

Big Data Analysis

AI, Data Science, and Big Data Analysis

1112BDA02

MBA, IM, NTPU (M6031) (Spring 2023)

Tue 2, 3, 4 (9:10-12:00) (B8F40)



<https://meet.google.com/paj-zhji-mya>



Min-Yuh Day, Ph.D,
Associate Professor

Institute of Information Management, National Taipei University

<https://web.ntpu.edu.tw/~myday>



Syllabus

Week	Date	Subject/Topics
1	2023/02/21	Introduction to Big Data Analysis
2	2023/02/28	(Day Off)
3	2023/03/07	AI, Data Science, and Big Data Analysis
4	2023/03/14	Foundations of Big Data Analysis in Python
5	2023/03/21	Case Study on Big Data Analysis I
6	2023/03/28	Machine Learning: SAS Viya, Data Preparation and Algorithm Selection

Syllabus

Week Date Subject/Topics

7 2023/04/04 (Children's Day) (Day off)

8 2023/04/11 Midterm Project Report

9 2023/04/18 Machine Learning: Decision Trees and Ensembles of Trees

**10 2023/04/25 Machine Learning: Neural Networks (NN) and
Support Vector Machines (SVM)**

11 2023/05/02 Case Study on Big Data Analysis II

12 2023/05/09 Machine Learning: Model Assessment and Deployment

Syllabus

Week	Date	Subject/Topics
13	2023/05/16	ChatGPT and Large Language Models (LLM) for Big Data Analysis
14	2023/05/23	Deep Learning for Finance Big Data Analysis
15	2023/05/30	Final Project Report I
16	2023/06/06	Final Project Report II
17	2023/06/13	Self-learning
18	2023/06/20	Self-learning

AI

Data Science

Big Data Analysis

FinTech ABCD

AI

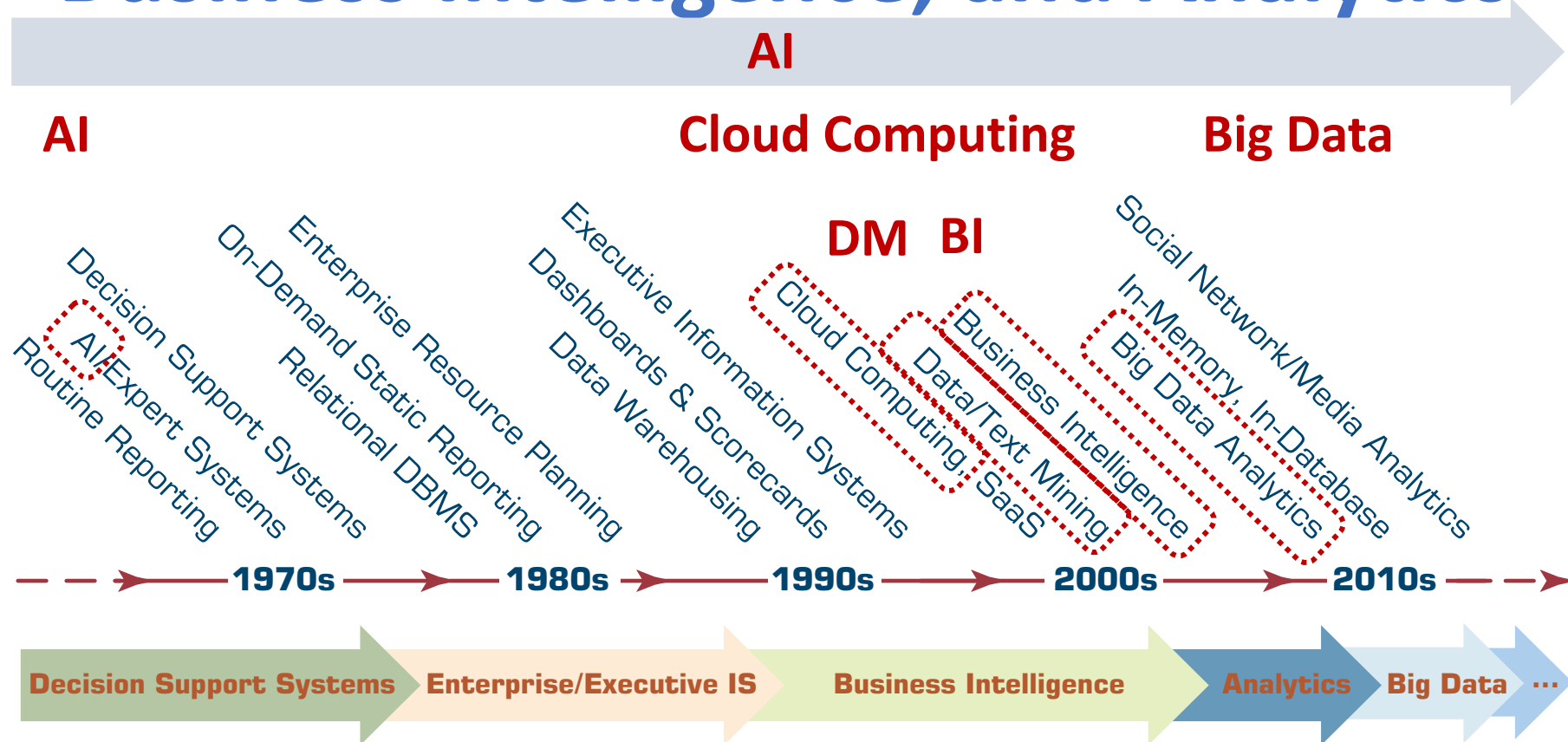
Block Chain

Cloud Computing

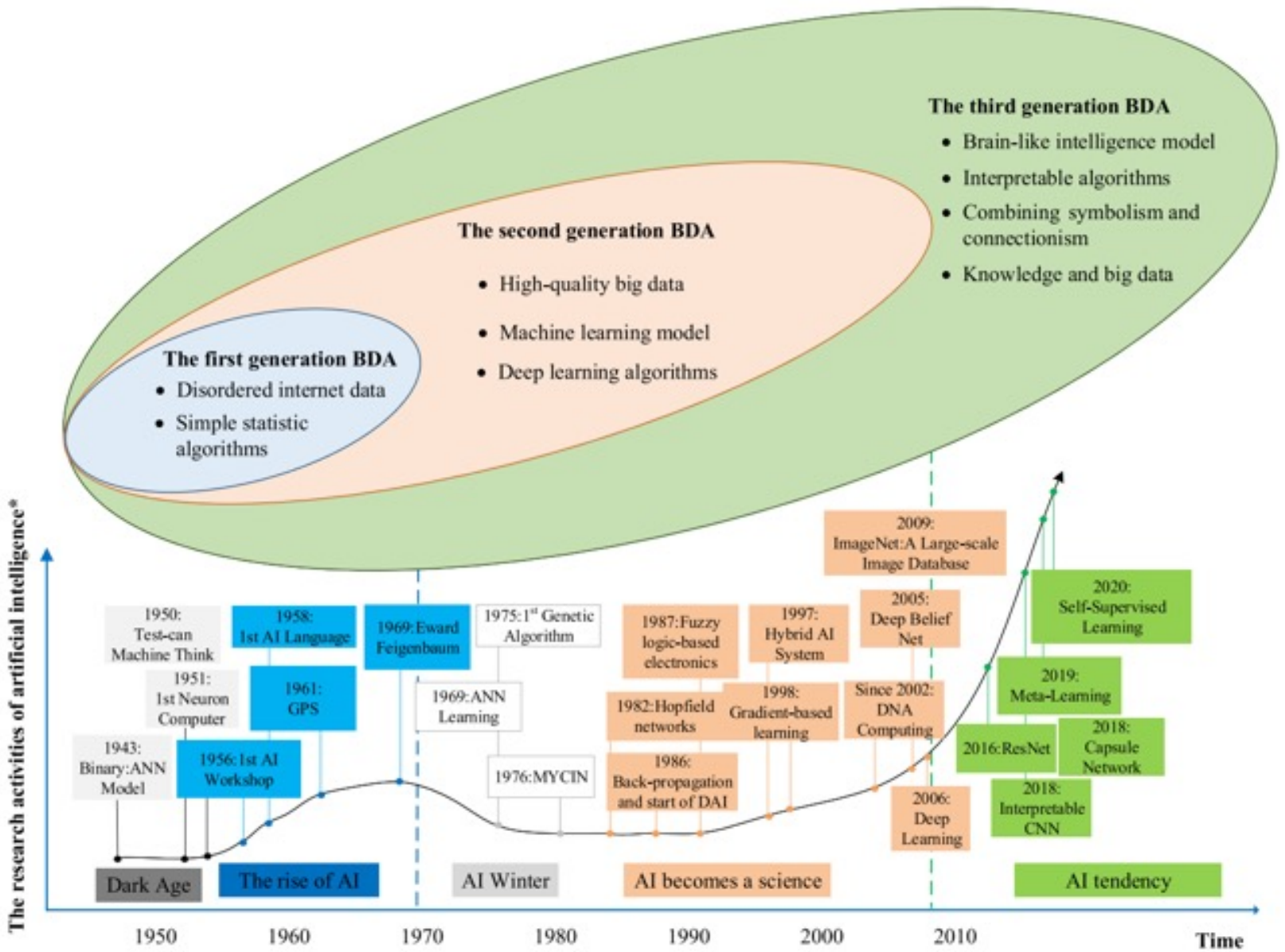
Big **D**ata

AI, Big Data, Cloud Computing

Evolution of Decision Support, Business Intelligence, and Analytics

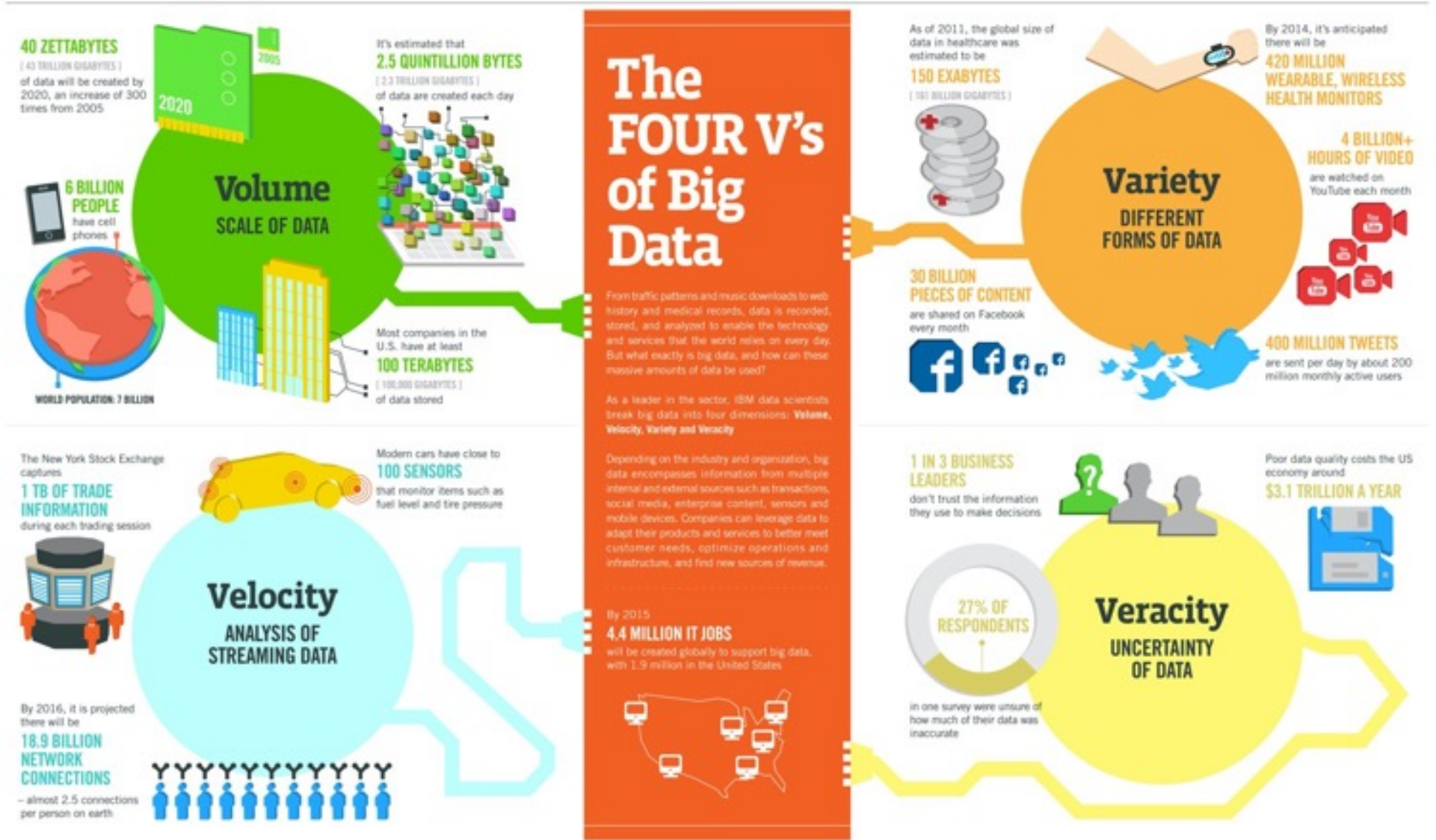


The Development of Big Data Analytics



Source: Wang, Junliang, Chuqiao Xu, Jie Zhang, and Ray Zhong (2022). "Big data analytics for intelligent manufacturing systems: A review." Journal of Manufacturing Systems 62 (2022): 738-752.

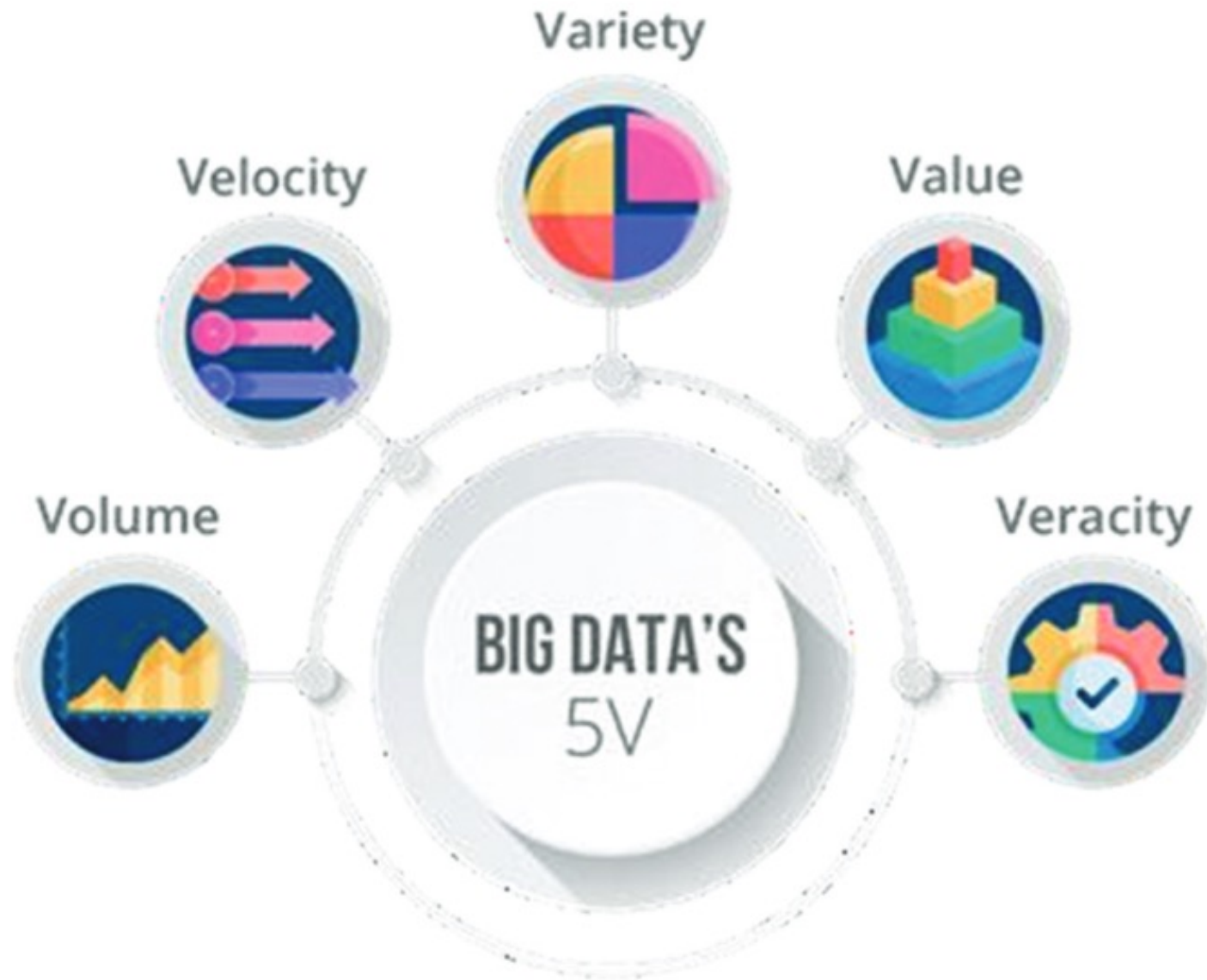
Big Data 4 V



Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, QAS



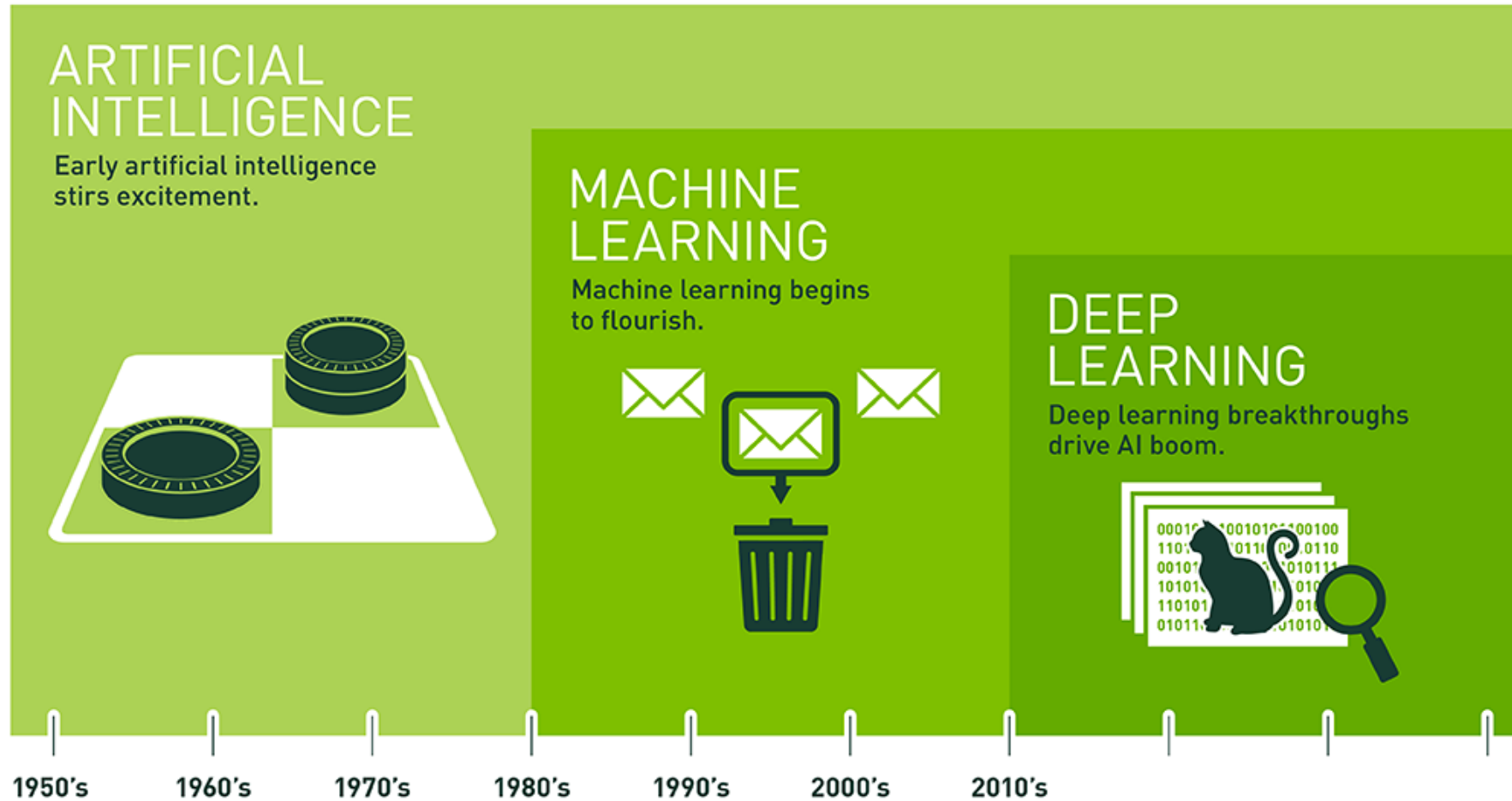
Big Data 5 V



Value

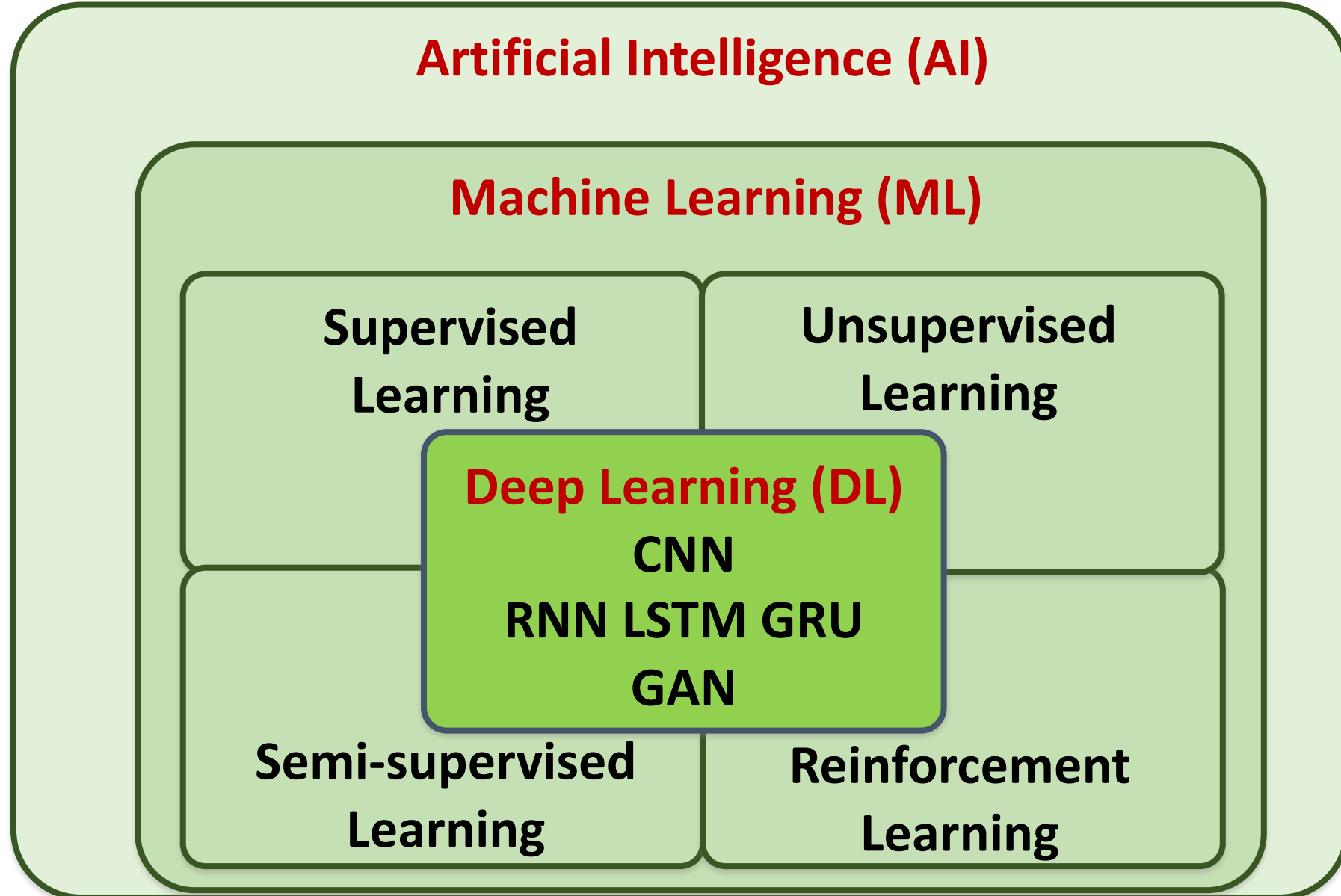
Artificial Intelligence

Machine Learning & Deep Learning



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

AI, ML, DL



Artificial Intelligence (AI)

Definition of Artificial Intelligence (A.I.)

Artificial Intelligence

**“... the science and
engineering
of
making
intelligent machines”**

(John McCarthy, 1955)

Artificial Intelligence

**“... technology that
thinks and acts
like humans”**

Artificial Intelligence

**“... intelligence
exhibited by machines
or software”**

4 Approaches of AI

Thinking Humanly	Thinking Rationally
Acting Humanly	Acting Rationally

4 Approaches of AI

<p>2. Thinking Humanly: The Cognitive Modeling Approach</p>	<p>3. Thinking Rationally: The “Laws of Thought” Approach</p>
<p>1. Acting Humanly: The Turing Test Approach (1950)</p>	<p>4. Acting Rationally: The Rational Agent Approach</p>

AI Acting Humanly: The Turing Test Approach

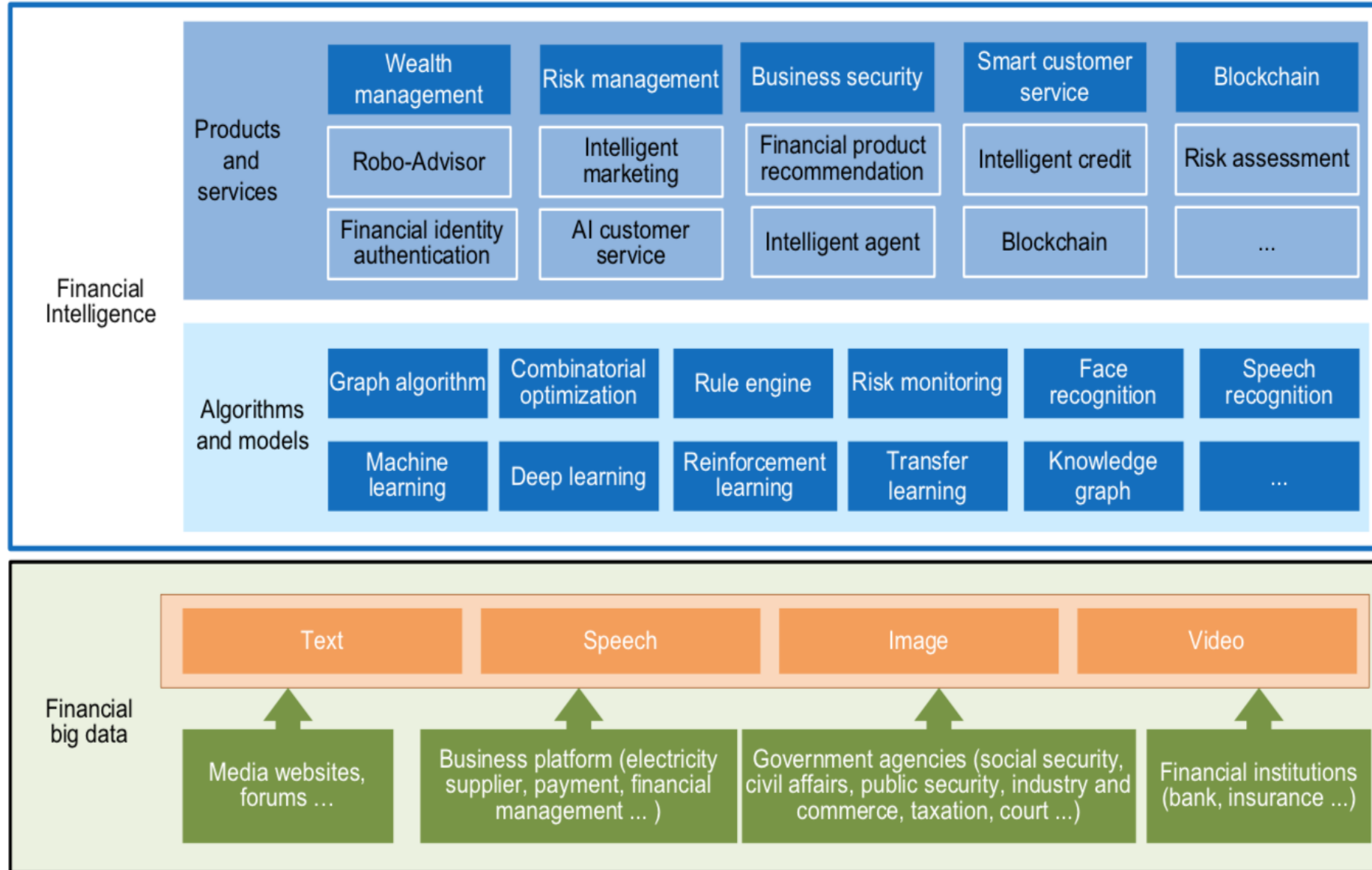
(Alan Turing, 1950)

- Knowledge Representation
- Automated Reasoning
- Machine Learning (ML)
 - Deep Learning (DL)
- Computer Vision (Image, Video)
- Natural Language Processing (NLP)
- Robotics

AI
in
FinTech

FinBrain: when Finance meets AI 2.0

(Zheng et al., 2019)



Source: Xiao-lin Zheng, Meng-ying Zhu, Qi-bing Li, Chao-chao Chen, and Yan-chao Tan (2019), "Finbrain: When finance meets AI 2.0." Frontiers of Information Technology & Electronic Engineering 20, no. 7, pp. 914-924

AI 2.0

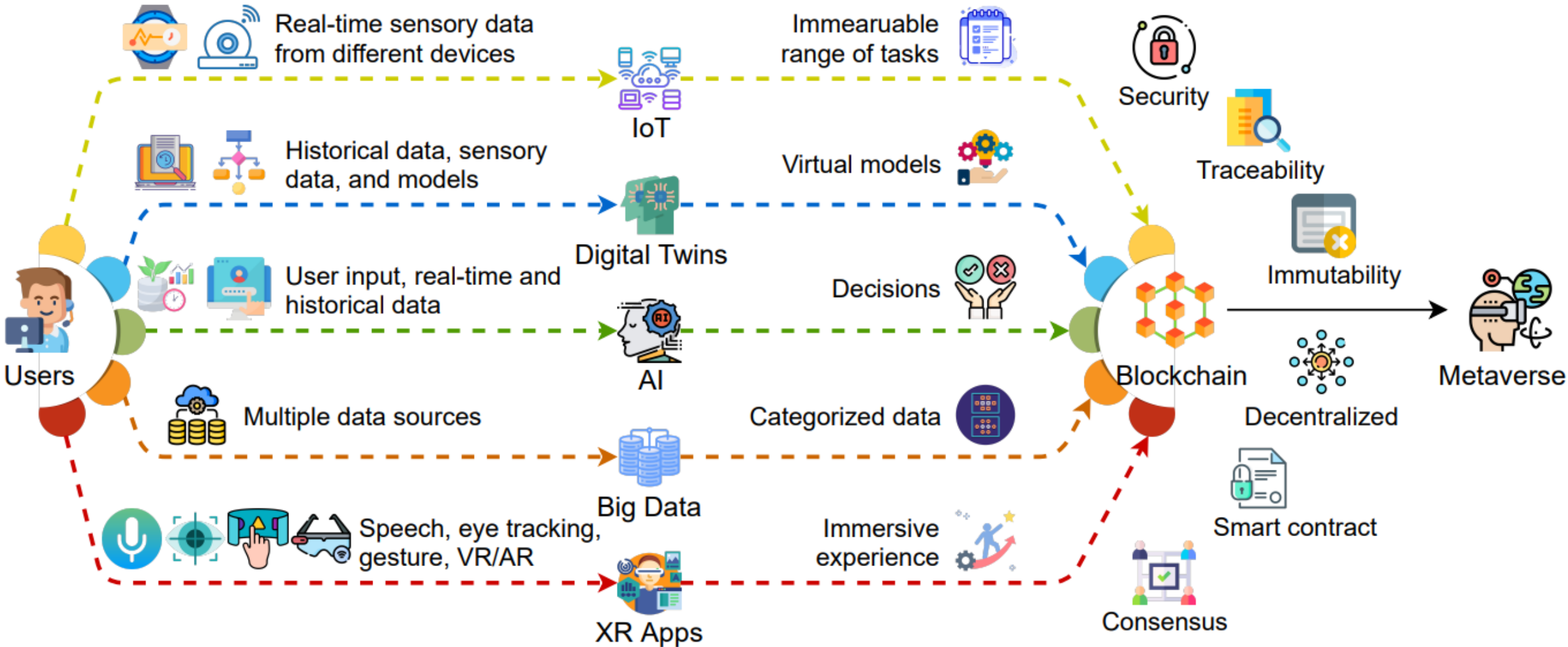
**a new generation of AI
based on the
novel information environment of
major changes and
the development of
new goals.**

Technology-driven Financial Industry Development

Development stage	Driving technology	Main landscape	Inclusive finance	Relationship between technology and finance
Fintech 1.0 (financial IT)	Computer	Credit card, ATM, and CRMS	Low	Technology as a tool
Fintech 2.0 (Internet finance)	Mobile Internet	Marketplace lending, third-party payment, crowdfunding, and Internet insurance	Medium	Technology- driven change
Fintech 3.0 (financial intelligence)	AI, Big Data, Cloud Computing, Blockchain	Intelligent finance	High	Deep fusion

AI and Blockchain

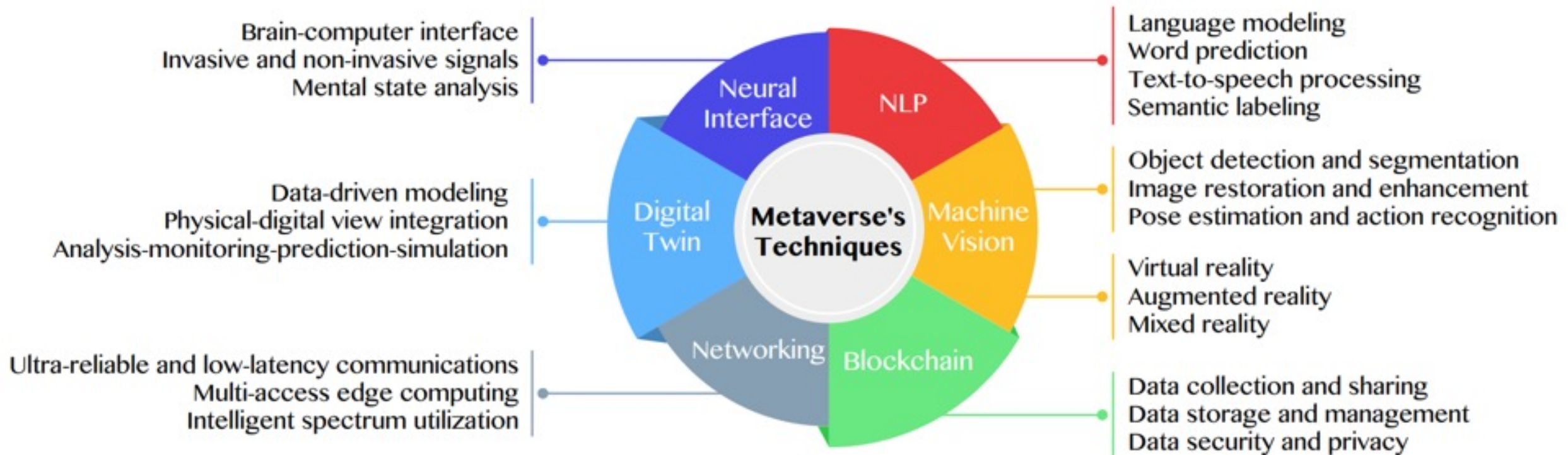
Key Enabling Technologies of the Metaverse



Source: Gadekallu, Thippa Reddy, Thien Huynh-The, Weizheng Wang, Gokul Yenduri, Pasika Ranaweera, Quoc-Viet Pham, Daniel Benevides da Costa, and Madhusanka Liyanage (2022). "Blockchain for the Metaverse: A Review." arXiv preprint arXiv:2203.09738..

