



# تحلیل‌های عظیم‌داده

نقشه راه پیاده‌سازی، فناوری و ابزارها

دیوید لوشین

مترجمین:

دکتر سعید روحانی

(عضو هیات علمی دانشکده مدیریت دانشگاه تهران)

سمیه حسینی

نیاز دانش

# پیشگفتار مترجمین

فناوری اطلاعات و ارتباطات در تمامی ارکان اجتماعی و ساخت‌های سازمانی نفوذ کرده و تبدیل به بستر نوین در حین حال مستحکم زندگی و کسب و کار گردیده است. این بستر شامل سیستم‌ها و فناوری‌ها، روز به روز از طریق داده‌ها و اطلاعات در حال عظیم و حجیم شدن است، لیکن راهبرد آینده برای برخورد و بهره‌برداری از این منابع همواره چالش برانگیز بوده است. مفهوم Big Data که ما آن را عظیم داده نامیدیم، یکی از رویکردهای در حال گسترش در مقابل این چالش است، که قطعاً در آینده‌ای نزدیک در کشور ما نیز به صورت مساله مدیریت پردازش‌ها، محاسبات و تحلیل‌های کسب و کار مورد توجه مدیریت فناوری اطلاعات سازمان‌ها قرار خواهد گرفت.

موارد فوق ما را بر آن داشت که کتاب ارزشمند "تحلیل‌های عظیم داده" جناب آقای پروفیسور دیوید لوشین را که بر اساس تجارب عملیاتی ایشان تدوین شده و یکی از معتبرترین کتاب‌های این موضوع است، به فارسی برگردان و ترجمه نماییم. این کتاب به نثر قابل فهم و در سطح کاربردی مدیریت فناوری اطلاعات، تمامی موضوعات مرتبط با عظیم داده را معرفی، بحث و ارائه نموده است. این کتاب می‌تواند مرجع علمی خوبی برای دانشجویان کارشناسی ارشد مدیریت فناوری اطلاعات، مهندسی فناوری اطلاعات، مهندسی نرم‌افزار و مهندسی صنایع باشد. همچنین این کتاب برای سازمان‌های کشور در حکم نقشه راه و راهنمای ورود به عصر پیاده‌سازی و تحلیل‌های عظیم داده می‌باشد.

گستره علم حوزه‌ای متغیر و پر جنب و جوش است و در این حوزه دانش منابع علمی به سرعت در حال تغییر و دگرگونی می‌باشند. از تمامی صاحب‌نظران و محققان در این حوزه دعوت می‌نمائیم که با پیشنهادها و توصیه‌های ارزنده خود ما را در ارائه آثار بعدی در این گستره یاری نمایند.

به امید ایرانی آگاه و آباد

دکتر سعید روحانی<sup>۱</sup> - سمیه حسینی

پاییز ۱۳۹۴

---

<sup>1</sup> SRouhani@ut.ac.ir

## پیشگفتار مولف

اخیرا، مفاهیم "عظیم داده" و "تحلیل‌های عظیم داده" تقریبا در همه جا به چشم می‌خورد. در هنگام مشاهده وبسایت‌ها، در حین مطالعه روزنامه و مجله می‌توان بدون اشاره مستقیم به این مفاهیم، آنها را به طور ضمنی مشاهده نمود. به هر حال فناوری‌های موجود و مجتمع شده در عظیم داده اعم از موازی‌سازی انبوه، حجم بسیار زیاد داده‌ها، توزیع داده، شبکه‌هایی با سرعت بالا، قابلیت محاسباتی بالا، مدیریت وظایف سیستمی، داده کاوی و تحلیل مفاهیمی جدید و تازه نمی‌باشند.

