

**تجزیه و تحلیل داده های بزرگ:**

**برنامه های کاربردی، چشم اندازها و چالش ها**

**چکیده**

در دوران انقلاب صنعتی چهارم (صنعت 4.0)، داده‌های بزرگ تأثیر زیادی بر کسب‌وکارها دارند؛ چراکه انقلاب شبکه‌ها، سیستم‌عامل‌ها، مردم و فناوری‌های دیجیتال، عوامل تعیین‌کننده نوآوری و رقابت شرکت‌ها را تغییر داده است. تحریم‌های شدید برای داده‌های بزرگ از دانشگاهیان و متخصصان به‌دست‌آمده است، زیرا تجزیه‌وتحلیل داده‌های بزرگ به دانش ارزشمند و پیشبرد فعالیت‌های نوآورانه شرکت‌ها و سازمان‌ها، تغییر اقتصاد در سطح محلی، ملی و بین‌المللی منجر می‌شود. در این زمینه، علم داده‌ها به‌عنوان مجموعه‌ای از اصول اساسی تعریف‌شده است که اطلاعات و دانش را از داده‌ها به دست می‌آورند. تکنیک‌ها و برنامه‌های کاربردی که مورداستفاده قرار می‌گیرند، به تجزیه‌وتحلیل داده‌های حیاتی برای پشتیبانی از سازمان‌ها در فهم محیط و در تصمیم‌گیری‌های بهتر در طی زمان کمک می‌کنند. امروزه افزایش چشمگیر داده‌ها از طریق اینترنت اشیا (افزایش مداوم دستگاه‌های متصل، سنسورها و گوشی‌های هوشمند) موجب افزایش دوران "داده‌ها" شده است که تجزیه‌وتحلیل داده‌ها در هر بخش (کشاورزی، سلامت ، انرژی و زیرساخت‌ها، اقتصاد و بیمه، ورزش، غذا و حمل‌ونقل) و هر اقتصاد جهانی. گسترش روزافزون داده‌های موجود یک روند شناخته‌شده در سراسر جهان است، درحالی‌که دانش ارزشمند برخاسته از اطلاعات حاصل فرایندهای تجزیه‌وتحلیل داده‌ها است. در این زمینه، بخش عمده سازمان‌ها به جمع‌آوری، ذخیره و تجزیه‌وتحلیل داده‌ها برای تصمیم‌گیری‌های کسب‌وکار استراتژیک می‌پردازند که به دانش ارزشمند می‌انجامد. توانایی مدیریت، تجزیه‌وتحلیل و عمل بر روی داده‌ها ("سیستم‌های تصمیم‌گیری مبتنی بر داده‌ها") برای سازمان‌ها بسیار مهم است و به‌ عنوان یک دارایی مهم شناخته می‌شود. چشم‌انداز تجزیه‌وتحلیل داده‌های بزرگ اهمیت دارد و منافع سازمان‌های هدایت داده تعیین‌کننده‌های مهمی برای رقابت و عملکرد نوآوری هستند. بااین‌وجود موانع قابل‌توجهی برای اتخاذ رویکرد مبتنی بر داده‌ها وجود دارد و از طریق داده‌های بزرگ دانش ارزشمندی دریافت می‌کنند.

**کلیدواژه‌ها:** داده‌های بزرگ ⋅ تجزیه‌وتحلیل داده‌های بزرگ ⋅ عملکرد ⋅ شرکت مدیریت دانش ⋅ اینترنت اشیا (IoT)

**1. مقدمه**

داده‌ها به‌عنوان منبع حیاتی تصمیم‌گیری و مواد خام برای پاسخگویی مشخص می‌شود. بدون داده‌های با کیفیت بالا فراهم کردن اطلاعات مناسب در مورد موارد مناسب در زمان مناسب ، طراحی، نظارت و ارزیابی سیاست‌های مؤثر تقریباً غیرممکن است [1]. در این زمینه، توجه دائمی دانشگاهیان و متخصصان به رویکرد داده‌ها و داده‌ها وجود دارد، زیرا دانش ناشی از فرایندهای تجزیه‌وتحلیل داده‌ها منجر به ترویج فعالیت‌های نوآورانه، تغییر سازمان‌ها، بنگاه‌ها و اقتصادهای ملی می‌شود.

امروزه، در چهارمین دوره انقلاب صنعتی، سازمان‌ها و دولت‌ها بر توسعه توانایی‌هایی تمرکز می‌کنند که دانش را از مجموعه داده‌های بزرگ و پیچیده‌ای که معمولاً "داده‌های بزرگ" نامیده می‌شوند،. فراهم می‌کند. داده‌های بزرگی در سال‌های اخیر درزمینه های تجارت و اقتصاد یک کلمه کلیدی است، زیرا نقش مهمی در فعالیت اقتصادی ایفا می‌کند و نقش آن در ایجاد ارزش اقتصادی با ایجاد راه‌های جدید برای تحریک نوآوری و رشد بهره‌وری را تقویت می‌کند. ازاین‌رو، توانایی مدیریت، تجزیه‌وتحلیل و عمل در چارچوب سرمایه مبتنی بر دانش (KBC) مهم است که با اطلاعات دیجیتالی، ظرفیت نوآوری و جنبه‌های اقتصادی ارتباط دارد [2].

در آن دوران، بسیاری از بنگاه‌های مستقل از ابتدا تا سازمان‌های بزرگ تلاش می‌کردند فرهنگ رقابتی با محوریت داده‌ها را برای مزیت رقابتی در برابر رقبا به دست بیاورد. شرکت‌ها بهره‌گیری از داده‌های تولیدشده در سازمان‌ها از طریق عملیات خود برای به دست آوردن بینش ارزشمند برای تصمیم‌گیری بهتر، سریع‌تر و دقیق در مسائل مهم کسب‌وکار را هدف خود قراردادند.

ظهور وب 2.0 به کاربران اجازه می‌دهد در سیستم‌عامل‌های اجتماعی با یکدیگر در ارتباط باشند، شرکت‌ها را قادر ساخت به مقدار زیادی اطلاعات به صورتی آسان‌تر و ارزان‌تر دسترسی داشته باشند. علاوه بر این، ظاهر وب 3.0 فرصت‌های قابل‌توجهی برای جمع‌آوری اطلاعات خارجی فراهم می‌کند. دستگاه‌های همراه (تلفن‌های هوشمند و تبلت ها) شرکت‌ها را برای سنجش دقیق‌تر ترغیب می‌کند، زیرا این دستگاه‌ها، هر دو از طریق اینترنت و موبایل فعال هستند، توانایی ارتقای پردازش‌ها و معاملات بسیار متحرک، محل آگاه و متمرکز را دارند. این قابلیت در طول سالها به ارائه چالش‌ها و فرصت‌های تحقیق منحصربه‌فرد ادامه می‌دهد [3].

شرکت‌های دیجیتال مانند گوگل، آمازون و فیس‌بوک اهمیت داده‌های بزرگ را برجسته می‌کنند و راه‌های مختلفی را نشان می‌دهد که می‌توانند از زنجیره عرضه تا رضایت مشتری استفاده شوند و مزایای شرکت‌ها را برجسته می‌کند. بسیاری از شرکت‌ها از فرصت‌های ناشی از توسعه گسترده‌ی فناوری‌های داده بزرگ بهره‌مند شدند. امروز، شرکت‌ها در هر بخش صنعت و نه محدود به بخش فناوری اطلاعات و ارتباطات بر روی بهره‌برداری از داده‌ها برای به دست آوردن مزیت رقابتی متمرکز شده است، درحالی‌که تصمیمات مدیریتی به تحلیل‌های مبتنی بر داده‌ها و کمترین استفاده از تجربه‌ی رهبر بستگی دارد [4]. بااین‌وجود، بهره‌برداری از داده‌های بزرگ نیاز به افراد بامهارت و تخصص دارد که قادر باشند از داده‌ها ارزش را به دست آورد و دانش قابل‌توجهی را برای مدیران و تصمیم گیران فراهم کند.

**1.1 تعریف داده‌های بزرگ**

انتظار می‌رود که در سال 2025 نسل فوق‌العاده‌ای از داده‌ها به 180 ZB از راه برسد و نقش مهمی در تغییر و رشد شکل‌گیری «جهان دیجیتالی جدید» قرن بیست و یکم، از طریق دگرگون کردن بازارها و کسب‌وکارها داشته باشد. اطلاعات دیجیتالی از داده‌های پیچیده و ناهمگن از هرکجا و هر زمان به دست می‌آیند، معرف عصر جدیدی، به نام عصر «بزرگ داده‌ها» [6] هستند.

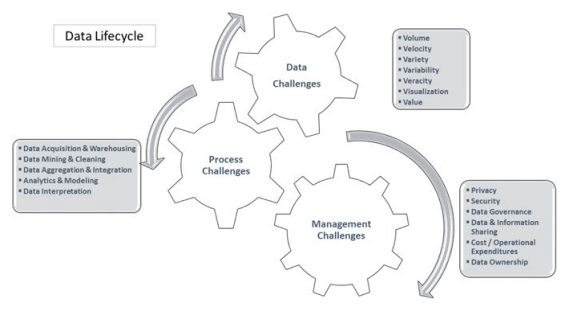
داده‌های بزرگ به مجموعه داده‌های بزرگی اشاره دارد که قادر به ضبط، ذخیره، مدیریت و تجزیه‌وتحلیل با استفاده از ابزارهای نرم‌افزاری معمول نیست [7]. مجموعه داده‌هایی که نه‌تنها در ابعاد بلکه در ناهمگونی و پیچیدگی (داده‌های ساختاری، نیمه ساختاریافته و غیر ساختاری) ازجمله عملیات، مبادلات، معاملات، بازاریابی و سایر اطلاعات، بزرگ هستند. علاوه بر این، داده‌های بزرگ شامل داده‌هایی می‌شوند که در فرمت‌های مختلف مانند متن، صدا، ویدئو، تصویر و موارد دیگر می‌آید. این داده‌های بدون ساختار رشد سریع‌تری نسبت به داده‌های ساختاریافته دارند و 90٪ از تمام داده‌ها را ذر برگرفته‌اند [8]. بنابراین، برای بیرون کشیدن بینش و محتوای درونی داده‌ها، جهت تصمیم‌گیری بهتر، اشکال جدیدی از قابلیت پردازش موردنیاز است.