Primary Technical Aspects in the Metaverse

AI with ML algorithms and DL architectures is advancing the user experience in the virtual world

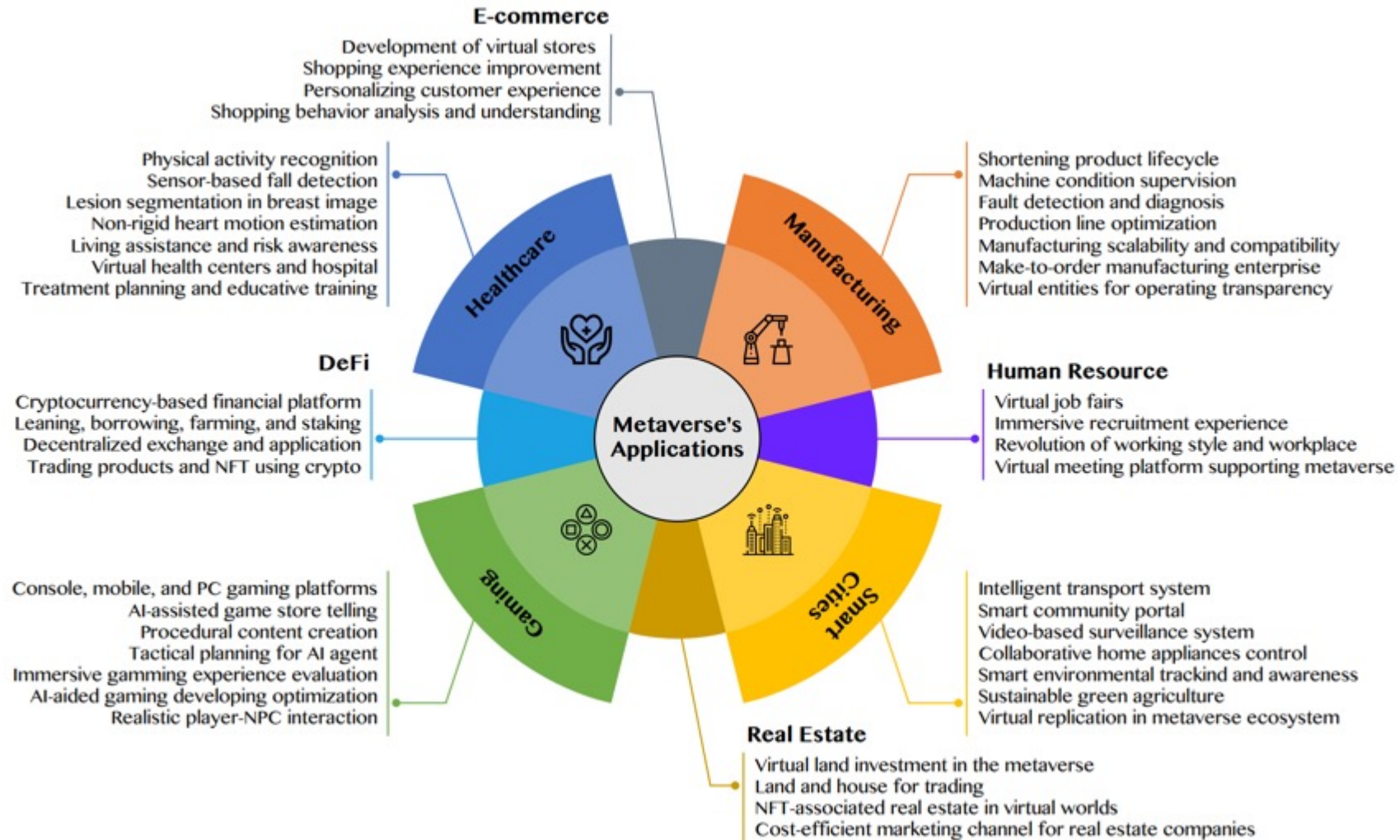


Source: Huynh-The, Thien, Quoc-Viet Pham, Xuan-Quy Pham, Thanh Thi Nguyen, Zhu Han, and Dong-Seong Kim (2022).

"Artificial Intelligence for the Metaverse: A Survey." arXiv preprint arXiv:2202.10336.

AI for the Metaverse in the Application Aspects

healthcare, manufacturing, smart cities, gaming
E-commerce, human resources, real estate, and DeFi

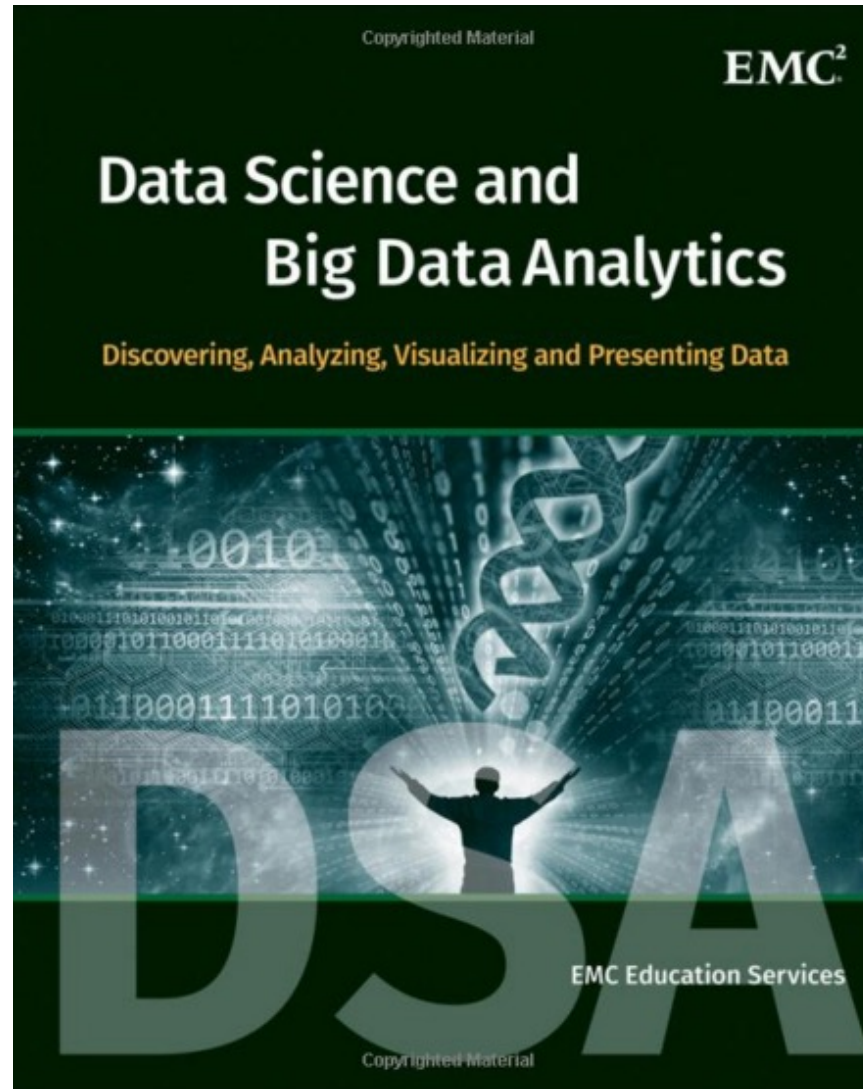


Source: Huynh-The, Thien, Quoc-Viet Pham, Xuan-Quy Pham, Thanh Thi Nguyen, Zhu Han, and Dong-Seong Kim (2022).

"Artificial Intelligence for the Metaverse: A Survey." arXiv preprint arXiv:2202.10336.

Data Science

EMC Education Services,
Data Science and Big Data Analytics:
Discovering, Analyzing, Visualizing and Presenting Data,
Wiley, 2015



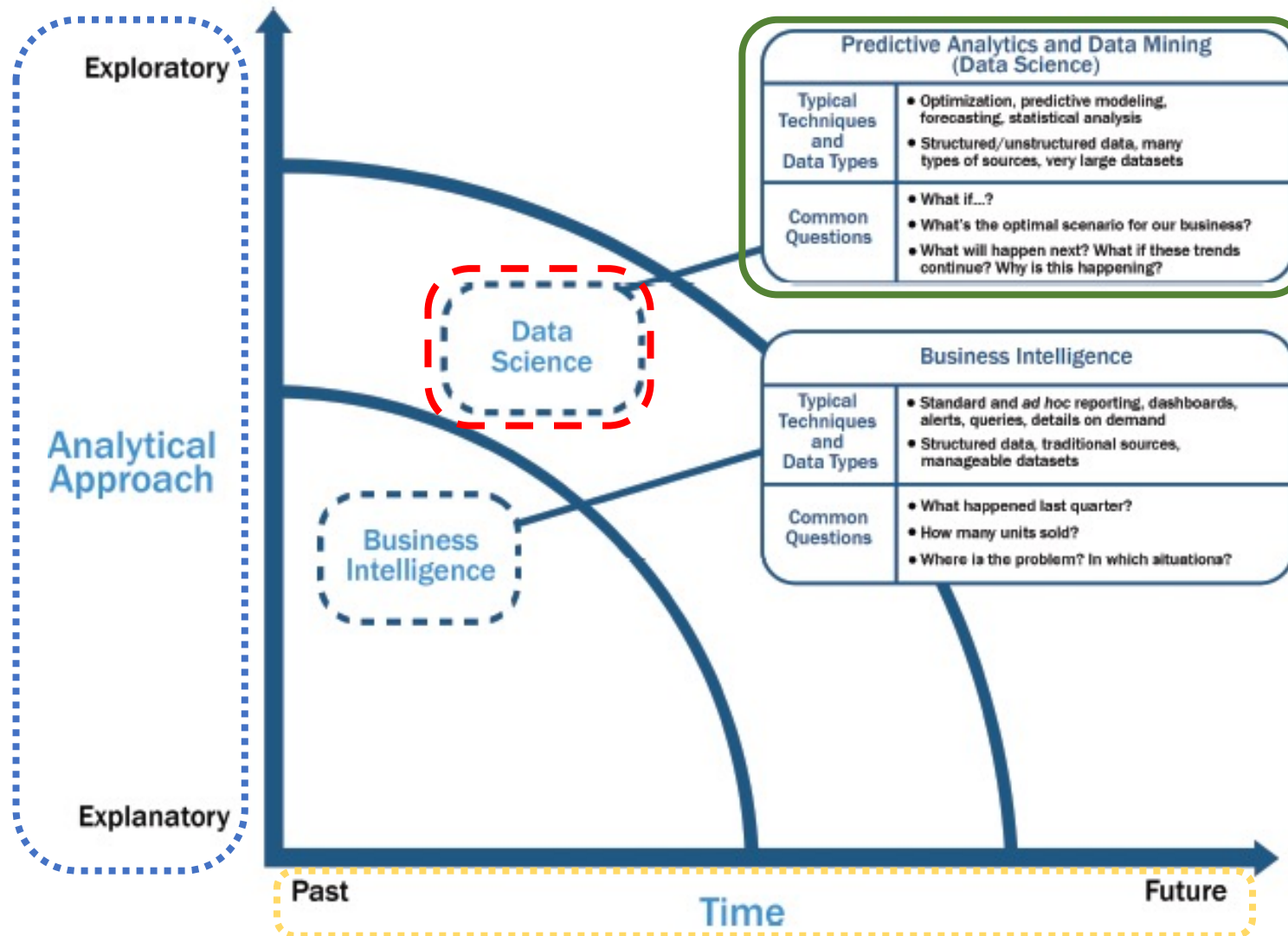
Data Analyst

- Data analyst is just another term for professionals who were doing **BI** in the form of **data compilation, cleaning, reporting**, and perhaps some **visualization**.
- Their skill sets included Excel, some SQL knowledge, and reporting.
- You would recognize those capabilities as **descriptive or reporting analytics**.

Data Scientist

- Data scientist is responsible for **predictive analysis, statistical analysis,** and more **advanced analytical tools and algorithms.**
- They may have a deeper knowledge of algorithms and may recognize them under various labels—**data mining, knowledge discovery, or machine learning.**
- Some of these professionals may also need deeper programming knowledge to be able to write code for data cleaning/analysis in current Web-oriented languages such as Java or Python and statistical languages such as R.
- Many analytics professionals also need to build significant expertise in **statistical modeling, experimentation, and analysis.**

Data Science and Business Intelligence



Data Science and Business Intelligence



Predictive Analytics and Data Mining (Data Science)

Past

Time

Future

Predictive Analytics and Data Mining (Data Science)

Structured/unstructured data, many types of sources,
very large datasets

Optimization, predictive modeling, forecasting statistical analysis

What if...?

What's the optimal scenario for our business?

What will happen next?

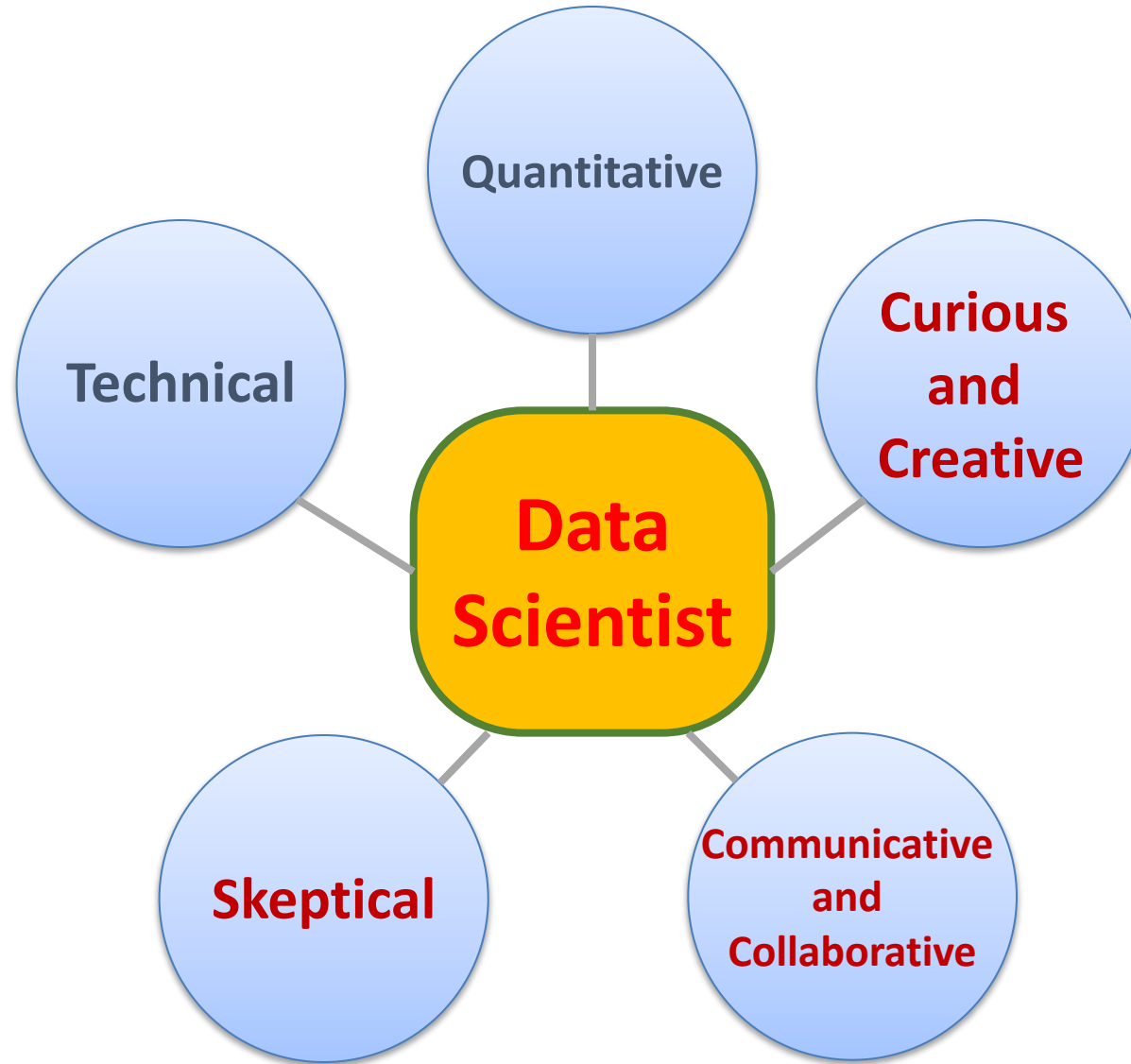
What if these trends continue?

Why is this happening?

Profile of a Data Scientist

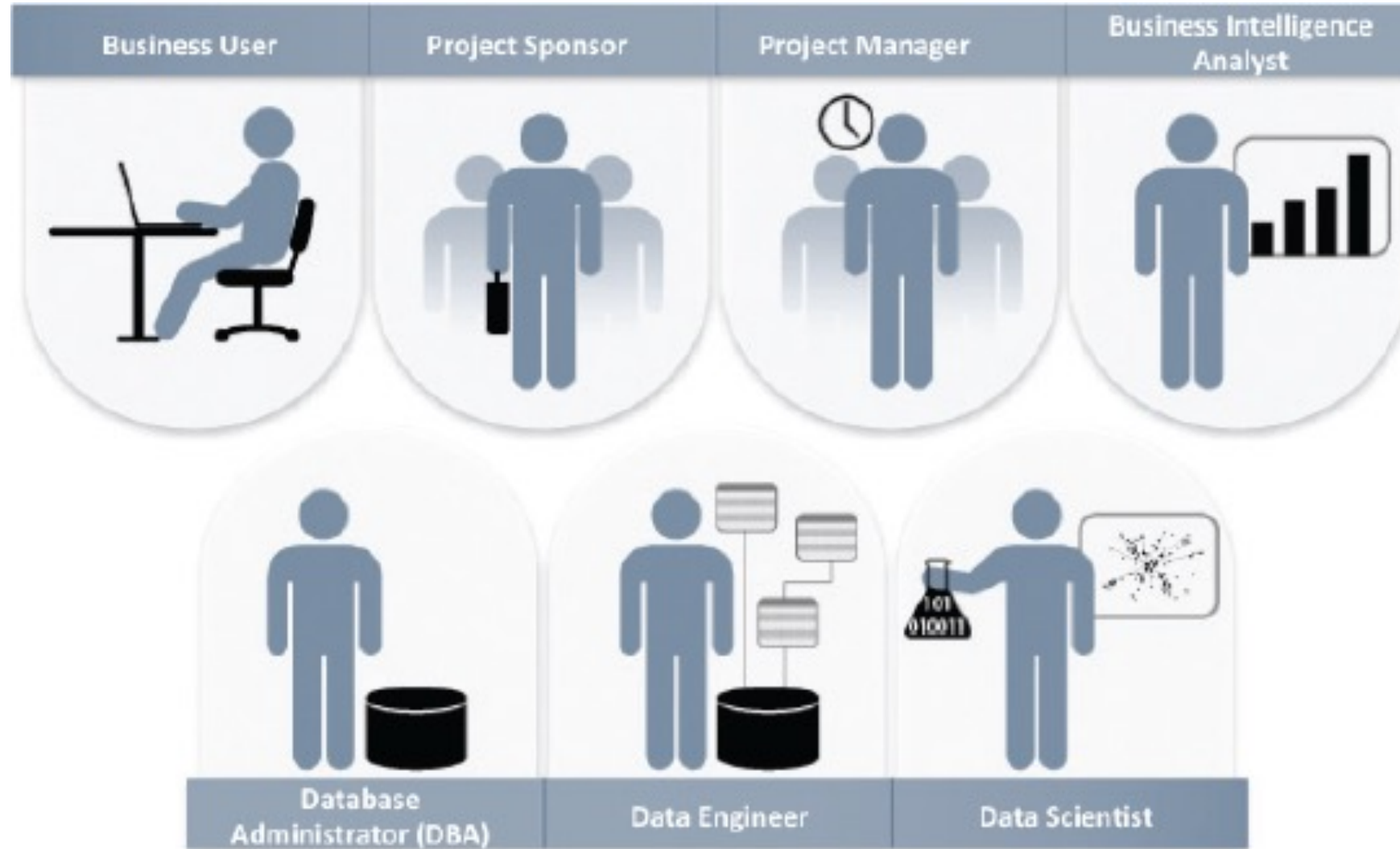
- **Quantitative**
 - **mathematics or statistics**
- **Technical**
 - **software engineering, machine learning, and programming skills**
- **Skeptical mind-set and critical thinking**
- **Curious and creative**
- **Communicative and collaborative**

Data Scientist Profile

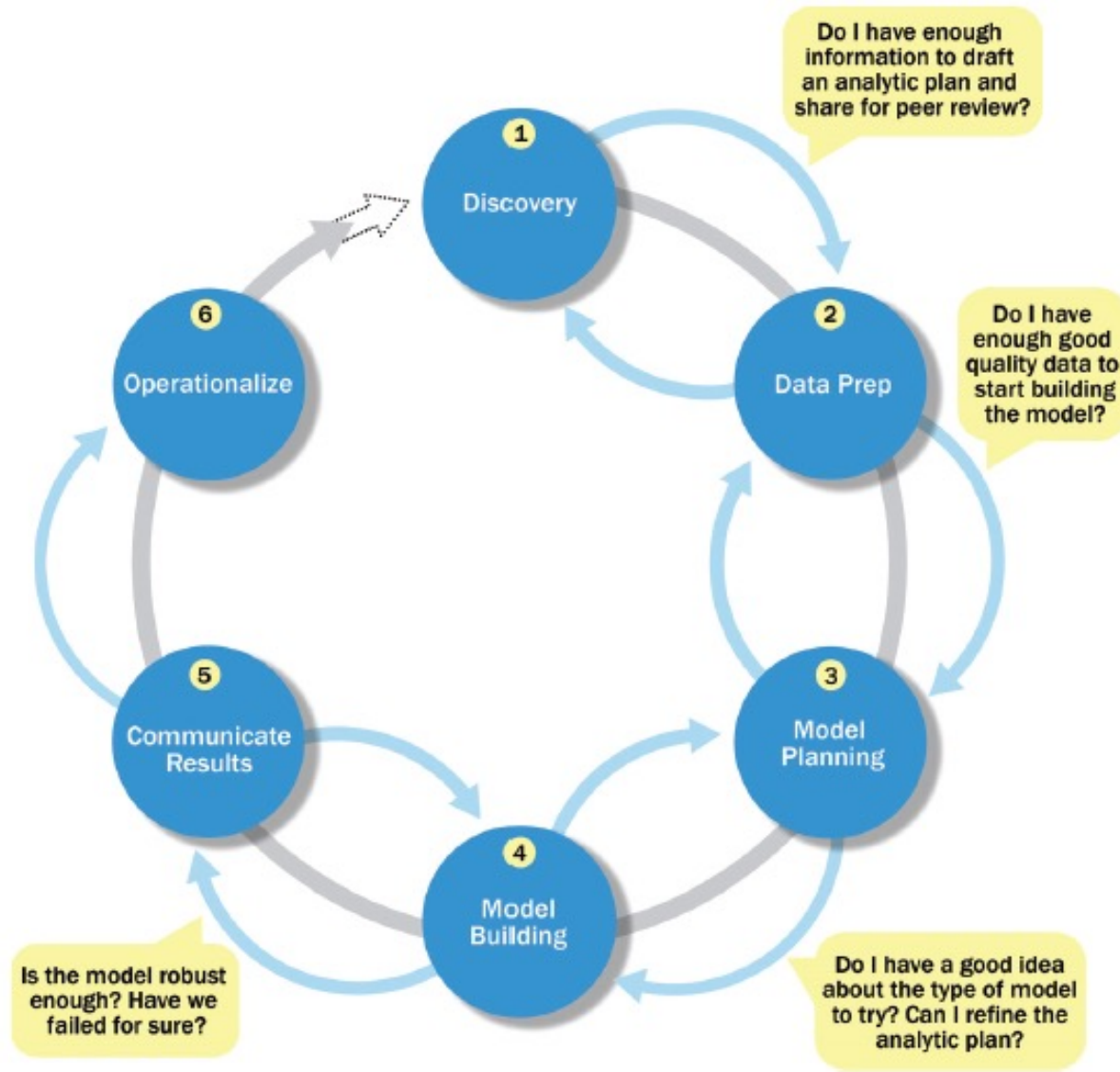


Big Data Analytics Lifecycle

Key Roles for a Successful Analytics Project



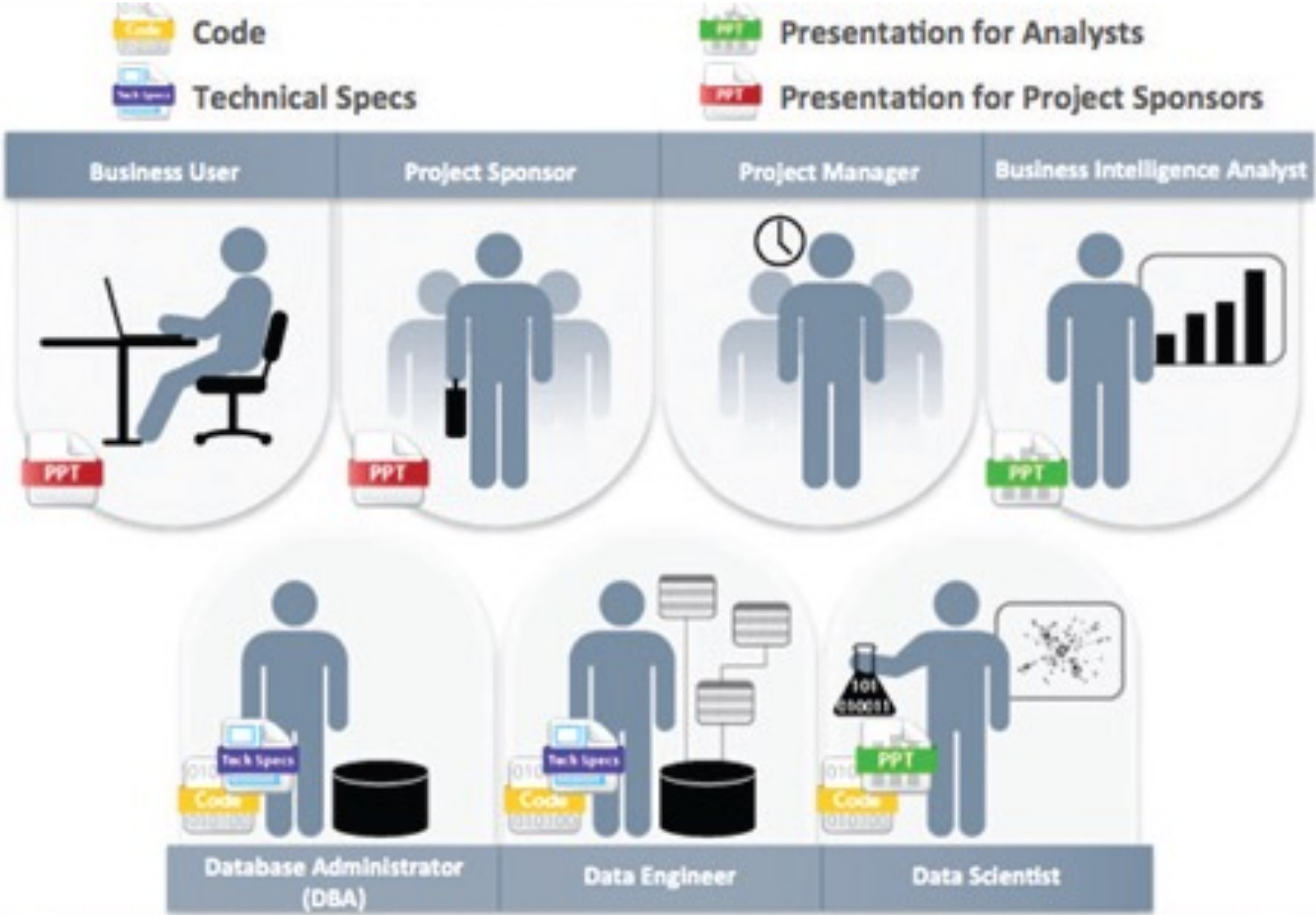
Overview of Data Analytics Lifecycle



Overview of Data Analytics Lifecycle

- 1. Discovery**
- 2. Data preparation**
- 3. Model planning**
- 4. Model building**
- 5. Communicate results**
- 6. Operationalize**

