



داده محوری؛ از مدیریت تا حکمرانی داده

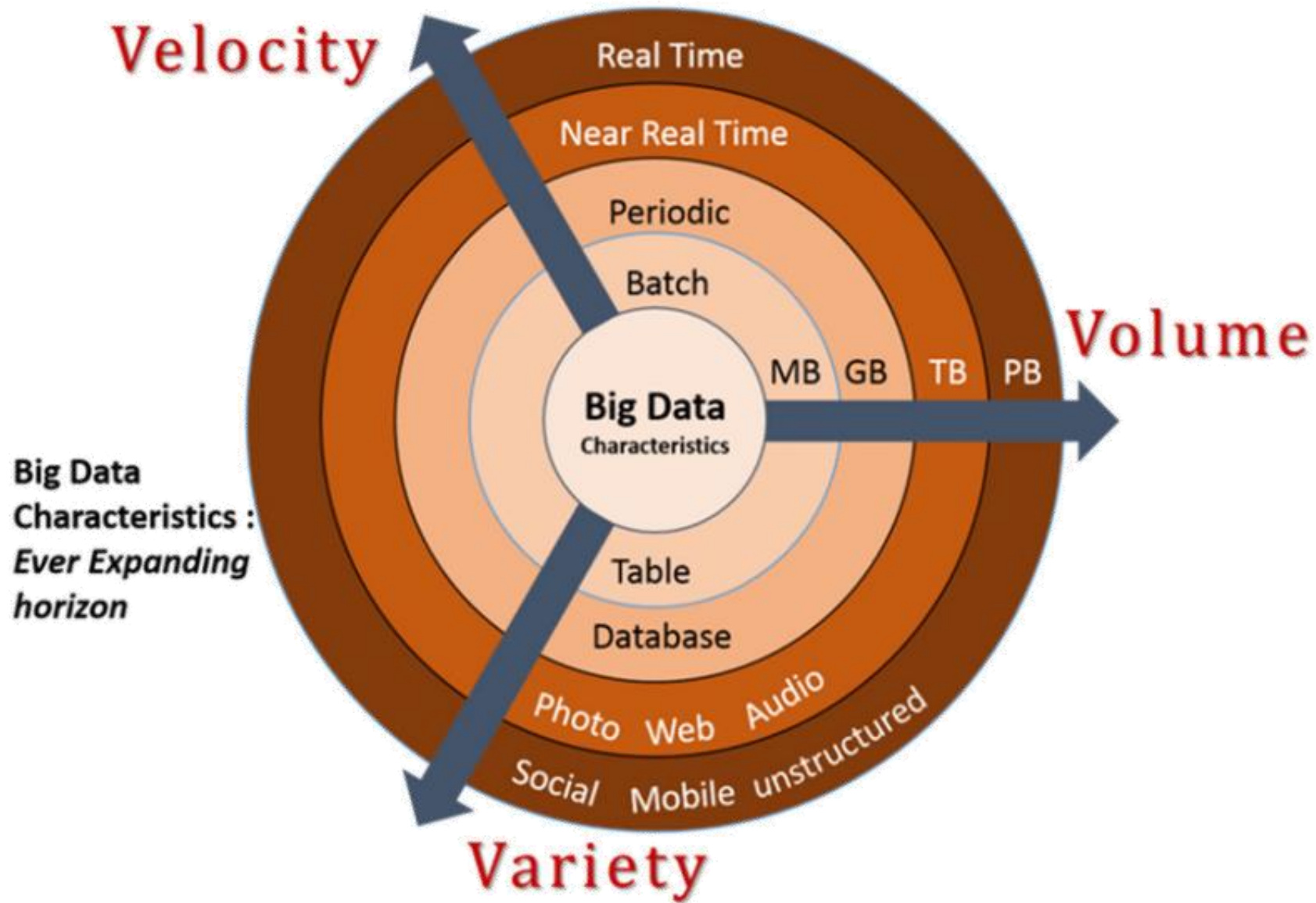
ارائه دهنده: محمدرضا محتاط

رئوس مطالب

- سازمان داده محور
- چالش های داده محوری
- مدیریت داده
- کسب و کار هوشمند
- طراحی و مدل سازی داده
- ذخیره سازی و عملیات داده
- داده مرجع و اصلی
- فراداده
- معماری داده
- یکپارچگی داده
- کیفیت داده
- مدیریت محتوا و اسناد
- امنیت داده
- حکمرانی داده
- ارزیابی بلوغ داده محوری
- جمع بندی

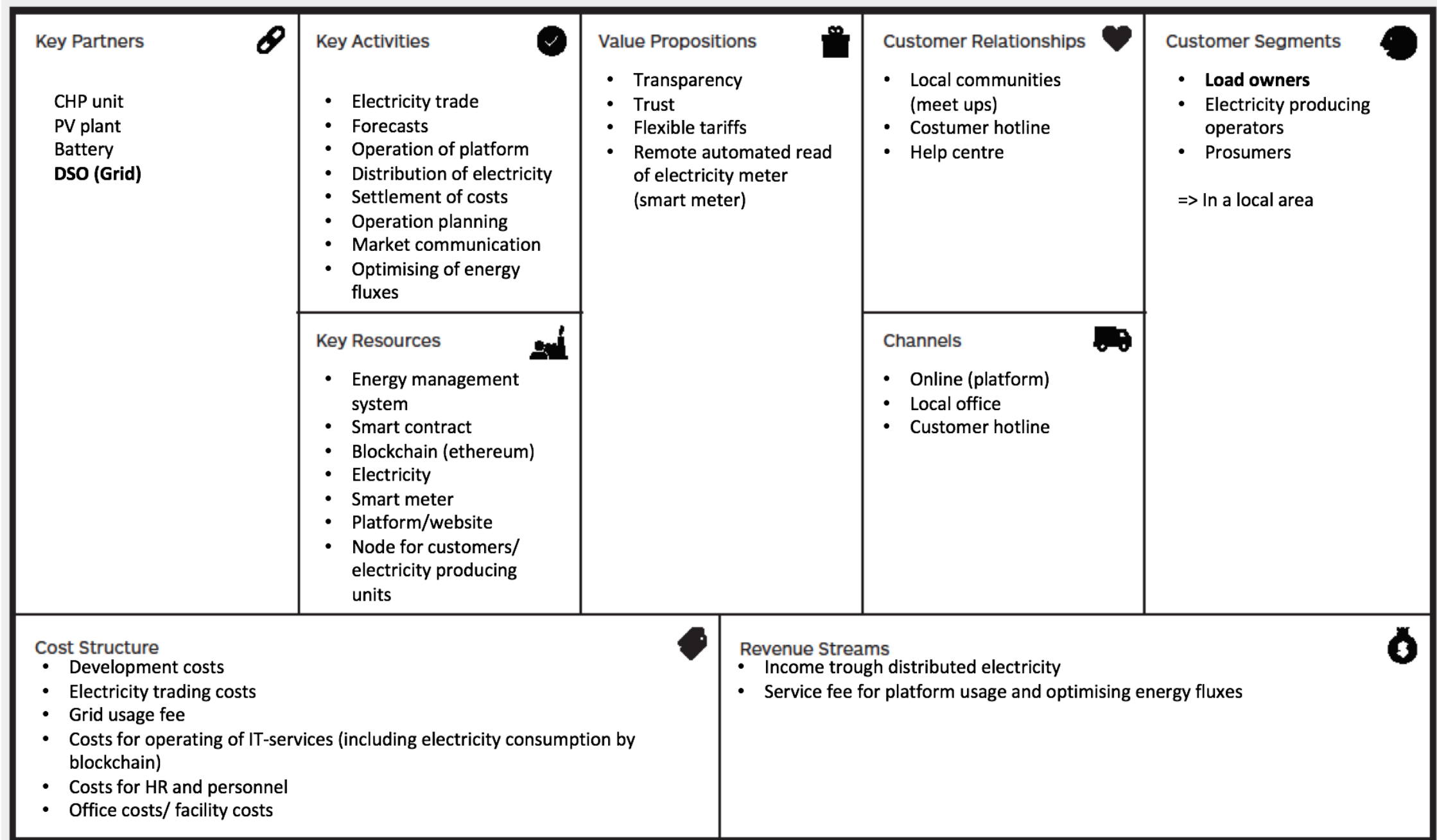


مقدمه



منابع داده‌ای کسب و کار

● ارزش مالی یک شرکت داده محور مبتنی بر اینترنت اشیا
در صنعت خانه هوشمند چگونه محاسبه میشود؟



داده بعنوان یک دارایی راهبردی

تمام نشدنی

بادوام

استراتژیک

هزینه تولید مجدد

ارزش بازار

بهای تمام شده

(اکتساب تا ذخیره سازی و بهره‌برداری)

هزینه از دست رفتن اطلاعات

هزینه فرصت استفاده از داده‌ها جهت خلق
ایده‌های نوآورانه

چالش‌های شرکت‌های ایرانی در عصر داده‌ها!

- دیدگاه سنتی در خلق ارزش (فرهنگ تصمیم‌گیری داده‌محور)
- مقاومت سازمانی کارشناسان و مدیران
- جذب استعدادهای مدیریت داده / تحلیل داده
- کیفیت داده
- تمرکز بیش از حد فنی
- حریم خصوصی، محرمانگی و امنیت داده
- مدیریت و حکمرانی داده
- داده جریانی و لحظه‌ای
- محدودیت الگوریتم‌های یادگیری ماشین
- فقدان داده

سازمان داده محور



نقش‌های متنوع داده‌ای در یک سازمان داده‌محور چیست؟



مدیریت داده

رسوایی بزرگ بانکی در آمریکا؛ اطلاعات ۱۰۰ میلیون نفر لو رفت

بانک «کپیتال وان فیننشال کراپ» از هک شدن اطلاعات خصوصی (شامل آدرس و نام) ۱۰۰ میلیون آمریکایی و ۶ میلیون کانادایی توسط هکری که در حال حاضر دستگیر شده، خبر داد.

پایگاه خبری تحلیلی انتخاب (Entekhab.ir):



چرایی مدیریت داده

چرایی مدیریت داده



اطلاعیه مهم شرکت آسیاتک درمورد حملات سایبری اخیر به سرورهای شرکت ابرآروان

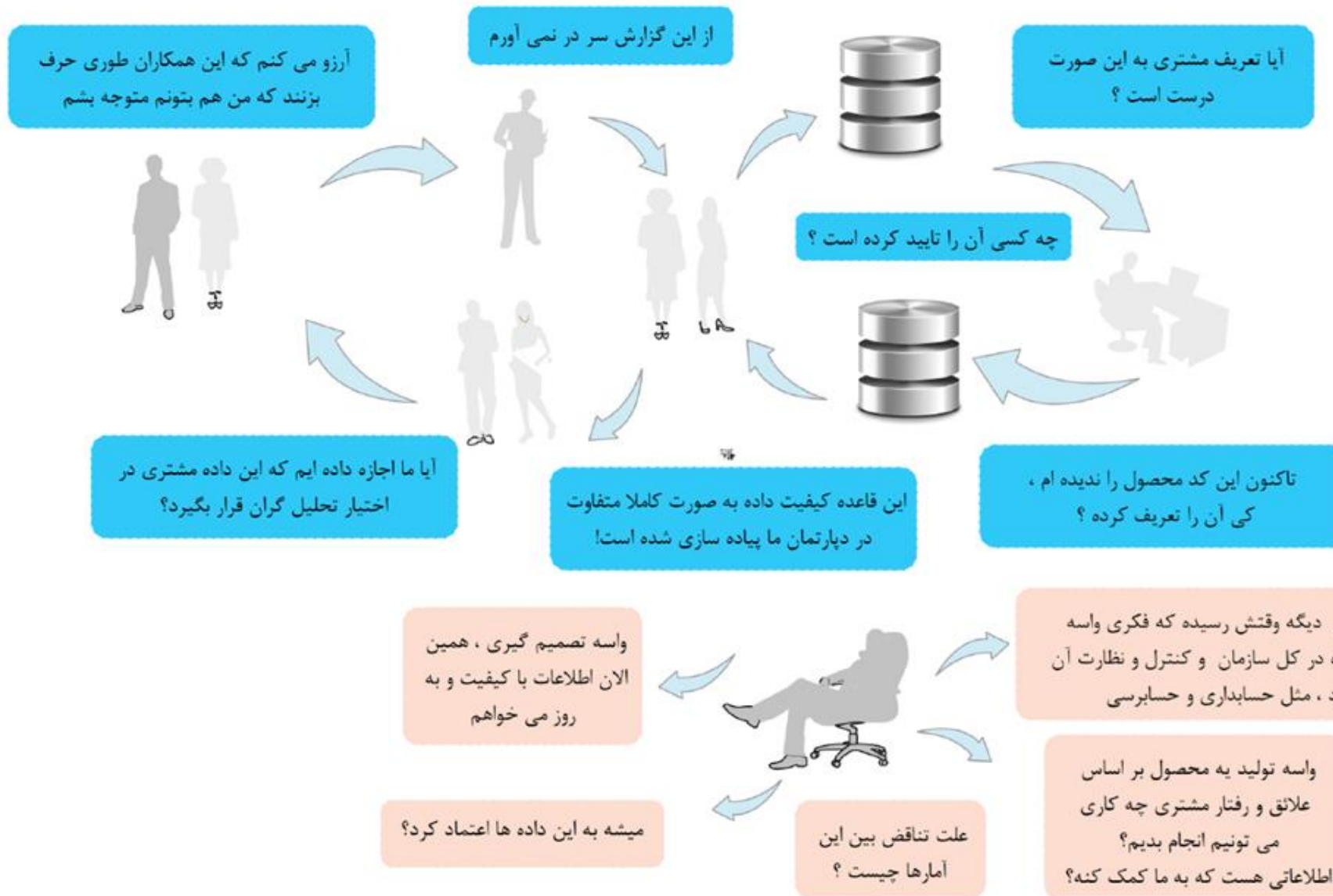
چرایی مدیریت داده

asiatech.ir



1544
91010000

چرایی مدیریت داده





سوال

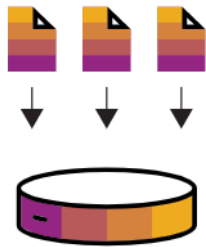
تعریف یک جمله‌ای از مدیریت داده!



مدیریت داده

تعریف: برنامه‌ریزی، اجرا، نظارت بر سیاست‌ها، فعالیت‌ها و پروژه‌هایی که ارزش سرمایه‌های داده‌ای و اطلاعاتی را کنترل، محافظت و افزایش می‌دهند.

رسالت: در دسترس‌پذیری داده‌ها برای ذی نفعان (مجاز)



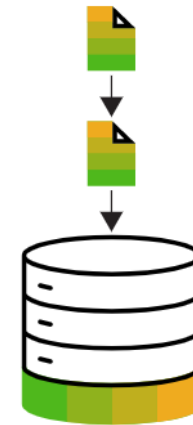
Data collecting



Data processing



Data validation



Data storage



اهداف پروژه مدیریت داده

	درک و پشتیبانی از نیازهای اطلاعاتی سازمان و ذی نفعان شامل مشتریان، کارکنان و شرکای تجاری
	کسب، ذخیره سازی و محافظت از دارایی‌های داده‌ای
	اطمینان و بهبود مداوم کیفیت داده و اطلاعات
	حفاظت و اطمینان از یکپارچگی دارایی‌های اطلاعاتی
	اطمینان از محرمانگی و حفظ حریم شخصی داده‌های ذی نفعان
	جلوگیری از دسترسی، دستکاری و استفاده غیرمجاز و نامناسب از داده‌ها
	اطمینان از بکارگیری موثر از داده‌ها جهت خلق ارزش برای سازمان‌ها

اصول مدیریت داده

داده یک دارایی با
ویژگی های
منحصربفرد است

مدیریت داده
مستلزم برنامه ریزی است

نوع های مختلف داده ، دارای نیازهای
چرخه عمر مختلف می باشند

ارزش داده می تواند و
می بایست به لحاظ اقتصادی
بیان شود

مدیریت داده ، مستقل از یک دپارتمان
بخصوص می باشد و وظیفه ای فرا دپارتمانی
است که نیازمند مجموعه ای
از مهارت و تخصص ها است

مدیریت داده
شامل مدیریت ریسک های
مرتبط با داده است

مدیریت داده
به معنی
مدیریت کیفیت داده است

مدیریت داده
نیازمند
یک دیدگاه سازمانی است

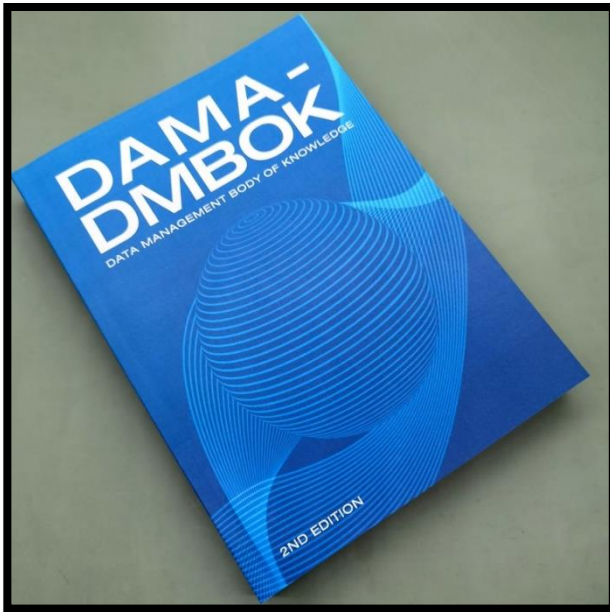
نیازهای مدیریت داده ،
می بایست محرک تصمیم گیری
فناوری اطلاعات باشد

از متادیتا جهت
مدیریت داده
استفاده می شود

مدیریت داده ،
مدیریت چرخه حیات داده است

مدیریت موثر داده
نیازمند
تعهد رهبری است

چيست DMBOK



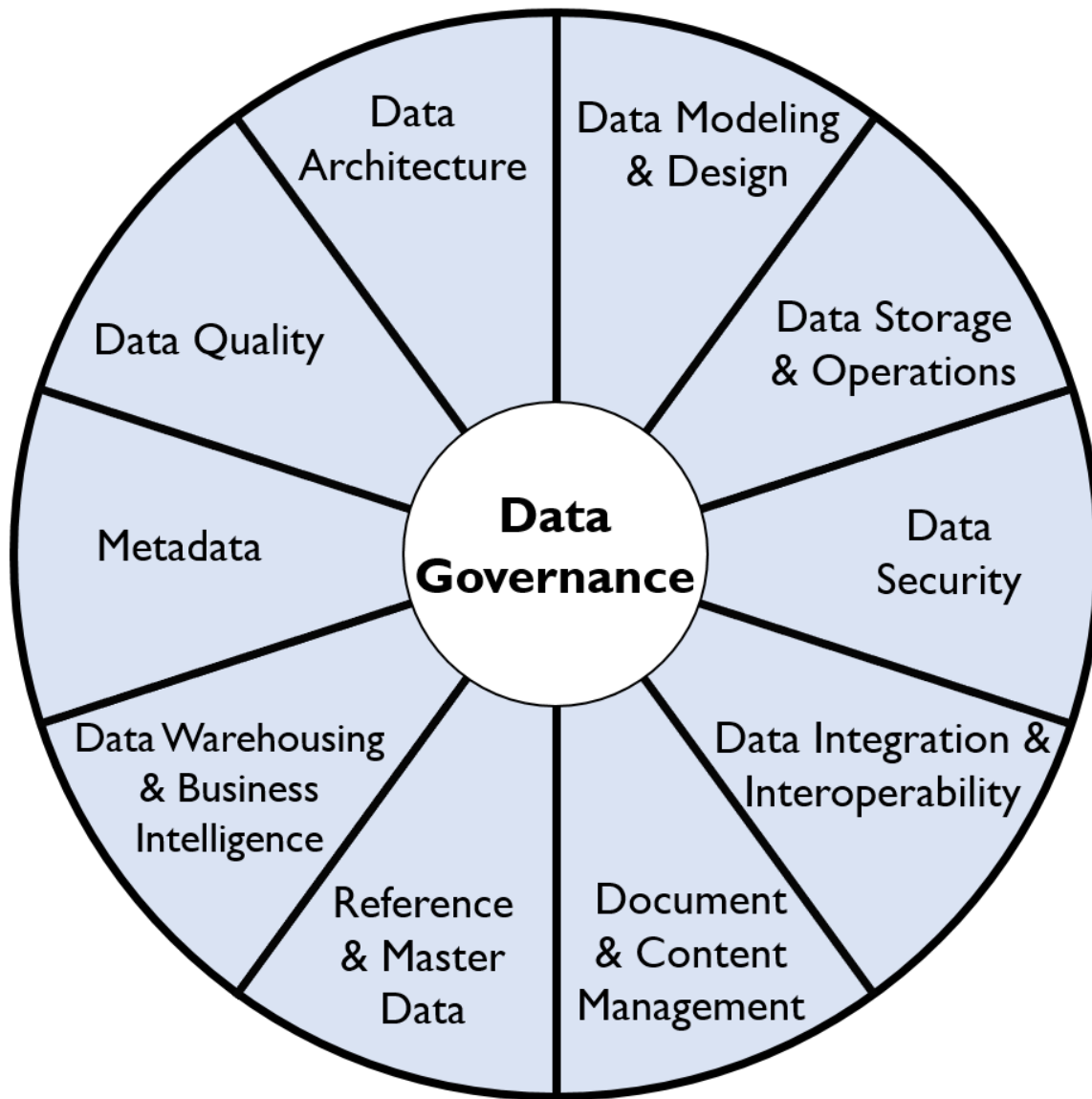
□ DAMA یک سازمان غیرانتفاعی بین المللی است که حدود ۷۵۰۰ عضو از ۴۰ شعبه در جهان دارد. هدف آن ارتقای و توسعه رویکردهای و مکانیزمهای مدیریت داده و اطلاعات به جهت حمایت از استراتژی های کسب و کار می باشد.

□ DMBOK یک راهنمای جامع برای مدیریت و حکمرانی داده است.



مخاطبان DAMA-DMBOK





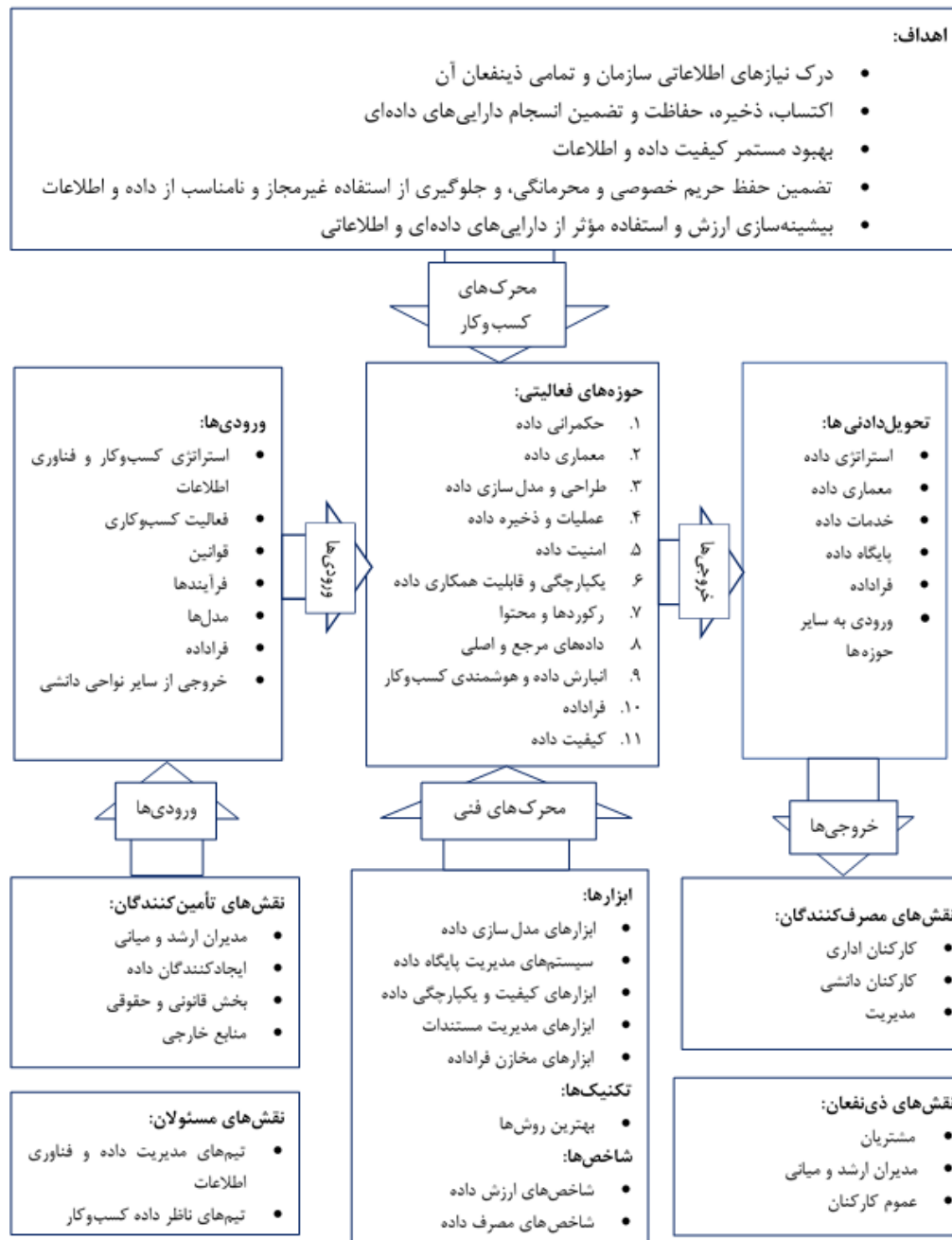
چارچوب DAMA DMBOK

Copyright © 2017 DAMA International



چارچوب DAMA DMBOK

معماری پروژه منطبق با چارچوب DAMA



خروجی پروژه مدیریت داده

چشم انداز مدیریت داده



طرح تجاری مدیریت داده



سنجش‌های پیشنهادی موفقیت مدیریت داده



اهداف کوتاه مدت برنامه مدیریت داده



توصیف ساختار و نقش‌های مدیریت داده به همراه خلاصه مسئولیت



برنامه اولویت بندی شده از کارها و محدوده مشخص آنها



نقشه پیاده سازی به همراه لیست اقدامات و پروژهها



PRODUCTIVITY

With good data management, your company will be more organized and productive. Employees will have an easier time finding, understanding, and relaying information.



COST EFFICIENCY

Data management can help your organization avoid unnecessary extra costs such as unneeded duplication. When data is easily accessible, You won't have to worry about employees conducting the same research over and over again.



OPERATIONAL NIMBLENESS

Great data management makes it easy for companies to respond quickly to the world around them. This means companies can respond efficiently to market changes and react appropriately to competitors.

WHY IS DATA MANAGEMENT IMPORTANT?



Source: <http://www.blue-pencil.ca/what-is-data-management-and-why-it-is-important/>



SECURITY RISKS

Proper data management helps ensure that your information stays secure and never ends up in the wrong hands. A strong data management system will help protect your information from theft and attacks.



REDUCED DATA LOSS

With a data management plan in place, you greatly reduce the risk of losing vital company information. It also ensures your important information is backed up and retrievable in case something happens to the original copies.