در چرخه زندگی داده، چالش‌ها را می‌توان به سه دسته تقسیم کرد: داده‌ها، فرایندها و چالش‌های مدیریت (شکل 1) [6]. چالش‌های داده به ویژگی‌های داده‌های بزرگ شامل حجم، سرعت، تنوع و درستی و صحت آن اشاره دارند. چالش‌های فرایند با تکنیک‌های موردنیاز برای کسب اطلاعات بزرگ، انعطاف‌پذیری، تحول و تجزیه‌وتحلیل به‌منظور به دست آوردن بینش از داده‌های بزرگ مرتبط است. چالش‌های مدیریت اطلاعات عبارت‌اند از چالش‌های مربوط به امنیت داده‌ها، حفظ حریم خصوصی، نظارت و هزینه‌ها / هزینه‌های عملیاتی.

داده‌های بزرگ را می‌توان با هفت V مشخص کرد: حجم، تنوع، صحت، سرعت، متغیر، تجسم و ارزش.

حجم (Volume) به‌اندازه‌ی بزرگ مجموعه داده‌ها اشاره دارد. این‌یک واقعیت است که اینترنت اشیا (IoT) از طریق توسعه و افزایش گوشی‌های هوشمند، حسگرها و سایر دستگاه‌های متصل شده به همراه فناوری اطلاعات و ارتباطات سریع (ICT) ازجمله هوش مصنوعی (AI)، به نسل فوق‌العاده‌ای از داده‌ها (شمارش رکوردها، معاملات، جداول، فایل‌ها و غیره) کمک کرده‌اند. سرعت داده‌ها فراتر از قانون مور است و حجم تولید داده‌ها اقدامات جدیدی را برای ذخیره‌سازی داده‌ها به وجود آورده است، به‌عنوان‌مثال exabytes، zettabytes و yottabytes.

تنوع (Variety) نشان‌دهنده تنوع در حال افزایش منابع داده‌ها و فرمت‌های داده است. وب 3.0 منجر به رشد شبکه‌های وب و شبکه‌های اجتماعی می‌شود که منجر به تولید نسل جدیدی از داده‌های متفاوت می‌شود. مانند پیام‌ها، به‌روزرسانی‌ها، عکس‌ها و فیلم‌هایی که در شبکه‌های رسانه‌های اجتماعی مانند فیس‌بوک یا توییتر ارسال می‌شوند، اس ام اس، سیگنال‌های GPS از گوشی‌های هوشمند، معاملات مشتری در بانکداری، کسب‌وکار الکترونیکی و خرده‌فروشی، داده‌های صوتی در مراکز تماس و غیره. خیلی از منابع حیاتی داده‌های بزرگ ، نسبتاً نو و جدید هستند، ازجمله دستگاه‌های تلفن همراه که جریان‌هایی گسترده از داده‌هایی که با رفتار انسانی از طریق فعالیت‌ها و مکان‌هایشان مرتبط است، را تولید می‌کنند؛ یا منابع اینترنتی که از طریق ثبت دامنه، جریان کلیک و اقدامات رسانه‌های اجتماعی، داده‌هایی را تولید می‌کنند. علاوه بر این، داده‌های بزرگ همچنین در انواع داده‌ای که تولید می‌شوند، متفاوت است. بنابراین داده‌های بزرگ ، از داده‌های ساختار‌یافته (جداول، پرونده‌ها)، داده‌های بدون ساختار (متن و صدا)، داده‌های نیمه ساختاریافته (XML، RSS feeds) و سایر داده‌هایی مانند داده‌های حاصل از صوتی، ویدئو و سایر لوازم که طبقه‌بندی‌شان دشوار است، تشکیل می‌شود.



شکل 1 چالش‌های چرخه حیات داده

اغلب **متغیر بودن (Variability)** را با انواع مختلف (تنوع) اشتباه می‌گیرند، اما تغییرات با تغییر سریع معنا مرتبط است. به‌عنوان‌مثال، کلمات در یک متن می‌توانند با توجه به زمینه‌ی اطلاعاتی متن، معنای متفاوتی داشته باشند، بنابراین برای تجزیه‌وتحلیل مفهومی دقیق ، لازم است الگوریتم‌ها و چارچوب‌ها معنی (احساس) یک کلمه را با توجه به کل متن و زمینه‌ای که در آن به‌کاربرده شده است، پیدا کنند.

**سرعت (Velocity)** داده‌های بزرگ با سرعت بالای تولید داده‌ها مشخص می‌شود. داده‌هایی که توسط دستگاه‌های متصل و وب تولیدشده‌اند به شرکت‌ها در زمان واقعی می‌رسند. این سرعت برای شرکت‌ها در انجام اقدامات مختلف که آن‌ها را در برابر رقبا قادر به به دست آوردن مزیت رقابتی و سرعت و چابکی بیشتر می‌کند، بسیار بااهمیت است. باوجوداین واقعیت که برخی از شرکت‌ها در حال حاضر داده‌های بزرگ (داده‌های جریان‌های کلیک) را برای پیشنهاد دادن به مشتریان توصیه‌های خود ، ارائه کرده‌اند، شرکت‌های امروزه، حتی تحلیل‌کننده‌های های داده‌های بزرگ، توانایی تجزیه‌وتحلیل و درک اقدامات گرفتن اطلاعات را در زمان واقعی دارند.

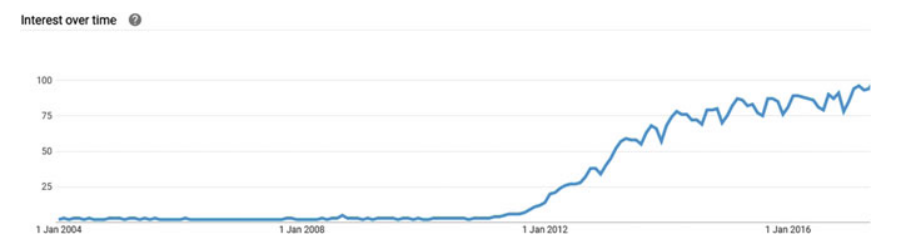
**صحت (Veraciy)** داده‌ها به اعتبار و دقت داده‌ها اشاره دارد. مجموعه‌ی داده‌ها دارای داده‌هایی است که تمیز و دقیق نیستند، بنابراین اعتبار و صحت داده‌ها به عدم قطعیت داده و سطح اطمینان مربوط به برخی از انواع داده‌ها اشاره می‌کند.

**تجسم (Visualization)** . تجسم داده یا مصور کردن داده‌ها ،علم تجسم بصری داده‌ها و اطلاعات است.این ویژگی اطلاعات کمی و کیفی را در برخی از اشکال تصویری مانند الگوها، روند، ناهنجاری، پایداری و تنوع که در فرم‌های دیگری مانند متن و جداول قابل ارائه نیستند، نمایش می‌دهد. [9].

نیرو و قدرت داده‌های بزرگ می‌تواند دانش ارزشمندی را ارائه دهد و بنابراین ارزش ارائه شده توسط فرایند تجزیه‌وتحلیل داده‌ها می‌تواند به نفع شرکت‌ها، سازمان‌ها، جامعه و مصرف کنندگان باشد.

بنگاه‌هایی که بر چالش‌ها غلبه می‌کنند و داده‌های بزرگ را به کار می‌گیرند، اطلاعات دقیقتری دارند و می‌توانند دانش جدیدی را ایجاد کنند تا به وسیله ی آن بتوانند استراتژی و عملیات تجاری خود را در مورد اهداف مشخص مانند بهره‌وری، عملکرد مالی و ارزش بازار بهبود بخشند. درحالی‌که داده‌های بزرگ نقش مهمی در تحول دیجیتالی شرکت‌های نوآورانه ایفا می‌کند. در نتیجه، افزایش علاقه به بهره‌برداری از اطلاعات بزرگ در میان نهادها و سازمان‌ها وجود دارد (شکل 2).