Key Outputs from a Successful Analytics Project

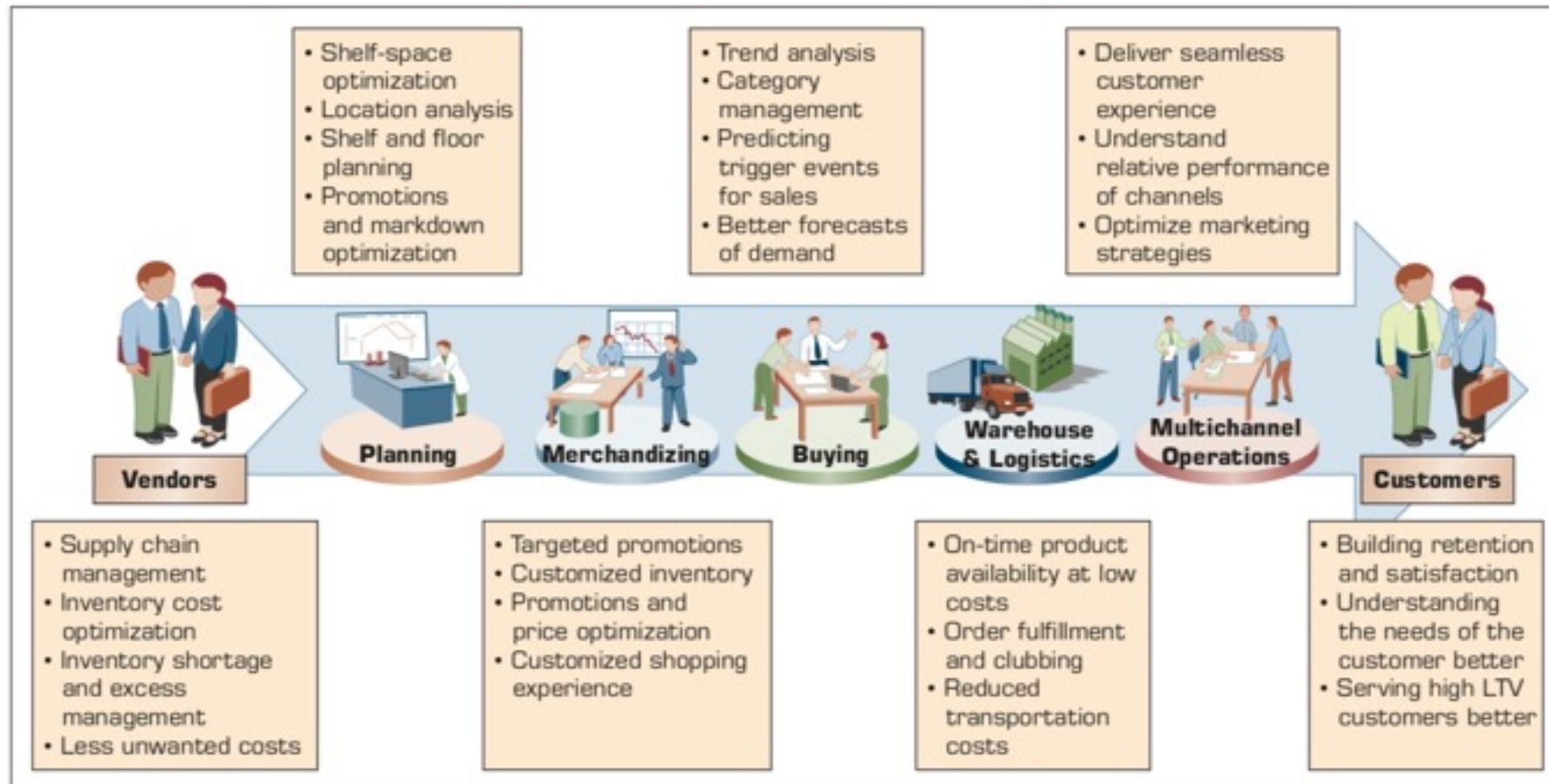


Source: EMC Education Services, Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, Wiley, 2015

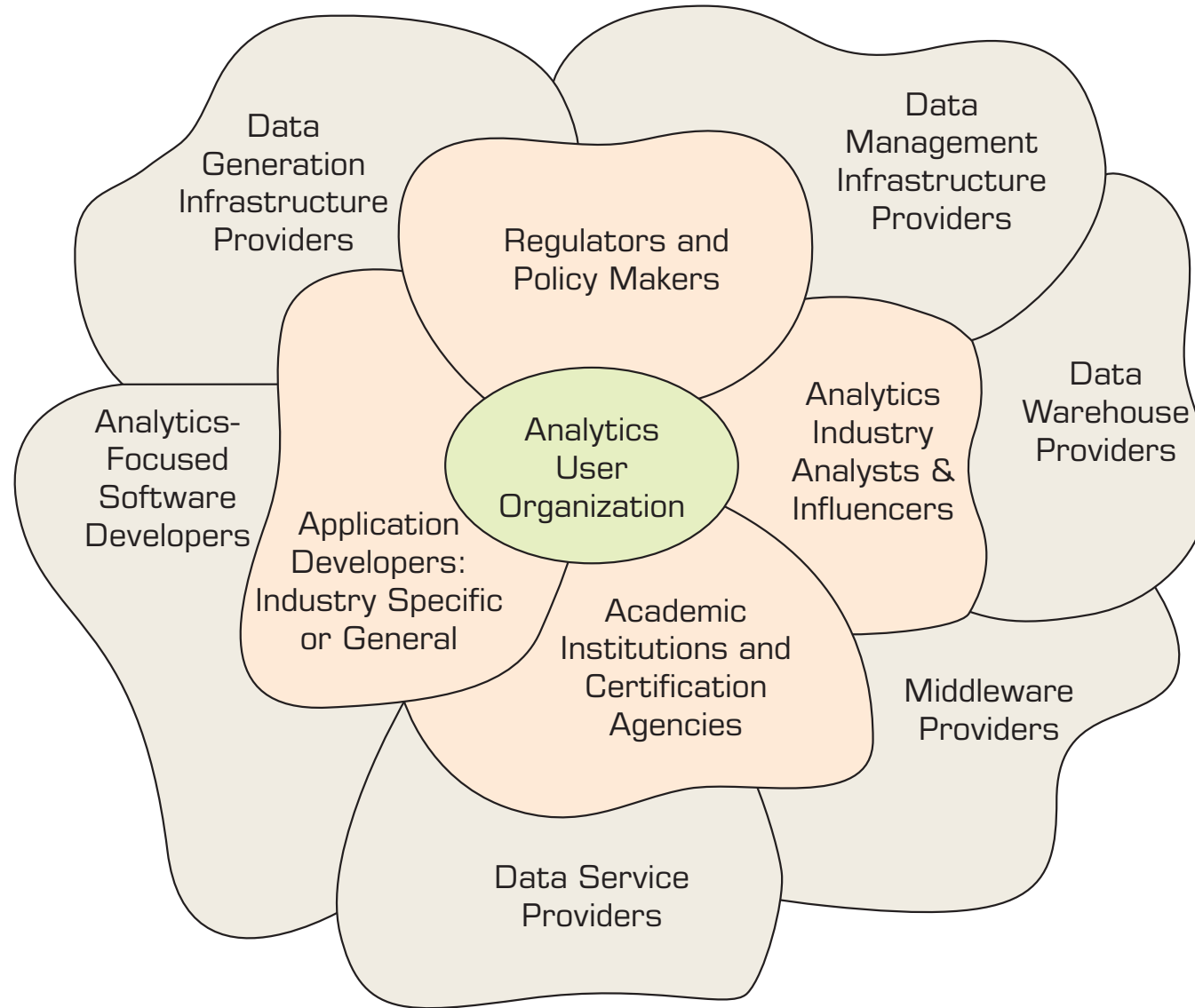
Example of Analytics Applications in a Retail Value Chain

Retail Value Chain

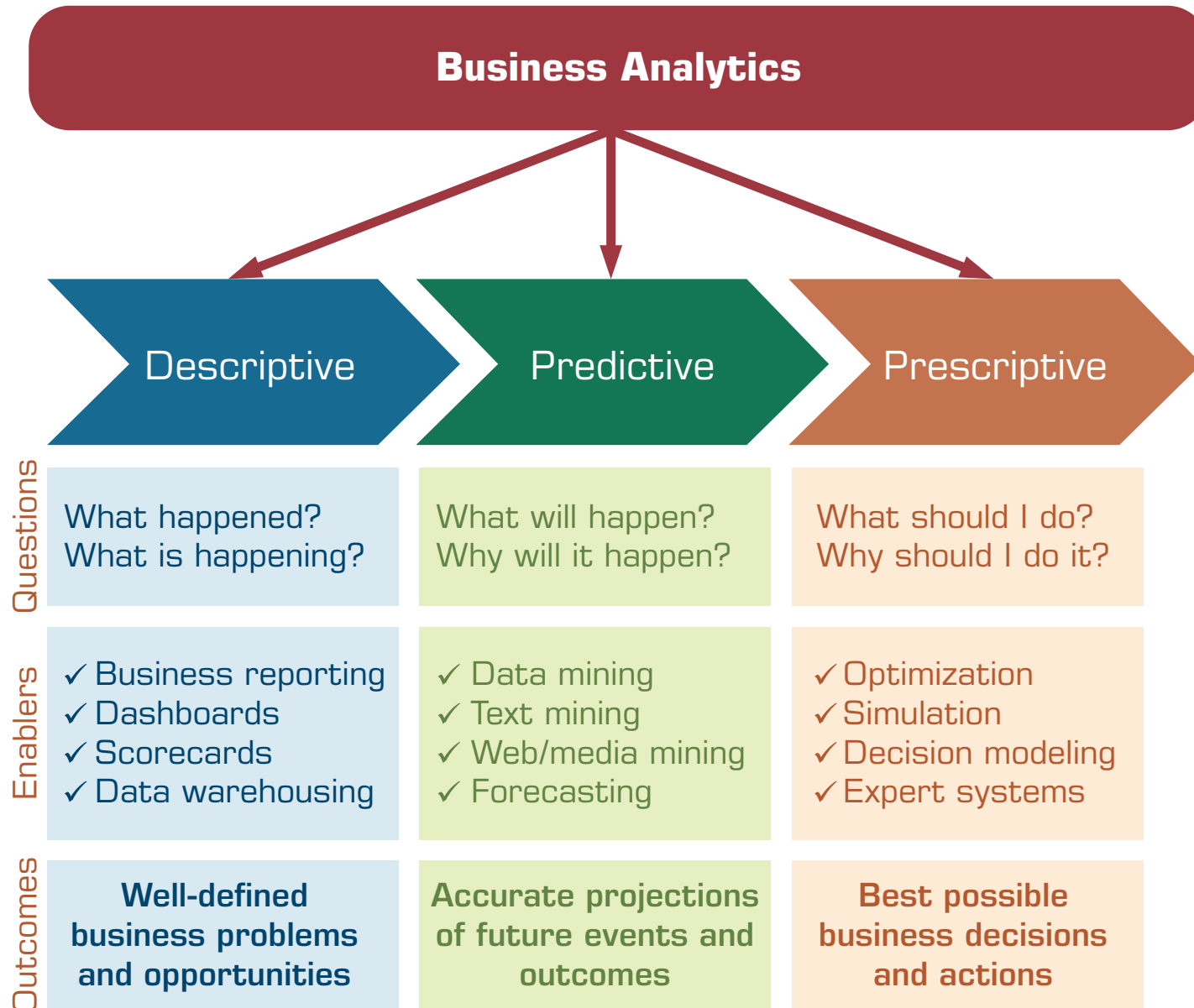
Critical needs at every touch point of the Retail Value Chain



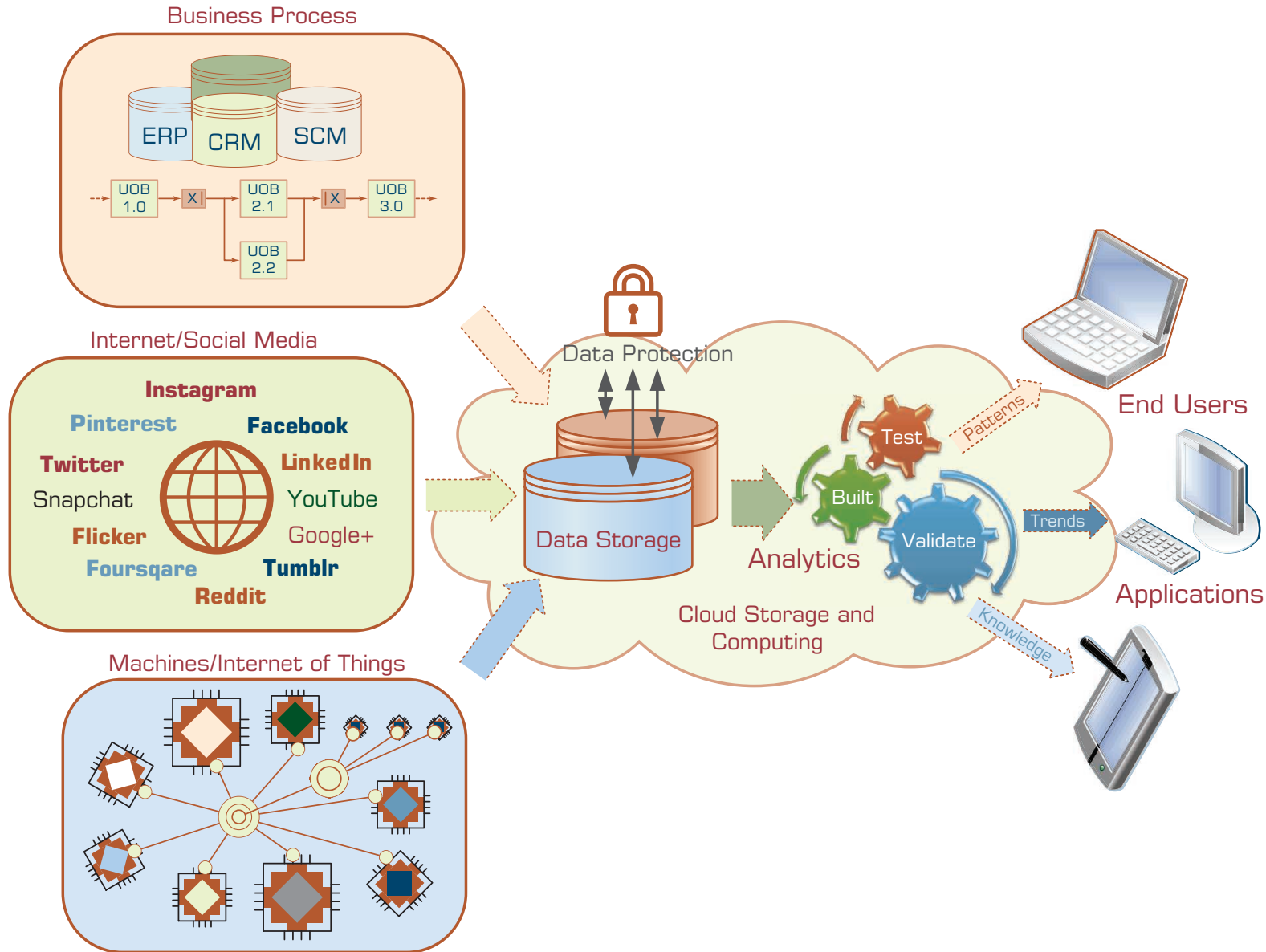
Analytics Ecosystem



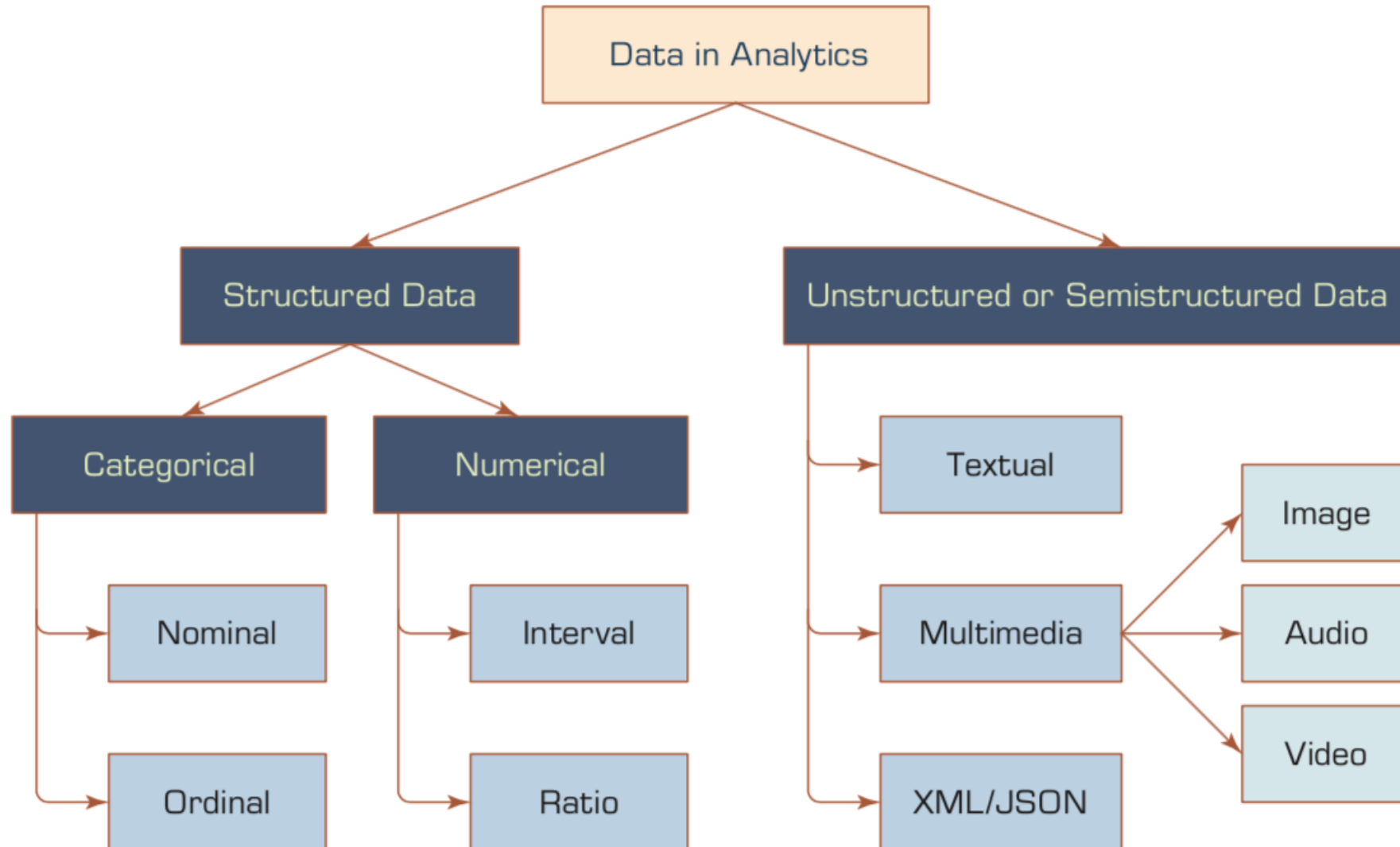
Three Types of Analytics



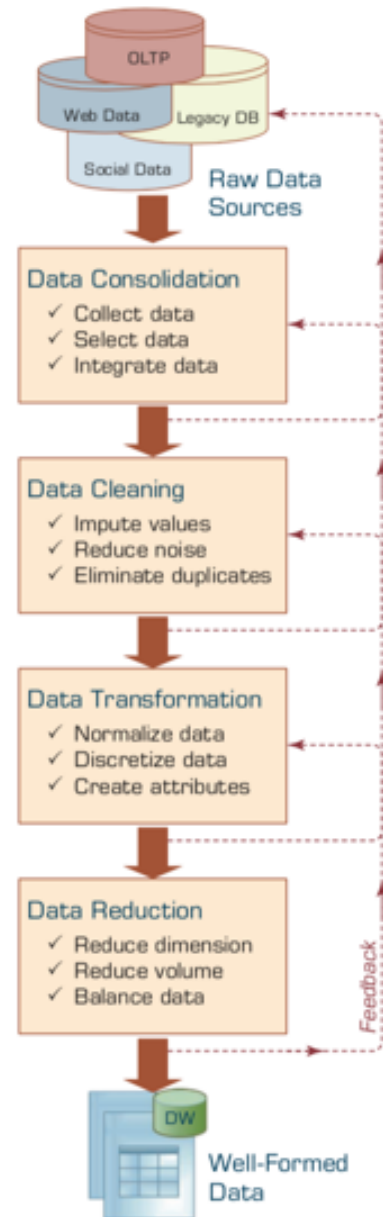
A Data to Knowledge Continuum



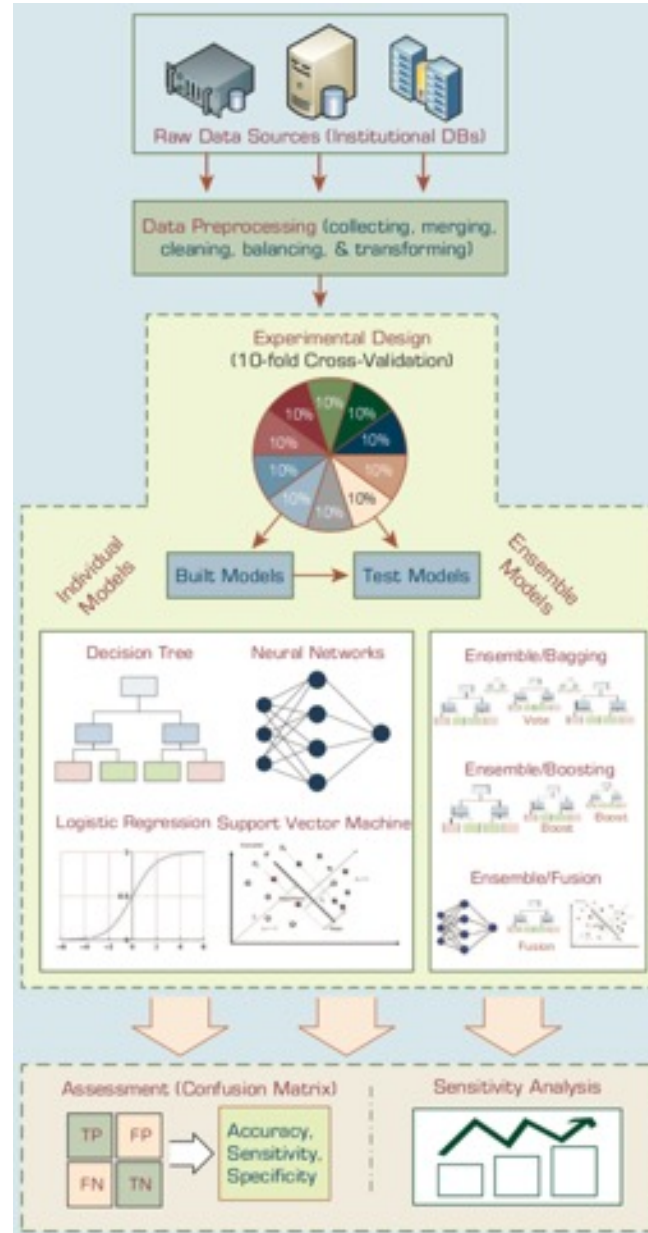
A Simple Taxonomy of Data



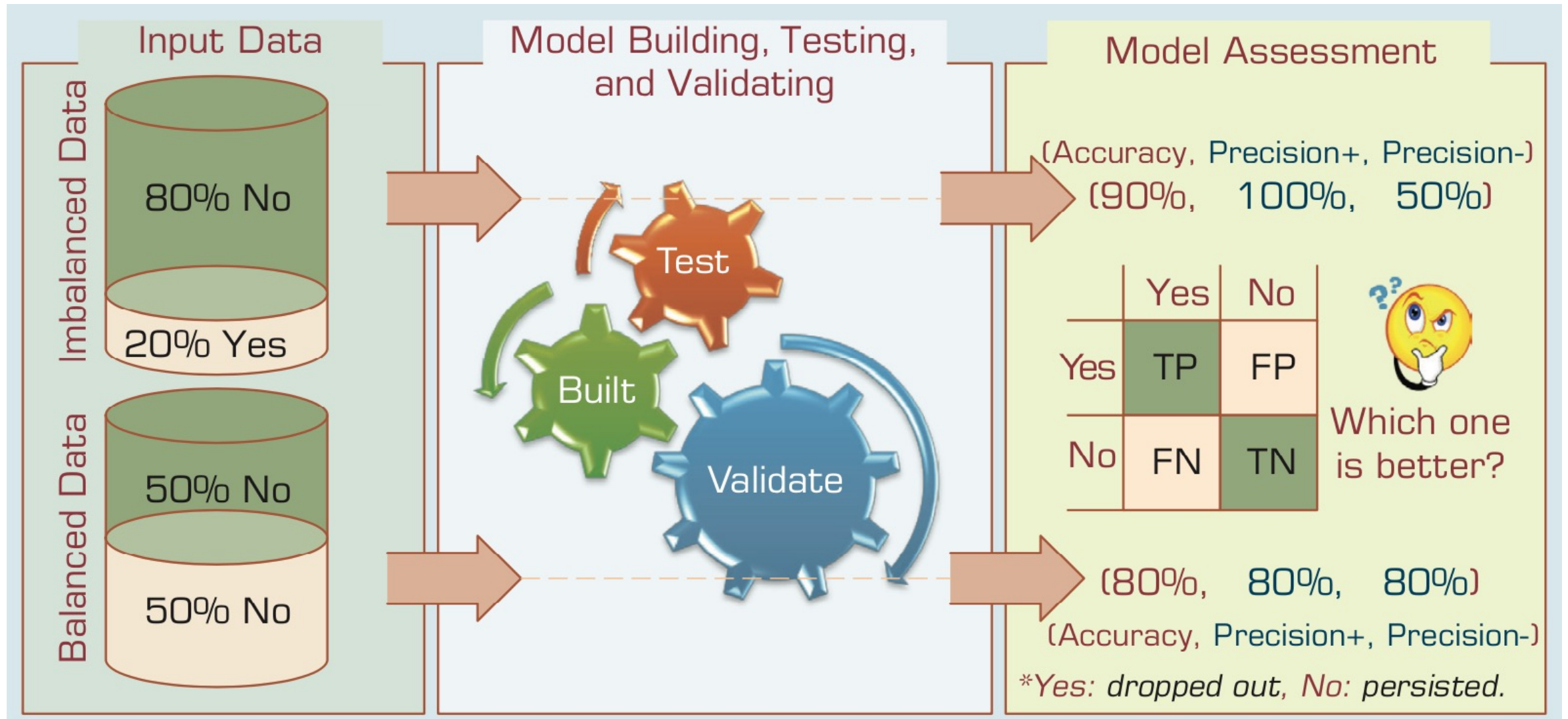
Data Preprocessing Steps



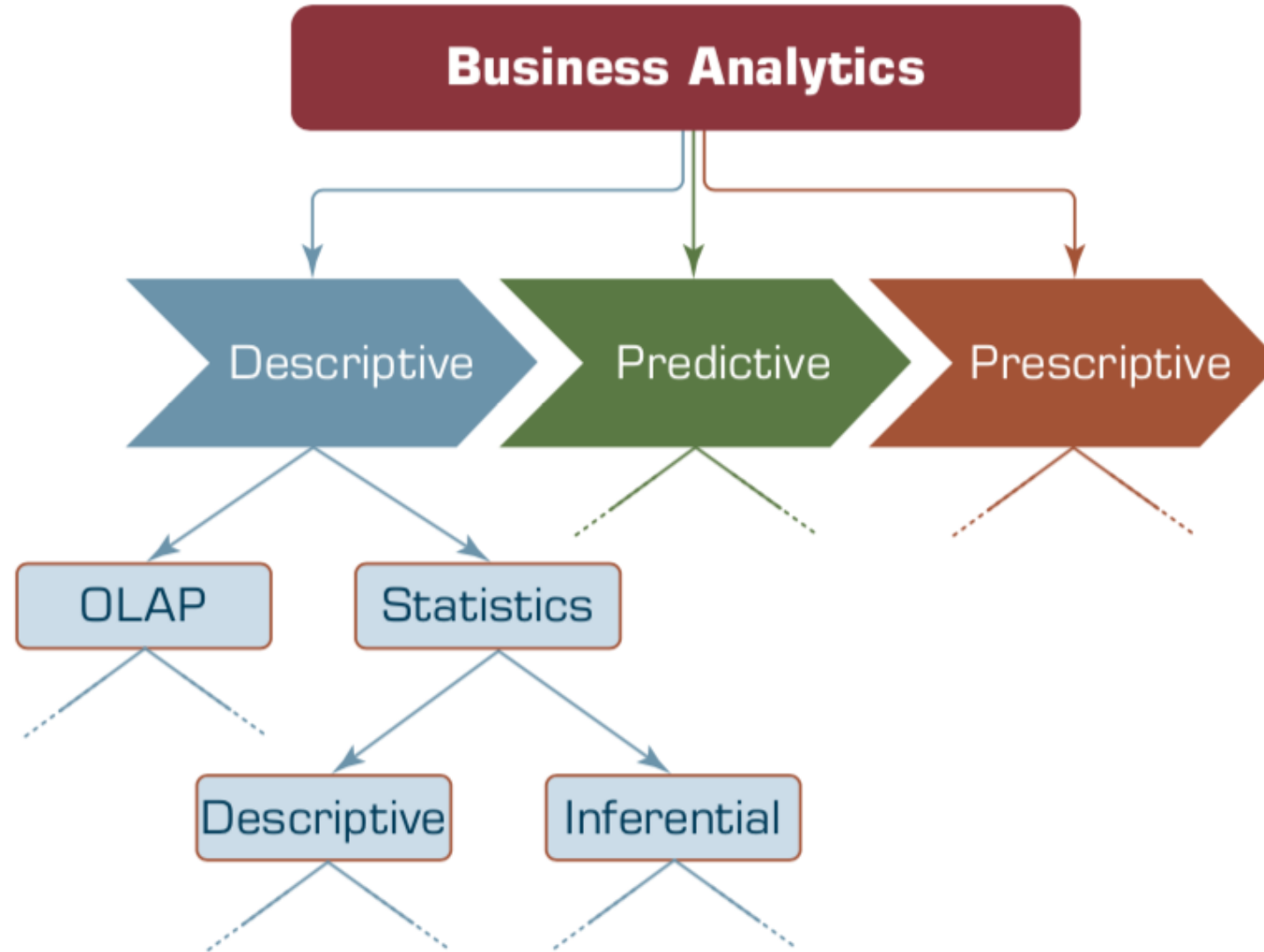
An Analytics Approach to Predicting Student Attrition



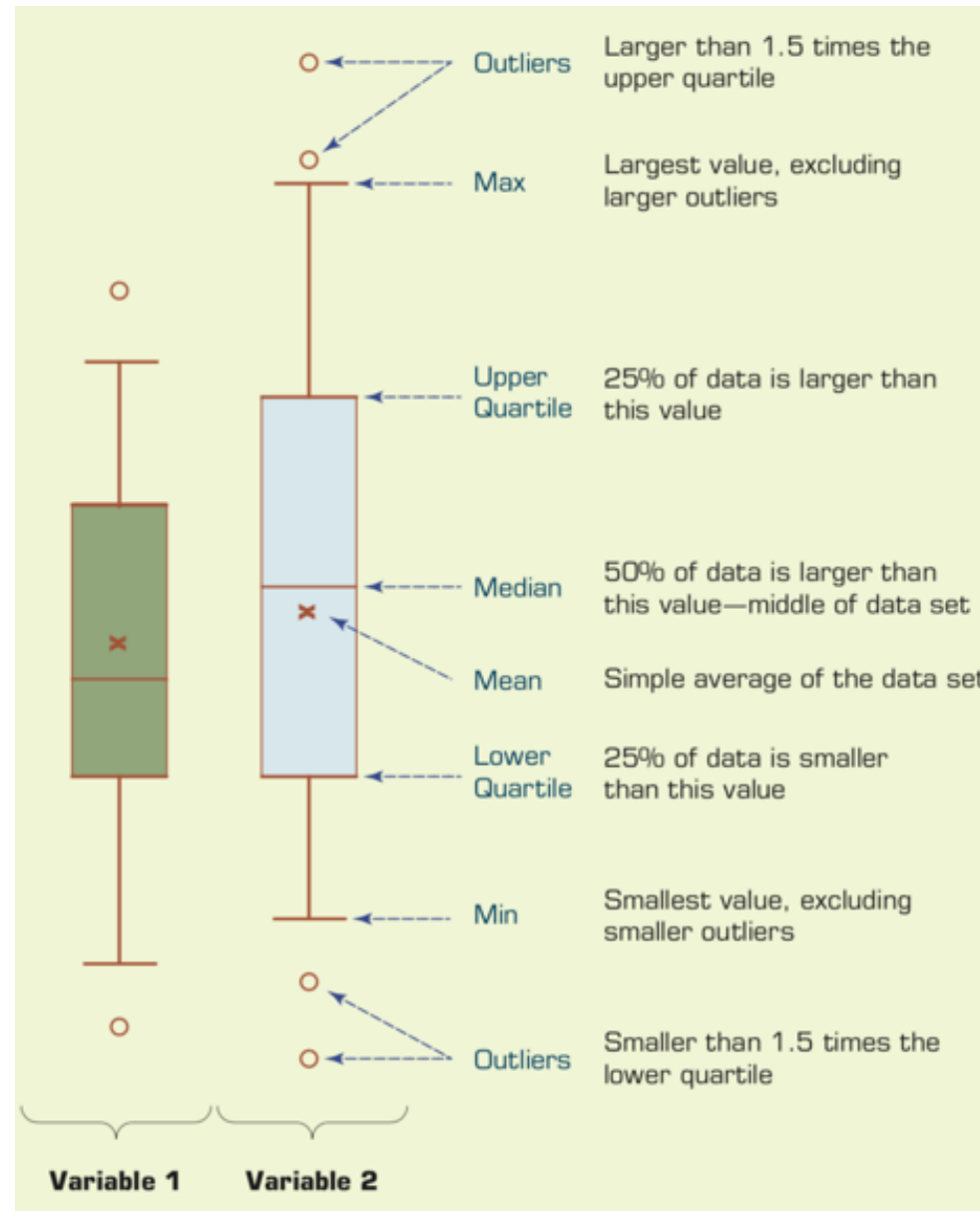
A Graphical Depiction of the Class Imbalance Problem



