ماهنامه همگانی دانش و مهندسی کامپیوتر

سال۳۰ شماره ۲۸۰ دی ۱۳۹۹ ۲۴صفحه بها: ۱۰۰۰۰تومان

ریزپردازنده

www.rizpardazandeh.com
ISSN: 2008-2088

• اینترنت آدم‌ها:

آنالیتیک، کلان‌داده‌ها،

و هوش مصنوعی برای پشتیبانی تصمیم

• آنالیتیک کلان‌داده‌ها و

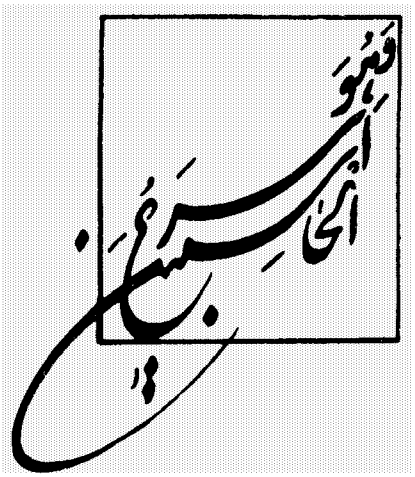
تعارض نظریه ولایت فقیه با بنیادهای دین

• تُنددادەها

• پشتیبانی هوش مصنوعی در تصمیم‌سازی

• تصمیم‌سازی خودکار

www.rizpardazandeh.com



ریزپردازنده

ماهنامه همگانی دانش و مهندسی کامپیوتر

سال۳۰، شماره ۲۸۰، انتشار دی ۱۳۹۹

شماره شاپا: ۲۰۸۸-۲۰۰۸ (ISSN: 2008-2088)

- صاحب امتیاز و مدیر مسئول: علیرضا محمدی‌فر
- تلفن ماهنامه ریزپردازنده: ۸۸٤۳٤۱٦۹
- تلفن همراه: ۰۹۱۲-٥٤۹٥٥٤٦
- نمابر: ۸۸٤۲۱۱۷۰
- نشانی: تهران، صندوق پستی ۱۵۸۷۵/٦٥۹۱، مجله ریزپردازنده (سهروردی، نیکان، پلاک۲۳)
- نشانی وب: http://www.rizpardazandeh.com
- ایمیل: rizpardazandeh@gmail.com
- کانال تلگرام: rizpardazandeh@
- چاپ: امین ۸۸٤۱۷۹٦۸ (سبلان، شهیدعلی‌اصغرنوری، شماره۱٦٦، کدپستی۱٦۳۷٦٤۹۷٤۷)
- صحافی: ایمان ۸۸٤۰۲٤۲٤
- اشتراک (شرکت‌ها، سازمان‌ها، ادارات): ۸۸۴۳۴۱۶۹ و ۰۹۱۲-٥٤۹٥٥٤٦

- اینترنت آدم‌ها (٤۳): توانا بود هر که دانا بود/۳ مدل DIKW برای فهم بهتر روش‌های تصمیم‌سازی و حل مسئله/٤
- آنالیتیک کلان‌داده‌ها و تعارض نظریه ولایت فقیه با بنیادهای دین /۱۱
- آنالیتیک، کلان‌داده‌ها، و هوش مصنوعی برای پشتیبانی تصمیم/۱۳

فرم اشتراک ریزپردازنده

برای شرکت‌ها، سازمان‌ها، ادارات

□ برای اشتراک، مبلغ ذکرشده را به حساب جاری سیبا شماره ۰۱۰۲۱۷۹۴۰۹۰۰۸ بانک ملی ایران به نام علیرضا محمدی‌فر (شناسه شبا IR86 0170 0000 0010 2179 4090 08) واریز و تصویر فیش را به همراه تصویر فرم پرشده زیر به نشانی rizpardazandeh@gmail.com ایمیل کنید.

□ بهای اشتراک یک‌ساله: یکصد و بیست هزار تومان

○ نام ونام خانوادگی: ○ تلفن:

○ تلفن همراه: ○ نمابر:

○ ایمیل:

○ نشانی:

○ کد پستی:

اطلاعیه مهم

ریزپردازنده

با توجه به شیوع ویروس کرونا، برای رعایت امور بهداشتی، تا اطلاع ثانوی نسخه چاپی برای مشترکان ارسال نخواهد گردید. نسخه PDF این شماره را به رایگان می‌توانید از وبگاه *ماهنامه ریزپردازنده* (http://www.rizpardazandeh.com) یا کانال rizpardazandeh@ در تلگرام دریافت کنید.

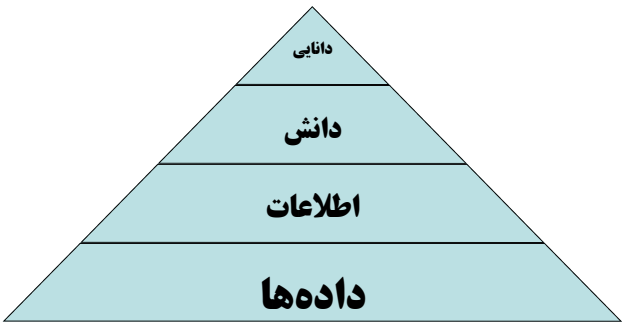
اینترنت آدم‌ها

آنالیتیک کلان‌داده‌ها و تعارض نظریه ولایت فقیه با بنیادهای دین

□ **علیرضا محمدی‌فر**

حل مسئله[1] و **تصمیم‌سازی**[2] کارآمد در نظام‌های سیاسی از گذشته‌های دور **دانش‌محور** و **دانایی‌محور** بوده است، بدین معنی که حکمرانان به عنوان مدیران کشورها از منابع **داده‌ها**، **اطلاعات**، و **دانش** برای تصمیم‌سازی و حل مسئله بهره می‌گرفته‌اند. به بیان دیگر، آنچه امروز به عنوان **هرم داده‌ها-اطلاعات-دانش-دانایی**[3] (DIKW) شناخته می‌شود که در آن از پردازش **داده‌ها** اطلاعات و دانش به دست می‌آید از گذشته‌های دور به عنوان یک مدل کارآمد برای تصمیم‌سازی و حل مسئله بهره گرفته شده است. جاده‌های باستانی مانند جاده ابریشم، قنات‌ها و سیستم‌های آب‌رسانی باستانی، بناهای بزرگ مانند **زیگورات چغازنبیل** همگی نشان می‌دهند که **دانش** از گذشته‌های دور نقشی اساسی در پیشرفت تمدن‌ها و تصمیم‌سازی‌ها داشته است. هرچند، نظر به این که انسان‌ها قرن‌های متمادی برای پیش‌بینی آینده ابزارهای علمی مورد نیاز را نداشتند برای این منظور گاه به جادوگران و پیش‌گویان مراجعه می‌کردند. اما همان‌گونه که در شماره ۲۷۹ این ماهنامه آمد از قرن شانزدهم به بعد که نظریه‌های آمار و احتمالات و **مدیریت ریسک** شکل می‌گیرد تصمیم‌سازی‌ها برای **آینده** بسیار کارآمدتر می‌شود. از دهه ۱۹۴۰ با ورود کامپیوترهای الکترونیک امکانات آمارگیری‌ها و نمونه‌گیری‌های گسترده‌تر و محاسبات پیچیده‌تر و کارآمدتر برای مدیریت ریسک فراهم می‌شود، اما از اوایل قرن بیست‌ویکم فناوری‌های مختلفی چون اینترنت، گوشی‌های هوشمند، **اینترنت چیزها**[4]، **اینترنت آدم‌ها**[5]، و مانند آن تحولی جدید را پدید می‌آورند و عصر **آنالیتیک کلان‌داده‌ها**[6] (**تحلیل کلان‌داده‌ها**) آغاز می‌شود. **کلان‌داده‌ها** این پتانسیل را فراهم می‌کند که سلامت بهتر، آموزش بهتر، رفاه بیشتر، و از همه مهم‌تر تصمیم‌سازی بهتر داشته باشیم. فناوری‌های آنالیتیک کلان‌داده‌ها و حل مسئله، **پشتیبان تصمیم‌سازی**[7]، و **تصمیم‌سازی خودکار در زمان واقعی**[8] (بلادرنگ) نشان داده‌اند که این پتانسیل را دارند که ریسک‌ها را به حداقل برسانند و حتی به صفر نزدیک کنند و افزایشی چشمگیر در **بهره‌وری** پدید بیاورند. **سیستم ترابری هوشمند** بر بنیاد **خودکاری تصمیم‌سازی** و خودران‌هایی که احتمال تصادفات را به صفر می‌رسانند و با بهره‌گیری بهینه از جاده‌ها و خیابان‌های هوشمند بهره‌وری را افزایش می‌دهند یک نمونه از پتانسیل‌های **خودکارسازی تصمیم** است، که توضیحات آن در ادامه مقاله در این شماره آمده است.

به بیان ساده، مدیریت مدرن امروز با تصمیم‌سازی روشمند بر بنیاد رهیافت‌های علمی و بر بنیاد آنالیتیک کلان‌داده‌ها در پی به‌حداقل‌رساندن ریسک‌ها و به‌حداکثررساندن بهره‌وری است. گذشته از اینها، همچنان که در شماره ۲۶۹ آمد فناوری‌های جدید این توان را به مردم جهان می‌دهند که با بهره‌گیری از **همکاری جمعی** و **هوشمندی جمعی**[9] و **سیستم‌های پشتیبانی تصمیم گروهی**[10] به ساخت ابزارهای تولید روی بیاورند و با حذف سود در تولید و خدمات _ که به معنی حذف یا کاهش چشمگیر ریسک‌های مختلف اقتصادی نیز هست _ به نابرابری‌های اقتصادی در جهان خاتمه بدهند.



یک مشکل بزرگ نظام حکمرانی در کشور ما برای **گذار به شهر هوشمند با امکانات تصمیم‌سازی خودکار و آنالیتیک کلان‌داده‌ها** آن است که تصمیم‌سازی کلان در آن مطابق قانون **شخص‌محور** است و نه **دانش‌محور**. بدین معنی که حکمران در کشور ما به جای آن که به عنوان مدیری که به منبع دانش، یعنی **مجموع دانش و دانایی بشری** _ که شامل مجموعهٔ علوم

[1] problem solving
[2] decision making
[3] Data-Information-Knowledge-Wisdom pyramid
[4] Internet of Things
[5] Internet of Humans
[6] big data analytics
[7] decision support
[8] real time
[9] collective intelligence
[10] group decision support systems

دینی و دانش همه فقها نیز می‌شود _ نگاه کند و حل مسئله و تصمیم‌سازی را بر اساس آن انجام دهد مطابق قانون باید به دانش فقهی خود مراجعه کند و با این فیلتر به **بخشی** از دانش بشری و نه کل آن بنگرد و بر این اساس تصمیم‌سازی کند یا مسائل کلان را حل کند. این که می‌گوییم بخشی از دانش بشری مورد نگرش قرار می‌گیرد و نه کل آن، از آن روست که بخش‌هایی از داده‌ها، اطلاعات، و دانش به دلیل فیلترینگ ناشی از اهداف **بهشت‌بَری** از **هرم DIKW** خارج می‌شوند.

توضیح: بهشت‌بَران یا **دوزخ‌دورسازان** هدف یا تکلیف حکومت را دورساختن مردم از دوزخ آسمانی و بردن مردم به بهشت آسمانی با یک **سبک زندگی** ویژه، یعنی از طریق اجرای احکام دینی یا **ارزش‌های** ایدئولوژیک می‌پندارند. از همین روی، بر این باورند که برای این که مردم از دوزخ دور شوند پاره‌ای از اطلاعات که ممکن است آنان را گمراه کند باید فیلتر شود. **(برای اطلاعات بیشتر مقاله «معماری شهر هوشمند آینده با دو دیدگاه بهشت‌سازان و بهشت‌بَران» در شماره ۲۷۰ ماهنامه ریزپردازنده را بخوانید.)**□

هرچه داده‌ها بیشتر باشد اطلاعات و دانش بیشتر می‌شود و تصمیم‌سازی کم‌ریسک‌تر. از همین روی، انسان از گذشته‌های دور با اختراع وسایلی که بتوانند به عنوان **حافظه جانبی** او عمل کنند تلاش فراوانی کرده است که حجم داده‌ها و دانش خود را افزایش بدهد. (اطلاعات بیشتر در شماره ۲۷۸ این ماهنامه که به موضوع **ذخیره و بازیابی اطلاعات** اختصاص دارد.)

هنگامی که حل مسائل و تصمیم‌سازی کلان به طور قانونی به دانش فقهی فقط یک فرد پیوند زده می‌شود به معنی بالارفتن ریسک تصمیمات خواهد بود. هرچه ریسک بالاتر باشد کارآمدی و بهره‌وری پایین‌تر خواهد بود. به بیان دیگر، اساس این نظام بر بنیاد دانش و یافته‌های بشر و مدیریت مدرن امروزی برای یک نظام کارآمد نیست، بلکه بر بنیاد یک تفسیر ویژه از متون مقدس شکل گرفته است و به قانون تبدیل شده است. چنین نظامی نه‌تنها ناکارآمدی در اداره کشور را قانونی کرده است، بلکه در یک ادعای مهم خود _ که بردن مردم به بهشت آسمانی است_ نیز ناکارآمد است، زیرا روش تصمیم‌سازی کلان آن با بنیادهای دین در تعارض است؛ چرا؟

پاسخ را می‌توانیم در **عدالت** سریع‌ترین حسابگر جهان بیابیم. **عدالت** خداوند بزرگ در **همه‌داده‌ها**بودن (**All Data**) و **همه‌دانابودن** اوست، بدین معنی که خداوند بزرگ در محاسبات خود **همهٔ داده‌ها** را در نظر می‌گیرد و نه بخشی از داده‌ها را. او بزرگ‌ترین **داده‌دان** (**data scientist**) کائنات است. آنالیز او از **همه‌داده‌ها** کاملاً دقیق است، **عدل** او در همین دقت نهفته است.

توضیح: وسایل موبایل، شبکه‌های اجتماعی، و کامپیوترهای متصل به وب و حس‌گرهای اینترنت چیزها و اینترنت آدم‌ها که در سال‌های اخیر برخط شده‌اند، در مجموع، نقشی عظیم در انفجار داده‌ها داشته‌اند. کلان‌داده‌ها (Big Data) به مجموعه‌های داده‌ای بسیار بزرگ (شامل داده‌های ساخت‌مند، ناساخت‌مند، و نیمه‌ساخت‌مند) اشاره دارد که این منابع تولید می‌کنند، و کسب‌وکارها، ادارات دولتی، و سایرین مایلند آنها را اخذ و استفاده کنند.

اما واضح است که کلان‌داده‌ها همهٔ داده‌ها نیست. خداوند بزرگ حساب‌گری است که سرعت هیچ حساب‌گر دیگری به پای آن نمی‌رسد و نخواهد رسید. توان چنین حساب‌گری در پردازش همهٔ داده‌های کائنات در زمان واقعی (real time) است. از همین روی، اصطلاح همه‌داده‌ها (All Data) را برای این حساب‌گر بی‌همتا به کار گرفته‌ایم.

