Helping Technology Help You Create Business Value

Felice Ho
MBA 656, October 2019
Long Island University

About me


J.P.Morgan


STATE STREET


Burgiss


eFront


EQUINOX

Overview: data journey in gaining business insights

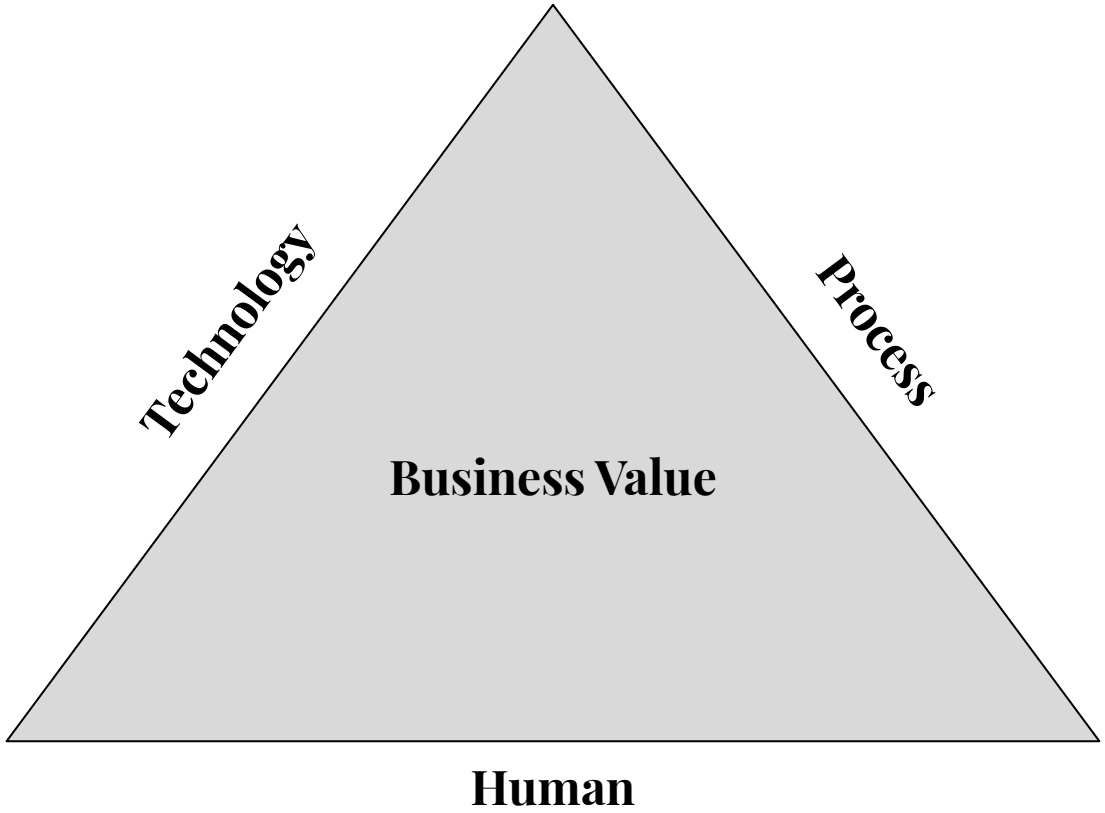
Data engineering

Business intelligence

Data warehouse

Data infrastructure/platform

Overview: helping tech help you create business value



Company

Luxury fitness and lifestyle

100+ fitness clubs

- US / UK / Canada

Personal training / Pilates

Group Fitness

Spa, Cycling, Hotel

Data Insights Team

Data engineering

- Develop data applications for data warehouse, data experience, data platform

Data experience

- Develop reports, data visualizations

Data insight

- Data analytics

Platform development

- Engineer data applications for reliability, deployment, operational support

Data science

- Build predictive modeling

Business systems

- Employee experience: provisioning, commissions, payroll

What is a data engineer?

Data workflows

- Build tools to help movement of data between systems

Data pipeline

- Send data from information systems to where it's needed

ETL processes

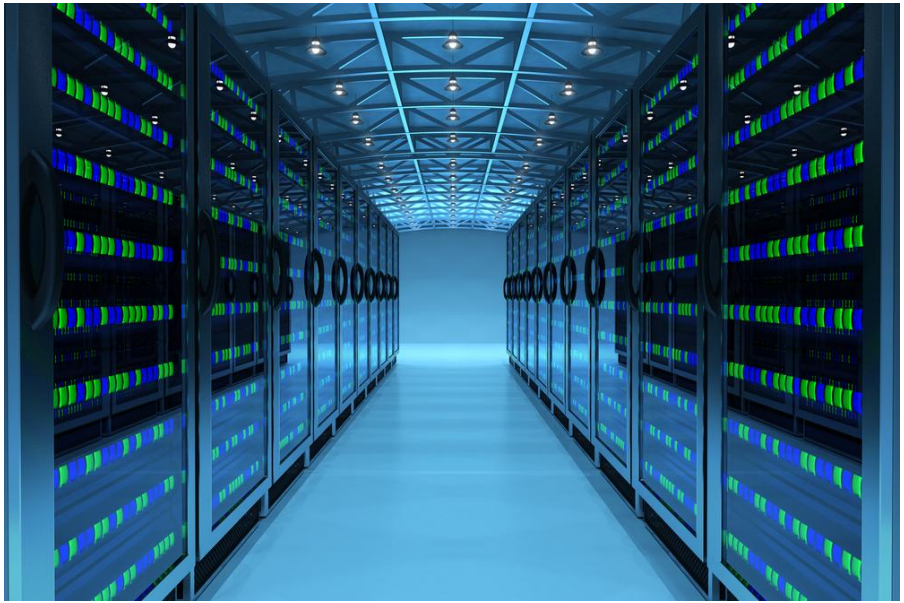
- **Extract** data, **transform** using business logic, **load** into data warehouse

Why is this important?

