

فناوری داده‌های عظیم و الزامات قانونی آن

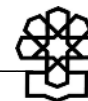
معاونت پژوهش‌های زیربنایی و امور تولیدی
دفتر: مطالعات ارتباطات و فناوری‌های نوین

کد موضوعی: ۲۸۰
شماره مسلسل: ۱۴۲۷۳
خردادماه ۱۳۹۴

به نام خدا

فهرست مطالب

۱.....	چکیده
۲.....	مقدمه
۲.....	۱. داده‌های عظیم - معنای عبارت
۵.....	۲. کاربرد داده‌های عظیم در بخش دولت
۱۱.....	۳. داده‌های عظیم، الزامات حقوقی و سیاستگذاری
۱۴.....	۴. مسائل قانونی پیش روی کشورهای مختلف و سیاست‌های دولت آنها
۱۸.....	جمع‌بندی
۱۹.....	پی‌نوشت‌ها



فناوری داده‌های عظیم و الزامات قانونی آن

چکیده

در سال‌های اخیر صنعت فناوری اطلاعات، با ظهور حجم گسترده‌ای از داده‌ها و اطلاعات مواجه بوده که این داده‌ها عمدتاً از منابع متفاوتی از جمله ابزارهای علمی، ماهواره‌ها، رسانه‌های دیجیتالی، ایمیل‌ها، وبلاگ‌ها، تراکنش‌های مالی، شبکه‌های اجتماعی و دیگر انواع شبکه؛ شبیه‌سازی‌های علمی، مدل‌ها و نظرسنجی‌ها؛ تجزیه و تحلیل‌های محاسباتی از داده‌های مشاهده‌ای و غیره پدید می‌آیند که موارد مذکور شامل متن، تصویر، صوت و غیره می‌شوند. نیاز به مدیریت و تجزیه و تحلیل این حجم گسترده از داده‌ها باعث پیدایش مفهوم جدیدی به نام «داده‌های عظیم»^۱ شده است. در واقع داده‌های عظیم اشاره به مجموعه داده‌های بزرگ و پیچیده‌ای دارد که با ابزارهای مدیریتی و پایگاه داده‌های سنتی قادر به مدیریت نیستند و برای جمع‌آوری، ذخیره‌سازی، به اشتراک‌گذاری، پردازش و مدیریت، نیازمند فناوری‌های جدیدند. داده‌های عظیم برای یادگیری سیستم‌های اطلاعاتی، داده‌کاوی، کشف الگوهای پنهان، همبستگی‌های ناشناخته و افزایش ضریب هوشمندی در سازمان‌های تجاری، بنگاه‌های اقتصادی و تحلیل‌های مالی، تحقیقات ژنتیکی، زیست‌شناسی، پیش‌بینی تغییرات آب و هوایی و غیره به کار گرفته می‌شوند. همچنین تجزیه و تحلیل داده‌های عظیم می‌تواند بهره‌وری و اثربخشی را در سراسر طیف گسترده‌ای از مسئولیت‌های دولت، از طریق بهبود فرآیندهای دولتی افزایش دهد.^(۱) دولت‌ها قادر خواهند بود با مدیریت و به‌کارگیری ابزارهای تجزیه و تحلیل این داده‌ها در جهت پیشگیری از جرم و جنایت، بهبود ترافیک و حمل‌ونقل، افزایش دفاع و امنیت ملی، مدیریت درآمد، ملاحظات زیست‌محیطی و خدمات اجتماعی، اقدامات مؤثری برای رفاه شهروندان انجام دهند. بدین‌منظور توسعه بهره‌گیری از داده‌های عظیم از سوی دولت‌ها علاوه بر تأمین زیرساخت‌های قدرتمند تجزیه و تحلیل داده‌ها، مستلزم تأمین الزامات حقوقی و قانونی مرتبط نیز است. این زیرساخت‌ها شامل حفظ حریم خصوصی یا حفاظت از داده‌های شخصی، رعایت مالکیت فکری داده‌ها، مقررات و الزامات قراردادی است. این گزارش با مروری بر مفاهیم و کاربردهای داده‌های عظیم در سطح ملی به تبیین الزامات حقوقی و قانونی بهره‌گیری از داده‌های عظیم به‌ویژه حفاظت از داده‌های شخصی می‌پردازد.

مقدمه

امروزه سازمان‌ها و شرکت‌ها با انواع متنوعی از داده‌ها سروکار دارند، از پژوهش‌های علمی و هوشمند گرفته تا داده‌های سرشماری و امور مالی که همه سه ویژگی حجم بزرگ، تنوع زیاد و سرعت بالا را دارند. این سازمان‌ها ملزم به مدیریت این داده‌ها با توجه به قوانین و مقررات موجودند. با به‌وجود آمدن شدت رقابت برای منابع، چه در بخش دولتی و چه در بخش خصوصی، سازمان‌ها به دنبال راهی برای متمایز شدن از رقبای خود به وسیله دستیابی به ثروت نهفته شده در اطلاعات در جهت بهبود رقابت، بهره‌وری و سودآوری بیشترند. داده‌های گردآوری شده در بخش دولتی با استفاده از فناوری داده‌های عظیم و ابزارهای تجزیه و تحلیل می‌توانند برای ایجاد ارزش افزوده به کار گرفته شوند.

۱. داده‌های عظیم - معنای عبارت

اصطلاح داده‌های عظیم برای توصیف حجم انبوهی از داده‌ها که توسط سازمان‌ها، شرکت‌های عظیم یا برخی پروژه‌های علمی تحقیقاتی کشورها برای ایجاد ارزش افزوده از مجموعه گسترده‌ای از اطلاعات با فرمت دیجیتال، به صورت ساختاریافته یا غیرساختاریافته مورد تجزیه و تحلیل قرار می‌گیرند، به کار می‌رود.^(۲)

داده‌های عظیم به عنوان یک حوزه جدید مطرح شده و در کنار فناوری‌های رایانش ابری، رایانش اجتماعی و رایانش موبایل، یکی از چهار محرک اصلی دگرگونی صنعت فناوری اطلاعات است که شامل ویژگی‌هایی از جمله یادگیری ماشین، چاپ سه‌بعدی، واقعیت مجازی، اینترنت و فناوری نانو است.^(۳)

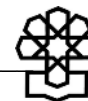
تعاریف گوناگونی از داده‌های عظیم در سراسر جهان توسط صاحب‌نظران ارائه شده و جامع‌ترین تعریف را می‌توان از گروه گارتنر^۱ ارائه داد که به صورت زیر است:

«دارایی‌های اطلاعاتی با حجم بالا، شتاب بالا و تنوع بالا که نیازمند اشکال جدید پردازشی است و تصمیم‌گیری پیشرفته، کشف دانش و بهینه‌سازی فرآیند را امکان‌پذیر می‌سازد». داده‌های عظیم با به‌کارگیری فناوری‌های پردازشی جدید، به چالش‌هایی از جمله پردازش چابک مجموعه‌های بسیار بزرگ داده‌ها یا مدیریت مجازی انواع گوناگون داده‌ها پاسخ می‌دهد.^{(۴) و (۵)}

دفتر اجرایی رئیس‌جمهور آمریکا^۲ نیز در سال ۲۰۱۴ موضوع داده‌های عظیم را به شرح زیر توصیف کرده است:

1. Gartner, Inc

2. Executive Office of the President of the United States



«اکثر تعاریف از داده‌های عظیم، منعکس‌کننده توانایی رو به رشد این فناوری در تصرف، جمعیت و پردازش هرچه بیشتر حجم، سرعت و تنوع داده‌هاست. به عبارت دیگر، داده‌ها در حال حاضر خیلی سریع‌تر با پوشش بیشتر و دامنه گسترده‌تر در دسترس بوده و شامل انواع جدیدی از فرمت‌ها می‌شوند که در گذشته قابل دستیابی نبودند. به بیان دقیق‌تر، مجموعه داده‌های عظیم، حجیم، متنوع، پیچیده و دارای طول جغرافیایی بالا می‌باشند که از منابع متعددی مانند حسگرها، تراکنش‌های اینترنتی، ایمیل، ویدئو، کلیک استریم یا دیگر منابع دیجیتال تولید شده‌اند»^(۶).