خداوند بزرگ در داوری‌هایش با احتمالات کار نمی‌کند. هیچ داده‌ای از نظر او پنهان نمی‌ماند. داوری او بر اساس **همه‌داده‌ها** و **همه‌دانایی** اوست. حساب‌گری بی‌همتاست که همهٔ داده‌ها را در **زمان واقعی** (کاملاً **بلادرنگ؛ real time**) محاسبه می‌کند. **عدل** خداوند بزرگ چنان است که **همه‌داده‌ها** محاسبه شود، یعنی حتی ذره‌ای از کل داده‌ها در محاسبه حذف نمی‌شود. حال چنین خدایی و چنین دانای بزرگی آیا امکان دارد که از انسان بخواهد که حل مسئله و تصمیم‌سازی را به جای آن که **دانش‌محور** و بر بنیاد **دانایی** (**wisdom**) و بر اساس **علم داده‌ها**[11]، یعنی براساس مجموع داده‌ها و دانش بشری باشد _ که کارآمدی خود را در عمل نشان داده است _ به حل مسئله و تصمیم‌سازی **شخص‌محور** و **ذره‌داده‌ها** روی بیاورد؟ یا به جای آن که از بیشترین داده‌هایی که پردازش آنها در توان انسان است مانند **کلان‌داده‌ها** بهره بگیرد و از قدرت‌مندترین برنامه‌های **آنالیتیک** مجهز به هوش مصنوعی که می‌تواند بسازد سود بجوید برای تصمیم‌سازی‌های کلان یک کشور به دانش تخصصی و داده‌های محدود یک نفر متوقف شود؟ (در اینجا از اصطلاح **ذره‌داده‌ها** بهره گرفتیم، چون داده‌های مورد استفاده در این نوع حل مسئله و تصمیم‌سازی قطره‌ای است حتی از آنچه امروز به **خُرده‌داده‌ها** یا **Small Data** مشهور است.) چنین خواسته‌ای از انسان برخلاف **عدل** خداوند بزرگ است.

از سوی دیگر، هنگامی که خواستهٔ پروردگار بزرگ را از انسان تشکیل یک نظام حکمرانی با نحوه تصمیم‌سازی شخص‌محور در نظر بگیریم بدین

[11] data science

معنی است که این داده‌دان بزرگ و این دانای بزرگِ **همه‌داده‌ها** که به ارزش و اهمیت داده‌ها کاملاً واقف است از بندگانش خواسته است که تصمیم‌سازی با ذره‌داده‌ها را بپذیرند. به این ترتیب، خداوند بزرگ را به یک خدای **ذره‌داده‌ها** تبدیل کرده‌ایم، که سبب می‌شود به جای پرستش خدای عادلِ **همه‌داده‌ها**، یک خدای ناموجود و تخیلی **ذره‌داده‌ها** را بپرستیم. به بیان ساده، یا باید بپذیریم که داده‌دان بزرگ و سریع‌ترین حساب‌گر کائنات با اصول علم داده‌ها و دانش و مهندسی کامپیوتر آشنا نیست، و یا بپذیریم که چنان خواسته‌ای با تفسیر نادرست از متون مقدس به دست آمده است.

از همین روی، تردیدی نیست که حل مسئله و تصمیم‌سازی شخص‌محور نمی‌تواند یک خواستهٔ خداوند بزرگ باشد، مگر آن که پیامبری را تعیین کرده باشد که امکان بهره‌گیری از منبع **همه‌داده‌های** خداوندی را داشته باشد. همچنان که در مقاله «**تز چرچ-تورینگ، هوش مصنوعی، و نظریه انتصاب الهی فقیه**» در شماره ۲۷۱ **ماهنامه ریزپردازنده** با اشاره به **تز چرچ-تورینگ** گفتیم، نظام حکمرانی در ایران نمی‌تواند نماینده خداوند بزرگ باشد و به منبع **همه‌داده‌های** او دسترسی داشته باشد.

بی‌گمان، **خداشناسی** یک معیار مهم در ارزیابی درستی یا نادرستی تفسیرها از متون مقدس است. اگر به خدای همه‌داده‌ها و همه‌دانای **عادل** باور داشته باشیم و اگر به خدایی که سریع‌ترین حساب‌گر و دقیق‌ترین **داده‌دان** کائنات است باور داشته باشیم، نمی‌توانیم تفسیری از متون مقدس را بپذیریم که **عدل** و **قدرت رایانش** و **دانسته‌های رایانشی** او را درک نمی‌کند و در نظر نمی‌گیرد و از ما می‌خواهد که بر خلاف **دانایی** عمل کنیم و در **ذره‌داده‌ها** و دانش یک نفر متوقف شویم. به عنوان داده‌دان برای این که بتوانیم تصمیم‌سازی‌های کارآمد را بیابیم باید به دقت به مهندسی کائنات نگاه کنیم و مهندس آن را و داده‌دان آن را بشناسیم. داده‌دان بزرگ کائنات که علم داده‌ها را دقیق‌تر از هر کسی می‌شناسد و به ارزش داده‌ها برای تصمیم‌سازی واقف است، مگر می‌شود که از بندگانش بخواهد که با داده‌های محدود داوری کنند؟

از اینها گذشته، هنگامی که اشخاص برای حل مسائل یا تصمیم‌سازی صرفاً به دانش خودشان اتکا می‌کنند بیشتر در معرض خطاهای شناختی‌ای قرار می‌گیرند که در پژوهش‌های **کانمن**[12] و **تورسکی**[13] آمده است، که گاه می‌تواند فاجعه‌بار باشد، به ویژه اگر تصمیم اتخاذشده برای یک کشور باشد، در حالی که مدیرانی که به منبع دانش و دانایی بشری مراجعه می‌کنند به‌ندرت در معرض چنین خطاهایی قرار می‌گیرند. (برای اطلاعات بیشتر به مقاله صفحه ۱۵ **ماهنامه ریزپردازنده** شماره ۲۷۹ مراجعه کنید.)

همچنان که در پاره‌ای از مقالات نظری این ماهنامه در دو سال گذشته گاه به طور ضمنی و گاه با صراحت گفته‌ایم با دو روش می‌توانیم حکمرانی در کشور را از ناکارآمدی‌ها خارج کنیم، یکی مشروطه‌سازی و دیگری برپاسازی شهرها یا مناطق هوشمند با سبک زندگی ویژه (برای اطلاعات بیشتر، مقاله «**خودتان با همکاری باز شهر هوشمند خودتان را بسازید**» در صفحه ۲۰ شماره ۲۷۰ را بخوانید).□

در دهه ۱۹۸۰ یکی از ذخیره‌گرهای اکسترنال پرطرفدار دیسکت‌های 3½ اینچی بودند. تعداد ۱۵۰ دیسکت داخل کیسه ظرفیتی معادل ۲۱۶ مگابایت دارد. اگر ۱۵۰ کیسه از این دیسکت‌ها درست کنیم که در مجموع وزنی حدود ۴۰۰ کیلوگرم دارند ظرفیت کل آن معادل ظرفیت مایکرواِس‌دی ۳۲ گیگابایتی نشان داده‌شده در تصویر است. یک مایکرواِس‌دی ۱۲۸ گیگابایتی معمول این روزها معادل ۶۰۰ کیسه دیسکت نشان‌داده‌شده به وزن 1.5 تن است؛ امروزه، یک مایکرواِس‌دی ۳۲ گیگابایتی به قیمت کمتر از ۲۰ دلار به فروش می‌رسد. یک گوشی متوسط امروزی ۱۲۸ گیگابایت حافظه دارد.

چنین پیشرفتی حدود ۳۰ سال زمان برده است، کافی است پیشرفت‌های فناوری‌های ذخیره داده‌ها و رایانش ۳۰ سال آینده را تصور کنید. آیا عدالت پروردگار همه‌داده‌ها و همه‌دانایی که به خوبی با علم رایانش آشناست و خالق کامپیوتر عظیم و حیرت‌انگیز کائنات است و عدل او در همه‌داده‌ها و همه‌دانایی اوست اجازه می‌دهد که یک کشور تصمیم‌سازی‌های کلانش را به دانش فقهی یک نفر متوقف کند که هر خطایی در تصمیم‌سازی‌هایش می‌تواند هزینه‌های فوق‌العاده سنگینی برای کشور به بار بیاورد؟ چنین شیوه‌ای از حکمرانی آن هم در زمانی که کلان‌داده‌ها و آنالیتیک کلان‌داده‌ها پتانسیل فراوانی را برای کاستن از ریسک‌ها نوید می‌دهند ناکارآمدی‌ها را روز به روز افزون‌تر می‌کند.□

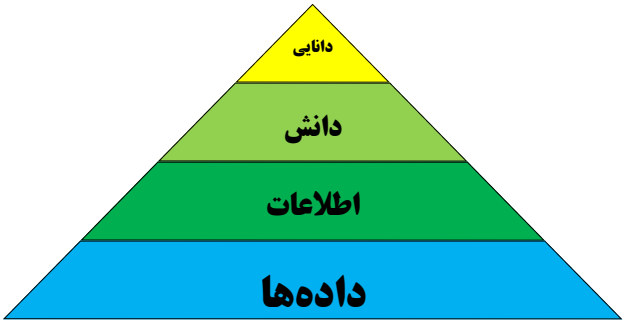
[12] Daniel Kahneman
[13] Amos Tversky

اینترنت آدم‌ها

توانا بود هر که دانا بود

ز دانش دل پیر برنا بود

می‌دانیم که ارزش داده‌ها کم نیست، بسیاری از صاحب‌نظران از عباراتی مانند عبارات زیر برای ارزش داده‌ها بهره می‌گیرند: داده طلاست، داده نفت است، یا داده جواهر است. اهمیت کشف و اختراع روش‌های مختلف ثبت و ذخیره داده‌ها بر سنگ یا چوب یا استخوان در حدود ۴۰۰۰۰ سال پیش به عنوان یک **حافظه اکسترنال** کمتر از اهمیت کشف آتش توسط انسان نبوده است. انسان در طول تاریخ با انواع حافظه اکسترنال، از سنگ، خشت، کاغذ، تا نوار مغناطیسی، دیسک سخت، و انواع دیگر حافظه‌ها توانست انبوهی از **داده‌ها** (Data)، **اطلاعات** (Information)، و **دانش** (Knowledge) را ذخیره‌سازی کند و با **بازیابی** و **پردازش** آنها بهره‌های بسیار فراوانی از آنها بگیرد و **دانایی** (Wisdom) را به بنیاد **حل مسئله**[1] و **تصمیم‌سازی‌های**[2] کارآمد و **مدیریت ریسک** تبدیل کند.



برای این که این بنیادهای **تصمیم‌سازی‌ها** را بهتر بشناسیم از هرمی بهره می‌گیریم که در دروس پایه رشته کامپیوتر به **هرم DIKW** یا **سلسله‌مراتب DIKW** شهرت پیدا کرده است. همچنان که پیداست DIKW سرواژه چهار واژه Data (**داده‌ها**)، Information (**اطلاعات**) یا **آگاهی**)، Knowledge (**دانش** یا **دانسته‌ها**)، و Wisdom (**دانایی**) است. گفته می‌شود این واژه‌ها از یک شعر مشهور **تی. اس. الیوت** (1888-1965)، شاعر نام‌دار آمریکایی-بریتانیایی گرفته شده است که در شعری با عنوان «**صخره**»[3] سروده است:

All our Knowledge brings us nearer to our ignorance.

All our ignorance brings us nearer to death.

But nearness to death no nearer to God.

Where is the Life we have lost in living?

Where is the wisdom we have lost in knowledge?

Where is the knowledge we have lost in the information?

این که این سلسله‌مراتب را چه کسی نخستین بار در محیط‌های علمی مطرح کرده است در ابهام باقی مانده است، اما پاره‌ای از صاحب‌نظران معتقدند که نخستین‌بار در یک محیط علمی از طرف **راسل آکُف**[4] در یک سخنرانی در سال ۱۹۸۹ در کنفرانس «**جامعه جهانی برای پژوهش‌های سیستم‌های عمومی**[5]» عنوان شده است.

هرم DIKW برای آموزش مفاهیم **داده‌ها** و **اطلاعات** در دروس مقدماتی رشته کامپیوتر و همچنین بسیاری از مقالات و کتب آموزشی مورد توجه قرار گرفته است. در این دوره‌های مقدماتی با بهره‌گیری از این هرم تعریف‌های ساده‌ای از داده، اطلاعات، دانش، و دانایی ارائه می‌شود: پردازش و پالایش **داده‌ها** به **اطلاعات** می‌انجامد؛ و یا **داده‌های پردازش‌شده**، بامعنی، قابل استفاده، یا به‌دردبخور **اطلاعات**

[1] problem solving
[2] decision making

[3] The Rock
[4] Russell Ackoff
[5] International Society for General Systems Research

نامیده می‌شود. به بیان دیگر، **دانش** مطالبی از نوع «**می‌دانیم که**» و «**چگونه می‌توانیم**» (تبدیل اطلاعات به دستورالعمل) است، مثلاً می‌دانیم که تهران پایتخت ایران است، یا چگونه می‌توانیم رانندگی کنیم.

میلان زلنی[6]، که پاره‌ای از منابع می‌گویند که چند سال پیش از **آکُف** نخستین بار رشته **DIKW** را به کار گرفت، **دانش** را دستور پختی می‌داند که به شما امکان می‌دهد از **اطلاعات** نان بسازید (یعنی اگر **اطلاعات** مخلوطی از آرد و مایه خمیر و آب باشد، **داده‌ها** مولکول‌ها و اتم‌های این مخلوط هستند).

از سوی دیگر، **کمیته اروپایی استانداردسازی** یک راهنما به نام «**راهنمای تمرین خوب در مدیریت دانش**» تدوین کرده است که در آن درباره **دانش** چنین نظر داده است: «**دانش** ترکیبی از **داده‌ها** و **اطلاعات** است که بر آنها نظرات، مهارت‌ها، و تجارب متخصصان اضافه می‌شود، به گونه‌ای که نتیجه مجموعه‌ای ارزشمند شود که می‌تواند برای تصمیم‌سازی به کار گرفته شود.»[7]

دانش مورد استفاده دانایان که از تعاملات و تجارب به دست می‌آید، برای یادگیری از اشتباهات و اشتباهات دیگران می‌تواند حاوی تجارب اشتباه و اشتباهات نیز باشد، تا در مرحله تصمیم‌سازی به کار گرفته نشوند.

دانایی بهره‌گیری از «**چگونه می‌توانیم**»ها برای رسیدن به اهداف است. **دانایی** به داوری اطلاعات و دانش و انتخاب تصمیم گفته می‌شود. تا یک یا دو دهه پیش **دانایی** مرز متمایزساختن انسان از ماشین بود، اما امروزه وضعیت متفاوت شده است و استثناهایی به وجود آمده است. به عنوان مثال، امروزه برای این که با خودرو از جایی به جای دیگر برویم تصمیم‌گیری درباره روش رسیدن به مقصد را به نرم‌افزارهایی مانند **Waze** می‌سپاریم.

با این همه، تعریف دقیق‌تر این مفاهیم بسیار پیچیده‌تر است و تعریف واحدی برای هر یک از این اصطلاحات وجود ندارد. برای اطلاعات و منابع بیشتر درباره **هرم DIKW** می‌توانید به وبگاه www.isko.org/cyclo/dikw مراجعه کنید.