Relationship between Statistics and Descriptive Analytics



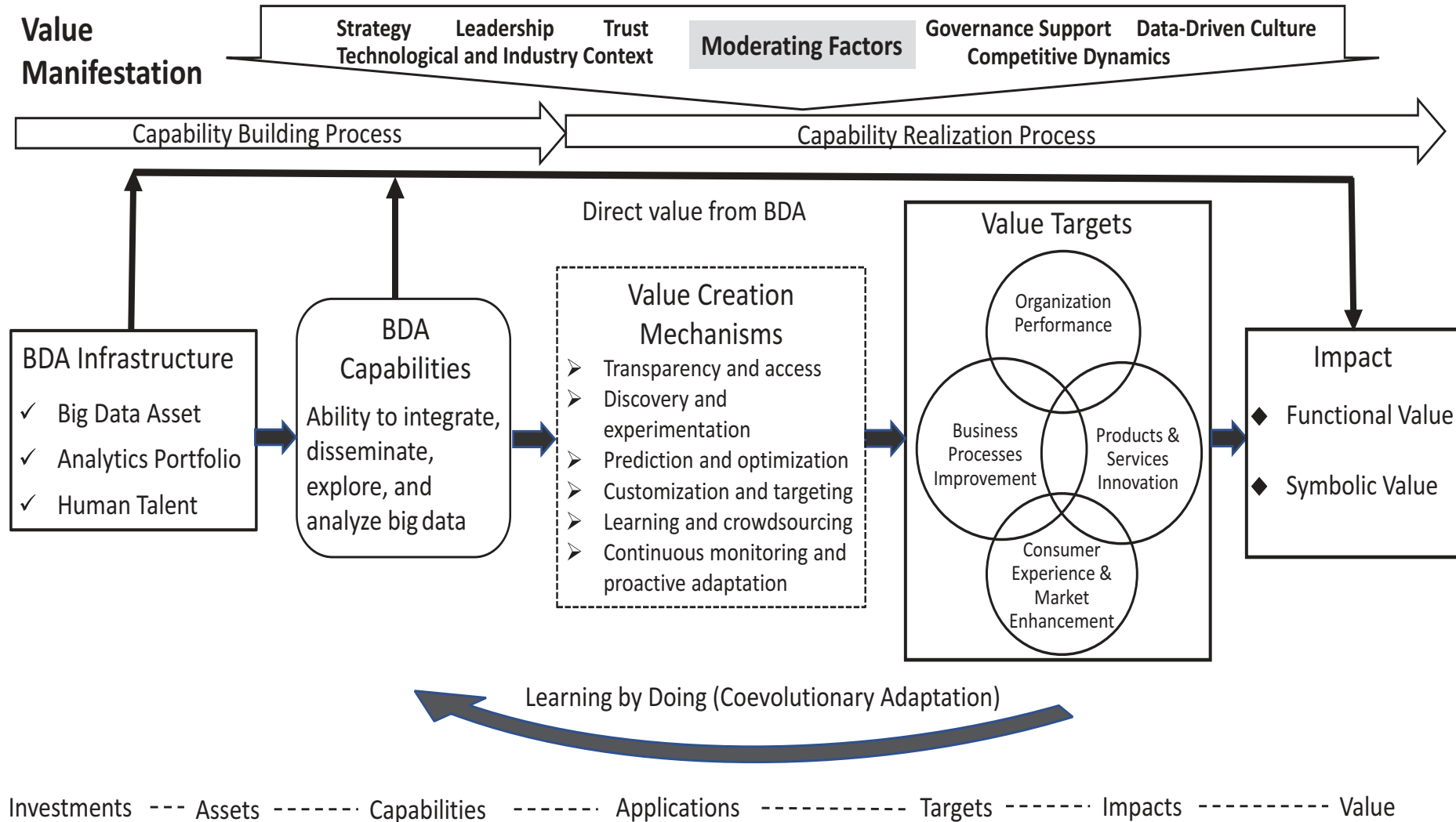
Understanding the Specifics about Box-and-Whiskers Plots



Big Data Analytics (BDA)

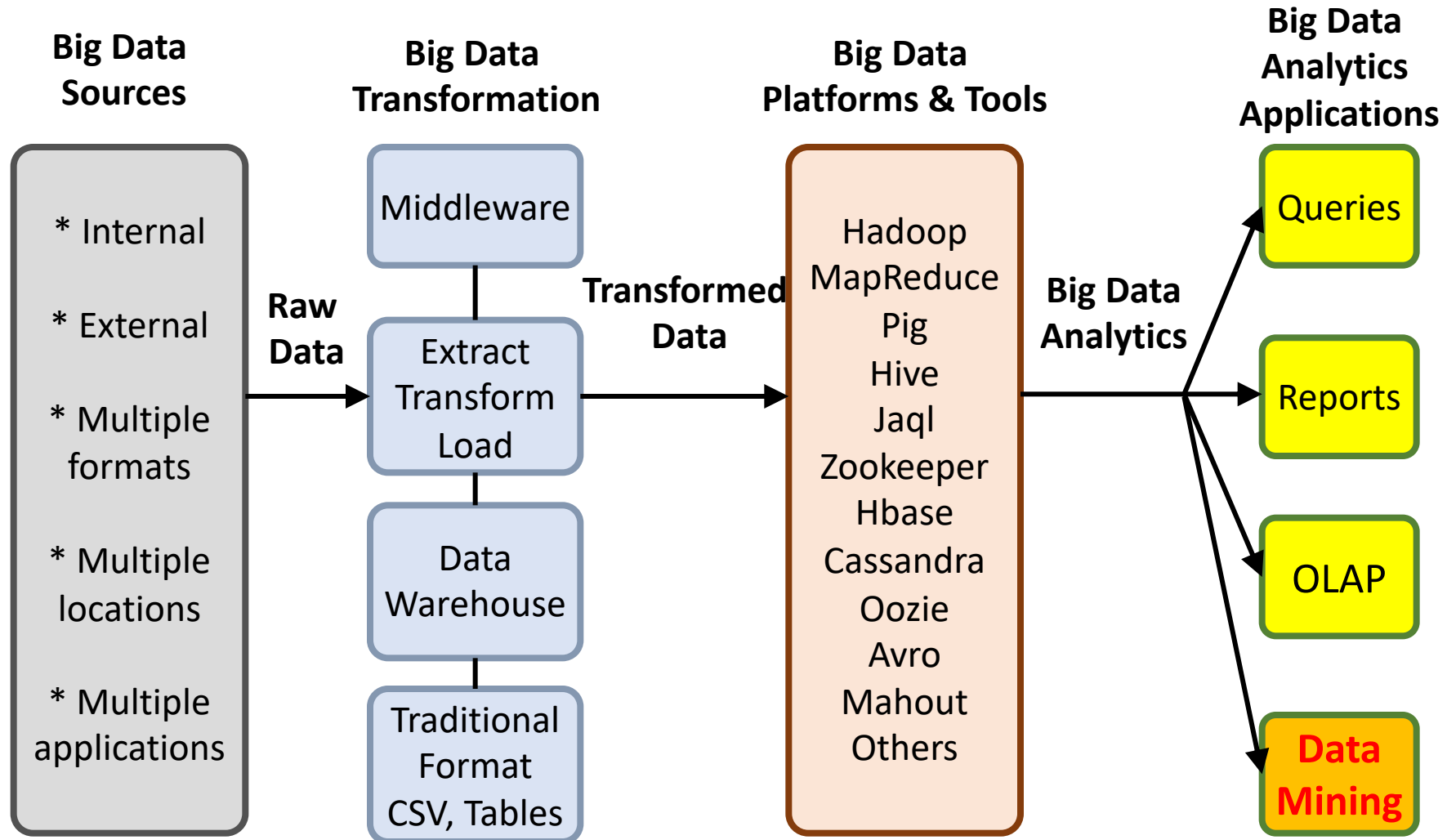
Value Creation by Big Data Analytics

(Grover et al., 2018)



Source: Varun Grover, Roger HL Chiang, Ting-Peng Liang, and Dongsong Zhang (2018), "Creating Strategic Business Value from Big Data Analytics: A Research Framework", Journal of Management Information Systems, 35, no. 2, pp. 388-423.

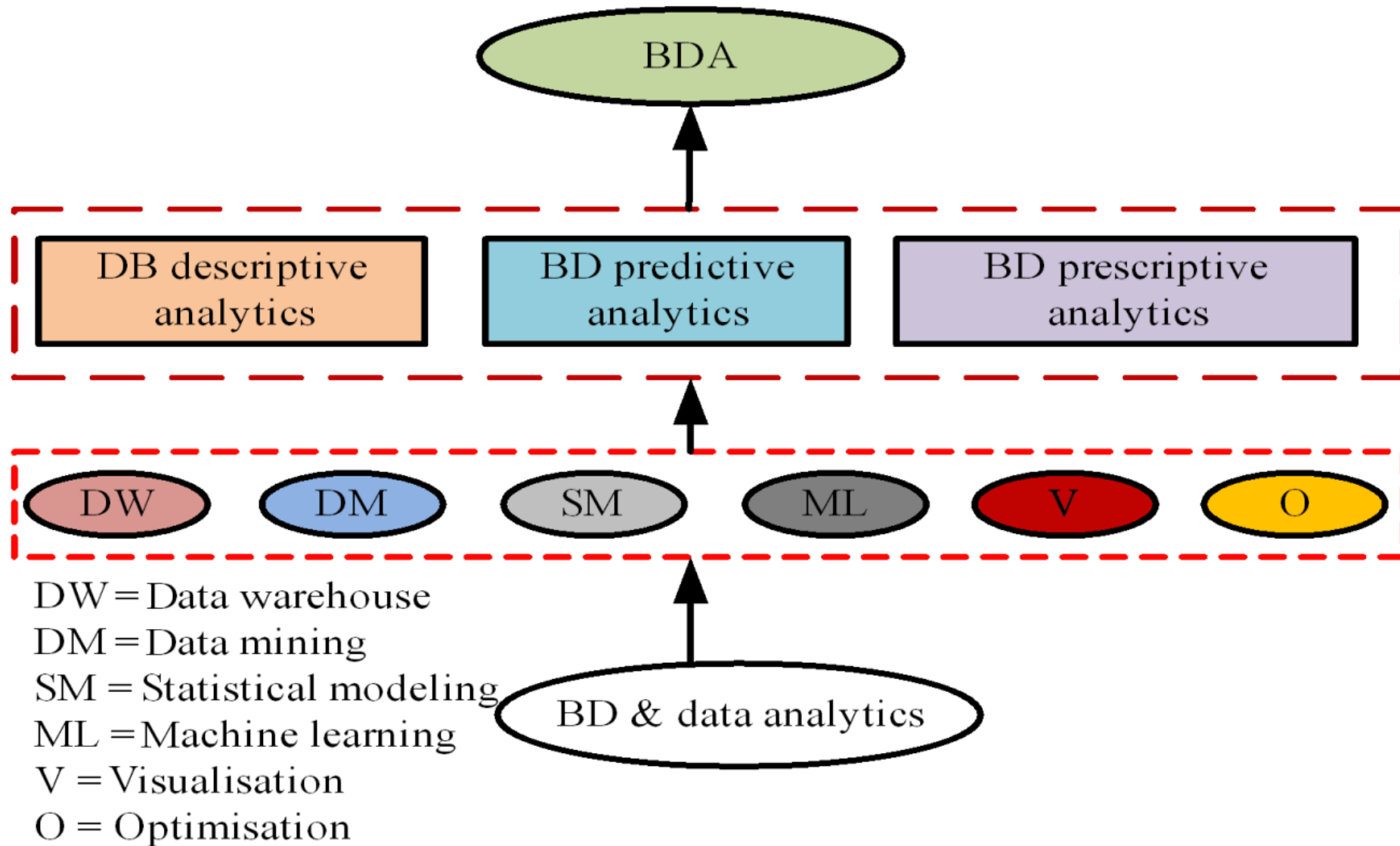
Architecture of Big Data Analytics



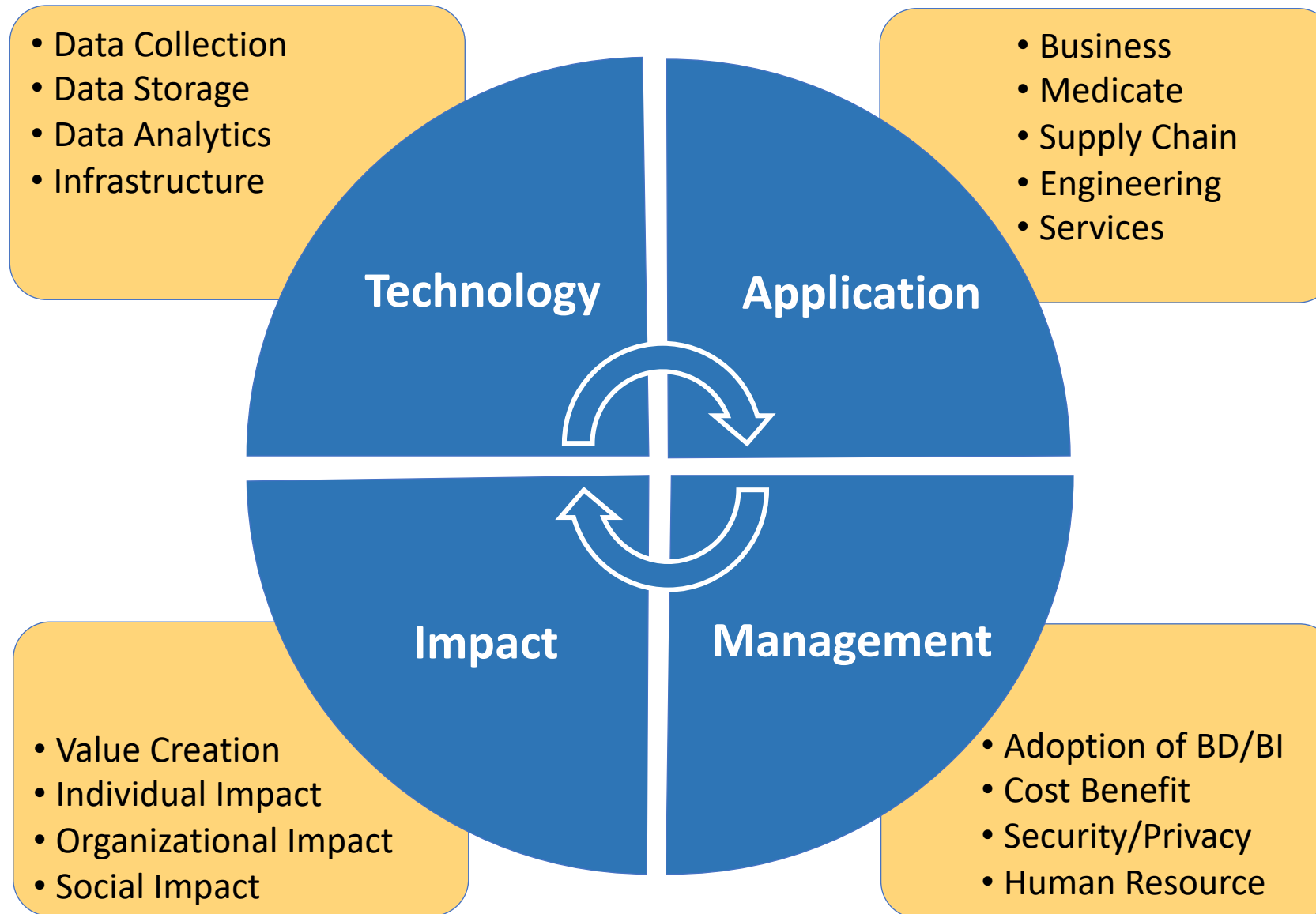
Architecture of Big Data Analytics



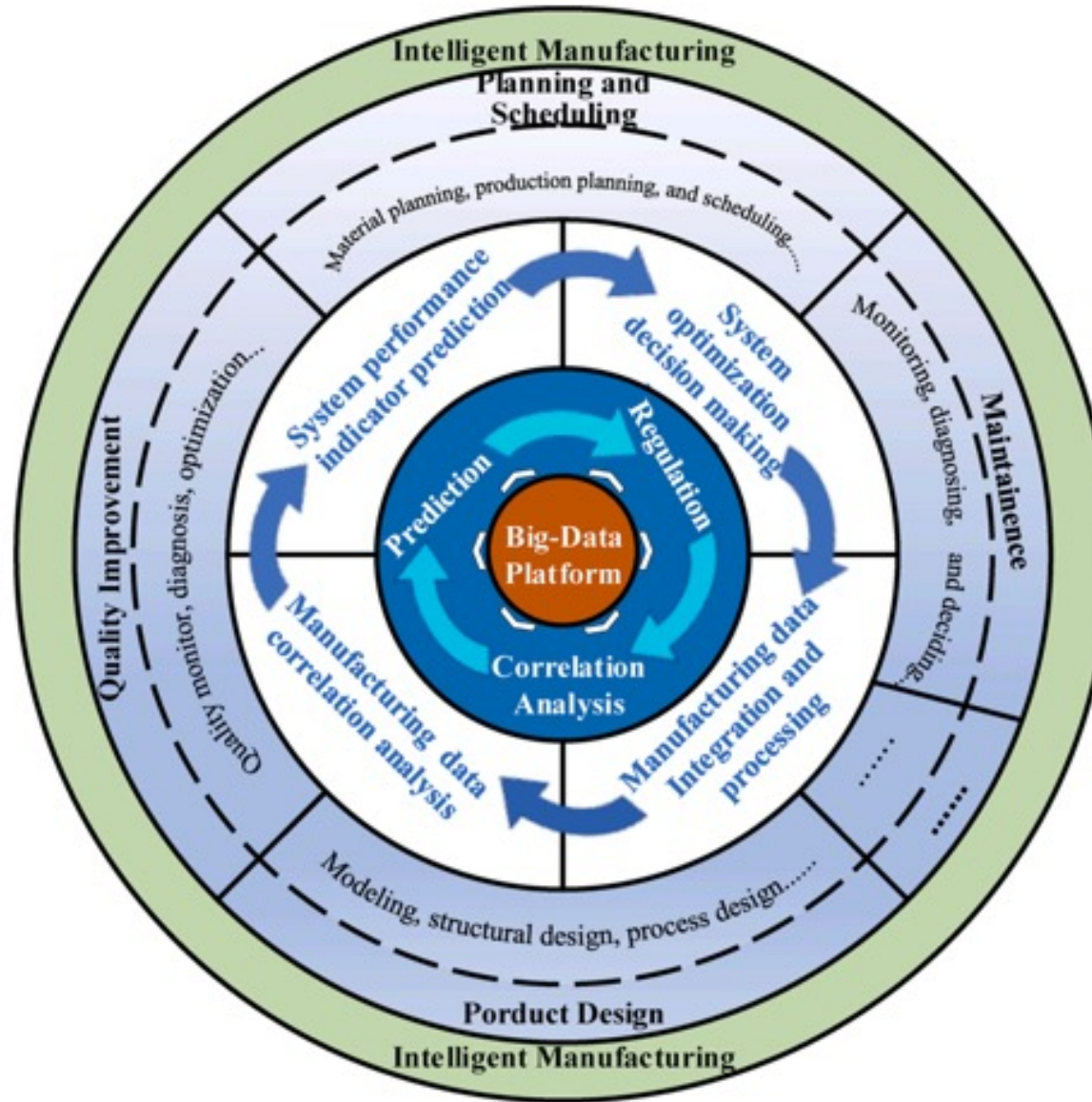
Taxonomy of Big Data Analytics (BDA)



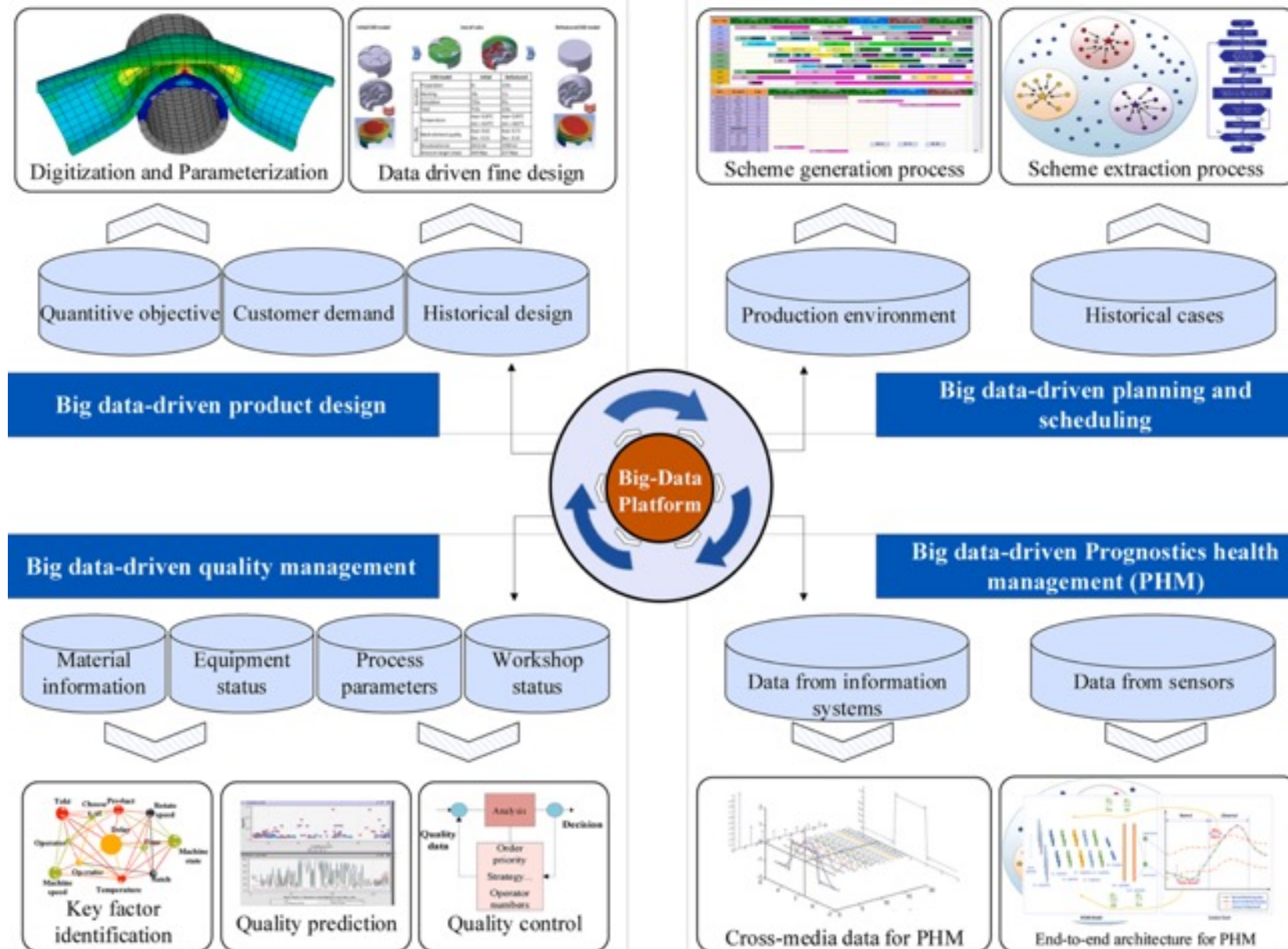
Framework for BD and BI Research



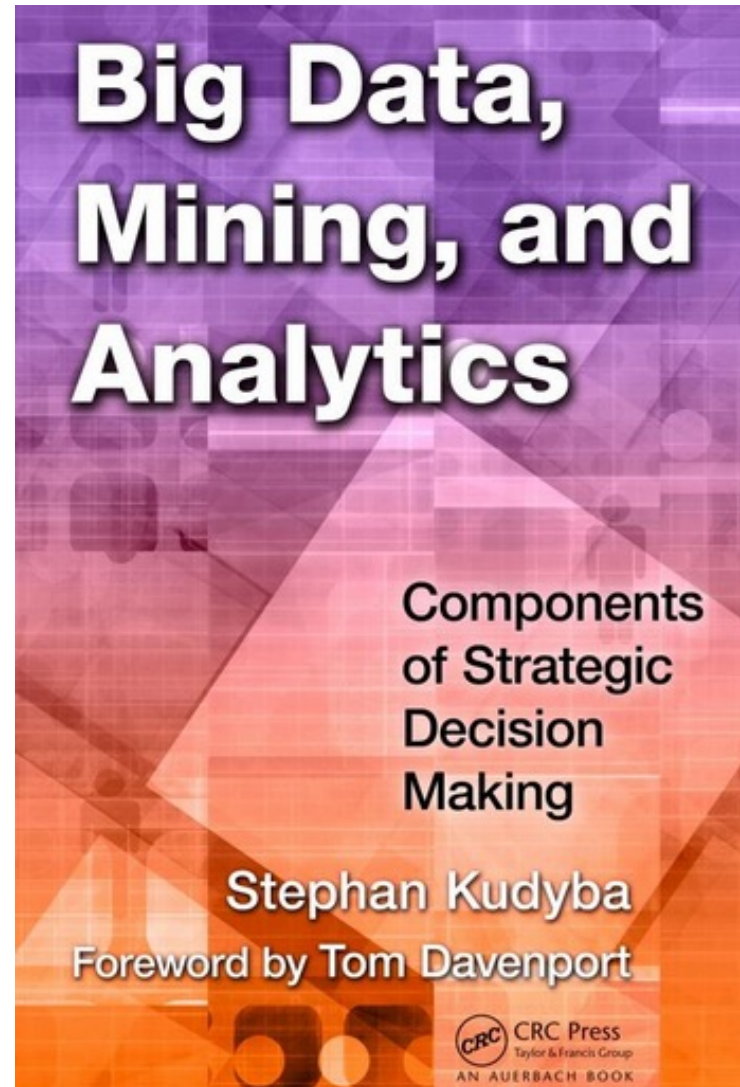
Big Data Driven Intelligent Manufacturing



Applications of BDA in Manufacturing Systems



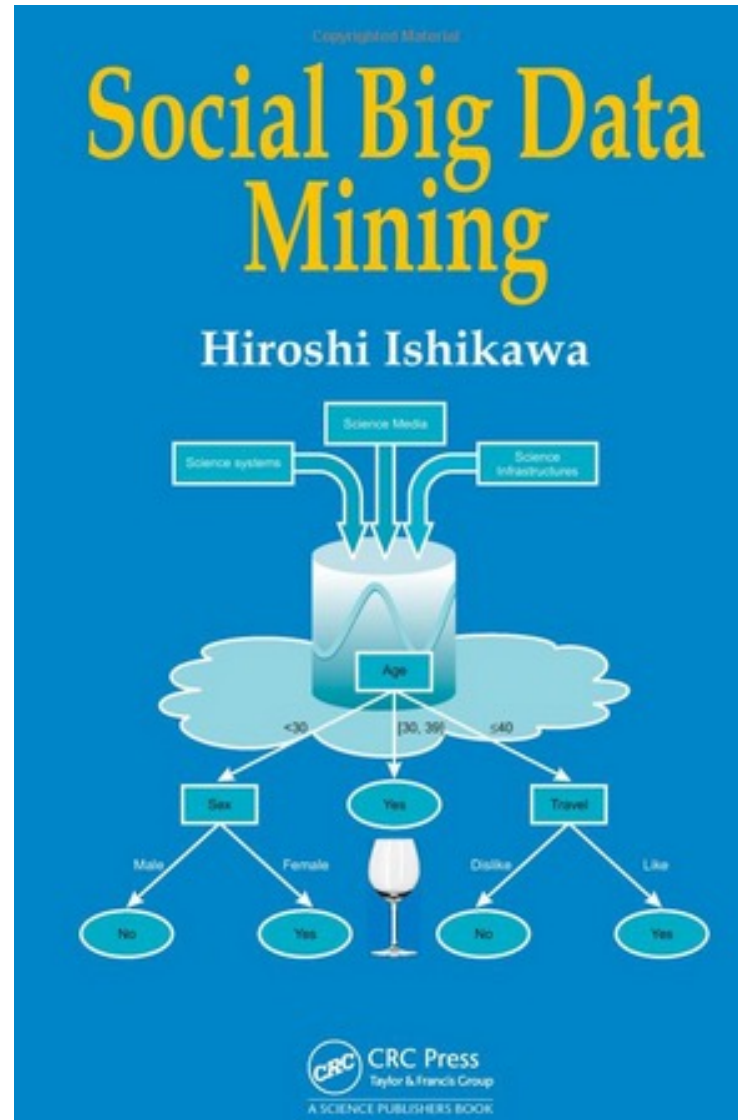
Stephan Kudyba (2014),
Big Data, Mining, and Analytics:
Components of Strategic Decision Making, Auerbach Publications



Source: <http://www.amazon.com/gp/product/1466568704>

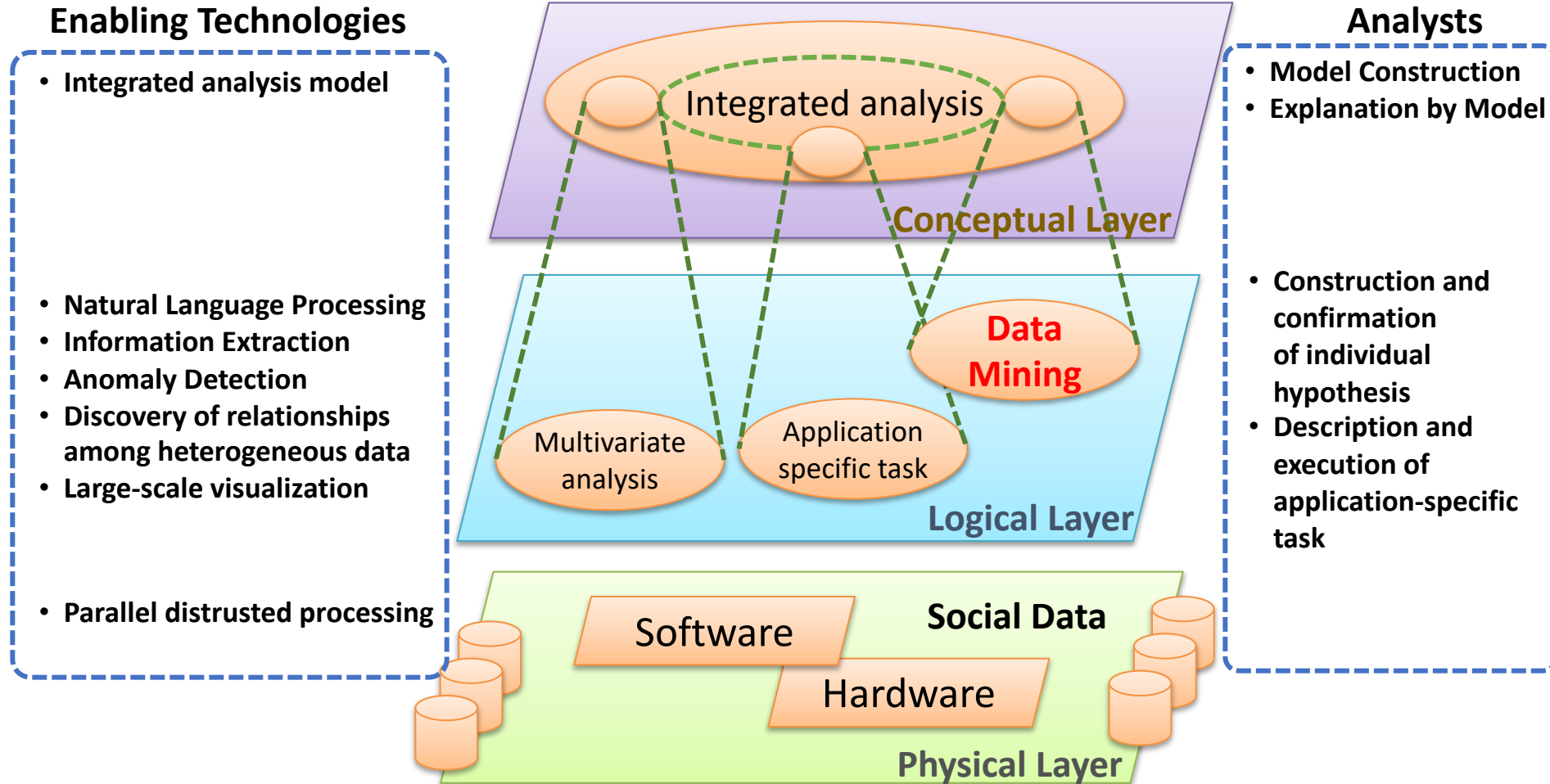
Social Big Data Mining

(Hiroshi Ishikawa, 2015)

