره‌سازی داده‌ها امری وقت‌گیر و پرهزینه بود. بنابراین، برای پاسخ به پرسش‌هایی که داشتیم تا جای ممکن داده‌های اندکی را جمع‌آوری کردیم. در واقع، در عصر داده‌های کوچک، مؤسسات ما و فرآیندهایی که ما طی آن از داده‌ها، کسب بینش می‌کردیم بر کمبود داده‌ها استوار بود. بنابراین ما بجای استفاده از کل داده‌ها از نمونه‌هایی از آن استفاده کردیم. بنابراین اگر موضوعی جالب می‌شد، به سرعت می‌توانستیم آن را از روی دقت نظاره کنیم زیرا ما به وضوح اطلاعات کامل را ذخیره نداشتیم.

پُر دردسرس‌تر آنکه با یک چنین روشی ما نمی‌توانستیم پرسش‌هایی را پاسخ گوئیم که از ابتدای امر با آن‌ها مواجه بودیم، نه اینکه، بتوانیم از داده‌ها به طرز مؤثر برای طرح پرسش‌هایی جدید استفاده کنیم که ممکن است به بینش‌های واقعاً بدیع منجر شود. ما همچنین کیفیت داده‌ها را ایمن کردیم، حتی کیفیت هزینه‌های بالا در راستای بهبود دقت و صحت اطلاعات اندکی که جمع‌آوری نمودیم، تلاش کردیم.

امروزه، بسیاری از موانع برای جمع‌آوری، تجزیه و تحلیل و ذخیره‌سازی مقادیر بزرگ داده‌ها تا حد بسیار زیادی کاهش یافته است. این امر بجای سامان بخشی به نمونه کوچکی از واقعیت، جمع‌آوری تمام داده‌ها را، یا نزدیک به تمامی داده‌ها را و نیز درک پیچیدگی‌های جهان ما را بسیار آسان‌تر می‌کند. در بسیاری از موارد بجای جمع‌آوری مقادیر بسیار اندک داده‌ها می‌توانیم داده‌های بیشتر و نیز نابسامان‌تری را گردآوری نماییم.

در نتیجه، اکنون ما می‌توانیم به دنیا با جزئیاتی بی‌سابقه بنگریم و تجزیه و تحلیل

کلان داده‌ها، پرسش‌های جدیدی را به وجود می‌آورد که ما را به سمت بینش‌هایی ارزشمند رهنمون می‌کند که در غیر این صورت ما به آن‌ها نمی‌رسیدیم. این همان روشی است که گوگل تلاش می‌کند تا برای پیش‌بینی گسترش آنفلوآنزا از طریق تجزیه و تحلیل میلیاردها درخواست جستجویی که دریافت می‌کند به آن دست یابد.

و این همان روشی است که شرکت‌های تازه راه‌اندازی شده همچون INRIX را قادر می‌سازد تا مشتریان خود را در اوج ترافیک و راه‌بندان در رفت و آمد روزانه خود هدایت کنند و همچنین به دولت‌های محلی درباره الگوهای در حال ظهور ترافیک مشاوره دهند.

اما تغییر سوم هم در چگونگی فهم ما از دنیا وجود دارد. ما انسان‌ها ترجیح می‌دهیم که دنیا را به عنوان تبدیلی از علت‌ها و معلول‌ها ببینیم. و بنابراین ما نمی‌توانیم جلوی خود را ببینیم. به دنبال دلایل هر چیزی بگردیم که مورد توجه قرار می‌دهیم سر و صداهای غیرمعمول، اتومبیلی که روشن نخواهد شد، تصمیم خریداران به خرید، نحوه‌ای که احساس می‌کنیم. اما همان‌طور که معلوم است، شناسایی علت‌ها امری دشوار است.