[6] Milan Zeleny
[7] https://www.isko.org/cyclo/dikw

مدل DIKW برای فهم بهتر روش‌های تصمیم‌سازی و حل مسئله

از همان زمان تشکیل تمدن‌های اولیه انسان به اهمیت **داده‌ها** و **اطلاعات** و **دانش** برای تصمیم‌سازی پی برد و ساخت **مراکز داده‌های** مالی و دانش به تدریج آغاز شد، تا تصمیم‌سازی‌ها **دانش‌محور** باشد و نه **شخص‌محور**. **کتابخانه اسکندریه** یا **کتابخانه دانشگاه جندی‌شاپور** دو نمونه از مراکز داده بزرگ دوران باستان هستند. با آن که اکثر حکومت‌های دوران باستان دارای پادشاهانی مستبد بودند برای این که امور کشور به گونه‌ای کارآمد اداره شود تا بتوان مالیات‌های بیشتری اخذ کرد در تصمیم‌گیری‌ها ناچار بودند **دانش‌محور** عمل کنند. هر چند، برای پیش‌بینی آینده راهی در پیش نداشتند جز مراجعه به پیش‌گویان و جادوگران. یک سند کم‌نظیر از نحوه تصمیم‌سازی در دوران باستان **گلدان داریوش** است که شرح آن در مقاله «**گلدان داریوش و ابزارهای محاسبه**» در شماره ۲۷۱ **ماهنامه ریزپردازنده** صفحه ۱۱ آمده است.

اما همچنان که در مقاله «**علیه خدایان: داستان شگفت‌انگیز ریسک**» در شماره ۲۷۹ به تفصیلِ بیشتر گفتیم با ترجمه کتاب‌های **خوارزمی** به لاتین و چاپ گسترده آنها در قرن پانزدهم در اروپا استفاده از **دستگاه عددنویسی هندی** گسترش یافت و سبب تحول علمی چشمگیری در غرب شد. این عامل و عامل **رنسانس** به پاره‌ای از ریاضی‌دانان بزرگ، مانند **بلز پاسکال**[8]، **دانیل برنولی**[9]، یا **کارل فردریش گاوس**[10] امکان داد که روی **نظریه‌های احتمالات** و **مدیریت ریسک** مطالعه کنند و برای نخستین بار به کاوش در **آینده** بپردازند، حوزه‌ای که پیش از آن در انحصار کلیسا و پیش‌گویان بود.

در شماره ۲۷۸ نیز به **تاریخ ذخیره و بازیابی داده‌ها** پرداختیم. تاریخ ذخیره و بازیابی را گفتیم تا هزاران سال تلاش خستگی‌ناپذیر انسان را برای گردآوری دانش و تصمیم‌سازی کارآمد بر اساس آن بازگو کنیم.

تا پیش از پیشرفت‌های خیره‌کننده سه دهه اخیر در ذخیره‌گرها و حافظه‌های داده‌ها و ریزپردازنده‌ها، گردآوری داده‌ها و ذخیره و بازیابی و پردازش آنها بسیار پرهزینه بود، از همین روی، مدیریت ریسک به **نمونه‌گیری** از داده‌ها، یعنی حجمی کوچک از داده‌ها اتکا داشت، و کنترل ریسک در حد آرمانی نبود.

[8] Blaise Pascal
[9] Daniel Bernoulli
[10] Carolus Fridericus Gauss

کلان‌داده‌ها وارد می‌شود

روند فناوری‌های تصمیم‌سازی و حل مسئله

جهان به گونه‌ای باورنکردنی در حال تغییر است. فناوری‌هایی که در پنجاه‌سال گذشته وارد زندگی انسان شده‌اند و زندگی انسان را دگرگون کرده‌اند کم نیستند، مانند فکس، اینترنت، گوشی هوشمند، کامپیوتر شخصی، **کلان‌داده‌ها**[11]، **اینترنت چیزها**[12]، **اینترنت آدم‌ها**[13]، و مانند آن. پیشتر در شماره ۲۷۷ گفتیم که این تغییراتِ دگرگون‌کننده سبب گذار از **عصر آدم** به **عصر سایبورگ**[14] شده است.

لازم است دولت‌ها، سازمان‌ها و شرکت‌ها برای چنین تغییراتی چابک عمل کنند و سریعاً و پیوسته تصمیمات استراتژیک، تاکتیکی، و عملیاتی مقتضی با آنها را اتخاذ کنند. اتخاذ چنین تصمیماتی به **داده‌ها**، **اطلاعات**، و **دانش**، و **دانایی** (DIKW) بسیار زیادی نیاز دارد.

داده‌دان (data scientist) کیست؟

داده‌دان تعریف دقیقی ندارد. داده‌دان به متخصص رشته **علم داده‌ها** (**data science**) گفته می‌شود. به بیان ساده، داده‌دان یک آنالیست خبره یا یک مهندس است که از ابزارهای کشف **کلان‌داده‌ها** برای یافتن **بینش‌های** (**insight**) جدید در داده‌ها با استفاده از تکنیک‌هایی که ماهیتاً آماری یا علمی هستند بهره می‌گیرد. به بیان ساده، اگر داده‌ها نفت باشد، داده‌دان کاشف نفت است. در آینده تیم‌های داده‌دان به جای یک یا چند داده‌دان کار خواهند کرد.□

در شماره گذشته، یعنی شماره ۲۷۹، **مدیریت ریسک** در **عصر خُرد-داده‌ها**[15] را بررسی کردیم. هیچ‌کدام از ریاضی‌دانان گذشته که در نظریه‌های آمار و احتمالات کار کردند و در آن مقاله معرفی کردیم از **پاسکال**، **دانیل برنولی**، **لاپلاس**، و **گالتون** گرفته تا **پوانکاره** از **کلان‌داده‌ها**، هوش مصنوعی، و برنامه‌های قدرتمند **آنالیتیک** خبر نداشتند.

کلان‌داده‌ها (Big Data) چیست؟

مقدار داده‌ها به این دلیل بسیار سریع‌تر از گذشته زیاد می‌شود که منابع مختلف تولید داده‌ها بسیار زیاد شده است. وسایل موبایل، شبکه‌های اجتماعی، و کامپیوترهای متصل به وب و حس‌گرهایی که در سال‌های اخیر **برخط** شده‌اند، که به **اینترنت چیزها** مشهور شده‌اند، نمونه‌هایی از این منابع تولید داده‌ها هستند. همه اینها در مجموع یک نقش عظیم در انفجار داده‌ها دارند. «**کلان‌داده‌ها**» به مجموعه‌های داده‌ای بسیار بزرگ اشاره دارد که این منابع تولید می‌کنند، که کسب‌وکارها، ادارات دولتی، و سایرین مایلند آنها را اخذ و استفاده کنند.

برای خلاصه‌کردن خصوصیات مهم کلان‌داده‌ها از اصطلاحات «**تراکم** یا volume» ، «**تنوع** یا variety»، «**تندی** یا velocity» (سه «**ت**» در فارسی و سه **v** در انگلیسی) استفاده می‌شود. «**تراکم**» به مقادیر بسیار زیاد داده‌ها اشاره دارد، «**تنوع**» به انواع گوناگون منبع‌های داده‌ها دلالت دارد، و «**تندی**» سرعت بسیار بالایی را توصیف می‌کند که داده‌ها ساخته می‌شوند.

ابزارهای جدید نسبت به گذشته مقادیر بسیار بیشتری داده در فرمت‌های متنوع تولید می‌کنند. این داده‌ها بیشتر در دو حالت هستند: **ساخت‌مند یا structured** که پردازش و تحلیل کامپیوتری آنها آسان است، مانند داده‌های مالی؛ **ناساخت‌مند یا unstructured** که پردازش آنها دشوارتر است، مانند عکس یا صدا.□

یکی از مهم‌ترین تغییراتی که در حال وقوع است تغییراتی است که **کلان‌داده‌ها** و **تُندداده‌ها**[16] در نحوه **تصمیم‌سازی** و **فناوری‌های تصمیم‌سازی** و **مدیریت ریسک** می‌توانند بر جای بگذارند. هنگامی می‌توانید درست تصمیم بگیرید که مجموعه‌ای کافی از داده‌ها داشته باشید، پس، **کلان‌داده‌ها** و **تندداده‌ها** با حجم بسیار بزرگی از داده‌ها می‌توانند کمک‌تان کنند و ریسک را به حداقل برسانند. همچنان که آینده **سیستم ترابری هوشمند** را به عنوان یک نمونه در بخش بعدی

[11] big data
[12] Internet of Things
[13] Internet of Humans
[14] cyborg
[15] small data
[16] fast data

توضیح داده‌ایم، فناوری‌های اطلاعات این پتانسیل را دارند که در سیستم‌های مختلف **ریسک را به سمت صفر برسانند**. فناوری‌های **هوش مصنوعی**، و **آنالیتیک** نوید می‌دهند که ریسک‌های بیشتری را می‌توانیم تحت کنترل در بیاوریم.

اختلاف بین «کلان‌داده‌ها» (Big Data) و «کوچک‌داده‌ها» (Small Data) چیست؟

کوچک‌داده‌ها فقط شامل داده‌های «**ساخت‌مند**» یا **structured** مانند صورت‌حساب‌های مشتری، سوابق فروش، اطلاعات پرداخت کارکنان، و غیره است که به طور مرسوم در **صفحه‌گسترده‌ها** (**spreadsheet**) و **پایگاه‌های داده** (**database**) مدیریت می‌شده‌اند. با آن که کلان‌داده‌ها می‌تواند شامل داده‌های «ساخت‌مند» هم باشد، به اندازه بسیار زیادی شامل داده‌های «**ناساخت‌مند**» یا **unstructured** مانند عکس، ویدئو، داده‌های روزآمدشده شبکه‌های اجتماعی، داده‌های وبلاگ‌ها، واقعه‌نگارهای حس‌گرهای راه دور، و سایر انواع گسترده اطلاعات است که برای پردازش، گروه‌بندی، و تحلیل با ابزارهای مرسوم دشوارترند.□

قیمت وسایل ذخیره داده‌ها همچنان رو به کاهش است، قیمت ابزارهای **اینترنت چیزها** و **اینترنت آدم‌ها** رو به کاهش است، یا قیمت کامپیوتر و پردازنده‌هایی که می‌توانند به هر چیزی و هر حس‌گری قدرت رایانش بدهند رو به کاهش است. از همین روی، با هزینه‌ای نسبتاً پایین می‌توانیم مقادیر بسیار عظیمی از داده‌ها را گردآوری کنیم و آنها را در عملیات تصمیم‌سازی به کار بگیریم.

پیش‌بینی‌هایی که از **کلان‌داده‌ها** به دست می‌آوریم نیز زمانی خوب هستند که فرصت کافی برای اقدامات مربوط به پیش‌بینی‌ها و تصمیم‌های ناشی از آنها را داشته باشیم. افزون بر این، داشتن داده‌ها کافی نیست. باید داده‌های باربط و به‌دردبخورِ کافی داشته باشیم، زمان کافی برای آنالیز داده‌ها داشته باشیم، و فرصتِ کافی برای اقدام. اگر رویدادی به زودی اتفاق می‌افتد به سرعت باید بتوانیم پیش‌بینی و اقدام کنیم.

تُندداده‌ها (fast data)

هر بار که روی یک لینک کلیک می‌کنید، یک پرسش می‌پرسید، یک ویدئوی برخط را تماشا می‌کنید، از یک فروشگاه اینترنتی خرید می‌کنید؛ یا هر ثانیه که خودروهای ما، کارخانه‌ها، و هواپیماها و سایر سیستم‌ها کار می‌کنند، جریان‌های داده‌ای انبوهی پدید می‌آید، که در هر ثانیه گیگابایت‌ها داده را جریان می‌دهند. این داده‌های در جریان که باید به صورت همزمان و در **زمان واقعی** پردازش شوند **تندداده** یا **داده‌های فوری** نامیده می‌شوند.

داده‌ها پیش از آن که بزرگ و کلان باشند تند و سریع هستند. ما در حال گذار به عصر جدیدی هستیم که در آن داده‌های در جریان یا داده‌های در حرکت گاهی کل داده‌های تاریخ را تحت‌الشعاع خود قرار می‌دهند. و به دلیل این گذار است که شرکت‌ها به دنبال روش‌هایی برای به‌کارگرفتن برنامه‌های تحلیل‌گر سریع‌تر برای کلان‌داده‌های کنونی خود هستند تا بتوانند از داده‌های در جریان در **زمان واقعی** شناخت پیدا کنند.

با وجود این، داده‌های در جریان فقط بخشی از تعریف کلی **تندداده‌ها** هستند. در حقیقت، علاوه بر سرعتی که داده‌ها به سیستم شما می‌رسند، «تند» در تندداده‌ها می‌تواند به تازگی و روزآمدبودن داده‌ها به هنگام اجرای برنامه تحلیل‌گر نیز اشاره کند، یا حتی به سرعت نگاه‌کردن شما به تحلیل‌های برنامه‌های آنالیتیک نیز اشاره دارد. فقط سرعت حرکت داده‌ها و یا سرعتی که شما آنها را به دست می‌آورید اهمیت ندارد، زمانی که طول می‌کشد که شما آن داده‌ها را آنالیز و تحلیل کنید و نتایج را به کار بگیرید نیز مهم است.□

اگر آنالیتیک خودکار باشد در «**زمان تصمیم**» اقدامات را به اجرا در می‌آورد. به این فناوری **خودکاری تصمیم**[17] گفته می‌شود.

[17] decision automation

زمان تصمیم به معنی آن است که آنالیز با سرعتی انجام می‌شود که برای فعال‌سازی اقدامات مربوط به تصمیم لازم است. در پاره‌ای از موارد، زمان تصمیم **زمان واقعی**[18] (یا خیلی نزدیک به آن) است. در موارد دیگر، زمان تصمیم ممکن است شامل چند دقیقه، یا چند ساعت یا چند روز تأخیر باشد. دانستن زمان تصمیم، برای موفقیت بسیار ضروری است، زیرا عملیات آنالیتیک (تحلیل) باید آماده باشد و در همان پنجره‌ای اجرا شود که برای تصمیم به کار گرفته می‌شود.

تفاوت «کلان‌داده‌ها» و «تحلیل کلان‌داده‌ها»

با آن که «**کلان‌داده‌ها**» و «**تحلیل کلان‌داده‌ها**» (**big data analytics**) اغلب به عنوان موضوع یک چیز در نظر گرفته می‌شوند، آنها چیزهای متفاوتی هستند. همچنان که پیشتر ذکر شد، **کلان‌داده‌ها** به مجموعه‌های انبوه داده‌هایی که به طور جهانی تولید می‌شود اشاره دارد. **تحلیل** یا **آنالیتیک کلان‌داده‌ها** اساساً پردازش و تحلیل انواع مختلف داده‌های داخل یک مجموعه داده‌ها به منظور آشکارکردن **الگوهای** ویژه مشتریان _ که پیشتر دیده نشده است _ است، تا بدین وسیله امکان بازاریابی بهتر فراهم شود.□

(عکس از مژده حمزه تبریزی؛ موزه BMW، مونیخ)

امروزه برای بخش بسیار بزرگی از **داده‌ها** که به طور جریان‌دار و با سرعت تولید می‌شوند و به **تندداده‌ها** (**fast data**) مشهورند این ضرورت وجود دارد که پردازش آنها سریع و حتی در **زمان واقعی** (**بلادرنگ**) یا **نزدیک به زمان واقعی** صورت بپذیرد، از همین روی، تقاضا برای تجهیزات قدرتمند کامپیوتری و نرم‌افزارهای **تحلیل** یا **آنالیتیک** قدرتمندی که به **سیستم‌های پشتیبان تصمیم**[19] شهرت یافته‌اند همواره رو به افزایش است. این تقاضا به پیشرفت‌ها در فناوری‌های آنالیتیک و حرکت به سوی ساخت **نرم‌افزارهای خودکارسازی تصمیم** پیشرفته‌ای می‌انجامد که می‌توانند با بررسی نحوه اجرای تصمیم، ریسک را در بسیاری از سیستم‌ها به صفر برسانند.