ACCURATE DECISIONS

Proper data management helps ensure all employees and workers view and analyze the same, most recent information. This helps ensure that your company will be making the most accurate decisions based on the most accurate information.

مزایا مدیریت داده



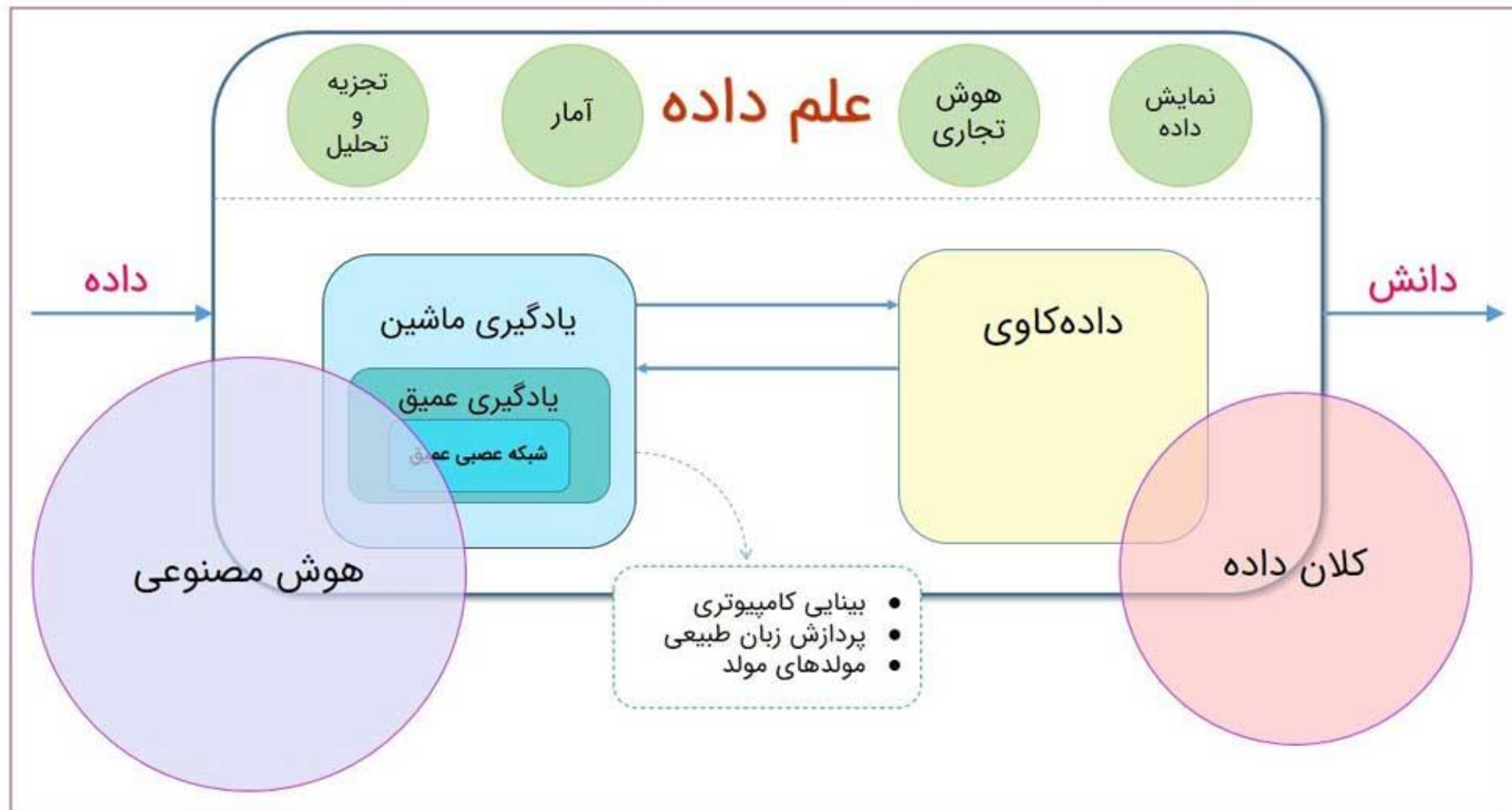
کسب و کار هوشمند

کسب و کار هوشمند

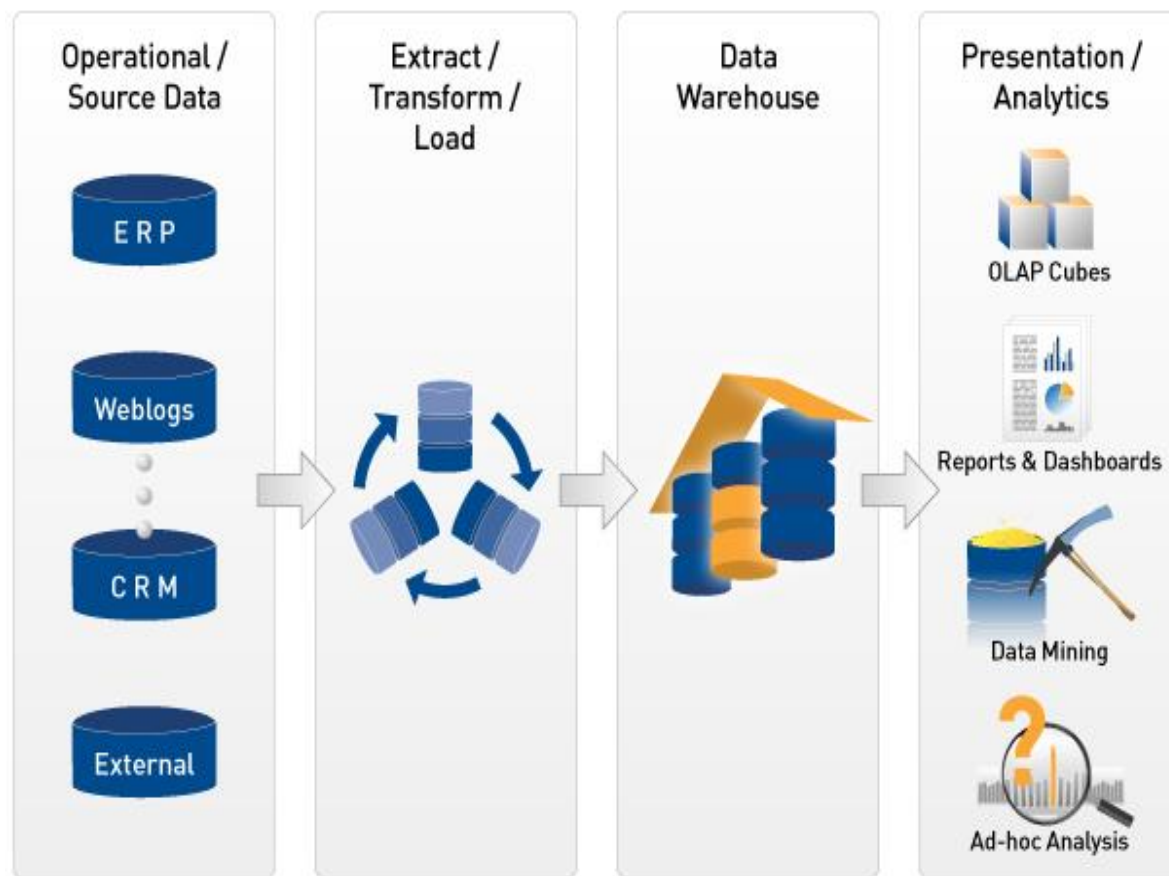
تعریف: برنامه ریزی و اجرای اقداماتی جهت ارائه سیستم تصمیم یار مدیران جهت کمک به تصمیم گیری های دقیق تر بر مبنای داده و اطلاعات

اهداف:

- فراهم آوردن زیرساخت فنی لازم جهت تجميع اطلاعات به منظور استفاده در سیستم های تصمیم یار یا سایر کارکردهای سازمان.
- فراهم آوردن امکان استفاده از سیستم های تصمیم گیری مبتنی بر داده برای مدیران سازمان به منظور افزایش سهولت و دقت و کارآمدی تصمیم گیری ها

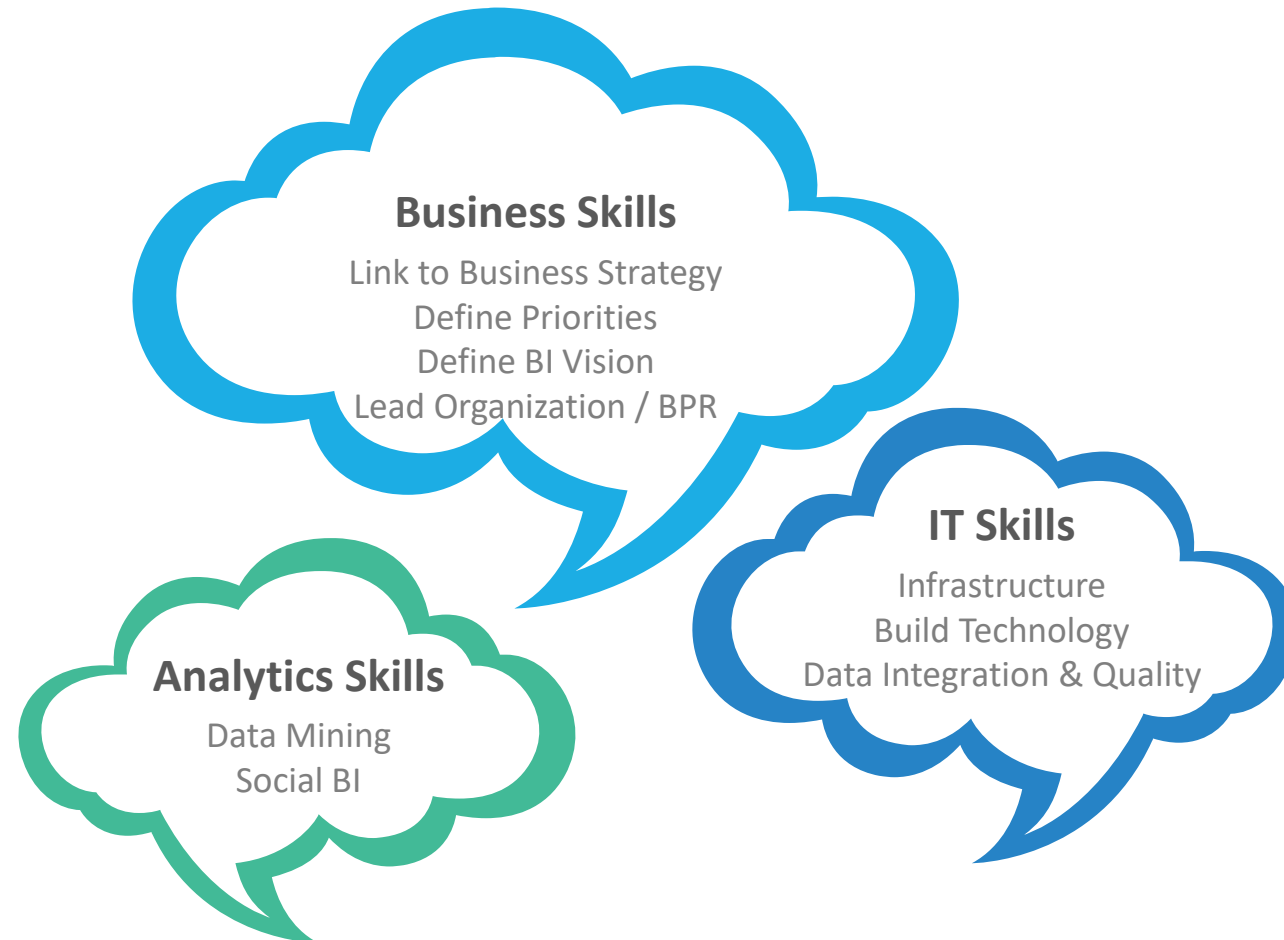


هوش تجاری



✓ هوش تجاری، نه بعنوان یک محصول و نه بعنوان یک سیستم، بلکه بعنوان یک معماری موردنظر است که شامل مجموعه ای از برنامه های کاربردی و تحلیلی است که به استناد پایگاه های داده عملیاتی و تحلیلی به اخذ و کمک به اخذ تصمیم برای فعالیت های هوشمند تجاری می پردازند.

دامنه پروژه‌های هوش تجاری





Executive Metrics Dashboard Share Dashboard

Ask a question about the data on this dashboard

Expected Revenue
THIS YEAR

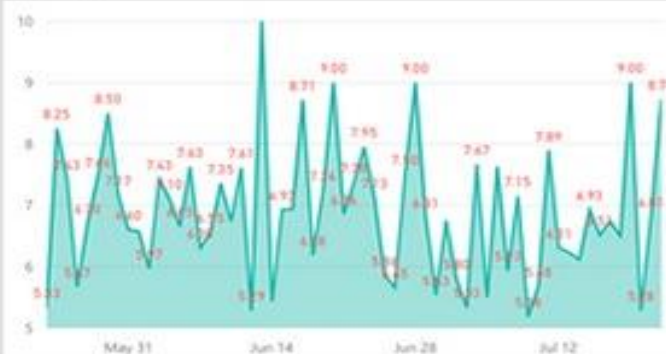
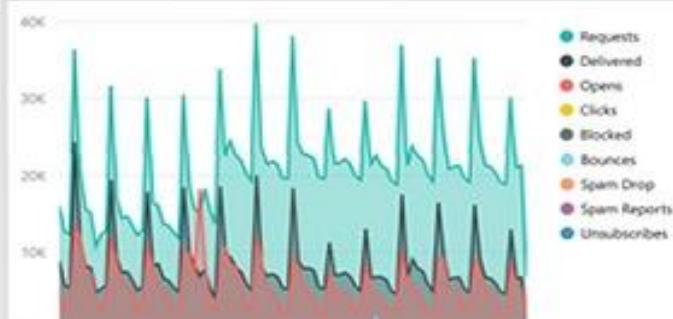
\$85.22M

Total Invoice
THIS MONTH

\$2.22M

Discount Savings
THIS MONTH

\$4.07M

Total Invoice
BY MONTH CATEGORYTotal Invoice
BY SUB CATEGORYMarketing Site Traffic
BY DATEFeedback Ratings
BY SCOREAccountCount
BY INDUSTRYCustomer Feedback Trend
AVERAGE SCOREVariance to Budget
BY MONTHSales Pipeline
BY STAGENAMECustomer Size vs Revenue
CUSTOMER GROUPAccountCount
BY SHIPPING CITY, INDUSTRYNumber of Visits, Expenditures (\$Mil)
BY MONTHMarketing Campaign Stats
BY DATE

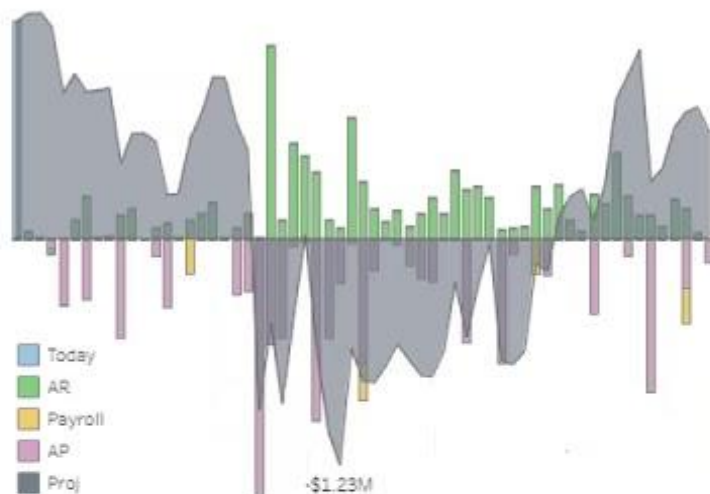
Marketing Campaign Metrics



Press Esc to exit full screen

10/5/2017 9:36:00 AM

60-day Cash Flow Projection



Orders to Ship: \$2,802K



Hot New Opportunities: \$4,045K



Over 60 Days

60 Days

30 Days

On Time

AR Aging: \$4,971K

Revenue YoY

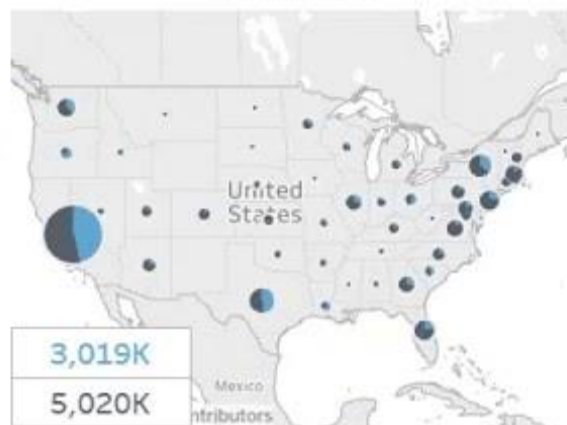
2016 \$7.90M
2017 \$8.84M
11.9% ↑



Avg Invoice Amt Trend



Website Visits



Gross Profit



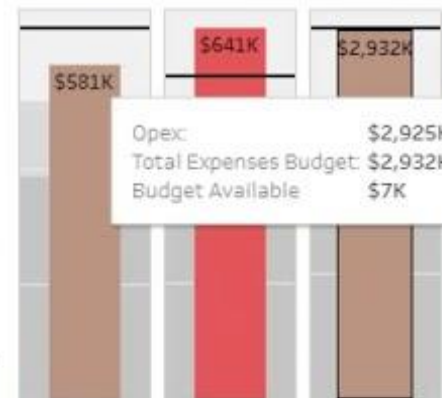
EBITDA



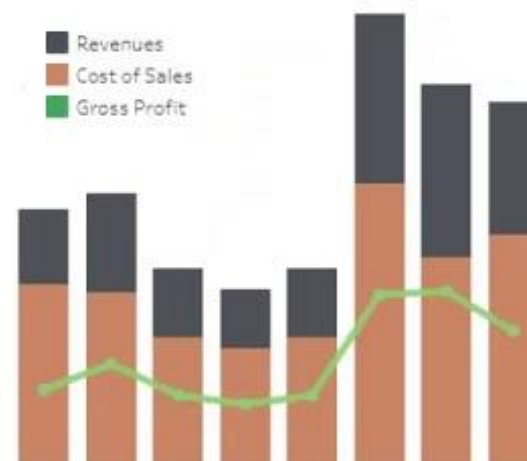
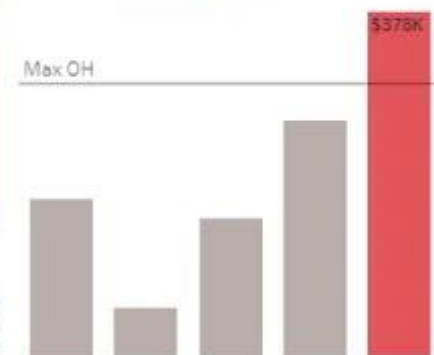
Admin

Mkt

Opex

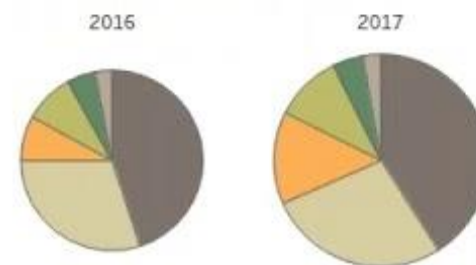


Inventory Assets



Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
\$1.0M	\$1.1M	\$0.8M	\$0.7M	\$0.6M	\$1.7M	\$1.5M	\$1.4M
\$0.7M	\$0.7M	\$0.5M	\$0.4M	\$0.5M	\$1.1M	\$0.8M	\$0.9M
\$0.3M	\$0.4M	\$0.3M	\$0.2M	\$0.3M	\$0.7M	\$0.7M	\$0.5M

Revenue Mix



Debt/Equity Ratio: 41%

Debt: \$617K Equity: \$1,519K



Product Analysis



Region

Europe	✓
Americas	
Asia	
Pacific	
Africa	

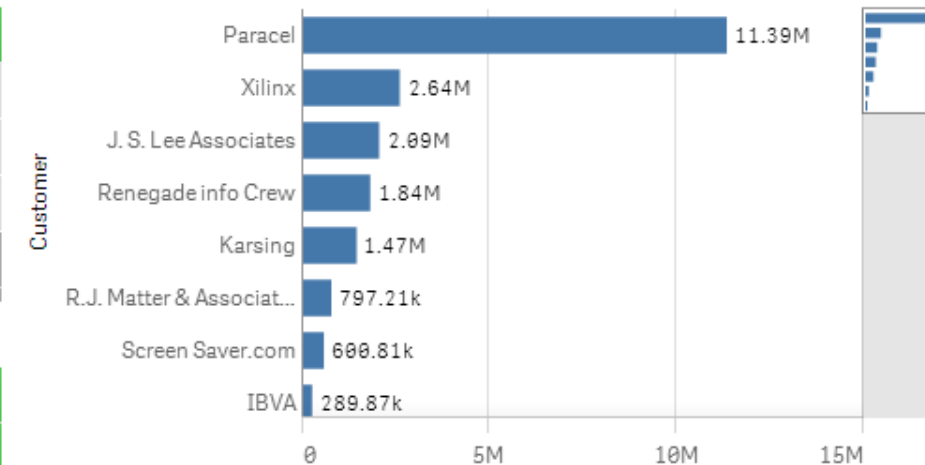
SegmentDesc

Bottle Shops	✓
Convenience Stores	✓
Restaurant & Cafes	✓
Supermarkets	
Alternative	

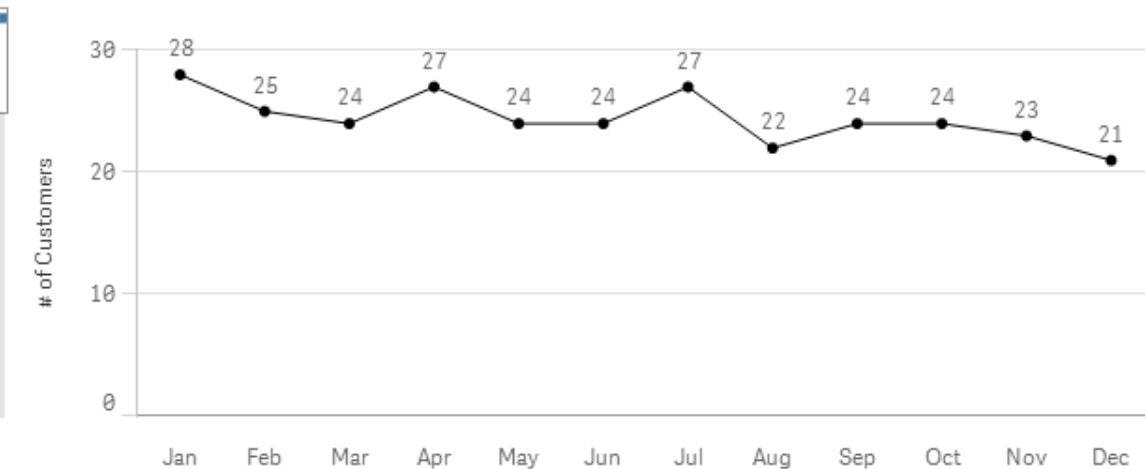
Product line

Anchovies
Bagels
Beer
Bologna
Canned Fruit

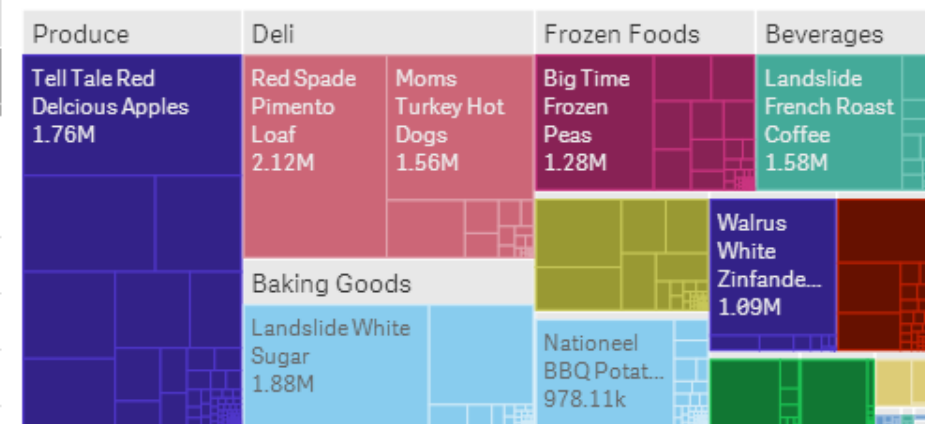
Revenue by Account



Customer count trend

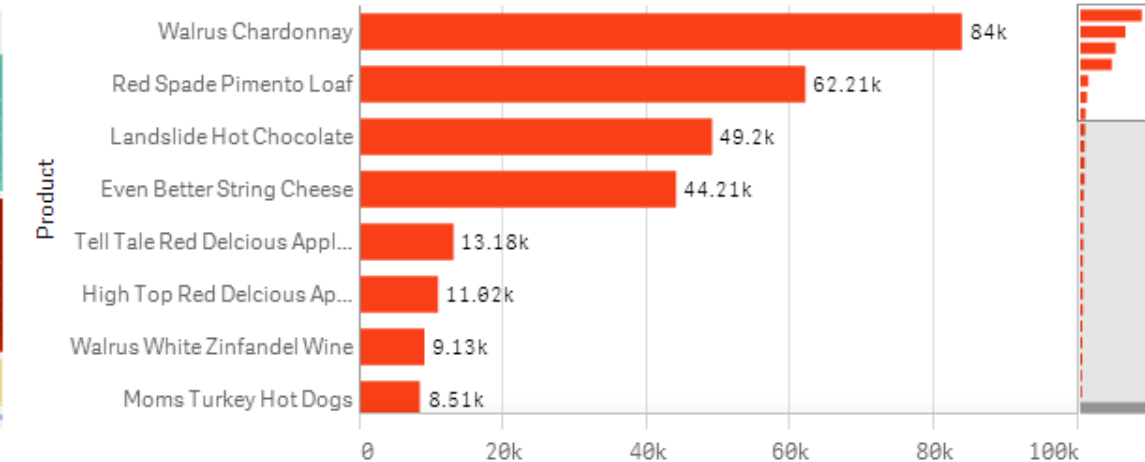


Revenue by Product *



* The data set contains negative or zero values that cannot be shown in this chart.