اکثراً تجزیه و تحلیل‌های کلان داده‌ها نمی‌توانند به علل حاصل از تأثیرات تجزیه و تحلیل دست یابند. غالباً این تحلیل‌ها نمی‌توانند چرایی یک موضوع را برای ما روشن کنند، آن‌ها فقط در مورد «چیستی» آن با ما سخن می‌گویند. اما، به شکلی شگفت‌انگیز، همین امر بیشتر وقت‌ها برای بهبود تصمیم‌گیری کافی است. این امر به گوگل کمک می‌کند تا ده‌ها زبان مورد استفاده را ترجمه کند و یک سوم فروش شرکت آمازون را محاسبه نماید. حتی این امکان را فراهم می‌کند که اتومبیل‌ها خود هدایت خود را در دست گیرند.

البته این موضوع به این معنی نیست که یک عمر تلاش ما در راستای فهم علت‌ها ناگهان باید به پایان برسد. به عبارت دیگر این موضوع به ما کمک می‌کند که بفهمیم پیش از آنکه بتوانیم بدویم، باید قادر به راه رفتن باشیم. با کلان داده‌ها برای

انجام درست این کار یک ابزار قدرتمند در اختیار خواهیم داشت. اگر از این ابزار به نحوی مناسب استفاده شود، تجزیه و تحلیل کلان داده‌ها «چیستی» امور را کشف خواهد کرد و به این ترتیب برای تجزیه و تحلیل‌های موشکافانه بعدی در باب علت‌ها به‌عنوان یک فیلتر ارزشمند عمل خواهد کرد.

این موضوع به ما کمک می‌کند تا بفهمیم دنیایی که در آن زندگی می‌کنیم بسیار پیچیده‌تر و جذاب‌تر و تفکر برانگیزتر از آن چیزی است که ما فکر می‌کردیم. همان‌گونه که ما متوجه می‌شویم اطلاعاتی که هم‌اکنون جمع‌آوری می‌کنیم ارزش بسیار بالقوه‌ای دارد. ما می‌توانیم از طریق کلان داده‌ها آن را آشکار کنیم، مشاغل نیز اینکه چگونه می‌توانند اینگونه چگونگی اقتصاد را تغییر شکل دهند و به تکامل جامعه کمک کنند را دوباره مورد نظر قرار می‌دهند.

این کتاب که به دست مصطفی در رشته تحریر درآمده است، درک درستی نسبت به کلان داده‌ها داشته و شما را به یک درک جدید از آن خواهید رسید و اینکه چگونه می‌توان کلان داده‌ها را برای اهداف خوب مهار کرد. این کتاب راهی را برای یک سفر هیجان‌انگیز ارائه می‌کند که می‌تواند به شما شگفت‌انگیز را موجب شود و ارزش قابل توجهی را تولید نماید. به شکلی با اهمیت بر شما خواهید شنید که چطور می‌توان قدرت حاصل از کلان داده‌ها را هدایت کرد به شکلی که بتوانیم از آن‌ها بهره‌مند شویم و از آسیب‌های حاصل از بعضی پیامدهای منفی آن اجتناب کنیم. کلان داده‌ها از اهمیتی بسیار برخوردار هستند. شما هم به آن‌ها دست بدهید.

ویکتور مایر شونبرگر (Viktor Mayer-Schönberger)

نویسنده همکار در کتاب داده‌های بزرگ: انقلابی که روش زندگی کردن، کار کردن و فکر کردن ما را تغییر خواهد داد.

فهرست مندرجات

۵	پیشگفتار مترجمان
۷	پیشگفتار نویسندگان
۱۷	مقدمه
۲۵	فصل ۱: مقدمه‌ای بر کلان داده‌ها
۲۷	خلاصه:
۲۸	۱. مقدمه‌ای بر کلان داده‌ها
۲۸	۱-۱. کلان داده‌ها چیست؟
۳۰	۱-۲. کاربرد کلان داده‌ها
۳۰	۱-۲-۱. بازاریابی
۳۳	۱-۲-۲. محاسبات فراگیر
۳۳	۱-۲-۳. اینترنت اشیا
۳۴	۱-۲-۴. شهرهای هوشمند
۳۵	۲. تجزیه و تحلیل داده‌های بزرگ
۳۷	۳. مطالعات موردی
۳۷	۳-۱. بهداشت و درمان
۳۸	۳-۲. خرده‌فروشی
۳۹	۳-۳. بیمه
۳۹	۳-۴. روند آنفلوآنزای گوگل
۴۰	۴. نگرانی‌های مربوط به امنیت و حریم خصوصی
۴۳	۴-۱. قانون حفاظت از اطلاعات شخصی و اسناد الکترونیکی (PIPEDA)
۴۴	۴-۲. قانون حفاظت از اطلاعات شخصی سلامت (PHIPA)
۴۵	۵. نتیجه‌گیری