در طول اولین فاز از دوران کاریم، از اواخر دهه ۸۰ تا اوایل دهه ۹۰، من یک توسعه‌دهنده نرم‌افزار در شرکتی بودم که وظیفه تولید زبان برنامه‌نویسی کامپایلرها را برای ابرکامپیوتر بر عهده داشت. اکثر این سیستم‌های high-end سیستم‌هایی چندپردازنده بودند که موازی سازی‌های انبوه را به همراه مجموعه‌های بزرگی از داده‌ها به کار می‌برد. وظیفه اصلی من بهینه‌سازی کدینگ، با در نظر گرفتن افزایش پهنای باند داده و بهره‌برداری از مزیت‌های سلسله مراتب حافظه که بر اساس این سیستم‌های طراحی و پیاده سازی می‌گردید، بود. و قسمت جالب ماجرا این است که بسیاری از این معماری‌ها و فناوری‌های مورد استفاده جهت طراحی سخت افزار و توسعه نرم‌افزار جدید نبوده، و بسیاری از آنها به ابرکامپیوترهایی نظیر Illiac IV (اولین سیستم‌های محاسباتی موازی حجیم که اولین با در ده ۷۰ توسعه یافتند) مربوط می‌باشد.

به همین دلیل است که پدیده عظیم داده برای من بسیار جذاب و جالب می‌باشد، زیرا فناوری جدیدی نبوده و بیشتر به چگونگی استفاده از فناوری‌های شناخته شده و بکارگیری آنها پرداخته می‌شود. هنگامی که جزییات فناوری که از ۲۰ سال گذشته ظهور کرده در نیویورک تایمز، وال استریت ژورنال و اکونومیست چاپ شدند تازه آن وقت شما متوجه خواهید شد که آن فناوری بلاخره رسیده است.

## چالش‌های اتخاذ فناوری جدید

بسیاری از مردم به طور طبیعی نسبت به فناوری جدید دارای کشش و جاذبه هستند. بطوری که می‌توان گفت، آنها به فناوری جدید به صورت برق و درخشش دکمه نقره‌ای رنگی درون خاک می‌نگرند که نه تنها تصور می‌کنند که برای سازمان مفید خواهد بود؛ بلکه با کاوش بیشتر و حفاری بیشتر می‌توانند از طریق آن دکمه به جریان و معادل طلا، برای حل کلیه مشکلات سازمان خود در حال و آینده دست یابند. در این گونه سازمان‌ها برای اتخاذ فناوری، انقلاب خاصی صورت نمی‌گیرد و همواره ترس طولانی و بلند مدتی نسبت به دور افتادن از مسیر تعالی به علت عدم اتخاذ و بکارگیری فناوری وجود دارد. حتی در برخی اوقات ممکن است دورنمای شفاف‌تری نیز از ارزش پیشنهادی حاصل از بکارگیری فناوری وجود نداشته باشد اما این احساس در بین افراد و کارکنان سازمان وجود دارد.

واضح است که، عاقلانه نیست که بدون ارزیابی مولفه‌های ارزشی هر فناوری جدید و بدون بررسی و مقایسه هزینه نهایی و کلی عملیات، مبادرت به پیاده‌سازی و بکارگیری فناوری‌های جدید نمود. ضرورتاً، آزمایش، تست فناوری‌های جدید برای حفظ و بقا، توان رقابتی و اطمینان از امکان‌سنجی و امکان‌پذیر بودن پروژه به لحاظ فنی بسیار حیاتی خواهد بود. اما در بسیاری از سازمان‌ها، فرایند سریع و پیشبرنده فنون و ابزارهای جدید، اغلب منجر به کنارگذاری برنامه‌های نظارتی و تجربیات برتر در سازمان می‌گردد. نتیجه این خواهد بود که، پروژه‌های پایلوتی که قبل از بررسی کامل به وضعیت "تولید (نهایی)" می‌رسند، راه حلی را برای جزایری از داده‌ها که از دیدگاه و چشم‌انداز عملکردی منطبق با عملکرد سازمان و معماری آن نمی‌باشند ارائه نمی‌کنند.

## هدف از نگارش این کتاب

هدف این کتاب آن است تا بستری را در سازمان برای شکل‌گیری استراتژی‌های اتخاذ فناوری‌های عظیم داده فراهم نمایند. به عبارت دیگر، مقصود این کتاب ارائه دید کلی در خصوص چستی عظیم داده و چرایی ارزش آفرینی آن برای سازمان و انواع مسائل و مشکلاتی که در سازمان می‌توان توسط عظیم داده حل و فصل نمود و چگونگی برنامه‌ریزی کامل و همه جانبه جهت تعیین نیازها، بکارگیری افراد درست و مناسب در سازمان و توسعه و

ایجاد یک برنامه استراتژیک برای ایجاد یکپارچگی می‌باشد.