Sales volume by Product



کارکردهای پروژه‌های هوشمندی کسب و کار



نقشه راه تولید یک پروژه هوش تجاری

این نقشه گام‌های لازم برای ساخت یک پروژه هوش تجاری را به شرح زیر بیان می کند:

- **مرحله هدف:**

گام ۱: ارزیابی وضعیت کسب و کار

- **مرحله طرح ریزی:**

گام ۲: ارزیابی زیر ساخت ساز کام

گام ۳: طرح ریزی پروژه

- **مرحله تحلیل تجاری:**

گام ۴: تعریف نیازمندی‌های پروژه

گام ۵: نمونه سازی برنامه کاربردی

گام ۶: تحلیل منابع داده‌ای و فراداده

نقشه راه تولید یک پروژه هوش تجاری

- مرحله طراحی:

- گام ۷: طراحی پایگاه داده
- گام ۸: طراحی ETL (استخراج/تبدیل/بارگذاری)
- گام ۹: طراحی مخزن فراداده

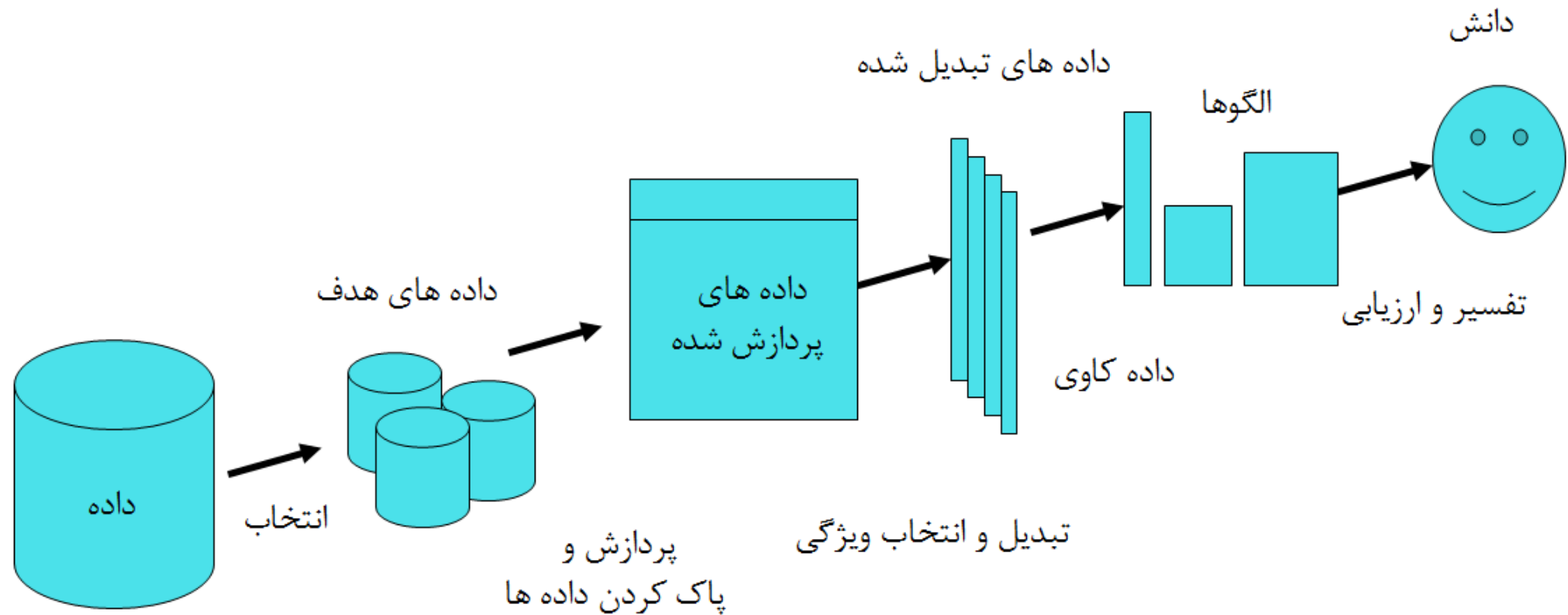
- مرحله ساخت:

- گام ۱۰: تولید ETL
- گام ۱۱: تولید برنامه کاربردی
- گام ۱۲: داده کاوی
- گام ۱۳: تولید مخزن فراداده

- مرحله نصب و استقرار:

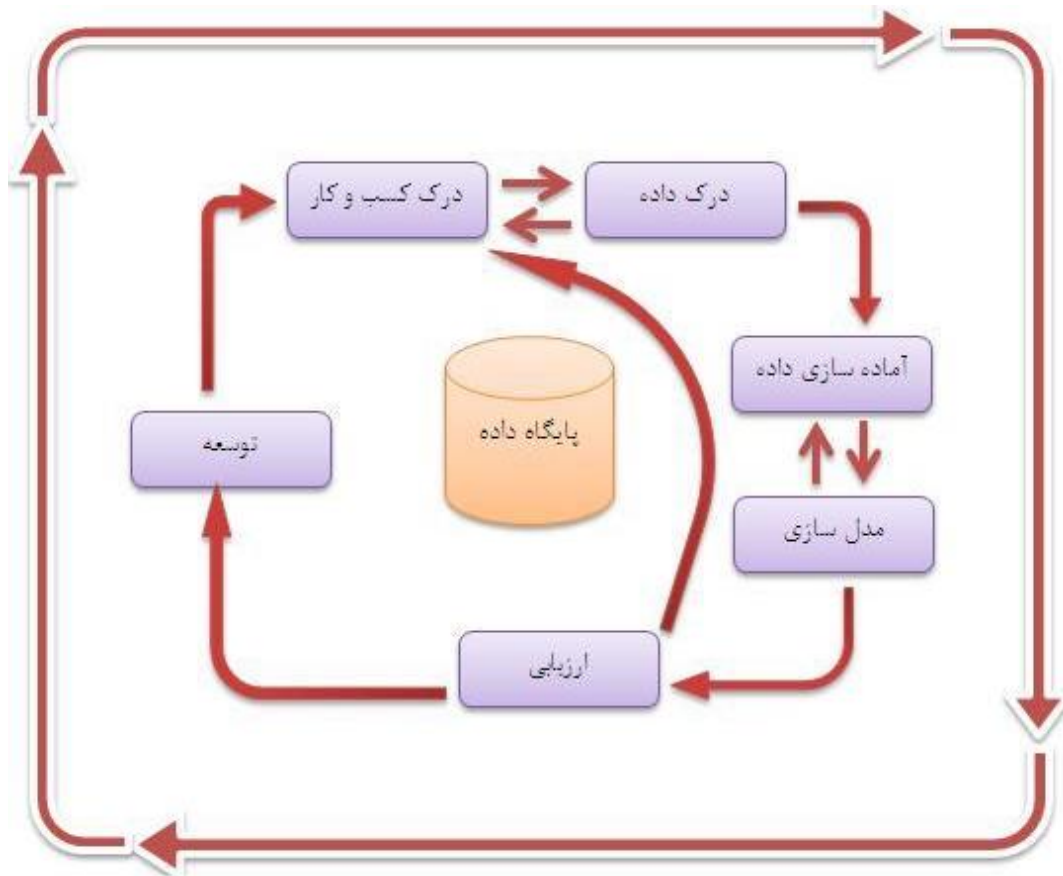
- گام ۱۴: پیاده سازی
- گام ۱۵: ارزیابی نسخه
- گام ۱۶: پشتیبانی

مراحل اجرایی پروژه‌های علم داده



علم داده عبارت است از استخراج اطلاعات و دانش و **کشف الگوهای پنهان** از پایگاه داده‌های بسیار بزرگ

مراحل اجرایی پروژه‌های علم داده



➤ شناخت سیستم و درک کسب و کار

➤ شناخت داده

➤ آماده سازی داده

➤ مدل سازی

➤ ارزیابی

➤ توسعه نهایی

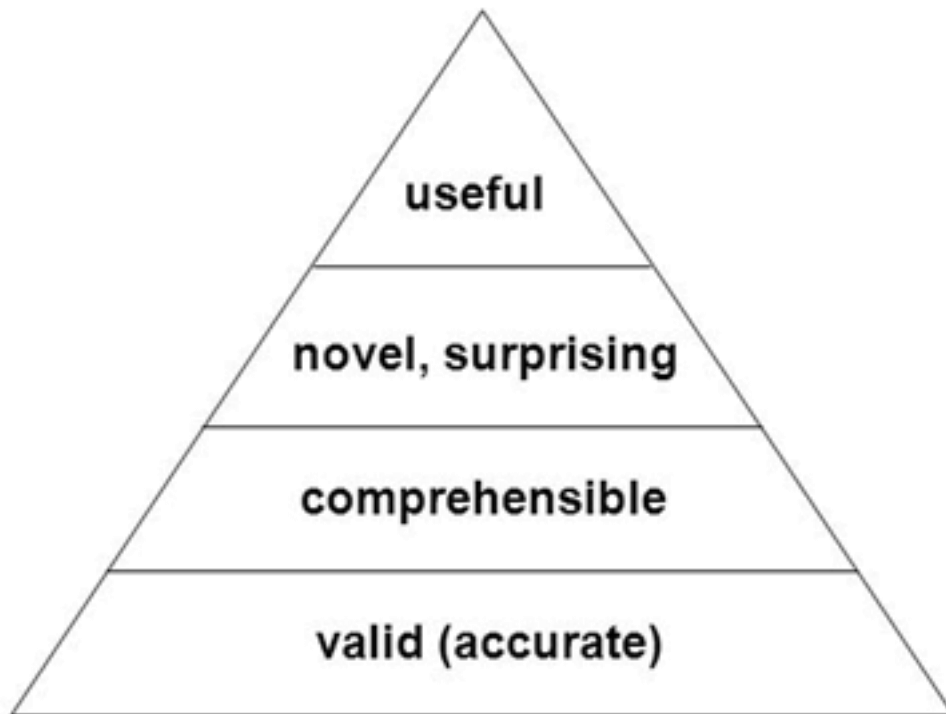
اهداف پروژه‌های علم‌داده

✓ معتبر باشد (Valid)

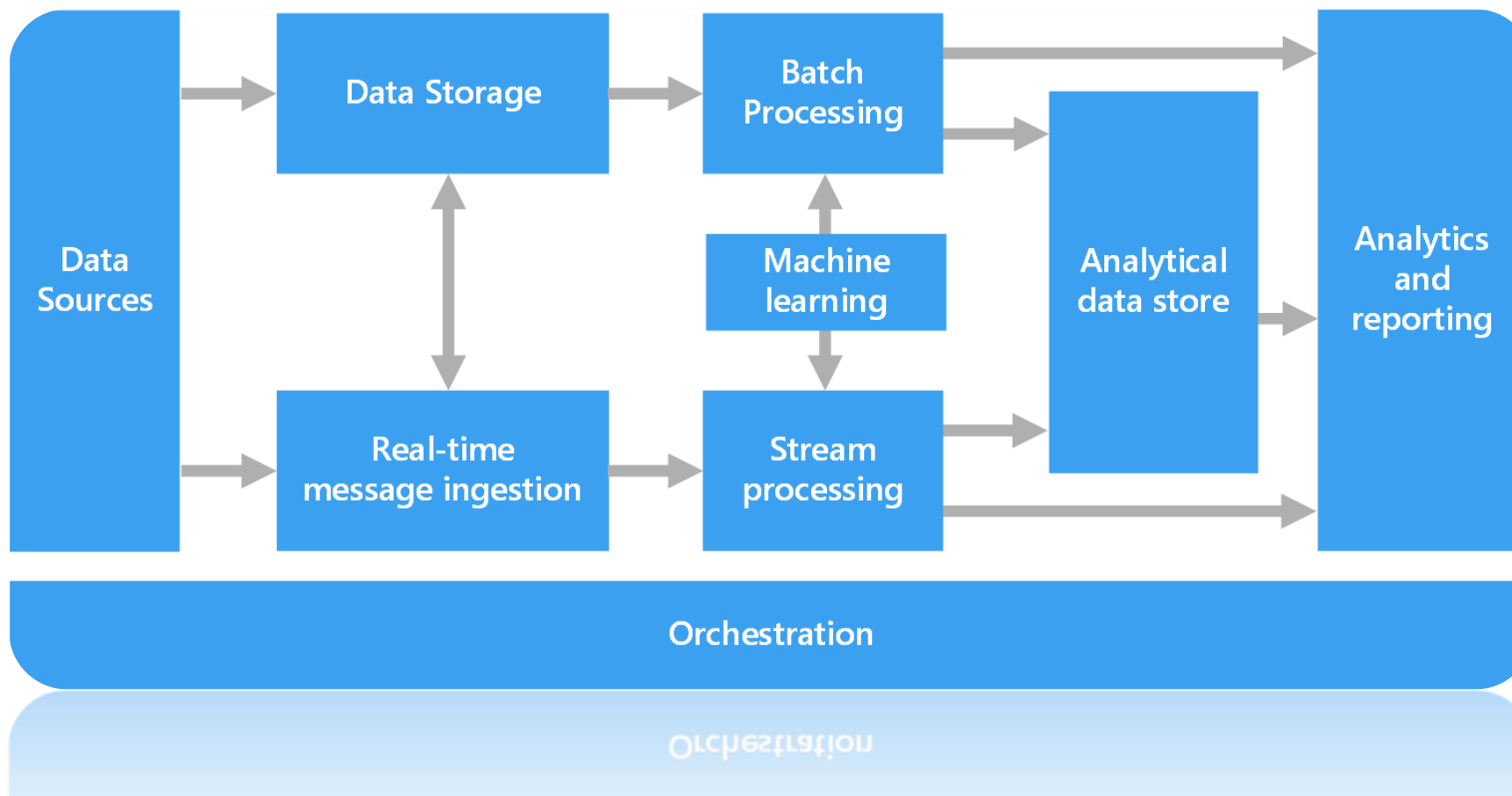
✓ قابل فهم باشد (Comprehensible)

✓ جدید باشد (Novel)

✓ مفید باشد (Useful)



طراحی و استقرار پروژه‌های هوش تجاری در مقیاس کلان داده





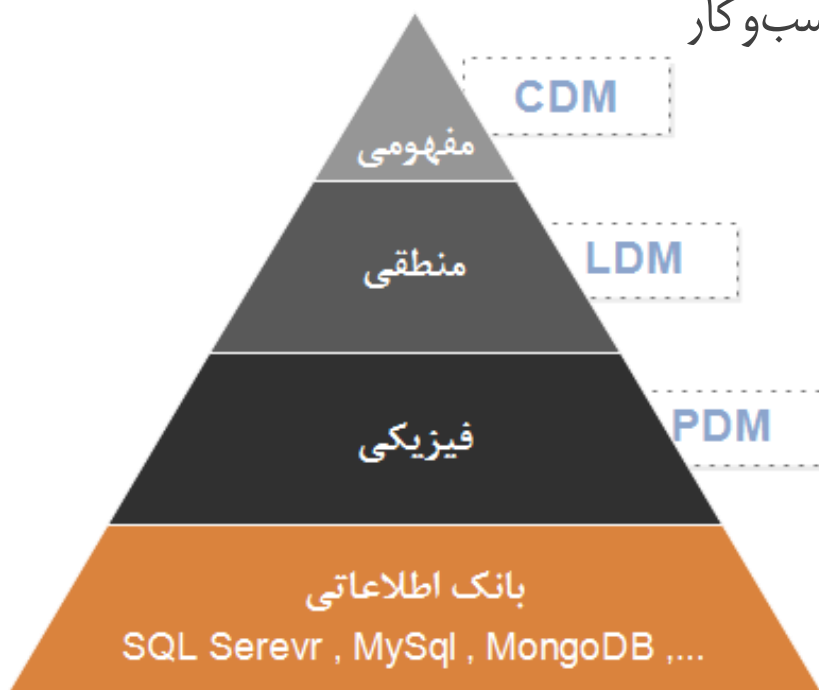


طراحی و مدل سازی داده

طراحی و مدل سازی داده

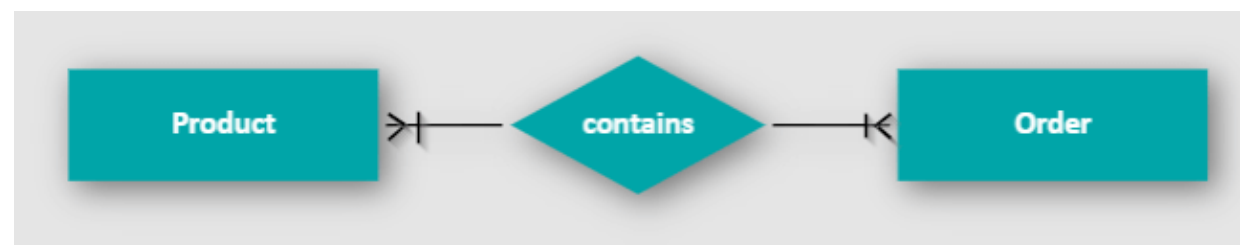
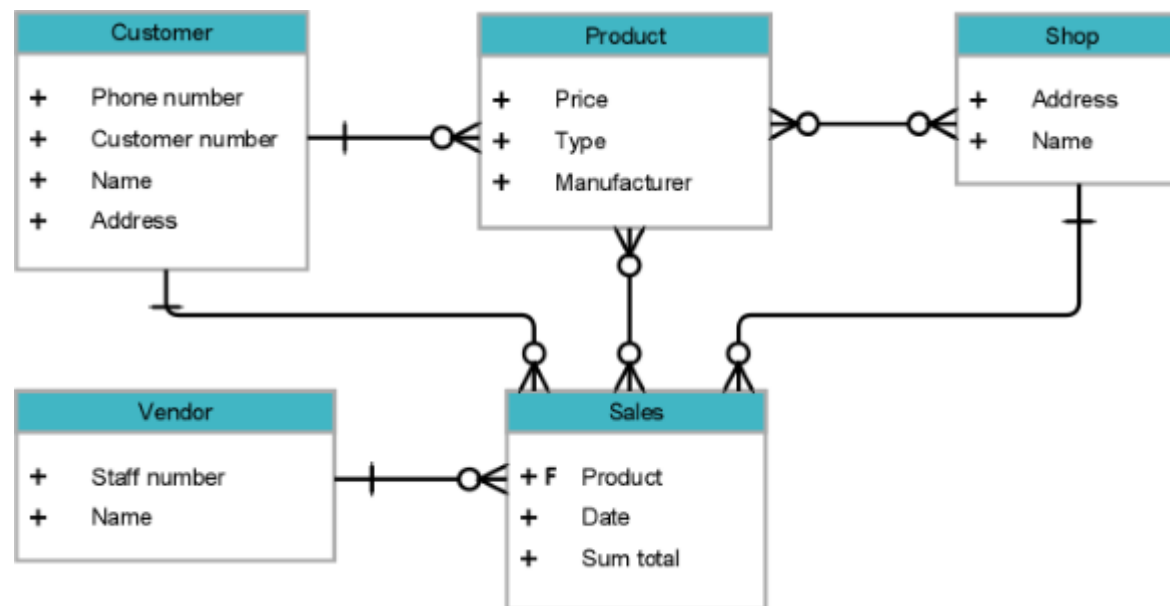
تعریف: شناسایی و تحلیل نیازمندی های اطلاعاتی و طراحی مدل داده ای آن

اهداف: شناسایی و مدل سازی ابعاد مختلف داده ها در راستای پاسخ دهی به نیازهای کسب و کار



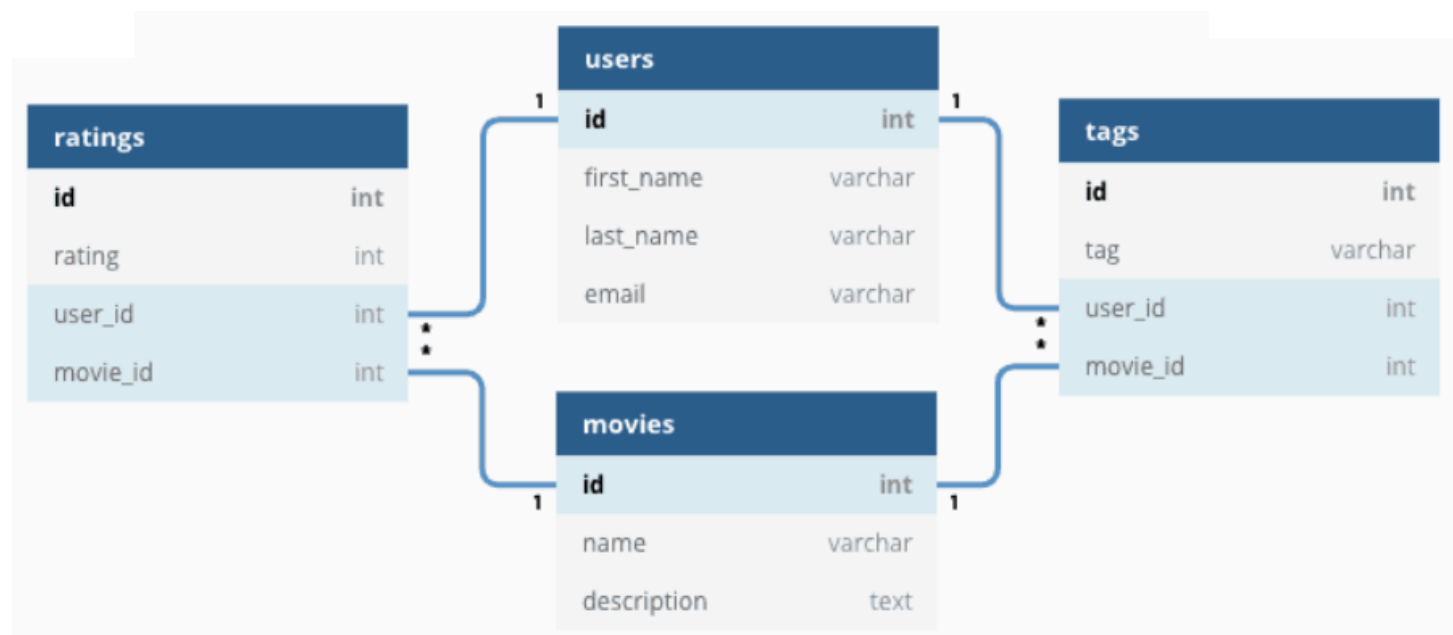
تفاوت مدل‌های داده‌ای

ویژگی	مفهومی	منطقی	فیزیکی
اسامی موجودیت‌ها	✓	✓	
روابط موجودیت‌ها	✓	✓	
ویژگی‌ها		✓	
کلیدهای اولیه		✓	✓
کلیدهای خارجی		✓	✓
اسامی جداول			✓
اسامی ستون‌ها			✓
نوع داده ستون‌ها			✓



تفاوت مدل‌های داده‌ای

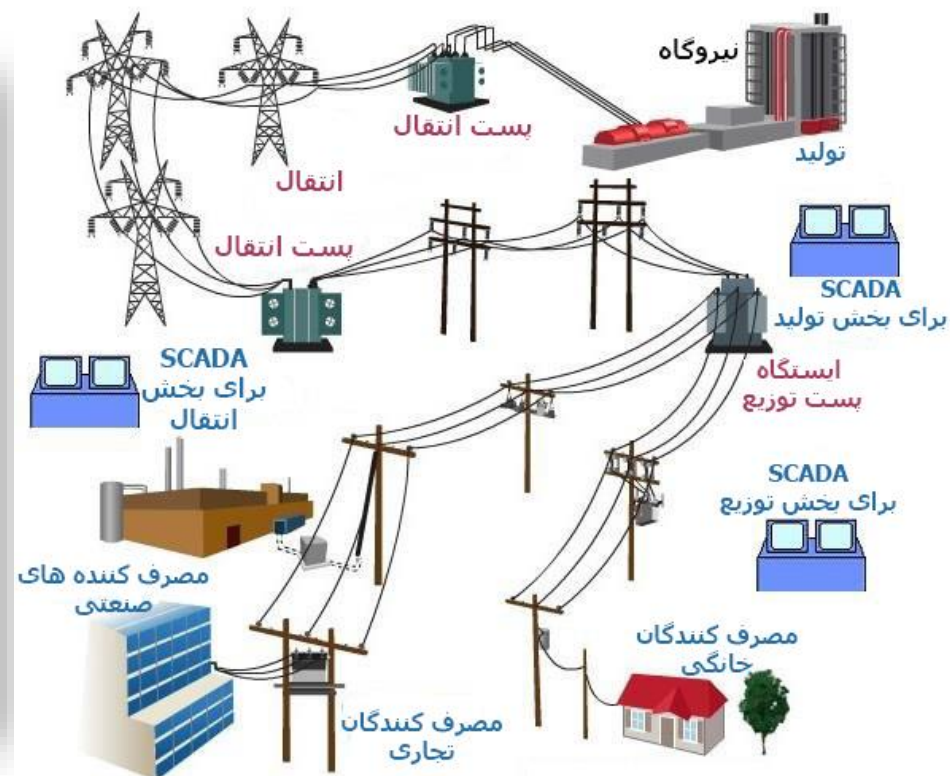
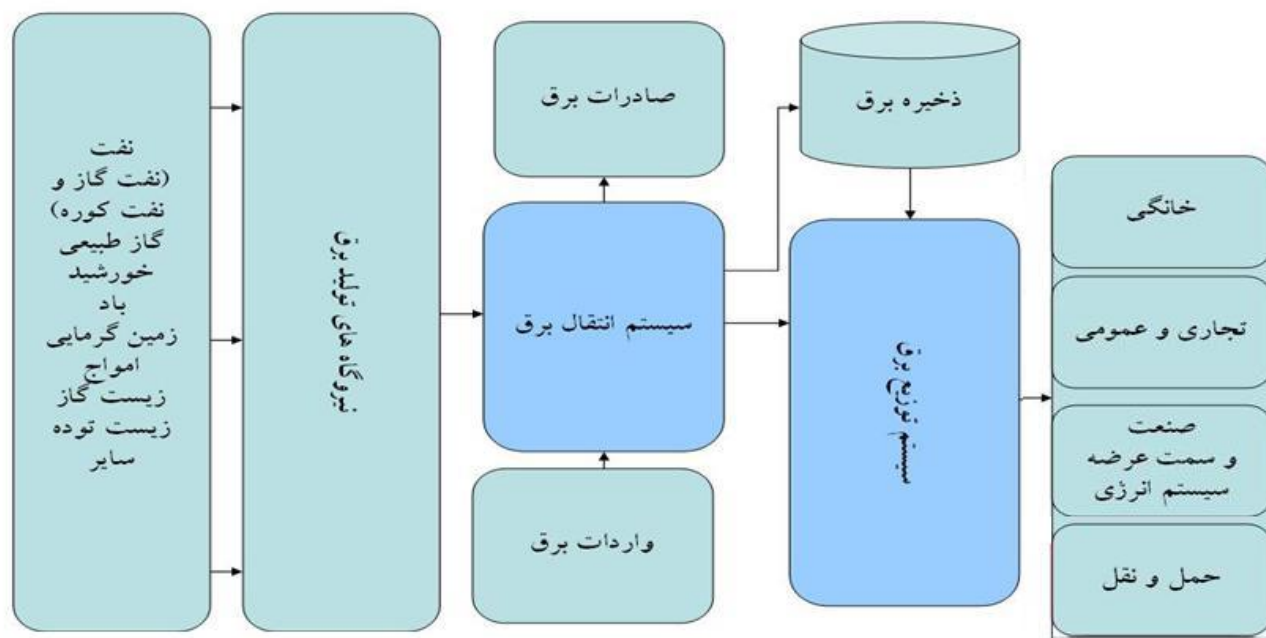
ویژگی	مفهومی	منطقی	فیزیکی
اسامی موجودیت‌ها	✓	✓	
روابط موجودیت‌ها	✓	✓	
ویژگی‌ها		✓	
کلیدهای اولیه		✓	✓
کلیدهای خارجی		✓	✓
اسامی جداول			✓
اسامی ستون‌ها			✓
نوع داده ستون‌ها			✓