Support daily updates in Enterprise Data Warehouse (EDW)

- critical component of business intelligence (BI) and data analysis

BI drives business decisions and strategic goals



Daily Management Report (DMR)

DMR - Summary

Open	New Sales						
Club	3D	Daily	MTD	Proj	Bud	%	LY
19TH							
92ND							
85TH							
63RD							
54TH							
50TH							
43RD							
44TH							
WALL							
TRIB							
GRNW							
CCL							
SOHO							
PARK							

Number of new sales

Metrics for key business lines

- Daily, month-to-date
- Project vs. Budgeted

Starting backwards on a data journey

Business Intelligence

Reporting is crucial in any business

Make sense of data stored in information systems from business operations

Helps determine how a business is doing

- Make management decisions

SAP Business Objects

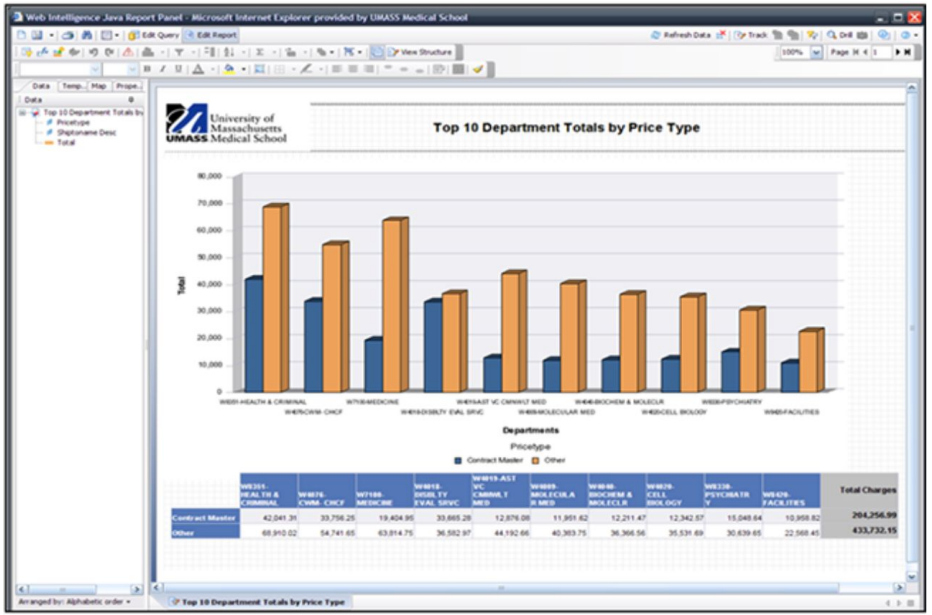
Forms of data analysis

Traditional reports

- Daily, weekly, monthly, quarterly, annual
- Snapshot, Inception to date