بنابر تعاریف فوق، آنچه داده‌های عظیم را از داده‌های سنتی متمایز می‌کند، ویژگی‌های زیرند:

۱- حجم وسیعی از داده‌های دیجیتال: براساس برخی گزارش‌ها در جهان هر دو روز به اندازه ۵ اگزابایت^۲ داده تولید می‌شود که در مجموع یعنی به اندازه اطلاعات ایجاد شده از آغاز تمدن تا سال ۲۰۰۳. بدیهی است که رشد این حجم وسیع از داده‌ها به سرعت اتفاق می‌افتد. برای مثال، بیش از ۷۰۰,۰۰۰ به‌روزرسانی در فیس‌بوک و ۲ میلیون ایمیل در هر دقیقه در سراسر جهان فرستاده می‌شوند.^(۷) در سال ۲۰۱۲ مقدار داده‌های ذخیره شده در یک سازمان ایالت متحده ۱/۶۱ پتابایت بوده است که در پایان سال ۲۰۱۴ این مقدار به ۲/۶۳ پتابایت افزایش یافته و نشان‌دهنده رشد ۶۳ درصدی است.^(۸) همچنین در طول ۵۰ سال گذشته، روزنامه نیویورک تایمز در کل، ۳ میلیارد تولید کلمات داشته، در حالی که کاربران توییت هر روز ۸ میلیارد کلمه تولید می‌کنند که همه این موارد بیان‌کننده رشد حجم داده‌ها در دنیای دیجیتال اند.^(۹)

۲- تنوع داده‌ها: داده‌های عظیم طیفی از فرمت‌های متنوع مانند متن، عکس، ویدئو، صدا و غیره را دربر می‌گیرند. از نظر ساختار، داده‌های تولید شده می‌توانند ساختاریافته، نیمه‌ساختاریافته، بدون ساختار یا چند ساختاری باشند.^(۱۰)

۳- سرعت بالا: داده‌ها با سرعت بالایی و به‌صورت بلادرنگ در حال تولیدند و همین امر باعث شده تا مکانیسمی برای واکنش و پاسخگویی به آنها مورد نیاز باشد.^(۱۱)

۴- صحت و اطمینان داده‌ای: این ویژگی اشاره به سطح قابلیت اطمینان به انواع خاصی از داده‌هاست. تلاش برای دستیابی به کیفیت بالای داده‌ها از مهمترین نیاز و چالش‌های داده‌های عظیم است، اما حتی بهترین روش‌های پاک‌سازی داده هم نمی‌تواند قابلیت ذاتی غیرقابل پیش‌بینی بودن برخی از داده‌ها، مانند آب‌وهوا، اقتصاد، تصمیم‌گیری خرید مشتری در آینده را حذف کنند. نیاز به

1. Volume

۲. یک واحد از اطلاعات که برابر ۱۰۲۴ پتابایت است. کوتاه شده آن (EB) است. این واژه تشکیل شده از پیشوند اگزا و کلمه بایت است. پیشوند اگزا به معنای ۱۰ به توان ۱۸ بایت است، اما در محاسبات که بایت برحسب توانی از دو محاسبه می‌شود، یک اگزابایت معادل ۲ به توان ۶۰ بایت است.

3. Variety

4. Velocity

5. Veracity

اذعان و برنامه‌ریزی برای عدم قطعیت، یکی از ابعاد دیگر داده‌های عظیم است که به مدیران برای درک بهتر پیچیدگی‌های محیطی معرفی می‌شود. این چهار ویژگی در شکل زیر نشان داده شده‌اند.^(۱۱)

شکل ۱. ویژگی‌های داده‌های عظیم

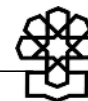


Source: IBM.

استفاده از نرم‌افزارهای تجزیه و تحلیل کمی، قابلیت تغییر از بینش گذشته‌نگر به آینده‌نگر، تجزیه و تحلیل به صورت بلادرنگ به جای پردازش دسته‌ای، بررسی اعتبار و نوسان و نمایش داده‌ها از دیگر ویژگی‌های داده‌های عظیم است. ایجاد تسهیلات کوچک اما پایدار، ایجاد تغییرات کسب‌وکاری به صورت سریع و تدریجی، افزایش توان رقابتی، نوآوری و ارزش داده مورد استفاده از ویژگی‌های ارزشی داده‌های عظیم به‌شمار می‌رود.

این فناوری شامل ذخیره‌سازی کم‌هزینه، حافظه پردازش بالا، پردازش مبتنی بر رایانش ابری و نرم‌افزارهای توسعه است. در این میان هیچ استاندارد جامعی برای تکنولوژی داده عظیم وجود ندارد. دلیل اصلی این است که شرکت‌های دارای فناوری تجزیه و تحلیل داده‌های عظیم معمولاً طبیعت پیچیده و متنوعی دارند. معروف‌ترین نرم‌افزار مدیریت داده‌های عظیم برای به‌کارگیری در توزیع، مدیریت و جستجوی داده در سراسر سرورهای چندگانه، تکنولوژی هادوپ^۱ است که توسط شرکت

۱. هادوپ (Hadoop) یک چارچوب برنامه‌نویسی رایگان و مبتنی بر جاواست. هادوپ، مسائلی را هدف گرفته که یافتن پاسخ آنها، به تحلیل کل داده‌های موجود در سیستم نیاز دارد. در این چارچوب عملیات روی داده‌های بسیار بزرگ و حجیم با مقیاس‌دهی افقی، پردازش روی تعداد بسیار زیادی از سرورها و با استفاده از روش MapReduce به انجام می‌رسد. تکنیک MapReduce در آغاز توسط دو تن از مهندسان گوگل و برای ایندکس کردن در کاربردهای مربوط به جستجوی وب معرفی شد و بعدها توسط بسیاری از پروژه‌ها مورد استفاده قرار گرفت.



آپاچی^۱ عرضه شده است. هادوپ چارچوبی برای پردازش، ذخیره‌سازی و تجزیه و تحلیل حجم انبوهی از داده‌های بدون ساختار و توزیع شده است.

به‌طور جامع و مختصر می‌توان گفت فناوری داده‌های عظیم به دو بخش عمده سخت‌افزار و نرم‌افزار تقسیم می‌شود که در شکل زیر نشان داده شده است. بخش سخت‌افزار به لایه‌های زیرساخت اشاره دارد؛ بخش نرم‌افزار نیز سازماندهی داده و مدیریت نرم‌افزار، تجزیه و تحلیل و کشف نرم‌افزار، سیستم‌های پشتیبانی تصمیم‌گیری و نرم‌افزارهای اتوماسیون را دربر می‌گیرد.^(۱۲)

شکل ۲. لایه‌های فناوری داده‌های عظیم



۲. کاربرد داده‌های عظیم در بخش دولت

گروه مک کینزی در سال ۲۰۱۱ پنج نوع ارزش جدید را که ممکن است از داده‌های عظیم به‌دست آورد را پیش‌بینی کرده که شامل موارد زیر است:

۱. ایجاد شفافیت در فعالیتهای سازمانی که می‌تواند برای افزایش بهره‌وری مورد استفاده قرار گیرد.

۲. قابلیت تجزیه و تحلیل دقیق‌تر از کارکنان و عملکرد سیستم که اجازه آزمایش و بازخورد را می‌دهد.

۳. بخش‌بندی جمعیت به‌منظور عملکردهای سفارشی.

۴. جایگزینی و پشتیبانی از تصمیم‌گیرهای انسانی با الگوریتم‌های خودکار.