فصل ۲: داده‌های بزرگ، بزرگ‌تر، بزرگ‌ترین	۴۹
خلاصه	۵۱
۱. کلان داده موضوع جدید نیست	۵۱
۲. رشد تصاعدی امکانات فناوری اطلاعات	۵۲
۳. تصمیم‌گیری بر اساس انجام کار نه بیان	۵۴
۴. کمیت بیشتر از کیفیت	۵۵
۵. احتمالات جدید	۵۷
فصل ۳: امواج اضطراب و غلبه بر آن	۶۷
۱. بهره‌یابی و سوءنویی از مردم	۶۹
۲. نقاط ضعف	۷۰
۳. رهیافت خطا نگر	۷۵
۴. سوءاستفاده	۷۶
۵. ممنوعیت	۷۹
۶. رفتار متناقض	۸۰
۷. اقبال عمومی هدایا	۸۱
۸. انسان، موجودی مدنی‌الطبع است	۸۳
۹. آژانس امنیت ملی آمریکا بهتر تحلیل می‌کند	۸۴
۱۰. عادت به داده‌های حریم خصوصی	۸۵
۱۱. تعارض بین موافقان و مخالفان کلان داده‌ها	۸۵
۱۲. ارائه ارزش‌های اجتماعی	۸۶
۱۳. ارائه ارزش اجتماعی در عمل	۸۷
۱۴. یک واکنش شدید عمومی ممکن است هیچ‌گاه دور نباشد	۸۸
۱۵. پلیدی داده‌ها	۸۹
فصل ۴: ادوات اطلاعاتی در دنیای نو	۹۲
۱. برخلاف تمام انتظارات	۹۴
۲. آیا زمان برای توسعه پایدار فرا رسیده است	۹۴

۳. زمان همکاری فرا رسیده است.....	۹۶
۴. آغاز عادت‌های نوین.....	۹۸
۵. ایجاد تغییر با ایجاد انگیزه یا توسل به زور.....	۹۹
۶. ارتباطات و بازخوردها.....	۱۰۱
۷. آیا تحقق یک تعادل جدید و پایدار امکان‌پذیر است؟.....	۱۰۲
۸. تمام هزینه‌ها را قابل مشاهده کنید.....	۱۰۴
۹. نظم، گزارش‌های جامع.....	۱۰۵
فصل ۵: تجربه و تحلیل داده‌ها در جامعه.....	۱۰۸
اخلاق در جامعه اطلاعاتی جدید.....	۱۱۰
۱. چگونگی توسعه مسئول فناوری.....	۱۱۰
۲. فناوری دو چهره دار.....	۱۱۱
۳. وابستگی شدید به فناوری.....	۱۱۱
۴. آیا تجزیه و تحلیل داده‌ها، ایمن‌تر است یا یک خطر محسوب می‌شود؟.....	۱۱۴
فصل ۶: پرسش و سؤال کردن با اهمیت‌تر از پاسخ.....	۱۱۸
نقش و تحول نظام آموزشی.....	۱۲۰
۱. بازشناسی سقراط.....	۱۲۰
۲. قابل اطمینان بودن اطلاعات.....	۱۲۱
۳. شما نمی‌توانید همه‌چیز را از دریچه کوچک دنیای خویش نگران کنید.....	۱۲۲
۴. نگاهی فرای حباب فیلتر.....	۱۲۵
۵. بیاموزید که در دنیای نوین اطلاعاتی چگونه عمل کنید.....	۱۲۷
۶. تغییر شکل و محتوا.....	۱۳۱
۷. زمان برای تغییر.....	۱۳۲
فصل ۷: شفافیت، ابزاری برای شهروندان.....	۱۳۸
تحولات شگرف توسط مردم صورت می‌گیرد.....	۱۴۰
۱. علائم راهنمایی برای ترافیک هوایی.....	۱۴۰
۲. قانون سستی.....	۱۴۰