مزایای اقتصادی داده‌های بزرگ در کسب‌وکارهای خصوصی و دولتی در بریتانیا از 25.1 میلیارد پوند در سال 2011 به 216 میلیارد پوند در سال 2017 افزایش خواهد یافت [11]. داده‌های بزرگ می‌توانند در راه‌های مختلف در شرکت‌ها ارزش بیشتری به دست آورند و قادر به افزایش تولید و رقابت پذیری شرکت‌ها هستند. داده‌های بزرگ به رشد مداوم داده‌ها و فن آوری هایی که برای جمع‌آوری، ذخیره‌سازی، مدیریت و تحلیل داده‌ها ضروری است، اشاره می‌کند. شیوه تفکر در مورد کسب‌وکار با داده‌های بزرگ تغییر کرده است، زیرا عناصر اصلی سازمان‌ها و نه‌تنها مدیریت را تغییر می‌دهد. داده‌های بزرگ می‌توانند یک منبع کلیدی برای کسب دانش جدید، ارزش افزوده و پرورش محصولات، فرایندها و بازارهای جدید باشند، بنابراین داده‌ها به‌عنوان دارایی مدیران شرکت‌ها نشان داده شده و نشان‌دهنده اهمیت رویکرد داده‌ها در شرکت‌ها است [12] . شرکت‌ها اطلاعات را سالها جمع‌آوری می‌کنند، اما امروزه بیشتر و بیشتر شرکت‌ها در حال تجزیه‌وتحلیل داده‌ها به جای نگه داشتن آن‌ها هستند. ازاین‌رو، شرکت‌های داده محور در شرایط مالی و عملیاتی بهتر عمل می‌کنند، 5٪ تولیدی تر و 6٪ سودآورتر از شرکت‌های غیر داده‌های محور هستند، و برتری رقابتی بهتری نسبت به رقبای خود به دست می‌آورند [13].



شکل 2 روند داده بزرگ

<https://trends.google.com/trends/exploredate=all&q=٪2Fm٪2F0bs2j8q>

**1.2 تجزیه‌وتحلیل داده بزرگ**

تجزیه‌وتحلیل مجموعه داده‌های بزرگ در شرکت‌ها، اصطلاح تجزیه‌وتحلیل داده‌های بزرگ با دانش داده‌ها، هوش کسب‌وکار و تجزیه‌وتحلیل کسب‌وکار مرتبط است. علم داده‌ها به‌عنوان مجموعه‌ای از اصول اساسی تعریف می‌شود که اطلاعات و دانش را از اطلاعات به دست می آورد [4]. در طی سالهای گذشته، رویکردهای داده محور نظیر هوش کسب‌وکار (BI) و تجزیه‌وتحلیل کسب‌وکار برای شرکت‌ها عاملی ضروری است. BI به‌عنوان روش ها، سیستم ها و برنامه‌های کاربردی برای جمع‌آوری، تهیه و تحلیل داده‌ها برای ارائه اطلاعات به تصمیم گیرندگان تعریف می‌شود. به عبارت دیگر، سیستم‌های BI سیستم‌های تصمیم‌گیری مبتنی بر داده‌ها هستند [14]، درحالی‌که تجارت تجزیه‌وتحلیل تکنیک‌ها، فن آوری ها، سیستم ها و برنامه‌های کاربردی است برای تجزیه‌وتحلیل داده‌های حیاتی کسب‌وکار برای حمایت از آن‌ها برای درک محیط کسب‌وکار خود و کسب‌وکار تصمیم‌گیری در زمان. قدرت تجزیه‌وتحلیل کسب‌وکار به‌منظور ساده کردن حجم گسترده داده‌ها برای افزایش ارزش آن، درحالی‌که BI عمدتا بر داده‌های تاریخی در نمودار ها و گزارش های جدول داده‌ها به‌عنوان راهی برای ارائه پاسخ به پرسشنامه ها بدون بهینه سازی داده‌ها و افزایش ارزش آن تمرکز می‌کند.

تجزیه‌وتحلیل کسب‌وکار جهت معرفی عنصر تحلیلی اصلی در BI در اواخر 2000 آغاز شد. پس از آن، شرایط داده‌های بزرگ و تجزیه‌وتحلیل داده‌های بزرگ مورداستفاده قرار گرفته است برای توصیف تکنیک‌های تجزیه‌وتحلیل برای مجموعه داده‌هایی که بسیار بزرگ و پیچیده اند، نیاز به ذخیره‌سازی پیشرفته داده‌ها، مدیریت، تجزیه‌وتحلیل و تکنیک‌های تجسم است. در این محیط به سرعت در حال رشد، سرعت داده‌ها تبدیل داده‌ها به دانش ارزشمند را به سرعت ضروری می سازد. تفاوت بین تجزیه‌وتحلیل‌های معمولی و تجزیه‌وتحلیل سریع با داده‌های بزرگ در مشخصات تجزیه‌وتحلیل (نوع، هدف و روش)، ویژگی‌های داده (نوع، سن / جریان، حجم) و هدف اولیه است(جدول 1) [15، 16].

جدول 1 تجزیه‌وتحلیل اطلاعات متداول و بزرگ

|  |  |  |
| --- | --- | --- |
|  | **تجزیه‌وتحلیل متعارف** | **تجزیه‌وتحلیل داده‌های بزرگ** |
| نوع تجزیه‌وتحلیل | پیش بینی، توصیفی | پیش بینی، توصیفی |
| روش های تجزیه‌وتحلیل | مبتنی بر فرضیه | یادگیری ماشین |
| اهداف اولیه | پشتیبانی اصلی تصمیم‌گیری داخلی و مدیریت عملکرد | فرآیندهای کسب‌وکار و  مبتنی بر داده‌های واقعی |
| نوع داده | ساختار و تعریف‌شده (فرمت شده در  ردیف ها و ستون ها) | بدون ساختار و نامشخص  (فرمت‌های غیر ساختاری) |
| داده‌ها سن / جریان | > 24 ساعت مخزن استاتیک داده‌ها | <حداقل جریان دائمی داده |
| حجم داده‌ها | ده ها ترابایت یا کمتر | 100 ترابایت به پتابایت |

توسعه اینترنت و سپس اتصال وب به افزایش حجم و سرعت داده‌ها کمک کرده است. از اوایل 2000، فناوری‌های اینترنتی و وب برای جمع‌آوری و تجزیه‌وتحلیل داده‌های منحصربه‌فرد ارائه شده است. سیستم‌های وب 1.0 شرکت‌ها را قادر به ایجاد حضور وب می‌کنند و محصولات / خدمات خود را به صورت آنلاین در ارتباط با مشتریان خود ارائه می دهند. سیستم‌های وب 2.0، ازجمله معرفی شبکه‌های اجتماعی مانند فیسیوک، اطلاعات بیشتری را در مورد شرکت‌ها، محصولات و مشتریان ارائه می دهند. روند در حال افزایش دستگاه‌های تلفن همراه در برابر تعدادی از رایانه های عصر جدید، تجزیه‌وتحلیل کسب‌وکار را آغاز کرد، ازجمله تجزیه‌وتحلیل محتوای تولیدشده توسط کاربر توسط کانال های رسانه‌های اجتماعی. دستگاه‌های موبایل توانایی ترویج برای مثال پردازش‌ها و معاملات بسیار متحرک، محل آگاه و متمرکز شده را داراست. بنابراین، تصمیم‌گیری مبتنی بر داده‌ها بر روی داده‌های حاصل از همه منابع شرکت‌ها است، درحالی‌که پیش بینی ها و یادگیری ماشین بر اساس داده‌های سنتی و منابع نوین جدید مانند IoT و AI است.

تجزیه‌وتحلیل داده‌ها فرایند بازرسی، تمیز کردن، تبدیل و مدل سازی داده‌ها به دست آوردن اطلاعات مفید برای پیشنهادات و پشتیبانی در تصمیم‌گیری است. این جنبه ها و رویکردهای متعدد شامل تکنیک‌های متنوع تحت نام های مختلف در طرح های مختلف کسب‌وکار، علم و علوم اجتماعی است، درحالی‌که "تحلیل داده‌های بزرگ" اشاره به تکنیک‌های پیشرفته تحلیلی، با توجه به انواع مختلفی از مجموعه داده‌ها برای بررسی و استخراج دانش از داده‌های بزرگ، متشکل از فرآیند فرعی در به دست آوردن بینش از فرآیند داده بزرگ است. با استفاده از فناوری پیشرفته، Big Data Analytics (BDA) شامل مدیریت داده‌ها، برنامه نویسی منبع باز مانند Hadoop، تجزیه‌وتحلیل آماری مانند احساسات و مقیاس زمانی، ابزارهای ویژوال است که به ساختار و اتصال داده‌ها کمک می‌کند تا الگوهای پنهان، ارتباطات ناشناخته و بینشهای عملی دیگر.

روند BDA یک منبع برای تصمیم‌گیری‌های استراتژیک است که منجر به بهبود قابل ملاحظه ای در عملکرد عملیات، جریان درآمد جدید و رقابت پذیری نسبت به رقبا می‌شود. در این زمینه، روند به دست آوردن بینش از داده‌های بزرگ را می‌توان به دو مرحله تقسیم کرد: مدیریت داده‌ها و تجزیه‌وتحلیل داده‌ها. مدیریت داده‌ها با فرایندها و فن آوری های مربوط به تولید، ذخیره‌سازی، استخراج و آماده سازی برای تجزیه‌وتحلیل داده‌ها مرتبط است، درحالی‌که تجزیه‌وتحلیل داده‌ها به روش ها و تکنیک‌های تحلیل و تفسیر بینش از داده‌های بزرگ اشاره می‌کند [17] (شکل3)



شکل 3 روند استفاده از داده‌های بزرگ

تجزیه‌وتحلیل را می‌توان به چهار دسته تقسیم کرد، از تجزیه‌وتحلیل توصیفی و تشخیصی به تجزیه‌وتحلیل پیش بینی پیشرفته و ریاضی پیشرفته تر.