۵. نوآوری در مدل‌های جدید کسب‌وکار، محصولات و خدمات.^(۱۳)

این ارزش‌های بالقوه باید توسط مدیران در صنایع و دولت پشتیبانی شوند. طبق مطالعات گروه گارتنر در سال ۲۰۱۲ در صنعت و از جنبه فناوری دولت مهم‌ترین بخش نسبت به سایر بخش‌های این فناوری است.^(۱۴)

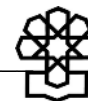
داده‌های عظیم در بسیاری از حوزه‌ها می‌توانند به توانمندسازی دولت در ارائه خدمات به شهروندان کمک کنند از جمله مدیریت ریسک، توسعه مدل‌های ریسک، پیگیری تبلیغات، نظارت بر استفاده از تلفن همراه، ارائه ابزارهایی برای توانمندسازی نیروی پلیس، بخش‌بندی بازار و مشتریان به‌طور هوشمند، ردیابی مکان، کشف روند رفتار مشتریان، تحلیل جریان‌های رسانه‌های اجتماعی، ردیابی محصولات از آغاز فرآیند تولید با به‌کارگیری داده‌های RFID، کنترل ترافیک، پیش‌بینی آب‌وهوا، پیشگیری از فاجعه امور مالی، کنترل تقلب، آموزش و پرورش، معاملات کسب‌وکار، امنیت ملی و بهداشت و سلامت.^(۲) و^(۳) بسیاری از برنامه‌های اولیه دولت الکترونیک در کشور ما بر این شیوه استوار است که اطلاعات را به‌صورت آنلاین از طریق برنامه‌های کاربردی وب در دسترس عموم شهروندان قرار دهد. اگر دولت، فناوری داده‌های عظیم را در آغوش نگیرد، به‌شدت مانع پیشرفت‌های علمی و حمایت مالی دولت در طرح‌های امنیتی خواهد شد. در ادامه به برخی از کاربردهای متنوع داده‌های عظیم اشاره می‌شود.

۱. امنیت اطلاعات

پیش‌بینی و پیشگیری از خطرات: امروزه چالش‌های پیش‌بینی و پیشگیری از تهدیدات در حال تبدیل شدن به مسئله‌ای پیچیده و نیازمند به رویکرد جدید در زمینه مدیریت اطلاعات است. تنوع، سرعت و حجم داده‌ها به‌طور فزاینده‌ای نیاز به پیش‌بینی و مقابله با روندهای در حال ظهور در زمینه فعالیت‌های غیرقانونی و تخلفات است. به‌منظور پیش‌بینی‌های مؤثرتر و جلوگیری از تهدیدات، سازمان‌های دولتی باید قادر به مدیریت داده‌های متعدد و مختلف باشند و از تجزیه و تحلیل پیشرفته برای تبدیل داده‌ها به دانش استفاده کنند.

۲. فرهنگی

جلوگیری از فریب‌های رسانه‌های اجتماعی: کشورهای مختلف به‌منظور کاهش آسیب‌های ناشی از برخی برنامه‌ها و تبلیغاتی که در رسانه‌های اجتماعی منتشر می‌شوند و کاهش هزینه‌های قضاوت و دادرسی و ایجاد بهبود یکپارچگی خدمات و برنامه‌های دولت از داده‌های عظیم استفاده می‌کنند. هدف از این استفاده، شناسایی و جلوگیری از تقلب و سوءاستفاده قبل از وقوع حادثه است. با ترکیب انواع



داده‌های ساختاریافته و غیرساختاریافته، شفاف ساختن هر چه بیشتر تصویر و تاریخچه شهروندان، کشف ارتباطات نامرئی و به اشتراک‌گذاری اطلاعات به صورت بلادرنگ برای تحلیل، می‌توان سازمان‌های دولت را در این زمینه توانمند ساخت.

۳. اقتصادی

ردیابی مالیاتی: سازمان‌های مالیات از طریق تقلب و سوءاستفاده مالیاتی شهروندان، هرساله مقدار زیادی درآمد خود را از دست می‌دهند. ایجاد یک رویکرد یکپارچه برای مبارزه با تقلب مالیاتی از اولویت‌های سازمان‌های مالیات کشورهاست. برخی دولت‌ها نظیر ایالات متحده آمریکا با ترکیب انواع زیادی از داده‌های عظیم از طریق تشخیص تاکتیک‌های جدید تقلب، کشف هویت چندگانه و شناسایی و بررسی رفتار مشکوک توانسته‌اند تقلب و سوءاستفاده مالیاتی را به حداقل برسانند.^(۱)

صرفه‌جویی قابل توجه در بودجه: استفاده از تجزیه و تحلیل یکپارچه، آنلاین و همزمان، صرفه‌جویی قابل توجهی در بودجه کشورها ایجاد می‌کنند. مقامات فناوری اطلاعات ایالت متحده مدعی‌اند که تجزیه و تحلیل همزمان داده‌های عظیم می‌تواند حداقل ۱۰ درصد از بودجه سالیانه را با تشخیص هزینه‌های اضافی در بخش سلامت کاهش دهد.

مدیریت بانک‌ها: به تدریج، حجم داده‌های عظیم بانک‌ها مهم‌ترین دارایی‌شان خواهد بود. این داده‌ها می‌توانند منشأ مزیت رقابتی و تمایز در چشم مشتریان باشند. بخش‌بندی وسیع مشتریان، تعریف سبدهای متنوع مشتریان، تجربه‌اندوزی دائمی از تغییرات خاستگاه مشتریان و بازتعریف مکرر بخش‌های مشتریان و سفارشی‌سازی خدمات بانکی خاص برای بخش‌های مشتریان، از جمله موارد قابل دستیابی با به‌کارگیری داده‌های عظیم خواهد بود.^(۱۵)

۴. سلامت

قابلیت بالقوه برای نجات زندگی افراد: با استفاده از نتایج تجزیه و تحلیل داده‌های عظیم سالیانه تعداد بسیار زیادی از افراد را می‌توان از مرگ نجات داد. به‌عنوان نمونه، پژوهشگران حوزه سلامت می‌توانند اطلاعات بهداشت و سلامت را برای آشکار ساختن الگوهایی که منجر به درمان‌های مؤثرتر و تشخیص شیوع بیماری‌اند جمع‌آوری کرده و مورد تجزیه و تحلیل قرار دهند.

افزایش کیفیت زندگی: بلادرنگ بودن داده‌های عظیم به بهبود کیفیت زندگی شهروندان کمک خواهد کرد. به‌عنوان مثال، با به‌دست آوردن حجم عظیمی از داده‌ها در سراسر سازمان‌های دولتی، به دولت در بهبود خدمات مورد نیاز شهروندان کمک می‌کند.^(۱۶)

پیش‌بینی وقوع بیماری: محققین قادرند داده‌های دقیقی که بازتاب الگوهایی از وقوع و گسترش یک بیماری‌اند را با جستجوی رفتار آنلاین مانند جستجوهای افراد در وب و داده‌های رسانه‌های اجتماعی به‌دست آورند. Google Dengue Trends در این زمینه با ادغام جستجوها، مدل دقیقی از روند شیوع بیماری تب استخوان‌شکن (Dengue) در کشورهای آمریکای جنوبی، آسیای جنوبی و جنوب شرقی ارائه داده است.^(۱۷)

همچنین محققان پزشکی با توانایی جمع‌آوری اطلاعات از بیماران مبتلا به سرطان در حد وسیع، می‌توانند مجموعه داده‌های ناشناسی که به‌طور بالقوه در شناسایی الگوهای درمانی سرطان‌ها مؤثرند را کشف کنند. ابزارهای بلادرنگ داده عظیم را می‌توان در جهت افزایش اثربخشی درمان و مراقبت‌های فردی در بیماری‌های خاص مورد استفاده قرار داد.

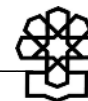
بخش امداد: امدادگران می‌توانند با بهره‌گیری از قابلیت‌های داده‌های عظیم برای پیش‌بینی مکان حادثه اضطراری قبل از وقوع، استفاده کنند. داده عظیم اجازه خواهد داد که کارکنان اورژانس برای شناسایی جایی که به آنها نیاز است اقدام کنند. همچنین اتصال به سابقه پزشکی کامل بیمار یک تشخیص دقیق‌تر و بهتر و مراقبت مقرون به‌صرفه‌تر را ارائه خواهد کرد.