Ad-hoc reporting

- Marketing campaigns
- Daily decisions



salesforce exacttarget marketing cloud

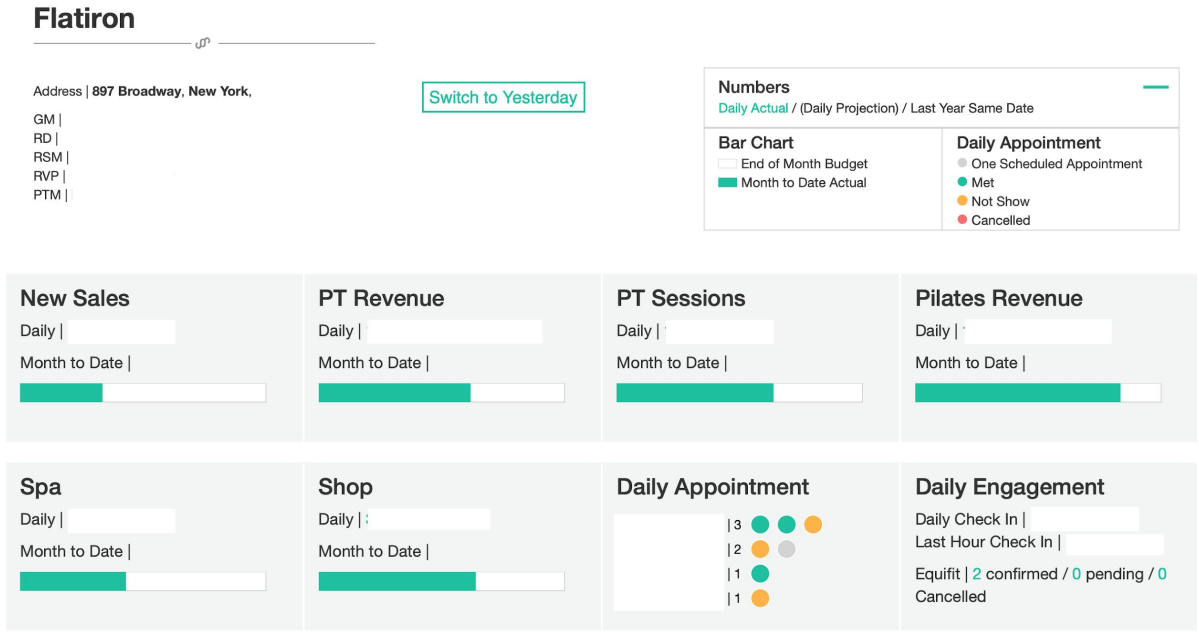
Forms of data analysis

Intraday

- Meet sales goals

Near real-time analysis

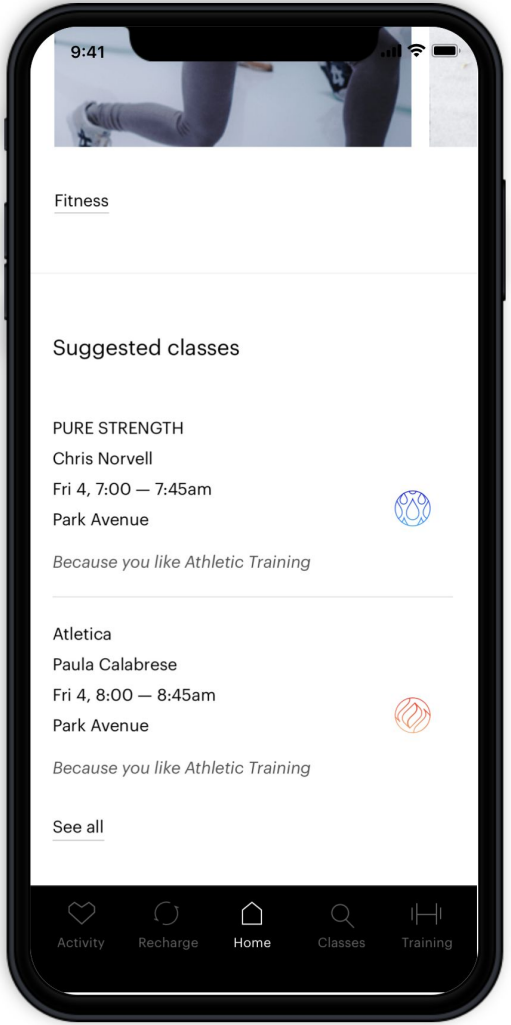
- Check-in alerts



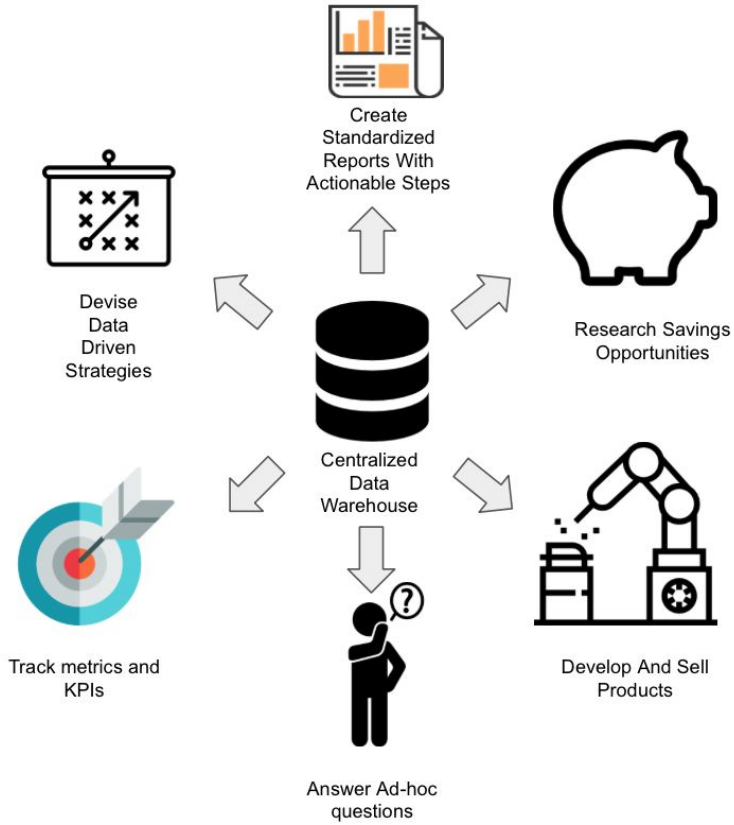
Forms of data analysis

Predictive analysis / data science

- Personalized class recommendations
- Forecasting daily sales
- Customer Lifetime Value
- Customer Segmentation



Data Warehouse



Makes reporting possible

Integrates data from information systems

Apply business logic to raw data into useful form

Materialized views for efficient access to common/key metrics, fields

Data Warehouse: data sources

Information systems

- Transactional / point of sale
- Website / mobile apps
- Customer Relationship Management (CRM)

Understand significance and data flow

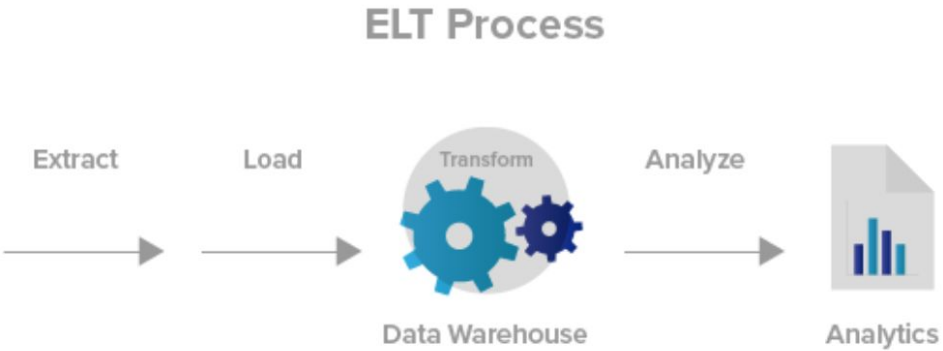
Data Warehouse: ELT (vs ETL)