**تجزیه‌وتحلیل توصیفی**، بر اساس داده‌های تاریخی و جاری، منبع مهم بینش در مورد آنچه در گذشته رخ داده است و همبستگی بین عوامل مختلف تعیین الگوهای با استفاده از سنجش های آماری مانند میانگین، دامنه و انحراف استاندارد. تجزیه‌وتحلیل‌های توصیفی با استفاده از تکنیک‌های پردازش تحلیلی آنلاین (OLAP) از تجربه گذشته برای بهره‌گیری از دانش استفاده می‌کنند تا پاسخ هایی را در آنچه که در سازمان‌ها اتفاق می افتد، پاسخ دهد. نمونه های معمول از تجزیه‌وتحلیل توصیفی شامل تجسم داده‌ها، داشبورد ها، گزارش ها، جدول ها و نمودار ها ارائه معیارهای کلیدی شرکت‌ها ازجمله فروش، سفارشات، مشتریان، عملکرد مالی و غیره.

**تجزیه‌وتحلیل‌های تشخیصی** همچنین بر اساس اطلاعات تاریخی، بینش هایی راجع به علت ریشه برخی نتایج گذشته، ارائه می‌دهد. بنابراین، سازمان‌ها می‌توانند تصمیمات بهتر بگیرند و از اشتباهات و نتایج منفی گذشته جلوگیری کنند.

**تجزیه‌وتحلیل پیش بینی** در مورد پیش بینی و ارائه برآورد برای احتمال یک نتیجه در آینده، تعیین فرصت ها و یا خطرات در آینده است. با استفاده از تکنیک‌های مختلف ازجمله داده کاوی، مدل سازی داده‌ها و یادگیری ماشین، اجرای تجزیه‌وتحلیل پیش بینی برای هر بخش سازمان مهم است. یکی از شناخته‌شده ترین برنامه‌های این نوع تحلیل، پیش بینی رفتار مشتری، تعیین عملیات، بازاریابی و جلوگیری از خطر است. با استفاده از داده‌های تاریخی و دیگر موجود، تجزیه‌وتحلیل پیش بینی قادر به کشف الگوهای و شناسایی روابط در داده‌ها است که می‌تواند برای پیش بینی استفاده شود [17]. تجزیه‌وتحلیل پیش بینی در دوران دیجیتال یک سلاح قابل توجه برای سازمان‌های مسابقه ای است. بنابراین سازمان‌هایی که از تجزیه‌وتحلیل پیش بینی استفاده می‌کنند می‌توانند روند و الگوهای آینده را شناسایی کنند، محصولات نوین / خدمات و نوآوری را در مدل های کسب‌وکار خود ارائه دهند.

**تجزیه‌وتحلیل‌های پیش بینی کننده** پیش بینی تأثیر فعالیت‌های آینده را پیش از آنکه به آن‌ها برسد پیش بینی می‌کند و به «نتیجه ی اقدامات سازمان» پاسخ می‌دهد. بنابراین، تصميم گيري با در نظر گرفتن پيش بينی نتايج آينده بهبود يافته است. تجزیه‌وتحلیل توصیفی با استفاده از ابزارهای مدل سازی سطح بالا می‌تواند به طور قابل‌توجهی به عملکرد و کارآیی سازمان‌ها کمک کند، از طریق تصمیم‌گیری دقیق‌تر و سریعتر با هزینه پایین و ریسک و شناسایی راه حل های بهینه برای تخصیص منابع [18].

تجزیه‌وتحلیل پیش بینی پیشرفته و پیشگیرانه می‌تواند نقش مهمی در تصمیم‌گیری کارآمد استراتژیک داشته باشد و با مشکلات مهمی از سازمان‌ها مانند طراحی و توسعه محصولات / خدمات، شکل‌گیری زنجیره تامین و غیره بازی کند. [19].

**1.2.1 کاربردهای تجزیه‌وتحلیل داده بزرگ**

امروزه، به‌عنوان نسل در حال رشد از داده‌های موجود، یک روند شناخته‌شده در سراسر شرکت‌ها، کشورها و بخش های بازار است، اکثر شرکت‌ها بدون در نظر گرفتن صنعت، به جمع‌آوری، ذخیره و تجزیه‌وتحلیل داده‌ها به‌منظور جمع‌آوری ارزش می‌پردازند. اقتصاد دیجیتال از طریق استفاده عظیم از اینترنت و خدمات دیجیتالی، تقریباً تمام بخش های صنعت، ازجمله کشاورزی و تولید را به خدمات بیشتر متمرکز کرده است [20]. بخش های بسیاری و مختلف مانند تجارت الکترونیک وجود دارد،سیاست، علم و فناوری، سلامت، خدمات دولتی و غیره، که در آن تجزیه‌وتحلیل داده‌های بزرگ استفاده می‌شود. شرکت‌های داده محور از صنایع مختلف، قدرت داده‌های بزرگ را روشن می‌کنند و پیش بینی دقیق‌تر را در تصمیم‌گیری‌های بهتر انجام می دهند.

جریانهای بزرگ داده‌ای که هر روز تولید می‌شوند، نیاز به زیرساخت‌های بهتر دارند تا بتوان آن‌ها را جمع‌آوری، ذخیره و تجزیه‌وتحلیل کرد. یک بازار با عرضه وسیع محصولات و ابزارهای جدید طراحی شده برای پوشش دادن تمام نیازهای داده‌های بزرگ ایجاد شده و به سرعت در حال توسعه است [21]. طیف وسیعی از ابزارهای تحلیلی وجود دارد که می‌تواند برای انجام BDA مورداستفاده قرار گیرد، ازجمله بر اساس پرس و جو SQL، تجزیه‌وتحلیل آماری، داده کاوی، خوشه سریع، پردازش زبان طبیعی، تجزیه‌وتحلیل متن، تجسم داده‌ها و هوش مصنوعی (AI ) این تکنیک‌ها و ابزار ها به آسانی و به سرعت در حال استفاده از داده‌های بزرگ است.

دانش حاصل از بهره‌برداری داده‌های بزرگ، شرکت‌ها ارزش افزوده را از طریق روش های جدید بهره‌وری، رشد، نوآوری و مازاد مصرف کننده به دست می‌دهد [7]، بنابراین داده‌های بزرگ به تعیین‌کننده اصلی رقابت می‌پردازند و شرکت‌ها به ظرفیت تجزیه‌وتحلیل داده‌ها برای بهره‌برداری کامل پتانسیل داده‌ها نیاز دارند.

شرکتهایی که یادگیری داده‌های بزرگ را با استفاده از اطلاعات در زمان واقعی از منابع مختلف مانند حسگرها، دستگاه‌های متصل و غیره می‌گیرند، می‌توانند جزئیات بیشتری را درک کنند و محیط های جدید را تعریف کنند، محصولات جدید و نوآورانه ای را ایجاد کنند، سریعا در تغییرات و بهینه سازی اقدامات بازاریابی اهرم داده‌های بزرگ می‌تواند به تخصیص و نظارت بر منابع کارآمد، کاهش ضایعات، تسهیل بینش جدید و سطح شفافیت بیشتر در بخش های مختلف از تولید و فروش کمک کند.

بنابراین، کاربردهای BDA تقریباً در هر بخش تجاری وجود دارد. برنامه‌های کاربردی نیز در سیاست و دولت الکترونیک، علم و فناوری، امنیت و ایمنی، بهداشت هوشمند و رفاه وجود دارد [3]. علاوه بر این، بسیاری از انواع مختلف برنامه‌های کاربردی داده‌های بزرگ در میان شرکت‌ها و بخش های صنعت وجود دارد. BDA می‌تواند در برنامه‌های تجارت الکترونیک و بازاریابی مانند تبلیغات آنلاین و فروش متقابل استفاده شود، درحالی‌که به شرکت‌ها کمک می‌کند تا رفتار مشتری را در شکل دادن مشخصات مشتری 360 درجه برای پیاده سازی اقدامات هدفمند و بهینه سازی شده برای تأثیر بر مشتری و رضایت مشتری. این ارائه می‌دهد درک بهتر از رفتار مشتری و ترجیحات و در نتیجه بهبود خدمات به مشتریان.

بعضی از نمونه روش های BDA که مورداستفاده قرار می‌گیرند اهمیت تجزیه‌وتحلیل را در موضوعات مختلف نشان می دهند [22]:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| بازار یابی | تجزیه‌وتحلیل بازار سبد | سیستم‌های توصیه شده | اطلاعات مشتری | مدل سازی حفظ | پیش بینی مشتری |
| فرایندها | تجزیه‌وتحلیل زنجیره تامین | پیش بینی تقاضا و عرضه | تجزیه‌وتحلیل فرآیندهای کسب‌وکار | تجزیه‌وتحلیل منابع انسانی |  |
| دولت | تشخیص تقلب | تشخیص تروریسم | اجتناب از مالیات | کاهش هزینه | امنیت اجتماعی |
| مدیریت ریسک | مدلسازی ریسک اعتباری | مدل ریسک بازار | تشخیص تقلب |  |  |
| وب و رسانه‌های اجتماعی | تجزیه‌وتحلیل وب | تجزیه‌وتحلیل رسانه‌های اجتماعی | تست چند متغیره |  |  |

شرکت‌ها و سازمان‌ها مقدار زیادی از اطلاعات مربوط به امنیت را جمع‌آوری می‌کنند، مانند برنامه‌های کاربردی نرم افزار، رویدادهای شبکه، رویدادهای فعالیت‌های مردم. تولید اطلاعاتی که از این اقدامات انجام می‌شود، به سرعت در حال افزایش است، زیرا سازمان‌ها قادر به ورود به منابع بیشتر، برنامه‌های نرم‌افزاری بیشتر، کارکنان بیشتر و حرکت به راه حل های ابر هستند. متاسفانه حجم و انواع داده‌های امنیتی به سرعت در حال شدت یافتن هستند و تکنیک‌های تحلیلی موجود نمی‌توانند کارآمد و قابل اعتماد عمل کنند. برنامه‌های BDA بخشی از مدیریت و نظارت امنیتی می‌شوند، زیرا این امر به طور مؤثر برای تمیز کردن، آماده سازی و تجزیه‌وتحلیل مجموعه داده‌های مختلف پیچیده و ناهمگن مؤثر است [23]. یکی از رایج ترین استفاده های BDA، تشخیص تقلب است، بنابراین موسسات مالی، دولت‌ها و شرکت‌های تلفن از فناوری‌های اطلاعاتی بزرگ برای رفع خطر و افزایش کارایی آن‌ها استفاده می‌کنند.