در شهر هوشمند **تُند-داده‌ها** در جریان است، که حجم بزرگی از داده‌ها را تولید می‌کنند. در شهرهای هوشمند تصمیم‌سازی **دانش‌محور** بر بنیاد **آنالیتیک کلان‌داده‌ها** و **خودکاری تصمیم‌سازی** آغاز شده است. دوره **خُرد-داده‌ها** و تصمیم‌سازی بر بنیاد خُرد-داده‌ها و دوره تصمیم‌سازی شخص‌محور به پایان خود نزدیک می‌شود.

کلان‌داده‌ها و آنالیتیک نوید داده‌اند که در دهه آینده کارکرد هر صنعت و کسب‌وکاری را تغییر بدهند. هر سازمانی که زودتر با کلان‌داده‌ها آغاز کند در رقابت‌ها امتیاز بیشتری کسب می‌کند. همان‌گونه که در دوره «خُرد-داده‌ها» شرکت‌هایی که آنالیتیک را به کار گرفتند از سایر رقبا جلو زدند، زمانه با سازمان‌هایی است که از فرصت کلان‌داده‌ها بهره بگیرند.

شهر هوشمند آینده با تصمیم‌سازی خودکار این توان را دارد که ریسک‌ها را به صفر نزدیک کند. از سوی دیگر، با آن که ریسک محرک اقتصادی در سرمایه‌داری است و ریسک‌کردن برای رشد اقتصاد بازار آزاد اساسی است و بدون آن رشد اقتصادی پدید نمی‌آید، **کمونیسم شهر هوشمند** ریسک در اقتصاد را با حذف سود و با همکاری شهروندان به صفر نزدیک خواهد کرد (اطلاعات بیشتر در **ماهنامه ریزپردازنده** شماره ۲۶۹).

خودروسازان جهان تلاش گسترده‌ای برای ساخت خودران‌های آینده انجام می‌دهند. درک صحنه برای خودران‌های آینده برای پیش‌گیری از هر نوع تصادف. (عکس از مژده حمزه تبریزی؛ موزه BMW، مونیخ)

[18] real time
[19] decision support system

تلاش‌ها برای حذف ریسک از سیستم ترابری هوشمند خودران

کمتر راننده‌ای هست که تجربه تصادف با یک خودروی دیگر را نداشته باشد. اصلاً اگر احتمال تصادف وجود نداشت کسی خودروی خود را بیمه شخص ثالث نمی‌کرد. تصمیم به رانندگی به معنی قبول یک ریسک است. حتی عبور پیاده از این سوی خیابان به آن سوی خیابان با ریسک همراه است.

راه‌داران و پلیس راهنمایی و رانندگی با بهره‌گیری از نظریه‌های آمار و احتمالات انواع تابلوی راهنمایی و سرعت‌گیر را در خیابان‌ها و جاده‌ها نصب می‌کنند تا بتوانند تعداد تصادفات را کمتر کنند. اما بازهم آمار تصادفات و مصدومان ناشی از تصادفات در سراسر جهان کم نیست، به ویژه آن که خطای انسانی همواره وجود دارد و معمولاً بیشترین عامل بروز تصادفات است.

فناوری این توان را دارد که ریسک تصادفات را به صفر برساند. چگونه؟

خودران‌ها

خودران‌ها وارد خیابان‌ها شده‌اند، هر چند، امروزه بیشتر دوره آزمایشی خود را می‌گذرانند. برای حفظ ایمنی خودران و جلوگیری از تصادفات انواع فراوانی از حس‌گرها در خودران نصب می‌شود. برای فاصله‌های طولانی **رادار** را دارید، **دوربین‌ها** را دارید که کاربردشان واضح است؛ و **حس‌گرهای فراصوتی** یا **آلتراسونیک** را دارید که از امواج صوتی برای تشخیص اشیاء نزدیک بهره می‌گیرند. حس‌گرهای فراصوتی در وضعیت‌های سرعت-پایین، مانند پارک‌کردن خودرو سودمند هستند، و می‌توانند اشیاء شفاف نزدیک مانند شیشه را تشخیص بدهند، که ممکن است برای حس‌گرهای تصویری محسوس نباشند. برای همه شرایط ترافیکی می‌توان از **رادار** سود جست، شامل تشخیص نقطه کور، کمک به تغییر مسیر، کمک به ترمزکردن، و **کروزکنترل تطبیقی**[20] (**تثبیت‌کننده هوشمند سرعت خودرو**). **لایدار**[21] را دارید، که با اندازه‌گیری مدت رفت و برگشت یک پالس لیزر _ که با برخورد به موانع باز می‌گردد _ فاصله‌ها را اندازه‌گیری می‌کند. **لایدار** یک سیستم فاصله‌یاب است که شبیه به رادار کار می‌کند، اما از لیزر بهره می‌گیرد، و میلیون‌ها نقطه را در نظر می‌گیرد، شبیه به یک ابرِ نقطه‌ای از لیزرها که تابش و بازتابش می‌کنند. اینها پاره‌ای از حس‌گرهایی هستند که از طریق آنها اطلاعات وارد خودران می‌شود.

این حس‌گرها برای تشخیص اشیاء و برای جلوگیری از تصادف و همچنین برای حصول اطمینان از این که خودران در جاده در جای درست یا در مسیر درست قرار گرفته است، به کار گرفته می‌شوند. مقادیر انبوهی عملیات کامپیوتری برای اطلاعات به‌دست‌آمده از این حس‌گرها و دوربین‌ها انجام می‌گیرد تا سفر خودران بی‌خطر انجام بگیرد. اما اینها تنها چیزهایی نیستند که امنیت خودران‌ها و مسافران آنها را تأمین می‌کنند.

انتقال داده‌ها و پردازش در خودران

با این همه داده‌هایی که در حس‌گرها، دوربین‌ها، و سیستم‌های مختلف متحرک در خودران تولید می‌شود، می‌توان تصور کرد که نوعی انتقال داده‌ها از خودران به یک **مرکز داده**[22] در جایی دیگر وجود داشته باشد. با وجود این، به جز چند مورد خاص، اکثر پردازش‌ها در داخل خودران روی می‌دهد. یک یا چند **اَبَرکامپیوتر** برای این محاسبات لازم است، که درون خودران‌ها تعبیه می‌شود. نرم‌افزار **آنالیتیک**[23] (تحلیل) برای تجزیه و تحلیل مقادیر هنگفتی از داده‌ها که لحظه به لحظه به سرعت می‌رسد (**تندداده‌ها**[24]) به یک **اَبَرکامپیوتر** نیاز دارد تا **بلادرنگ** (**زمان واقعی**[25]) تصمیم‌سازی کند، یعنی بداند که کجا ترمز کند و کجا سرعت را افزایش بدهد.

افزون براین، این **اَبَرکامپیوتر** آنچه را از رانندگی انسان‌ها و خودران‌های دیگر با فناوری **یادگیری ماشین**[26] فرا گرفته است در خود دارد و آن را پیوسته در حافظه خود به‌هنگام می‌کند و در تجزیه و تحلیل عملیات کامپیوتری می‌تواند همواره یاد بگیرد که چگونه یک راننده بهتر بشود.

از لحاظ کلی، تمام عملیات و پردازش‌ها تقریباً در داخل خودران انجام می‌گیرد. البته کمی ارتباطات به ویژه برای نقشه لازم است. اگر جاده در دست تعمیر باشد یا تصادفی در جلو رخ داده باشد، سیستم از لحاظ نظری این اطلاعات را دریافت یا ارسال خواهد کرد. اطلاعاتی که آنها گردآوری می‌کنند به **ابر**[27] ارسال می‌شود، نقشه اصلاح می‌شود، و داده‌های نقشه جدید به خودران باز می‌گردد، به گونه‌ای که حالا در حافظه بداند که مثلاً در جایی که پیشتر مسیر باز بوده است حالا بسته شده است.

[20] adaptive cruise control
[21] LIDAR

[22] data center
[23] analytics
[24] fast data
[25] real time
[26] machine learning
[27] cloud

این قدرت عظیم کامپیوتری در حوزه‌های دیگر نیز به خودران کمک می‌کند. می‌توانید از عملکردهای قطعات درونی خودران همواره آگاه باشید، چون بسیاری از قطعات آن توسط یک کامپیوتر مرکزی مدیریت می‌شود. از طریق مانیتور داشبورد می‌توانید بگویید که کدام قطعات درست کار می‌کنند و کدام قطعات معیوب هستند. اما مهم‌تر، این واحد کنترل مرکزی آن‌قدر اطلاعات دارد که همه مسائل را ردیابی کند و از امنیت مسافران محافظت کند.

حذف بقیه ریسک‌ها با فناوری‌های V2V و V2I

به کمک فناوری **V2V**[28] (**ارتباطات خودرو با خودرو**) می‌توان ریسک تصادفات را بازهم کاهش داد. فناوری ارتباط **خودرو با خودرو** (**V2V**) می‌خواهد صحبت‌کردن خودروهای داخل جاده با یکدیگر را ممکن کند. هر خودرویی با این فناوری یک گیرنده و یک فرستنده دارد که اطلاعات را به خودروهای در بُرد ارتباطی خودش ارسال می‌کند. بزرگ‌ترین فایده V2V آن است که طوری طراحی می‌شود که خصوصیات ایمنی خودکار خودروها را بسیار بهتر کند. V2V به خودران‌ها امکان می‌دهد که به طور بی‌سیم پیام‌هایی را با خودران‌های نزدیک رد و بدل کنند، مانند اطلاعات سرعت و مکان، برای کمک به جلوگیری از وقوع تصادفات و دادن هشدارهای به‌موقع به رانندگان.

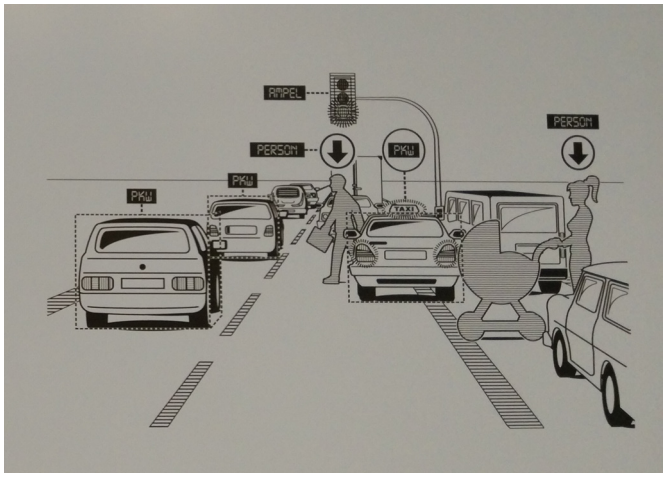
به عنوان مثال، فرض کنیم در یک ترافیک سنگین (که به ترافیک سپر به سپر مشهور است) هستید و خودروهای جلویی پیوسته ترمز می‌کنند. ایست ناگهانی در رانندگی‌های انسان‌ها گاه سبب برخورد زنجیره‌ای از خودروها به یکدیگر می‌شود. حال، فرض کنید که هر خودران پشت سر خودرانی که ناگهانی ترمز می‌کند بلافاصله از ترمزکردن خودروی جلویی آگاه شود، مثلاً از طریق یک هشدار. چند ثانیه وقت اضافی می‌تواند جلوی برخورد دومینویی خودروها را بگیرد. واضح است که کمی وقت هشدار اضافی می‌تواند در انواعی از شرایط رانندگی دیگر مفید باشد، مانند جلوگیری از تصادفات در تقاطع‌ها یا در زمانی که خودروها لاین یا مسیر خود را تغییر می‌دهند. V2V می‌تواند به خودران شما و به خودران‌های دیگر در مورد احتمال برخورد هشدار بدهد.

ارتباطات **V2I**[29] (**ارتباطات وسیله نقلیه با زیرساخت**) نیز یک عامل دیگر برای حذف ریسک تصادف است، که انتقال پیام‌ها بین وسایل نقلیه و چیزهایی مانند علائم راهنمایی و رانندگی، باجه عوارضی، و خط‌کشی عابر پیاده، هشدارهای سیگنال‌های ترافیکی، منطقه مدرسه، و منطقه کار را فعال می‌کند. به خودران‌ها پیش از رسیدن به بخش‌های در دست تعمیر خیابان یا بزرگراه و حضور راهبان‌ها در خیابان یا بزرگراه آگاهی داده می‌شود. حتی ممکن است زمان تغییرکردن چراغ راهنمایی به خودران‌ها اطلاع داده شود تا بی‌جهت سرعت را زیاد نکند. این فناوری به خودران‌ها امکان می‌دهد که به مناطق پرترافیک که ممکن است بر روی یک نقشه GPS مشخص نشده باشد وارد نشوند.

چگونه ریسک خیابان و جاده به صفر نزدیک می‌شود؟

همچنان که گفته شد انواع گوناگونی از **حس‌گرها** در خودران‌ها تعبیه می‌شود، فناوری‌های **V2V** و **V2I**، به کار گرفته می‌شود، و یک **اَبَرکامپیوتر** در داخل خودران‌ها از طریق یک **برنامه آنالیتیک** (تحلیل) که از **هوش مصنوعی** بهره می‌گیرد همه داده‌هایی را که برای ایمنی حرکت خودروها و عابران لازم است و از طریق حس‌گرها با سرعت زیاد تولید می‌شود در **زمان واقعی** (بلادرنگ) پردازش می‌کند و تصمیم‌ها در زمان واقعی به اجرا در می‌آید و در نهایت یک سیستم تا جای ممکن بدون ریسک پدید می‌آید.