Data **extracted** from source systems

Integrated and **loaded** into database tables

Business rules applied to **transformed** data

Summarized data available in data marts / views for data analysis



Data infrastructure

Backend requirements to allow for **storing**, **sharing**, **building** system data needs

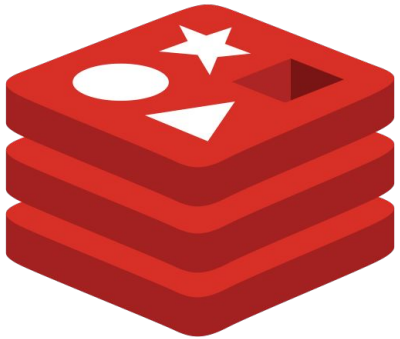
Data infrastructure: data storage

SQL Databases

relational

NoSQL Databases

non-relational



Redis

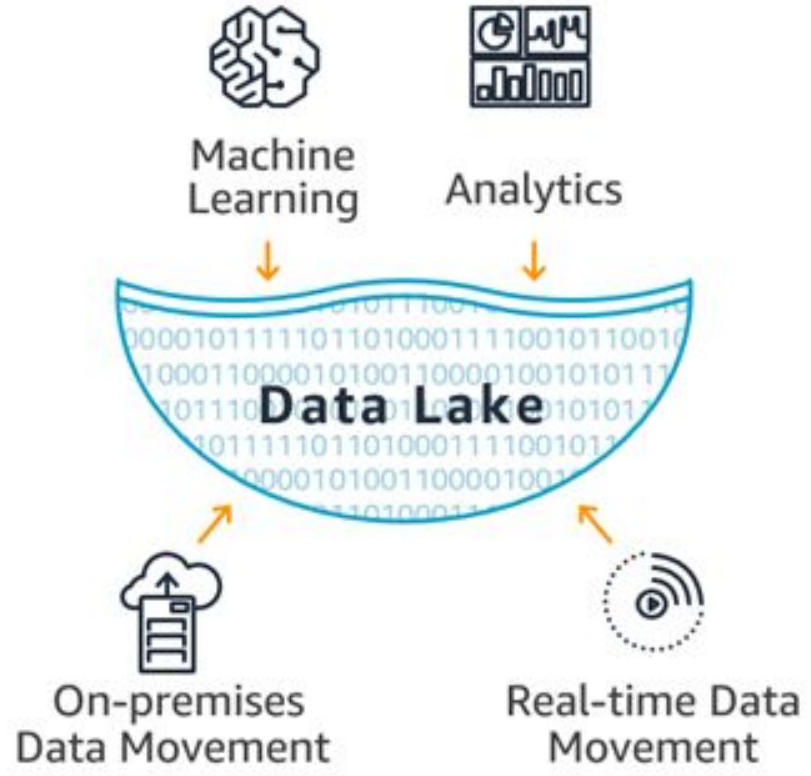
Data infrastructure: data storage

Files

- Logging, vendor data

Data lakes

- AWS S3



Data infrastructure: data sharing

Data warehouse

Application Program Interface (APIs)