علاوه بر این، BDA به طور گسترده ای در زنجیره تامین و عملیات تدارکات نقش مهمی در توسعه استراتژی های زنجیره تامین و مدیریت عملیات زنجیره تامین بازی می‌کنند. BDA می‌تواند تصمیم‌گیری را از طریق شناخت تغییرات شرایط بازاریابی، شناسایی خطرات زنجیره تامین و بهره‌برداری از توانایی های زنجیره تامین برای تدوین استراتژی های زنجیره ای عرضه حمایت کند، به طوری که انعطاف‌پذیری و سود دهی زنجیره تامین را بهبود بخشد. BDA همچنین به تصمیم‌گیری در سطح عملیاتی کمک می‌کند، با توجه به اینکه اندازه گیری و تجزیه‌وتحلیل عملکرد زنجیره تامین را با توجه به برنامه ریزی تقاضا، منابع، تولید، موجودی و تدارکات، مورد توجه قرار می‌دهد. بدین ترتیب، بهبود کارایی عملیات، اندازه گیری عملکرد زنجیره تامین، کاهش تغییرات فرآیند و کمک به اجرای بهترین راهبردهای زنجیره تامین در سطح عملیاتی را بهبود می بخشد [24].

صحبت کردن در مورد شرکت‌های دیجیتال و مبتنی بر داده‌ها، نخستین کسانی هستند که گوگل، آمازون، اپل و فیسبوک دارند. آمازون که دیجیتال متولد شد، داده‌های بزرگ را به دست آورد که باعث شکستن بازار کتاب سنتی شد و رهبر خرید دیجیتال شد. یک مثال دیگر از یک شرکت مشهور متولد شده از دیجیتال گوگل است که داده‌ها را از طریق موتور جستجو به بازاریابی دیجیتال جهت ارائه و جستجو در اختیار کاربران خود قرار می‌دهد، درحالی‌که Google و Facebook جمع‌آوری داده‌ها برای بازاریابی شخصی و سفارشی را فراهم می‌کنند.

بااین‌وجود، شرکت‌های سنتی غیر فناورانه نیز تلاش می‌کنند که مزایای داده شده را به دست آورند. شرکت ژنرال الکتریک (GE) یک پلتفرم مبتنی بر ابر برای نرم افزار صنعتی اینترنتی به نام "Predix" را توسعه داده است که در زمان واقعی برای مهندسان بینش فراهم می‌کند تا برنامه ریزی های تعمیر و نگهداری را انجام دهد، کارایی دستگاه را بهبود می بخشد و خرابی را کاهش می‌دهد. به این ترتیب GE پیشنهادات خدمات ارزشی جدید در بازار محافظه کار صنعت نفت و گاز را ارائه داد، درحالی‌که با موانع جدی خود مواجه است: بهبود دارایی ها و بهره‌وری عملیاتی و از بین بردن هزینه دانش ضمنی از نیروی کار پیرامون [25].

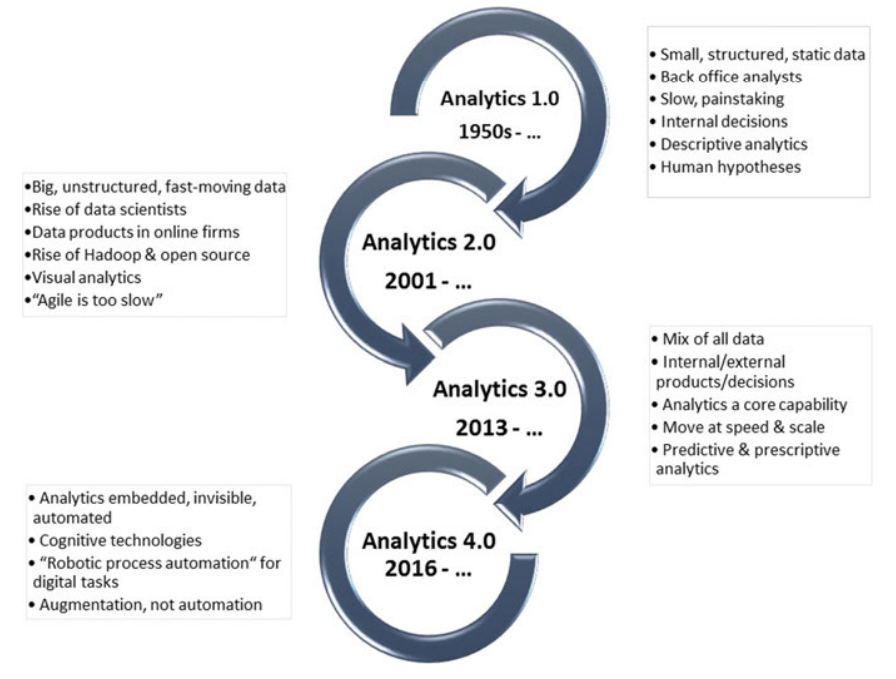
Walmart و سایر خرده فروشان عمده که از BDA در کل فرایند کسب‌وکار استفاده می‌کنند، از مدیریت زنجیره تامین تا بازاریابی، از اطلاعات سود می برند. استفاده از BDA در همه جا و نه‌تنها در بخش های دیجیتال، بلکه در بخش های غیر مبتنی بر وب شامل تولید، کشاورزی، مراقبت های بهداشتی، انرژی، مسافرت و دیگران نیز است. در بخش های مراقبت های بهداشتی، کاربرد های مختلف BDA وجود دارد، از کیفیت خدمات درمان و بهره‌وری هزینه بیمارستان تا بهبود و پیش بینی وضعیت سلامت بیمار. در سفر و خرده‌فروشی، برنامه‌های BDA قادر به ارائه اطلاعات مشتری از طریق تجزیه‌وتحلیل وب و شبکه‌های اجتماعی هستند، در نتیجه شرکت‌ها می‌توانند محصولات / خدمات شخصی را ارائه دهند. علاوه بر این، در مدیریت انرژی اکثریت شرکت‌ها از تجزیه‌وتحلیل داده‌ها برای ردیابی و کنترل دستگاه‌های دستیابی به مدیریت انرژی کارآمد بدون انحراف خدمات استفاده می‌کنند.

**1.2.2 بزرگ داده‌ها تجزیه‌وتحلیل چشم‌انداز**

تجزیه‌وتحلیل در روش تصمیم‌گیری عنوان جدیدی نیست، تجزیه‌وتحلیل کسب‌وکار در اوایل اواسط دهه 1950 - تجزیه‌وتحلیل 1.0 ساله ظاهر شد با ظهور ابزارهایی که قادر به تولید مقادیر زیادی داده در انبار داده‌های شرکت‌ها و کشف الگوهای سریع‌تر از ذهن انسان با ابزار هوش تجاری بودند. در آن دوران ابتدا، مدیران، تصمیم‌گیری مبتنی بر داده‌ها را فراتر از شهود در تصمیم‌گیری به دست آوردند. تا اواسط سال 2000، رشد سریع تولید داده‌ها و رسیدن داده‌های بزرگ، یک ارزیابی جدید 2.0 را نشان دادند که شرکت‌ها می‌توانند این داده‌ها را با استفاده از ابزارهای جدید قدرتمند تر استفاده کنند. نیاز به فن آوری های جدید نوآورانه ظاهر شد و شرکت‌ها به سرعت برای دستیابی به توانایی ها و دانش لازم برای به دست آوردن بینش از داده‌های بزرگ، با تفاوت عمده بین دوره ها در مهارت های موردنیاز برای تجزیه‌وتحلیل داده‌ها حرکت کردند[26]. در عصر بعدی، تجزیه‌وتحلیل یک بخش جدایی ناپذیر از شرکت‌های حمایت کننده تصمیم‌گیری است و شرکت‌ها به ایجاد محصولات / خدمات مبتنی بر تجزیه‌وتحلیل حرکت می‌کنند. حرکت بعدی، دوران بعدی Analytics 3.0 یا "اقتصاد اطلاعات"، با افزایش قابل‌توجهی از تولید داده‌ها از رشد اینترنت اشیاء (IoT) با دستگاه‌های 8.4 میلیارد دستگاه متصل در سال 2017 در سطح جهانی و 20.4 میلیارد در سال 2020 مشخص می‌شود. [27]

جدیدترین دوره Analytics 4.0 شامل تکنولوژی های شناختی ازجمله یادگیری ماشین است، جایی که اقدامات و تصمیم‌گیری به افزایش اتوماسیون با ماشین پویا منتقل می‌شود. ویژگی‌های اصلی تمام این دوره ها در شکل 4 [28] ظاهر می‌شود.