نظر به این که این سیستم از **تُنددادده‌ها** (یعنی انبوهی از داده‌هایی که بلاانقطاع از طریق حس‌گرها تولید می‌شود و به پردازش و تصمیم‌سازی در **زمان واقعی** نیاز دارند) و همچنین از **کلان‌داده‌ها** (داده‌های تاریخی ترافیک، نقشه‌های از پیش تهیه‌شده، اطلاعات مسیرهای پیشترپیموده‌شده، نظرات کاربران، و مانند آن) به گونه‌ای بهره می‌گیرد که ریسک تصادفات را به صفر برساند مجموع کل داده‌ها برای سیستم هوشمند ترابری را **همه‌داده‌ها** (**All Data**) نام نهاده‌ایم تا تمایزی بین این نوع از سیستم‌ها با سیستم‌هایی که در آنها ریسک همچنان وجود دارد ایجاد کرده باشیم. به بیان ساده، هنگامی که **همهٔ داده‌های** مورد نیاز برای حذف ریسک در یک سیستم را داشته باشیم برای داده‌های چنین سیستمی می‌توانیم از اصطلاح **همه‌داده‌ها** بهره بگیریم.□



[28] vehicle-to-vehicle communication
[29] vehicle-to-infrastructure

آنالیتیک، کلان‌داده‌ها، و هوش مصنوعی برای پشتیبانی تصمیم

این مقاله چکیده‌ای است از کتاب «آنالیتیک، علم داده‌ها، و هوش مصنوعی»:

Sharda, Ramesh. Delen, Dursun. Turban, Efraim.
***Analytics, Data Science, & Artificial Intelligence: Systems for Decision Support.* 2020, Pearson Education.**

آنالیتیک[1] (**تحلیل**) فناوری پیشتاز این دهه شده است. شرکت‌هایی چون **اُراکل، مایکروسافت، آی‌بی‌ام،** و دیگران ساخت واحدهای سازمانی جدیدی را آغاز کرده‌اند که بر **آنالیتیک** تأکید دارد. این واحدهای سازمانیِ **آنالیتیک** به کسب‌وکارها کمک می‌کنند که در عملیات خود مؤثرتر و کارآمدتر عمل کنند.

تصمیم‌سازان از ابزارهای داده‌ای و کامپیوتری برای تصمیم‌گیری‌های بهتر بهره می‌گیرند. حتی مشتریان از ابزارهای آنالیتیک مستقیم یا غیرمستقیم برای تصمیم‌سازی در مورد اموری مانند خرید، سلامت، و سرگرمی بهره می‌گیرند. حوزه‌های **آنالیتیک کسب‌وکار**[2] (**BA**)، **علم داده‌ها**[3] (**DS**)/ **سیستم‌های پشتیبان تصمیم**[4] (**DSS**)/ **هوشمندی کسب‌وکار**[5] (**BI**) به سرعت در حال تکامل هستند.

فناوری‌های **هوش مصنوعی** با فراهم‌کردن امکان اتخاذ تصمیم‌های خودمختارانه و پشتیبانی از مراحل مختلف عملیات تصمیم‌سازی اثر بسیار بزرگی بر رویه‌های تصمیم‌سازی دارند. هوش مصنوعی و آنالیتیک برای فراهم‌ساختن امکانات تصمیم‌سازی بهتر از یک‌دیگر پشتیبانی می‌کنند.

هدف این مقاله معرفی فناوری‌هایی است که در مجموع **آنالیتیک** (**تحلیل**) نامیده می‌شوند، هر چند ممکن است برای آن از نام‌هایی دیگر مانند **سیستم‌های پشتیبان تصمیم**، **سیستم‌های اطلاعات اجرایی**، و **هوشمندی کسب‌وکار** نیز استفاده شده باشد. در این مقاله از همه این نام‌ها بهره گرفته‌ایم.

[1] analytics
[2] business analytics
[3] data science
[4] decision support system
[5] business intelligence

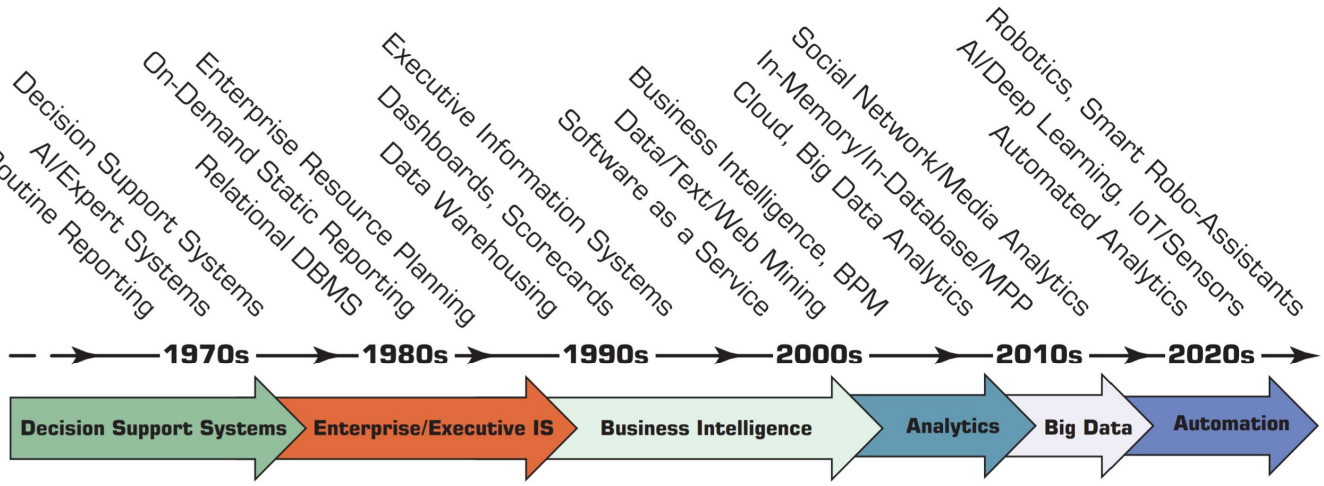
عمل تصمیم‌سازی

مدیران سال‌ها **تصمیم‌سازی** را اساساً یک **هنر** در نظر می‌گرفتند _ استعدادی که از طریق تجربه‌هایی طولانی و با فراست به دست می‌آمد (مثلاً یادگیری با آزمون‌وخطا). نظر به این که انواعی از سبک‌های فردی می‌توانست در یافتن و حل‌کردن موفقیت‌آمیز انواع مسائل مدیریتی مشابه به کار گرفته شود **مدیریت** را یک **هنر** در نظر می‌گرفتند. این سبک‌ها به جای روش‌های کمّی بر بنیاد رهیافت‌های علمی، اغلب بر بنیاد خلاقیت، قضاوت، فراست، و تجربه بودند. با این همه، امروزه بر تصمیم‌سازی روش‌مند و بر بنیاد آنالیتیک تأکید می‌شود.

مدیران معمولاً با مراحل زیر تصمیم‌سازی می‌کنند:

۱. درک درست از مسئله و تصمیمی که لازم است بگیرید؛
۲. گردآوری **همه اطلاعات**؛
۳. مشخص‌سازی راه‌حل‌ها؛
۴. ارزیابی مزایا و معایب؛
۵. انتخاب بهترین راه‌حل؛
۶. اتخاذ تصمیم؛
۷. ارزیابی اثر تصمیمی که گرفته‌اید.

تغییرات و روندها چنان سریع اتفاق می‌افتند که جایی برای روش آزمون و خطا باقی نمی‌ماند. از این روی، مدیران ناچارند از ابزارها و تکنیک‌های جدید **آنالیتیک** بهره بگیرند.



داده‌ها و آنالیز آنها در تصمیم‌سازی

سازمان‌ها می‌توانند از **آنالیتیک** برای تدوین گزارش درباره آنچه اتفاق می‌افتد و پیش‌بینی آنچه احتمال دارد اتفاق بیفتد، بهره بگیرند و بهترین تصمیم‌سازی ممکن را برای موقعیت به‌وجودآمده انجام بدهند. سازمان‌ها برای این مراحل مقادیر انبوهی از داده‌ها را گردآوری و آنالیز می‌کنند. از لحاظ کلی، مقدار داده‌ها هر سال دو برابر سال پیش می‌شود.

برنامه‌های کامپیوتری شرکت‌ها و مؤسسات تجاری از عملیات **پردازش تراکنش‌ها و نظارت** به **برنامه‌های آنالیز مسئله و راه‌حل** تکامل پیدا کرده‌اند، و امروزه بسیاری از عملیات این برنامه‌ها با فناوری‌های **ابر-بنیاد**[6] انجام می‌شوند، که در بسیاری از موارد از طریق وسایل همراه قابل دستیابی است. ابزارهای آنالیتیک و **BI** (**هوشمندی کسب‌وکار**) مانند **انبار داده‌ها**[7]، **داده‌کاوی**[8]، **پردازش برخط آنالیتیکی**[9] (**OLAP**)، داشبوردها، و برنامه‌های کاربردی **ابر-بنیاد**[10] برای پشتیبانی از تصمیم زیربنای مدیریت مدرن امروزی را تشکیل می‌دهند. مدیران باید سیستم‌های اطلاعات شبکه‌ای پرسرعت (باسیم و بی‌سیم) داشته باشند تا در اجرای مهم‌ترین وظیفه‌شان به آنها کمک کنند: **تصمیم‌سازی**[11]. در بسیاری از موارد چنین تصمیم‌هایی کاملاً خودکار انجام می‌شوند، بدون نیاز به دخالت مدیریتی.

تکامل پشتیبانی تصمیم کامپیوتری به هوشمندی کسب‌وکار/آنالیتیک/علم داده‌ها

نمودار نشان‌داده‌شده در بالای این صفحه **تاریخچه آنالیتیک** را از دهه ۱۹۷۰ به بعد نشان می‌دهد. در دهه ۱۹۷۰، توجه اصلی سیستم‌های پشتیبانی تصمیم بر روی فراهم‌ساختن **گزارش‌های ساخت‌یافته** به طور دوره‌ای بود، و مدیران می‌توانستند برای تصمیم‌سازی آنها را به کار بگیرند (یا آنها را نادیده بگیرند). در نتیجه، کسب‌وکارها به تدوین گزارش‌ها برای آگاهی‌دادن به تصمیم‌سازان (مدیران) درباره آنچه در دوره پیشین رخ داده بوده است (مثلاً روزانه، هفتگی، ماهانه، یا فصلی) روی آوردند. با آن که دانستن آنچه در گذشته رخ داده بوده است مفید بود مدیران به بیش از چنین گزارش‌هایی نیاز داشتند: آنها به انواعی از گزارش‌ها در سطوح مختلف در بخش‌های مختلف نیاز داشتند تا نیازها و چالش‌های در حال تغییر کسب‌وکار را درک و حل کنند. اینها را **سیستم‌های اطلاعات مدیریت**[12] (MIS) می‌نامیدند. در اوایل دهه ۱۹۷۰، **اسکات-مورتون**[13] نخستین بار مفاهیم اصلی **سیستم پشتیبان تصمیم** (**DSS**) را تعریف کرد. او **DSS** را به صورت «سیستم‌های دوسویه کامپیوتری‌ای که به تصمیم‌گیران کمک می‌کنند

[6] cloud-based
[7] data warehousing
[8] data mining
[9] online analytical processing
[10] cloud-based system
[11] decision making

[12] management information system
[13] Scott-Morton

که **داده‌ها** و **مدل‌ها** را برای حل‌کردن مسائل ناساخت‌مند به کار بگیرند» تعریف کرد.

توجه داشته باشید که اصطلاح **سیستم پشتیبان تصمیم** مانند **سیستم مدیریت اطلاعات** و تعداد زیاد دیگری اصطلاح در حوزه **IT** اصطلاحی **با معنی آزاد**[14] هستند (یعنی برای افراد مختلف معانی مختلفی دارند). از همین روی، یک تعریف جامع از **DSS** که مورد توافق عمومی باشد وجود ندارد.

در روزهای اولیه آنالیتیک، داده‌ها از طریق متخصصان هر حوزه‌ای به صورت دستی (مثلاً از طریق مصاحبه یا بررسی) برای ساخت مدل‌های ریاضیاتی یا دانش-بنیاد به منظور بهینه‌سازی و حل مسائل تهیه می‌شد. چنین مدل‌های پشتیبان تصمیمی به **تحقیق عملیات**[15] (**OR**) مشهور بود.

در اواخر دهه ۱۹۷۰ و اوایل دهه ۱۹۸۰، افزون بر مدل‌های پیشرفته **OR** که در بسیاری از صنایع و سیستم‌های دولتی به کار گرفته می‌شد، یک خط جدید و جالب از مدل‌ها پدیدار شد: **سیستم‌های خبره قاعده-بنیاد**[16]. این سیستم‌ها نوید می‌دادند که دانش مدیران را در فرمتی ذخیره کنند که کامپیوترها بتوانند آنها را پردازش کنند (از طریق مجموعه‌ای از قاعده‌های **if-then-else**)، به گونه‌ای که بتوان از خبرگان هر حوزه برای مشخص‌کردن یک مسئله **ساخت‌مند** و تجویز محتمل‌ترین راه حل بهره گرفت. **سیستم‌های خبره** این امکان را فراهم ساختند که دانش خبرگان کم‌شمار حوزه‌های مختلف در هر جا و هر زمانی که لازم باشد از طریق سیستم پشتیبان تصمیمِ «هوشمند» در دسترس عموم قرار بگیرد.

دهه ۱۹۸۰ شاهد یک تغییر مهم در روش کسب داده‌های سازمان‌ها بود. در روش قدیمی چندین سیستم اطلاعات جدا از هم داده‌های تراکنشی واحدهای مختلف یک سازمان را ذخیره می‌کردند (مثلاً داده‌های حساب‌داری، بازاریابی و فروش، مالی، و تولید). در دهه ۱۹۸۰، این سیستم‌ها به شکل سیستم‌های اطلاعات مجتمع کل سازمان در آمدند که حالا به **سیستم‌های برنامه‌ریزی منابع سازمانی**[17] (**ERP**) مشهور هستند. طرح‌های ذخیره داده‌ای **ترتیبی** و غیراستاندارد جای خودشان را به **سیستم‌های مدیریت پایگاه داده رابطه‌ای**[18] (**RDBM**) دادند. این سیستم‌ها داده‌گیری و ذخیره‌سازی داده‌ها را و همچنین رابطه‌های بین فیلدهای داده‌ای سازمانی را بهینه کردند و در ضمن به اندازه چشم‌گیری جلوی تکرار را در ذخیره اطلاعات گرفتند.

نیاز به **RDBM** و **ERP** هنگامی بود که یک‌پارچگی و همخوانی داده‌ها به یک نیاز مهم تبدیل شده بود. از سوی دیگر، یک نیاز دیگر سازمان‌ها **گزارش‌گیری فوری** بود. تصمیم‌سازان هر وقت می‌خواستند تصمیمی بگیرند بلافاصله گزارش‌گیری می‌کردند.

در دهه ۱۹۹۰ نیاز به گزارش‌های انعطاف‌پذیر به توسعه **سیستم‌های اطلاعات اجرایی**[19] (**EIS**ها؛ **DSS**هایی که اختصاصاً برای مدیران اجرایی و نیازهای تصمیم‌سازی آنان طراحی و ساخته می‌شد) انجامید. این سیستم‌ها به شکل داشبوردهای گرافیکی و ورقه‌های ثبت امتیازات طراحی می‌شدند تا عوامل مهم برای تصمیم‌سازان برجسته‌تر دیده شوند. برای این که گزارش‌ها را بتوان با انعطاف بیشتری ساخت، ساخت یک لایه میانی داده‌ها ضروری شد. این لایه به **انبار داده‌ها**[20] مشهور شد. در مدتی کوتاه اکثر کسب‌وکارهای بزرگ تا متوسط **انبار داده‌ها** را به عنوان پلت‌فرم تصمیم‌سازی انتخاب کردند. داشبوردها داده‌های خود را مستقیماً از **انبار داده‌ها** می‌گرفتند، و در نتیجه به کارایی سیستم‌های تراکنش مؤسسه خود _ که سیستم‌های ERP نامیده می‌شدند _ فشاری وارد نمی‌کردند.