Automated emails

Event driven alerts

Cached data

Data pipelines

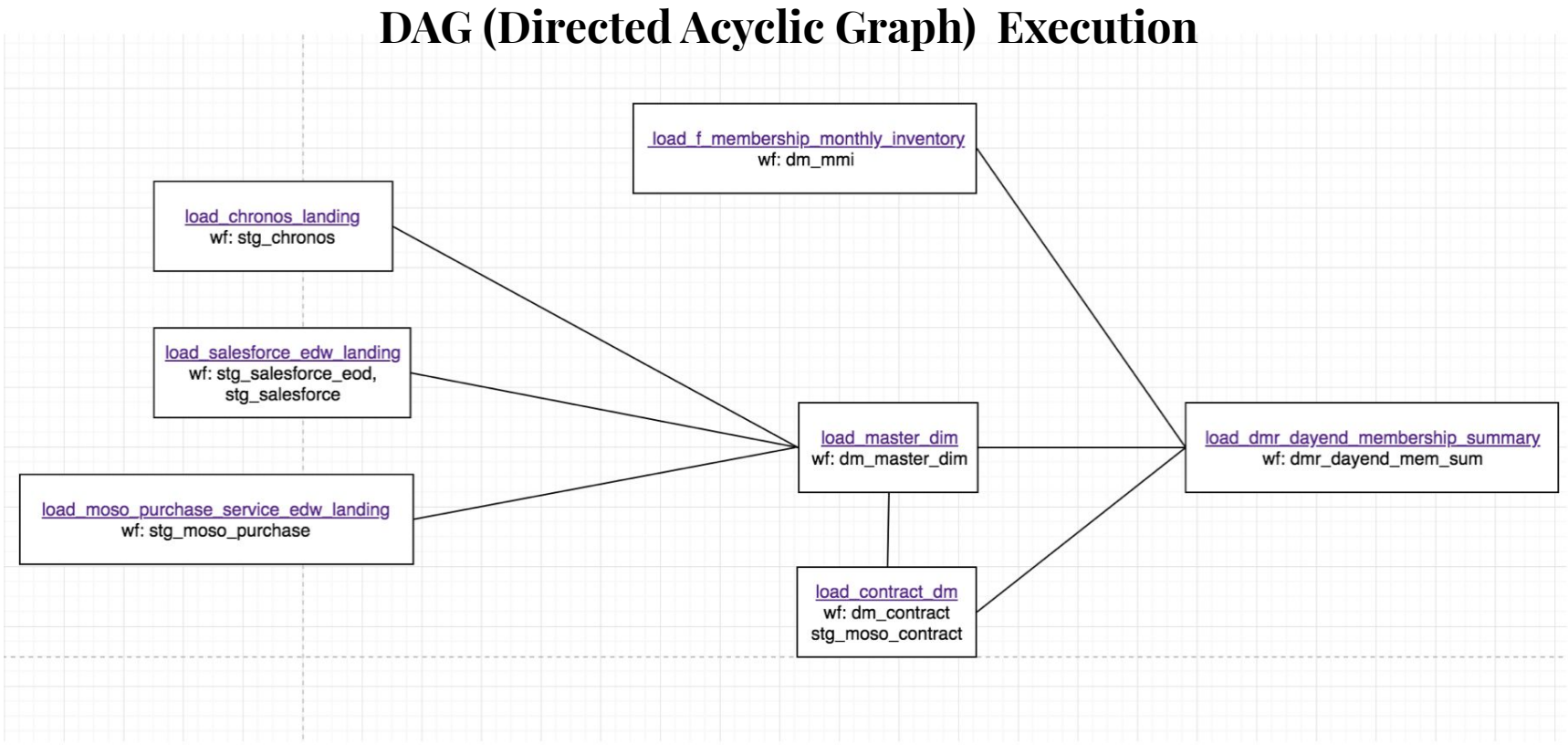
Data pipelines

In order to share the data

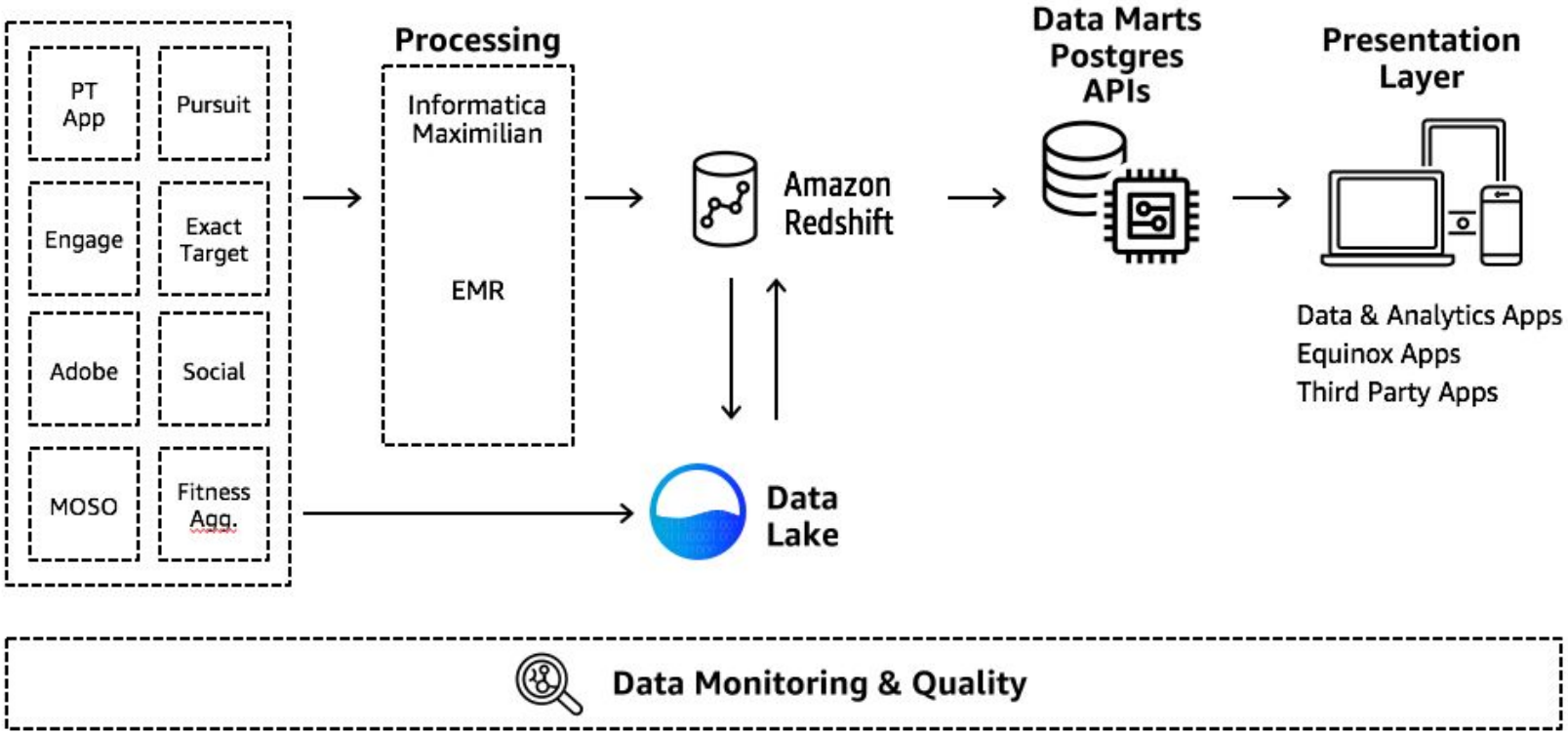
- need to get data from its sources and deliver to where it’s needed

What is required?