این کتاب به بیان "چگونه می‌توان" برنامه‌های کاربردی عظیم داده، برنامه‌نویسی MapReduce یا پیاده‌سازی Hadoop را توسعه و اجرایی نمود، می‌پردازد. در حقیقت تمایل نویسنده آن است تا دورنمای کلی از برخی جنبه‌های مربوط به اکوسیستم یا فرایند اتخاذ و بکارگیری عظیم داده را در هر فصل برای خوانندگان به شرح زیر ارائه نماید:

- فصل اول: ما به بررسی شرایطی از بازار که می‌تواند منجر به پذیرش وسیع و کلی نسبت به تحلیل‌های عظیم داده گردد می‌پردازیم. این شرایط عبارت خواهند بود از: ایجاد و تولید محصولات نرم‌افزاری و سخت‌افزاری، افزایش حجم داده، گسترش تنوع در انواع دارایی‌های دانشی جهت تحلیل‌ها، متدها و روش‌های متفاوت در تحویل و ارائه داده و در نهایت انتظارات روبه رشد جهت رسیدن به یکپارچگی بهنگام و بلادرنگ نسبت به نتایج تحلیل‌ها و فرایندهای عملیاتی.
- فصل دوم: در این فصل، ما به مشخصات و ویژگی‌های مشکلات و مسائل کسب و کار که از گذشته نیازمند منابع بوده و منجر به گسترش دامنه کسب و کار شده و برای حل آنها می‌توان از بسترها و پلتفرم‌های عظیم داده (یا حتی سخت‌افزارهای تخصیص داده شده یا مجازی سازی / مبتنی بر ابر) استفاده نمود پرداخته می‌شود.
- فصل سوم: در این فصل به این موضوع پرداخته می‌شود که، چه کسانی در سازمان باید در فرایند اکتساب، اثبات و استقرار راه‌حل‌های عظیم داده در سازمان شرکت جوید؟ مسئولیت‌ها و نقش‌های افراد مزبور چگونه تعریف می‌شوند؟ در این فصل به اتخاذ فناوری جدید و چگونگی همراستایی سازمان برای یکپارچگی و مطابقت با چرخه عمر توسعه سیستم پرداخته خواهد شد.
- فصل چهارم: این فصل ادامه فصل قبلی بوده و برخی موارد کلیدی که اغلب در اتخاذ فناوری جدید مطرح هستند و نشان می‌دهند که موضوعات کلیدی مزبور جدید نبوده و تقریباً بخشی از دانش سازمانی هستند که می‌تواند در پروراندن برنامه استراتژیک معقول و منطقی تاثیرگذار باشد، تشریح می‌گردد.

- فصل پنجم: در این فصل ما به دنبال بینش، درک و نظارت بر داده مخصوصاً هنگامی که توسعه برنامه‌های کاربردی عظیم داده، منجر به کنارگذاشتن فناوری اطلاعات سنتی و کانال‌های سنتی مدیریت داده قدیمی می‌گردد، خواهیم بود.
- فصل ششم: در این فصل ما به طراحی سخت‌افزارهای لازم برای اجرای تحلیل‌ها و چگونگی مهندسی سخت‌افزارهای مزبور برای پذیرش مجموعه داده‌های بزرگ می‌پردازیم.
- فصل هفتم: این فصل به بررسی و تشریح درونمایی کلی و سطح بالایی از ابزارها، نظیر Hadoop می‌پردازد.
- فصل هشتم: در این فصل مدل برنامه نویسی MapReduce مورد بررسی و تجزیه و تحلیل قرار خواهد گرفت.
- فصل نهم: در این فصل، به تشریح گستردگی روش‌های مدیریت داده پرداخته و روش‌هایی را که برای توسعه برنامه‌های کاربردی عظیم داده مناسب هستند، معرفی می‌گردد.
- فصل دهم: در این فصل به مسائل و مشکلاتی از کسب و کاری که تحلیل‌های گراف، راه حل آنها هستند، پرداخته شده و آنچه را که منجر به تمایز مسائل در رویکردهای سنتی شده و ملاحظات را که باید برای اکتشاف در مقابل تحلیل‌های جستجو<sup>1</sup> در نظر گرفت، بررسی می‌گردند.
- فصل یازدهم: در این فصل کوتاه و پایانی، تجربیات برتر برای اتخاذ عظیم داده در سازمان تدریجاً بررسی و بازنگری می‌گردد.