در دهه ۲۰۰۰ DSSهای توانمندشده با DW (انبار داده‌ها) به **سیستم‌های هوشمندی کسب‌وکار** یا **BI** شهرت پیدا کردند.

DWها بسیار بزرگ و غنی هستند، در نتیجه «**کاویدن**» داده‌ها برای «**کشف**» دانش جدید و مفید به تدریج رونق گرفت و اصطلاحات

[14] content free expression
[15] operations research
[16] rule-based expert system

[17] enterprise resource planning
[18] relational database management systems
[19] executive information system
[20] data warehousing

داده‌کاوی و **متن‌کاوی**[21] رایج شدند. با افزایش **حجم** یا **تراکم** (volume) و **تنوع** (variety) داده‌ها نیاز به ذخیره‌گرهای بیشتر و قدرت پردازشی بیشتر احساس شد. با آن که شرکت‌های بزرگ از پس این مسئله بر می‌آمدند، شرکت‌های کوچک و متوسط مشکل داشتند. این نیاز به **معماری خدمات-محور** و **نرم‌افزار به عنوان یک سرویس**[22] و **زیرساخت به عنوان یک سرویس**[23] برای مدل‌های کسب‌وکار آنالیتیک منجر گردید. به این ترتیب، شرکت‌های کوچک به امکانات آنالیتیک دست پیدا کردند و به جای این که منابع گران‌قیمت سخت‌افزاری و نرم‌افزاری را خریداری کنند بر اساس میزان استفاده از هر سرویس هزینه‌ای می‌پرداختند.

در دهه ۲۰۱۰، یک **جابه‌جایی پارادایم** دیگر را در روشی می‌بینیم که داده‌گیری انجام می‌شود. علت تا اندازه بسیار زیادی از یک سو در استفاده گسترده مردم از **اینترنت**، و از سوی دیگر، در پدیدارشدن رسانه‌های تولید داده جدید است. از میان همه منابع داده‌ای جدید (مانند برچسب‌های **RFID**[24]، انرژی‌سنج‌های دیجیتال، اطلاعات کلیک‌زنی‌ها بر روی وب، وسایل خانه هوشمند، یا پوشیدنی‌های هوشمند)، شاید از همه جالب‌تر و پرمسئله‌تر داده‌های شبکه‌های اجتماعی/رسانه‌های اجتماعی باشد. این **داده‌های ناساخت‌مند** از لحاظ محتوای اطلاعاتی بسیار غنی هستند، اما آنالیز چنین منابعی از داده‌ها مسائل فراوانی را برای سیستم‌های کامپیوتری، چه نرم‌افزار چه سخت‌افزار مطرح می‌کرد. اخیراً اصطلاح **کلان‌داده‌ها**[25] برای مشخص‌کردن این مسائل ابداع شد. پیشرفت‌های فراوانی در سخت‌افزار (مانند **پردازش موازی انبوه** با مقادیر بسیار فراوانی حافظه و سیستم‌های کامپیوتری با تعداد بسیار زیادی ریزپردازنده موازی) و نرم‌افزار/الگوریتم (مانند **Hadoop** با **MapReduce** و **NoSQL**، **Spark**) به وجود آمده است تا مسائل کلان‌داده‌ها را حل کنند.

در چند سال اخیر رشد خوبی در فناوری‌ها داشته‌ایم و در دهه پیش رو رشد بسیار بزرگ‌تری را شاهد خواهیم بود. به عنوان مثال، **اینترنت چیزها**[26] را **آنالیتیک جریان‌دار**[27] و فناوری‌های حس‌گر توانمندتر کرده‌اند. هوش مصنوعی با به‌کارگرفتن روش‌های جدید آنالیز تصاویر از طریق **یادگیری ژرف**[28] در حال متحول‌کردن **هوشمندی کسب‌وکار** (BI) است. **یادگیری ژرف** و هوش مصنوعی همچنین در حال متحول‌کردن **بازشناسی صدا** و **ساخت صدا** هستند، که رابط‌های جدیدی را در ارتباط با فناوری‌ها فراهم خواهند کرد.

در نتیجه، در میان رقابت‌های جهانی، یک فرصت بزرگ برای اتخاذ تصمیم‌های بهتر با بهره‌گیری از داده‌ها و آنالیتیک به منظور افزایش‌دادن درآمد ضمن کاستن از هزینه‌های تولید پدید آمده است، با ساختن کالاهایی بهتر، بهینه‌سازی تجارب مشتریان، بهینه‌کردن درگیری مشتریان از طریق سفارشی‌سازی، و تأسیس کسب‌وکارهای کاملاً جدید، همگی با **آنالیتیک** و **داده‌ها**.

با آن که بسیاری از مردم معتقدند که **BI** به **آنالیتیک** و **علم داده‌ها** تکامل پیدا کرده است، بسیاری از فروشندگان و پژوهش‌گران هنوز از این اصطلاح بهره می‌گیرند.

EISها شروع کردند به ارائه نمودارها و تصاویر بیشتر، هشدارها، و امکانات اندازه‌گیری کارایی. این کارایی‌ها EISها را دگرگون کردند. از سال ۲۰۰۶ اکثر محصولات و خدمات مهم تحت اصطلاح **هوشمندی کسب‌وکار** یا **BI** به میدان آمدند.

تعریف BI

هوشمندی کسب‌وکار یک اصطلاح چتری است که چندین معماری، ابزار، ابزارهای آنالیتیک، برنامه کاربردی، و روش‌شناسی مختلف را تحت پوشش قرار می‌دهد. همچون **DSS** یک اصطلاح با معنی آزاد است، یعنی برای افراد مختلف می‌تواند معانی مختلفی داشته باشد.

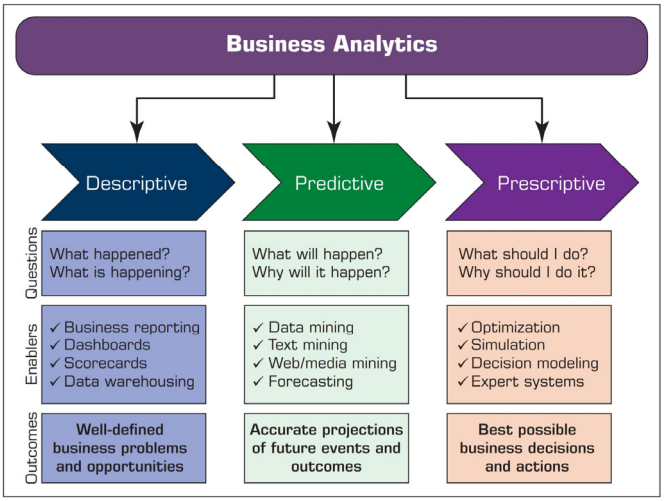
[21] text mining
[22] software as a service
[23] infrastructure as a service
[24] radio-frequency identification
[25] big data

[26] Internet of Things
[27] streaming analytics
[28] deep learning

هدف اصلی BI فراهم‌کردن دستیابی دوسویه داده‌ها (گاهی در **زمان واقعی**[۲۹])، فراهم‌کردن امکان دست‌کاری داده‌ها، و فراهم‌کردن امکان آنالیز برای مدیران است. مدیران با آنالیز داده‌های تاریخی و کنونی، وضعیت‌ها، و عمل‌کردها، **بینش** خوبی از کسب‌وکار خود به دست می‌آورند و می‌توانند تصمیمات آگاهانه و خوبی اتخاذ کنند. **پردازش BI** تبدیل **داده‌ها** به **اطلاعات** و سپس به **تصمیم** و در نهایت به **اقدام** است.

معماری BI

یک سیستم BI چهار بخش اصلی دارد: یک **انبار داده‌ها** (DW)؛ **آنالیتیک کسب‌وکار**؛ یک مجموعه از ابزارها برای دست‌کاری، کاویدن، و آنالیز داده‌های موجود در انبار داده‌ها؛ **BPM**[۳۰] (مدیریت بازدهی کسب‌وکار) برای پایش و آنالیز کارایی؛ و یک رابط کاربری (مثلاً یک **داشبورد**).



آنالیتیک

واژه **آنالیتیک** به اندازه زیادی جای فناوری‌های مختلف **پشتیبان تصمیم** کامپیوتری را که در گذشته تحت برچسب‌های مختلف حضور داشته است گرفته است. در حقیقت، بسیاری از عمل‌گراها و دانشگاهیان امروزه از واژه **آنالیتیک** به جای **BI** بهره می‌گیرند. با آن که بسیاری از نویسندگان و مشاوران آن را اندکی متفاوت تعریف می‌کنند، **آنالیتیک** را می‌توان به عنوان عملیات ساخت تصمیم‌های قابل اجرا یا توصیه‌هایی برای عمل بر بنیاد بینش‌های به‌دست‌آمده از داده‌های تاریخی در نظر گرفت. مطابق تعریف **مؤسسه تحقیق عملیات و علوم مدیریت**[۳۱] (**INFORMS**)، **آنالیتیک** ترکیب فناوری کامپیوتر، تکنیک‌های علوم مدیریت، و آمار برای حل مسائل واقعی است. مؤسسه INFORMS سه سطح از آنالیتیک را پیشنهاد کرده است: **توصیفی**[۳۲]، **آینده‌نگر**[۳۳]، و **تجویزی**[۳۴].

آنالیتیک توصیفی (descriptive)

آنالیتیک توصیفی (یا **گزارشی**) به دانستن آنچه در یک سازمان در حال روی‌دادن است و درک روندهای در بطن آن و علت‌های چنان رویدادهایی اشاره دارد. اول از همه، این آنالیتیک شامل ادغام منابع داده‌ای و در دسترس قرارگرفتن همه داده‌های باربط و به‌دردبخور به شکلی است که **گزارش‌دهی** و **آنالیز** مورد نیاز را ممکن بسازند. معمولاً ساخت این نوع زیرساختار داده‌ها در بخشی از برنامه‌های DWها (انبار داده‌ها) انجام می‌گیرد. با چنین زیرساختاری می‌توانیم گزارش‌ها، بررسی‌ها، هشدارها، و روندهای وضعیت موجود را با استفاده از ابزارها و تکنیک‌های گزارش‌سازی بسازیم.

یک فناوری مهم که یک بازی‌گر مهم در این حوزه است فناوری **بصری‌سازی**[۳۵] است. با استفاده از آخرین فناوری‌های **بصری‌سازی** می‌توانیم به **بینش‌های** عمیقی درباره عملیات سازمانی‌مان دست پیدا کنیم.

[29] real time
[30] Business Performance Management
[31] Institute for Operations Research and Management Science
[32] descriptive
[33] predictive
[34] prescriptive
[35] visiualization

آنالیتیک آینده‌نگر (predictive)

هدف **آنالیتیک آینده‌نگر** آن است که آنچه را ممکن است در آینده روی بدهد تعیین کند. این آنالیتیک بر بنیاد تکنیک‌های آماری و همچنین تکنیک‌های به تازگی ساخته‌شده‌ای است که تحت گروه **داده‌کاوی** قرار می‌گیرند. هدف این تکنیک‌ها آن است که بتوانند پیش‌بینی کنند که آیا این احتمال وجود دارد که مشتری برای خرید بعدی به یک رقیب مراجعه کند (churn) یا نه، مشتری در خرید بعدی خود چه چیزی و چه مقداری خواهد خرید، چه تشویق‌هایی می‌تواند مشتری را به خرید بعدی تشویق کند، ریسک اعتبار مالی مشتری چقدر است، و مانند آن. انواعی از تکنیک‌ها برای برنامه‌های آنالیتیک آینده‌نگر به کار گرفته می‌شود، شامل انواع **الگوریتم‌های طبقه‌بندی**[۳۶] (مانند تکنیک‌های **شبکه عصبی**[۳۷] یا مدل‌های **درخت تصمیم**[۳۸]).

آنالیتیک تجویزی (prescriptive)

هدف **آنالیتیک تجویزی** درک آن چیزی است که در حال روی‌دادن است و همچنین درک آن چیزی است که پیش‌بینی می‌شود روی بدهد، و اتخاذ تصمیم برای دست‌یافتن به بالاترین کارآمدی. این گروه از تکنیک‌ها به طور تاریخی تحت چتر **تحقیق عملیات** یا **علم مدیریت** مطالعه می‌شد و معمولاً هدف آن بهینه‌سازی کارایی یک سیستم بوده است.

تصمیم‌ها ممکن است از طریق یک گزارش برای یک تصمیم‌ساز ارائه شود و یا به طور مستقیم در یک سیستم قواعد تصمیم خودکار به کار گرفته شود.

به کاربستن آنالیتیک برای قلمروهای مختلف

آنالیتیک در صنایع و قلمروهای مختلف به کار گرفته می‌شود. در نتیجه این واژه را به این قلمروها یا صنایع می‌توان وصل کرد تا به یک حوزه تخصصی ویژه ارتباط پیدا کند، مانند **آنالیتیک متن، آنالیتیک وب، آنالیتیک ورزش، آنالیتیک بازاریابی، آنالیتیک سلامت، آنالیتیک خرده‌فروشی، آنالیتیک رفتاری**، و مانند آن.

آنالیتیک یا علم داده‌ها

با آن که مفهوم آنالیتیک در صنایع و دانشگاه‌ها توجهات بیشتری را به خود جلب کرده است، یک اصطلاح دیگر نیز مدتی است معرفی و شایع شده است. این اصطلاح جدید **علم داده‌ها**[۳۹] است. از این روی، به عمل‌گرایانِ علم داده‌ها **داده‌دان**[۴۰] گفته می‌شود. گاهی از **دی. جی. پاتیل**[۴۱] در LinkedIn به عنوان ابداع‌کننده اصطلاح **علم داده‌ها** نام برده می‌شود.

با این همه، پاره‌ای از صاحب‌نظران بین **آنالیست‌های داده‌ها** و **داده‌دان‌ها** تفاوت‌هایی را برشمرده‌اند. یک دیدگاه آن است که **آنالیست داده‌ها** اصطلاحی است که برای متخصصانی به کار می‌رود که کار BI به شکل گردآوری، تمیزسازی، گزارش‌سازی، و شاید بصری‌سازی داده‌ها انجام می‌دهند. در مجموعه مهارت آنها مهارت در برنامه **اِکسل**، معلوماتی از SQL، و گزارش‌سازی نیز وجود دارد. این توانمندی‌ها در جهت **آنالیتیک تجویزی** یا **گزارش‌سازی** (**توصیفی**) است.

افزون بر این، آنالیتیک تجویزی شامل خبرگی در **تحقیق عملیات** شامل بهینه‌سازی، شبیه‌سازی، و آنالیز تصمیم است.