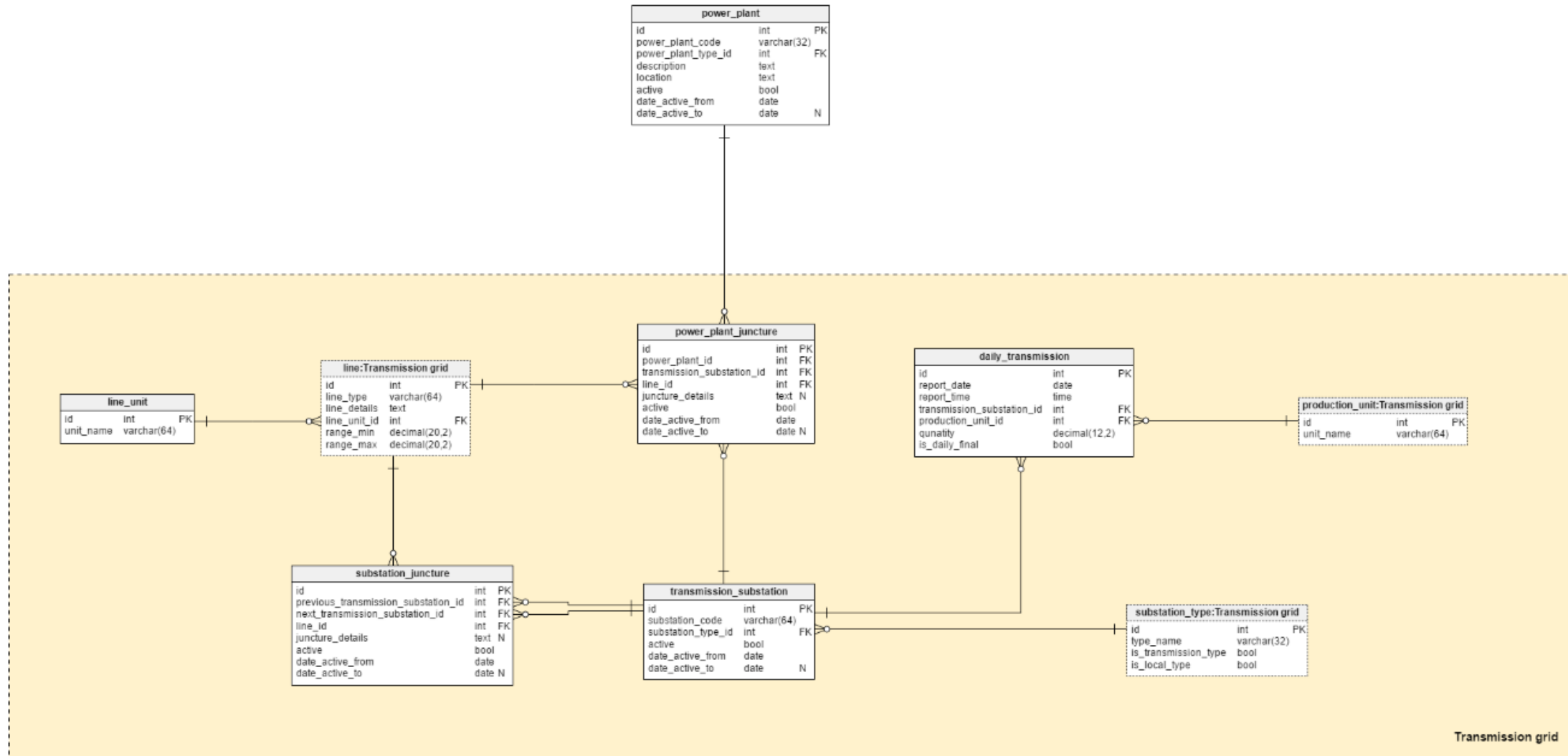
طراحی و مدل سازی داده

طراحی و مدل سازی داده مفهومی سیستم توزیع نیروی برق

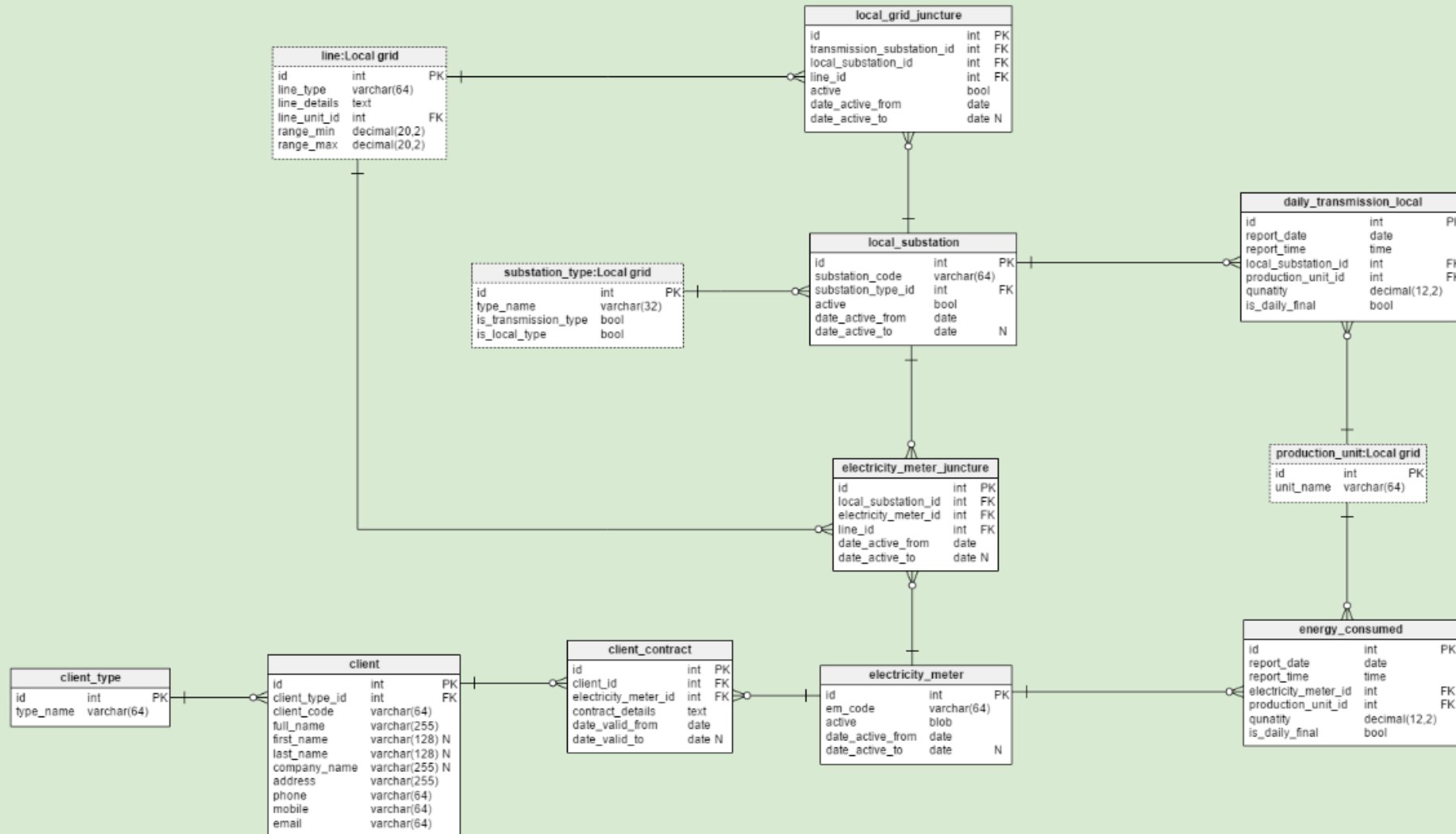
طراحی و مدل سازی داده



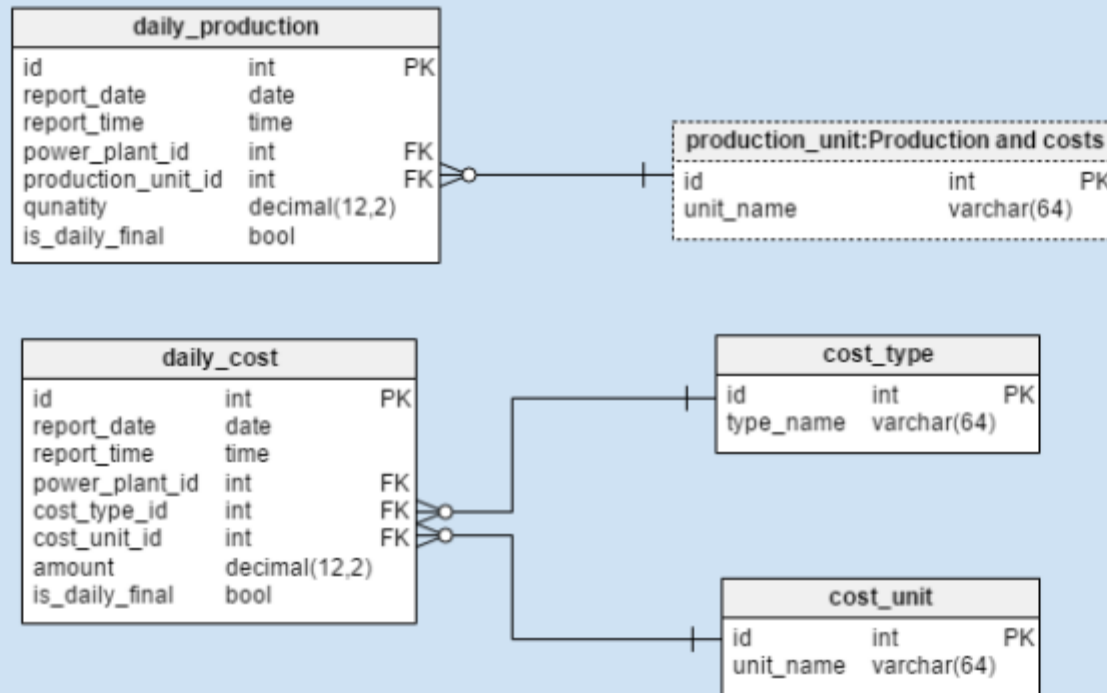
بخش ۱: شبکه انتقال از نیروگاه به شبکه‌های محلی



بخش ۲: شبکه محلی



بخش ۳: مالی



Production and costs



ذخیره سازی و عملیات داده

ذخیره‌سازی و عملیات داده

در کسب‌وکار خود از چه پایگاه داده‌ای استفاده می‌کنید؟
چرا؟!

رویکردهای ذخیره سازی داده

NoSQL



Gaming



Social



IoT



Web



Mobile



Enterprise



Key/value store



Document database



Column family store

SQL



Web



Mobile



Enterprise



Data mart



Relational table storage



Relationships use joins

شناخت انواع تکنولوژی پایگاه داده

ارزیابی تکنولوژی‌های مختلف

مدیریت تکنولوژی‌های مختلف

پایش مداوم تکنولوژی

مدیریت عملیات پایگاه داده

شناسایی نیازمندی‌ها

تهیه طرح تداوم کسب و کار

مدیریت کارایی پایگاه داده

انتقال یا مهاجرت داده‌ها

فعالیت‌های ذخیره‌سازی و عملیات داده

خروجی ذخیره سازی و عملیات داده

معیارهای ارزیابی فناوری پایگاه داده

محیط پایگاه داده

برنامه تداوم کسب و کار

داده ها و کدهای مربوط به انتقال داده ها

توافق نامه سطح سرویس های پایگاه داده

شاخص‌های ارزیابی عملکرد ذخیره‌سازی و عملیات داده



تعداد پایگاه داده
میزان زمان بالا بودن پایگاه داده
آمار تراکنش‌های ورودی (حجم، سرعت و...)
میانگین زمان پاسخ‌دهی به پرس‌وجوها
برنامه پشتیبان‌گیری (زمان، تعداد دفعات، مدت نگهداری و...)
تحلیل مشکلات (نرخ وقوع، مدت زمان اصلاح و...)



داده‌های مرجع و اصلی

• ویژگی‌های لازم جهت یک مشترک (آب یا برق) را بنویسید
(رکورد طلایی)

• ویژگی‌های لازم جهت مدیریت شبکه برق کشور را بنویسید
(رکورد طلایی)

تهیه منبع معتبر و جامع از موجودیت‌های حیاتی
داده برای عملیات و تحلیل

داده اصلی

#	تاریخ	زمان	نیاز مصرف برق در بیک	نیاز مصرف برق در بیک (سال گذشته)	ذخیره نیروگاه ها	مصرف برق صنایع	تولید نیروگاه های تجدید پذیر (مجموع 24 ساعت)	میانگین فرکانس
1	1400/05/15	21:37	56494	55052	434	3,827	6185	49.93
2	1400/05/15	13:29	58350	55558	650	3,745	6185	49.93
3	1400/05/14	21:04	57238	56109	629	3,227	6403	49.91
4	1400/05/14	12:48	62495	56315	525	3,011	6403	49.91
5	1400/05/13	21:19	58238	56897	684	3,240	6118	49.90

مدیریت داده‌های مرجع و اصلی

مدیریت داده‌های مشترک و اشتراک‌گذاری داده‌ها میان واحدهای کسب‌وکار

ارائه یک منبع داده مطمئن

کاهش افزونگی داده

کاهش هزینه یکپارچه‌سازی

اطمینان از کیفیت داده بالا

استانداردسازی تفسیر دارایی‌های اطلاعاتی

مدیریت داده‌های مرجع و اصلی



انواع داده در یک سازمان

Reference Data

نوع مشتری	کد
تجاری	۱
خانگی	۲
تجاری و خانگی	۳

Meta Data

نام فیلد	شرح	نوع فیلد	طول فیلد
کد محصول	یک کد انحصاری برای هر محصول که توسط مدیر محصول در نظر گرفته می شود	عددی	۲۰
نام محصول	نام انحصاری در نظر گرفته شده برای محصول	حرفی	۴۰

Master Data (رکورد محصول)

کد محصول: ۱۳۹۴

نام محصول: سیستم الف

Master Data (رکورد مشتری)

کد مشتری: ۱۰۰۶۱

نوع مشتری: تجاری

نام مشتری: محمدی

آدرس: ایران - تهران - ...

Transactional Data (ثبت سفارش)

شماره سفارش: ۹۹۹۹۹

کد مشتری: ۱۰۰۶۱

نام مشتری: محمدی

نوع مشتری: تجاری

آدرس: ایران - تهران - ...

کد محصول: ۱۳۹۴

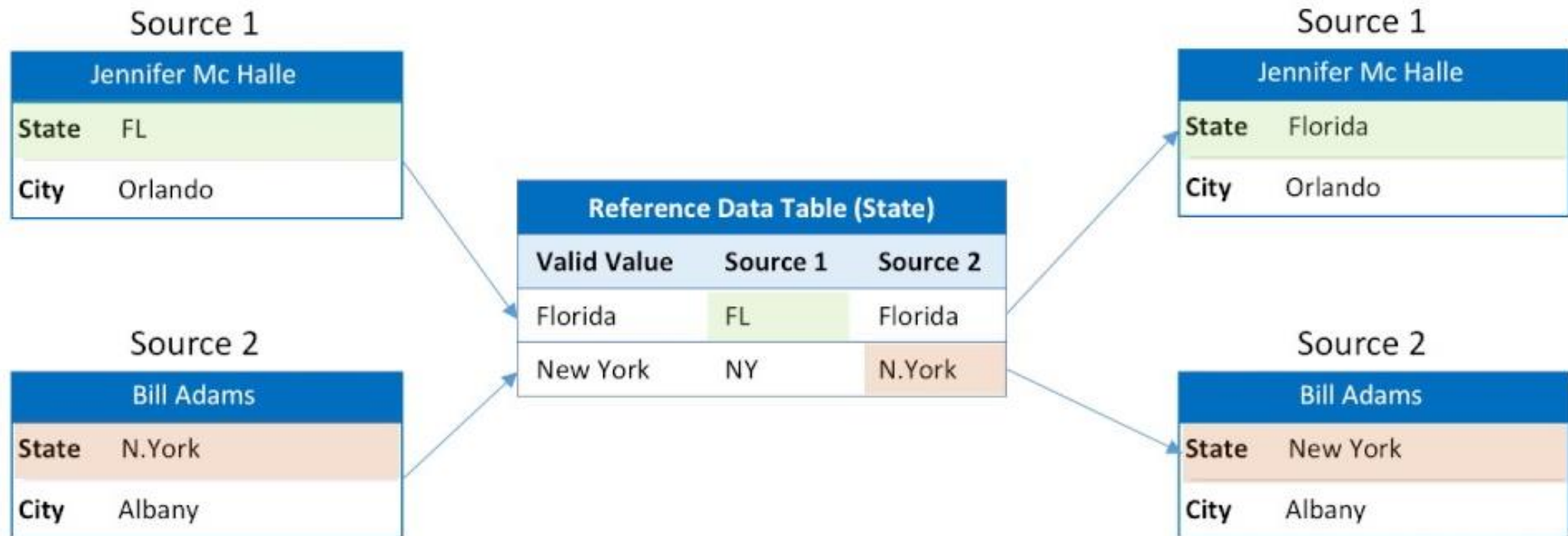
شرح: سیستم الف

تعداد: ۲

فی: ۱۰۰

جمع: ۲۰۰

کد محصول	نوع	قیمت فی
۱۳۹۴	تجاری	۱۰۰
۱۳۹۵	خانگی	۵۰
۱۳۹۶	تجاری و خانگی	۷۵





فراداده

فراداده

به داده‌هایی گفته می‌شود که جزئیات یک داده دیگر را تشریح می‌کند.
به عبارت دیگر فراداده‌ها، داده‌هایی هستند درباره داده‌های دیگر

employee_id	first_name	last_name	nin	department_id
44	Simon	Martinez	HH 45 09 73 D	1
45	Thomas	Goldstein	SA 75 35 42 B	2
46	Eugene	Comelsen	NE 22 63 82	2
47	Andrew	Petculescu	XY 29 87 61 A	1
48	Ruth	Stadick	MA 12 89 36 A	15
49	Barry	Scardelis	AT 20 73 18	2
50	Sidney	Hunter	HW 12 94 21 C	6
51	Jeffrey	Evans	LX 13 26 39 B	6
52	Doris	Bemdt	YA 49 88 11 A	3
53	Diane	Eaton	BE 08 74 68 A	1
54	Bonnie	Hall	WW 53 77 68 A	15
55	Taylor	Li	ZE 55 22 80 B	1

Data

Metadata

Column	Data Type	Description
employee_id	int	Primary key of a table
first_name	nvarchar(50)	Employee first name
last_name	nvarchar(50)	Employee last name
nin	nvarchar(15)	National Identification Number
position	nvarchar(50)	Current position title, e.g. Secretary
department_id	int	Employee department. Ref: Departments
gender	char(1)	M = Male, F = Female, Null = unknown
employment_start_date	date	Start date of employment in organization.
employment_end_date	date	Employment end date. Null if employee sti

مدیریت فراداده

تعریف: برنامه‌ریزی، پیاده‌سازی و کنترل فعالیت‌ها با هدف دستیابی به فراداده یکپارچه و با کیفیت بالا

اهداف:

❑ جمع‌آوری و یکپارچه‌سازی فراداده‌های از منابع مختلف

❑ ارائه روشی استاندارد جهت دستیابی و دسترسی به فراداده‌ها

❑ اطمینان از کیفیت فراداده

❑ ارائه شناخت سازمانی از واژه‌ها کسب‌وکار

فراداده مدیریت شبکه برق کشور

داده اصلی

#	تاریخ	زمان	نیاز مصرف برق در بیک	نیاز مصرف برق در بیک (سال گذشته)	ذخیره نیروگاه ها	مصرف برق صنایع	تولید نیروگاه های تجدید پذیر (مجموع 24 ساعت)	میانگین فرکانس
1	1400/05/15	21:37	56494	55052	434	3,827	6185	49.93
2	1400/05/15	13:29	58350	55558	650	3,745	6185	49.93
3	1400/05/14	21:04	57238	56109	629	3,227	6403	49.91
4	1400/05/14	12:48	62495	56315	525	3,011	6403	49.91
5	1400/05/13	21:19	58238	56897	684	3,240	6118	49.90

نام ستون	نوع	واحد	شرح
تاریخ	رشته	—	سال، ماه و روز
زمان	رشته	—	شامل ساعت، دقیقه و ثانیه
نیاز مصرف در پیک	عدد	مگاوات	مقدار مصرف در لحظه پیک
نیاز مصرف در پیک (در سال گذشته زمان مشابه)	عدد	مگاوات	—
ذخیره نیروگاه‌ها	عدد	مگاوات	برای اطمینان از پوشش نیاز مصرف کشور و تامین فرکانس مطلوب، تولید همیشه اندکی بیشتر از پیش‌بینی مصرف برنامه‌ریزی می‌شود. از این مقدار مازاد که غالباً به صورت سوخت در نیروگاه‌های حرارتی کنار گذاشته می‌شود با عنوان ذخیره نیروگاه‌ها یاد می‌شود. به بیان دقیق‌تر ذخیره نیروگاه عبارت است از تفاضل توان قابل تولید و توان تولید شده در لحظه پیک است.
مصرف برق صنایع	رشته	مگاوات	بخشی از برق که به صورت اختصاصی یا اشکال دیگر برای صنایع تولید می‌شود.
تولید نیروگاه‌های تجدیدپذیر	عدد صحیح	مگاوات	مقدار تولید در نیروگاه‌های برق آبی (سدها)، خورشیدی و ...
میانگین فرکانس	عدد اعشاری	هرتز	ایران مانند کشورهای آسیا، آفریقا، اروپا و آمریکای جنوبی از استاندارد ولتاژ ۲۲۰ و فرکانس ۵۰ هرتز استفاده می‌کند. میانگین فرکانسی که در جدول ثبت می‌شود مقداری است که در آن لحظه محقق شده است.

دلایل رشد مدیریت فراداده



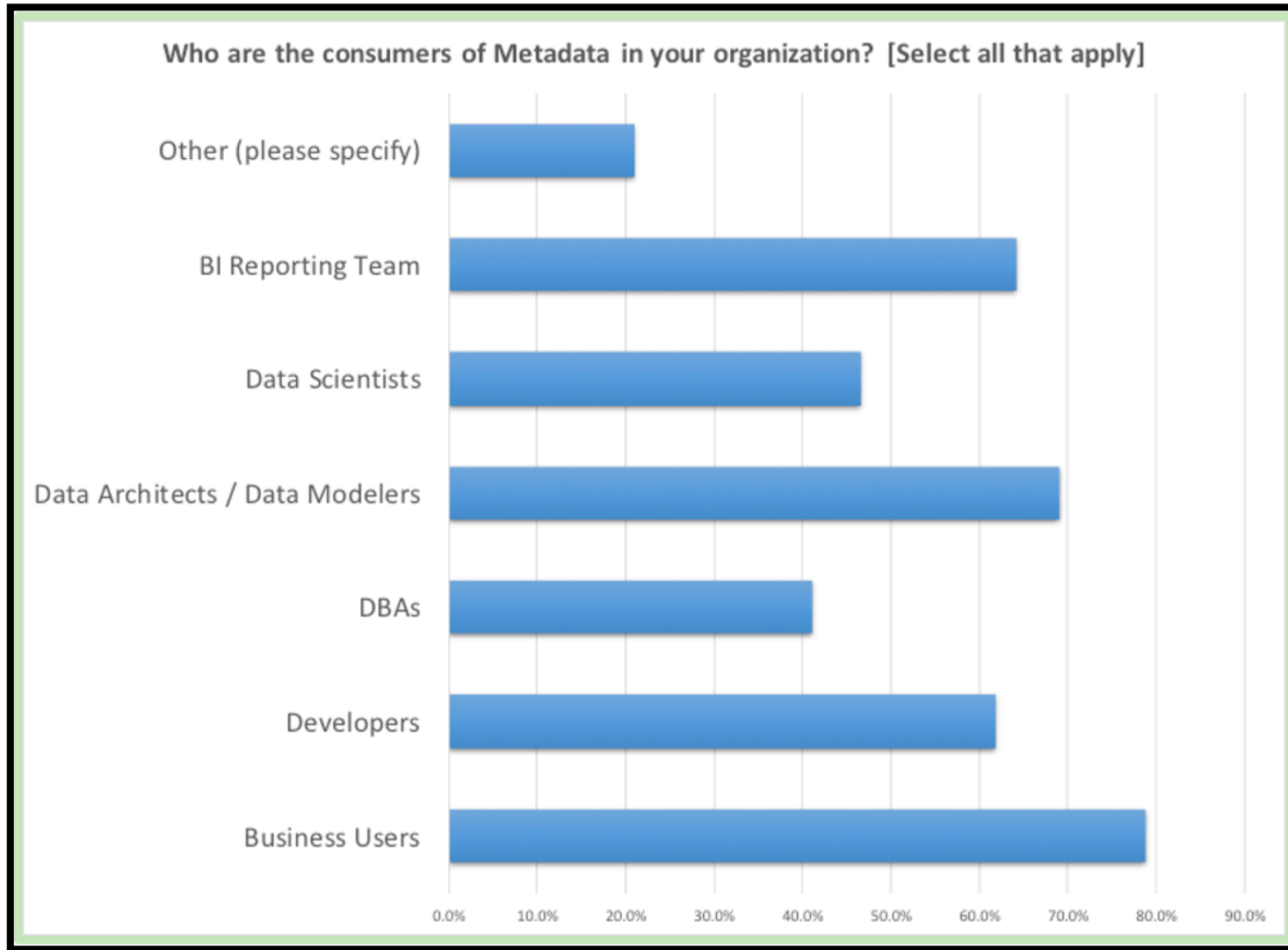
افزایش الزامات مقرراتی و انطباق از قبیل شفافیت