۵. نیروی انتظامی

نیروی انتظامی برای کاهش جرم و جنایت: امروزه با استفاده از ابزارهای بلادرنگ داده‌های عظیم برای توسعه مدل‌های پیش‌بینی در مورد زمان و مکان جنایت، می‌توان به طرز چشمگیری نرخ جرم و جنایت را کاهش داد.^(۱۸) در شهر لس‌آنجلس آمریکا پلیس با همکاری دانشگاه کالیفرنیا، مدلی از میلیون‌ها آرشیو جرائم ایجاد کرده است که کمک به پیش‌بینی زمان و مکان وقوع جرائم می‌کند. تست این مدل در حومه شهر لس‌آنجلس گزارش روزانه‌ای از مکان‌هایی با احتمال وقوع جرائم بالا را تولید کرد و نتیجه آزمایش ثابت کرد که سرقت تا ۲۶ درصد و سایر جرائم تا ۱۲ درصد، در این مناطق کاهش یافته است.^(۱۷)

۶. انرژی

ابزارهای بلادرنگ داده‌های عظیم اجازه می‌دهند با بهره‌گیری از شبکه‌ای از حسگرها، خطرات بالقوه را پیش از تبدیل شدن آنها به خطرات جدی، کشف کرده و به مسئولین هشدار داد. این بدان معناست که لوله‌های گاز طبیعی و خطوط برق را می‌توان به‌منظور پیشگیری از خطرات احتمالی، شناسایی کرد و تکنسین‌ها می‌توانند برای بررسی یک مشکل قبل از وقوع حادثه اعزام شوند. داده‌های عظیم توانایی دیدن الگوهای شکست بالقوه‌ای را می‌دهد که پیش از آن قابل مشاهده نبوده است.



همچنین ابزارهای بلادرنگ داده‌های عظیم، قادر به اندازه‌گیری مصرف انرژی خانه‌های شخصی است. این داده‌ها، اطلاعات ارزشمندی را در اختیار بخش صنعت قرار می‌دهند تا درک بهتری از دوره‌های اوج تقاضا داشته باشند و همچنین به‌عنوان یک هشداردهنده زودهنگام از شبکه در مواقعی که ممکن است در معرض استفاده بیش‌ازحد قرار گیرند، برای جلوگیری از خاموشی به‌کار گرفته می‌شوند و پیش‌بینی و تعدیل تقاضای انرژی را میسر می‌شوند.

۷. هواشناسی

پیش‌بینی آب‌وهوا: مسئولین هواشناسی با استفاده از ابزارهای داده‌های عظیم می‌توانند مدل‌های هواشناسی را برای اماکن مسکونی به‌کار گیرند تا به‌سرعت مردم در معرض خطر را از شرایط اضطراری مربوط به آب‌وهوا مطلع کنند.

۸. آموزش

شناسایی دانش‌آموزان دارای تصمیم‌بالمقوه برای ترک تحصیل، قبل از اینکه آنها به صورت قطعی از حضور در کلاس درس خودداری کنند، یکی از قابلیت‌های داده عظیم در بخش آموزش است. با استفاده از این ابزار، مدیران مدرسه می‌توانند به دانش‌آموزان در معرض خطر پیش از اینکه خیلی دیر شده باشد، توجه بیشتری داشته باشند.

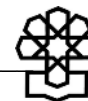
تجزیه و تحلیل بلادرنگ داده‌های عظیم می‌تواند ابزار موردنیاز برای بهبود مداوم تجربیات آموزشی دانش‌آموزان را به مقامات آموزش و پرورش ارائه کند. مدیران می‌توانند به‌سرعت تشخیص دهند که کدام روش کارآمد عمل می‌کند و کدام مفید واقع نمی‌شود. به‌عنوان مثال، برای برنامه‌ریزی در زمان کلاس ریاضی، مشخص کردن ساعاتی از روز که دانش‌آموزان برای یادگیری ریاضی بالاترین آمادگی را دارند یا پیدا کردن اندازه بهینه سائز کلاس‌ها براساس تجزیه و تحلیل داده‌ها.

۹. راهنمایی و رانندگی

کنترل ترافیک: داده‌های عظیم به‌وسیله تجزیه و تحلیل آب‌وهوا، رویدادها، ترافیک و سایر متغیرها، بینش موردنیاز را برای طراحی و مدیریت کارآمدترین الگوی حمل‌ونقل ارائه می‌دهد و مسئولین راهنمایی و رانندگی را قادر به تغییر چراغ توقف برای تنظیم جریان ترافیک می‌سازد. همچنین توانایی پالایش اطلاعات کاربران تلفن همراه را برای افزایش سرعت در شناسایی حوادث جاده‌ای و دیگر چالش‌های ترافیک می‌دهد؛ بنابراین به‌جای نظارت بر چند مکان خاص با دوربین‌های ترافیک، مقامات می‌توانند کل شبکه ترافیک را بلادرنگ نظارت کنند.

علاوه بر این با استفاده مؤثر از داده‌های عظیم، سازمان‌های دولتی به توانایی‌های زیر تجهیز می‌شوند:

۱. تصمیم‌گیری بهتر و سریع‌تر؛ با شناسایی روندها و پیش‌بینی‌های بالقوه داده‌های عظیم، سازمان‌ها می‌توانند تصمیم‌گیری خود را بهبود بخشند. با استفاده از داده‌های عظیم و با ابزارهای تجزیه و تحلیل و دیگر فناوری‌های پردازش داده‌ها، تصمیم‌گیری‌ها سریع‌تر انجام می‌شوند. بدون این ابزارها و فناوری‌ها، تصمیم‌گیرندگان ممکن است صرفاً به حدس زدن یا در نهایت اجتناب از تصمیم‌گیری قناعت کنند.
۲. بهبود نتایج مأموریت سازمان‌ها؛ داده‌های عظیم توانایی پیش‌بینی نتایج و ایجاد مدل و سناریوهای مختلف براساس داده‌ها را به ارمغان می‌آورد.
۳. شناسایی و کاهش ناکارآمدی؛ با استفاده از داده‌های عظیم، سازمان می‌تواند براساس داده‌های فرآیندهای تولیدی خود، برداشتن گام‌های غیرضروری را شناسایی کند.
۴. از بین بردن تقلب و سوءاستفاده؛ با شناسایی ناکارآمدی، سازمان‌ها می‌توانند عناصر مخرب را از سازمان بیرون کنند. با توجه به مأموریت‌های خود، سازمان همچنین می‌تواند به‌طور بالقوه، تقلب و سوءاستفاده توسط افراد یا احزاب را شناسایی کرده و از بین ببرند.
۵. بهبود بهره‌وری؛ با ابزار مناسب، حتی کاربران غیرفنی می‌توانند با مجموعه داده‌های عظیم، برای پیدا کردن اطلاعات، ارائه خدمات بهتر و یا تصمیم‌گیری، در جهت حمایت از مأموریت سازمان اقدام کنند.
۶. افزایش نرخ بازگشت سرمایه و کاهش هزینه مالکیت؛ در قلب راه‌حل‌های داده‌های عظیم استفاده بهتر از داده‌هایی که سیستم‌های فناوری اطلاعات تولید می‌کنند و بهبود بازده سرمایه‌گذاری در فناوری اطلاعات، است. با به‌کارگیری مناسب از منابع داده و ابزارهای تجزیه و تحلیل، سازمان‌ها می‌توانند برای تأسیسات خود کل هزینه مالکیت را کاهش دهند.
۷. افزایش شفافیت و خدمات؛ اداره و پردازش مناسب داده‌ها اجازه می‌دهد تا سازمان‌ها داده‌ها را نه تنها در بخش شرکای خصوصی دولتی و خصوصی، بلکه به عموم مردم ارائه کنند و شهروندان را قادر می‌سازد تا بفهمند دولت چه اطلاعاتی را جمع‌آوری می‌کند. پردازش و به اشتراک‌گذاری داده‌های عظیم نیز اجازه می‌دهد تا سازمان‌ها، اطلاعات را به‌عنوان یک سرویس ارائه کنند.
۸. کاهش تهدیدات امنیتی، جرم و جنایت؛ تجزیه و تحلیل داده‌های عظیم کلیدی برای کمک به پلیس، مقامات امنیتی کشور و تحلیلگران اطلاعاتی است و با دقت در کشف الگوهای پنهان و سایر اطلاعات برای کمک به شناسایی تهدیدهای خاص به کار می‌رود.^{(۱۸) و (۱۹)}