۱۴۱.....	۳. قابل حصول بودن اطلاعات در جهان نوین و شفاف
۱۴۲.....	۴. در جامعه DIY، سازمان‌دهی امری بسیار سهل است
۱۴۳.....	۵. مشکل هر چه پیچیده‌تر، فرایند حل آن زودتر
۱۵۴.....	فصل ۸: رفتار ما را سیستم تعیین می‌کند
۱۵۶.....	قانون‌گذار ناتوان نیست
۱۵۶.....	۱. برد بودن نظام ترابری
۱۵۷.....	۲. اثر کرومیز
۱۵۹.....	۳. قدرت گوگل
۱۶۱.....	۴. پیشگیری سوءاستفاده از قدرت
۱۶۳.....	۵. فضای سایبری به‌عنوان جد موری مستقل؟
۱۶۵.....	۶. بی‌طرفی شبکه
۱۶۶.....	نتیجه‌گیری
۱۶۸.....	فصل ۹: داده‌های بدون ساختار، NoSQL، و حالت‌های تجزیه و تحلیل
۱۷۰.....	خلاصه
۱۷۰.....	۱. مقدمه
۱۷۳.....	۲. داده‌های بدون ساختار و NoSQL
۱۷۶.....	۳. تجزیه و تحلیل داده‌های بدون ساختار
۱۷۶.....	۳-۱. تجزیه و تحلیل به‌عنوان سرویس
۱۷۸.....	۳-۲. موضوعات آشکار
۱۷۹.....	۳-۳. کاوش داده‌های بدون ساختار
۱۸۰.....	۳-۴. استخراج واژگان
۱۸۲.....	۳-۵. موضوع‌ها
۱۸۴.....	۴. معماری بسترهای نرم‌افزاری AAAS
۱۸۴.....	۴-۱. جریان اجرای جستجو
۱۸۸.....	۵. مدل برنولی
۱۸۹.....	۵-۱. فضای برنولی

۱۹۰	۲-۵. مدل تک‌واژه برنولی
۱۹۱	۳-۵. مدل دو واژه برنولی
۱۹۲	۶. مدل هیدن مارکوف
۱۹۷	۷. APRIORI
۲۰۰	۸. روش‌های تحلیل گراف جهت‌دار
۲۰۱	۸-۱. جستجوی خطی
۲۰۲	۸-۲. موازی‌سازی کردن
۲۰۴	۸-۳. "تورتم‌های جستجوی بدینانه و جستجوی خوش‌بینانه"
۲۰۵	۹. بررسی الگوریتم
۲۰۵	۹-۱. مدل برنولی HMM
۲۰۷	۹-۲. مدل Apriori و HMM
۲۱۴	۹-۳-۱. مدت زمان جستجو
۲۱۵	۹-۳-۳. روش‌های گراف جهت‌دار
۲۱۸	۱۰. نتیجه‌گیری
۲۲۳	سخن آخر
۲۲۸	نمایه
۲۳۰	واژه‌نامه انگلیسی به فارسی
۲۳۴	واژه‌نامه فارسی به انگلیسی

مقدمه

فناوری، یکی از جدیدترین و پویاترین موضوعات باورنکردنی را ممکن کرده است. فناوری جدید اکنون بخش مهمی از زندگی ما است به شکلی که ما بیست سال پیش حتی ممکن نبود خواب آن را ببینیم.

ما می‌توانیم برخی از ابعاد و ارقام شگفت‌انگیز درباره ظرفیت فوق‌العاده محاسباتی مورد نیاز برای جستجوی موفق برای ذره هیگزبوزون^۱ را اعلام کنیم که در سازمان اروپایی سرن، واقع در ژنو انجام شد. این تحقیق یک تحقیق بنیادین بود که در رابطه با ذرات بنیادین انجام شد. نتیجه در لیگ قهرمانان تجزیه و تحلیل داده‌ها وارد شد. سندر^۲، دست‌اندرکار آزمایش‌ها در این سازمان بود که داده‌های به دست آمده از طریق شتاب‌دهنده ذرات را تجزیه و تحلیل می‌کرد. ما می‌توانیم توضیح دهیم که محصولات جدید با متن‌باز، پایه و اساس بسیاری از نرم‌افزارهای اطلاعاتی نوین و پیشگام هستند. ما می‌توانیم برای نشان دادن الگوریتم‌هایی تلاش کنیم که برای گوگل این امکان را ایجاد می‌کنند که چنین ساج خوبی را برای جستجوهای شما نمایش دهد.