---

<sup>1</sup> Search analyses

## چرا باید به مطالعه این کتاب پرداخت؟

- احتمالا شما به یک یا تعدادی از دلایل خوب و منطقی زیر باید این کتاب را مطالعه نمایید:
- اگر شما یک مدیر ارشد بوده و به دنبال جستجوی مزایا از اطلاعات سازمان خود جهت خلق یا افزودن ارزش به سازمان از طریق افزایش درآمد، کاهش هزینه‌ها، بهبود بهره‌وری، کاهش ریسک یا بهبود تجربیات مشتریان می‌باشید.
  - اگر شما یک مدیر ارشد اطلاعات<sup>۱</sup> یا یک مدیر ارشد داده در یک سازمان باشید که نیازمند استفاده بهینه از دارایی‌های اطلاعاتی سازمان خود باشد.
  - اگر شما مدیری باشید که از او خواسته شده تا به توسعه یک برنامه عظیم داده مبادرت ورزد.
  - اگر شما مدیری باشید که از او خواسته شده تا چیزی بیش از اطلاعات اولیه در خصوص برنامه‌های کاربردی عظیم داده ارائه نماید.
  - اگر شما مدیری باشید که از او خواسته شده باشد تا در اجرای یک برنامه عظیم داده موفق گردد.
  - اگر شما یک مدیر ارشد کسب و کار باشید که موظف است تا ارزشی را که عظیم داده می‌تواند به سازمان شما بدهد را مشخص و کشف نماید.
  - اگر شما بعنوان یکی از پرسنل سازمان علاقمند به درک و کسب بینش عمیق، نسبت به راهی که کسب و کار شما در حال گذر از آن است باشید.
  - اگر شما یک مهندس پایگاه داده یا نرم‌افزار بوده و برای مدیریت فنی برنامه عظیم داده انتخاب و تعیین شده باشید.
  - اگر شما یک مهندس نرم‌افزار بوده و تمایل داشته باشید که برای مدیریت برنامه عظیم داده انتخاب شوید.
  - اگر شما تحلیلگر مهندسی باشید که بر روی چارچوب عظیم داده کار می‌کند و مایل به جایگزینی با مدیر فعلی خود هستید.

---

<sup>۱</sup> Chief Information Officer

- اگر شما یک تحلیلگر کسب و کار بوده و از او خواسته شده باشد تا به تیم برنامه کاربردی عظیم داده ملحق گردد.
- اگر شما یک مدیر ارشد باشید و مدیران مستقیم تحت نظارت شما صحبت در خصوص عظیم داده با واژگان خاص آن را آغاز کرده و شما فکر می‌کنید که آنها از شما انتظار دارند که متوجه منظورشان شوید.
- شما یک مدیر میانی یا یک مهندس هستید و مدیر شما صحبت در خصوص عظیم داده با واژگان خاص آن را آغاز کرده و شما فکر می‌کنید که او از شما انتظار دارد تا متوجه منظورش شوید.
- شما فقط مایل و مشتاق نسبت به مطالعه در زمینه عظیم داده هستید.

من چقدر از شما می‌دانم و واقعا چقدر شما را می‌شناسم؟ من در بسیاری از مراحل زندگی در حقیقت مثل شما بوده‌ام به طوری که مثلاً یا بر روی یک پروژه کار کردم یا مدیر یک پروژه بوده‌ام که خلاهای دانشی زیادی در آن حوزه داشته‌ام، یا در سازمانی کار کرده‌ام که پر بوده از افرادی که مطمئن نبودند به کجا خواهند رسید، چه کاری انجام می‌دهند. در آن شرایط من واقعاً نیازمند کتاب راهنمایی بودم که بتوانم بدون صرف هزینه زیاد که اغلب برای کتاب‌های فنی و تخصصی باید پرداخت، بدان رجوع کرده و بتوانم یک نگاه کوتاه یا شاید هم یک مطالعه طولانی داشته باشم.