در مقابل، **داده‌دان‌ها** مسئول **آنالیز آینده‌نگر**، آنالیز آماری، و کاربرد ابزارهای آنالیتیکی و الگوریتم‌های پیشرفته هستند. آنها ممکن است معلومات عمیق‌تری درباره الگوریتم‌ها داشته باشند و ممکن است این الگوریتم‌ها را تحت برچسب‌های مختلفی مانند **داده‌کاوی، کشف دانش**[۴۲]، یا **یادگیری ماشین**[۴۳] استفاده کنند. برای بعضی از این متخصصان همچنین لازم است که معلومات برنامه‌سازی خوبی برای نوشتن برنامه‌های **تمیزسازی داده‌ها**[۴۴] یا آنالیز داده‌ها به زبان‌های

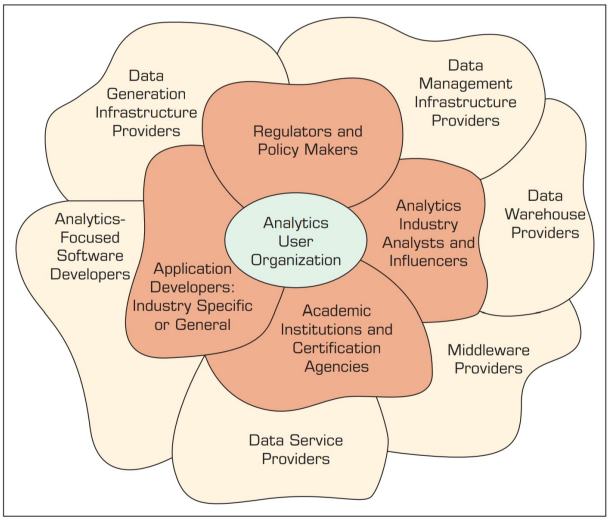
[36] classification
[37] neural network
[38] decision tree model

[39] data science
[40] data scientist
[41] D. J. Patil
[42] knowledge discovery
[43] machine learning
[44] data cleaning

وب-محور مانند **جاوا** یا **پایتون**[45] و به زبان‌های آماری مانند **R** داشته باشند.

با این همه، کسانی که در این حوزه‌ها تسلط داشته باشند بیشتر **داده‌دان** خوانده می‌شوند تا **آنالیست داده‌ها**. در مجموع، **داده‌دان‌ها** معلومات فنی خوبی در دانش و مهندسی کامپیوتر، آمار، و برنامه‌های ریاضیات کاربردی دارند.

از سوی دیگر، صاحب‌نظران فیزیک کاربردی پیشنهاد کرده‌اند که از اصطلاح **علم شبکه**[46] برای توصیف آنالیتیک مربوط به گروه‌های مردم _ شبکه‌های اجتماعی، شبکه‌های زنجیره عرضه، و مانند آن _ بهره گرفته شود[47].



کلان‌داده‌ها چیست؟

امروزه **آنالیتیک** و **علم داده‌ها** در اصل به معنی **آنالیتیک کلان‌داده‌ها** هستند. مغزهای ما بسیار سریع و کارآمد عمل می‌کنند و در پردازش مقادیر زیادِ همه انواع داده‌ها انعطاف خوبی دارند: تصویر، متن، صدا، بو، و ویدئو. ما همه انواع مختلف داده‌ها را به آسانی پردازش می‌کنیم. از سوی دیگر، کامپیوترها در همگامی با داده‌هایی که لحظه به لحظه تولید می‌شوند هنوز مشکل دارند، چه برسد به آنالیز سریع آنها. از همین روست که ما برای **کلان‌داده‌ها** مشکلات فراوانی داریم. اما **کلان‌داده‌ها** چیست؟ به بیان ساده، **کلان‌داده‌ها** به داده‌هایی اشاره دارد که در یک واحد ذخیره‌گرِ تنها نمی‌تواند ذخیره شود. کلان‌داده‌ها به داده‌هایی اشاره دارد که در شکل‌های بسیار مختلفی هستند: **ساخت‌مند**، **ناساخت‌مند**، در یک **جریان داده‌ای**[48]، و مانند آن. منابع اصلی چنین داده‌هایی عبارتند از کلیک‌زنی‌ها بر روی وبگاه‌ها، پست‌های رسانه‌های اجتماعی مانند **فیس‌بوک**، و داده‌های تولیدی وسایل ترافیک، حس‌گرها، یا سیستم‌های پیش‌بینی آب‌وهوا. یک **موتور جست‌وجو** مانند **گوگل**، به جست‌وجو و نمایه‌گذاری میلیاردها صفحه وب نیاز دارد تا بتواند در جست‌وجوهایی که شما با این موتور انجام می‌دهید در کسری از ثانیه داده‌های به‌دردبخور و باربط ارائه بدهد. با آن که نمایه‌سازی از همه صفحات وب بر روی اینترنت در **زمان واقعی** (بلادرنگ) انجام نمی‌شود و گاه روزها طول می‌کشد، کار آسانی نیست. خوشبختانه **گوگل** توانسته است این مسئله را حل کند. یکی از ابزارهای مهم **گوگل** برای حل این مسئله **تکنیک‌های آنالیتیک کلان‌داده‌ها** بوده است.

اداره این حجم از داده‌ها با دو توان‌مندی مهم انجام می‌شود: **ذخیره‌سازی** و **پردازش**. اگر ما یک راه‌حل ذخیره‌سازی بسیار گران‌قیمت برای ذخیره‌سازی همه این داده‌ها در **یک واحد** را داشتیم ایجاد امکانات **تحمل خطا**[49] برای چنین واحدی به هزینه بسیار گزافی نیاز پیدا می‌کرد. یک راه حل هوشمندانه پیشنهاد شد که شامل ذخیره‌سازی این داده‌ها در قسمت‌های مختلف بر روی ماشین‌های مختلف متصل به شبکه بود _ قراردادن یک یا دو کپی از هر قسمت در مکان‌های مختلف بر روی شبکه، هم به صورت **منطقی** و هم به صورت **فیزیکی**. در ابتدا توسط **گوگل** به کار گرفته شد (که مدتی بعد **Google File System** نام گرفت) و بعدها توسط یک پروژه **Apache** به عنوان **Hadoop Distributed File System (HDFS)** ساخته و منتشر شد.

[45] Python
[46] network science
[47] http://barabasi.com/networksciencebook

[48] data stream
[49] fault tolerant

با وجود این، **ذخیره‌سازی داده‌ها** نیمی از مسئله بود. داده‌ها اگر در کسب‌وکارها ارزشی نیافرینند ارزشی ندارند، و برای این که در کسب‌وکارها ارزش بیافرینند، باید آنالیز شوند. این حجم بسیار انبوه داده‌ها چگونه پردازش شود؟ دادن کل محاسبات کامپیوتری به یک کامپیوتر قدرتمند عملی نیست، این حجم از داده‌ها سرریز فوق‌العاده بزرگی را بر روی یک چنین کامپیوتر قدرتمندی وارد خواهد کرد. یک راه حل هوشمندانه دیگر پیشنهاد شد: محاسبات کامپیوتری را به **گره**[50] **داده‌ها** محول (push) کنید، به جای آن که داده‌ها را به یک **گره کامپیوتر** در شبکه بدهید. این یک پارادایم جدید بود و به یک روش کاملاً جدید پردازش داده‌ها انجامید. این چیزی است که امروزه ما آن را پارادایم **برنامه‌سازیِ MapReduce** (نگاشت-کاهش) می‌شناسیم، که پردازش کلان‌داده‌ها را ممکن ساخت. MapReduce ابتدا در **گوگل** ساخته شد، و یک نگارش بعدی آن را پروژه Apache تحت نام **Hadoop MapReduce** منتشر کرد.

در چند سال گذشته آنچه **کلان‌داده‌ها** نامیده می‌شد با ظهور کاربردهای کلان‌داده‌ها تغییر کرد. نیاز به پردازش داده‌هایی که با سرعت زیاد می‌آمدند **تُندی** (velocity) را به معادله اضافه کرد. **معاملات الگوریتمی** یک نمونه از پردازش **داده‌های سریع** است. معاملات الگوریتمی از پلت‌فرم‌های الکترونیک بر بنیاد الگوریتم برای معاملات سهام در بورس اوراق بهادار بهره می‌گیرد که در میکروثانیه عمل می‌کنند.

نیاز به پردازش انواع مختلف داده‌ها **تنوع** (variety) را نیز به معادله اضافه کرد. نمونه‌ای دیگر از تنوع گسترده داده‌ها **آنالیز نظرات** است، که از انواع مختلف شکل‌ها در پلت‌فرم‌های رسانه‌های اجتماعی و نظرات مشتریان کالاها و خدمات بهره می‌گیرد. امروزه، کلان‌داده‌ها تقریباً به همه مجموعه‌های بزرگ داده‌ها گفته می‌شود که خصوصیات **تراکم** (volume)، **تندی**(velocity)، و **تنوع** (variety) را داشته باشند (مشهور به **سه v** یا **سه ت**). همچنان که پیشتر گفته شد در کنار اینها با **آنالیتیک داده‌های جریان‌دار**، **اینترنت چیزها**، **رایانش ابری**[51]، و هوش مصنوعی توانمندشده با **یادگیری ژرف** نیز سروکار داریم.

[50] data node
[51] cloud computing

هوش مصنوعی

هوش مصنوعی که سال‌ها یک شگفتی مردم بوده است به سرعت در حال تبدیل‌شدن به یک فناوری کاربردی با کاربردهای گوناگون است. پاره‌ای از صاحب‌نظران معتقدند که هوش مصنوعی مهم‌ترین فناوری‌ای است که بشر ساخته است. تعریف‌های مختلفی برای هوش مصنوعی شده است. به بیان ساده، می‌توان گفت که هدف هوش مصنوعی ساخت ماشین‌هایی است که هوشمندی‌ای را به نمایش می‌گذارند که نزدیک به هوش انسان باشد، با این امید که به نفع انسان باشد.

سرمایه‌گذاری‌های سنگینی روی هوش مصنوعی انجام می‌شود. دولت چین امیدوار است که با صرف ۶۰ میلیارد دلار تا سال ۲۰۲۵ به رهبری هوش مصنوعی در دنیا برسد.

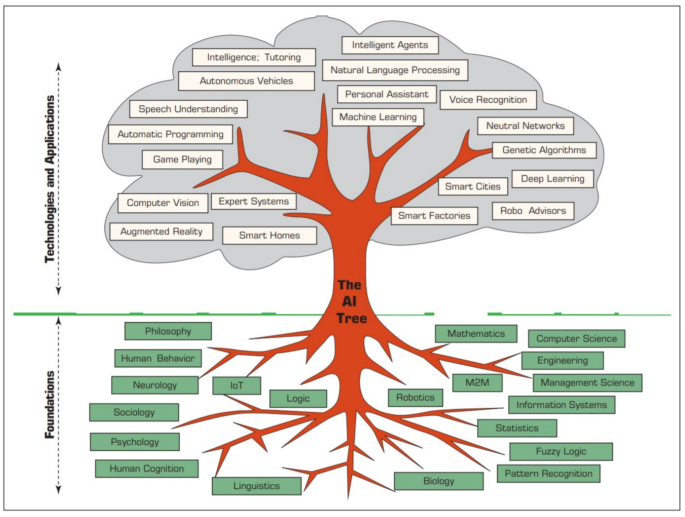


FIGURE 2.1 The Functionalities and Applications of Artificial Intelligence.

هوش مصنوعی در تصمیم‌سازی

ما دوست داریم که تصمیم‌سازی کامپیوتری ساده‌تر شود، طرز کار آسان‌تری داشته باشد، فراست بیشتری داشته باشد، و تهدیدکننده نباشد؛ و در حقیقت، تلاش‌های فراوانی در این مسیر در طول زمان برای ساده‌ترکردن و خودکارسازی تکالیف مختلف در عملیات

تصمیم‌سازی انجام شده است. کافی است روزی را تصور کنید که در آن یخچال‌ها قادر باشند که محتویات خود را اندازه‌گیری و ارزیابی کنند و برای مواد غذایی مورد نیاز سفارش خرید بدهند. رسیدن به چنین روزی دیر نخواهد پایید، و چنین تکلیفی را هوش مصنوعی پشتیبانی خواهد کرد.

خصوصیات مهم هوش مصنوعی

یک روند رو به فزونی برای ساخت هوشمندتر کامپیوترها وجود دارد. برنامه‌های کاربردی گوناگونی بر بنیاد تکنیک‌های مختلف هوش مصنوعی ساخته شده است. به عنوان مثال، حوزه **ترجمه ماشینیِ** زبان‌ها به مردم کمک می‌کند که به زبان‌های مختلف برای همکاری و همچنین خرید محصولاتی با هم صحبت کنند که به زبان‌هایی غیر از زبان مادری‌شان آگهی شده باشند. به طور مشابه، ترجمه ماشینی می‌تواند به مردمی که فقط با زبان مادری‌شان آشنایند کمک کند که با مردمی صحبت کنند که به زبان‌های دیگر صحبت می‌کنند و در **زمان واقعی** (بلادرنگ) با هم تصمیم‌گیری کنند.

پشتیبانی هوش مصنوعی در تصمیم‌سازی

تقریباً از همان بدو تولد هوش مصنوعی، پژوهش‌گران امکان استفاده از آن را برای پشتیبانی از عملیات تصمیم‌سازی و برای خودکارسازی کامل عمل تصمیم‌سازی تشخیص دادند. این که **جف بزوس**[52]، مدیر عامل **آمازون** در ماه مه ۲۰۱۷ گفت که هوش مصنوعی در یک عصر طلایی قرار دارد و مسائلی را حل می‌کند که در گذشته در قلمروی داستان‌های علمی-تخیلی قرار داشته است سخنی به گزافه نگفته است. **بزوس** درباره علت این اظهارات خود گفت که شرکت **آمازون** تقریباً در صدها کاربرد و برنامه مختلف از هوش مصنوعی بهره می‌گیرد و هوش مصنوعی در عمل یک یاور واقعی شرکت **آمازون** شده است. به عنوان مثال، **آمازون** بیش از ۲۰ سال است که از هوش مصنوعی برای توصیه‌های خرید بهره می‌گیرد. این شرکت همچنین برای قیمت‌گذاری محصولات و برای بسیاری از مسائل دیگر از هوش مصنوعی بهره می‌گیرد. در واقع، **آمازون** از ابتدای تأسیس از هوش مصنوعی برای حل مسئله و تصمیم‌سازی بهره گرفته است. فناوری‌های هوش مصنوعی به مردم امکان می‌دهند که تصمیم‌های بهتری بگیرند. حقیقت آن است که هوش مصنوعی این توانایی‌ها را دارد:

- **حل مسائل پیچیده‌ای** که مردم قادر به حل‌کردن آنها نیستند. (توجه داشته باشید که حل مسائل معمولاً به تصمیم‌سازی نیز ربط پیدا می‌کند.)؛
- **تصمیم‌گیری بسیار سریع‌تر**. به عنوان مثال، **آمازون** میلیون‌ها تصمیم قیمت‌گذاری و توصیه را در کسری از هر ثانیه انجام می‌دهد؛
- **پیداکردن بسیار سریع اطلاعات به‌دردبخور و باربط**، حتی در منابع داده‌ای بسیار بزرگ؛
- **اجرای سریع محاسبات پیچیده**؛
- **اجرای مقایسه‌ها و ارزیابی‌های پیچیده در زمان واقعی (بلادرنگ)**.