جامعه اطلاعاتی نوین

هرکدام از این موضوعات یک موضوع فناورانه جالب است، اما در این کتاب تنها بحث مختصری درباره آن‌ها می‌کنیم. ترجیح می‌دهیم که به تأثیر این نوع از فناوری بر جامعه پردازش و به شکلی خاص به این موضوع پردازیم که این فناوری

1. Higgs boson particle

2. Sander

مرتبط با داده‌ها، پردازش داده‌ها، اطلاعات و ارتباطات چگونه می‌تواند جامعه ما را بهبود بخشد. دنیای جدیدی در حال ظهور است، دنیای جدیدی که ما در این کتاب نام آن را جامعه اطلاعاتی نوین نامیده‌ایم. این جهان، جهانی است که در آن همه چیز قابل اندازه‌گیری است و در آن مردم و تقریباً هر وسیله‌ای که فکرش را بکنید از طریق اینترنت با یکدیگر در ارتباط هستند. این شبکه‌ای از ارتباطات و حسگرها، میزان فوق‌العاده‌ای از داده‌ها را تولید می‌کنند و امکانات جدید و جذابی را ارائه می‌نمایند که روی هم رفته آن را اغلب کلان داده‌ها می‌نامند. بدون شک، این امکانات وسیع هم‌زمان با آن‌ها بهره‌برداری خواهد شد. هرچند بهره‌برداری از این امکانات قبلاً آغاز شده است.

آگاهی از این تغییر از طریق مبالغه و دولت‌هایی که از خود می‌پرسند چگونه باید به این تحولات پاسخ دهند گسترش می‌یابد. این موضوعی فراتر از پرداختن به محصولات یا خدمات جدید است. این موضوع درباره چگونگی پاسخ به جهانی کاملاً متفاوت است.

ما در میان یک پدیده اقتصادی کلاسیک هستیم: «نابودی سازنده». این امر به آن معنی است که ما تنها زمانی می‌توانیم به نوآوری دست یابیم که نظم کهن برچیده شده باشد، امری که تقریباً در تمامی بخش‌ها در حال وقوع است. کلان داده‌ها برای این نابودی سازنده، همچون هرمون رشد عمل می‌کنند. رشد غیرقابل انکار است. تمام حرف‌های ما درباره رسانه‌های اجتماعی، موسیقی آنلاین و انجام کارهای آنلاین و غیره، منجر به ایجاد مقادیری از ترافیک داده‌ها می‌شود که به‌ندرت قابل فهم هستند. و این کوه داده‌ها، هم‌اکنون که «اینترنت اشیا»^۱ میلیاردها وسیله را که پیش از این آفلاین بودند متصل می‌کند، به سرعت در حال رشد است. وسیله‌هایی همچون تلویزیون‌ها، یخچال‌ها، ابزارهای امنیتی، ترموستات‌ها، و آشکارسازهای دود، اکنون همگی داده تولید می‌کنند و به اشتراک می‌گذارند.

پخش این ارقام قابل توجه در اینجا امری و سوسه‌انگیز است اما این موضوع هدف این کتاب نیست. یک رقم واقعاً کفایت می‌کند: در دو سال گذشته میزان داده‌های ثبت شده ۱۰ برابر بزرگ‌تر از تمام داده‌هایی بوده است که در تمام طول تاریخ بشریت جمع‌آوری شده بود. این امر به نظر شگفت‌انگیز می‌رسد اما آنچه که واقعاً شگفت‌انگیز است این است که شما با این حجم از داده‌ها چکار می‌توانید انجام دهید. این همان موضوعی است که ما در این کتاب درباره آن بحث می‌کنیم.