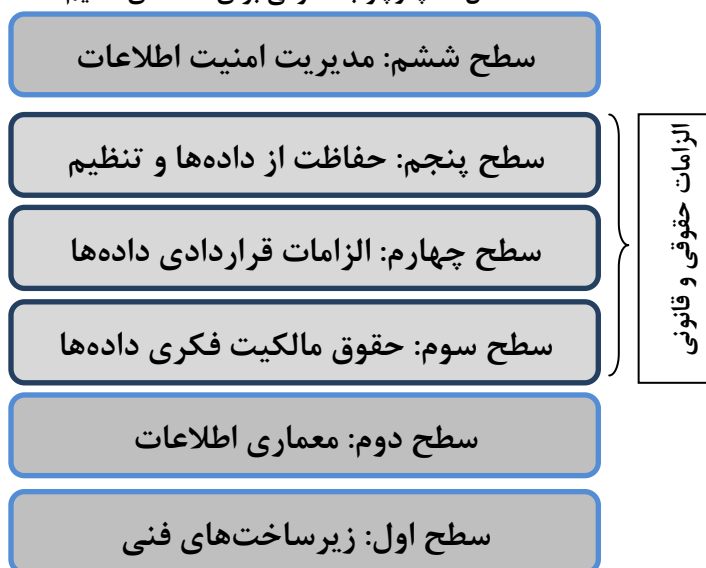


۳. داده‌های عظیم، الزامات حقوقی و سیاستگذاری

الزامات توسعه بهره‌گیری از داده‌های عظیم را می‌توان در سیستمی شامل زیرساخت‌های فنی، معماری اطلاعات، زیرساخت‌های حقوقی و قانونی و مدیریت امنیت اطلاعات در نظر گرفت. الزامات حقوقی و قانونی را می‌توان در قالب حقوق مالکیت فکری، حفاظت از داده‌ها، الزامات قراردادی داده‌ها و تنظیم قوانین تبیین کرد. امنیت داده‌ها نیز به چگونگی محافظت از اطلاعات حساس رقابتی و یا دیگر اطلاعاتی اطلاق می‌شود که باید خصوصی نگهداری شوند. با افزایش تعارض به حریم داده‌ها، پرداختن به موضوع امنیت داده‌ها از طریق ابزارهای فناورانه و سیاستگذاری‌ها امری ضروری به‌شمار می‌رود.^(۱۳)

موارد فوق را می‌توان در یک مدل تحلیلی حقوقی همانند شکل زیر بیان کرد. این چارچوب با رعایت موارد حقوقی می‌تواند برای سازمان‌ها و دولت‌ها در جهت به‌کارگیری داده‌های عظیم مورد استفاده قرار گیرد. چارچوب پیشنهادی دارای ۶ سطح است که در ادامه به شرح جزئیات آن پرداخته شده است.

شکل ۳. چارچوب حقوقی برای داده‌های عظیم



سطح اول: زیرساخت‌های فنی

این سطح شامل زیرساخت‌های فیزیکی از جمله سرورها، ذخیره‌سازی، دستگاه‌های کاربر و غیره است. همچنین نرم‌افزارهایی مانند سیستم‌عامل، میان‌افزارها، نرم‌افزارهای ارتباطی و دسترسی به داده‌ها برنامه‌های کاربردی مانند تجزیه و تحلیل و هوشمندی کسب‌وکار^۱ در این لایه قرار می‌گیرند. تجزیه و

تحلیل حقوقی در این سطح در حوزه مسائل کپی‌رایت نرم‌افزار و روابط حقوقی میان کپی‌رایت و پایگاه داده‌ها، در ارتباط با دسترسی و استخراج اطلاعات از نرم‌افزارهای پایگاه داده است.

سطح دوم: معماری اطلاعات

معماری اطلاعات^۱ در مرکز استانداردسازی جریان داده‌ها در سازمان‌ها قرار گرفته است. جایگاه حقوق مالکیت فکری معماری اطلاعات در عمل به راحتی نادیده گرفته شده است. استانداردسازی معماری اطلاعات و چگونگی اتخاذ سیاست‌های مرتبط با حقوق مالکیت فکری، به طور معمول توسط شورای سیاستگذاران در این محدوده تعیین می‌شود.

سطح سوم: حقوق مالکیت فکری داده‌ها

مهم‌ترین حقوق مالکیت فکری در رابطه با داده‌ها، کپی‌رایت، قوانین پایگاه داده و محرمانه بودن است. حق امتیازات و قوانین اختراعات می‌تواند برای نرم‌افزار و فرآیندهای کسب‌وکار که داده را پردازش می‌کنند، پذیرفته شوند اما شامل قوانین مرتبط با خود داده نمی‌شوند. کپی‌رایت از شکل یا چگونگی بیان اطلاعات محافظت می‌کند. کپی‌رایت در زمینه داده‌ها از جمله انتشار پژوهش‌ها، تجزیه و تحلیل بازار سهام و اسناد تکنیکی مرتبط با نرم‌افزارهای کامپیوتری و معماری اطلاعات به کار برده می‌شود.

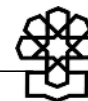
سطح چهارم: الزامات قراردادی داده‌ها

قوانین قرارداد در رابطه با داده‌ها به لحاظ فنی به طور کامل از حقوق مالکیت فکری جدا هستند. نواحی کلیدی برای قرارداد داده‌های عظیم مشابه مواردی است که در دیگر قراردادهای مرتبط با داده‌ها منعقد می‌شود. از جمله نکات خاصی که در الزامات قراردادی باید توجه کرد:

- چگونگی استفاده از داده‌ها،
- اهدافی که داده‌ها می‌توانند مورد استفاده قرار گیرند،
- مالکیت حقوق اساسی و حقوق اطلاعات به دست آمده و مشتق شده،
- ضمانت‌نامه انطباق با قوانین و مقررات، حفاظت از داده‌ها، بخش مقررات خاص و حسابرسی،
- تخصیص ریسک،
- مدت زمان تعلیق و فسخ تأمین خدمات،
- امکان استفاده از داده‌ها پس از پایان یافتن مدت عرضه خدمات و چگونگی برخورد با اطلاعات به دست آمده است.

سطح پنجم: حفاظت از داده‌ها و تنظیم قوانین

منطقه حقوقی سوم که اهمیت آن برای داده عظیم افزایش می‌یابد تنظیم قوانین است. قانون حفاظت داده‌ها (که در داخل این قانون، قوانین کپی‌رایت، پایگاه داده و محرمانگی وجود دارد)، قرارداد و



آیین‌نامه‌های مقرراتی، مجموعه‌ای گسسته از هنجارها را تشکیل می‌دهند که هرکدام دارای قوانین فنی مرتبط با خودند.

حفظ حریم خصوصی در حفاظت از داده‌ها عمدتاً به اطلاعات هویتی شخصی اشاره دارد که می‌تواند برای شناسایی فرد مورد استفاده قرار گیرد. تعریف حریم خصوصی بستگی کامل به فرهنگ و زمینه‌های اجتماعی و محیطی دارد. در عین حال محافظت از حریم خصوصی معمولاً به‌عنوان ابزاری برای ترسیم محدوده خطوطی که جامعه می‌تواند در امور افراد دخالت کند تلقی می‌شود. آنچه افراد در اینجا دنبال می‌کنند آن است که بر اطلاعات شخصی‌شان کنترل و نظارت داشته باشند و به این ترتیب، تنها بحث افشای آنها مطرح نیست و استفاده‌های آتی یا مکرر را نیز دربر می‌گیرد.^(۲۰)

در حقیقت دو حوزه حمایت داده‌ها و حریم خصوصی با یکدیگر همپوشانی دارند و آن حوزه‌ای که تحت شمول هر دو قرار می‌گیرد را می‌توان حریم اطلاعات یا بهتر از آن حریم داده‌ها نامید اما در اینکه چه داده‌هایی «خصوصی» تلقی می‌شود دیدگاه‌های بسیار متنوعی وجود دارند.^(۲۱)