ما به طور مرتب از سوی فروشندگان در خصوص تولید انواع تحلیل‌ها و رهبری در بسیاری از حوزه‌های مدیریت داده مورد پرسش و درخواست قرار می‌گرفتیم. در بخش دولتی برای بخش فدرال و دولت و سایر آژانس‌های جهانی دولتی مشاوره داده‌ایم. ما پروژه‌های مدیریت داده بسیاری را در صنایع مختلف اعم از دولتی، خدمات مالی، خدمات بهداشتی سلامتی، تولید، سرویس‌های انرژی، بیمه و خدمات اجتماعی و غیره راهنما و مشاور بودیم.

هنگامی که ما کار خود را آغاز نمودیم، آگاهی از ارزش مدیریت اطلاعات به عنوان یکی از مهمترین موضوعات در مدیریت ارشد خودنمایی می‌کرد. با ظهور فعالیت‌هایی که شامل



بهره‌برداری از داده‌های سازمان نظیر برنامه‌ریزی منابع سازمان<sup>۱</sup>، مدیریت زنجیره تامین<sup>۲</sup> و مدیریت ارتباط با مشتری<sup>۳</sup>، نیاز به نمایش تلفیقی از داده‌ها با کیفیت بالا از اطلاعات حیاتی کسب و کار بوجود آمد. با افزایش نظارت‌ها، نیاز به تبادل اطلاعات، مدیریت عملکرد کسب و کار و ارزش معماری مبتنی بر خدمات و حرکت به سمت تمرکز بیشتر بر مدیریت مبتنی بر عملکرد داده‌های سازمان با در نظر گرفتن، قابلیت دسترسی، ثبات، به روز بودن، تازگی و قابلیت استفاده از دارایی‌های اطلاعاتی مشترک ضروری می‌باشد.

از آنجا که نیت بنده در این کتاب ایجاد یک راهنما برای برنامه‌ریزی استراتژیک در حوزه عظیم داده می‌باشد، ممکن است شرایطی ایجاد شود که برخی متخصصین نیازمند کمک و راهنمایی بیشتر از سوی بنده باشند. تمرینات و رویکردهای تشریح شده در این کتاب خلاصه و چکیده‌ای از تجربیات واقعی کاربران و مشتریان بوده و تجربیات گسترده ما می‌تواند شما را در پرداختن به مأموریت خود در قالب استقرار برنامه کاربردی عظیم داده یاری رساند. با توجه به روحیه آزاداندیش بنده، من همواره مشتاق پاسخگویی به سوالات، ارائه توضیحات بیشتر و شنیدن بازخوردهای حاصل از پیاده‌سازی رویکردها و دانش یکپارچه در این کتاب که از سال ۱۹۹۹ بر اساس تجربیات موفق ما با مشتریان و کاربران خود در این کتاب جمع آوری شده است، می‌باشم.

ما همواره به دنبال فرصت‌هایی برای کمک به سازمان‌هایی هستیم که به دنبال کسب ارزش، توسعه طرح‌های اولیه، نقشه راه، طرح برنامه، کمک به پیاده‌سازی هوشمندی کسب و کار، استراتژی استفاده و بهره‌برداری از اطلاعات و هرگونه فرصت برای توزیع و به اشتراک‌گذاری ایده به سازمان‌ها هستیم. معنای حرفه‌ای بالای من، همین یک جمله است: من واقعا می‌خواهم که حرف‌های شما را بشنوم.

دیوید لوشین

<sup>۱</sup> Enterprise Resource Planning (ERP)

<sup>۲</sup> Supply Chain Management (SCM)

<sup>۳</sup> Customer Relationship Management (CRM)