افزایش منابع داده‌ای و پیچیدگی سیستم‌ها

نیاز به درک و تحلیل داده

توجه روزافزون واحدهای کسب و کار به قابلیت‌های داده‌ای

مخاطبین فراداده



خروجی‌های مدیریت قراردادها

استانداردهای قراردادها

معماری قراردادها

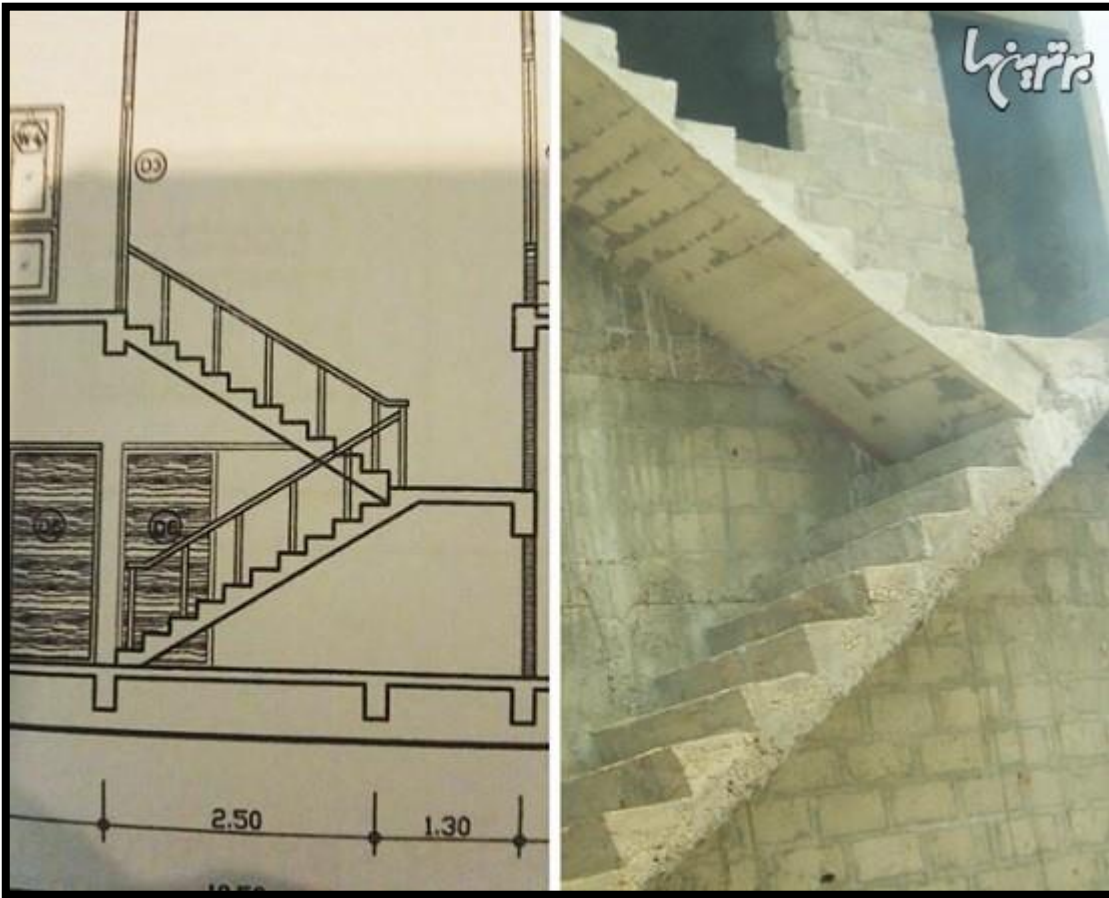
مخزن قراردادها

فرایند کنترل و ارزیابی قراردادها



معماری داده

چالش نبود معماری

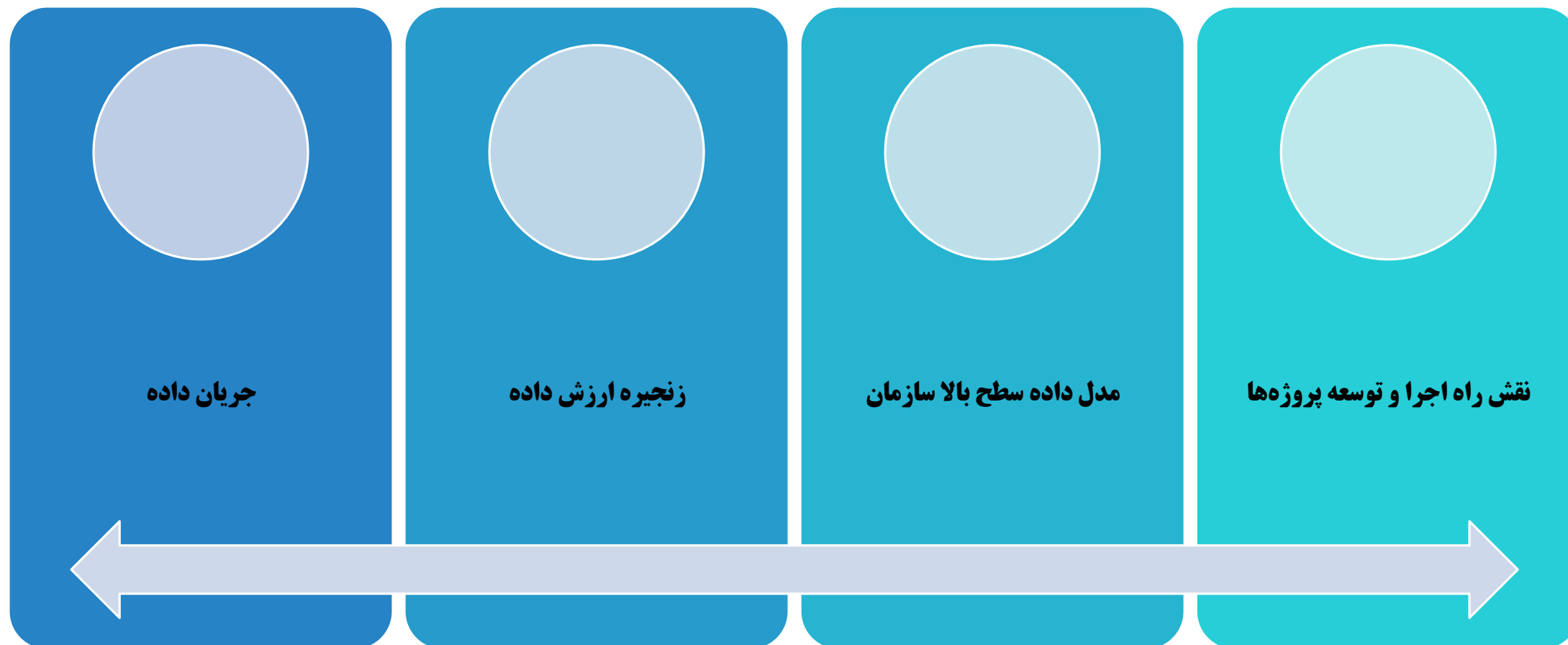


تعریف یک مدل جامع برای داده‌های کلیدی کسب و کار

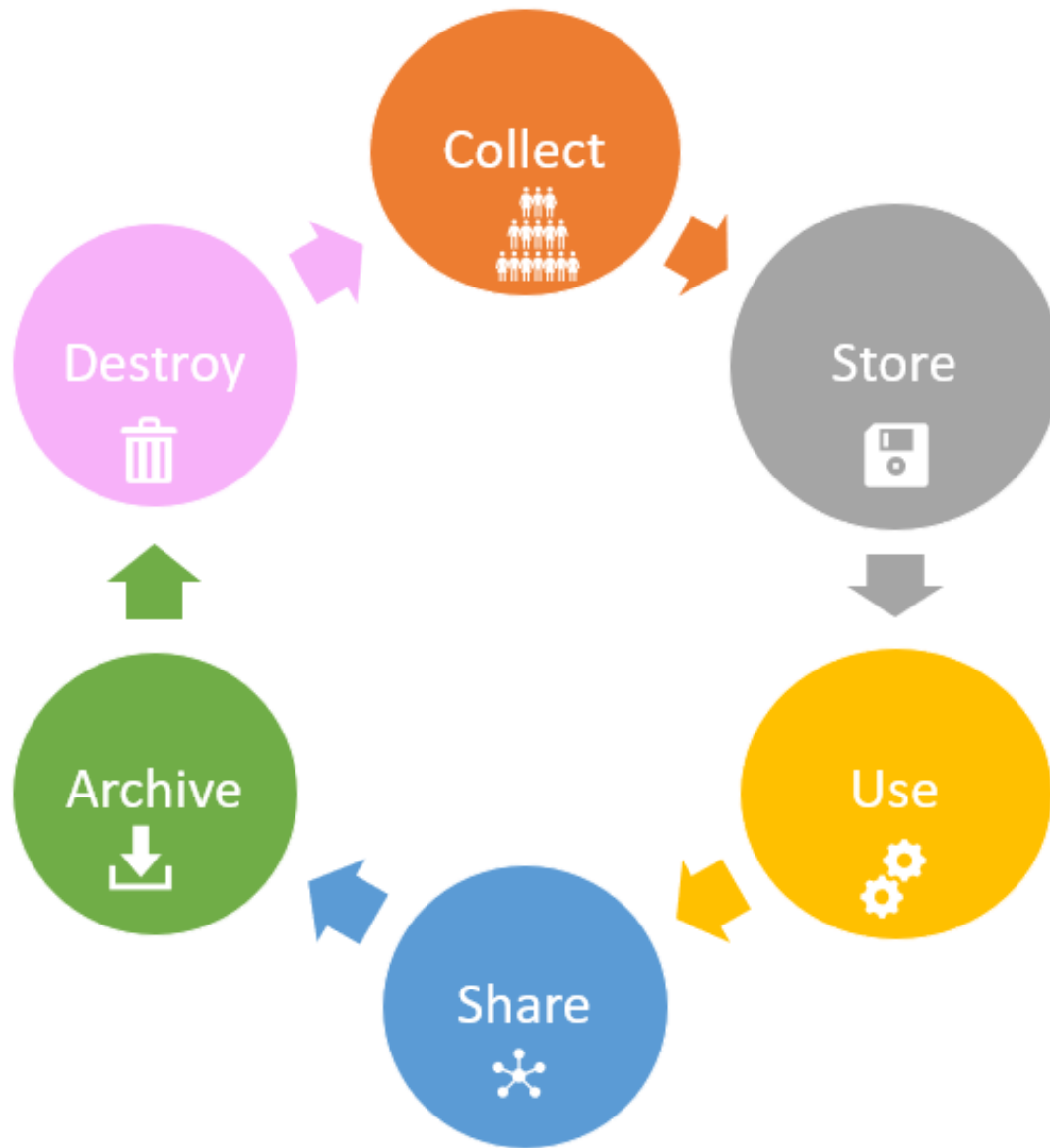
معماری داده

مدل‌ها، خط‌مشی‌ها، قوانین یا استانداردهایی که بر داده‌های جمع‌آوری شده و نحوه ذخیره‌سازی، ترتیب، یکپارچه‌سازی و استفاده در سیستم‌های داده و سازمان‌ها حاکم است.

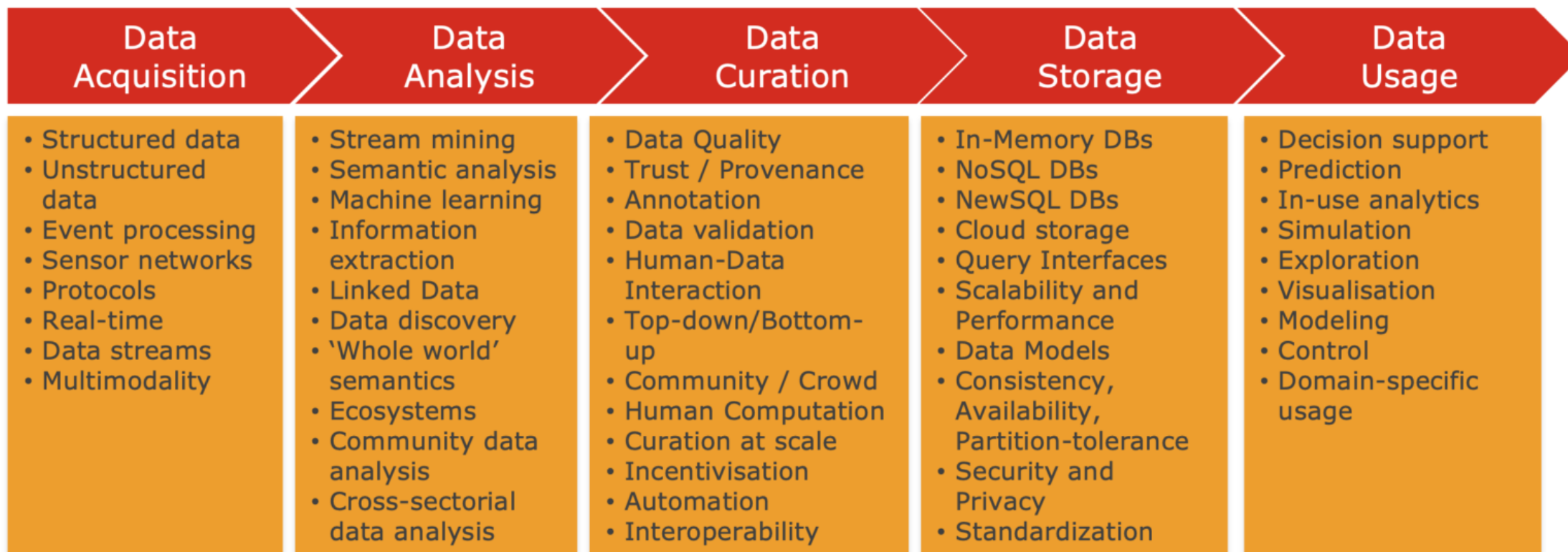
خروجی‌های معماری داده



چرخه حیات داده



زنجیره ارزش کلان داده



اکوسیستم داده

نقش تولیدکننده / تامین کننده داده



نقش جمع کننده داده



نقش توزیع کننده داده



نقش مصرف کننده / مشتری / کاربر داده



نقش تحلیلگر / پردازشگر داده

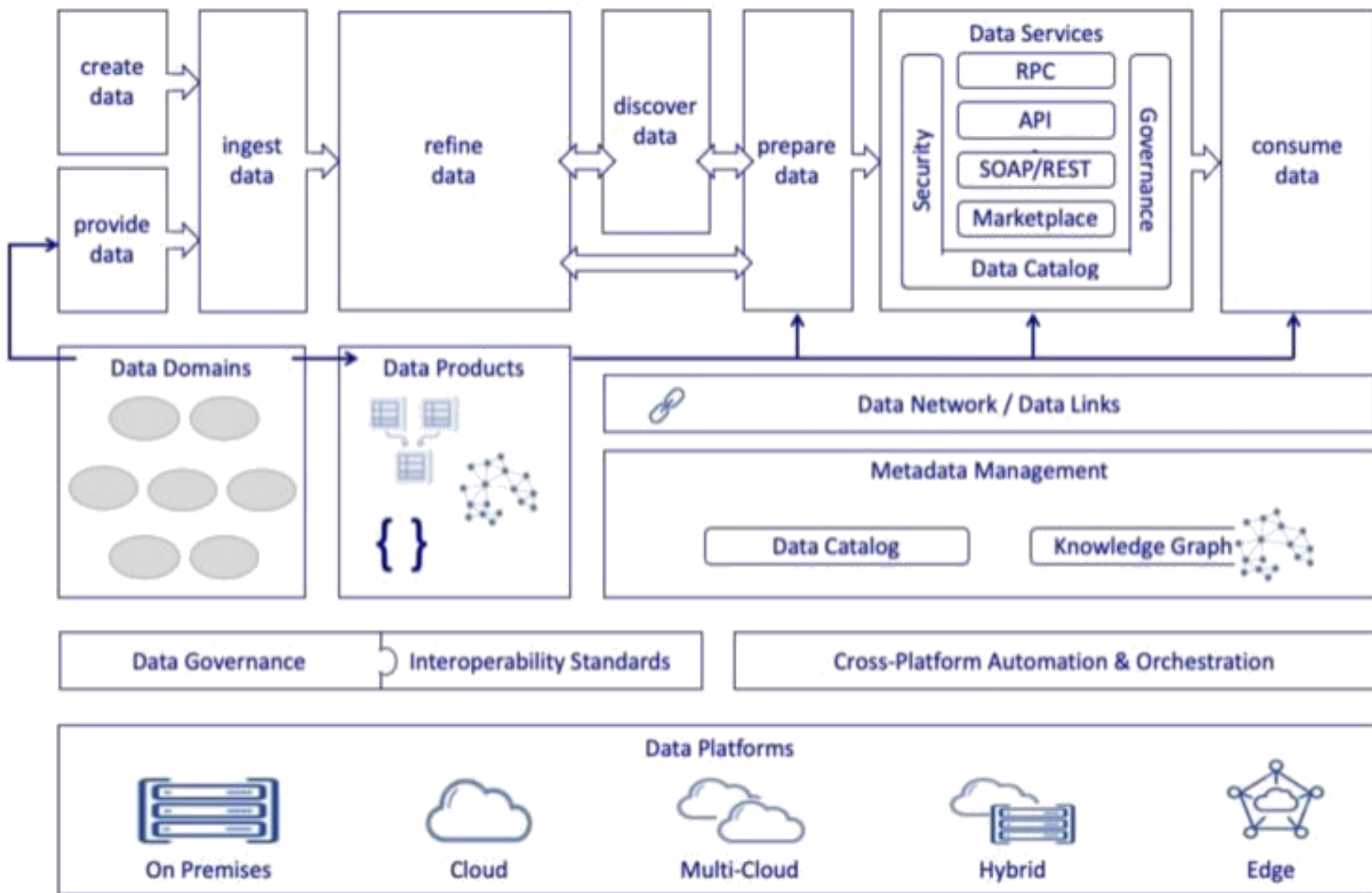


نقش فراهم کننده زیرساخت داده‌ای

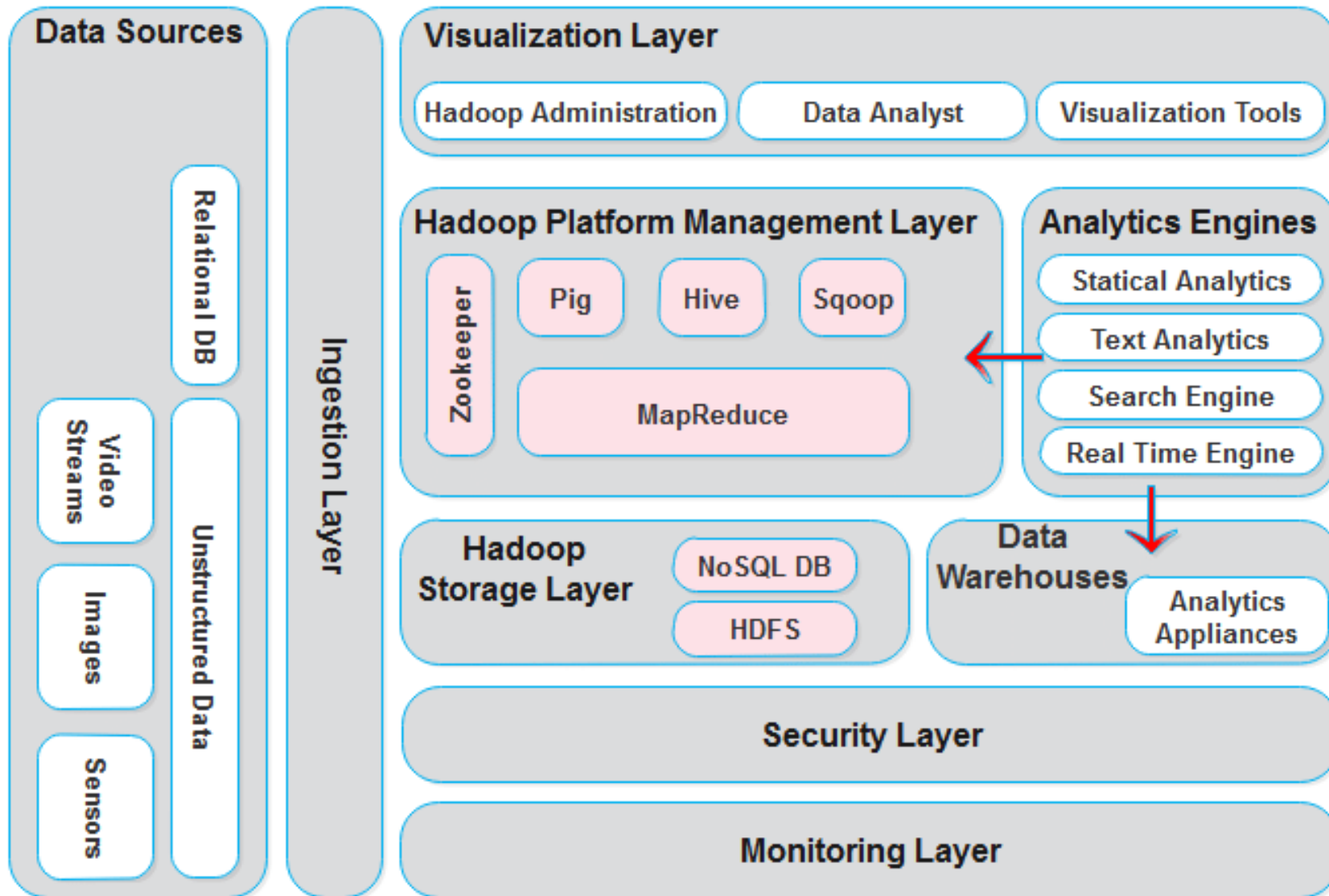


نقش فراهم کننده سرویس داده‌ای





معماری داده



معماری داده مبتنی بر کلان داده

وظایف و مسئولیت‌های معمار داده

ترجمه‌ی نیازمندی‌های کسب‌وکار به پایگاه‌های داده، انبارهای داده و جریان‌های داده

تجزیه و تحلیل، برنامه‌ریزی و تعریف چارچوب معماری داده‌ها

ایجاد و پیاده‌سازی فرایندها و رویه‌های مدیریت داده

همکاری با دیگر تیم‌ها در سازمان برای طراحی و اجرای مدیریت داده

توسعه‌ی API برای بازیابی داده‌ها



یکپارچگی داده

چالش‌های نگرش جزیره‌ای

هزینه و مشکلات انتقال داده

وجود داده تکراری در منابع مختلف

کاهش امنیت داده

عدم نگاه کل نگر و تقاطعی به مسائل

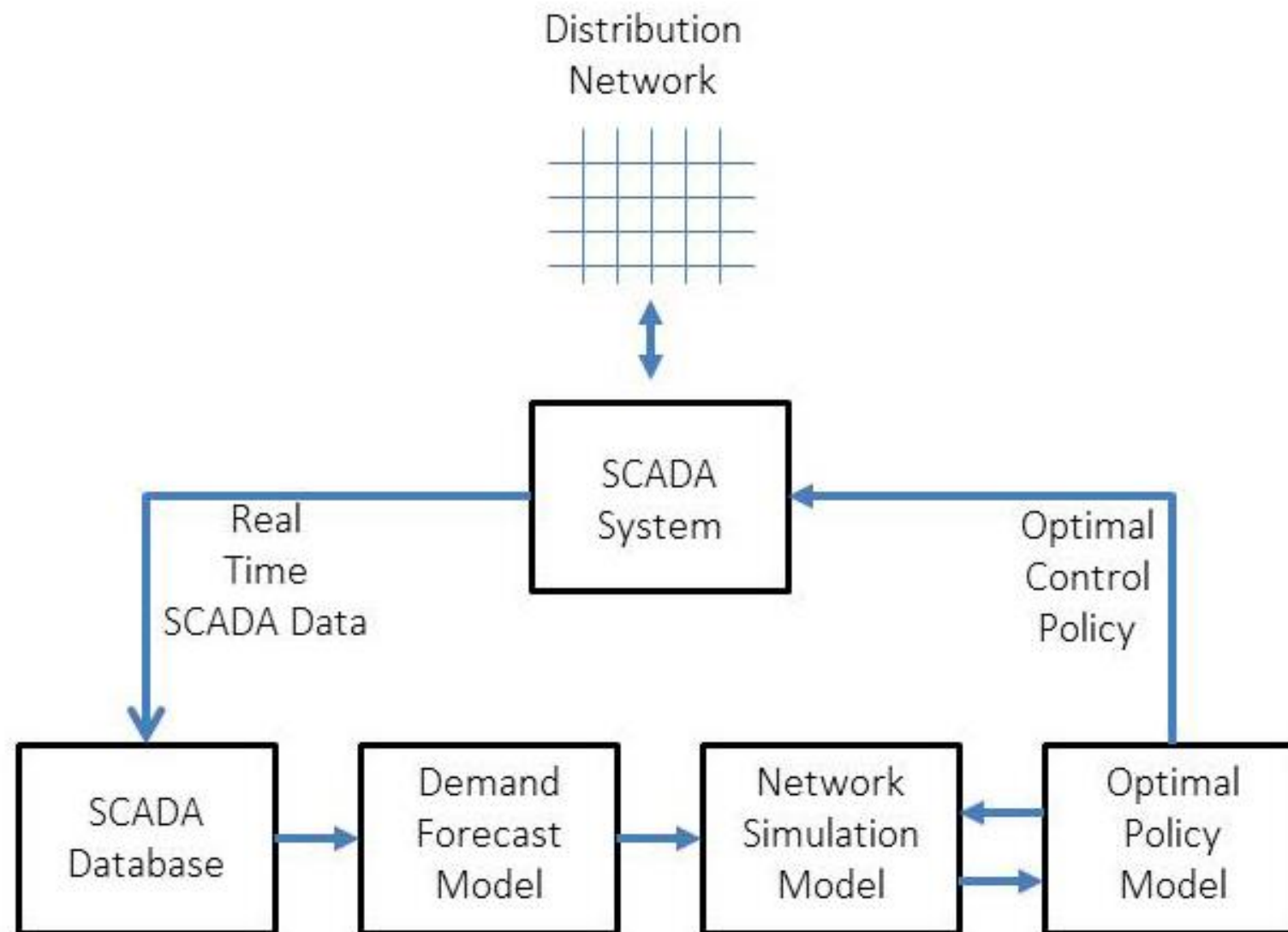
کاهش کیفیت تصمیم گیری

افزایش هزینه‌های عملیاتی

تعارض و هم‌پوشانی نقش‌ها و مسئولیت‌ها

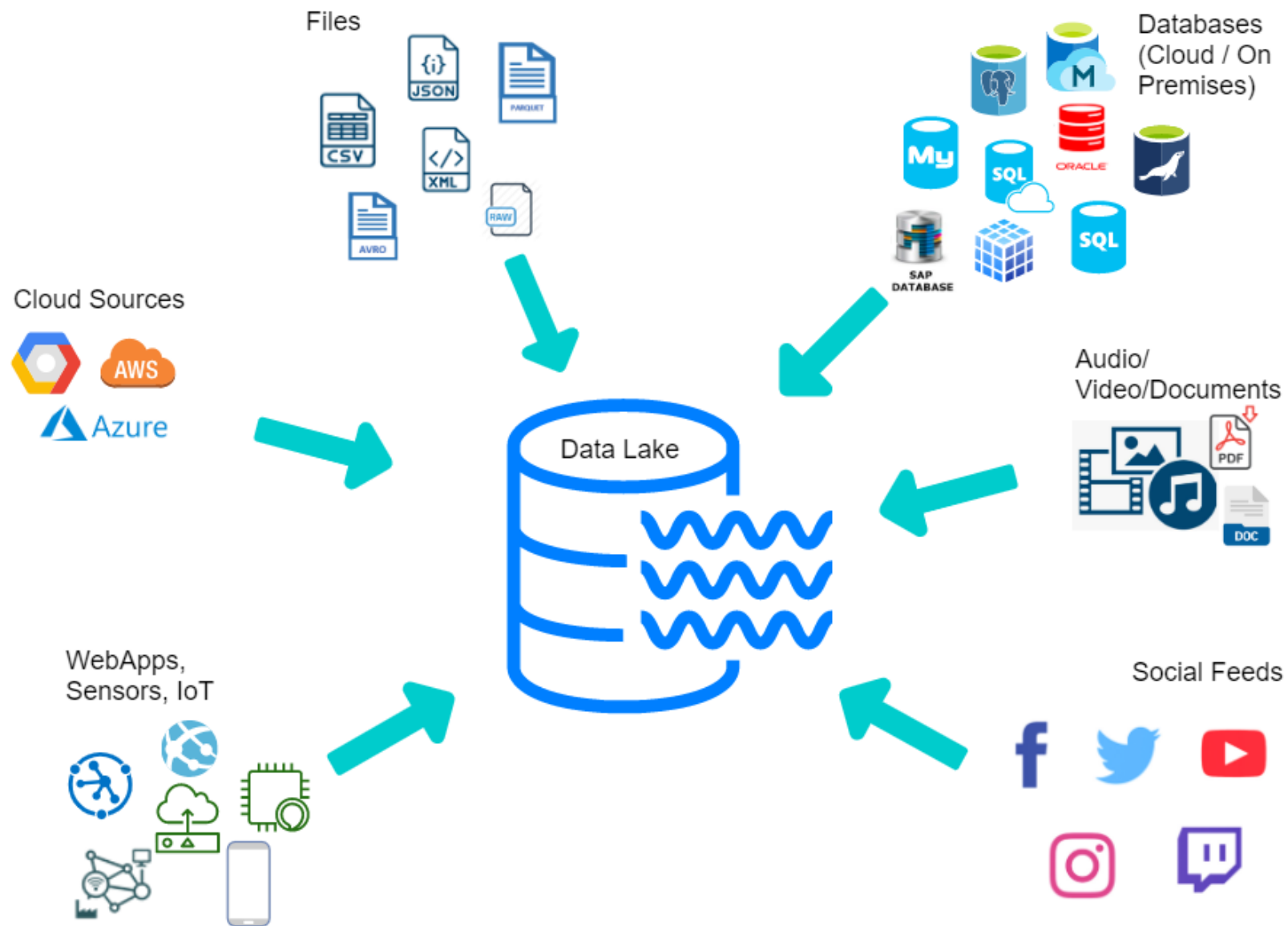


برای پیش‌بینی و بهینه‌سازی مصرف و کیفیت توزیع آب به چه منابع داده‌ای جهت ادغام و یکپارچه‌سازی نیاز دارید؟