در عصر کنونی تجزیه‌وتحلیل، فناوری‌های جدید در حال ظهور تولید داده‌ها را افزایش خواهد داد، در نتیجه نهاد ها و سازمان‌ها باید با چالش‌های فنی مواجه شوند تا دسترسی به داده‌های بیشتر و بهتر داشته باشند. درآمد جهانی داده‌های بزرگ و تجزیه‌وتحلیل کسب‌وکار (BDA) در سال 2020 بیش از 203 میلیارد دلار خواهد بود و بانک های دولتی، تولید، دولت و خدمات حرفه ای به‌عنوان شرکت‌های بین‌المللی (IDC) [29] صنایع برتر در سرمایه گذاری های BDA خواهد بود.



شکل 4: تحولات تحلیلهای تاریخی

بنابراین، شرکت‌ها باید بر استفاده از تکنیک‌ها و ابزارهای تحلیلی جهت دریافت ارزش داده‌ها تمرکز کنند. BDA می‌تواند به شرکت‌ها کمک کند تا روند را بررسی و نکات جدید برای به دست آوردن مزیت رقابتی، ارائه محصولات جدید و پیشرفته را کشف کند. در میان دیگران، تجسم داده‌ها و شبیه سازی فرآیند، تجزیه‌وتحلیل متن و صدا، تجزیه‌وتحلیل رسانه‌های اجتماعی، تکنیک‌های پیش بینی کننده و پیشنهادی می‌توانند دانش ارزشمندی برای شرکت‌ها فراهم کنند، در حالیکه می‌توانند بینش شفاف تر و تأثیر بخش هر سازمانی را داشته باشند.

فن آوری داده‌ها و فناوری‌های اطلاعاتی بزرگ، تسهیل تصمیم‌گیری مبتنی بر داده‌ها را ارتقا می بخشد و در نتیجه به عملکرد سازمانی بهتر کمک می‌کند، زیرا هدف نهایی علوم داده بهبود تصمیم‌گیری است. بنابراین، سازمانهایی که نمیتوانند از تصمیم‌گیری مبتنی بر داده‌ها به‌عنوان استراتژی خود استفاده کنند، آن‌ها شکست خورده اند [4]. شواهدی وجود دارد مبنی بر اینکه تصمیم‌گیری مبتنی بر داده‌ها به طوری قابل توجه و مثبت به عملکرد سازمانی از لحاظ بهره‌وری و سودآوری کمک می‌کند [30]. رویکرد مبتنی بر داده‌ها می‌تواند فرصت‌های خوبی برای به دست آوردن مزیت رقابتی ارائه دهد؛ زیرا اندازه گیری و مدیریت دقیق‌تر تجزیه‌وتحلیل کسب‌وکار می‌تواند سازمان‌ها را به پیش بینی بهتر و تصمیم‌گیری دقیق‌تربرای هدف قرار دادن مداخلات مؤثر بیشتر قادر سازد.

نهادها و سازمانها در یک دوره کاملا جدید در تجزیه‌وتحلیل داده‌ها در حال بررسی استراتژی های نوآورانه و تکنیک‌های جدید برای حفظ رقابت در بازار خود هستند. استفاده از BDA به آن‌ها کمک می‌کند تا محصولات و خدمات جدید و یا بهبود یافته را مدیریت کنند، زنجیره های عرضه و فرآیندهای خود را به طور مؤثر مدیریت کنند، خطر را از طریق شناسایی تقلب و بهبود امنیت و بهره‌برداری از اطلاعات مشتری از بین ببرند.

کاربردهای BDA می‌تواند مزایای متعددی در سازمان‌ها و شرکت‌هایی داشته باشد که یک رویکرد مؤثر بر داده‌ها دارند. تجزیه‌وتحلیل داده‌های بزرگ قادر به ارائه دانش عمیق در مورد بخش های مختلف یک سازمان است و در نتیجه با استفاده از تجزیه‌وتحلیل داده‌های بزرگ برای پیش بینی ساخت کمک خواهد کرد به افزایش عملکرد و بازده بیشتر در سرمایه گذاری با هزینه پایین و خطر، درحالی‌که شفافیت بیشتری به دست آورد.

برخی از چشم‌اندازهای تجزیه‌وتحلیل داده‌های بزرگ عبارت‌اند از:

به دست آوردن بینش از تجزیه‌وتحلیل داده‌های بزرگ از تمام بخش های یک سازمان برای ایجاد یک استراتژی جامع کسب‌وکار، و یا کل سازمان. این استراتژی قادر خواهد بود به سطح بالاتری از بهره‌وری و کارایی در بخش ها، و همچنین در کل سازمان با کاهش هزینه و از بین بردن فرآیندها کمک کند.

سازمانها از تکنولوژی های هوش مصنوعی(AI) بیشتر بهره‌مند خواهند شد که می‌توانند سازمان‌ها را به روش های مختلف مجددا راه اندازی کنند. با این حال، سازمانها باید قبل از اینکه به تصویب AI پیشرفته حرکت کنند، اتوماسیون و تجزیه‌وتحلیل ساختاری را توسعه دهند. ادغام تجزیه‌وتحلیل داده‌های ساختاریافته و بدون ساختار با سیستم‌های AI امکان بررسی، توضیح و پیش بینی ترجیحات و رفتار مشتری را فراهم می‌کند [31].

نوآوری مبتنی بر داده‌ها (DDI) بر سرمایه مبتنی بر دانش، به نوآوری هایی ناشی از فرایندهای تصمیم‌گیری مبتنی بر داده‌ها می پردازد [2] که منجر به نادیده گرفتن مدل های کسب‌وکار جدید و مخرب، افزایش هوش مشتری [32] و معرفی محصولات / خدمات جدید / بهبود یافته می‌شود. پتانسیل داده‌های نوآورانه برای نوآوری در بخش خصوصی و دولتی در انگلستان به 24.1 میلیارد پوند کمک به اقتصاد انگلستان در سال‌های 2012-2017 منجر خواهد شد. [11]

تجزیه‌وتحلیل در زمان واقعی یک روند بزرگ است که شرکت‌ها باید در آینده نزدیک به آن توجه کنند. باوجود چالش‌ها و مسائلی که در نظر گرفته شده است، اثبات شده مدیریت تحلیلی مبتنی بر پیامدهای مهمی در شرکت‌ها وجود دارد، چه اینکه آن‌ها به دنبال رشد، کارایی یا تمایز رقابتی هستند. بنابراین، تجزیه‌وتحلیل داده‌های بزرگ پتانسیل به ظاهر نامحدودی برای کمک به یک شرکت برای رشد و به آشکار ساختن پتانسیل داده‌های خود نشان می‌دهد.

رشد سریع تقاضا برای تجزیه‌وتحلیل داده‌ها در کنار فقدان استعداد به همکاری و ابتکارات بین دانشگاه ها و صنعت به‌منظور پر کردن شکاف استعداد انجامید. در این زمینه، بسیاری از دانشگاه ها در حال آماده سازی و شروع دوره های آکادمیک مرتبط با علوم داده هستند. علاوه بر این، شرکت‌ها با بهره‌گیری از امکان بالقوه داده‌های بزرگ، آموزش کارکنان خود را ارائه می دهند. به تازگی، AirBnB دانشگاه خودش را به نام "Data University" برای دموکراتیزه کردن علوم داده و کمک به تصمیم‌گیری در مورد تصمیم‌گیری‌های اطلاعاتی آغاز کرد.

انتظارات متفاوتی از شرکت‌ها در خصوص تجزیه‌وتحلیل داده‌های بزرگ وجود دارد. رهبران سازمان می خواهند از تحلیلی استفاده کنند تا هوشمندانه تر و نوآورانه تر از قبل عمل کنند، درحالی‌که مدیران ارشد می خواهند از تصمیم‌گیری مبتنی بر داده‌ها برای عملکرد کارآمد خود استفاده کنند [33]. مدیران با استفاده از یک سیستم تصمیم‌گیری مبتنی بر داده‌ها (DSS)، به داده‌های تاریخی و جدیدی را که از آن‌ها پشتیبانی می‌کنند، برای دستیابی به اطلاعات برای فرایندهای سازماندهی و عملکرد منابع، دسترسی دارند. DDS نه‌تنها برای سازمان‌های جهانی بلکه برای سازمان‌های کوچک و متوسط ​​که می‌توانند به نفع آن‌ها بهره‌برداری شوند نیز مهم است [10].

**1.2.3 چالش‌ها و موانع تجزیه‌وتحلیل داده‌های بزرگ**

چالش‌های عمده در اتخاذ تحلیل داده‌های بزرگ از شرکت‌ها نسبت به داده‌ها و تکنولوژی بیشتر مدیریتی و فرهنگی هستند، درحالی‌که موانع اصلی عدم درک نحوه استفاده از تجزیه‌وتحلیل داده‌های بزرگ برای افزایش کسب‌وکار و عدم طیف مدیریتی از اولویت های رقابتی هستند [33]. مطالعات در میان بخش های مختلف صنعت نشان می‌دهد که سازمان‌ها از کمتر از نیمی از داده‌های ساختاریافته خود در فرایند تصمیم‌گیری استفاده می‌کنند، درحالی‌که کمتر از 1٪ از داده‌های غیر ساختاری آن‌ها مورد تجزیه‌وتحلیل قرار گرفته و یا مورد بهره‌برداری قرار می‌گیرند، 70٪ کارکنان دسترسی به داده‌هایی دارند که نباید و 80٪ وقت تحلیلگران صرف این می‌شود که اطلاعات را کشف و تهیه کنیم [34].