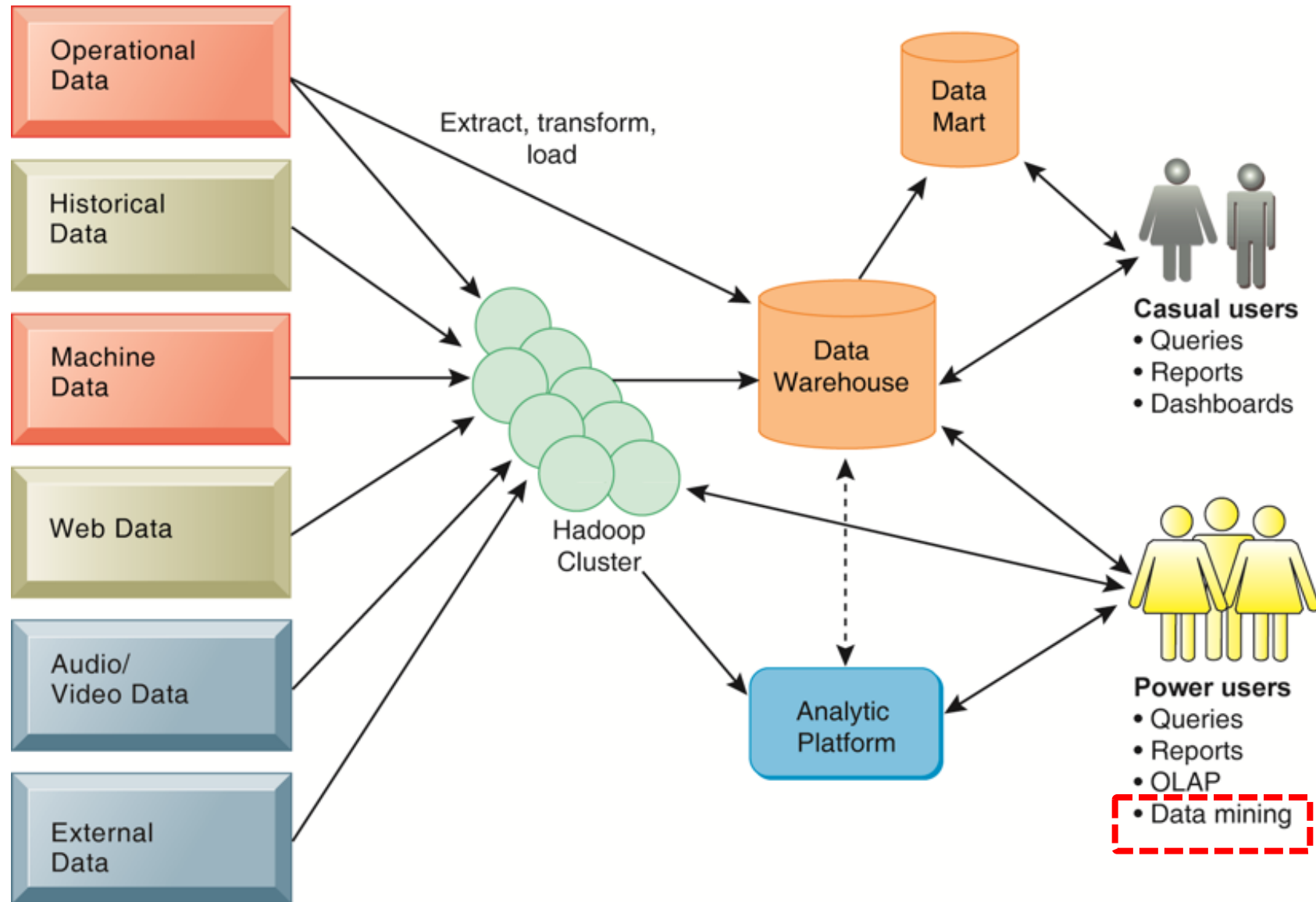


Architecture for Social Big Data Mining

(Hiroshi Ishikawa, 2015)

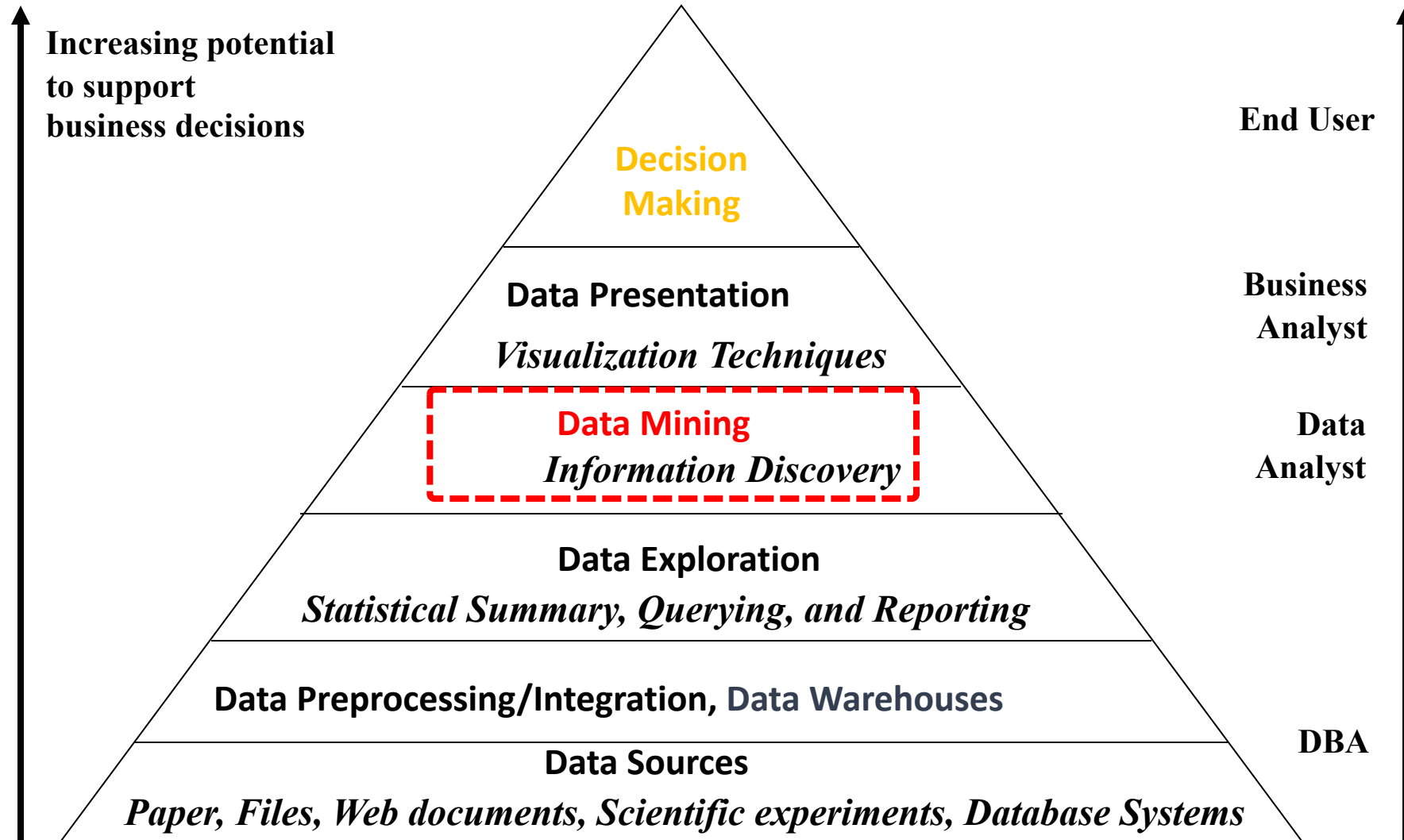


Business Intelligence (BI) Infrastructure

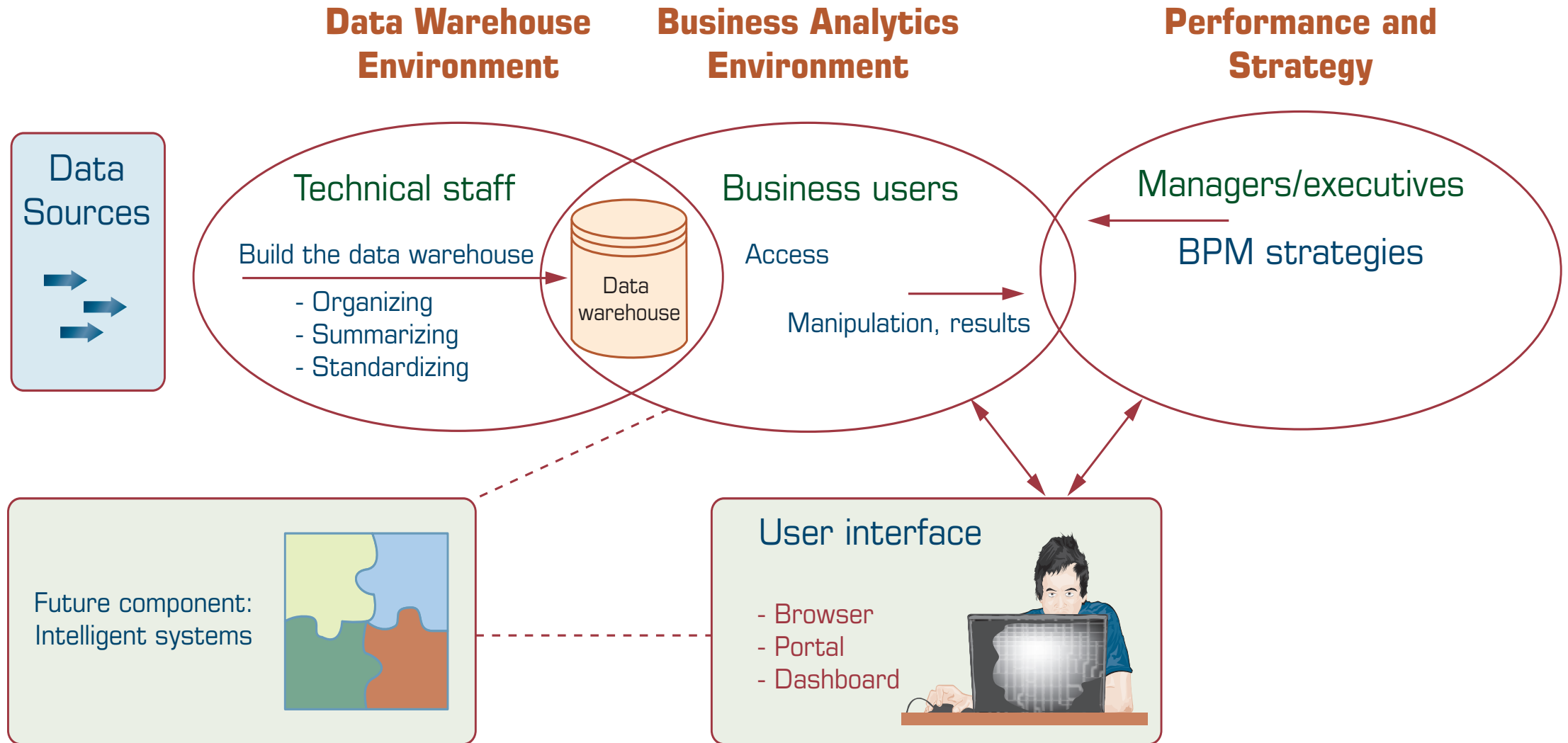


Data Warehouse

Data Mining and Business Intelligence



A High-Level Architecture of BI



The Evolution of BI Capabilities

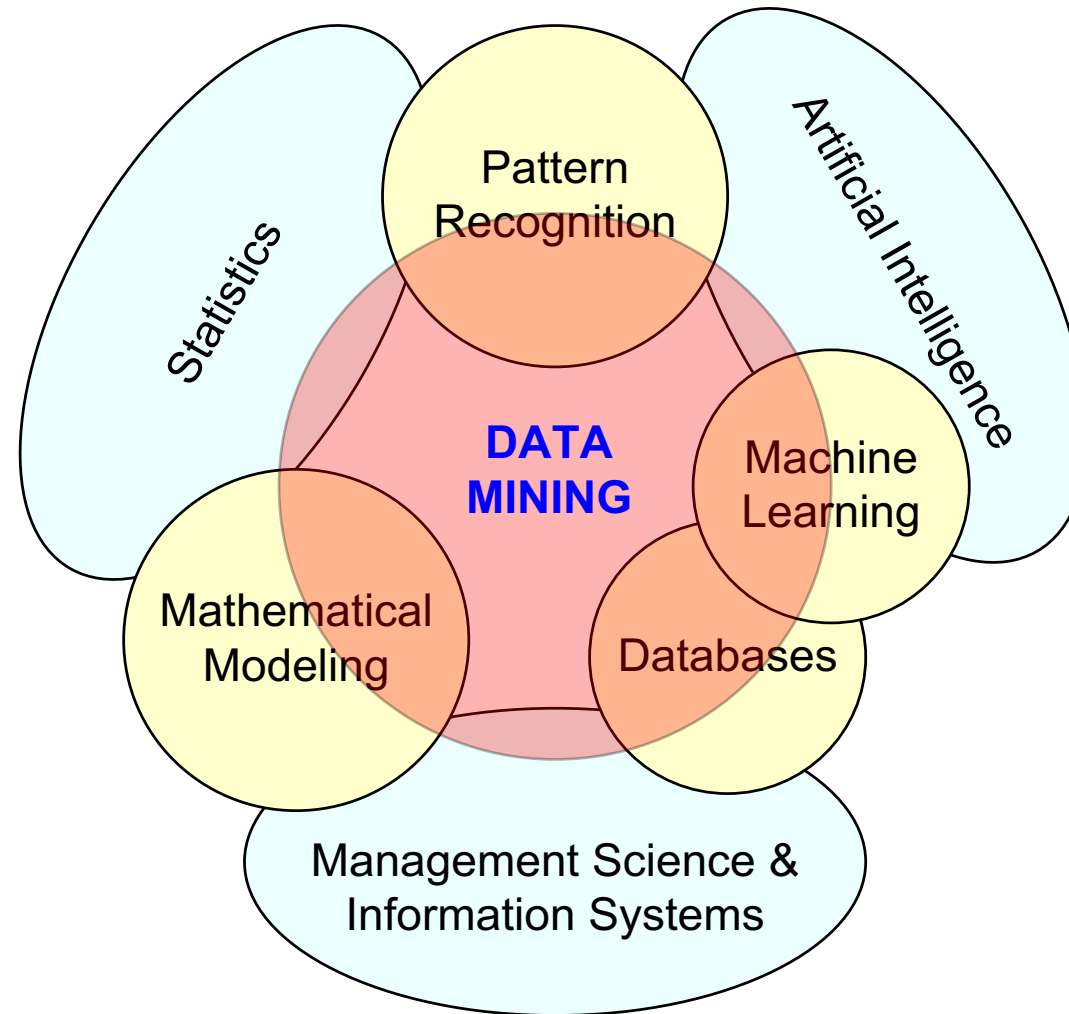


Data Mining

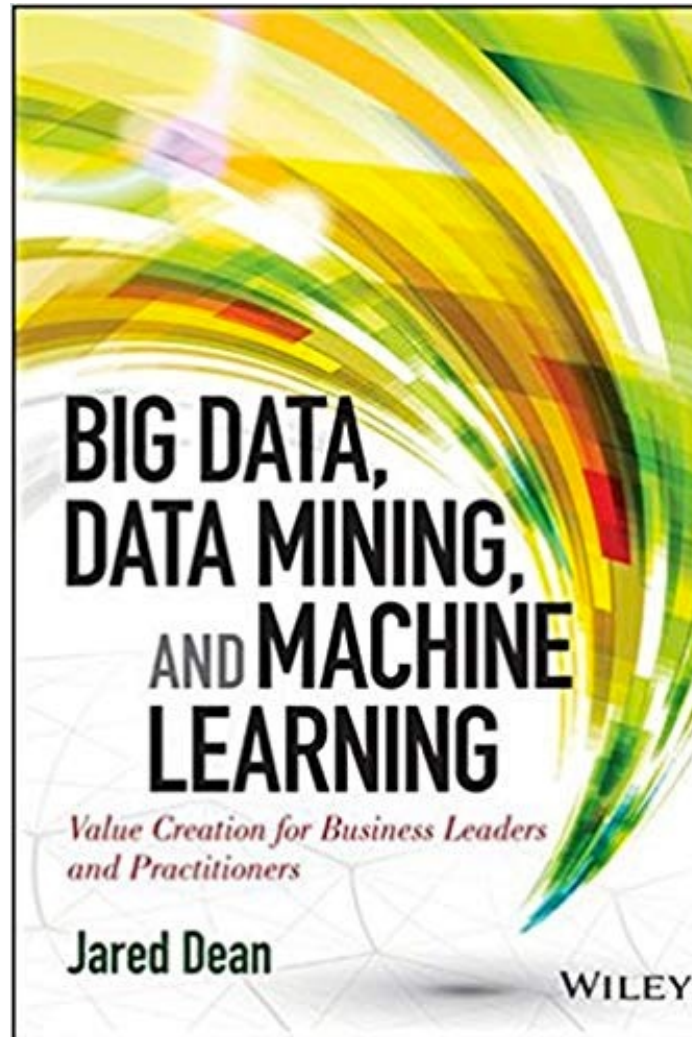
Is a Blend of Multiple Disciplines

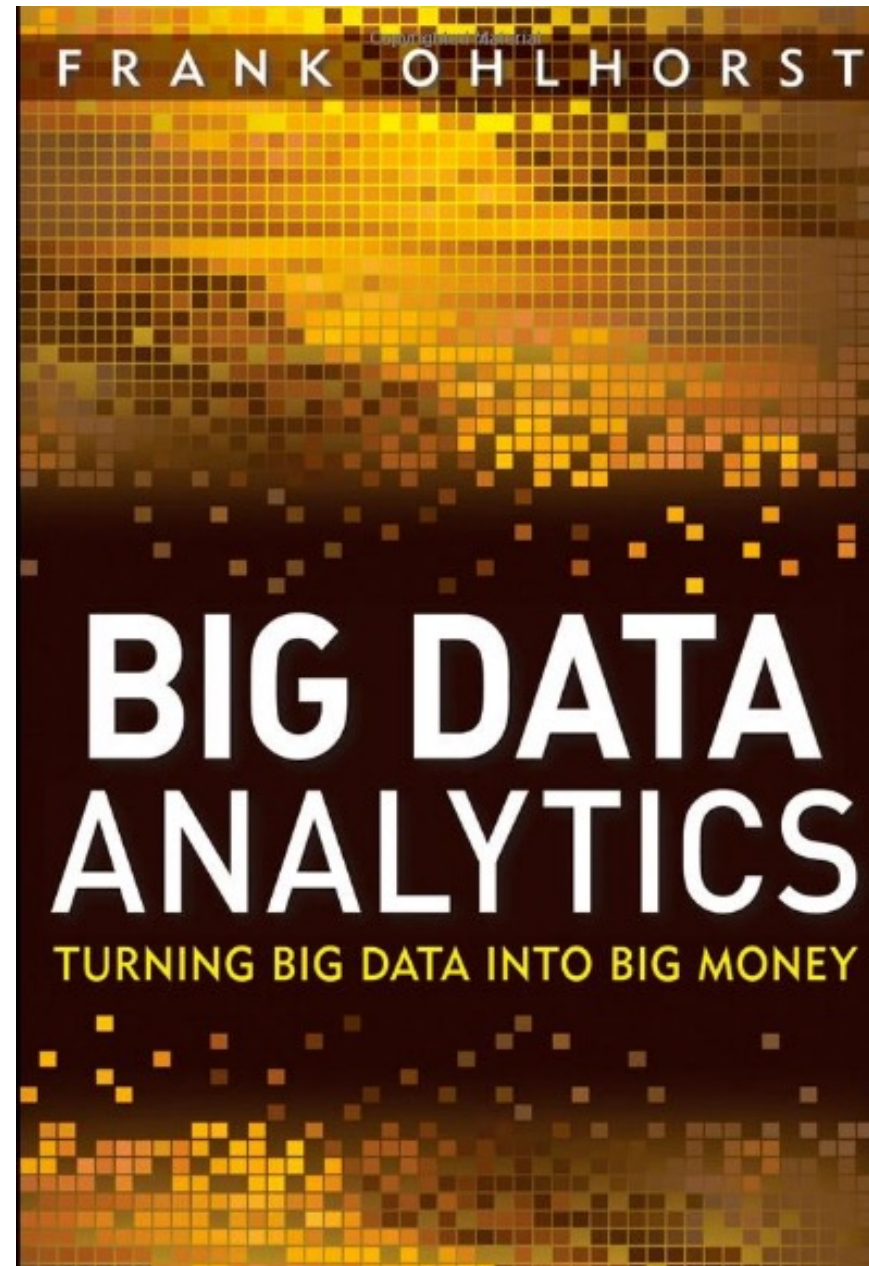


Data Mining at the Intersection of Many Disciplines



**Big Data, Data Mining, and Machine Learning: Value Creation for
Business Leaders and Practitioners,
Jared Dean,
Wiley, 2014.**









VISUAL ANALYTICS

<p>DYNAMIC & INTERACTIVE</p> <p>Dashboard Graph Map</p>	<p>ENHANCE</p> <p>Understanding Investigation User Experience</p>
--	--

BIG ANALYTICS

<p>QUERY & FILTER</p> <p>Complex queries R²I²</p>	<p>DETECT</p> <p>Anomalies Communities Typologies</p>	<p>PREDICT</p> <p>Tending Real-time Prediction</p>	<p>DECIDE</p> <p>Simulation Optimization</p>
--	--	---	---

BIG DATA – Batch

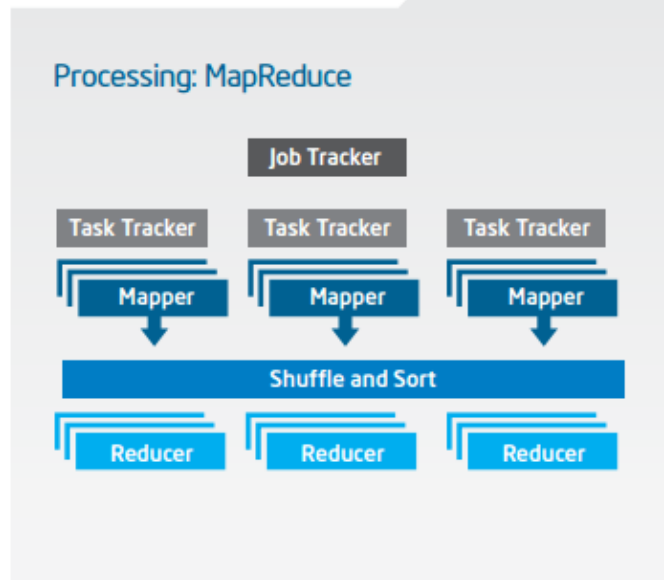
BIG DATA – Real Time

DATA

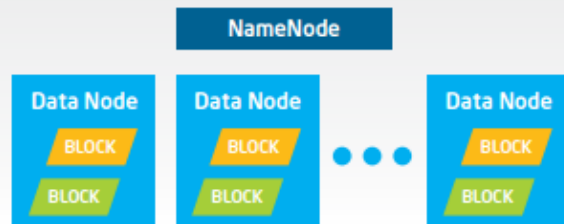
<p>Complex by nature</p>		<p>Complex by structure</p>
--------------------------	--	-----------------------------