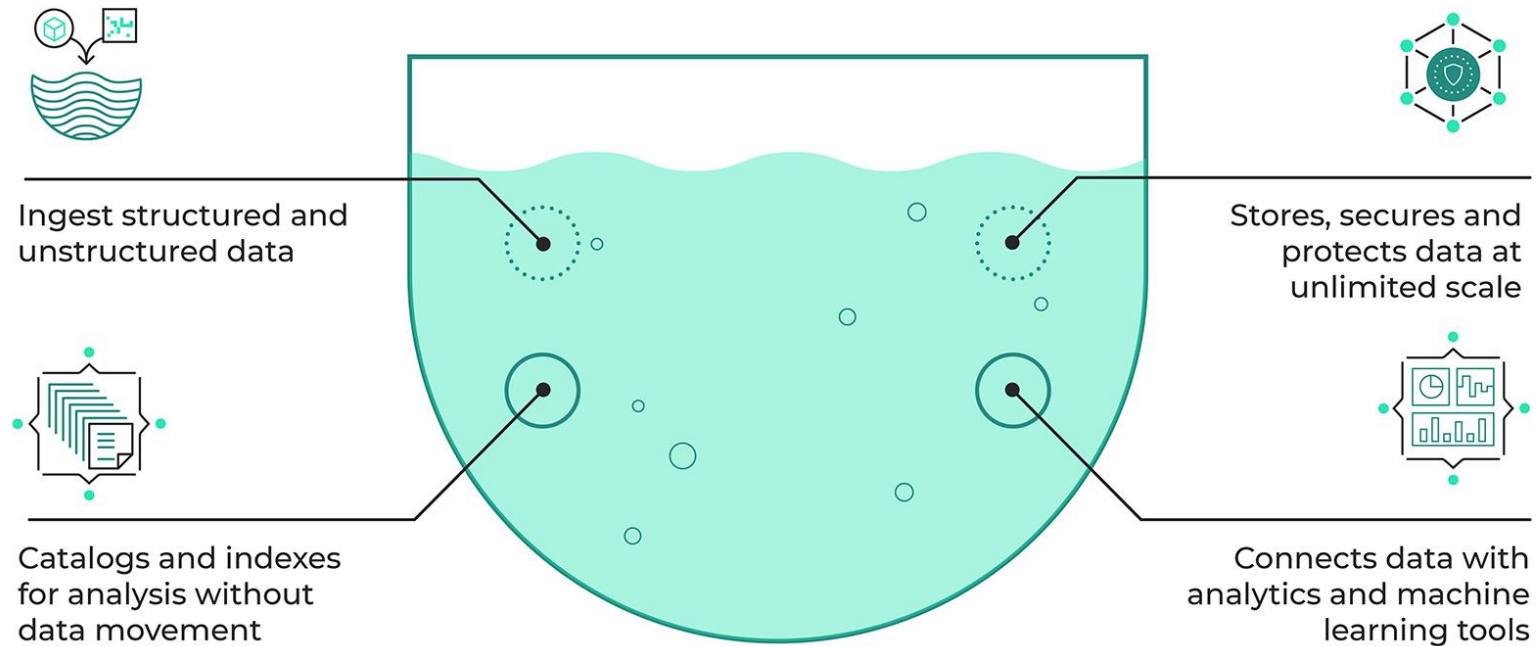
خروجی‌های یکپارچگی داده





دریاچه داده

Data Lakes Features

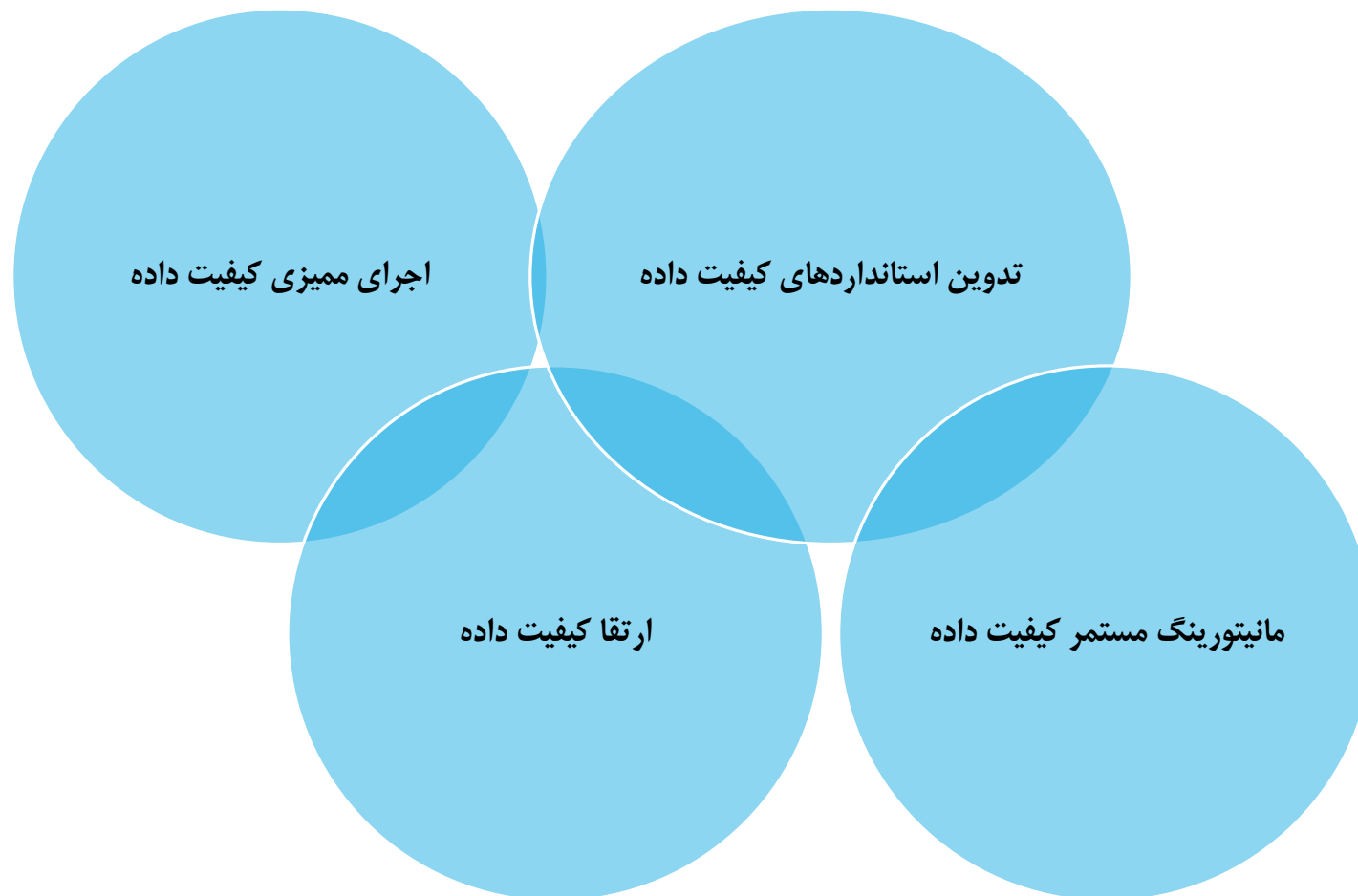


دریاچه داده



کیفیت داده

مدیریت کیفیت داده



خروجی‌های مدیریت کیفیت داده

استراتژی و چارچوب کیفیت داده

ساختار تیم کیفیت داده

توصیه اقدامات اصلاحی

عوامل ریشه‌ای مشکلات

تدوین روال‌ها، سیاست‌ها و راهنماهای کیفیت داده

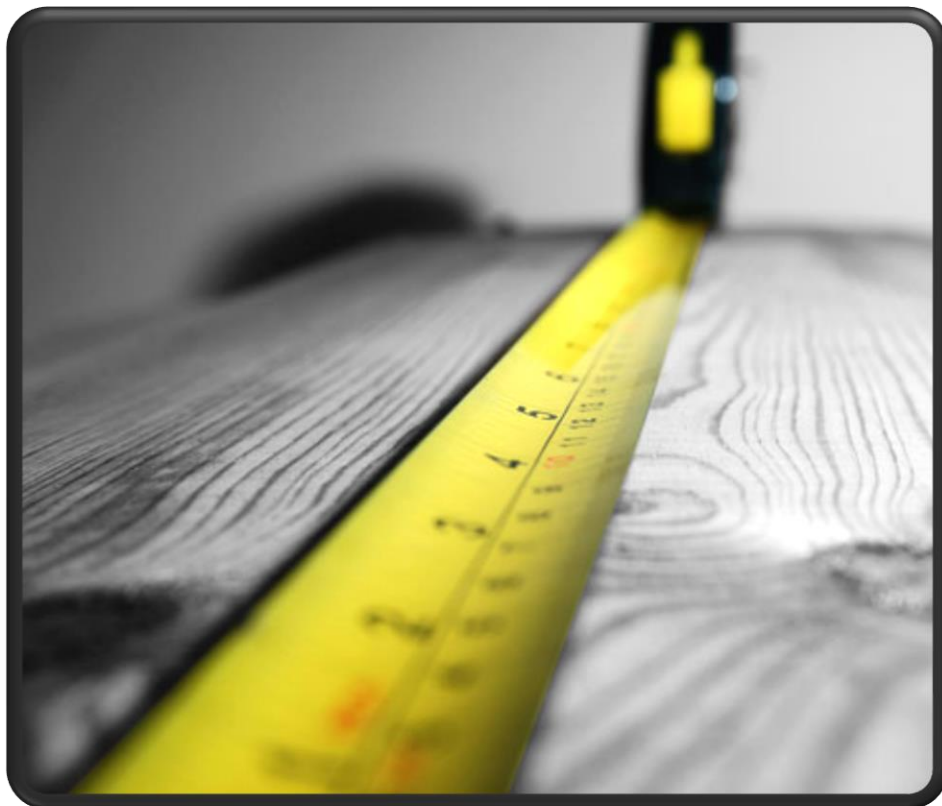
گزارش‌های کیفیت داده

توافق نامه سطح سرویس کیفی داده

سیاست‌ها و راهنمای کیفیت داده

گزارش‌های حاکمیتی کیفیت داده

ابعاد اندازه‌گیری کیفیت داده



- صحت
- کامل بودن
- سازگاری در منابع مختلف
- به هنگام بودن
- ارزش افزوده
- قابلیت تفسیر
- در دسترس بودن

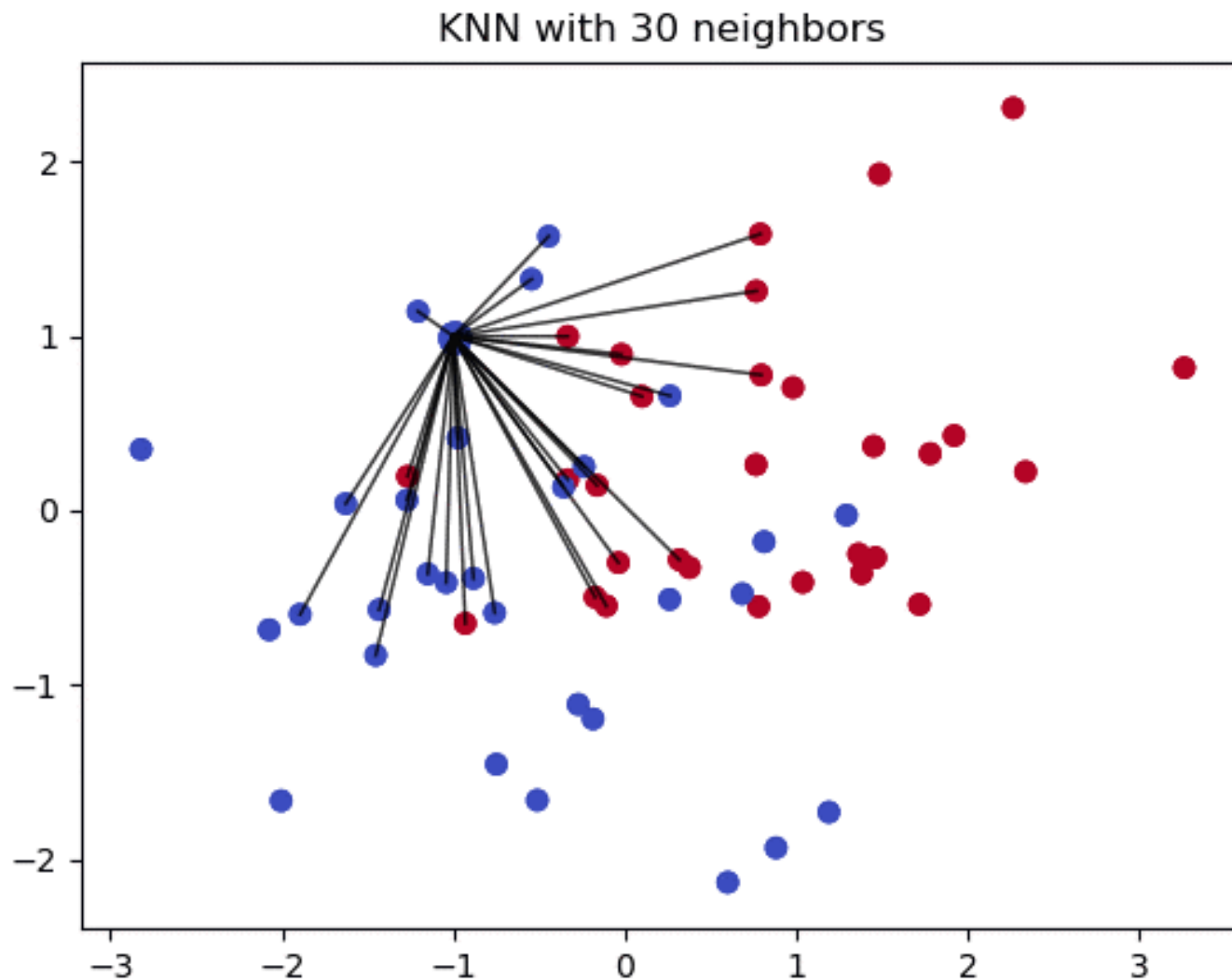
مدیریت کیفیت داده

- مدیریت داده‌های مفقوده
- مدیریت داده‌های نویزی
- مدیریت داده‌های پرت و ناهنجار
- تبدیل و نرمال سازی داده
- کیفی سازی دادگان کیفی
- ناسازگاری
- داده تکراری

مدیریت داده‌های مفقود

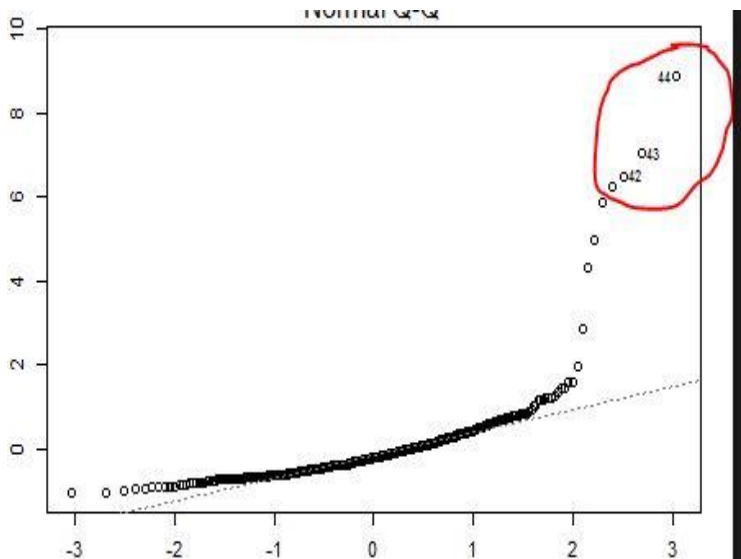
- استفاده از یک معیار مرکزی (میانگین / میانه) ویژگی مورد نظر در سایر رکوردها
- میانگین ویژگی مورد نظر در سایر رکوردهای متعلق به طبقه ای که ویژگی مفقوده آن باید جایگزین شود (مانند فشار خون برای رده های سنی مختلف)
- استفاده از محتمل ترین مقدار (مانند مد): مانند فشار خون محتمل ۱۲ که بیشتر شایع است
- مقادیر با احتمال بیشتر؛ فرمول‌های بیز، درخت تصمیم، KNN یا رگرسیون



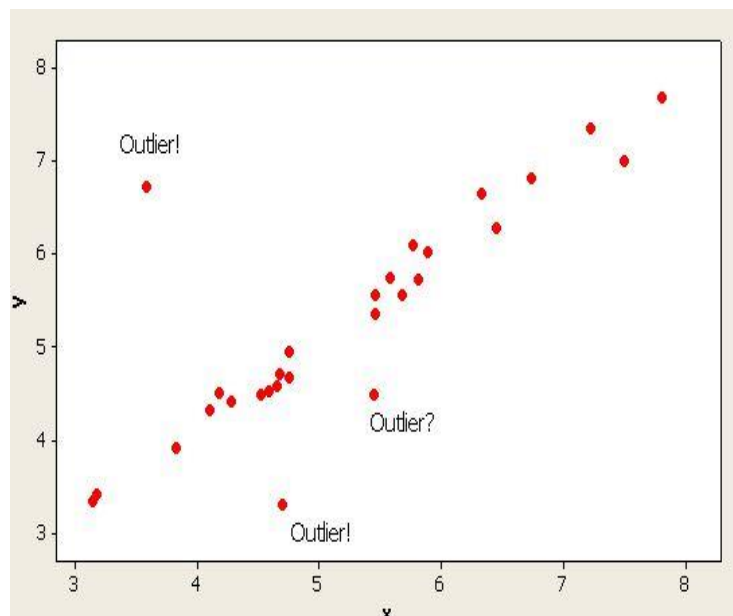


مدیریت داده‌های مفقود
با روش KNN

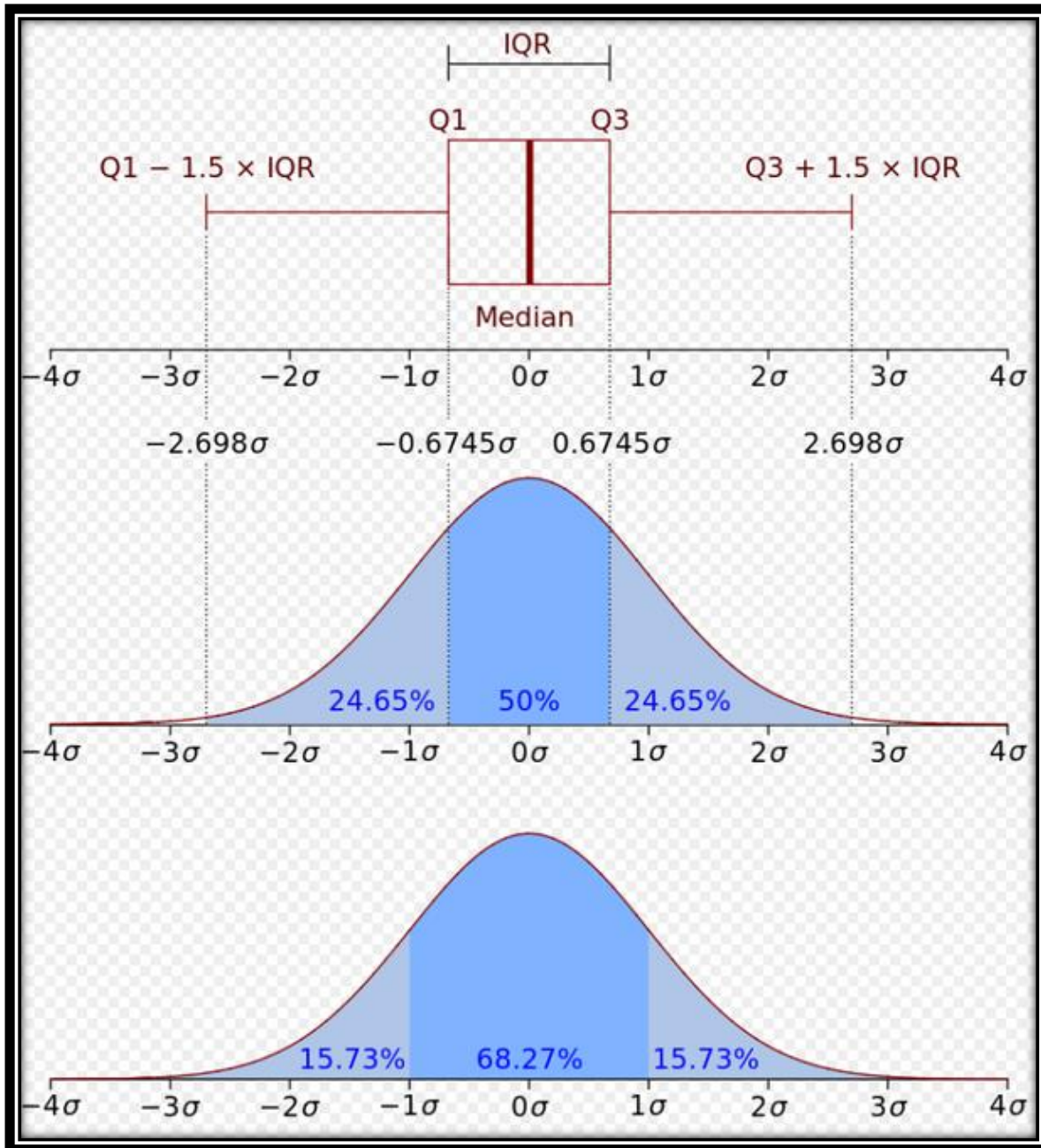
داده پرت (Outlier)



➤ یک داده پرت مشاهده‌ای است که به طور غیرعادی یا اتفاقی از وضعیت عمومی داده‌ها انحراف داشته باشد.



➤ داده‌های پرت داده‌هایی هستند که با سایر داده‌ها تفاوت دارند و این تفاوت به حدی است که به نظر می‌رسد این داده‌ها توسط توزیع یا مکانیزم متفاوتی از سایر داده‌ها تولید شده‌اند.



مدیریت داده‌های پرت

نرمال سازی

نرمال سازی متغیرها، به این معناست که مقادیر مختلف مشاهدات در یک مقیاس مشخص تنظیم می شوند. رایج ترین مقیاس در نرمال سازی مقیاس [۰-۱] است.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

- نرمال کردن در اینجا به معنای بی مقیاس کردن است.
- با نرمال کردن متغیرهای عددی ضریب اهمیت هریک از متغیرها برابر خواهد شد.

نرمال سازی

ID	Gender	Age	Salary
1	F	27	19,000
2	M	51	64,000
3	M	52	100,000
4	F	33	55,000
5	M	45	45,000

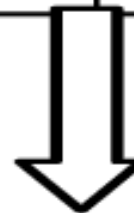
ID	Gender	Age	Salary
1	1	0.00	0.00
2	0	0.96	0.56
3	0	1.00	1.00
4	1	0.24	0.44
5	0	0.72	0.32

Person	Marital status
xxx	Single
yyy	Married
zzz	Divorcee



Categorical to binary

Person	Single	Married	Divorcee
xxx	1	0	0
yyy	0	1	0
zzz	0	0	1



Binary to numeric

Person	Martital status
xxx	4
yyy	2
zzz	1

تبدیل داده‌ها



مدیریت محتوا و اسناد

چالش‌های مدیریت اسناد و محتوا

عدم وجود متولی
اسناد

عدم وجود مکان
مناسب برای نگهداری
اسناد

عدم وجود ابزارهای
مناسب جهت شناسایی
اسناد

عدم دسترسی به اسناد

عدم وجود ابزارهای
حفاظت از محتوای
اسناد

اسناد در حال تخریب

ورودی‌های سیستم مدیریت اسناد و محتوا

قوانین بالادستی

استراتژی‌ها

فایل‌های متنی

ملزومات قانونی

فایل‌های الکترونیکی

مستندات کاغذی

جریان داده شبکه‌های اجتماعی

ابزارهای مدیریت اسناد و محتوا

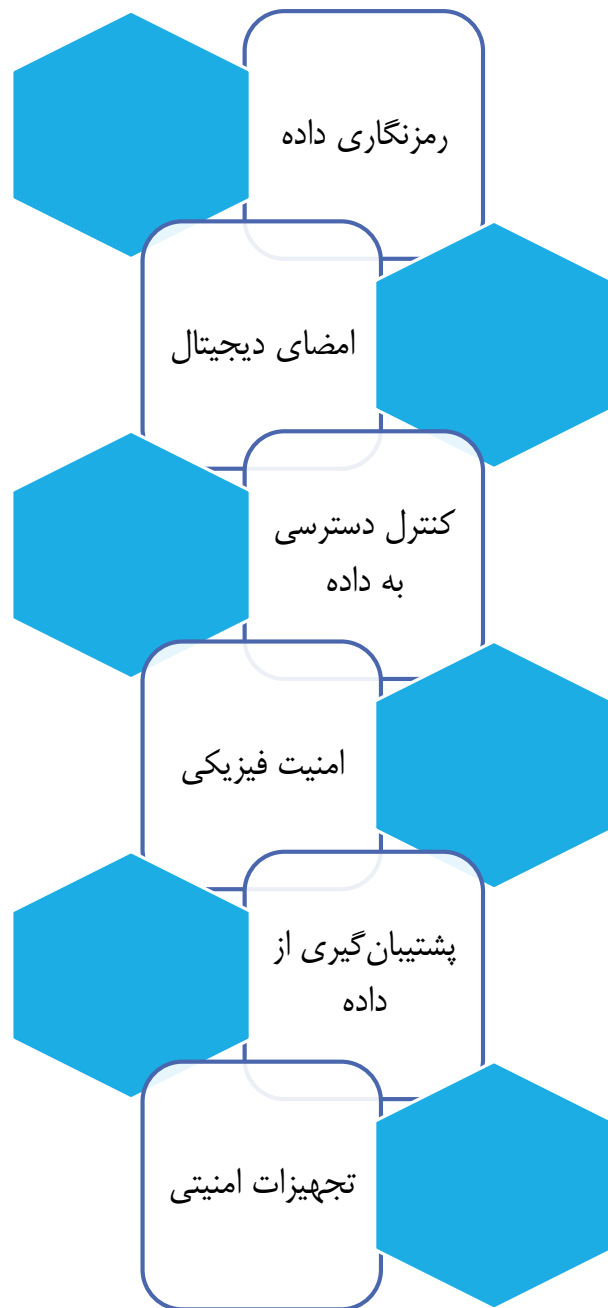
سیستم‌های پردازش تصویر

سیستم‌های مدیریت اسناد

سیستم‌های مدیریت دانش



امنیت داده



اصول مدیریت امنیت داده

فرایند مدیریت امنیت داده

مدیریت انطباق داده

- تأمین قوانین، مقررات و استانداردهای داخلی و خارجی

مدیریت امنیت داده

- تأمین الزامات امنیتی
- تعیین مالکیت داده

ممیزی امنیت داده و انطباق آن

- بررسی روش‌مند و ادواری ممیزی امنیت و انطباق داده
- ارائه پیشنهادهای لازم جهت انجام اقدام‌های مناسب توسط گزارش‌های ممیزی
- بررسی جنبه‌های امنیتی داده به منظور اطمینان از تأمین نیازمندی‌های امنیتی

مدیریت گزارش‌های ممیزی امنیت و انطباق

- مدیریت گزارش‌های تولید شده حاصل فرایند ممیزی امنیت و انطباق
- اطمینان از در دسترس بودن گزارش‌های ممیزی توسط افراد مجاز و در زمان مناسب

خروجی‌های مدیریت امنیت داده

معماری امنیت داده‌ها

سیاست امنیت داده‌ها

استانداردهای محرمانگی داده‌ها

کنترل دسترسی

تطابق با مقررات

طبقه بندی امنیتی

شناسایی کاربران و سطوح دسترسی

گزارش ارزیابی امنیت داده

قانون GDPR

- GDPR مخفف عبارت General Data Protection Regulation به معنی «مقررات محافظت از داده‌های عمومی» است. قانون GDPR در پارلمان اتحادیه اروپا به تصویب رسیده و به همین خاطر در کشورهای عضو این اتحادیه و کشورهای وابسته لازم‌الاجرا است.
- طبق این قانون، شهروندان اروپایی مالک داده‌های خود هستند و شرکت‌هایی که این داده‌ها را جمع‌آوری می‌کنند، برای این داده‌ها از حق مالکیت بهره نمی‌برند.
- با اجرایی شدن این قانون، شرکت‌هایی که مرتکب خلاف شوند، تا ۲۰ میلیون یورو یا ۴ درصد از گردش مالی جهانی خود (هرکدام که بیشتر باشد) جریمه می‌شوند.