Data pipelines: dependency management



Data pipelines: data quality and monitoring checks



Data pipelines: operations monitoring

Data operations

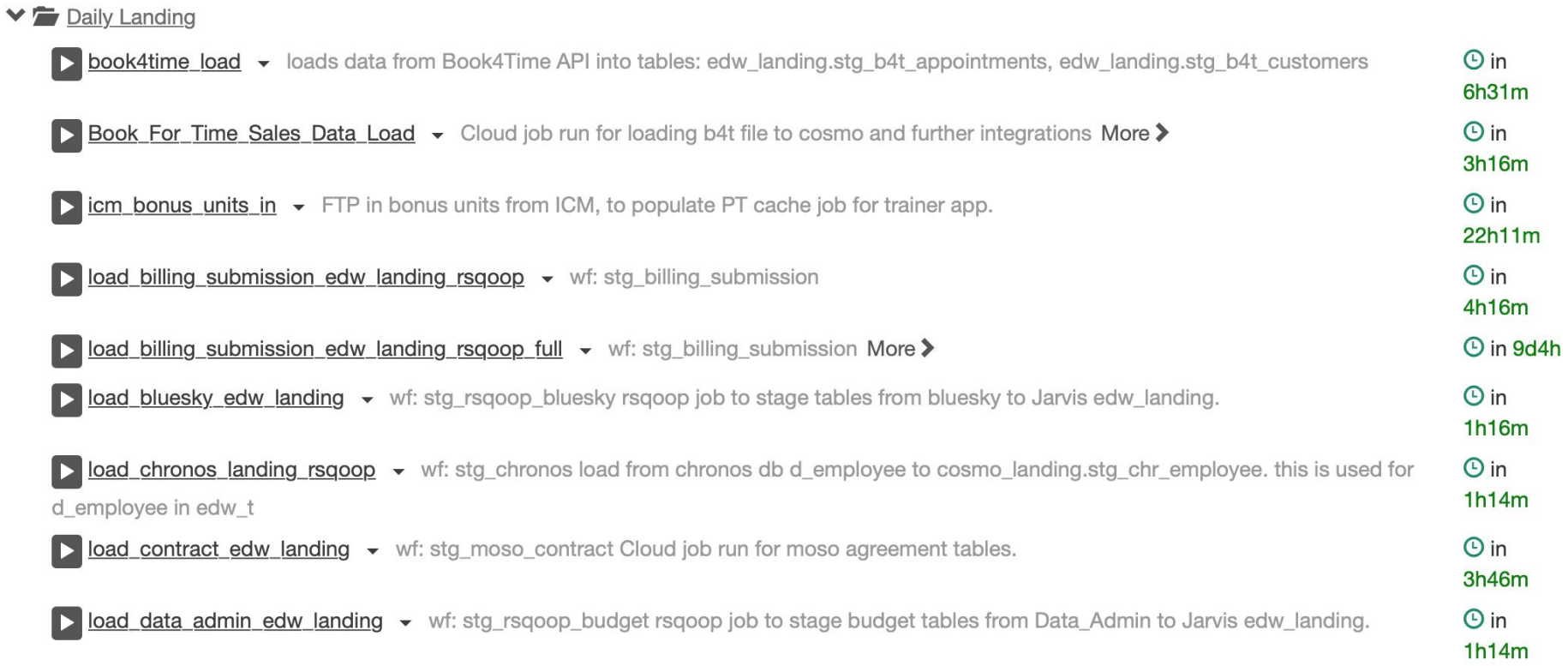
Handle when jobs fail

ELT process run throughout the night for DMR to get updated by next morning

Due to dependencies - need to resolve any failures or report will be late!

pagerduty

Data pipelines: scheduling tool



Data pipelines: deployment of job updates

Continue integration, continuous deployment

Changes to jobs reflected immediately

Automated and no manual updates needed

Jenkins

Data platform

In order to build data requirements

- need tools to support data pipelines and its processes

Ways to bring business insights to the stakeholders

- presentation layer (ie: reports, dashboards, emails, etc)

What is required?