Big Data with Hadoop Architecture

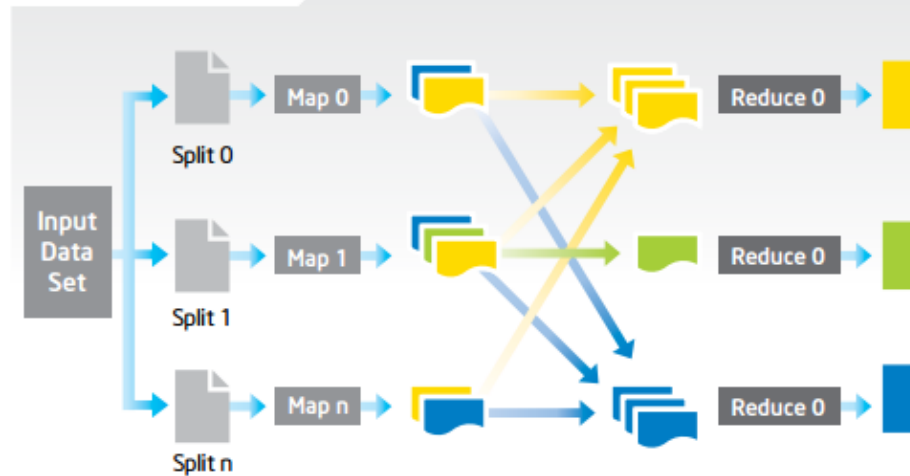
LOGICAL ARCHITECTURE



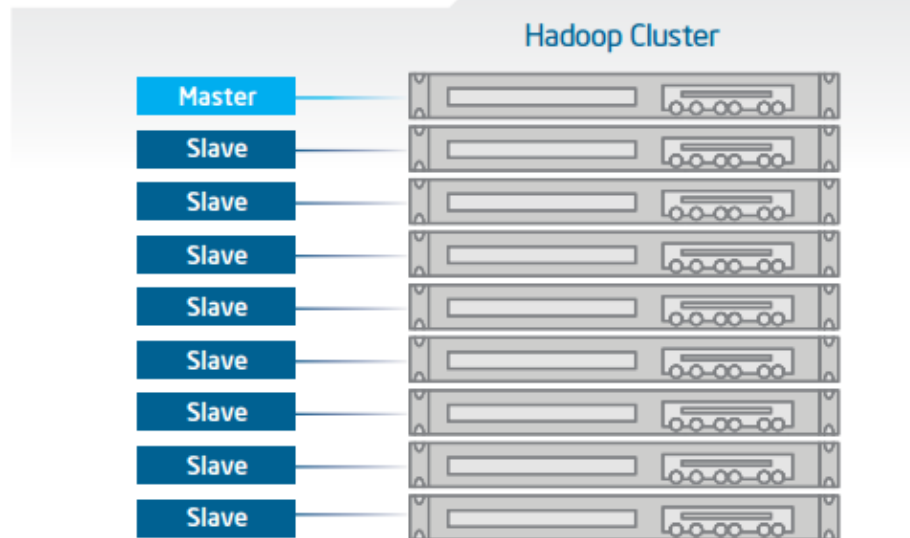
Storage: HDFS



PROCESS FLOW



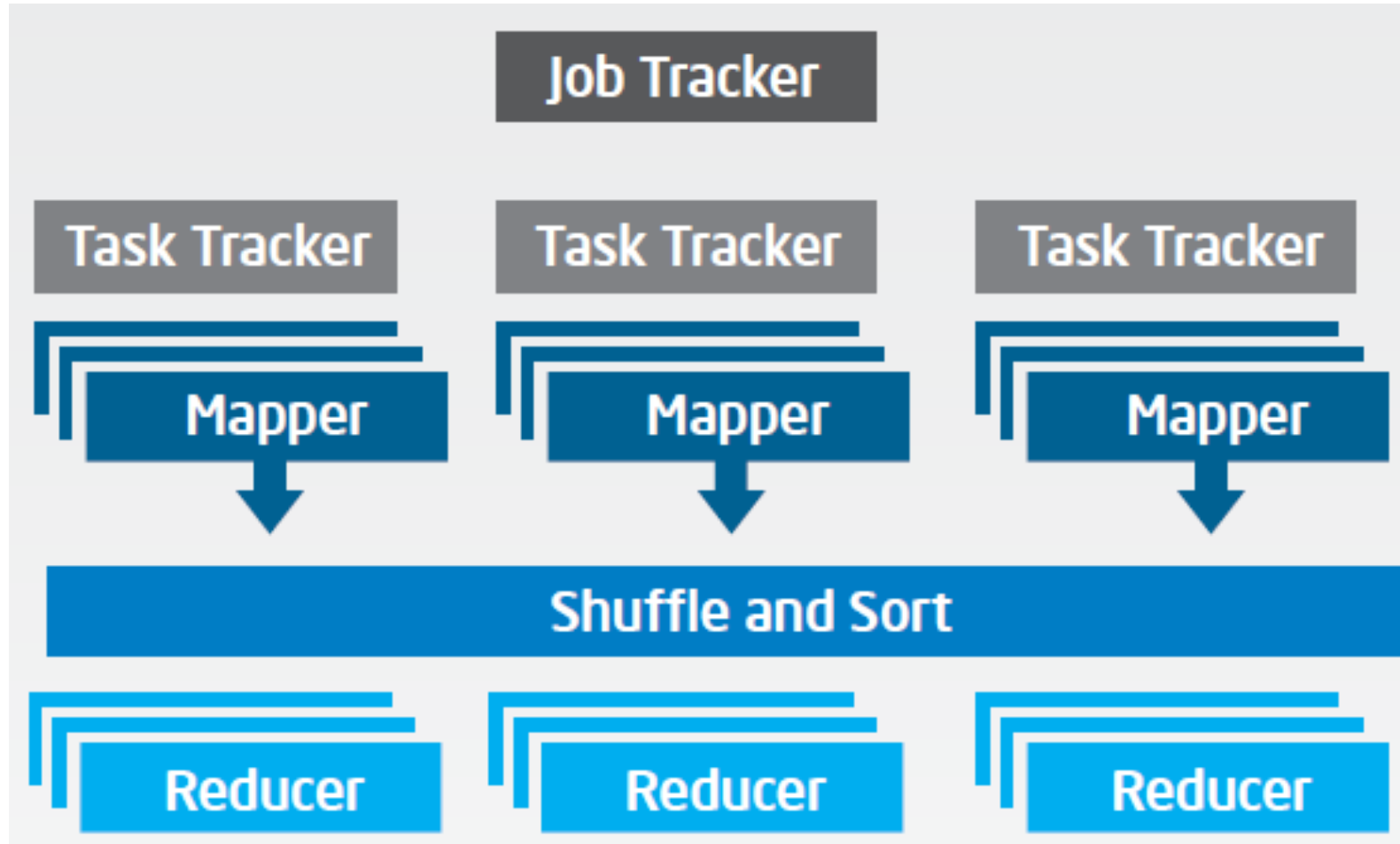
PHYSICAL ARCHITECTURE



Big Data with Hadoop Architecture

Logical Architecture

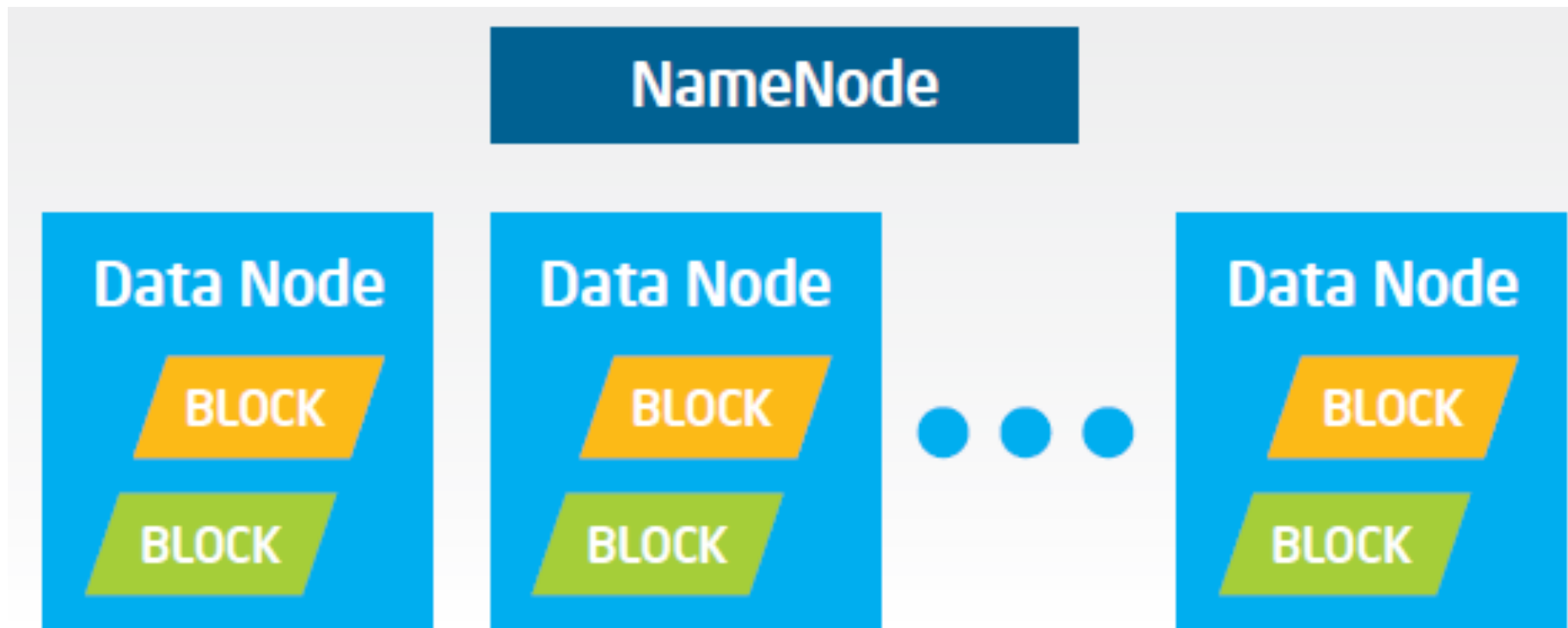
Processing: MapReduce



Big Data with Hadoop Architecture

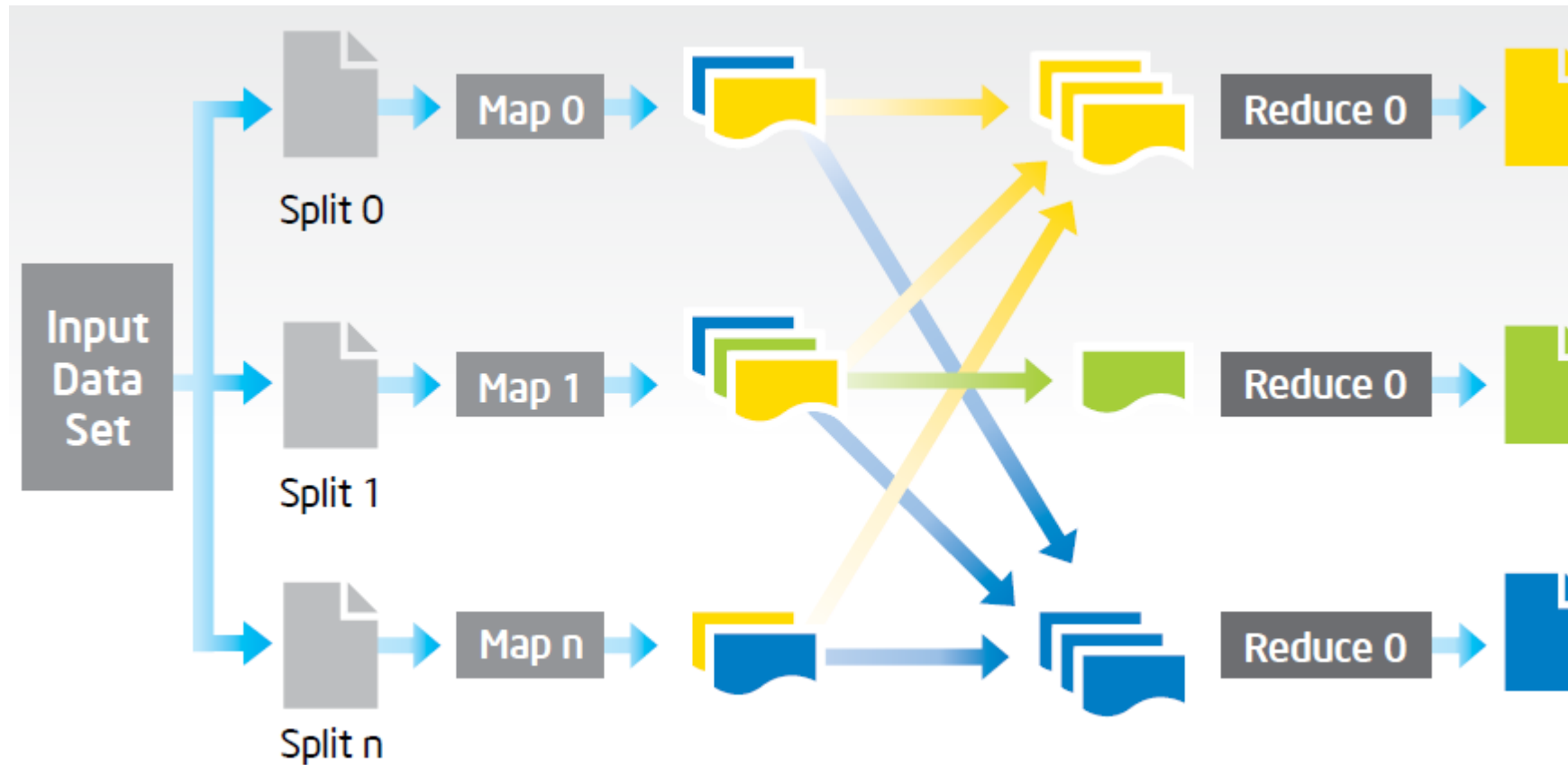
Logical Architecture

Storage: HDFS



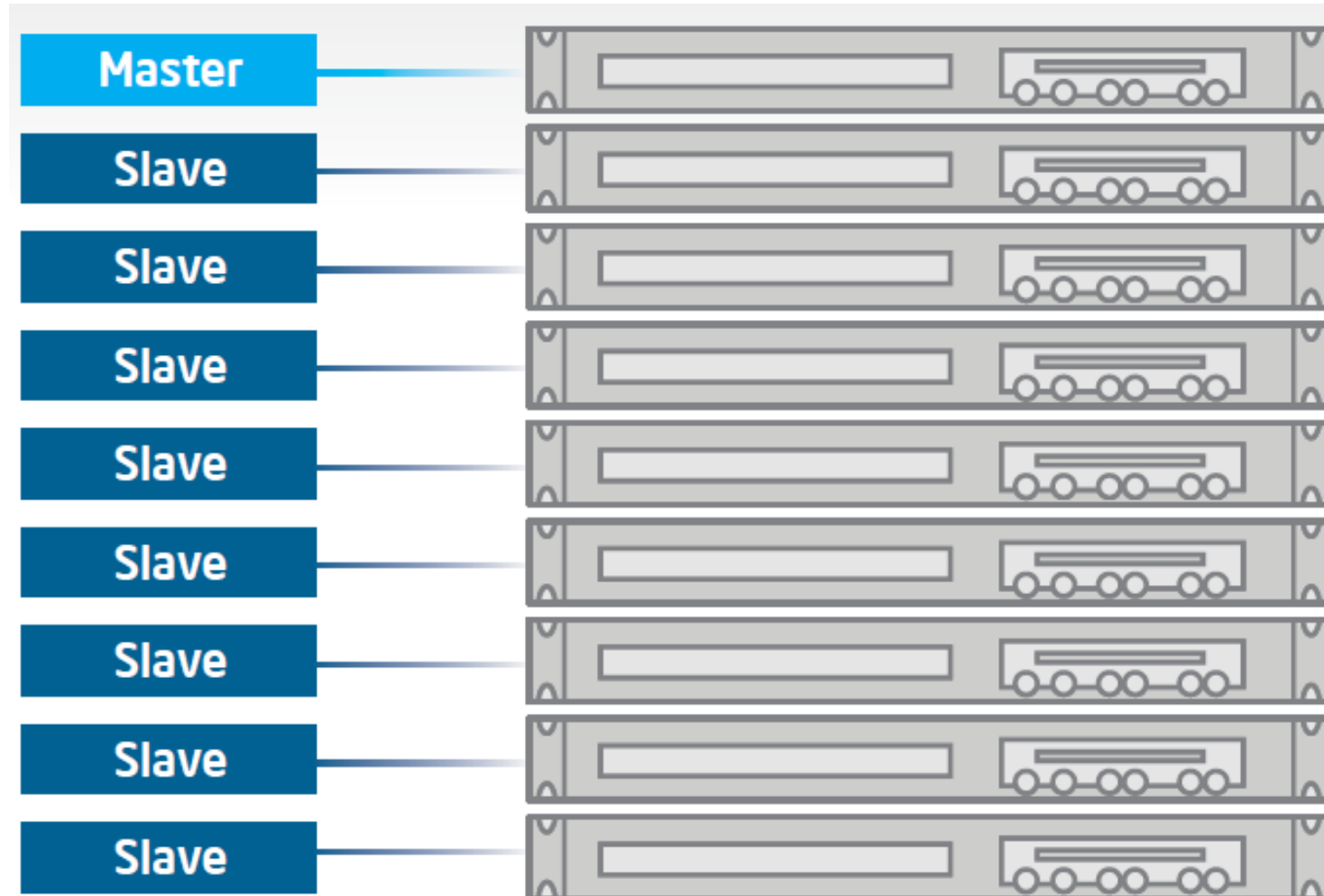
Big Data with Hadoop Architecture

Process Flow

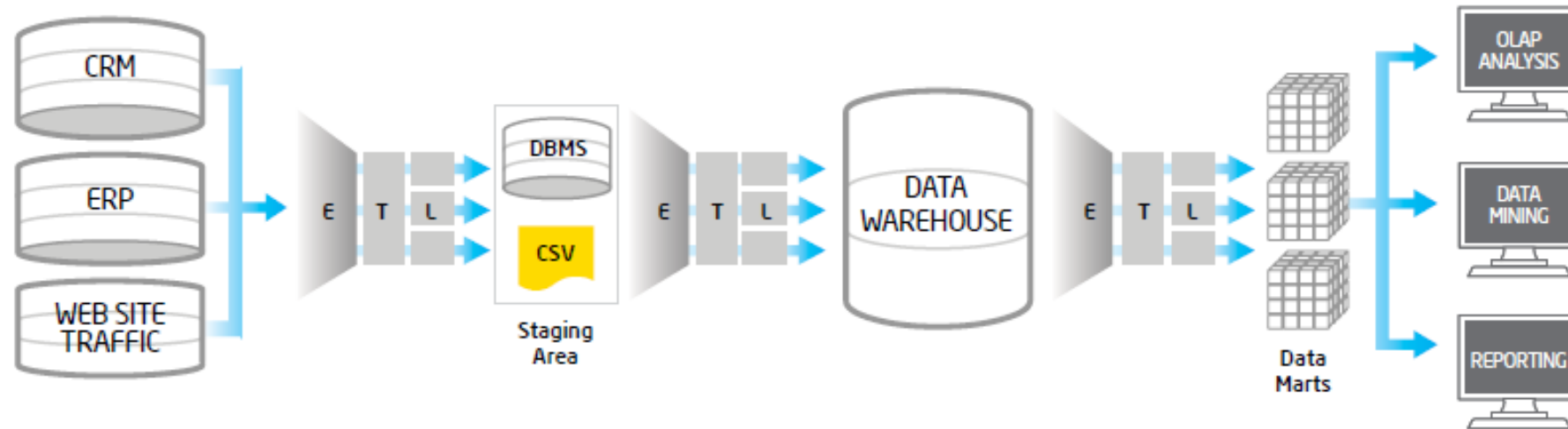


Big Data with Hadoop Architecture

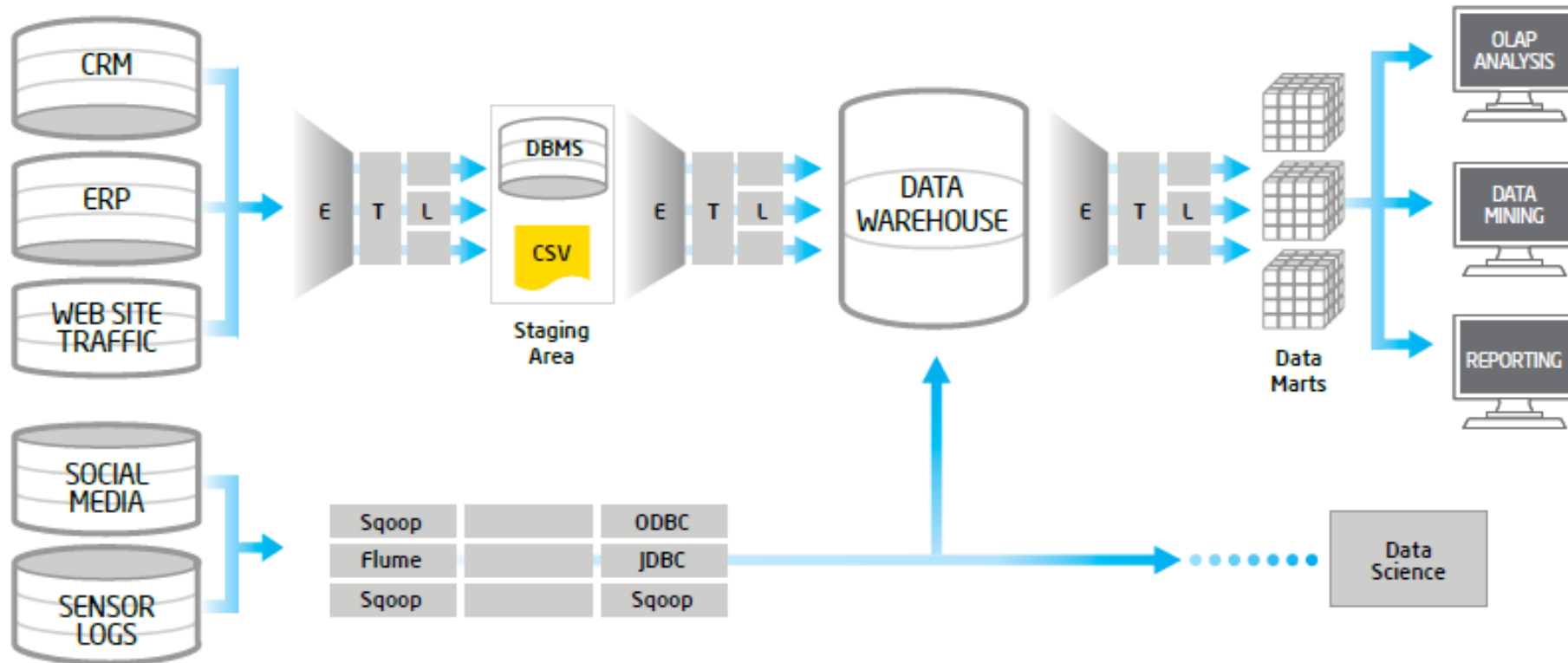
Hadoop Cluster



Traditional ETL Architecture



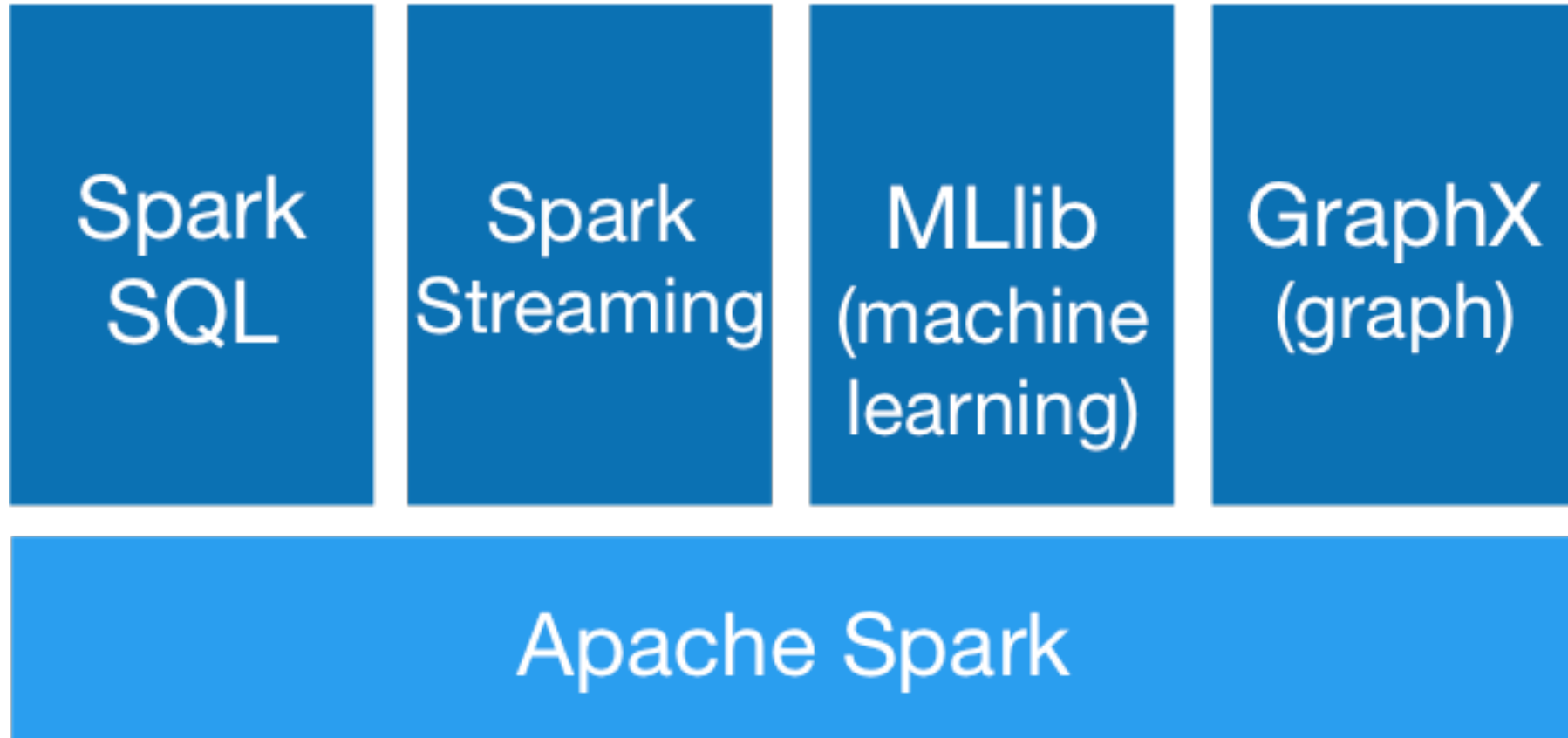
Offload ETL with Hadoop (Big Data Architecture)



Spark and Hadoop



Spark Ecosystem



OpenAI ChatGPT



API

RESEARCH

BLOG

ABOUT

ChatGPT: Optimizing Language Models for Dialogue

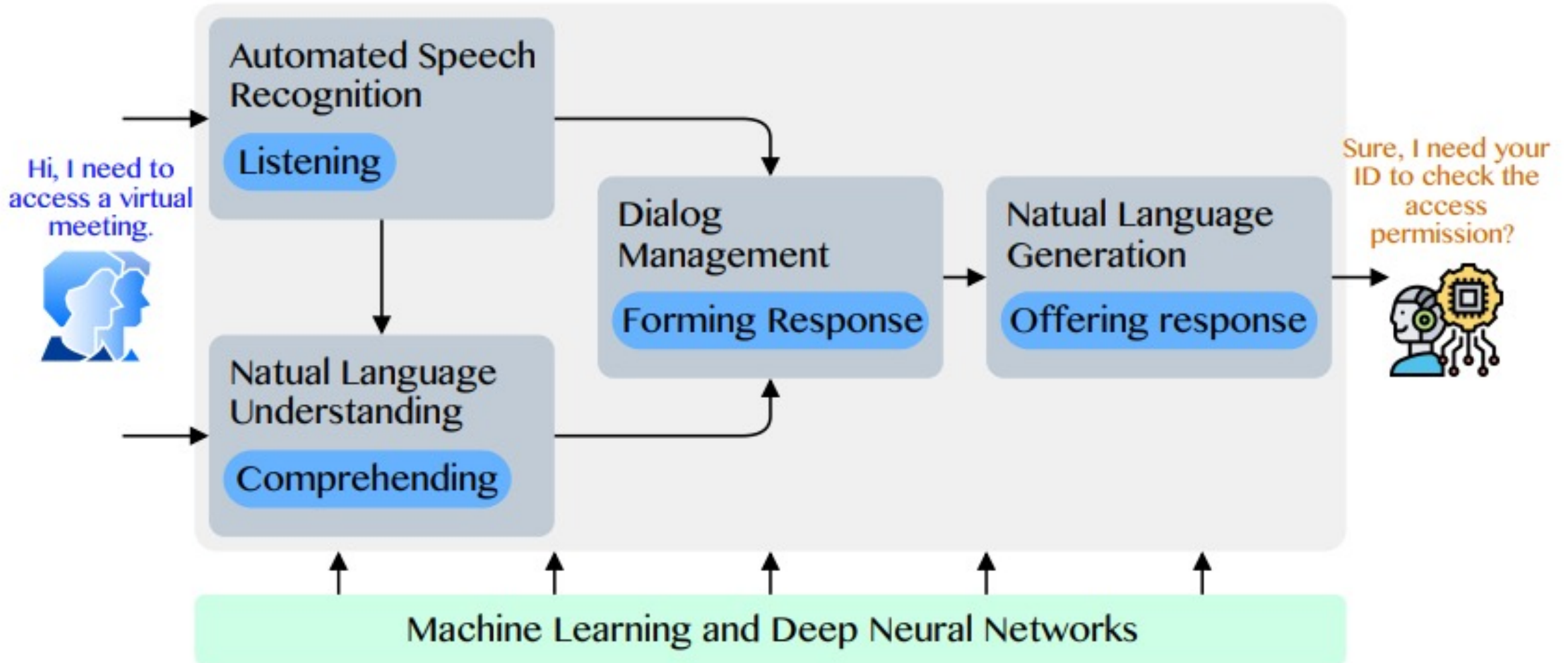
We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response.



Source: <https://openai.com/blog/chatgpt/>

Conversational AI

to deliver contextual and personal experience to users



Source: Huynh-The, Thien, Quoc-Viet Pham, Xuan-Quy Pham, Thanh Thi Nguyen, Zhu Han, and Dong-Seong Kim (2022). "Artificial Intelligence for the Metaverse: A Survey." arXiv preprint arXiv:2202.10336.

OpenAI ChatGPT

ChatGPT



Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021



Free Research Preview: ChatGPT is optimized for dialogue. Our goal is to make AI systems more natural to interact with, and your feedback will help us improve our systems and make them safer.

Source: <https://chat.openai.com/chat>

OpenAI ChatGPT



What is the future of AI?



It is difficult to predict the exact future of AI, but many experts believe that it will continue to play an increasingly important role in our lives. Some potential future developments in the field of AI include:



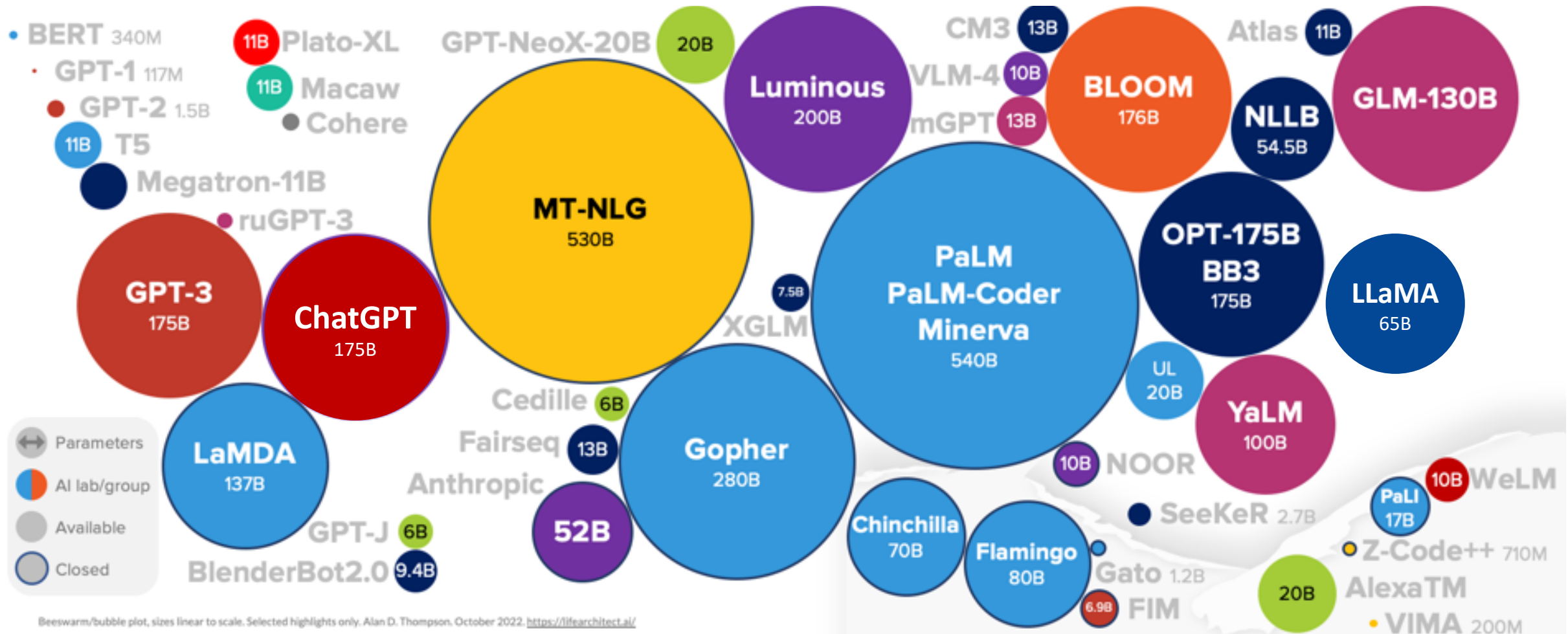
- Improved natural language processing, allowing AI systems to better understand and respond to human speech and writing.
- Enhanced ability to learn and adapt, allowing AI systems to become more autonomous and capable of handling a wider range of tasks.
- Greater integration with the internet of things (IoT), allowing AI systems to control and interact with a wide range of connected devices in the home and workplace.
- Continued advances in machine learning, allowing AI systems to make more accurate predictions and decisions based on data.
- Increased use of AI in healthcare, transportation, and other industries, leading to more efficient and effective processes and services.
- Continued development of ethical and regulatory frameworks to ensure that AI is developed and used responsibly.

Ultimately, the future of AI will depend on how it is developed and used by individuals, organizations, and society as a whole.

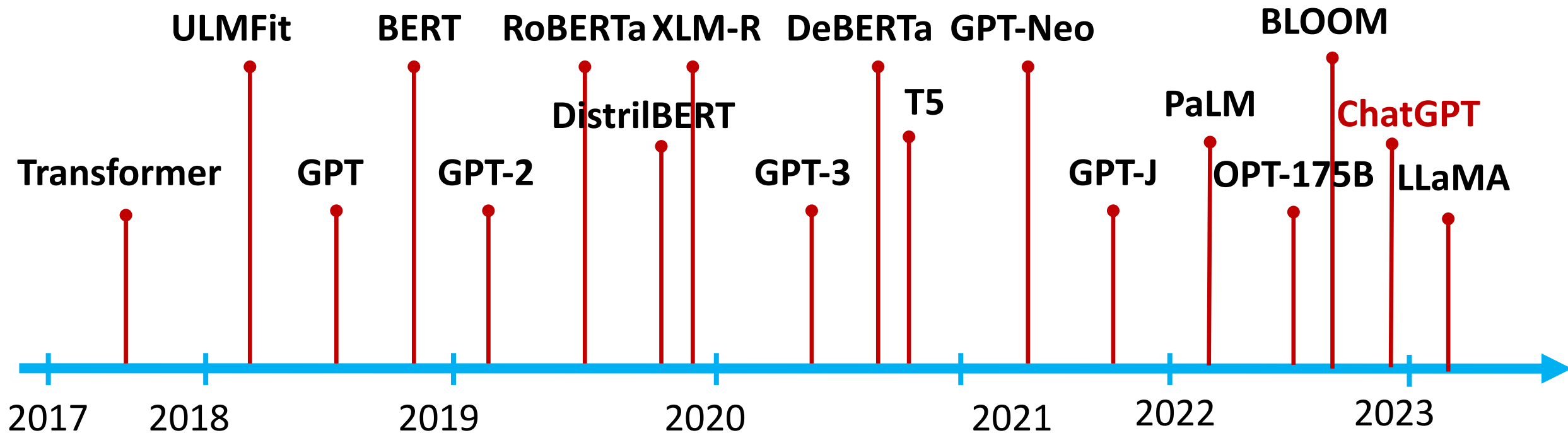
Source: <https://chat.openai.com/chat>

Large Language Models (LLM)

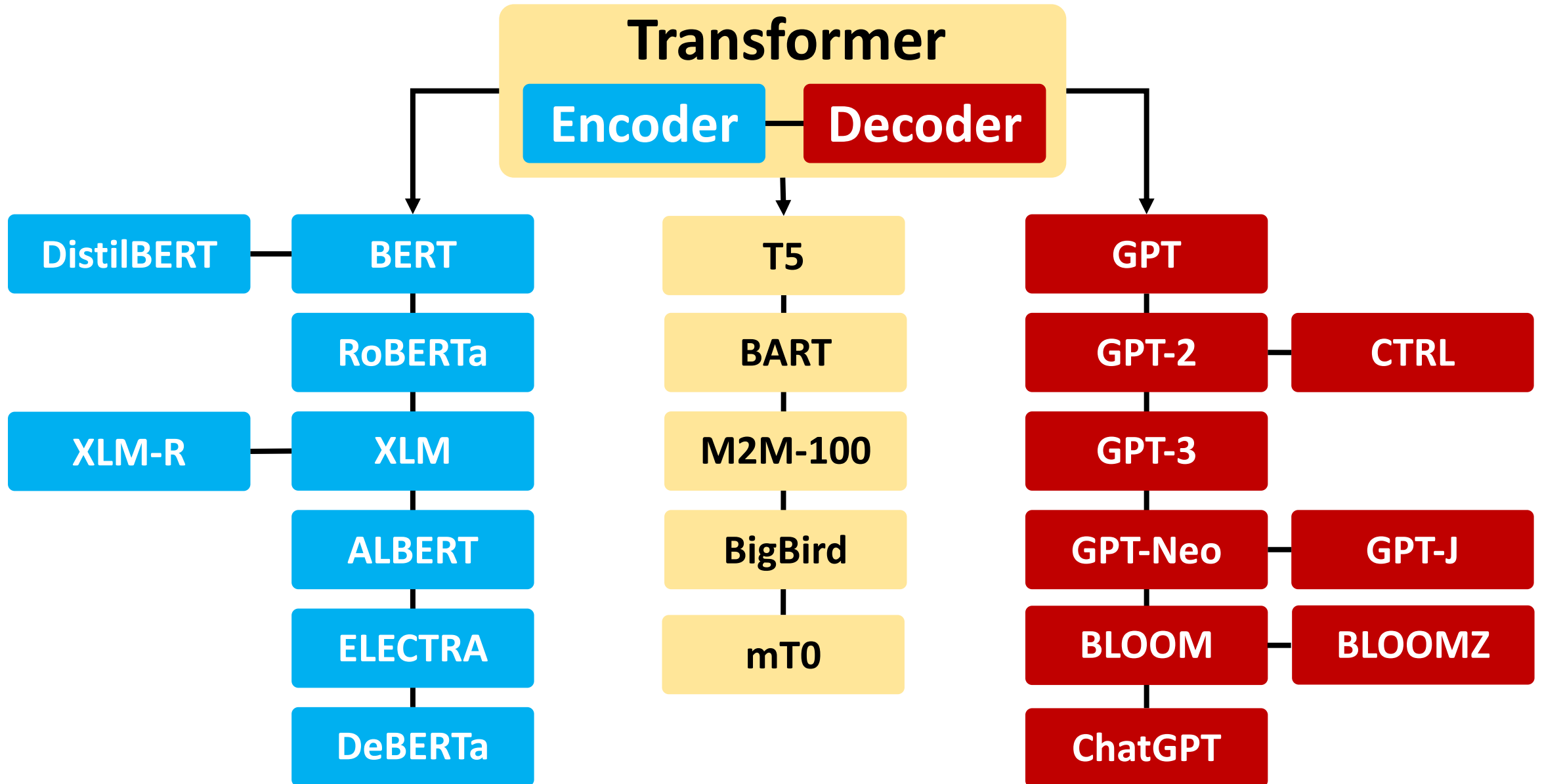
(GPT-3, ChatGPT, PaLM, BLOOM, OPT-175B, LLaMA)



The Transformers Timeline

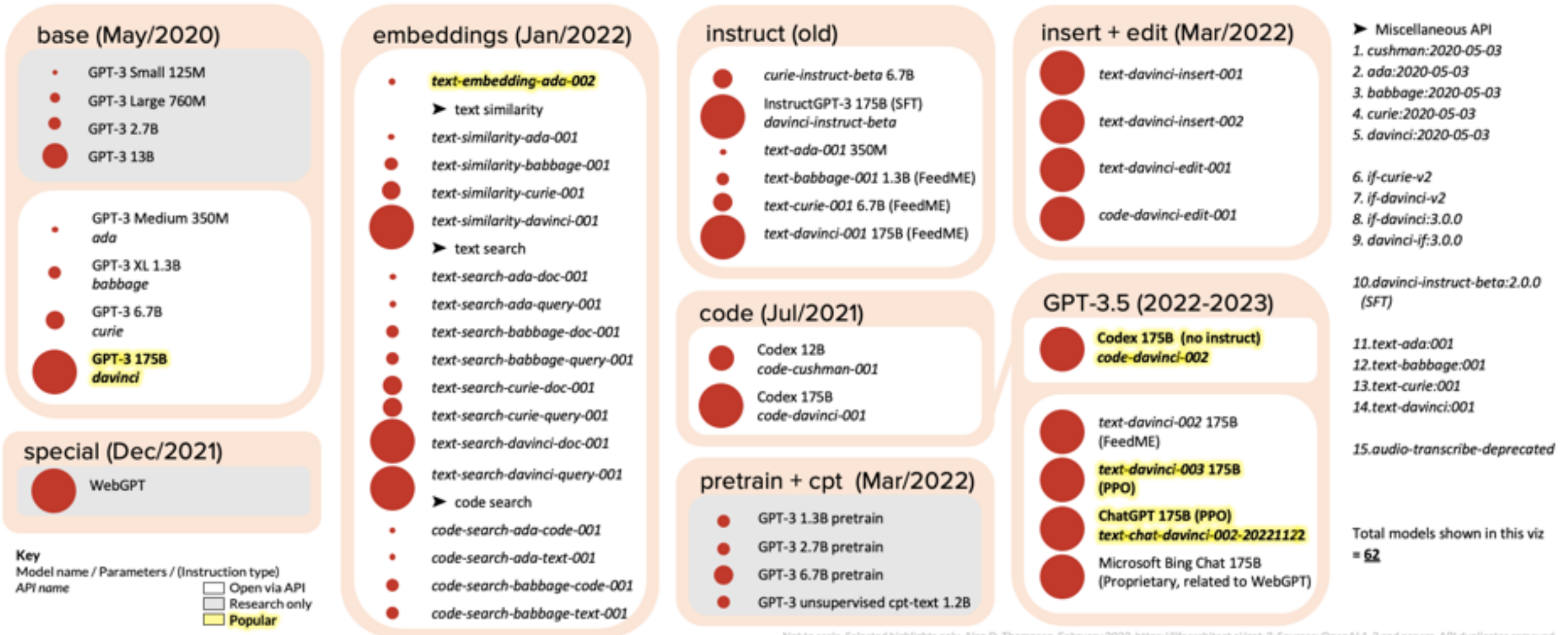


Transformer Models



ChatGPT and GPT-3 Family

(GPT-3, InstructGPT, GPT-3.5, ChatGPT)



Not to scale. Selected highlights only. Alan D. Thompson, February 2023. <https://lifearchitect.ai/gpt-3/> Sources: OpenAI 1, 2 and papers, API [duplicates](#) removed.