در ایالات متحده قوانین حفظ حریم خصوصی فدرال بر استفاده از اطلاعات مالی و پزشکی حاکم است، اما اینکه چگونه این قوانین برای داده‌های عظیم اعمال می‌شوند همیشه روشن نیست. برای مثال اخیراً، طرح دعوی در دادگاه علیه گوگل توسط کاربران اندروید، مدعی است که این شرکت با نقض قوانین فدرال و تخلف از قانون و همچنین نقض قانون رقابت غیرمنصفانه کالیفرنیا،^۱ داده‌های جغرافیایی افراد و دیگر اطلاعات شخصی را به توسعه‌دهندگان برنامه‌ها انتقال داده است.^(۲۲)

صرف‌نظر از اینکه چگونه یک چارچوب قانونی برای حفاظت از حریم خصوصی وجود دارد، سازمان‌ها باید در مورد واکنش مشتریان نسبت به چگونگی رفتار با داده‌هایشان نگران باشند. باید توجه داشت که با وجود خلأهای قانونی و مقرراتی داده‌های عظیم، این فناوری در حال حاضر به‌طور گسترده‌ای در دسترس است و حتی به‌طور فزاینده‌ای به یک ابزار حیاتی برای کمک به مدیریت مصرف‌کنندگان و کمک به کسب‌وکار و دولت برای بهبود ارائه محصولات و خدمات به‌صورت مؤثرتر و ایمن‌تر تبدیل می‌شود.^(۱۸)

سطح ششم: مدیریت امنیت اطلاعات

در بالای چارچوب حقوقی داده‌های عظیم، مدیریت امنیت اطلاعات قرار گرفته است. سازمان استانداردهای بین‌المللی،^۲ سری ۲۷۰۰۰ از استانداردهای سیستم‌های مدیریت امنیت اطلاعات^۴ را منتشر کرده است و در ایالات متحده آمریکا، استانداردهایی را برای چگونگی اعلام گزارش خدمات

1. California's Unfair Competition Law
2. ISO
3. ISO/ IEC 27000-Series
4. Information Security Management System (ISMS)

شرکت در مورد انطباق کنترل‌ها و امنیت اطلاعات ارائه داده که قابل به‌کارگیری در پروژه‌های داده‌های عظیم است.^(۳)

مجموعه قوانین امنیت اطلاعات، قوانین ملی اعطا شده توسط قانون ملی هر کشور است که قابل اجرا در دادگاه‌های ملی همان کشور است. در میان قوانین کشورهای مختلف تفاوت‌هایی از طیف جزئی تا بزرگ وجود دارند. این تفاوت‌ها در مقررات فنی و همچنین کاربردهای متقارنی که از به‌کارگیری قوانین مختلف با داده‌های مشابه وجود دارد و از طرف دیگر چند لایه بودن قوانین در چرخه عمر داده‌ها، سرعت فرآیندها و تفاوت‌های میان قوانین ملی، هرکدام به پیچیدگی‌های قانونی داده‌های عظیم کمک کرده و چالش‌های حقوقی در زمینه پروژه‌های داده‌های عظیم ایجاد می‌کند.

۴. مسائل قانونی پیش روی کشورهای مختلف و سیاست‌های دولت آنها

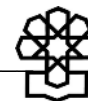
۴-۱. مسائل قانونی داده‌های عظیم در آمریکا

جمع‌آوری اطلاعات شخصی در ایالات متحده بدون محافظت کافی از حفظ حریم خصوصی منجر به افزایش سرسام‌آور سرقت هویت، نقض امنیت و کلاهبرداری مالی شده است و نارسایی‌هایی در عدالت و دادرسی و فرآیندهای ناشی از آن به‌وجود آورده است. محیط کنونی داده‌های عظیم در ایالت متحده به وسیله نقض داده‌ها و استفاده تبعیض‌آمیز از تجزیه و تحلیل آینده‌نگر داده‌ها رنج می‌برد. استفاده از تجزیه و تحلیل آینده‌نگر توسط بخش‌های دولتی و خصوصی در حال تضعیف آزادی در این کشور است. ارتباطات اجتماعی آنلاین، شرکت در بحث‌های آنلاین و دریافت علائق شهروندان از طریق دنبال کردن فعالیت‌های آنلاین، هم‌اکنون می‌تواند توسط دولت و شرکت‌ها در مواردی مانند تعیین حدود توانایی مالی شهروندان، توانایی به‌دست آوردن یک شغل، صلاحیت افراد یا اطلاعات کارت اعتباری، مورد استفاده قرار گیرند.^(۲۲)

عدم امنیت داده‌ای کافی در مؤسسات تجاری برای محافظت از حریم خصوصی شهروندان در آمریکا منجر به نقض حقوق آنها می‌شود. در طول فصل تعطیلات، میلیون‌ها نفر از مشتریان آمریکایی که از سازمان تجاری Target و Neiman Marcus خرید کرده‌اند از نقض حفاظت داده رنج می‌برند. به علت هک شدن پایانه‌های فروش و نبود استانداردهای امنیتی کافی در حدود ۷۰ میلیون نقض حفاظت داده صورت گرفته است. این اتفاق برای دارندگان حدوداً چهل میلیون حساب که دربردارنده شماره کارت اعتباری، تاریخ انقضا، کد امنیتی سه رقمی^۱ و حتی شماره شناسایی شخصی^۲ بودند، رخ داده است.^(۲۳)

1. Cvv number

2. Personal Identification Number (PIN)



نقض امنیتی حریم شخصی در داده‌های عظیم در مؤسسات آموزشی ایالات متحده باعث به خطر افتادن حریم خصوصی دانش‌آموزان و استادان نیز شده است. در دانشگاه مریلند، پایگاه داده حاوی ۳۰۹,۰۷۹ رکورد از دانشجویان، کارکنان و پرسنل هک شد. این داده‌ها شامل نام، شماره امنیت اجتماعی، تاریخ تولد و شماره شناسایی دانشگاه بود که پرونده‌های بیست سال گذشته دانشگاه را پوشش می‌داد. پس از دانشگاه مریلند در دانشگاه ایندیانا مشخصات کامل حدوداً ۱۴۶,۰۰۰ دانش‌آموز و فارغ‌التحصیل در معرض سرقت هویت و دیگر اشکال تقلب قرار گرفت.^(۲۳)

- چارچوب خط‌مشی‌گذاری مرکز اطلاعات حریم خصوصی الکترونیکی^۱ ایالات متحده

در سال ۲۰۱۴، اوباما طرحی را با نگاه جامع به مسائل حریم خصوصی در داده‌های عظیم اعلام کرد. تقریباً بلافاصله پس از اعلام کاخ سفید، مرکز اطلاعات حریم خصوصی الکترونیکی ایالات متحده، در ائتلاف با حریم خصوصی مصرف‌کنندگان و منافع عمومی، علمی و سازمان‌های آموزشی، درخواستی مبنی بر درگیر کردن عموم مردم با پذیرش نظرات عمومی، درباره حریم خصوصی در آینده و داده‌های عظیم به دفتر سیاست‌های علمی و فناوری^۲ داد. علت این درخواست آن است که از دیدگاه این ائتلاف، حفظ حریم خصوصی افراد توسط دولت و شرکت‌های خصوصی بزرگ در خطر است.

مرکز اطلاعات حریم خصوصی الکترونیکی از اصول تنظیم شده‌ای، به‌عنوان نقطه شروع برای ایجاد یک چارچوب که بتواند با چالش‌های داده‌های عظیم روبرو شود، پشتیبانی می‌کند و گنجانند این اصول را در گزارش کاخ سفید توصیه کرده که به‌شرح زیر است:

- **شفافیت:**^۳ نهادهایی که اطلاعات شخصی را جمع‌آوری می‌کنند در مورد اینکه چه اطلاعاتی را و چگونه آنها را جمع‌آوری کرده، چه کسانی به آن دسترسی خواهند داشت و چگونه مورد استفاده قرار می‌گیرند، باید به‌طور شفاف توضیح دهند. علاوه بر این، الگوریتم به‌کار گرفته شده در داده‌های عظیم باید در دسترس عموم قرار داده شود.

- **نظارت:**^۴ مکانیسم‌هایی که اطلاعات افراد را مورد بررسی قرار می‌دهند، باید در محلی قرار داده شوند تا از یکپارچگی داده‌ها و الگوریتم‌های به‌کار گرفته شده در تجزیه و تحلیل داده‌ها اطمینان حاصل شود. این مکانیسم‌ها باید در تضمین دقت و عدالت در تصمیم‌گیری یاری‌رسان باشند.