مشارکت داده‌های تولید شده به شکلی تصاعدی به سطوحی نجومی افزایش می‌یابد، با این حال کلمه «بزرگ» در رابطه با کلان داده‌ها اندکی گمراه‌کننده است. بسیاری از نرم‌افزارها برای وی‌اچ‌اس یا تفسیر مقادیر عظیمی از داده‌ها ساخته نشده‌اند همان به اصطلاح «کلان داده‌ها» گفته می‌شود اما بیشتر برای تلفیق هوشمند داده‌ها طراحی شده‌اند. به عنوان مثال، ظهور «رایده خود سنجی» است، روندی که در آن تعداد رو به افزایشی از مردم، جنبه‌های محدود، از زندگی خود را اندازه‌گیری می‌کنند. آن‌ها می‌خواهند درباره آنچه که می‌خوانند و می‌نویسند، زمانی که حرکت می‌کنند یا اتفاقی که برای ضربان قلبشان در حین ورزش می‌افتد، اطلاعات کسب کنند. آن‌ها خودشان سطح جزئیاتی را که می‌خواهند اندازه‌گیری کنند، تعیین می‌نمایند. مثلاً آن‌ها می‌توانند میزان قهوه‌ای را که نوشیده‌اند یا مقدار شکلاتی را که خوردند به ثبت برسانند. از طریق برنامه‌هایی که بر روی گوشی‌های هوشمند یا ابزارهای دیگر نصب می‌شوند، آن‌ها به سادگی می‌توانند اطلاعات مرتبط با رفتار خود را، سلامتی و سبک زندگی‌شان را ثبت کنند و سپس با استفاده از این داده‌ها می‌توانند به عنوان مثال در خود این انگیزه را ایجاد کنند که به سطح بعدی تناسب اندام دست یابند. به لحاظ معنایی این امر، ارتباط اندکی با کلان داده‌ها دارد. با این حال، این امر بخشی از آن چیزی را شکل می‌دهد که جامعه آن را کلان داده در نظر می‌گیرد.

کلان داده‌ها در این کتاب

اصطلاح کلان داده‌ها به عنوان یک واژه عمومی و کلی برای استفاده، انتخاب

شده است. کلان‌داده‌ها در این کتاب شامل تمام فرصت‌های جدید، امکانات، فن‌ها و تهدیدات مرتبط با این واقعیت می‌شود. اینکه اکنون ما با داده‌ها می‌توانیم به شکلی متفاوت روبرو شویم، یا به بیان دیگر، جنبه‌های مثبت و منفی «داده‌سازی» از جامعه، جنبه‌های اجتماعی همچون حفظ حریم خصوصی و تأثیرات اجتماعی را شامل می‌شود. این موضوع نظارت گسترده رفتار شخصی توسط سازمان‌های اطلاعاتی، حل معضل ترافیک، توقف بیماری‌های همه‌گیر و غیره را شامل می‌شود. یک بخش مهم از داده‌های بزرگ، تجزیه و تحلیل داده‌ها است: جستجو برای الگوهایی در مقادیر بزرگی از داده‌ها که توان رسیدن به نتیجه‌گیری‌های مناسب را ایجاد می‌کند.

ارتباط با فناوری

ما بر این باوریم که کلان‌داده می‌تواند برای این دنیا بسیار ارزشمند باشد. ما برخی از مزایای شگفت‌انگیز این موضوع را هر روز تجربه می‌کنیم درحالی‌که فناوری تمام جنبه‌های زندگی ما را تحت تأثیر قرار داده است. ما همگی با فناوری مورد استفاده برای برقراری ارتباط با دوستان و اعضای خانواده در سرتاسر جهان آشنا هستیم. ما می‌توانیم از طریق گوشی‌های هوشمند، برای آن‌ها پیام‌های متنی ارسال کنیم، بر روی تبلت‌های خود با آن‌ها به صورت آنیمر بازی کنیم یا از طریق رسانه‌های اجتماعی با آن‌ها در ارتباط باشیم.