ChatGPT: Optimizing Language Models for Dialogue

Step 1

Collect demonstration data and train a supervised policy.

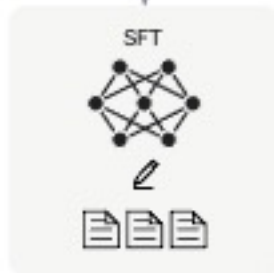
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



This data is used to fine-tune GPT-3.5 with supervised learning.



Step 2

Collect comparison data and train a reward model.

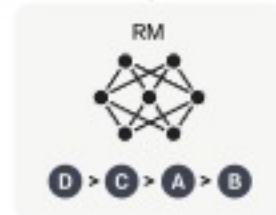
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.



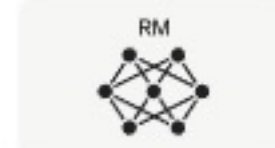
The PPO model is initialized from the supervised policy.



The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



Training language models to follow instructions with human feedback

InstructGPT and GPT 3.5

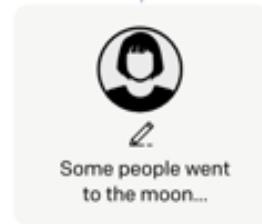
Step 1

Collect demonstration data, and train a supervised policy.

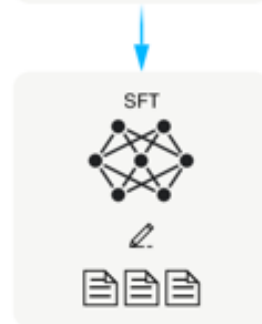
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

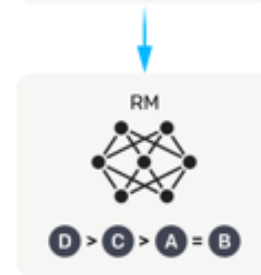
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

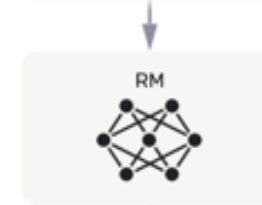
A new prompt is sampled from the dataset.



The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.

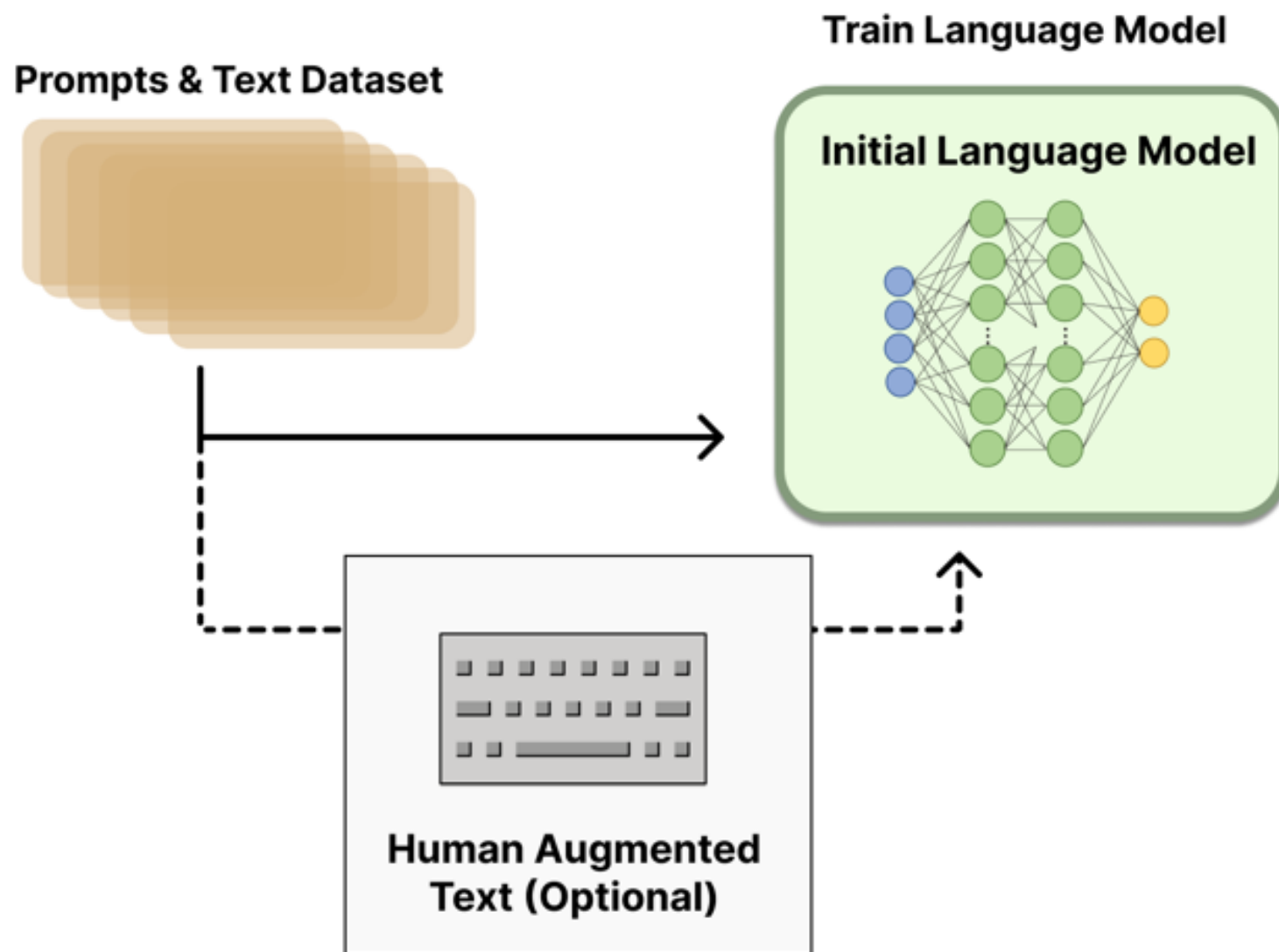


Reinforcement Learning from Human Feedback (RLHF)

- 1. Pretraining a Language Model (LM)**
- 2. Gathering Data and Training a Reward Model**
- 3. Fine-tuning the LM with Reinforcement Learning**

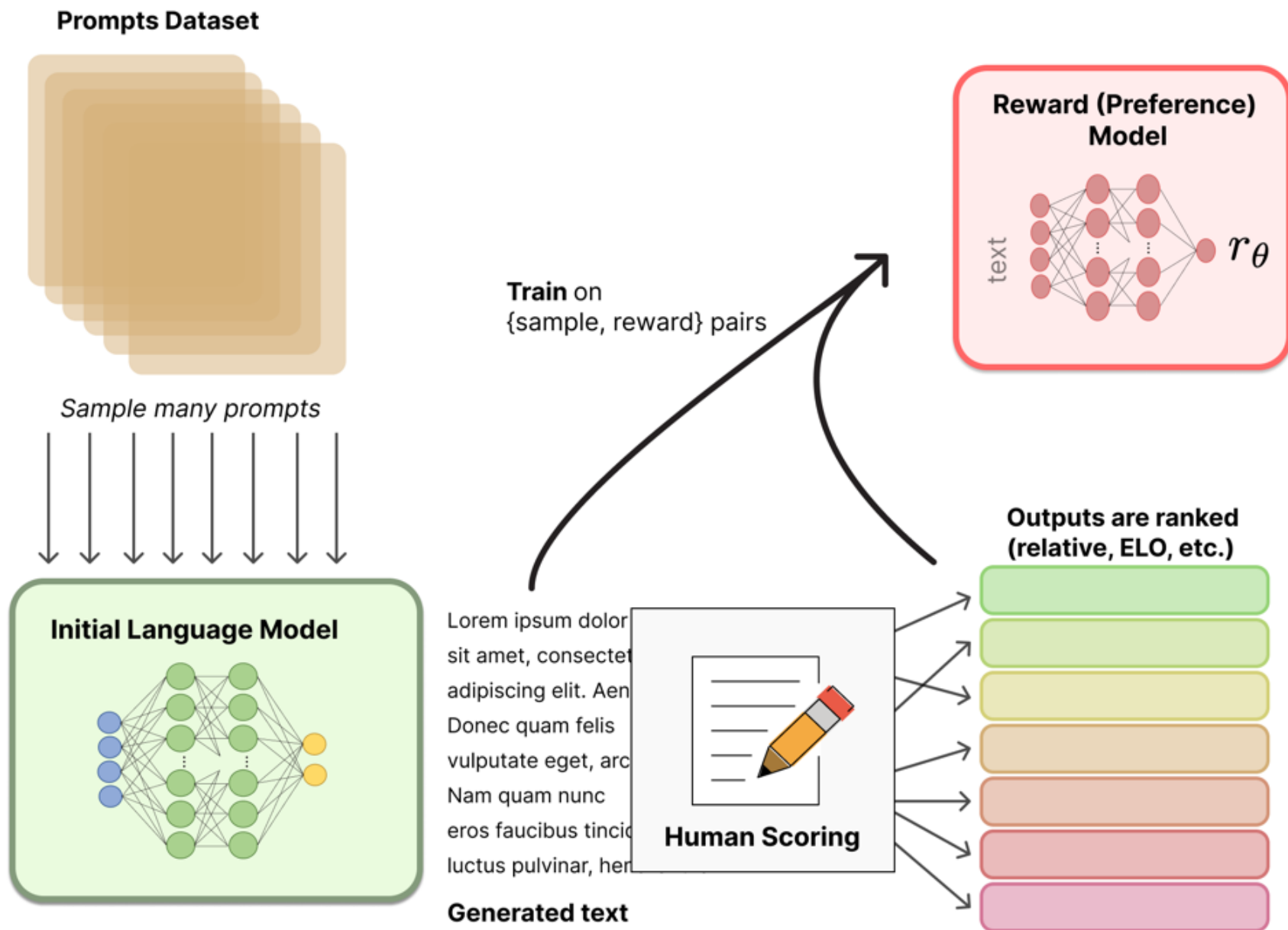
Reinforcement Learning from Human Feedback (RLHF)

Step 1. Pretraining a Language Model (LM)



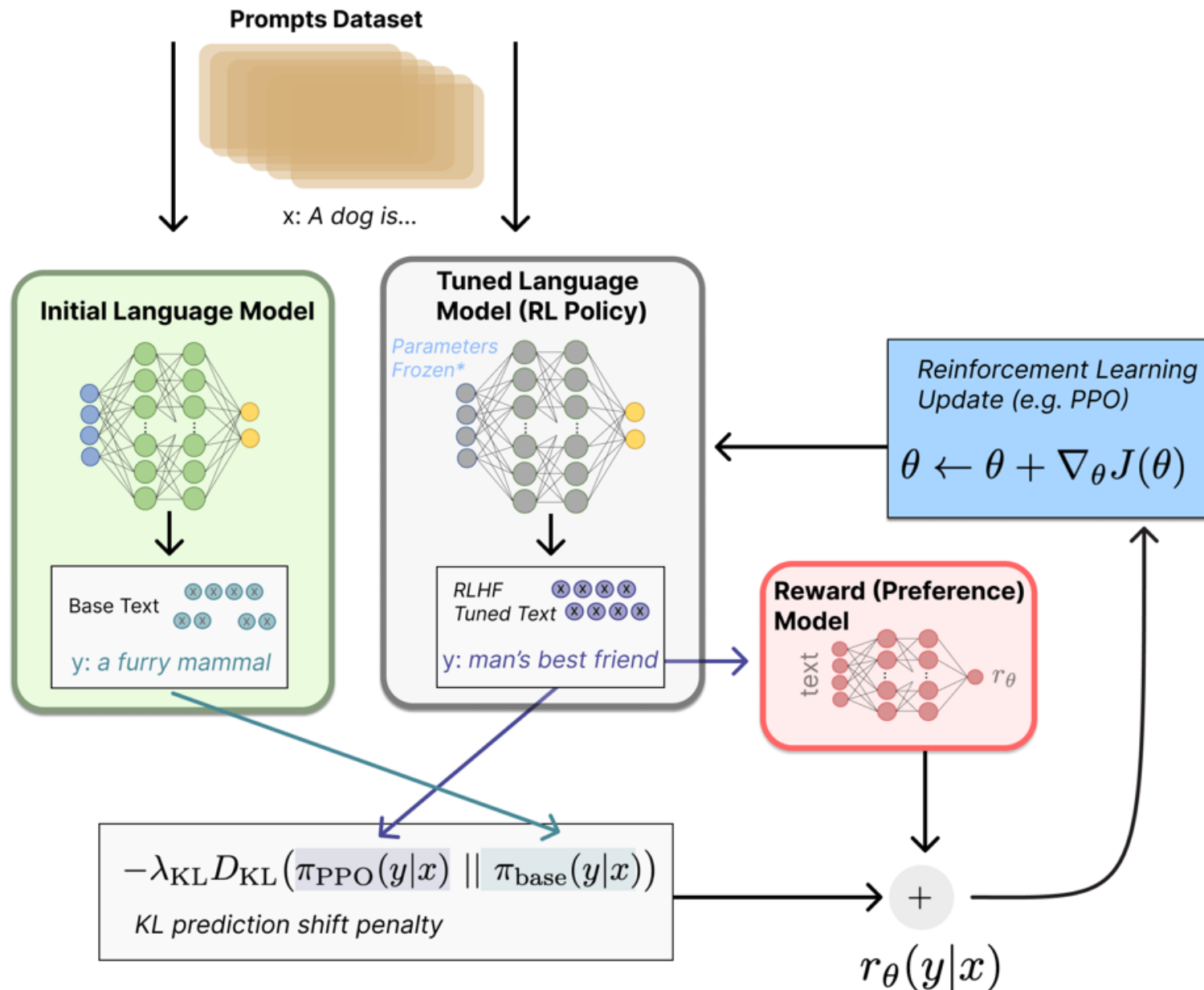
Reinforcement Learning from Human Feedback (RLHF)

Step 2. Gathering Data and Training a Reward Model



Reinforcement Learning from Human Feedback (RLHF)

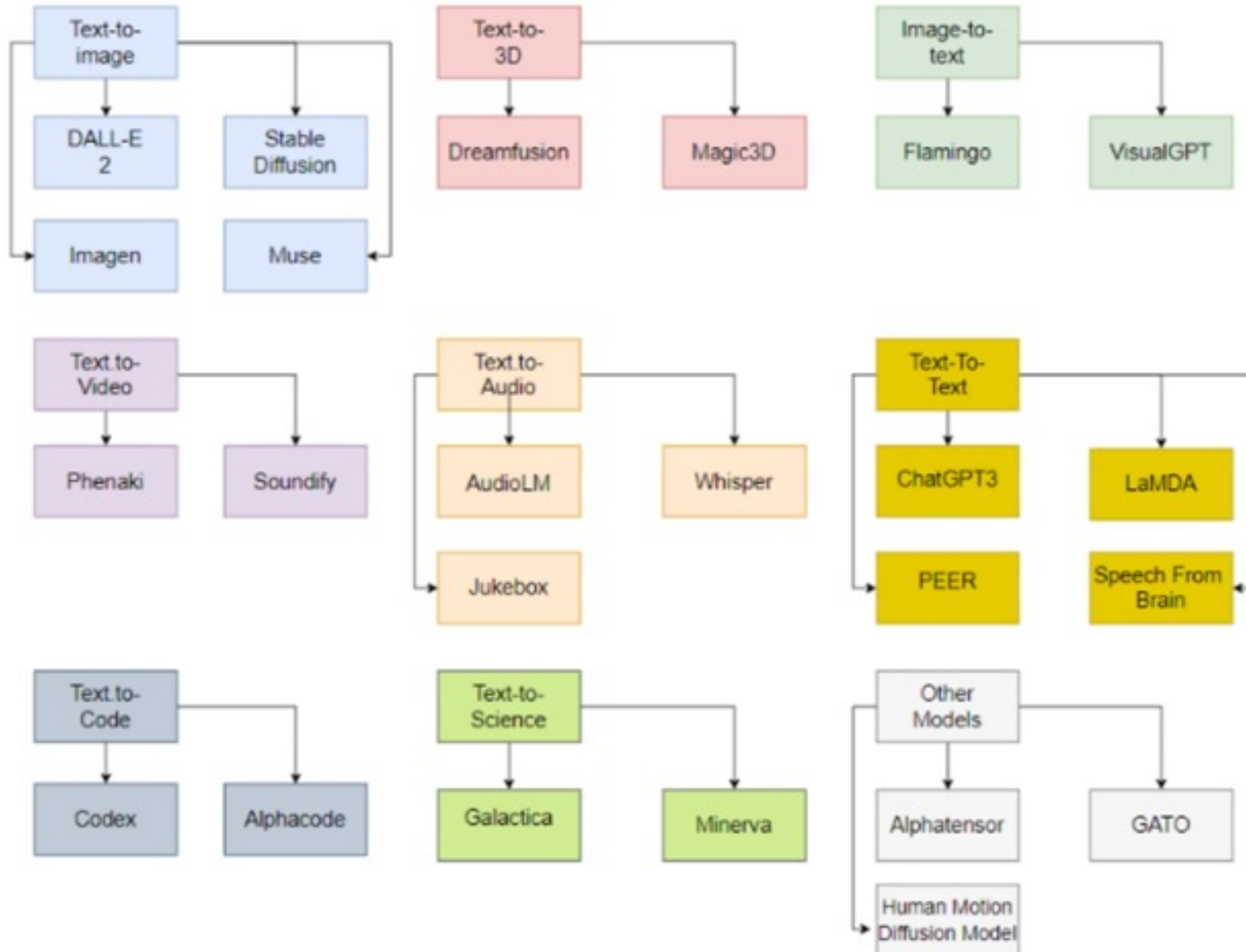
Step 3. Fine-tuning the LM with Reinforcement Learning



Generative AI

Gen AI

Generative AI Models



**ChatGPT
is not
all you need**

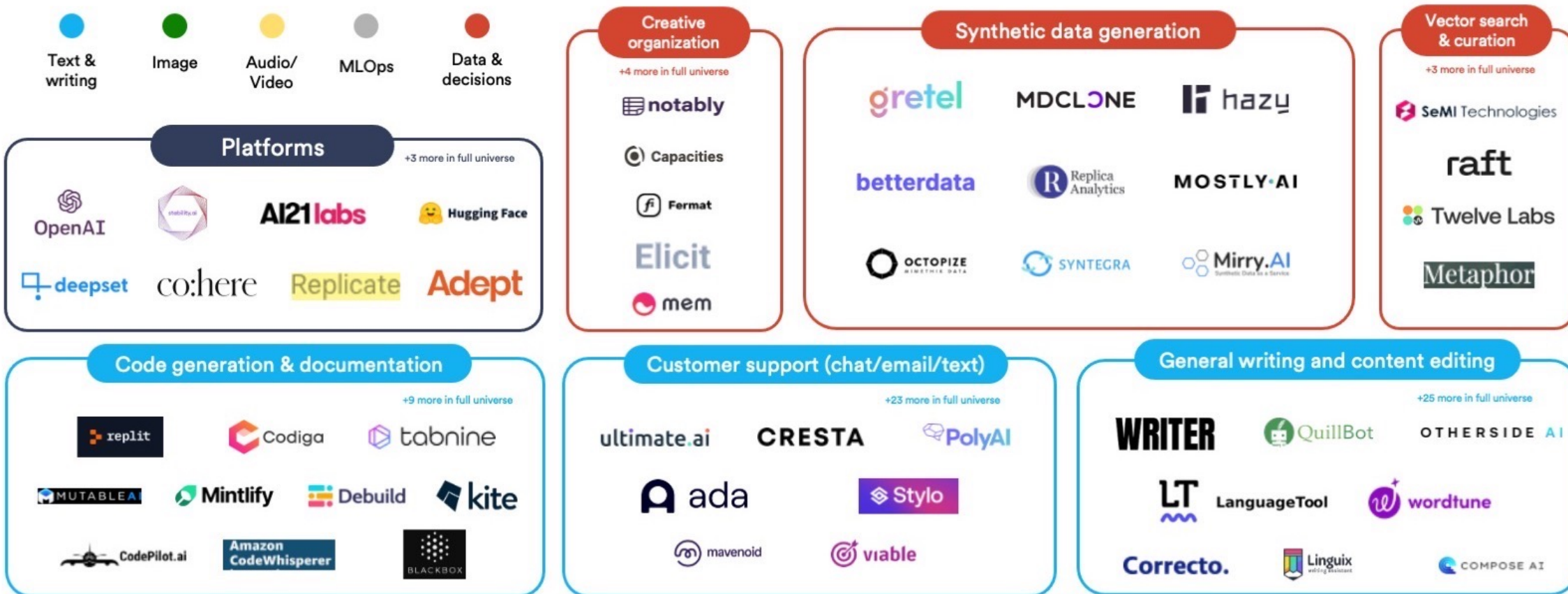
**Attention
is
all you need**

Generative AI

BASE10 TREND MAP: GENERATIVE AI

Companies are grouped based on medium produced and segmented by use case within each medium. Companies that offer products across segments are grouped in the segment of the core product offering.

Base¹⁰



Generative AI

Marketing & sales copy



Text & data summarization

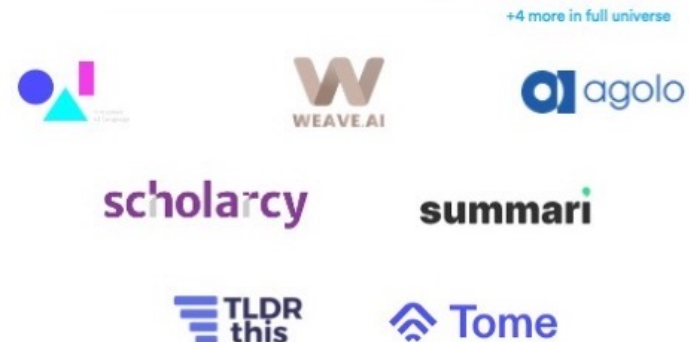


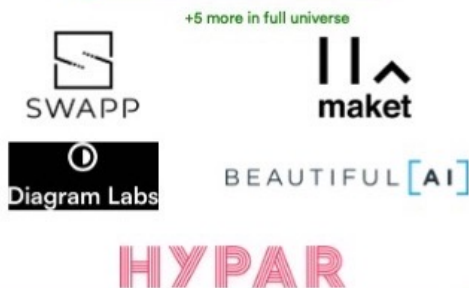
Image editing



Ad collateral



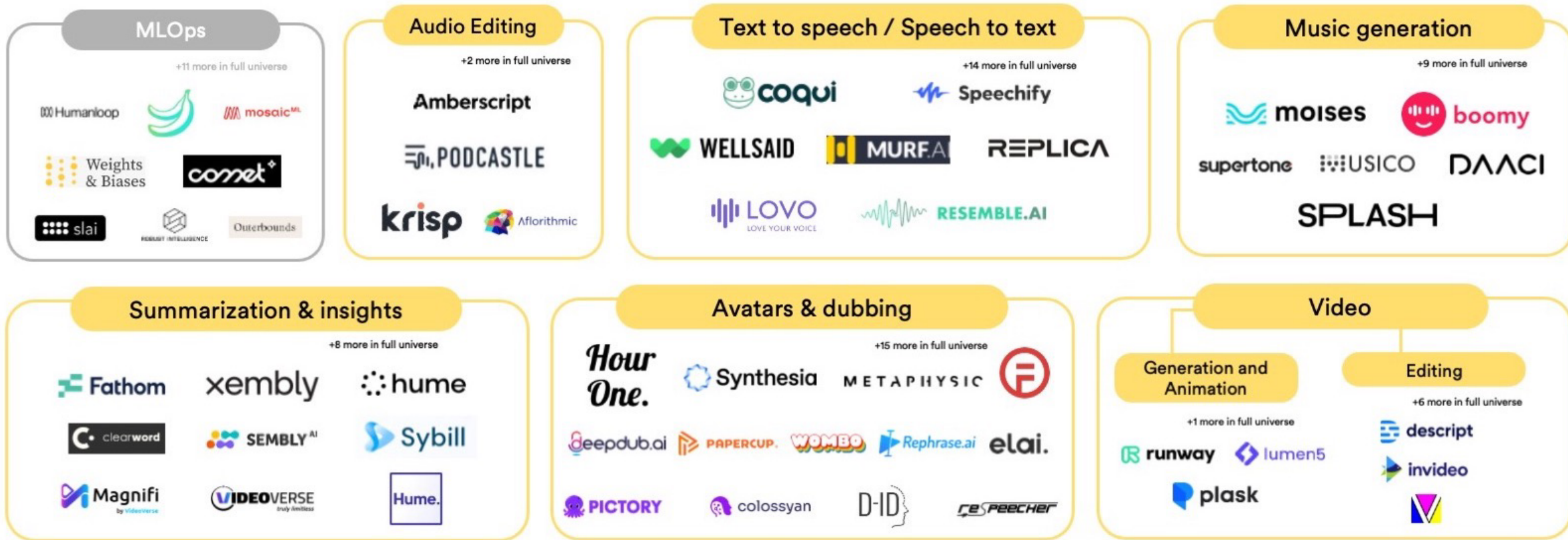
Design



Text to image



Generative AI



DALL·E 2

Create original, realistic images and art from a text description.
It can combine concepts, attributes, and styles.

TEXT DESCRIPTION

An astronaut Teddy bears A
bowl of soup

riding a horse lounging in a
tropical resort in space playing
basketball with cats in space

in a photorealistic style in the
style of Andy Warhol as a pencil
drawing




DALL·E 2



<https://openai.com/dall-e-2/>

Stable Diffusion

 **Hugging Face** [Models](#) [Datasets](#) [Spaces](#) [Docs](#) [Solutions](#) [Pricing](#) ☰


Spaces: [stabilityai/stable-diffusion](#) 👍 like 1.89k 🟢 Running

[App](#) [Files](#) [Community](#) 241 ⋮ [Linked Models](#)

🤖 Stable Diffusion Demo

Stable Diffusion is a state of the art text-to-image model that generates images from text.
For faster generation and forthcoming API access you can try [DreamStudio Beta](#)

[Generate Image](#)



The image shows two side-by-side generated images. The left image depicts a silver robot with a small white and black insect-like headpiece, holding a large, multi-layered burger. The right image shows a large, metallic, fly-like insect with transparent wings on a table with scattered red and green food items.

<https://huggingface.co/spaces/stabilityai/stable-diffusion>

Stable Diffusion Colab

woctezuma / [stable-diffusion-colab](#) Public

Notifications Fork 7 Star 31

Code Issues Pull requests Actions Projects Wiki Security Insights

main 1 branch 0 tags

Go to file Code

About

Colab notebook to run Stable Diffusion.

[github.com/CompVis/stable-diffusion](#)

- deep-learning
- colab
- image-generation
- text-to-image
- diffusion
- text2image
- colaboratory
- google-colab
- colab-notebook
- google-colaboratory
- google-colab-notebook
- text-to-image-synthesis
- huggingface
- diffusion-models
- text-to-image-generation
- latent-diffusion
- stable-diffusion
- huggingface-diffusers
- diffusers
- stable-diffusion-diffusers

Readme
MIT license
31 stars
2 watching

woctezuma	README: add a reference for sampler schedules	37bc02d 24 days ago	🕒 18 commits
LICENSE	Initial commit		27 days ago
README.md	README: add a reference for sampler schedules		24 days ago
stable_diffusion.ipynb	Allow to choose the scheduler		25 days ago

☰ README.md

Stable-Diffusion-Colab

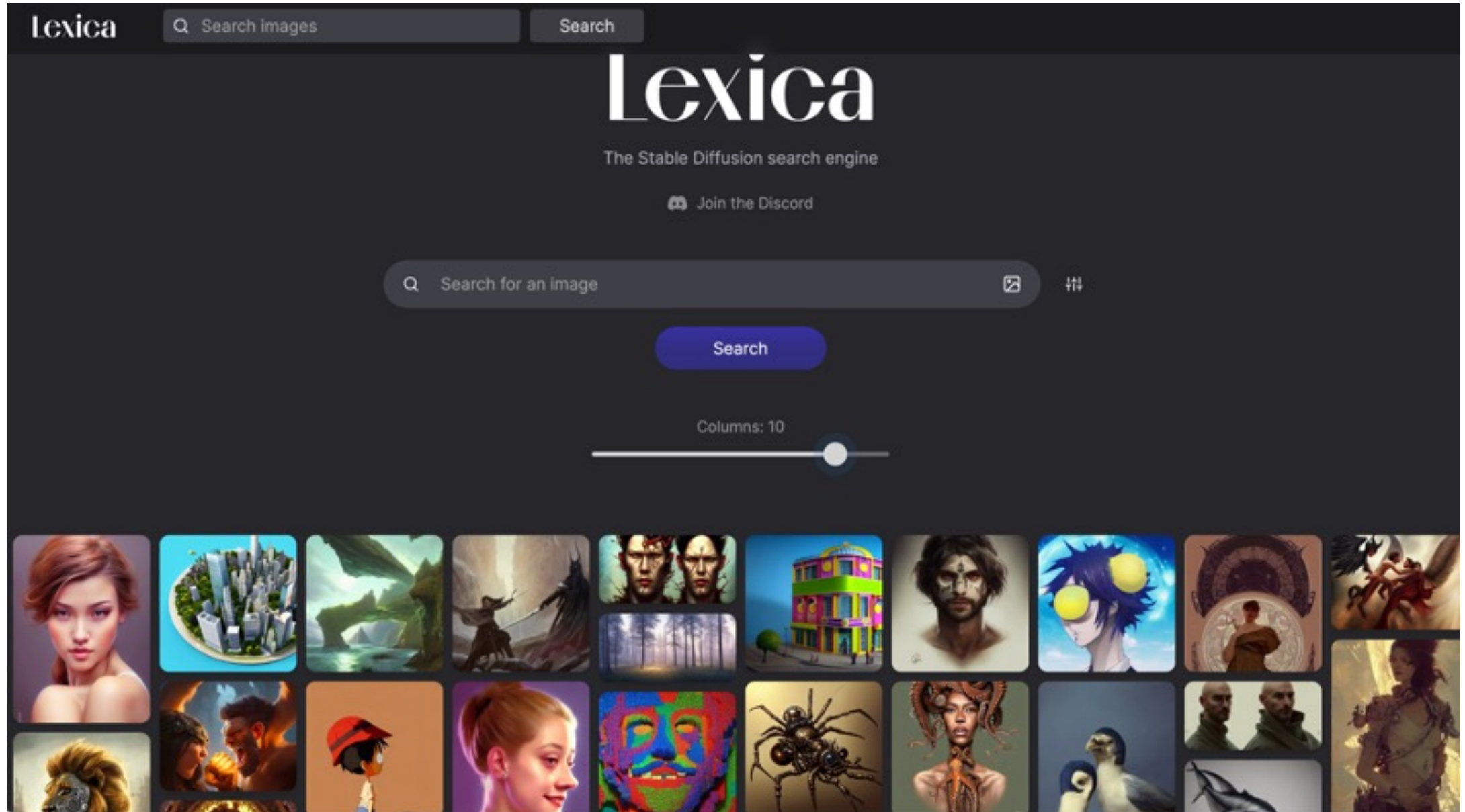
The goal of this repository is to provide a Colab notebook to run the text-to-image "Stable Diffusion" model [1].