**رهبری**. با توجه به چالش‌های مدیریتی، شرکت‌هایی که در دوران داده‌ها موفق به موفقیت می‌شوند، تیم های رهبری هستند که اهداف را تعیین می‌کنند، دستاوردها را مدوله می‌کنند و از طریق بینش اطلاعات می‌توانند به سوالات حقوقی پاسخ دهند. علیرغم رویکرد تکنولوژیکی، قدرت داده‌های بزرگ نمی‌تواند بدون بینش و توانایی انسان مورد سوء استفاده قرار گیرد. بنابراین، رهبران شرکت‌هایی که دارای دید و توانایی آشکار شدن روند و فرصت‌های آینده هستند، توانایی کار نوآورانه دارند، تیم های خود را در دستیابی به اهداف خود برمی انگیزانند.

**مدیریت استعداد**. شرکت‌ها به‌منظور استفاده از داده‌ها از طریق تحلیل داده‌های بزرگ نیاز به سرمایه انسانی با سطح بالایی از مهارت های فنی برای استفاده و بهره‌برداری از این سیستم ها برای دستیابی به دانش بهره‌برداری برای کاربران نهایی، به طور عمده C-suite دارند. مهارت های ویژه افراد شامل آمار، داده کاوی بزرگ، ابزار تجسم استاد، ذهنیت کسب‌وکار و یادگیری ماشین است. این ها برای دستیابی به اطلاعات مهم در تصمیم‌گیری ضروری است [13]. با این حال، یافتن این افراد (دانشمندان داده، تحلیلگران داده‌ها و غیره) بسیار دشوار و بنابراین تقاضا برای آن‌ها زیاد است. چالش در پیدا کردن دانشمندان داده بامهارت در هر دو زمینه تحلیل و در حوزه دانش وجود دارد. به طور کلی، نسبت به نیاز دانشمندان اطلاعات کمتری وجود دارند [35].

**روش تصمیم‌گیری** در موسسات کارآمد، تصمیم گیرندگان و دانش حاصل از بهره‌برداری داده‌ها در یک مکان قرار دارند. بااین‌وجود، برای تصمیم گیرندگان رسیدگی به مقادیر زیاد داده‌ها دشوار است. بنابراین، نیاز به تصمیم گیرندگان دارای مهارت حل مسئله و توانایی ارائه پاسخ به مشکلات با داده‌های مناسب یا همکاری افراد مختلف در حل مسئله از طریق استفاده از داده‌های بزرگ وجود دارد [13].

**کیفیت تصمیم‌گیری**. کیفیت تصمیم‌گیری با اتخاذ یک رویکرد مبتنی بر داده‌ها، یک عامل مهم برای استفاده از امکاناتی است که تجزیه‌وتحلیل داده‌های بزرگ ارائه می‌دهد. در این زمینه، اطمینان از کیفیت تصمیم‌گیری با عواملی مانند کیفیت داده‌ها از منابع داده بزرگ، توانایی تجزیه‌وتحلیل داده‌ها، کارکنان و کیفیت تصمیم گیرنده ارتباط دارد [36]. دقت منابع داده بزرگ در ارائه ارزش بالا در تصمیم‌گیری حذف اقدامات اشتباه مشهود است، درحالی‌که قابلیت های تجزیه‌وتحلیل داده‌های بزرگ مرتبط با استفاده از تکنیک‌های مناسب و ابزار متخصصان با دانش تجزیه‌وتحلیل داده‌های بزرگ است.

**فرهنگ مبتنی بر داده‌ها** یکی دیگر از چالش‌های مهم برای پذیرش رویکرد مبتنی بر داده‌ها، فرهنگ سازمانی است. مبنای کسب فرهنگ مبتنی بر داده‌ها، توانایی سرعت بخشیدن، تجزیه‌وتحلیل و توزیع اطلاعات مهم کسب‌وکار به تصمیم گیرندگان است. این مبنای مهم برای افزایش عملکرد کسب‌وکار بسیار مهم است، درحالی‌که توسعه و بهبود این توانایی ها، شرکت‌ها را به پیشرفت در تمام بخش های کسب‌وکار قادر می سازد و به بازده بیشتر درسرمایه گذاری منجر می‌شود. در این زمینه، شرکت‌ها باید تصمیم‌گیری مبتنی بر داده‌ها را در تمام مسائل اتخاذ کنند و بر اساس به نفوذ و غریزه عمل نکنند. بنابراین، مدیریت باید به طور کامل اهمیت به دست آوردن بینش از استثمار داده‌ها درک کند. علاوه بر این، برای یک شرکت مبتنی بر داده، افرادی که در فرایند تصمیم‌گیری مبتنی بر داده دخیل هستند، نیاز به تعدیل دارند. مدیران باید بر مدیریت تیم و پروژه های کارآزمایی داده‌های تجزیه‌وتحلیل کارآمد باشند، درحالی‌که بازاریابان باید بتوانند معیارها و تجزیه‌وتحلیل ها را درک کنند تا مدیریت فعالیت‌های بازاریابی کارآمد را مدیریت کنند.

**استفاده از تکنولوژی جدید** بسیاری از شرکت‌ها قدرت داده‌ها را درک می‌کنند، مهارت های فناوری را در هوش تجاری و / یا انبارداری داده توسعه داده اند، اما فناوری‌های تجزیه‌وتحلیل داده‌های بزرگ متفاوت و جدید هستند. بنابراین، شرکت‌ها باید تکنیک‌ها و فن آوری هایی را که برای به دست آوردن ارزش از داده‌های بزرگ در دسترس هستند را به کار گیرند. از آنجا که این فن آوری ها به سرعت در حال تحول هستند، بخش های فناوری اطلاعات باید بتوانند ظرفیت خود را توسعه دهند و به نوآوری های مداوم متصل شوند. به‌عنوان‌مثال، وقتی نرم افزار پایگاه داده از گزینه های تجزیه‌وتحلیل داده بزرگ پشتیبانی نمی‌کند، مشکلات ظاهر می‌شوند.

**حریم خصوصی داده‌ها** جمع‌آوری داده‌ها به وسیله بسیاری از افراد به شدت مشکوک است. برای آن‌ها داده‌های بزرگ یک حمله به حریم خصوصی آن‌ها است. بازاریابان با درک مصرف کنندگان از داده‌ها درگیرند؛ 71٪ از آن‌ها بر این باورند که مارک های دارای دسترسی به داده‌های شخصی بی پروا از آن استفاده می‌کنند، درحالی‌که 58٪ از آن‌ها به دلیل نگرانی های امنیتی که منجر می‌شود به تصمیم‌گیری در مورد برنامه‌های کاربردی که دانلود می‌کنند، آدرس های ایمیل که به اشتراک می گذارند و سایت های رسانه‌های اجتماعی برای اتصال به وب سایت های دیگر استفاده می‌کنند، از هیچ سرویس دیجیتالی لستفاده نمی‌کنند[37]. بنابراین، شرکت‌ها باید از ضمانت نامه ها استفاده کنند تا اطمینان حاصل شود که داده‌ها برای نقض حریم خصوصی شخصی مشتری مورداستفاده قرار نمی‌گیرند [7]. در این راستا، سیاست‌های داده‌ها شامل حفظ حریم خصوصی، امنیت، مالکیت معنوی و مسؤولیت مسئولیت باید به‌منظور بهره‌برداری از داده‌های بزرگ داده تعیین شود.

**2. نتیجه گیری**

رشد اینترنت با آغاز عصر وب 2.0 باعث شد شرکت‌ها را قادر ساخت به مقادیر زیادی اطلاعات آسان‌تر و ارزان‌تر دسترسی یابند، درحالی‌که امکان جمع‌آوری داده‌های خارجی حتی با ظهور وب 3.0 نیز افزایش یافته است. شرکت‌ها و سازمان‌ها از هر بخش شروع به تمرکز بر استفاده از داده‌ها برای به دست آوردن مزیت رقابتی کردند.

امروزه دوران داده‌ها ی بزرگ بی سر و صدا تقریباً در هر شرکتی ایجاد شده است، زیرا آن‌ها متوجه شدند تصمیمات مبتنی بر داده‌ها تصمیمات بهتر و دقیق‌تری هستند. با این حال، بسیاری از شرکت‌ها در صنایع مختلف از تجزیه‌وتحلیل‌های عملیاتی ازجمله تجزیه‌وتحلیل داده‌های بزرگ استفاده می‌کنند، این بدان معنا نیست که همه آن‌ها از طریق درک ارزشمندی و ارزش واقعی کسب‌وکار از داده‌های موجود بهره‌مند شوند.

تبدیل شدن به یک شرکت مبتنی بر داده‌ها بیش از استفاده از تکنیک‌های تحلیلی و ابزار است. شرکت‌ها باید افرادی را که مجهز به تفکر سیستماتیک هستند، برای ترویج موفقیت در تصمیم‌گیری مبتنی بر داده‌ها استخدام کنند. موفقیت در محدوده کسب‌وکار اطلاعات محرمانه امروزه شامل توانایی تحلیل داده‌ها از طریق داده‌ها می‌شود. از آنجا که مقدار داده‌ها به طور مداوم در حال رشد است، دانش و تحلیل دامنه را نمی‌توان به‌عنوان مناطق جداگانه در نظر گرفت. انتظار می‌رود که متخصصان علمی و حرفه ای شرکت‌ها مهارت های تحلیلی و فرآیندهای کسب‌وکار را درک کنند.