قانون GDPR

□ طبق این قانون، شرکت‌ها باید بدانند که چه اطلاعاتی را و چرا جمع‌آوری می‌کنند. همچنین برای استفاده از داده‌های شخصی کاربران، باید به زبان روشن و ساده از کاربران اروپایی خود اجازه بگیرند. علاوه بر این، کاربران می‌توانند از شرکت موردنظر بخواهند که داده‌های شخصی آن‌ها را ارائه دهد و در صورت درخواست کاربران، شرکت مذکور باید توضیح دهد که چنین داده‌هایی در چه مواردی مورد استفاده قرار گرفته است.



هک بزرگ



حکمرانی داده

حکمرانی داده



اهداف پروژه حکمرانی داده

کنترل و هدایت (برنامه ریزی، پایش، الزام) بر مدیریت دارایی‌های اطلاعاتی

نظارت، ابلاغ و اجرای اصول، قواعد، رویه‌ها، سنجه‌ها، ابزارها و مسئولیت‌های مدیریت داده

نظارت و هدایت تطابق با سیاست‌ها، استفاده از داده‌ها

حاکمیت تولید، مدیریت، ذخیره، و در دسترس گذاشتن داده‌ها

حکمرانی داده

چه کسی مسئول تصمیم گیری است؟

چه سیاست‌ها بایست تدوین گردد؟

چه تصمیماتی لازم هست اتخاذ شود؟

تفاوت مدیریت داده و حکمرانی داده

DATA MANAGEMENT	VS	DATA GOVERNANCE
Logistics or methods of how data is organized	WHAT	Policies, controls, or rules for how data is governed and data quality is achieved
Primarily the responsibility of IT to implement framework to manage data	WHO	Multiple members of an enterprise holistically build a framework (data stewards, data citizens, and more) for data management
An umbrella term, covering all aspects (including data governance) of how an enterprise uses and manages its data	PURPOSE	The first building block of data management, focusing on the framework to achieve business goals and reduce risk
Logistical, focused on technology	GOAL	Philosophical, focused on an overall business strategy

PEERNova

محدوده پروژه حکمرانی داده

استراتژی داده

تعریف، اطلاع رسانی و هدایت اجرایی استراتژی‌های داده و استراتژی‌های حکمرانی داده

سیاست‌ها

تعیین سیاست‌های مرتبط با مدیریت، دستیابی، استفاده، امنیت و کیفیت داده و فراداده و الزام اجرای آن

کیفیت و استانداردها

تعیین و الزام استانداردهای کیفیت داده و معماری داده

نظارت

مشاهده، ممیزی، اصلاح حوزه‌های کلیدی کیفیت، سیاست‌ها و مدیریت داده

انطباق

حصول اطمینان از اینکه سازمان مقررات قانونی و بالادستی مربوط به داده‌ها

مدیریت مشکلات

شناسایی، تعریف، طبقه بندی و حل مشکلات مرتبط با فرایندهای مدیریت داده

پروژه‌های مدیریت داده

تعریف و حمایت پروژه‌های بهبود مدیریت داده

ارزش گذاری دارایی‌های اطلاعاتی

تدوین استانداردها و فرایندهایی برای تعیین ارزش دارایی‌های اطلاعاتی به شیوه‌ای سازگار

استراتژی داده

طراحی سیاست‌های در جمع‌آوری و تکمیل منابع داده‌ای

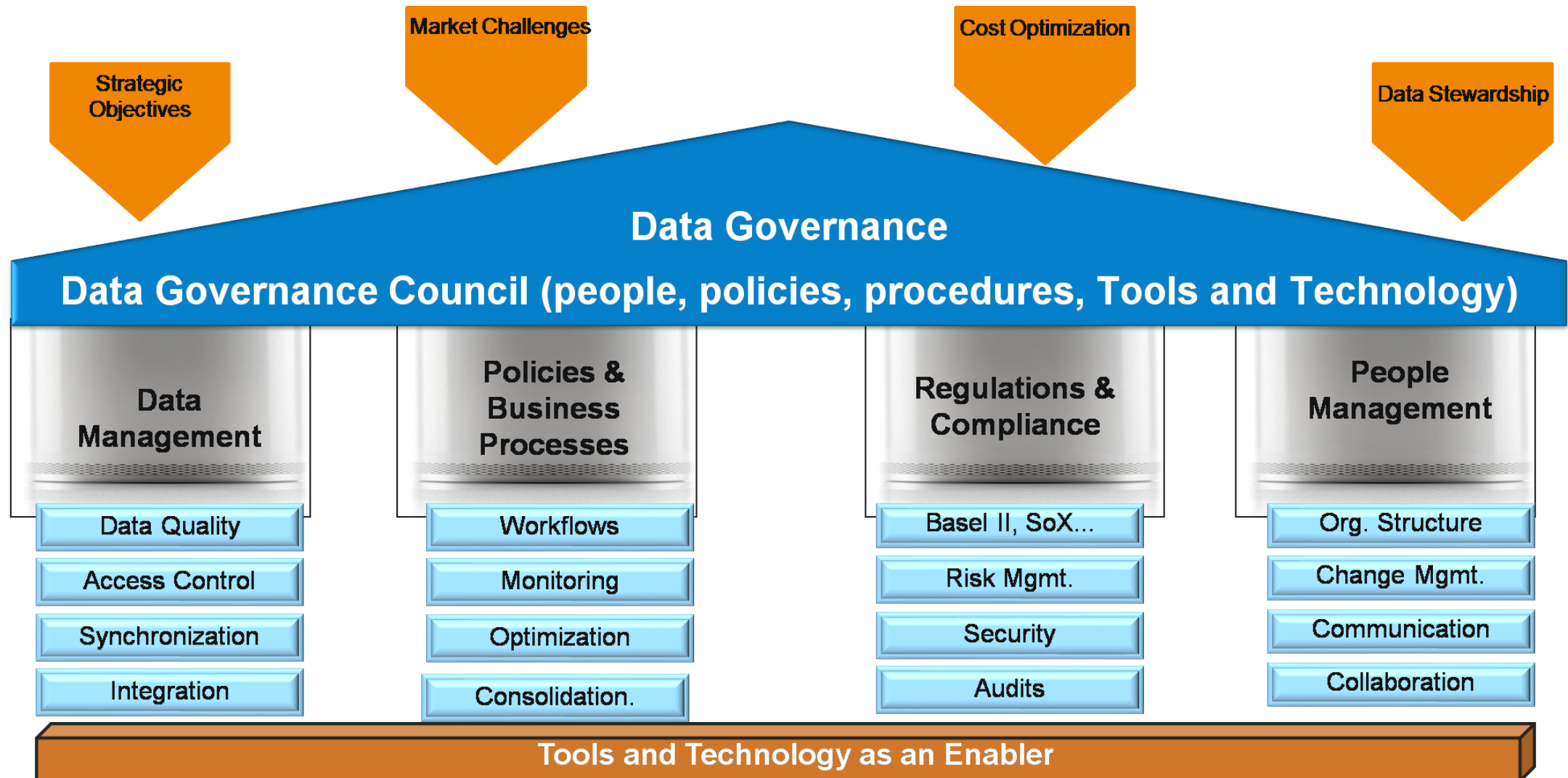
تدوین برنامه طرح رشد کارکنان و جذب استعدادهای داده / تحلیل داده

بررسی طرح‌های اجرایی مبتنی بر کسب درآمد از داده

شناسایی و اولویت‌بندی پروژه‌های کلیدی زیرساخت، جمع‌آوری و تحلیل داده

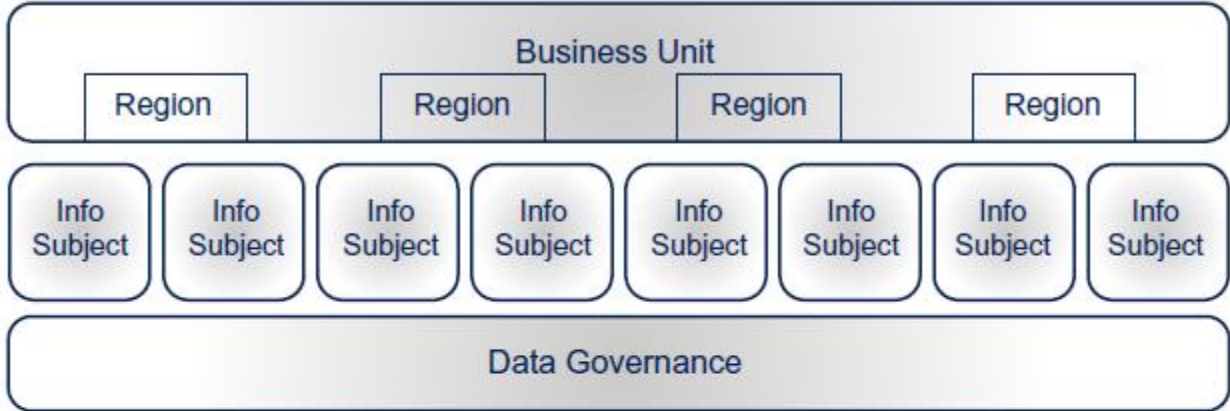
تدوین نقشه‌راه استقرار فرهنگ تصمیم‌گیری داده‌محور

ارکان حاکمیت داده

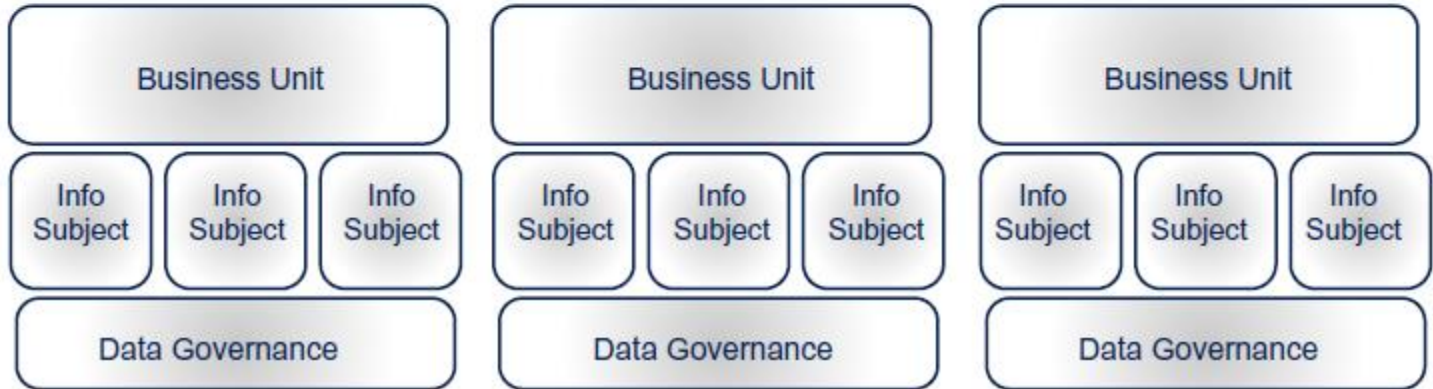


ساختار حکمرانی داده

Company A

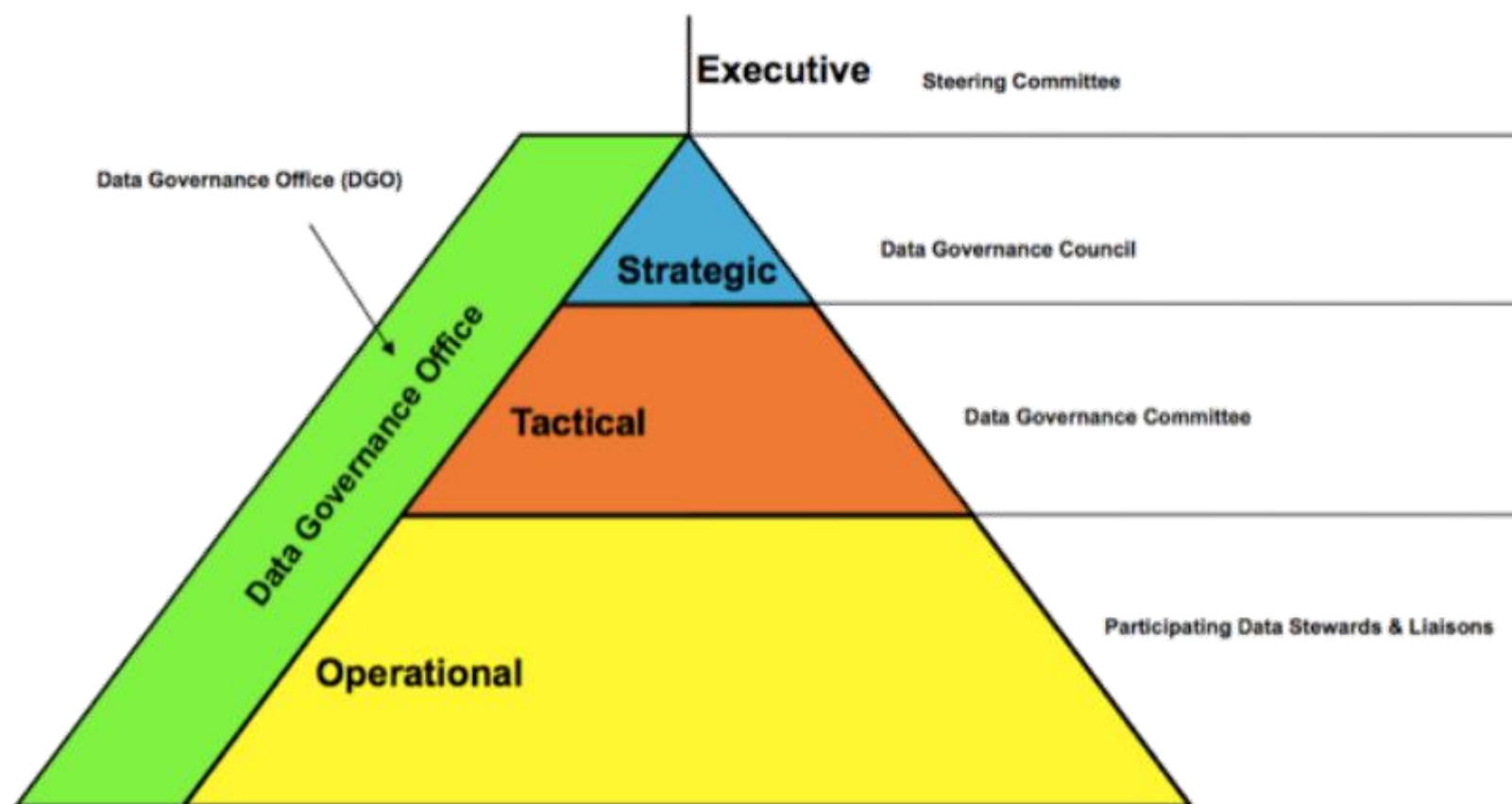


Company B



ساختار شورا و کمیته حکمرانی داده

Data Governance Body



ساختار شورا و کمیته حکمرانی داده

کمیته راهبری حکمرانی داده

مسئولیت نظارت، پشتیبانی و تامین بودجه فعالیت ها و پروژه‌های مدیریت داده و بالاترین سطح تصمیم‌گیر اقدامات حکمرانی داده



شورای حکمرانی داده

تدوین سیاست‌ها، شاخص‌ها و مدیریت چالش‌های مدیریت و حکمرانی داده



دفتر حکمرانی داده

تمرکز بر استقرار استانداردها مدیریت داده در سطح سازمان

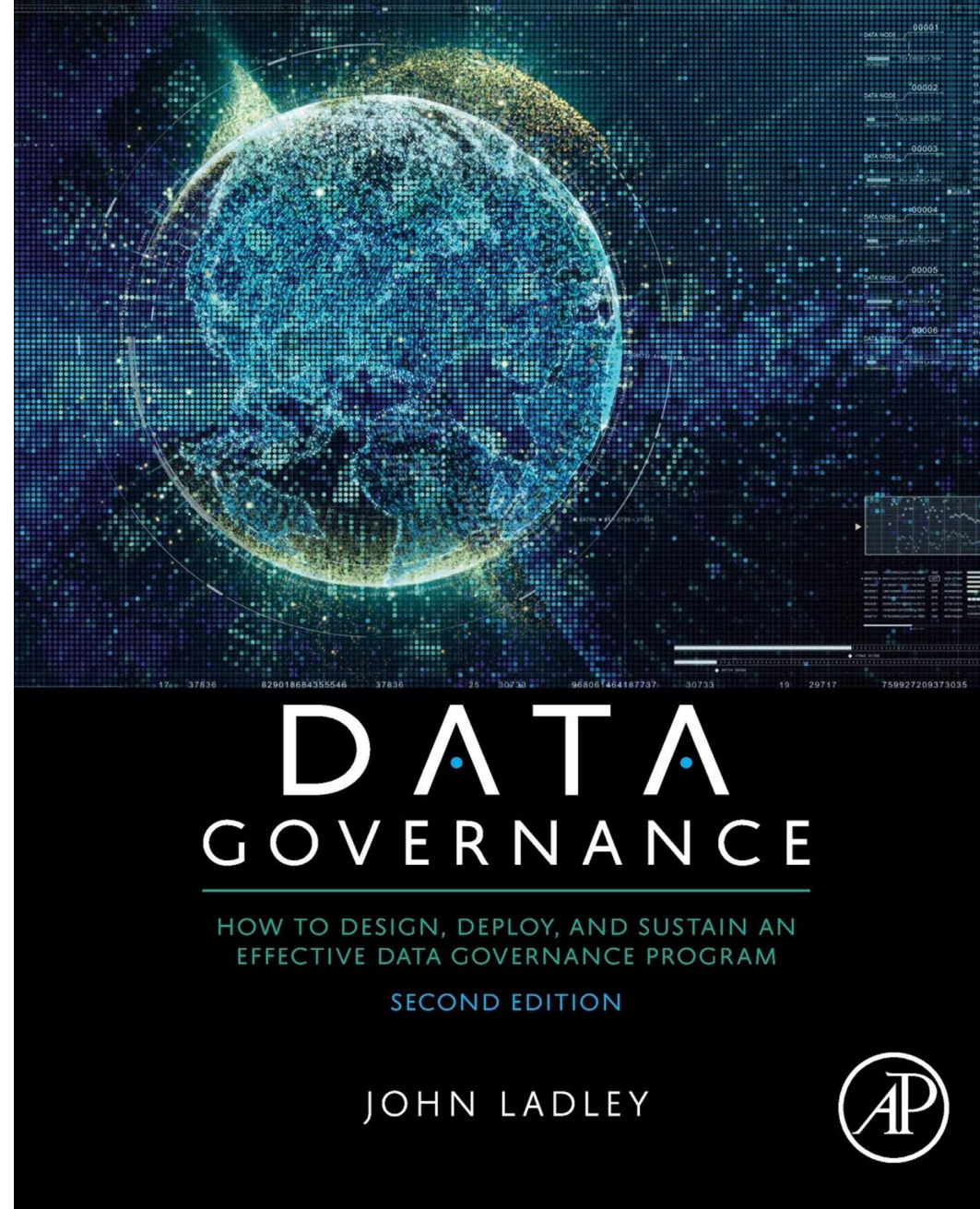


تیم داده‌بان

تمرکز بر جنبه‌های محدود یا پروژه‌های خاص مدیریت داده (بررسی کیفیت و اطمینان از داده‌های موضوعی)