در یک کلام، هوش مصنوعی می‌تواند سبب سریع‌ترشدن بسیار زیاد بسیاری از انواع تصمیمات شود. برای اطلاعات بیشتر ویدئوی **youtube.com/watch?v=DrjeRy9whQ** را ببینید. همان گونه که پیشتر گفته شد ماهیت تصمیم‌سازی، به ویژه انواع غیرمعمول آنها پیچیده است. انواع گوناگون تصمیم و سطوح مدیریتی گوناگون برای اتخاذ آن تصمیمات وجود دارد. تصمیم‌سازی، که بسیاری از آنها برای حل مسئله اتخاذ می‌شوند، به هوش و خبرگی نیاز دارد. هدف هوش مصنوعی آن است که هر دو را فراهم کند. در نتیجه روشن است که استفاده از هوش مصنوعی برای تصمیم‌سازی شامل فرصت‌ها، منافع، و گوناگونی‌هاست. به عنوان مثال، هوش مصنوعی می‌تواند به گونه‌ای موفقیت‌آمیز از انواع خاصی از تصمیم‌سازی‌ها پشتیبانی کند و بقیه آنها را کاملاً خودکار کند.

[52] Jeff Bezos

مسائل تصمیم‌سازی

مسائل مختلفی در موفقیت هوش مصنوعی در تصمیم‌سازی دخیل هستند:

●**ماهیت تصمیم**. به عنوان مثال، احتمال خودکارسازی تصمیمات تکراری یا روتین بسیار بیشتر است، به ویژه اگر ساده باشند؛

●**روش پشتیبانی، فناوری(هایی) که به کار گرفته می‌شود**. در آغاز، پشتیبانی خودکار تصمیمات بر **بنیاد قاعده** (rule-based) بودند. در عمل، **سیستم‌های خبره** برای تولید راه‌حل برای وضعیت‌های تصمیم‌سازی ویژه در قلمروهای به دقت تعریف‌شده ساخته می‌شدند. یک فناوری پرطرف‌دار دیگر فناوری «**توصیه‌گر**[53]» بود که به همراه تجارت الکترونیک در دهه ۱۹۹۰ پدیدار شد. امروزه، با کاربرد رو به فزونی **یادگیری ماشین** و **یادگیری ژرف** رو به رو هستیم. یک فناوری مرتبط با اینها فناوری **بازشناسی الگو**[54] است. امروزه، توجهات به انواع **بازشناسی زیست‌سنج**[55] زیاد شده است؛

●**آنالیزهای هزینه-فایده و ریسک**. برای تصمیم‌های بزرگ-مقیاس این دو عامل بسیار مهم هستند، اما محاسبات کامپیوتری چنین عواملی به دلیل دشواری اندازه‌گیری هزینه‌ها، ریسک‌ها، و منافع با مدل‌های هوش مصنوعی ساده نیستند؛

●**الگوریتم‌های هوش مصنوعی**. تعداد بسیار زیادی الگوریتم هوش مصنوعی وجود دارد که بنیادهای تصمیم‌سازی‌های خودکار و پشتیبانی از تصمیم هستند. کیفیت تصمیم‌ها به ورودی الگوریتم‌ها بستگی دارد، که ممکن است تحت تأثیر محیط کسب‌وکار قرار بگیرند؛

●**سرعت**. **خودکارسازی تصمیم** به سرعتی که لازم است تصمیم‌ها تحت اقدام قرار بگیرند نیز بستگی دارد. بعضی از تصمیم‌ها به این دلیل نمی‌توانند خودکار بشوند که به‌دست‌آوردن همه داده‌های باربط ورودی برای آنها وقت‌گیر است. از سوی دیگر، تصمیم‌های دستی (غیرکامپیوتری) ممکن است برای بعضی از شرایط بسیار آهسته باشند.

پشتیبانی هوش مصنوعی در گام‌های مختلف عملیات تصمیم‌سازی

امروزه برای گام‌های مختلف عملیات تصمیم‌سازی از پشتیبانی هوش مصنوعی می‌توان بهره گرفت. تصمیم‌های کاملاً خودکار برای وضعیت‌های روتین معمول هستند. گام‌های مختلف عملیات تصمیم‌سازی در زیر آمده است:

● **تشخیص مسئله**. برای تشخیص مسائل امروزه به طور گسترده‌ای از سیستم‌های هوش مصنوعی بهره گرفته می‌شود، مثلاً در تشخیص کژکاری دستگاه‌ها و مسائل پزشکی، یافتن روزنه‌های امنیتی، تخمین سلامت مالی، و مانند آن. فناوری‌های گوناگونی به کار گرفته می‌شود. به عنوان مثال، داده‌های تحت کنترل حس‌گرها را می‌توان با الگوریتم‌های هوش مصنوعی به کار گرفت. سطوح کارآمدی ماشین‌ها با استانداردها مقایسه می‌شود، و آنالیز روندها را می‌توان به فرصت‌ها یا تهدیدها نسبت داد.

● **تولید یا یافتن راه‌حل‌ها**. انواعی از فناوری‌های هوش مصنوعی با انطباق‌دادن خصوصیات مسائل با بهترین تجربه‌ها یا راه‌حل‌های پیشتر آزموده‌شده و اثبات‌شده که در **پایگاه‌های داده** ذخیره شده‌اند راه‌حل ارائه می‌دهند. هم **سیستم‌های خبره** و هم **چَت‌بات‌ها**[56] این رهیافت‌ها را به کار می‌گیرند. آنها می‌توانند راه‌حل‌های توصیه‌شده را تولید کنند، یا چندین گزینه فراهم کنند تا از میان آنها یک راه‌حل انتخاب شود. ابزارهای هوش مصنوعی مانند **استدلال مبتنی بر موارد خاص**[57] و **رایانش عصبی** برای این منظور به کار گرفته می‌شود.

● **انتخاب یک راه‌حل**. مدل‌های مختلف هوش مصنوعی برای ارزیابی راه‌حل‌های پیشنهادشده به کار گرفته می‌شوند، مثلاً با

[53] recommender
[54] pattern recognition
[55] biometric recognition

[56] chatbot
[57] case-based reasoning

پیش‌بینی اثر آینده آنها (**آنالیز آینده‌نگر**)، برآورد موفقیت آنها، یا پیش‌بینی پاسخ یک شرکت به اقدامی که توسط کامپیوتر انجام می‌گیرد.

● **به‌کاربستن راه‌حل‌ها**. هوش مصنوعی می‌تواند برای پشتیبانی از **عملیات اجرای** راه‌حل‌های پیچیده به کار گرفته شود. به عنوان مثال، در مرحله اجرا می‌تواند برای نمایش پیشنهادهای برتر و برای نمایش برآورد مقاومت‌ها در برابر تغییرات به کار گرفته شود.

تصمیم‌سازی خودکار

هرچه قدرت فناوری‌های هوش مصنوعی بیشتر شود، توان آن در خودکارسازی کامل تعداد بیشتری از وضعیت‌های تصمیم‌سازی بیشتر می‌شود.

پشتیبانی از تصمیم هوشمند و خودکار. از دهه ۱۹۷۰، تلاش‌های فراوانی برای خودکارکردن تصمیم‌سازی به عمل آمده است. این تلاش‌ها به طور معمول با بهره‌گیری از **سیستم‌های خبرهٔ قاعده-بنیاد** انجام گرفته است که راه‌حل‌های توصیه‌شده را برای مسائل مدیریتی تکراری فراهم می‌کنند. نمونه‌هایی از تصمیم‌هایی که به طور خودکار انجام می‌شوند در زیر آمده است:

- تصویب وام‌های کوچک؛
- مراحل ابتدایی درخواست‌های استخدام؛
- قیمت محصولات و خدمات (زمان و چگونگی تغییردادن آنها)؛
- توصیه خرید محصولات (مثلاً در **آمازون**).

ماهیت داده‌ها، مدل‌سازی آماری، و بصری‌سازی

در عصر کلان‌داده‌ها و آنالیتیکِ کسب‌وکار که در آن زندگی می‌کنیم اهمیت داده‌ها را نمی‌توان نادیده گرفت. عباراتی که به تازگی ابداع شده‌اند مانند «داده‌ها نفت است»، «داده‌ها ارز جدید است»، و «داده‌ها پادشاه است» روی اهمیت بیشتر داده‌ها تأکید می‌کنند. اما نوع داده‌هایی که درباره آنها صحبت می‌کنیم واضح است که منظورمان هر داده‌ای نیست. مفهوم و اصل «**ورودی آشغال خروجی آشغال**»[58] یا **GIGO** برای پدیده کلان‌داده‌های امروزی بیشتر صادق است.

برای این که ارزش ایجاد شود، به **بینش** و **دانایی** برسیم، داده‌ها باید به دقت ساخته، شناسایی، گردآوری، مجتمع، تمیزسازی، و تبدیل شوند، و در بافتار درست قرار بگیرند تا بتوانند تصمیم‌سازی دقیقی را به دست بدهند.

ماهیت داده‌ها

داده‌ها مواد خام **علم داده‌ها**، **هوشمندی کسب‌وکار** (BI)، و **آنالیتیک** هستند. در حقیقت آنها را می‌توان مواد خام برای **فناوری‌های تولید تصمیم _ اطلاعات، بینش**[59]، و **دانش**[60]_ در نظر گرفت. بدون داده‌ها هیچ‌کدام از این فناوری‌ها ممکن نبود که به وجود بیایند. با وجود این، در گذشته ما مدل‌های مختلفی از آنالیتیک را بر بنیاد دانش خبرگان و تجارب، شاید با کمی داده یا اصلاً بدون داده ساخته‌ایم؛ هرچند، مربوط به گذشته بوده است، اما امروز داده‌ها اساس تصمیم‌سازی و آنالیتیک هستند. گردآوری، ذخیره‌سازی و مدیریت داده‌ها در گذشته بسیار دشوار بوده است. امروزه داده‌ها باارزش‌ترین جزء سرمایه یک سازمان و یک کشور هستند. داده‌ها این پتانسیل را دارند که **بینش‌های** باارزشی به وجود بیاورند تا مشتریان، رقبا، و عملیات کسب‌وکار را بهتر بشناسیم.

روش‌های گردآوری داده‌ها در گذشته دستی بوده است، اما امروزه انواع مکانیسم‌های گردآوری خودکار داده‌ها که از اینترنت، حس‌گر و RFID بهره می‌گیرند به وجود آمده است. این سیستم‌های گردآوریِ خودکار داده‌ها نه‌تنها ما را در گردآوری حجم‌هایی هنگفت از داده‌ها توانمند می‌کنند بلکه **کیفیت داده‌ها**[61] و صحت و یک‌پارچگی داده‌ها را بهتر می‌کنند.

[58] garbage in garbage out
[59] insight
[60] knowledge
[61] data quality

هنر و دانش پیش‌پردازش

داده‌ها در شکل اولیه (داده‌های دنیای واقعی) معمولاً برای استفاده در تکالیف آنالیتیک آماده نیستند. آنها اغلب کثیف، نامنظم، پیچیده، و نادقیق هستند. یک عملیات وقت‌گیر مشهور به **پیش‌پردازش** لازم است تا داده‌های دنیای واقعی را به داده‌های پالایش‌شده برای الگوریتم‌های آنالیتیک تبدیل کند.

مرحله اول پیش‌پردازش **پیداکردن داده‌های باربط** است. مرحله دوم **تمیزسازی داده‌ها**ست.

کلان‌داده‌ها از کجا می‌آید؟

پاسخ ساده است: از همه جا. منابعی که پیشتر به دلیل محدودیت‌های فنی نادیده گرفته می‌شد حالا به عنوان معدن طلا به آنها نگاه می‌شود. پاره‌ای از منابع داده‌ای کلان‌داده‌ها به قرار زیرند: وبلاگ‌ها، GPSها، RFIDها، متن‌های اینترنتی، داده‌های **اینترنت چیزها**، نمایه‌های جست‌وجوی وب، سابقه پزشکی، بایگانی‌های عکس و ویدئو، پست‌های رسانه‌های اجتماعی، خریدوفروش‌های اینترنتی، نظرات مشتریان کالاها و خدمات و ...□

تصمیم‌سازی گروهی، سیستم‌های همکاری، و پشتیبانی هوش مصنوعی

مردم با هم کار می‌کنند و تعداد زیادی تصمیم گروهی می‌گیرند. پشتیبانی از کار گروهی که در آن اعضای گروه ممکن است در نقاط مختلف باشند و در زمان‌های مختلف کار کنند ممکن شده است. پشتیبانی از گروه‌ها یک خصوصیت مهم و اساسی سیستم‌های پشتیبانی از تصمیم است. برای کارِ گروهی انواعی از ابزارها و روش‌ها ابداع شده است: **هوشمندی جمعی، جمع‌سپاری،** و انواع گوناگون الگوریتم‌های هوش مصنوعی. **همکاری انسان-ماشین و ماشین-ماشین قدرت همکاری و حل مسئله را بالا می‌برد.**

هنگامی که کاری را بیش از یک نفر انجام می‌دهند با **کار گروهی** یا **group work** سروکار داریم. تصمیم‌سازی گروهی یک خصوصیت کارهای گروهی است. **گروه‌افزارها** انواعی از نرم‌افزارها هستند که از کار گروهی پشتیبانی می‌کنند. گروه‌افزارها مکانیسم‌های ویژه‌ای را برای اعضای گروه فراهم می‌کنند تا بتوانند نظرات، داده‌ها، اطلاعات، دانش، و منابع دیگرشان را به اشتراک بگذارند.

سیستم‌های پشتیبانی تصمیم گروهی (GDSS) یکی از انواع گروه‌افزار است. که معمولاً می‌تواند بهره‌وری تصمیم‌سازی‌ها را بالا ببرد.

معیار بهره‌وری برای گروه‌افزارها **هوشمندی جمعی**[۶۲] است. «**هوشمندی جمعی**» (**CI**) به هوشمندی کل گروه اشاره دارد. همچنین به **دانایی جمع** یا **دانایی توده‌ها** نیز اشاره دارد. مردم در یک گروه از مهارت‌ها و دانسته‌های خودشان برای حل مسائل و فراهم‌سازی بینش‌ها و نظرات تازه بهره می‌گیرند. فایده‌های مهم بالا بودن **هوش جمعی** عبارت است از حل مسائل پیچیده و/یا طراحی کالاها و خدمات جدید که از نوآوری‌ها سرچشمه می‌گیرند. یک مرکز پژوهشی مهم **هوشمندی جمعی** (CI) **مرکز MIT برای هوشمندی جمعی**[۶۳] (CCI) است (cci.mit.edu). یکی از پژوهش‌های مهم CCI آن است که چگونه ترکیب مردم و کامپیوترها می‌توانند به همراه هم کار کنند، به گونه‌ای که بتوانند خلاقانه‌تر از هر فرد تنها، گروه تنها، یا کامپیوتر تنها عمل کنند. CI در علوم مختلف از جامعه‌شناسی تا علوم سیاسی مورد مطالعه قرار می‌گیرد.

جمع‌سپاری به عنوان یک روش برای پشتیبانی از تصمیم

جمع‌سپاری[۶۴] به برون‌سپاری تکالیف به گروه بزرگی از مردم اشاره دارد. یکی از دلائل مهم اجرای این کار بهره‌گیری از پتانسیل دانایی یک جمع به منظور بهینه‌کردن تصمیم‌سازی‌ها و کمک به حل‌کردن مسائل دشوار است. بنابراین، جمع‌سپاری را می‌توان یکی از روش‌های بهره‌گیری از **هوشمندی جمعی** در نظر گرفت.□

---

62 collective intelligence
63 MIT Center for Collective Intelligence
64 crowdsourcing