# فهرست مطالب

۱	فصل ۱ محرک‌های بازار و کسب و کار برای تحلیل‌های عظیم داده
۱	۱-۱- واقعیت و فرضیات عظیم داده
۵	۲-۱- محرک‌های کسب و کار
۹	۳-۱- کاهش موانع جهت ورود
۱۱	۴-۱- ملاحظات
۱۲	۵-۱- تمرینات فکری
۱۵	فصل ۲ عظیم داده و راه‌حل‌های آن برای کسب و کار
۱۶	۱-۲- اعتبار (درمقابل) مبالغه: برازش سازمانی
۱۹	۲-۲- ترویج ارزش عظیم داده
۲۰	۳-۲- موارد استفاده از عظیم داده
۲۲	۴-۲- مشخصات برنامه‌های کاربردی عظیم داده
۲۵	۵-۲- ادراک و تشریح ارزش
۲۶	۶-۲- نگاه پیش‌تاز نسبت به ارزش
۲۶	۷-۲- تمرینات فکری
۲۷	فصل ۳ همراستایی سازمانی در جهت تحلیل‌های عظیم داده
۲۸	۱-۳- دو سوال کلیدی
۲۹	۲-۳- چشم انداز تاریخی مربوط به گزارش‌گیری و تحلیل
۳۰	۳-۳- چالش مقاومت‌های انسانی
۳۱	۴-۳- تشریح جنبه‌ها و ابعاد اتخاذ فناوری عظیم داده
۳۳	۵-۳- مشارکت تصمیم‌گیران مناسب
۳۴	۶-۳- نقش‌های سازمانی
۳۶	۷-۳- تمرینات فکری

## فهرست مطالب

۳۹	فصل ۴ راهبرد یکپارچه سازی تحلیل‌های عظیم داده با سازمان
۴۰	۱-۴ چه زمان فناوری عظیم داده برای سازمان شما مناسب است؟
۴۱	۲-۴ طرح راهبردی برای اکتساب فناوری
۴۱	۳-۴ تجارب استاندارد برای برآورد انتظارات کاربران
۴۳	۴-۴ پذیرش انطباق فناوری: تشریح شاخص‌های رد و تأیید
۴۴	۵-۴ تدارک محیط داده جهت مقیاس‌های انبوه
۴۶	۶-۴ ترویج استفاده مجدد از داده‌ها
۴۶	۷-۴ نظارت و مدیریت بر سطوح دسترسی
۴۷	۸-۴ فرایند حاکمیتی جریان گسترش فناوری
۴۸	۹-۴ ملاحظات در یکپارچگی سازمانی
۴۹	۱۰-۴ تمرینات فکری

۵۱	فصل ۵ حاکمیت داده
۵۱	۱-۵ تکامل و شکل‌گیری حاکمیت داده‌ها
۵۳	۲-۵ عظیم داده و حاکمیت داده
۵۴	۳-۵ تفاوت‌های مجموعه‌های عظیم داده
۵۶	۴-۵ نظارت بر عظیم داده: پنج مفهوم کلیدی
۵۷	۱-۴-۵ مدیریت انتظارات و توقعات مصرف‌کنندگان داده
۵۸	۲-۴-۵ تعریف و تعیین ابعاد حیاتی کیفیت داده
۵۹	۳-۴-۵ انسجام و یکپارچگی فراداده و داده مرجع برای استخراج موجودیت
۶۰	۴-۴-۵ هدف‌گذاری و تفسیر مجدد
۶۱	۵-۴-۵ بهبود و غنی‌سازی داده
۶۳	۵-۵ ملاحظات
۶۳	۶-۵ تمرینات فکری

# فهرست مطالب

۶۵	فصل ۶ معرفی برنامه‌های کاربردی مدیریت عظیم داده
۶۷	۱-۶ موارد استفاده
۶۹	۲-۶ ملاحظات ذخیره سازی: زیر ساخت جهت مدیریت چرخه عمر داده
۷۱	۳-۶ تجهیزات عظیم داده
۷۲	۴-۶ گزینه‌های معماری
۷۳	۵-۶ در نظر گرفتن مشخصات عملکردی
۷۶	۶-۶ چیدمان داده مبتنی بر ردیف-در مقابل - ستون و عملکرد نرم افزار
۸۰	۷-۶ ملاحظات مربوط به گزینه‌های پلتفرم
۸۱	۸-۶ تمرینات فکری
۸۳	فصل ۷ روش‌ها و ابزارهای عظیم داده
۸۳	۱-۷ درک و شناخت ذخیره‌سازی عظیم داده
۸۴	۲-۷ نگاه کلی به معماری با قابلیت بالا
۸۶	۳-۷ سیستم‌های فایلی توزیع شده Hadoop
۸۹	۴-۷ MAPREDUCE و YARN
۹۱	۵-۷ ارتقاء اکوسیستم برنامه‌های کاربردی عظیم داده
۹۱	۶-۷ ZOOKEEPER
۹۲	۷-۷ HBASE
۹۳	۸-۷ HIVE
۹۴	۹-۷ PIG
۹۵	۱۰-۷ MAHOUT
۹۶	۱۱-۷ ملاحظات
۹۸	۱۲-۷ تمرینات فکری