وظایف کلیدی داده بان



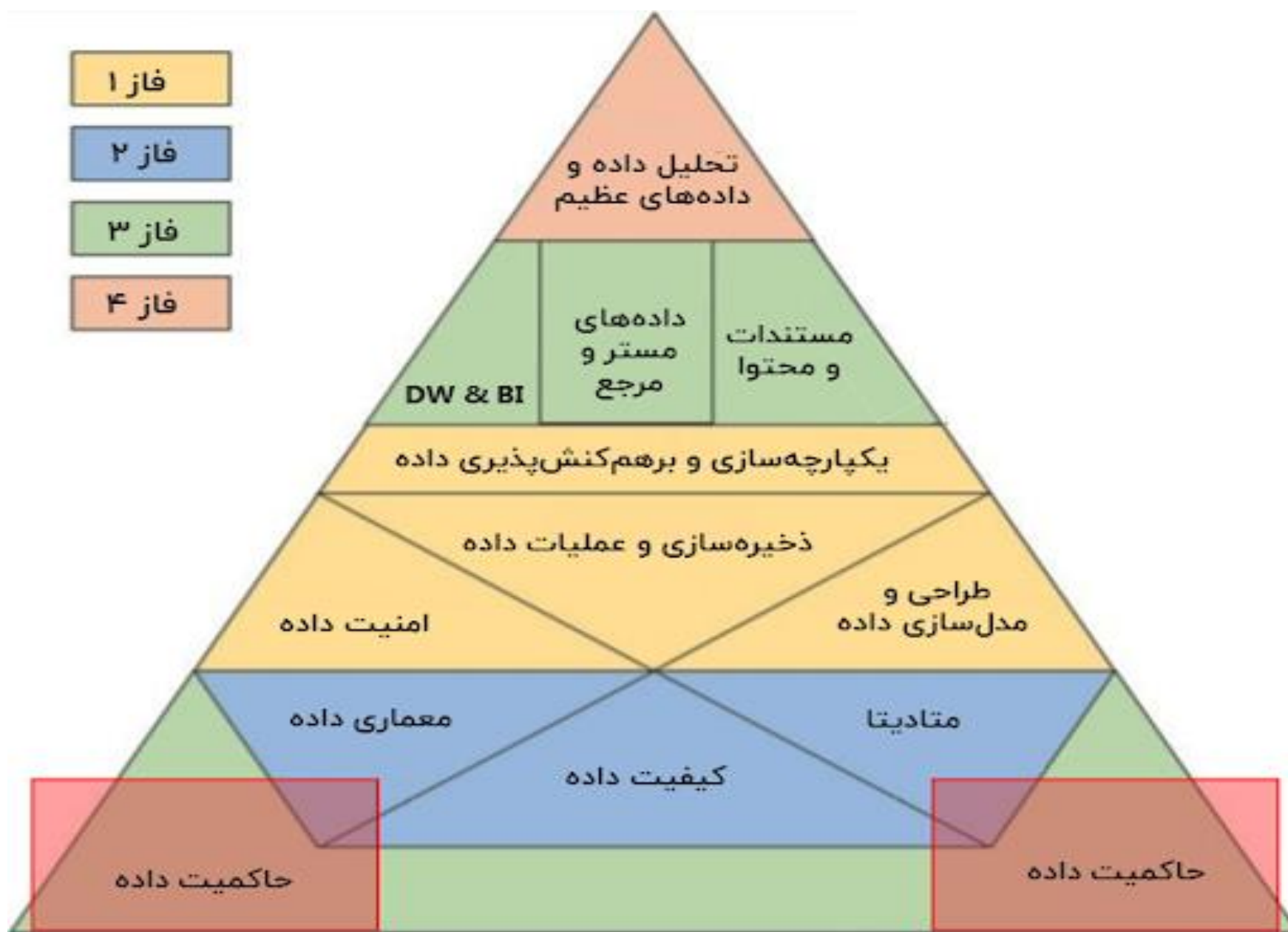


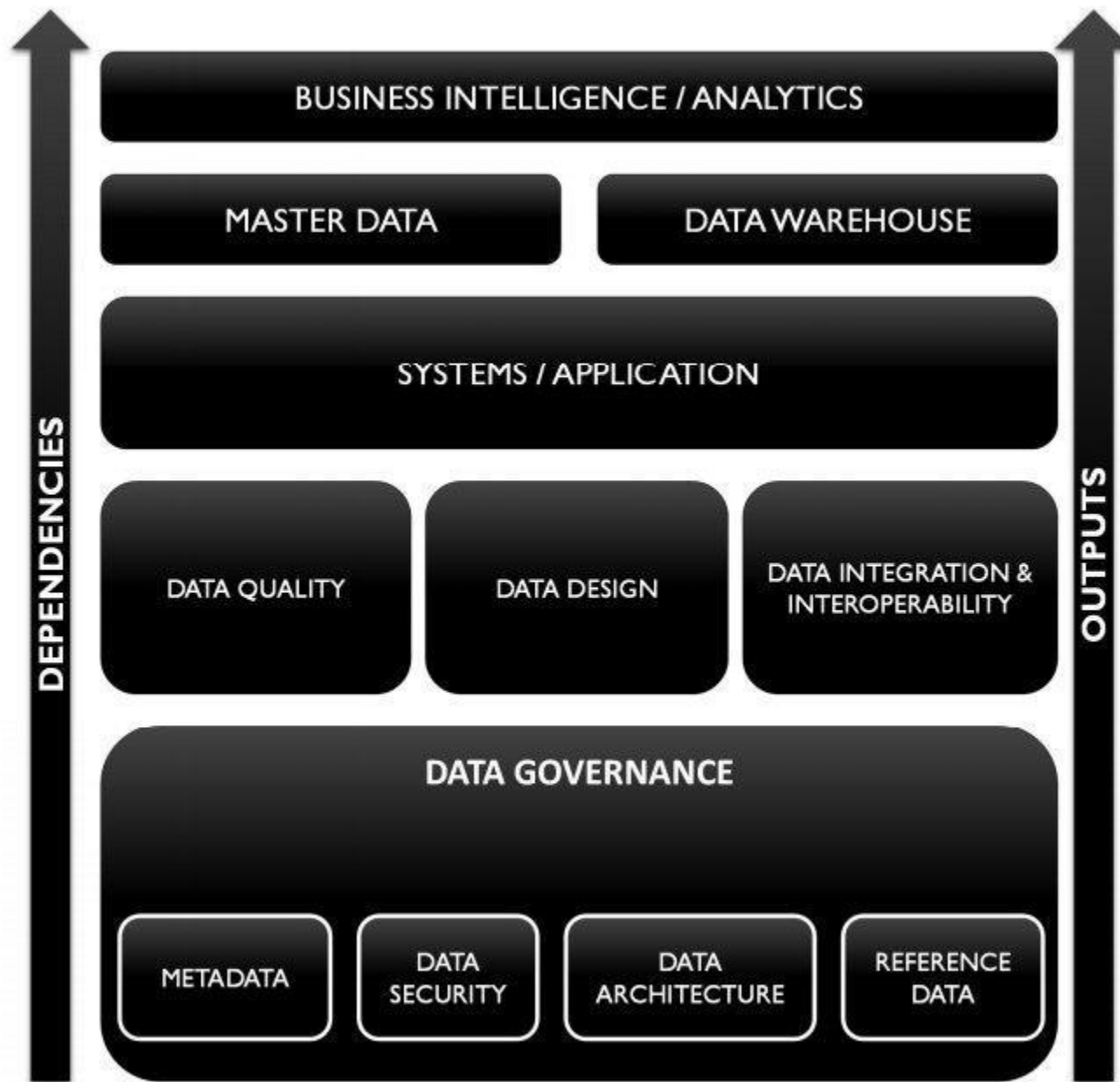


ارزیابی بلوغ

سفری به سوی تعالی سازمانی

فازبندی





چارچوب کلان معماری پروژه مدیریت داده

TDWI

IDC

Hortonworks

Academia

DMM

IBM

DCAM

Knowledgent

Booz & Company

Temporal

Van Veenstra

Coleman

Infotech

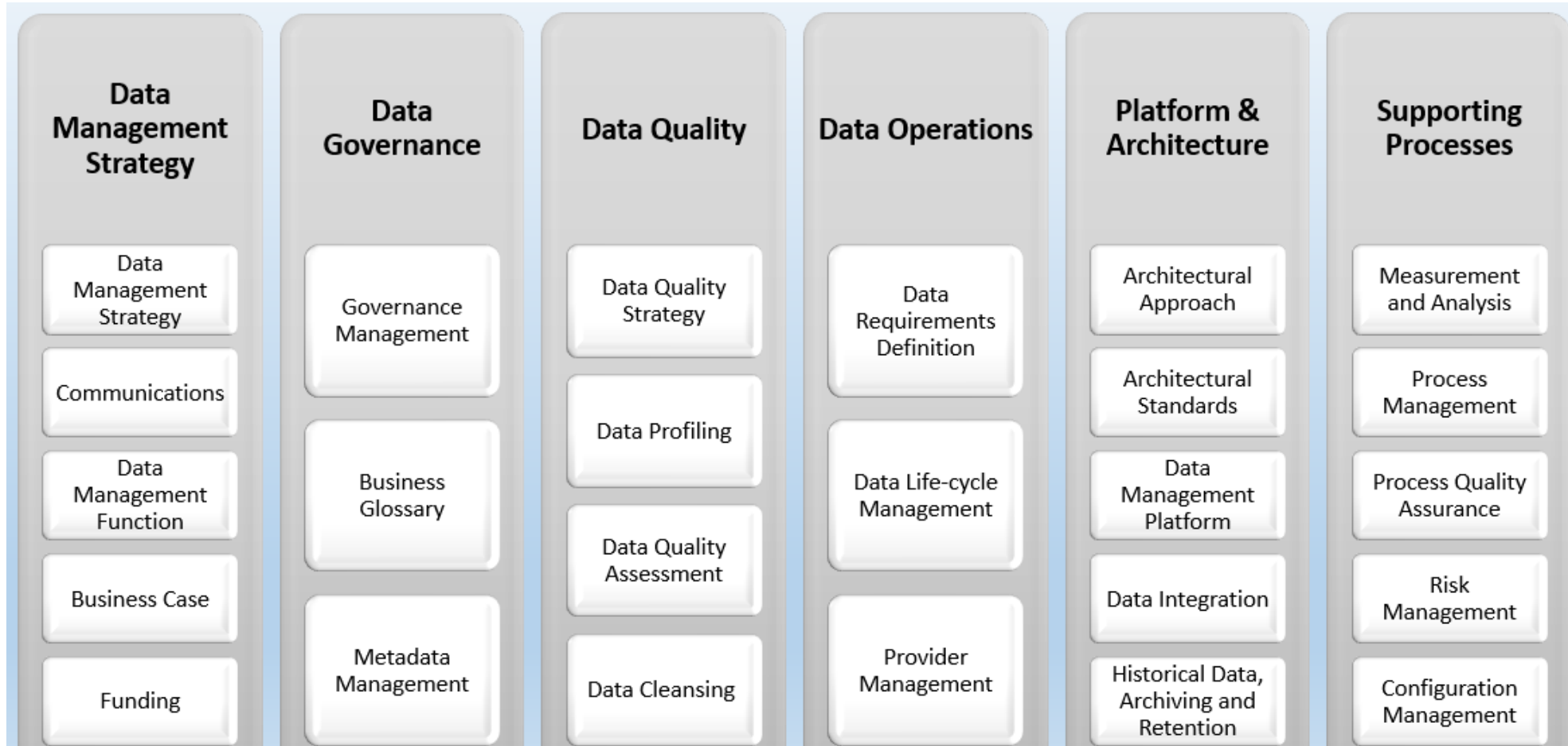
Radcliffe Advisory

Alejandro Pena

EMC

مدل‌های ارزیابی بلوغ بین‌المللی

مدل بلوغ DMM



مشکلات مدل‌های معمول ارزیابی بلوغ داده

جعبه‌سیاه بودن مدل‌های ارزیابی بلوغ (عدم ارائه پرسشنامه یا ارائه سوالات محدود)

عدم جامعیت و بررسی مولفه‌های محدود

لحاظ کردن وزن یکسان به تمامی مولفه‌ها در ارزیابی بلوغ

مدل ارزیابی بلوغ داده‌محوری

خلاصه مدیریتی:

- ❑ مدل بلوغ داده‌محوری یک مدلی جامع جهت ارزیابی ۳۶۰ درجه کسب‌وکار در ۳ جنبه کلان و ۱۶ مولفه تاثیرگذار نرم‌افزاری و سخت‌افزاری است.
- ❑ مدل فوق براساس بومی‌سازی و تقاطع ۱۶ مدل مرجع ارزیابی بلوغ داده‌های عظیم و علوم داده بین‌المللی در دانشگاه تهران توسعه داده شده است.
- ❑ شناسایی نقاط گلوگاه و ضعف سازمانی، رتبه‌بندی کسب‌وکار و ارائه راهکارهای تحول از مزیت‌های کلیدی مدل فوق است.



اهداف مدل ارزیابی بلوغ داده محوری

01. ارزیابی وضعیت بلوغ داده محوری کسب و کار

02. شناسایی نقاط ضعف و فرصت‌های بهبود کسب و کار

03. ارائه گزارش بازخور در راستای ارتقا وضعیت بلوغ

04. بهره‌مندی از مشاوره متخصصین

مزیت‌های مدل ارزیابی بلوغ داده‌محوری



مدل ارزیابی بلوغ داده‌محوری چگونگی رشد قابلیت‌های داده‌محور سازمان را مرحله به مرحله و در امتداد یک مسیر پیش بینی شده مطلوب نشان می‌دهد.

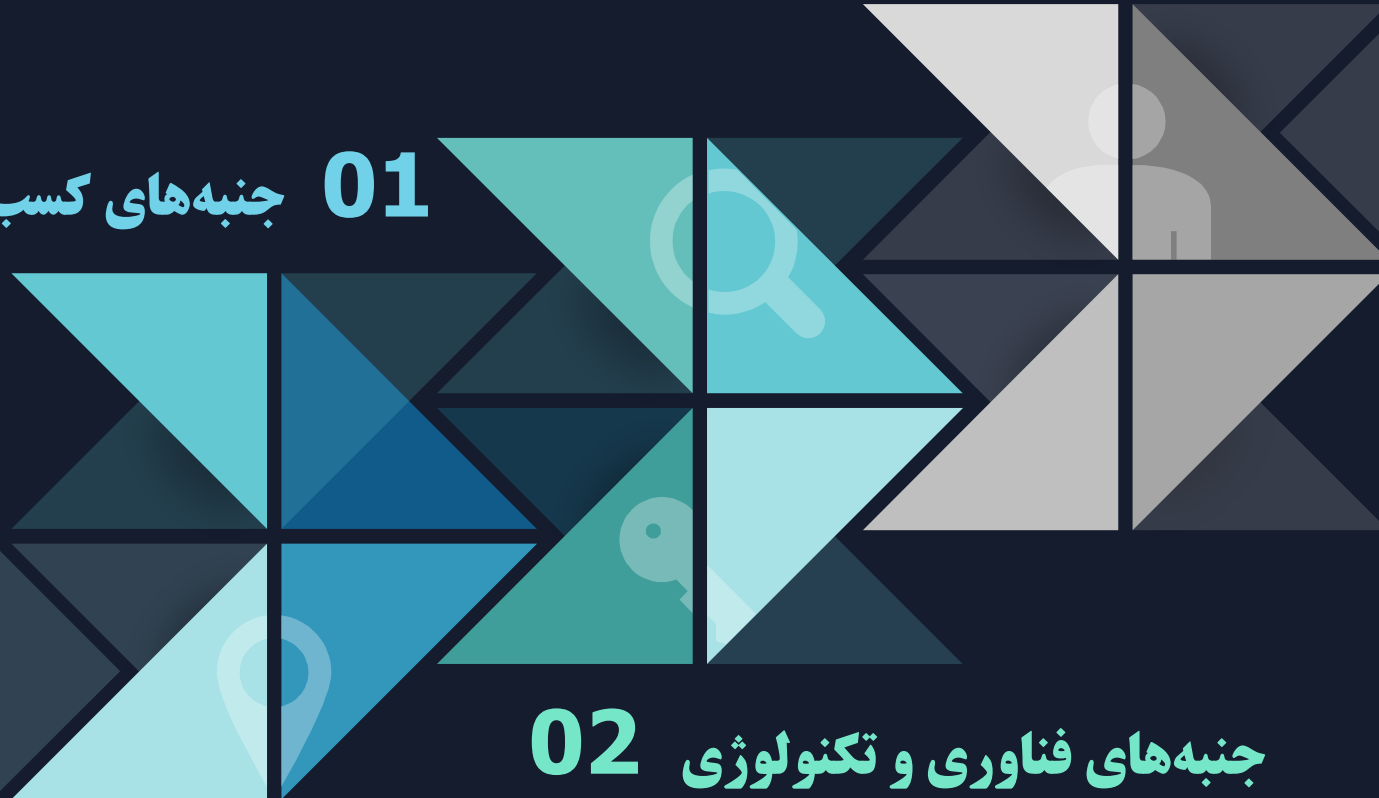
منظرهای کلان مورد پیمایش در ارزیابی بلوغ داده محوری

ارزیابی ۳۶۰ درجه زیرساخت‌های نرم‌افزاری و سخت‌افزاری کسب و کار

03 جنبه‌های مدیریت و حکمرانی داده

01 جنبه‌های کسب و کار و مدیریت

02 جنبه‌های فناوری و تکنولوژی



BUSINESS <i>Strategy</i>	DATA <i>Strategy</i>	BUSINESS <i>Process</i>	Organization
DATA <i>Analytics</i>	DATA <i>Governance</i>	Privacy	Leadership
Technology	DATA <i>Architecture</i>	DATA <i>Security</i>	People.
Infrastructure	DATA <i>Management</i>	DATA <i>Quality</i>	Culture

۱۶+۱ مولفه مورد پیمایش در ارزیابی بلوغ

□ امروزه سازمان‌های اثربخش به دنبال استفاده استراتژیک از داده‌ها همسو با کسب‌وکار خود می‌باشند. داشتن یک استراتژی جامع علوم‌داده و داده‌های عظیم که روح خود را از استراتژی کسب‌وکار می‌گیرد، یکی از الزامات اولیه و حیاتی جهت نیل به موفقیت پایدار در عصر حاضر است.

□ استراتژی کسب‌وکار، استراتژی داده، فرایندهای کسب‌وکار، ساختار سازمانی، رهبری، کارکنان، فرهنگ، تحلیل‌داده، حکمرانی داده، حریم خصوصی، امنیت داده، تکنولوژی، زیرساخت، مدیریت داده، معماری داده، کیفیت داده **به همراه مولفه داده** همگی از مولفه‌هایی هستند که برای ارتقا بلوغ علوم داده و داده‌های عظیم حائز اهمیت می‌باشند. در مدل ارزیابی بلوغ داده‌محوری به تفکیک این ۱۷ مولفه تاثیرگذار مورد ارزیابی بلوغ قرار خواهند گرفت.

مؤلفه‌های فناوری و تکنولوژی مورد پیمایش در مدل ارزیابی بلوغ

01

منابع داده‌ای کسب‌وکار

02

تکنولوژی‌های علوم داده و داده‌های عظیم

03

قابلیت‌ها و زیرساخت‌های تحلیل داده

04

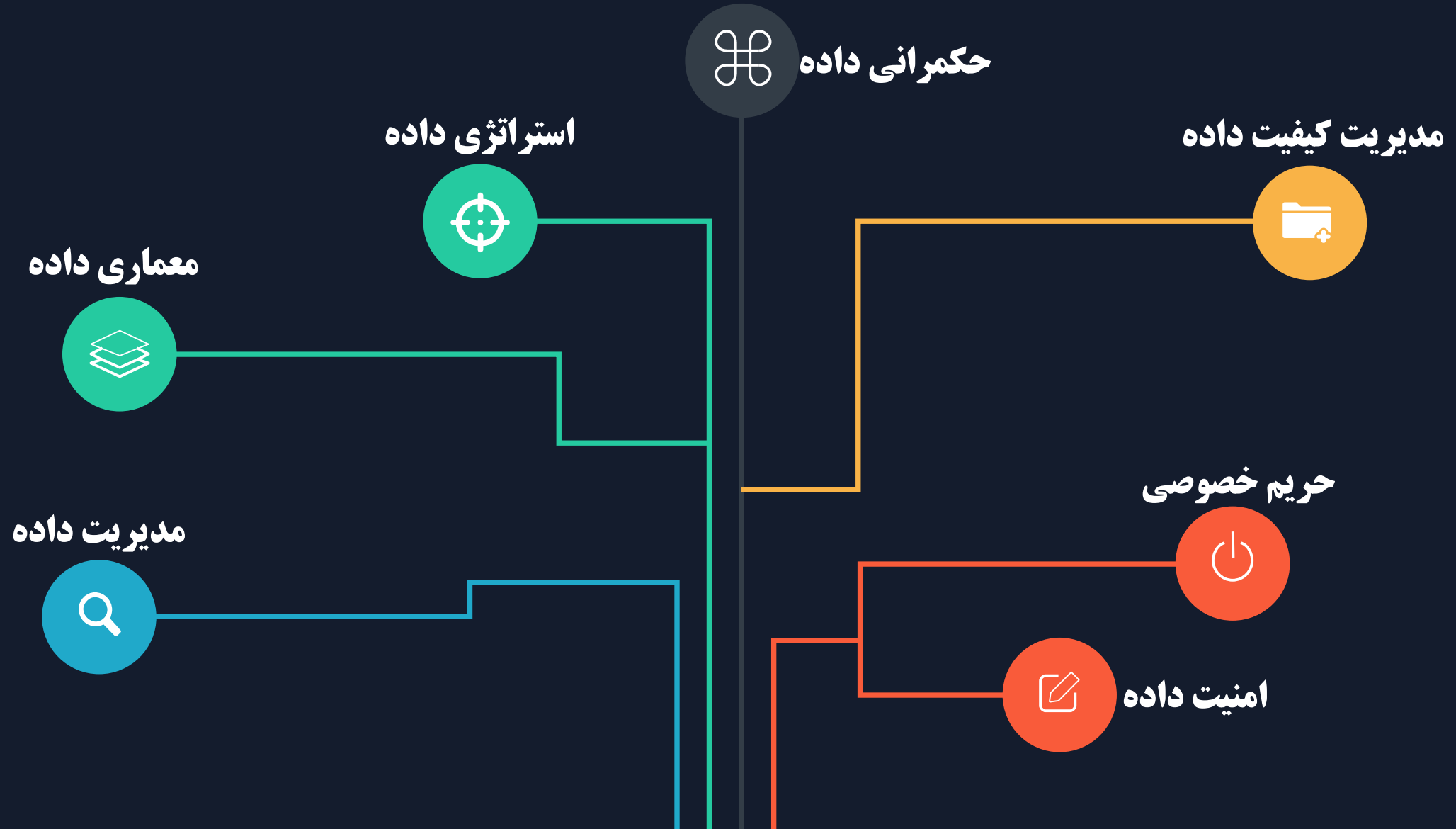
زیرساخت سخت‌افزاری



مولفه‌های کسب‌وکار و مدیریتی مورد پیمایش در مدل ارزیابی بلوغ



مولفه‌های مدیریت و حکمرانی داده مورد پیمایش در مدل ارزیابی بلوغ



سطوح مدل ارزیابی بلوغ داده محوری

ارزیابی کلان سطح بلوغ کسب و کار به همراه ارزیابی تفکیکی ۱۷ مولفه تاثیر گذار در ۵ سطح ذیل

بلوغ بهینه سازی شده

استراتژی واحد در سراسر شرکت

مشارکت فعالانه تمامی مدیران و پرسنل

دانش و مهارت بهینه سازی شده در سراسر سازمان

بهینه سازی فرایندها

بهینه سازی تکنولوژی ها

اختصاص بودجه براساس سند استراتژی داده

بهینه سازی و ارتقا مدیریت داده یکپارچه

بهینه سازی قابلیت ها و فرایندهای تحلیل داده

بلوغ مدیریت شده

استراتژی در سطح چند واحد کسب و کار

مشارکت مدیران فنی و کسب و کار

دانش و مهارت در تمامی واحدها

ارزیابی و توسعه فرایندها

ارزیابی و توسعه تکنولوژی ها

اختصاص بودجه با هدف توسعه

استقرار مدیریت داده

استقرار قابلیت ها و فرایندهای تحلیل داده

بلوغ تعریف شده

استراتژی در سطح واحد کسب و کار

مشارکت مدیران فنی

دانش و مهارت در چند واحد

مانیتورینگ و مستندسازی فرایندها

بهره برداری اولیه از تکنولوژی ها

اختصاص بودجه برنامه ریزی شده

تدوین فرایندها و مستندات مدیریت داده

توسعه قابلیت های تحلیل داده در سطح چند

واحد

بلوغ اولیه

استراتژی در سطح واحدهای فنی

بررسی اولیه مدیریت

دانش و مهارت محدود در واحد فنی

توسعه اولیه فرایندها

توسعه اولیه تکنولوژی ها

اختصاص بودجه با بروز مشکل

مدیریت داده مقدماتی و در سطح واحد فنی

تحلیل داده محدود در سطح واحد فنی

بلوغ آشوب

بدون استراتژی علوم داده و داده عظیم

عدم علاقه مندی مدیریت

خلا دانش و مهارت

عدم وجود فرایندها

عدم وجود تکنولوژی

عدم اختصاص بودجه

عدم وجود مدیریت داده

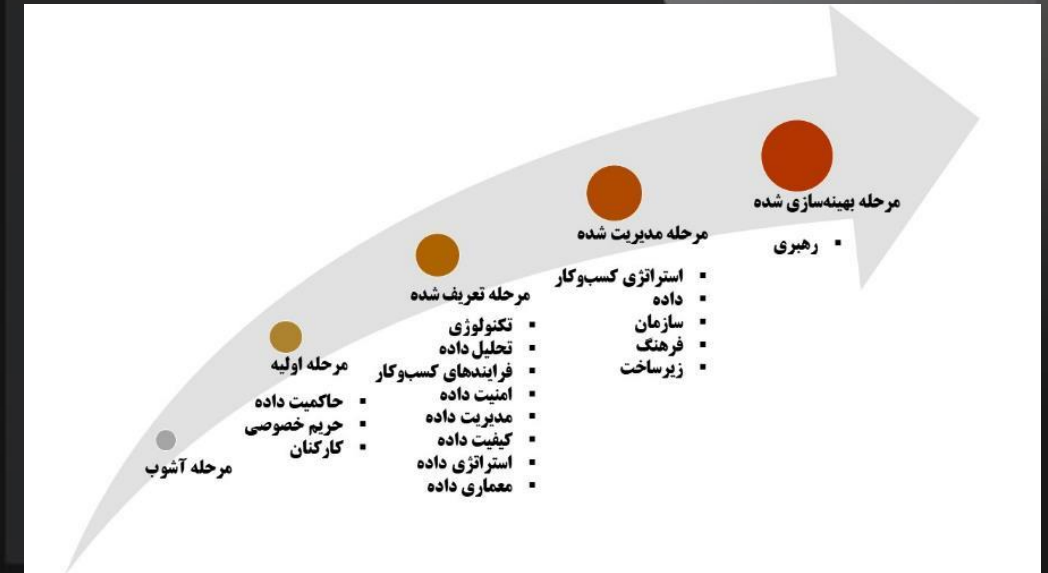
عدم وجود قابلیت تحلیل داده

نمونه ارزیابی بلوغ داده محوری

شناخت نقاط قوت، ضعف و گلوگاه‌های کسب و کار



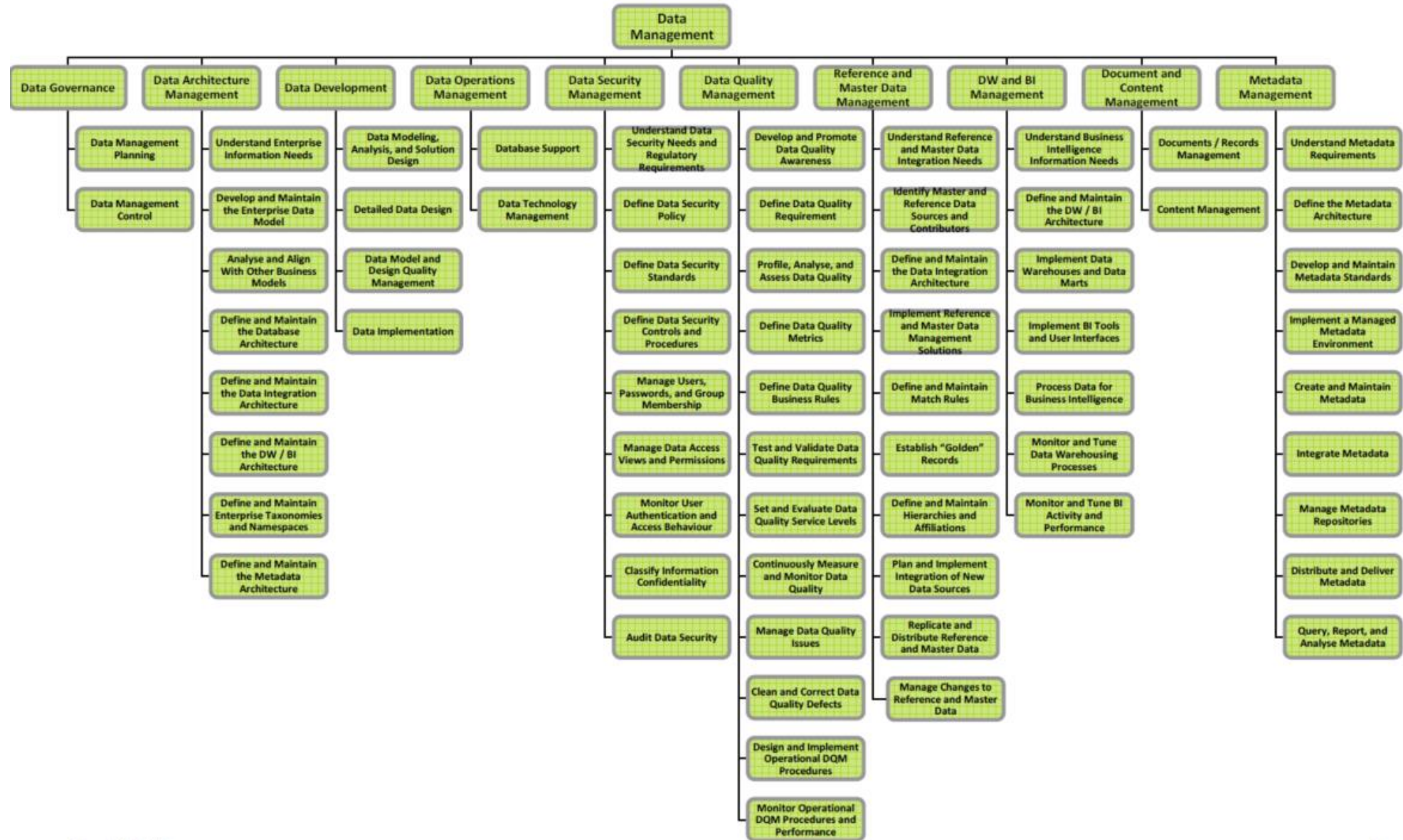
MacBook





جمع بندی

DMBOK Function and Activity Structure



ریسک‌ها!

سپردن سکان کار به
دست فناوری اطلاعات!

عدم شناخت مناسب
سطح بلوغ سازمانی

اجرای حاکمیت داده
به عنوان یک پروژه!

عدم همسویی
با استراتژی کسب و کار

عدم درک مناسب
چشم انداز داده

جلب رضایت صرفاً یک
گروه از ذینفعان
(رعایت قوانین و مقررات)

عدم بکارگیری
یک چارچوب مناسب

شروع کار را
بزرگ دیدن

ارایه راه حل های
ابزار محور

مفروضات کلیدی پروژه‌های حکمرانی داده

تسهیل در برگزاری جلسات موضوعی با مدیران و کارشناسان

تسهیل در برگزاری جلسات فنی

دسترسی به اسناد بالادستی

دسترسی به مستندات معماری داده سامانه‌ها و پایگاه داده‌ها

دسترسی مدیریتی و کاربری به سامانه‌ها و منابع داده‌ای جهت بررسی قابلیت‌ها و اجزای آن

دسترسی به مستندات مدیریت فرایند (BPM) سامانه‌ها و فرایندهای تعامل با ذی‌نفعان

THE

END