ما عکس‌ها و ویدئوها را در فضای ابری ذخیره می‌کنیم - که بسیار مفید است در آن هنگام که بخواهیم آن‌ها را به اشتراک بگذاریم - و ما از فضای ابری استفاده می‌کنیم تا از ذخیره تمامی اطلاعات مرتبط با کسب و کار خود بر روی هارد درایوهای خود اجتناب کرده باشیم. نگهداری اطلاعات به صورت آنلاین آسان‌تر است و خطر از دست رفتن اطلاعات را کاهش می‌دهد. ما کارهای بانکی خود را با یک نرم‌افزار، مرتب می‌کنیم. همان‌طور که پیترو هینسن به عنوان کارآفرین فناوری در یکی از کتاب‌هایش بیان می‌کند: «در هر کاری که ما انجام می‌دهیم دیجیتال یک بهنجار جدید است. زندگی دیجیتالی فقط این نیست که ما کاری را به‌طور معمول

انجام دهیم، ما همچنین به زندگی دیجیتالی معتاد شده‌ایم و انتظارات ما در مورد امکانات این زندگانی دیجیتالی به شکل فوق‌العاده‌ای افزایش یافته است.»

انتظارات

انتظارات در کجا به پایان می‌رسند؟ زندگی دیجیتالی در آینده چگونه خواهد بود؟ آیا همگی ما به عنوان جایگزینی برای گوشی‌های هوشمند خود عینک‌های گوگل یا سل‌های بعدی آن را بر چشمانمان خواهیم زد؟ یا آیا ساعت هوشمند غلبه خواهد کرد؟ شاید. آیا اشتغال در پنجاه سال آینده به نیم کاهش پیدا خواهد کرد؟ وظایف بیشتر و بیشتری همچون نگارش این کتاب را کامپیوترها به عهده خواهند گرفت؟ شاید. اما با سخ صادقانه این است که هیچ‌کس نمی‌تواند با قطعیت بگوید که استفاده از فناوری چگونه تمام خواهد یافت.

ما هیچ تردیدی نداریم که فناوری به تدریج در تمام جنبه‌های زندگی ما نفوذ خواهد کرد و حتی این امر در مهره‌های پیش رو بیشتر خواهد شد. این امر نه تنها به این خاطر است که اجتناب از آن غیرممکن است، بلکه به این دلیل است که ما فعالانه خواستار آنیم. زیرا ما زندگی عادی از راحتی، آسایش و ارزش افزوده‌ای را که فناوری به ارمغان می‌آورد، نمی‌خواهیم.

فناوری منجر به ایجاد انتظار می‌شود، ما باید به انجام کارهایی قادر باشیم؛ و برای همه چیز یک نرم‌افزار داشته باشیم. زمان انجام کار تمام حد ذات باید ۱۰۰ درصد باشد. شرکت‌ها ترجیحاً با هزینه‌ای برابر با صفر نیاز پیدا خواهند کرد که محصولات سفارشی تولید کنند. افزون بر این، تمام این‌ها باید سرگرم‌کننده و بادوام باشد.

تمام این موضوعات خلاصه کردن چالش جدید پیش روی شرکت‌ها، دولت‌ها، و دیگر سازمان‌ها را آسان می‌کند. غیرممکن باید فوراً در دسترس باشد.

احتمالات

کلان داده‌ها بیش از یک فناوری یا مفهومی صرفاً جدید است. کلان داده‌ها برای ما در تمام بخش‌ها و صنایع، فرصت‌هایی را ارائه می‌کند تا خودمان را به شکلی