- **پاسخگویی:**^۵ اشخاصی که به‌طور نادرست از داده‌های شخصی برای برجسته‌سازی یا تبعیض استفاده می‌کنند باید پاسخگو باشند. مردم باید به راحتی قادر به دسترسی و تصحیح اطلاعات نادرستی که در مورد آنها جمع‌آوری شده باشند.

1. Electronic Privacy Information Center (EPIC)
2. Office of Science and Technology Policy (OSTP)
3. Transparency
4. Oversight
5. Accountability

– تکنیک‌های محرمانگی مقاوم:^۱ تکنیک‌هایی که کمک به دست آوردن مزایای استفاده از داده‌های عظیم می‌کنند، در حالی که ریسک‌ها و خطرات حفظ حریم خصوصی را به حداقل می‌رسانند باید تشویق شوند؛ اما این تکنیک‌ها باید مقاوم، مقیاس‌پذیر، قابل اثبات و کاربردی باشند.

– ارزیابی معنی‌دار:^۲ مؤسسات و سازمان‌هایی که از داده‌های عظیم استفاده می‌کنند باید مفید بودن آن را با توجه به اساس کار خود ارزیابی کنند و از جمع‌آوری و حفظ اطلاعاتی که برای هدف مورد نظر خود لازم ندارند، خودداری کنند. مردم ایالت متحده آموخته‌اند که برنامه فراداده گسترده^۳ ایجاد شده توسط آژانس امنیت ملی ایالات متحده^۴ عملاً هیچ نقش قابل توجهی در تجسس تروریسم بازی نکرده است، آنها همچنین بر این باورند این مسئله برای بسیاری از برنامه‌های داده‌های عظیم نیز صادق باشد.

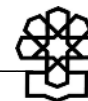
– کنترل:^۵ شهروندان باید قادر باشند به اعمال کنترل بر داده‌هایی که توسط خودشان ایجاد شده یا به آنها مرتبطند و بتوانند در مورد اینکه چه داده‌هایی باید جمع‌آوری شود و در صورت جمع‌آوری چگونه باید از آنها استفاده کرد، تصمیم بگیرند.^(۲۲)

۲-۴. مسائل قانونی داده‌های عظیم در استرالیا

در کشور استرالیا مأموریت داده‌های عظیم، فراهم آوردن درک بهتر از رفتارها، ترجیحات و انتظارات مشتری برای دولت و سازمان‌های بخش دولتی و خصوصی برای به حداکثر رساندن بهره‌وری، کاهش هزینه کسب‌وکار و مصرف‌کنندگان و بهبود محصولات و خدمات است. با این حال، بسیاری از شهروندان استرالیایی، درک عمیقی از چگونگی مورد استفاده قرار گرفتن داده‌های عظیم ندارند و تاکنون تعداد کمی از آنها منافع داده‌های عظیم را درک کرده‌اند. تحریک رسانه‌های عمومی و نگرانی در مورد از دست دادن حریم خصوصی، مصرف‌کنندگان را به بدگمانی در مورد پدیده داده‌های عظیم سوق داده است.

قانون حفظ حریم خصوصی کشور استرالیا از سال ۱۹۸۸ میلادی می‌باشد و داده‌های عظیم از این قوانین مستثنا نیست. برای داده‌های عظیمی که شامل هرگونه اطلاعات شخصی است، همان قوانین حریم خصوصی ۱۹۸۸ جهت جمع‌آوری، استفاده و آشکارسازی اعمال می‌شود. جمع‌آوری اطلاعات شخصی با رضایت مشتریان، آگاه ساختن افراد از هدف جمع‌آوری، امکان دسترسی افراد به اطلاعات خودشان و تأمین موارد امنیتی از سوی دولت برای جلوگیری از افشای غیرمجاز از جمله آن قوانین است.

-
1. Robust Privacy Techniques
 2. Meaningful Evaluation
 3. massive metadata program
 4. United States National Security Agency (NSA)
 5. Control



علاوه بر رعایت قوانین حریم خصوصی ۱۹۸۸، اصول جدید حفظ حریم خصوصی در استرالیا از ماه مارس سال ۲۰۱۴ با تصویب مجموعه‌ای از «برنامه‌ها» مؤثرتر شد. این برنامه‌ها به‌طور خلاصه در ادامه بیان شده‌است.^(۱۷)

برنامه ۱: شفافیت

مدیریت باز و شفاف از اطلاعات شخصی: شفافیت در مورد اطمینان از آگاهی مصرف‌کنندگان از داده‌های جمع‌آوری شده و هدف استفاده از آنها امری ضروری است، در نتیجه به مصرف‌کنندگان اجازه می‌دهد انتخاب آگاهانه‌ای در مورد به اشتراک‌گذاری اطلاعات شخصی خود داشته باشند.

برنامه ۲: گمنامی و جعل هویت

این قانون در حال حاضر برای چگونگی تصمیم‌گیری در مورد افراد یا شرکت‌هایی است که ناشناس‌اند یا با نام مستعار در جامعه حضور پیدا کرده‌اند.

برنامه ۳: جمع‌آوری اطلاعات شخصی

سازمان‌ها تنها به جمع‌آوری اطلاعات شخصی به‌طور مستقیم از فرد یا از یک شخص ثالث (به‌عنوان مثال جمع‌آوری داده‌ها به‌صورت آنلاین) مجازند. جمع‌آوری داده‌های شخصی تنها در مواردی امکان‌پذیر است که به‌منظور تحقق اهداف معین و از پیش تعیین شده‌ای به‌کار گرفته شوند، یا مربوط به یکی از عملکردها یا فعالیت‌های صریح سازمان باشند. جمع‌آوری اطلاعات حساس نیز می‌بایست با رضایت اشخاص باشد.

برنامه ۴: اطلاعات شخصی ناخواسته

این قانون اظهار می‌کند اگر سازمانی اطلاعات شخصی را دریافت کرد که مورد درخواست نبود، باید آن را نابود یا حذف کند، مگر اینکه سازمان بتواند ثابت کند که این اطلاعات مطابق قوانین برنامه ۳، جمع‌آوری شده‌اند.

برنامه ۵: اطلاع‌رسانی

افراد را باید قبل از جمع‌آوری یا در زمان جمع‌آوری از اینکه چگونه و چرا اطلاعات شخصی آنها جمع‌آوری می‌شود، آگاه کرد.

برنامه ۶: استفاده و آشکارسازی

اطلاعات شخصی نباید استفاده و یا آشکارسازی شوند مگر اینکه با رضایت فرد باشد یا برای هدف جدیدی مرتبط با اهداف اولیه جمع‌آوری اطلاعات باشد.

برنامه ۷: بازاریابی مستقیم

نهادها و سازمان‌ها نباید از اطلاعات شخصی به‌منظور اهداف بازاریابی مستقیم استفاده کنند مگر اینکه فرد منطقیاً انتظار آن را داشته باشد که سازمان به استفاده از اطلاعات برای بازاریابی مستقیم اقدام کند.

برنامه ۸: افشای مرزی

افشای مرزی زمانی صورت می‌گیرد که اطلاعات شخصی جمع‌آوری شده در استرالیا توسط شخص یا اشخاصی در یک کشور دیگر مشاهده شود. باید در نظر داشت اطلاعات شخصی ذخیره شده در سرورهای کشورهای دیگر که امکان دسترسی به داده‌ها را داشته باشند، خطرپذیری «افشا» شدن را در پی خواهد داشت، بنابراین استفاده از خدمات مبتنی بر رایانش ابری نیاز به مدیریت چنین خطرپذیری‌هایی در زمینه افشای مرزی را دارد.

برنامه ۹: کیفیت اطلاعات شخصی

سازمان‌ها باید گام‌های معقولی برای اطمینان از دقت در جمع‌آوری، استفاده یا آشکار کردن اطلاعات، بردارند. دقت شامل حصول اطمینان از به‌روز، کامل و مرتبط بودن اطلاعات است.