Data platform

Tools to connect to data storage

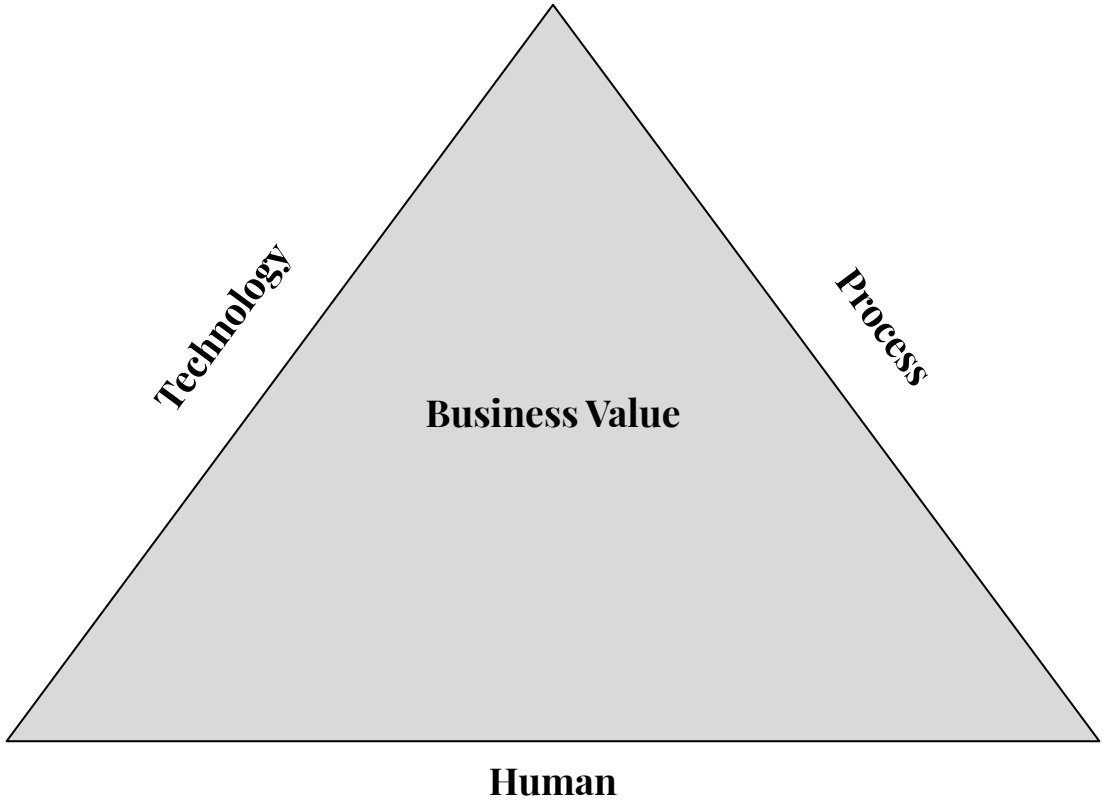
Data integration tools

- movement/sharing of data between systems

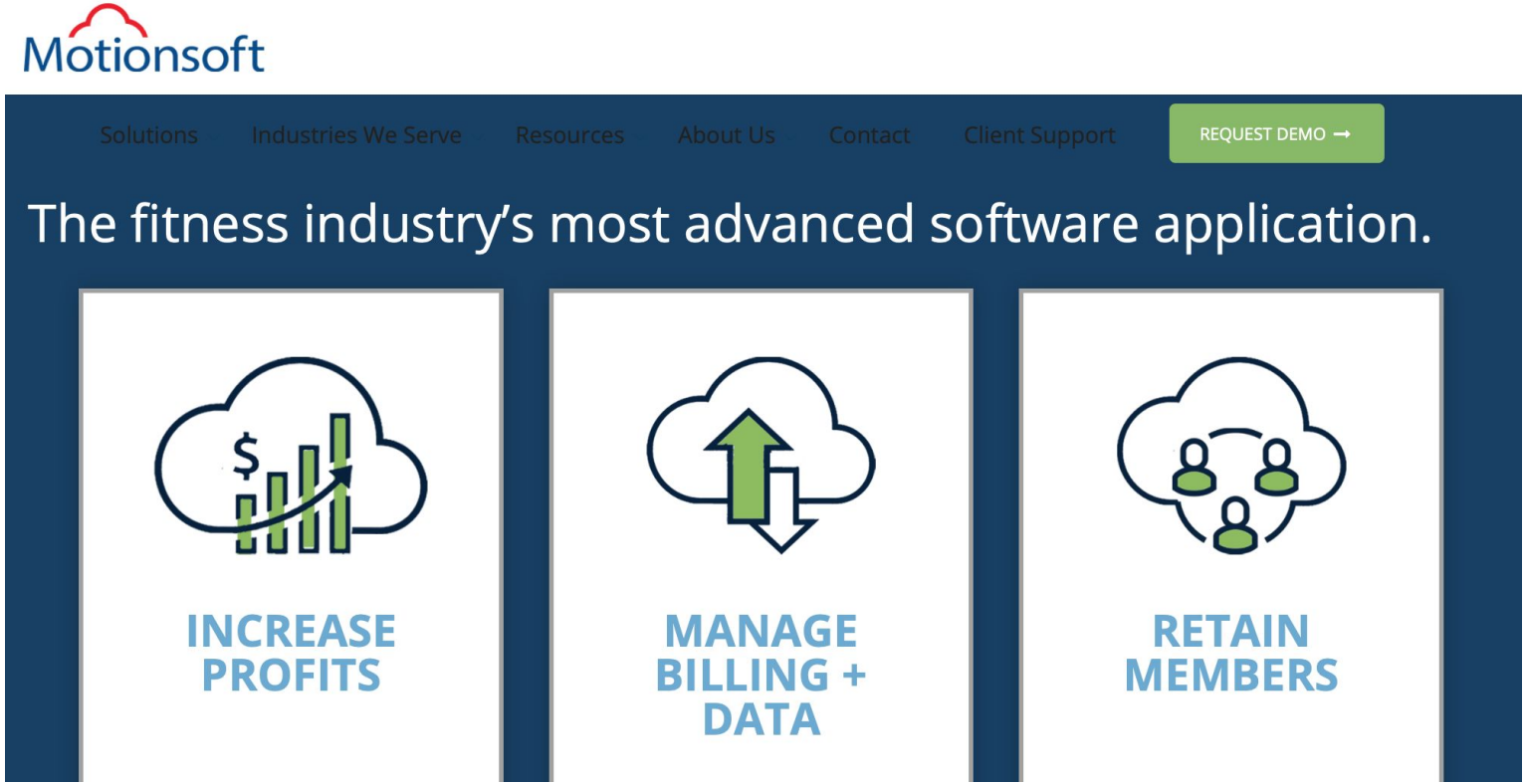
Framework for presenting business intelligence

- BI, data visualization, data delivery tools

Helping Technology Help You Create Business Value


Technology
Process
Business Value
Human

The tech side: modernizing our systems



The tech side: modernizing our data warehouse

Server moved to the cloud (internet)

- Scalable for big data processing/storage
- Disaster recovery, business continuity

Data lake strategy

- Store data in raw, unprocessed form for analysis later

“Data is the oil, some say the gold, of the 21st century”

- Joe Kaesar, Siemens CEO

Part of tech innovation is moving to the cloud

Hybrid Cloud

Chapter 2 of the cloud is about moving core business applications across multiple clouds. IBM is announcing new capabilities designed to simplify and speed the journey to hybrid cloud.