متفاوت سازمان‌دهی کنیم، تا پیشرفت کنیم و کارهایی را به انجام برسانیم که تا همین اواخر به‌سادگی امکان‌پذیر نبود. بسیاری از مردم واژه «کلان داده» را با شرکت‌هایی مرتبط می‌دانند که می‌خواهند به مشتریان اقلام حتی بیشتری را بفروشند. با دانستن هر چیزی که برای دانستن درباره مشتریان وجود دارد. با این حال، این تنها یک روی سکه است، همان روی سکه که ما به یقین در این کتاب به آن خواهیم پرداخت. ما همچنان بر آنیم که نشان دهیم کلان داده‌ها این توانایی را دارند که مشکلات اجتماعی ما را حل کنند و زندگی‌هایمان را بهبود بخشند. ما می‌توانیم از کلان داده‌ها برای نشان دادن جان انسان‌ها استفاده کنیم. ما می‌توانیم پایداری و طرح‌های مقرون به‌صرفه را بهبود بخشیم. ما می‌توانیم ایمنی ترافیک را افزایش دهیم و از تراکم آن بکاهیم. ما می‌توانیم تولید محصولات کشاورزی را افزایش دهیم. ما می‌توانیم برای افراد سالخورده یک زندگی همراه با کیفیت بهتر فراهم کنیم. ما می‌توانیم جهان را پایدارتر کنیم. ما می‌توانیم در تشخیص و روش‌های درمان پزشکی یک انقلاب ایجاد کنیم. ما می‌توانیم ایمنی در پرداخت‌ها را افزایش دهیم. ما حتی می‌توانیم عملکرد یک تیم ورزشی را بهبود بخشیم.

ساختار

ما این کتاب را با یک مرور کلی بر آنچه که ما می‌توانیم با کلان داده‌ها انجام دهیم آغاز می‌کنیم و هم مزایا و هم اشکالات موجود در این راستا را به تفصیل شرح می‌دهیم. ما بحث می‌کنیم که چگونه به‌عنوان شهروند و مصرف‌کننده نظارتی غیرمتعارف داریم، اما همچنین در برابر اشتباهات سازمان‌ها برای اطلاعات شخصی خودمان احساس مقاومت می‌کنیم و به برای حفظ حریم خصوصی خویش بیمناک هستیم. در انجام این کار، ما به این نتیجه می‌رسیم که ما نمی‌توانیم در انزوا به کلان داده‌ها بپردازیم، زیرا کلان داده‌ها اساساً با پیشرفت اجتماعی در پیوند هستند.

به همین دلیل است که ما در بخشی‌هایی از کتاب درباره جامعه اطلاعاتی جدیدی بحث می‌کنیم که در حال ظهور است، از جامعه‌ای که کلان داده‌ها یک

بخش ضروری از آن است. ما نشان می‌دهیم که افراد به شکلی فزاینده می‌خواهند که عملکرد سازمان‌ها بجای آنکه صرفاً در راستای کسب منفعت باشند به جامعه نیز فایده‌ای برساند. ما نشان می‌دهیم که ما به عنوان جمعی از افراد بیش از هر وقت دیگر قدرتمند خواهیم شد و اینکه به عنوان یک جمع، ما این توانایی را خواهیم داشت که بر اقدامات سازمان‌ها تأثیری قابل توجه بگذاریم. در این دنیای جدید، ما باید برای قرار دادن موفقیت‌آمیز کلان داده‌ها در درون جامعه، خود راهی پیدا کنیم.

پس بر آنکه محدوده عمل خود را گسترش دادیم تا جامعه اطلاعاتی را به عنوان یک کل ببینیم، این عمل در بخشی از این کتاب به بحث در این مورد منجر شد که چگونه می‌توان کلان داده‌ها را در جامعه وارد کرد. مسائلی که مورد بررسی قرار می‌گیرند اصول اخلاق، امور ش و پرورش و نقش دولت را شامل می‌شود.

کلان داده‌ها در قالب سیوهای کنترل شده ممکن است به تشکیل یک جامعه بهتر کمک کند. این دقیقاً همان دلیل است که ما را بر آن داشت تا این کتاب را به رشته تحریر در آوریم. ما می‌خواهیم نشان دهیم که کمپانی‌ها و دولت‌ها هنوز نمی‌دانند که چگونه باید به مقوله کلان داده‌ها بپردازند. این موضوع با نادانی و سوءتعبیر دنبال می‌شود. پروژه‌های کلان داده‌ها اغلب پراکنده و بدون ساختار هستند. این موضوع به پروژه‌های ناموفق، ناآمیدی قابل توجه و مقاومت منجر می‌شود. این وضعیت باید تغییر کند. در سرتاسر این کتاب، هدف ما این است که نشان دهیم موضوعات تا چه اندازه ممکن است متفاوت باشند.