کارکنانی که درک پایه ای از تفکر تحلیلی داده ندارند، واقعا نمی دانند چگونه بر کسب‌وکار یک سازمان کار می‌کند. اگر آن‌ها قادر به درک روند و مراحل آن باشند، برای آن‌ها آسان‌تر خواهد بود تا راه حل های مناسب برای نقاط ضعف مرحله مربوطه را پیدا کنند. اما سازمان‌ها می‌توانند با اجرای داده - محوری، با برخی از چالش‌ها، مدیریتی و فنی مواجه شوند.

داده‌های بزرگ فقط در مورد حجم داده‌ها نیست، بلکه در مورد انواع و سرعت نیز هست. تجزیه‌وتحلیل داده‌های بزرگ توانایی کمک به شرکت‌ها در درک شرایط محیط کسب‌وکار، رفتار مشتریان و نیازهای آن‌ها و فعالیت‌های رقباشان است. با توجه به تجزیه‌وتحلیل داده‌های بزرگ، شرکت‌ها قادرند محصولات و اقدامات خود را به‌منظور برآورده ساختن نیازهای مشتریان و نوآوری در برابر رقبا از طریق پیش بینی های بهتر و تصمیم‌گیری‌های هوشمندانه بر اساس شواهد به جای شهود بسازند. سازمان‌هایی که برای مدیریت چالش‌ها و اتخاذ یک فرهنگ مبتنی بر داده‌ها دست به دست می‌شوند، می‌توانند چشم‌انداز خوبی داشته باشند. شواهد قوی وجود دارد که عملکرد کسب‌وکار را می‌توان با تصمیم‌گیری مبتنی بر داده‌ها، ابزارهای تحلیلی فناوری اطلاعات و تکنیک‌های بزرگ بر روی داده‌های بزرگ بهبود داد. همانطور که شرکت‌های بیشتری یاد می‌گیرند مهارت های ضروری برای استفاده از داده‌های بزرگ و نحوه تعامل با تکنولوژی های فعلی که به طور مداوم در حال توسعه هستند، ممکن است به زودی از رقبای خود جدا شوند و مزایای رقابتی قابل‌توجهی داشته باشند.

**References**

1. United Nations: A world that counts. Mobilizing the data revolution for sustainable development. United Nations, New York (2014)

2. OECD: Data-driven innovation big data for growth and well-being: big data for growth and well-being. OECD Publishing (2015)

3. Chen, H., Chiang, R., Storey, V.C.: Business intelligence and analytics: from big data to big impact. Miss. Q. 36(4), 1165–1188 (2012)

4. Provost, F., Fawcett, T.: Data science and its relationship to big data and data-driven decision making. Big Data 1(1), 51–59 (2013)

5. Economist, T.: Data is giving rise to a new economy. In: The Economist, 05 Jun 2017. https:// [www.economist.com/news/briefing/21721634-how-it-shaping-up-data-giving-rise-neweconomy. Accessed 06 Oct 2017](http://www.economist.com/news/briefing/21721634-how-it-shaping-up-data-giving-rise-neweconomy.%20Accessed%2006%20Oct%202017)

6. Sivarajah, U., Kamal, M.M., Irani, Z., Weerakkody, V.: Critical analysis of big data challenges and analytical methods. J. Bus. Res. 70, 263–286 (2017)

7. Manyika, J., et al.: Big data: the next frontier for innovation, competition, and productivity (2011)

8. Gantz, J., Reinsel, D.: Extracting Value from Chaos, IDC (2011)

9. Friendly, M.: The golden age of statistical graphics. Stat. Sci. 23(4), 502–535 (2008)

10. Power, D.J.: Understanding data-driven decision support systems. Inf. Syst. Manag. 25(2), 149–154 (2008)

11. Cebr: Data equity: unlocking the value of big data Report for SAS, April (2012). https://www. cebr.com/wp-content/uploads/2013/03/1733\_Cebr\_Value-of-Data-Equity\_report.pdf. Accessed 06 Nov 2017

12. Website. [https://www.news.microsoft.com/europe/2016/04/20/go-bigger-with-big-data/sm. 0008u654e19yueh0qs514ckroeww1/XmqRHQB1Gcmde4yb.97. Accessed 15 Jun 2017](https://www.news.microsoft.com/europe/2016/04/20/go-bigger-with-big-data/sm.%200008u654e19yueh0qs514ckroeww1/XmqRHQB1Gcmde4yb.97.%20Accessed%2015%20Jun%202017)

13. McAfee, A., Brynjolfsson, E.: Big data: the management revolution. Harv. Bus. Rev. 90(10) 60–66, 68, 128 (2012)

14. Burstein, F., Holsapple, C.: Handbook on Decision Support Systems 1: Basic Themes. Springer Science & Business Media (2008)

15. Larson, D., Chang, V.: A review and future direction of agile, business intelligence, analytics and data science—Science Direct. Int. J. Inf. Manage. 36(5), 700–710 (2016)

16. Davenport, T.: Big Data at Work: Dispelling the Myths. Harvard Business Review Press, Uncovering the Opportunities (2014)

17. Gandomi, A., Haider, M.: Beyond the hype: big data concepts, methods, and analytics. Int. J. Inf. Manage. 35(2), 137–144 (2015)

18. How to leverage the power of prescriptive analytics to maximize the ROI. In: IBM Big Data and Analytics Hub. [http://www.ibmbigdatahub.com/blog/how-leverage-power-prescriptiveanalytics-maximize-roi. Accessed 16 Jun 2017](http://www.ibmbigdatahub.com/blog/how-leverage-power-prescriptiveanalytics-maximize-roi.%20Accessed%2016%20Jun%202017)

19. Demirkan, H., Delen, D.: Leveraging the capabilities of service-oriented decision support systems: putting analytics and big data in cloud. Decis. Support Syst. 55(1), 412–421 (2013)

20. Lodefalk, M.: Servicification of manufacturing—evidence from Sweden. Int. J. Econom. Bus. Res. 6(1), 87 (2013)

21. Davenport, T.H., Barth, P., Bean, R.: How ‘big data’ is different. MIT Sloan Manag. Rev 54(1), 22–24 (2012)

22. Baesens, B.: Analytics in a Big Data World: The Essential Guide to Data Science and its Applications. Wiley (2014)

23. Big Data Analytics for Security—IEEE Xplore Document. http://ieeexplore.ieee.org/abstract/ document/6682971/?reload=true. Accessed 18 Jun 2017

24. Wang, G., Gunasekaran, A., Ngai, E.W.T., Papadopoulos, T.: Big data analytics in logistics and supply chain management: certain investigations for research and applications—science direct. http://www.sciencedirect.com/science/article/pii/S0925527316300056?via%3Dihub. Accessed 18 Jun 2017

25. GE’s big bet on data and analytics|MIT sloan management review. In: MIT Sloan Management Review. http://sloanreview.mit.edu/case-study/ge-big-bet-on-data-and-analytics/. Accessed 14 Jun 2017

26. Analytics 3.0: Harvard Business Review, 01 Dec 2013. [https://hbr.org/2013/12/analytics-30. Accessed 21 Jun 2017](https://hbr.org/2013/12/analytics-30.%20Accessed%2021%20Jun%202017)

27. Gartner Says 8.4 Billion Connected. [http://www.gartner.com/newsroom/id/3598917. Accessed 21 Jun 2017](http://www.gartner.com/newsroom/id/3598917.%20Accessed%2021%20Jun%202017)

28. Davenport, T.: Analytics and IT new opportunity for CIOs. In: Harvard Business Review (2016)

29. Double-digit growth forecast for the worldwide big data and business analytics market through 2020 led by banking and manufacturing investments, according to IDC. http://www. idc.com, http://www.idc.com/getdoc.jsp?containerId=prUS41826116. Accessed 21 Jun 2017

30. Brynjolfsson, E., Hitt, L.M., Kim, H.H.: Strength in numbers: how does data-driven decision making affect firm performance?. SSRN Electron. J.

31. If your company isn’t good at analytics, it’s not ready for AI. In: Harvard Business Review, 07 Jun 2017. [https://www.hbr.org/2017/06/if-your-company-isnt-good-at-analytics-its-notready-for-ai. Accessed 22 Jun 2017](https://www.hbr.org/2017/06/if-your-company-isnt-good-at-analytics-its-notready-for-ai.%20Accessed%2022%20Jun%202017)

32. Ryan, L.: The Visual Imperative: Creating a Visual Culture of Data Discovery. Morgan Kaufmann (2016)

33. Lavalle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N.: Big data, analytics and the path from insights to value. MIT Sloan Manag. Rev. 52(2), 3–22 (2010)

34. The 2 types of data strategies every company needs. In: Harvard Business Review, 01 May 2017. [https://hbr.org/2017/05/whats-your-data-strategy. Accessed 18 Jun 2017](https://hbr.org/2017/05/whats-your-data-strategy.%20Accessed%2018%20Jun%202017)

35. Waller, M.A., Fawcett, S.E.: Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. J. Bus. Logist. 34(2), 77–84 (2013)

36. Janssen, M., van der Voort, H., Wahyudi, A.: Factors influencing big data decision-making quality. J. Bus. Res. 70, 338–345 (2017)

37. Chahal, M., et al.: Marketers overestimate consumers’ attitude to data—Marketing Week. In: Marketing Week, 23 Jun 2016. https://www.marketingweek.com/2016/06/23/marketersoverestimate-consumers-attitude-to-data/. Accessed 18 Jun 2017