Businesses today have completed the first 20% of their cloud journey, which was Chapter 1. Chapter 2 is about moving mission-critical apps (the other 80%) to the cloud. But as companies do this they have to take into account "unique needs around compliance, security or location," Ginni said.

The result is a complex web of hybrid, multi-cloud environments. And, already, 94 percent of enterprises use multiple clouds. To pioneer their move to Chapter 2, businesses need the ability to move apps and data between clouds—efficiently and securely.

Source: https://newsroom.ibm.com/think-spotlight?item=30994

The tech side: benefits of cloud computing - reduced costs

Focus on what you're good at and outsource the rest

No need to manage on-premise infrastructure

No more expensive commercial licenses

Pay only for what you use - serverless computing

Upgrade/downgrade IT requirements as needed - scalability

You need to know what you want out of a system before you can build it

The process side: requirements analysis and design

Software Development Lifecycle (SDLC)

- Process to manage a project

Project scope

- what is expected and included in a phase of work

Requirements gathering

- understand important data points

The process side: planning and milestones communication

Waterfall methodology - sequential

- heavy, rigid design up front
- requirements well documented, solidified before development
- months / years of implementation
- key milestones defined
- longer feedback loop
- difficult to go back in design and make changes

The process side: planning and milestones communication

Agile methodology - iterative

- flexible design, changes easy to implement
- 2 - 3 week implementation
- smaller, definable pieces of completed work
- deliver sooner and get feedback early
- requirements can change over time

There are many actors in play during the implementation of any system

The human side: different roles on a team/project

Project manager

Data scientist

Report developer

Data engineer

Stakeholders

Data architect

Data analyst

QA Analyst

The human side: communication

Be adaptable to environment

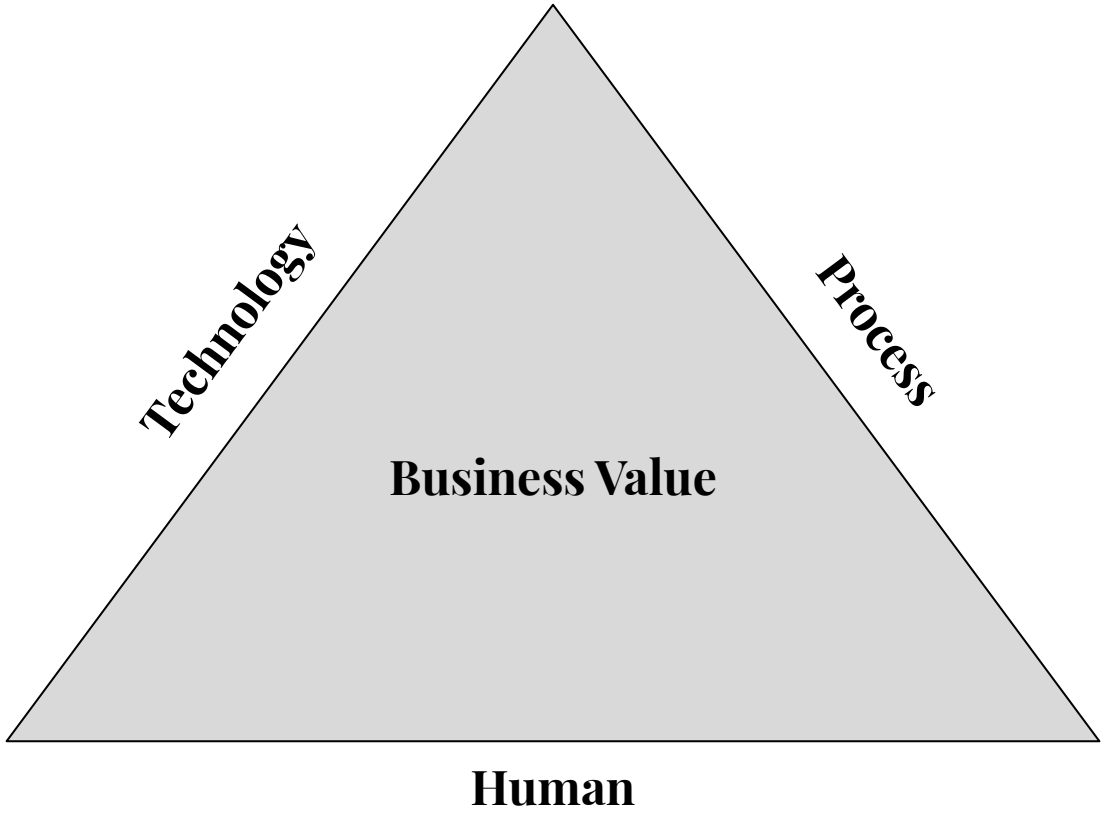
Communication, empathy, honesty important

Develop trust

Ask for frequent feedback

- there are different mediums to convey ideas

Successful tech innovation



Thank you

LinkedIn /feliceho