برنامه ۱۰: امنیت اطلاعات شخصی

سازمان‌ها باید سیاست‌های ویژه‌ای برای محافظت از اطلاعات شخصی تدوین کنند. این سیاست‌ها شامل اتخاذ اقدامات امنیتی در زمینه فیزیکی و فناوری است. همچنین اطلاعات شخصی باید از سوءاستفاده، مفقود شدن، افشا شدن و دسترسی‌های غیرمجاز محافظت شوند.

برنامه ۱۱: دسترسی به اطلاعات شخصی

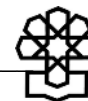
سازمان‌ها باید اجازه دسترسی افراد به اطلاعات شخصی‌شان را در صورت درخواست بدهند.

برنامه ۱۲: اصلاح اطلاعات شخصی

سازمان‌ها موظفند اطلاعات شخصی را بنا به درخواست افراد تصحیح کنند.^(۱۶)

جمع‌بندی

قابلیت‌های رو به رشد فناوری داده‌های عظیم در تصرف، تجمیع و پردازش هرچه بیشتر داده‌های حجیم، در حال حرکت و متنوع موجب شده است تا کاربردهای گسترده‌تری از داده‌ها فراهم شود. تجزیه و تحلیل داده‌های عظیم می‌تواند بهره‌وری و اثربخشی را در طیف گسترده‌ای از کسب‌وکارها و مسئولیت‌های دولت افزایش دهد. دولت‌ها می‌توانند با تجزیه و تحلیل داده‌های عظیم، اقدامات مؤثری را برای رفاه شهروندان انجام دهند. پیشگیری از جرم، جنایت، بهبود ترافیک و حمل‌ونقل افزایش امنیت ملی، مدیریت درآمد، خدمات اجتماعی و بهبود وضعیت محیط زیست نمونه‌ای از این اقدامات است. در بخش خصوصی نیز به‌کارگیری و تحلیل این داده‌ها، بهبود و اثربخشی کسب‌وکارها را به همراه خواهد داشت.^(۲۴)



توسعه بهره‌گیری از داده‌های عظیم علاوه بر تأمین زیرساخت‌های فنی، امنیتی و اجتماعی نیازمند تأمین الزامات حقوقی و قانونی جهت به‌کارگیری داده‌های عظیم است. ابعاد مهم الزامات حقوقی و قانونی داده‌های عظیم شامل حقوق مالکیت فکری، حفاظت از داده‌ها و الزامات قراردادی است. حفاظت از داده‌های شخصی شهروندان یکی از مهم‌ترین اولویت‌های اصلی در تأمین الزامات حقوقی داده‌های عظیم است، زیرا حفظ حریم خصوصی به‌عنوان یک حق اساسی شهروندان شمرده می‌شود و تصویب قوانین حفاظت از داده‌های شخصی در فضای مجازی کلید مؤثری در جهت توسعه بهره‌گیری از داده‌های عظیم است. بدون وجود چنین قانونی بالطبع شهروندان از به اشتراک‌گذاری اطلاعات در فضای مجازی اجتناب می‌ورزند یا اطلاعات صحیحی ارائه نمی‌دهند. در کشور ما با توجه به خلأ قانونی در زمینه حفاظت از داده‌ها در فضای مجازی، تصویب قانون در این موضوع از اولویت‌های اصلی قانونگذاری حوزه فناوری اطلاعات و ارتباطات است. توجه به الزامات قراردادی استفاده از داده‌ها در کاربردهای داده‌های عظیم نیز از عوامل دیگر کاهش چالش‌های حقوقی مرتبط است. حقوق مالکیت فکری داده‌ها در کشور تا مشخص نشدن موضع کشور در موضوع کپی‌رایت در بلا تکلیفی به‌سر می‌برد که لازم است این موضع به روشنی در قالب لایحه مالکیت فکری که از طرف دولت به مجلس ارسال شده است، شفاف شود.

پی‌نوشت‌ها

1. IBM, "IBM Big Data Success Stories", 2011.
2. Kemp,R., " Legal Aspects of Managing Big Data", September 2014.
3. Yan,J., " Big Data, Bigger Opportunities ", April 9, 2013.
4. Information Commissioner's Office, "Big Data and Data Protection", 2014.
5. Big Data: Seizing Opportunities, Preserving Value", <http://www.whitehouse.gov/issues/technology/big-data-review>. The report focuses on 'how big data will transform the way we live and work and alter the relationships between government, citizens, businesses, and consumers.'
6. NASSCOM. " Big Data: The Next Big Thing. International Youth Center, NASSCOM, New Delhi, India, 2012, <http://ebookbrowse.com/big-data-report-2012-pdf-d410599517>.
7. NetApp, " The big data gap", May 7, 2012,<http://www.meritalk.com/bigdatagap>.
8. Bobley, B., "Digging into Data", Presentation at the Big Data Conference, Digital Government Institute, Washington, DC, October 11, 2012.
9. GeekInterview, November 17, 2008, <http://www.learn.geekinterview.com/it/data-modeling/semi-structured-model/1.html>.
10. IBM, "Analytics: The real-world use of big data", 2012.
11. Juniper NETWORKS, Introduction to Big Data: Infrastructure and Networking Considerations, 2012.
12. McKinsey Global Institute, "Big data: The next frontier for innovation, competition, and productivity", May 2011.

13. Columbus, L., "Roundup of Big Data Forecasts and Market Estimates ", 2012, Forbes. <http://www.forbes.com/sites/louiscolumbus/2012/08/16/roundup-of-big-data-forecasts-and-market-estimates-2012>.
14. Electronic Privacy Information Center (EPIC), " Request for Information: Big Data and the Future of Privacy", April 4, 2014.
15. <http://www.techamericafoundation.org/big-data-can-save-money-and-lives-say-government-it-officials>.
16. Association for Data-driven Marketing & Advertising, "BEST PRACTICE GUIDELINE: BIG DATA", 2013.
17. SAP Solutions for Analytics, " Big Data Analytics Guide: Better technology, more insight for the next generation of business applications", 2012.
18. statetechmagazine, "8 Benefits of Big Data for State and Local Governments", 2013, <http://www.statetechmagazine.com/article/2013/05/8-benefits-big-data-state-and-local-governments>.
19. ADVISEN Insurance Intelligence, the Privacy Implications of Big Data, June 2013.
۲۰. فقیهی، مهدی، معمارزاده طهران، غلامرضا و رفوگر آستانه، حسین، حفظ حریم خصوصی بیماران، پیش نیاز توسعه سلامت الکترونیک، فصلنامه اخلاق پزشکی، سال چهارم، شماره دوازدهم، تابستان ۱۳۸۹.
۲۱. فولادی، محمد، اخلاق روزنامه‌نگاری، آسیب‌ها و چالش‌ها، ماهنامه معرفت، شماره ۱۲۳.
22. Comments of the Electronic Privacy Information Center to the Office of Science Request for Information: Big Data and the Future of Privacy and Technology Policy, April 4, 2014.
23. TechAmerica, " DEMYSTIFYING BIG DATA ", A Practical Guide To Transforming The Business of Government.
۲۴. فقیهی، مهدی، جلیلیان عطار، مریم، بررسی الزامات حقوقی و قانونی بهره‌گیری از داده‌های عظیم با تأکید بر حفاظت از داده‌ها و الزامات قراردادی داده‌ها، همایش داده‌های عظیم، مرکز تحقیقات مخابرات ایران، دی‌ماه ۱۳۹۳.



مرکز پژوهش‌ها
مجلس شورای اسلامی

شماره مسلسل: ۱۴۲۷۳

شناسنامه گزارش

عنوان گزارش: فناوری داده‌های عظیم و الزامات قانونی آن

نام دفتر: مطالعات ارتباطات و فناوری‌های نوین (گروه ارتباطات و فناوری اطلاعات)

تهیه و تدوین‌کنندگان: مهدی فقیهی، مریم جلیلیان عطار

ناظر علمی: محمدرضا محمدخانی

همکار: مهدی معافی

متقاضی: معاونت پژوهش‌های زیربنایی و امور تولیدی

ویراستار تخصصی: محمدمین فصیحی

ویراستار ادبی: قاسم میرخانی

واژه‌های کلیدی:

۱. داده‌های عظیم

۲. فناوری ارتباطات و اطلاعات

۳. Big Data



تاریخ انتشار: ۱۳۹۴/۳/۴