## فهرست مطالب

۹۹	فصل ۸ توسعه برنامه‌های کاربردی عظیم داده
۹۹	۱-۸- موازی‌سازی
۱۰۰	۲-۸- افسانه مقیاس‌پذیری آسان
۱۰۱	۳-۸- چارچوب توسعه برنامه کاربردی
۱۰۲	۴-۸- مدل برنامه‌نویسی MAPREDUCE
۱۰۴	۵-۸- یک مثال ساده
۱۰۵	۶-۸- توضیحات بیشتر در خصوص MAP REDUCE
۱۰۷	۷-۸- سایر چارچوب‌های توسعه عظیم داده
۱۰۸	۸-۸- مدل اجرایی
۱۱۰	۹-۸- تمرینات فکری
۱۱۱	فصل ۹ عظیم داده و NoSQL
۱۱۱	۱-۹- NOSQL
۱۱۲	۲-۹- "مدل‌های Schema-Less": افزایش انعطاف‌پذیری جهت دستکاری داده
۱۱۳	۳-۹- مخازن کلید-ارزش
۱۱۶	۴-۹- مخازن مستندات
۱۱۷	۵-۹- مخازن جدولی
۱۱۸	۶-۹- مخازن داده‌ای شی‌گرا
۱۱۸	۷-۹- پایگاه‌های داده نموداری
۱۱۸	۸-۹- ملاحظات
۱۱۹	۹-۹- تمرینات فکری
۱۲۱	فصل ۱۰ تحلیل‌های گراف در عظیم داده
۱۲۲	۱-۱۰- تحلیل‌های گراف

## فهرست مطالب

۱۲۳	۲-۱۰- سهولت مدل گراف
۱۲۵	۳-۱۰- بازنمایی سه گانه
۱۲۶	۴-۱۰- گراف‌ها و ساختار شبکه
۱۲۷	۵-۱۰- انتخاب تحلیل‌های گراف
۱۲۸	۶-۱۰- موارد استفاده از تحلیل‌های گراف
۱۳۰	۷-۱۰- الگوریتم‌های تحلیلی گراف و رویکردهای حل مساله
۱۳۲	۸-۱۰- پیچیدگی‌های فنی تحلیل گراف‌ها
۱۳۴	۹-۱۰- ویژگی‌های یک پلتفرم تحلیلی گراف
۱۳۷	۱۰-۱۰- ملاحظات: برنامه‌های کاربردی تحلیل گراف
۱۳۸	۱۱-۱۰- تمرینات فکری

### فصل ۱۱ نقشه راه عظیم داده

۱۳۹	۱-۱۱- مقدمه
۱۴۰	۲-۱۱- طوفان ذهنی: ارزیابی نیازمندی و ارزش عظیم داده
۱۴۱	۳-۱۱- توافق سازمانی
۱۴۲	۴-۱۱- تیم‌سازی
۱۴۷	۵-۱۱- اثبات یک مفهوم
۱۴۹	۶-۱۱- ارزیابی فناوری و انتخاب اولیه
۱۵۳	۷-۱۱- فرایند توسعه، آزمایش و پیاده‌سازی برنامه کاربردی
۱۵۴	۸-۱۱- دامنه پروژه و پلتفرم
۱۵۵	۹-۱۱- برنامه یکپارچگی تحلیل‌های عظیم داده
۱۵۶	۱۰-۱۱- مدیریت و نگهداری
۱۵۷	۱۱-۱۱- ارزیابی
۱۵۸	۱۲-۱۱- خلاصه و ملاحظات
۱۶۰	۱۳-۱۱- تمرینات فکری