Usage

- Run `stable_diffusion.ipynb` . Open in Colab

<https://github.com/woctezuma/stable-diffusion-colab>

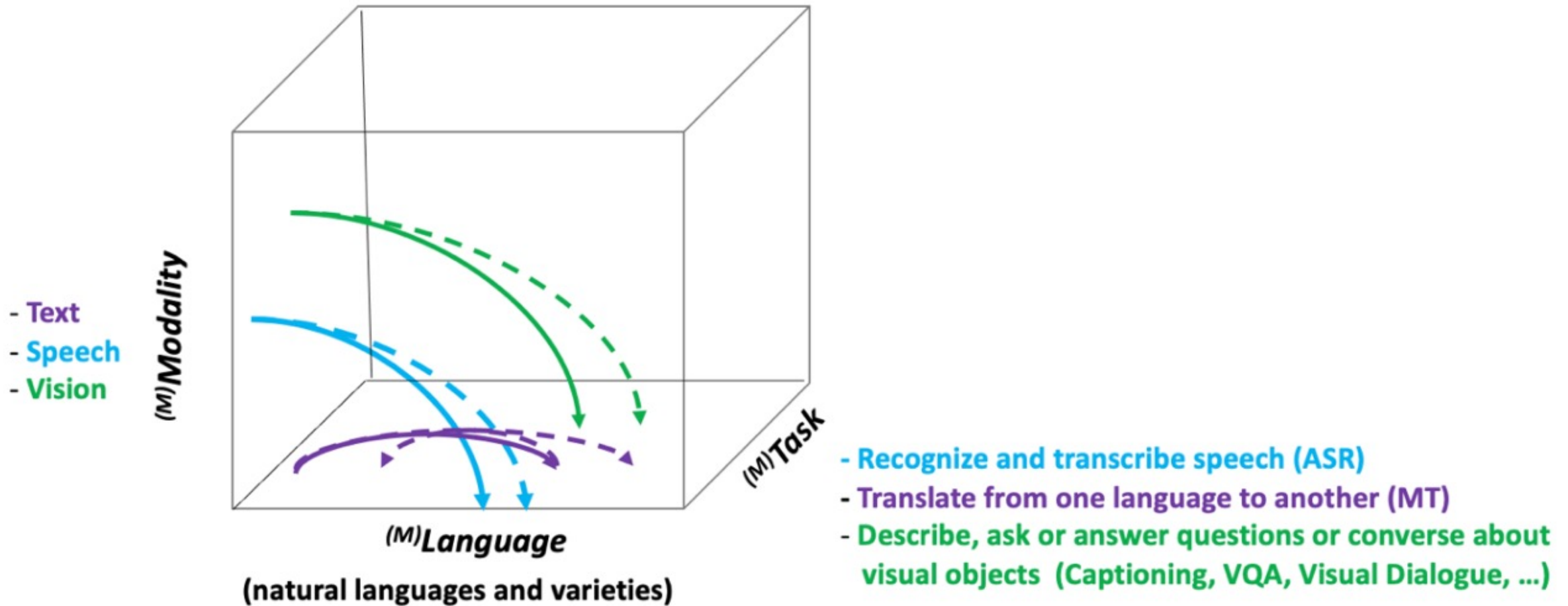
Lexica Art: Search Stable Diffusion images and prompts



<https://lexica.art/>

NLG from a Multilingual, Multimodal and Multi-task perspective

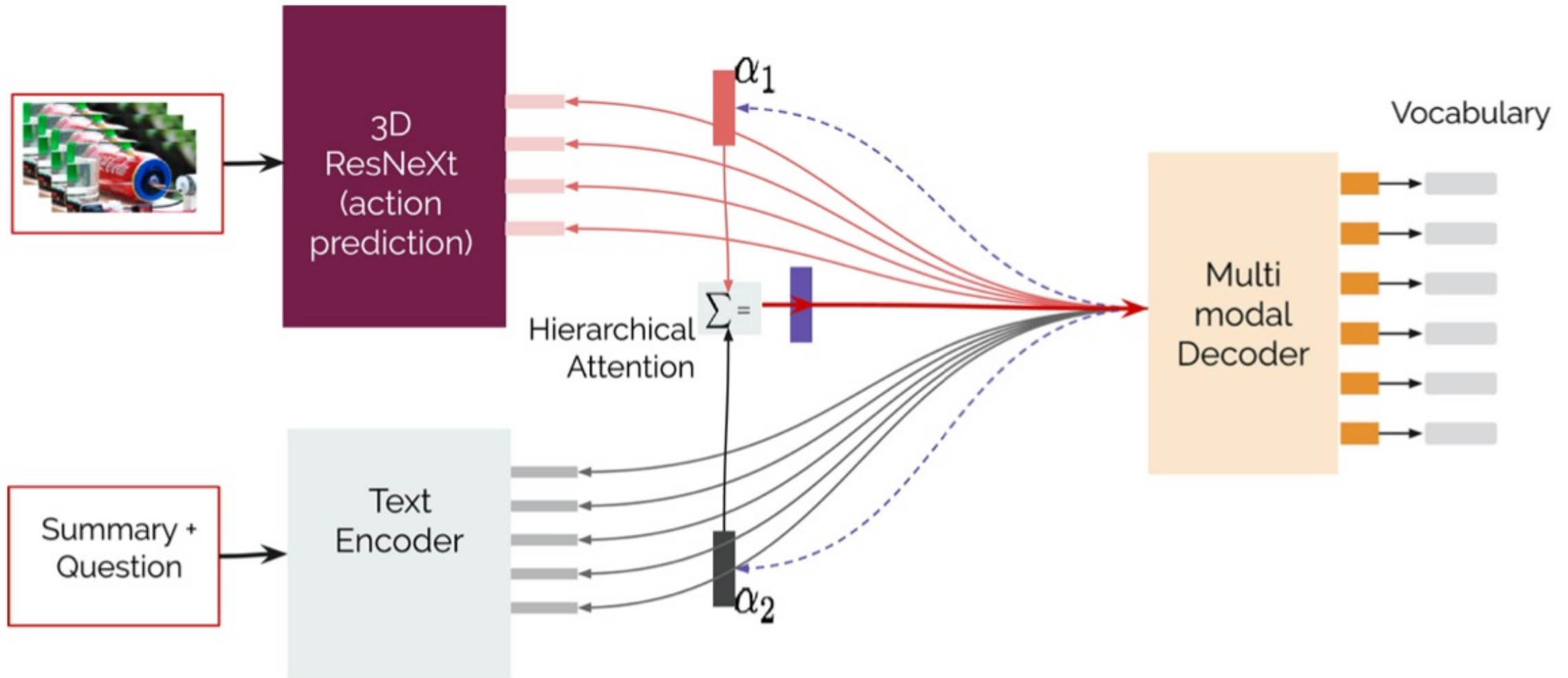
Multi³(Natural Language) Generation



Source: Erdem, Erkut, Menekse Kuyu, Semih Yagcioglu, Anette Frank, Letitia Parcalabescu, Barbara Plank, Andrii Babii et al.

"Neural Natural Language Generation: A Survey on Multilinguality, Multimodality, Controllability and Learning." *Journal of Artificial Intelligence Research* 73 (2022): 1131-1207.

Text-and-Video Dialog Generation Models with Hierarchical Attention



Source: Erdem, Erkut, Menekse Kuyu, Semih Yagcioglu, Anette Frank, Letitia Parcalabescu, Barbara Plank, Andrii Babii et al.

"Neural Natural Language Generation: A Survey on Multilinguality, Multimodality, Controllability and Learning." Journal of Artificial Intelligence Research 73 (2022): 1131-1207.

Multimodal Few-Shot Learning with Frozen Language Models

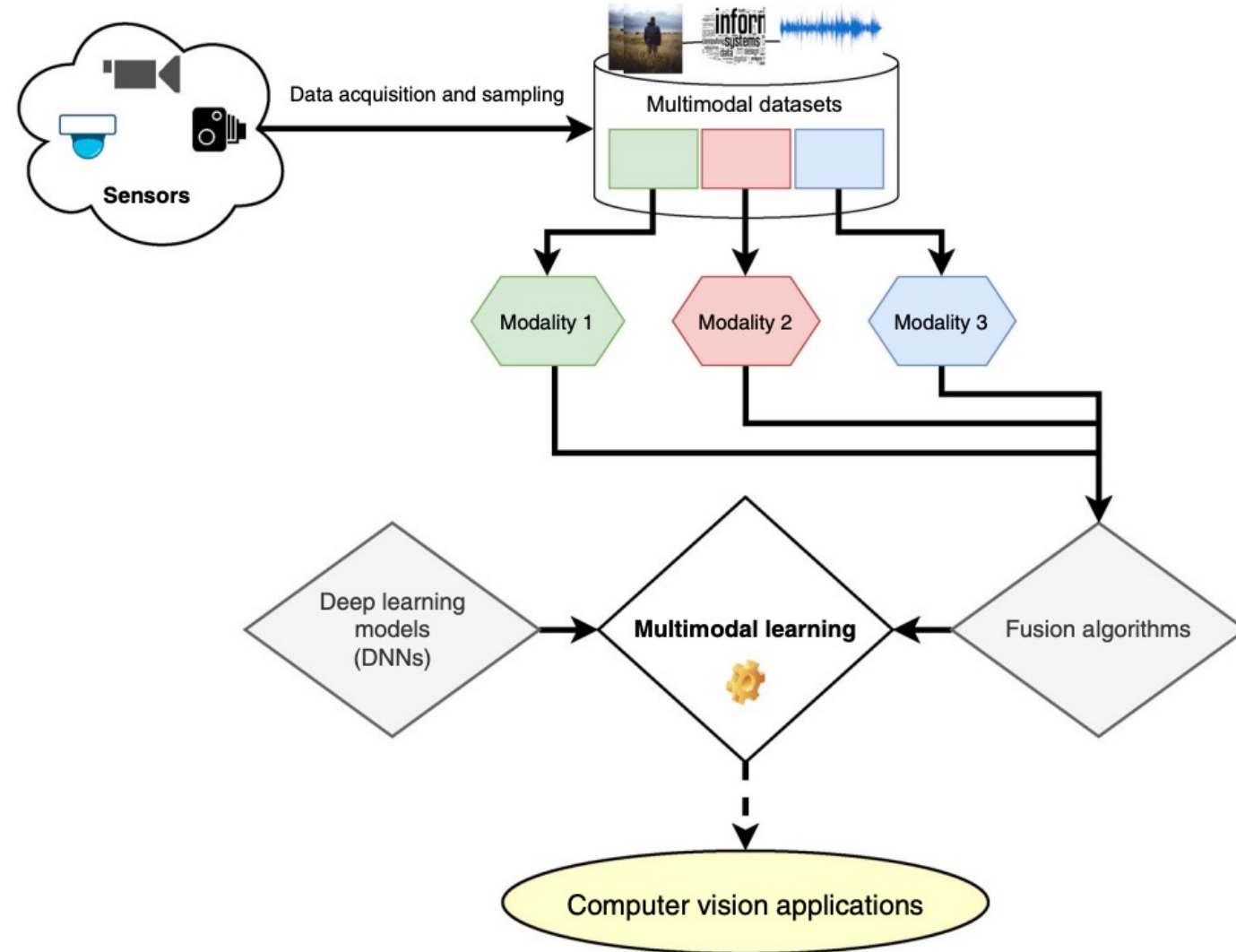


Curated samples with about five seeds required to get past well-known language model failure modes of either repeating text for the prompt or emitting text that does not pertain to the image.

These samples demonstrate the ability to generate open-ended outputs that adapt to both images and text, and to make use of facts that it has learned during language-only pre-training.

Multimodal Pipeline

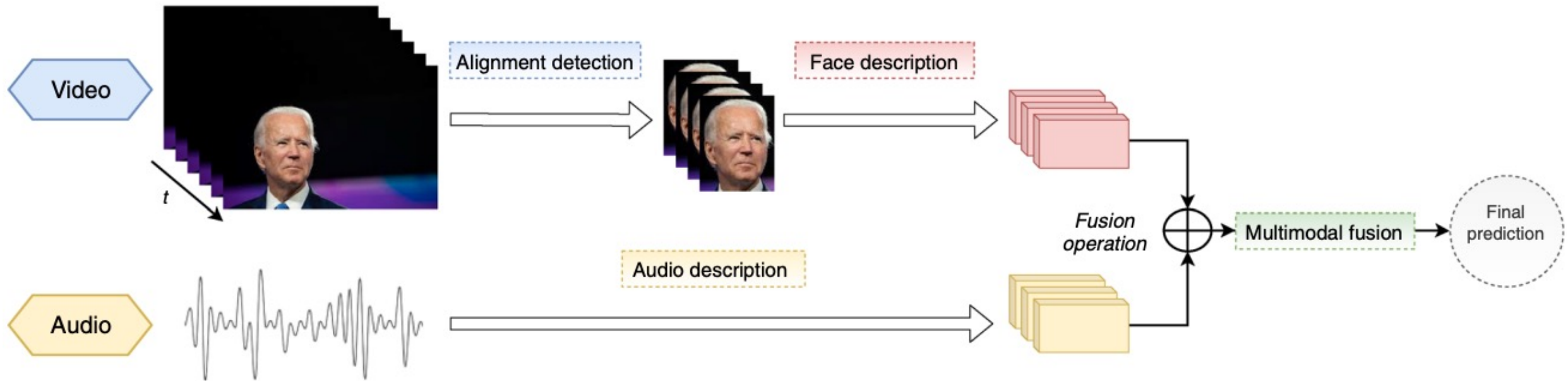
that includes three different modalities (Image, Text, Audio)



Source: Bayoudh, Khaled, Raja Knani, Fayçal Hamdaoui, and Abdellatif Mtibaa (2022).

"A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets." The Visual Computer 38, no. 8: 2939-2970.

Video and Audio Multimodal Fusion



Source: Bayoudh, Khaled, Raja Knani, Fayçal Hamdaoui, and Abdellatif Mtibaa (2022).

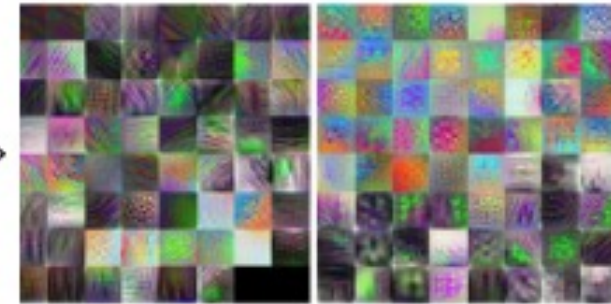
"A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets." The Visual Computer 38, no. 8: 2939-2970.

Visual and Textual Representation

Image



Visual representations (Dense)



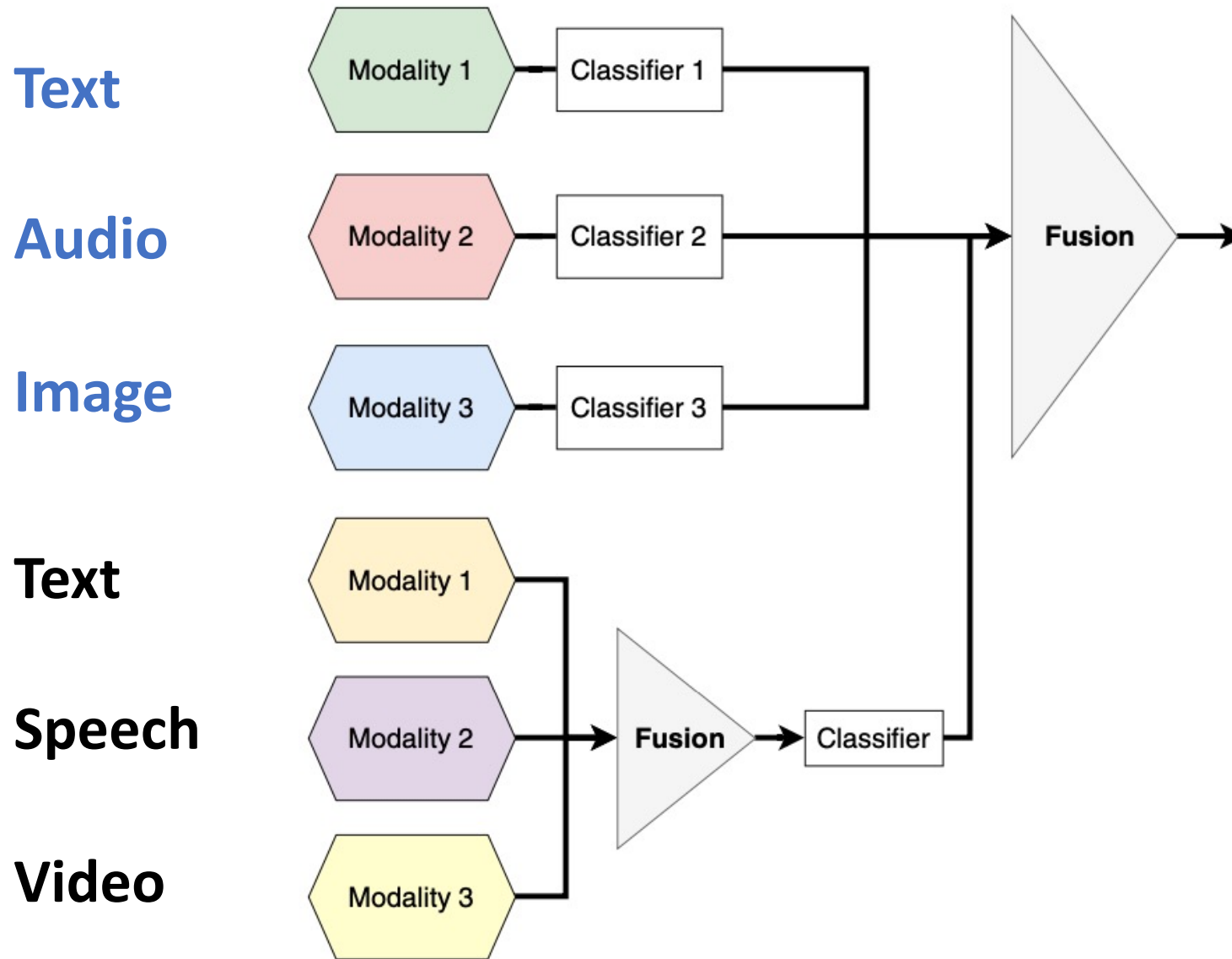
Text

This is the oldest and most important defensive work to have been built along the North African coastline by the Arab conquerors in the early days of Islam. Founded in 796, this building underwent several modifications during the medieval period. Initially, it formed a quadrilateral and then was composed of four buildings giving onto two inner courtyards.

Textual representations (Sparse)



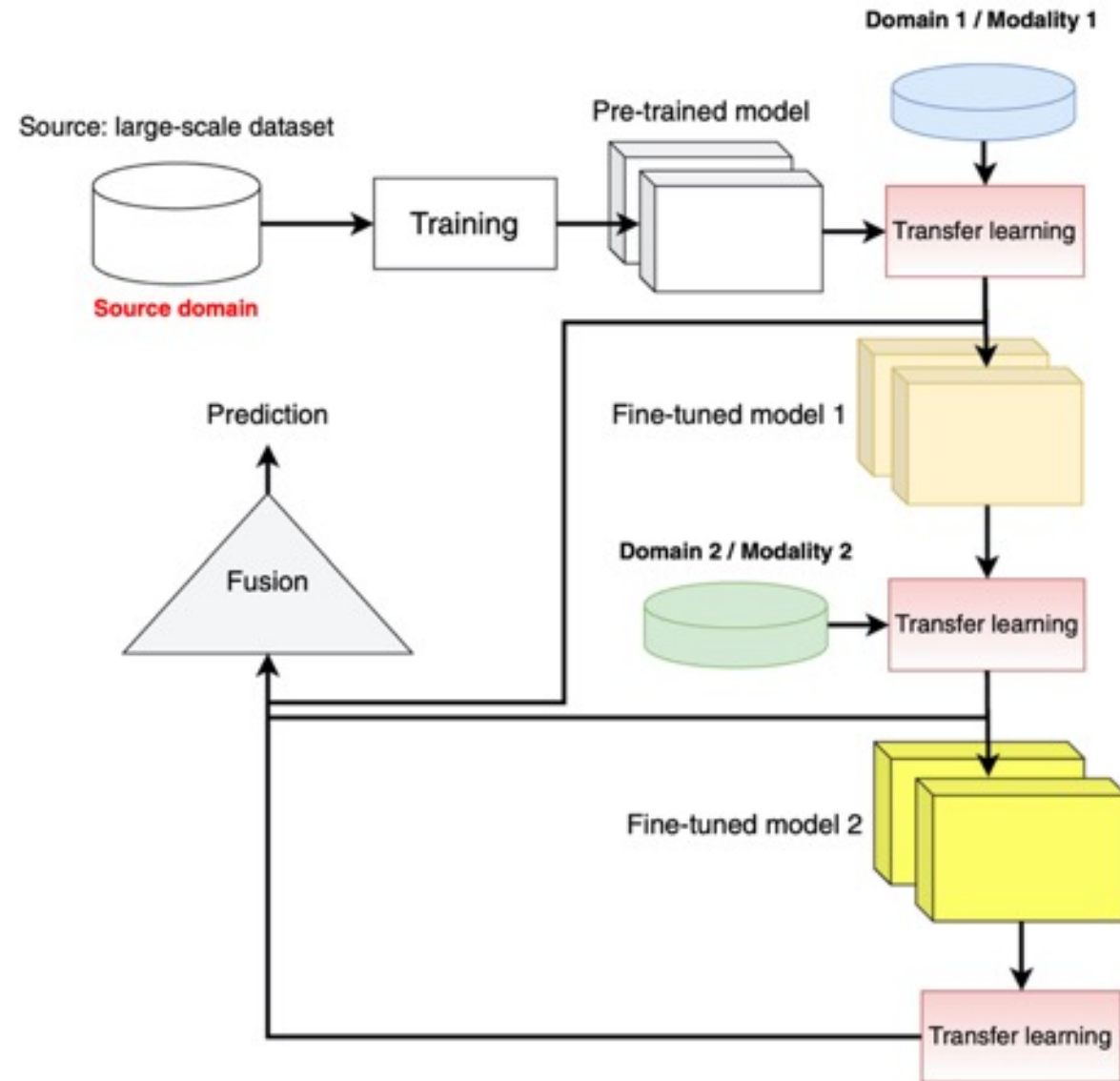
Hybrid Multimodal Data Fusion



Source: Bayouhd, Khaled, Raja Knani, Fayçal Hamdaoui, and Abdellatif Mtibaa (2022).

"A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets." The Visual Computer 38, no. 8: 2939-2970.

Multimodal Transfer Learning

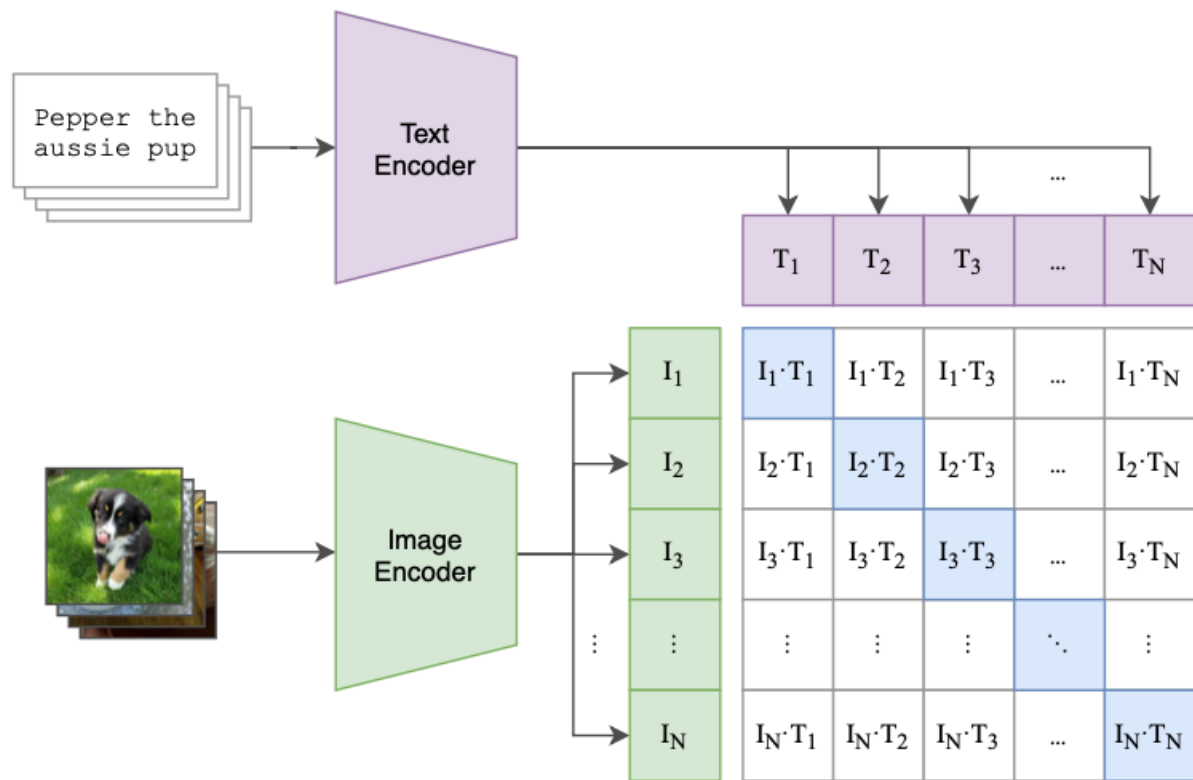


Source: Bayoudh, Khaled, Raja Knani, Fayçal Hamdaoui, and Abdellatif Mtibaa (2022).

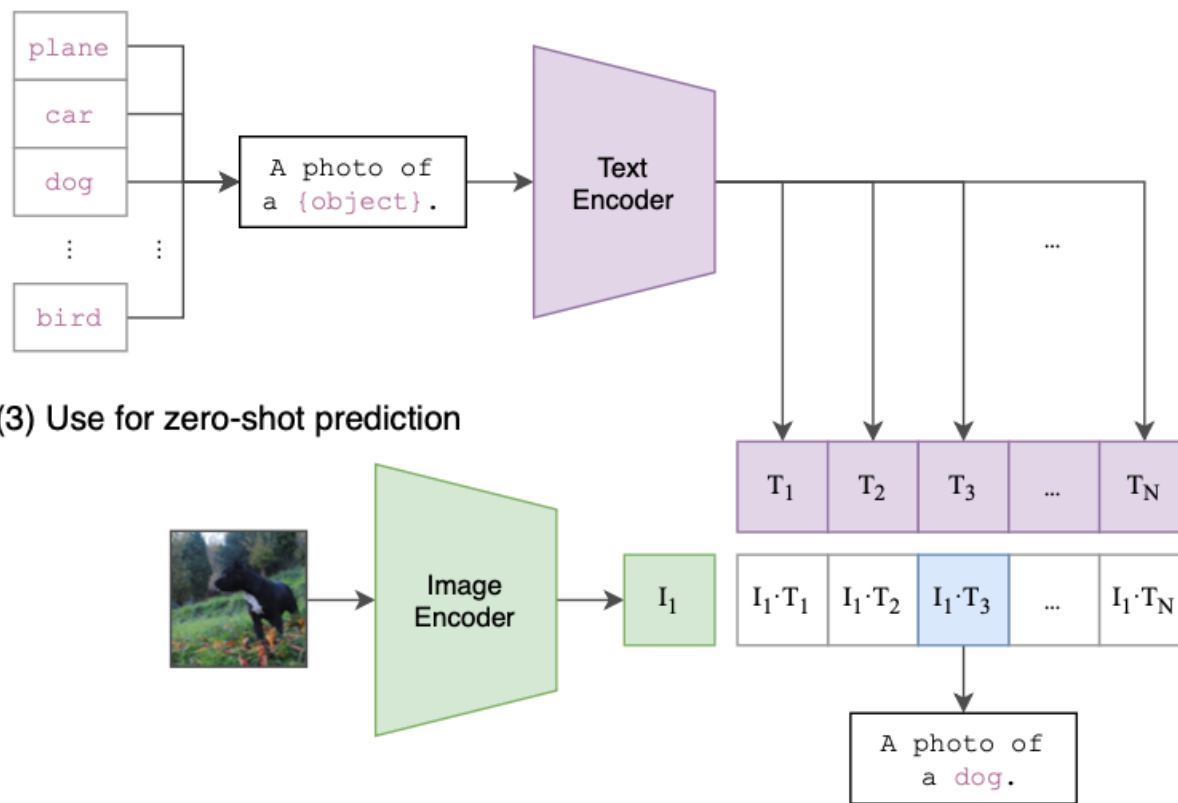
"A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets." The Visual Computer 38, no. 8: 2939-2970.

CLIP: Learning Transferable Visual Models From Natural Language Supervision

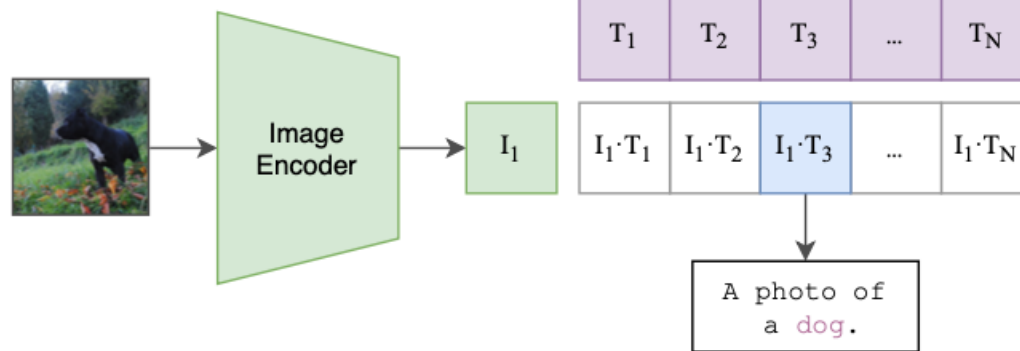
(1) Contrastive pre-training



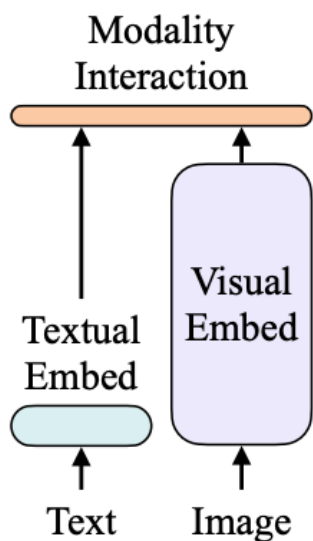
(2) Create dataset classifier from label text



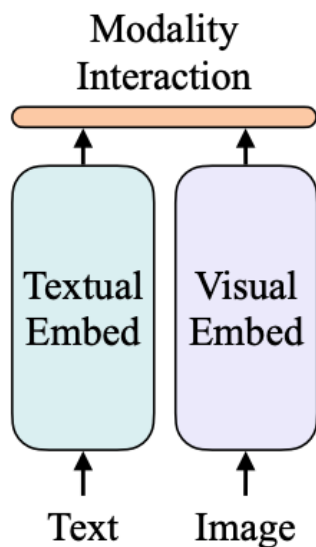
(3) Use for zero-shot prediction



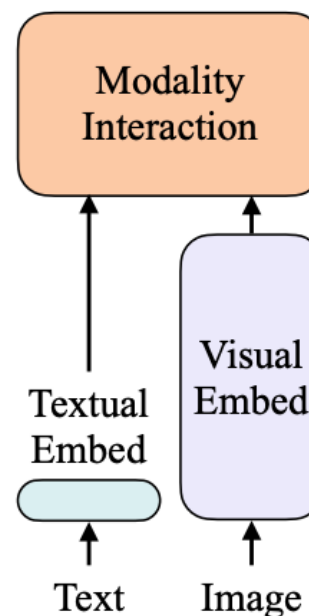
ViLT: Vision-and-Language Transformer Without Convolution or Region Supervision



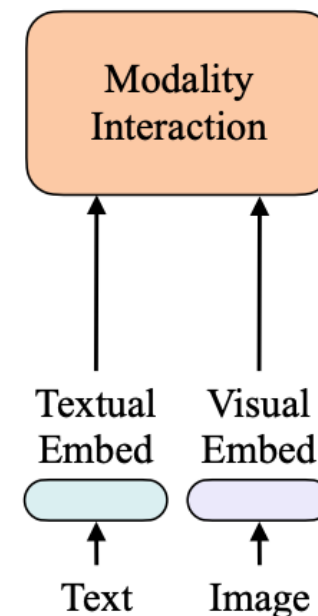
(a) $VE > TE > MI$



(b) $VE = TE > MI$



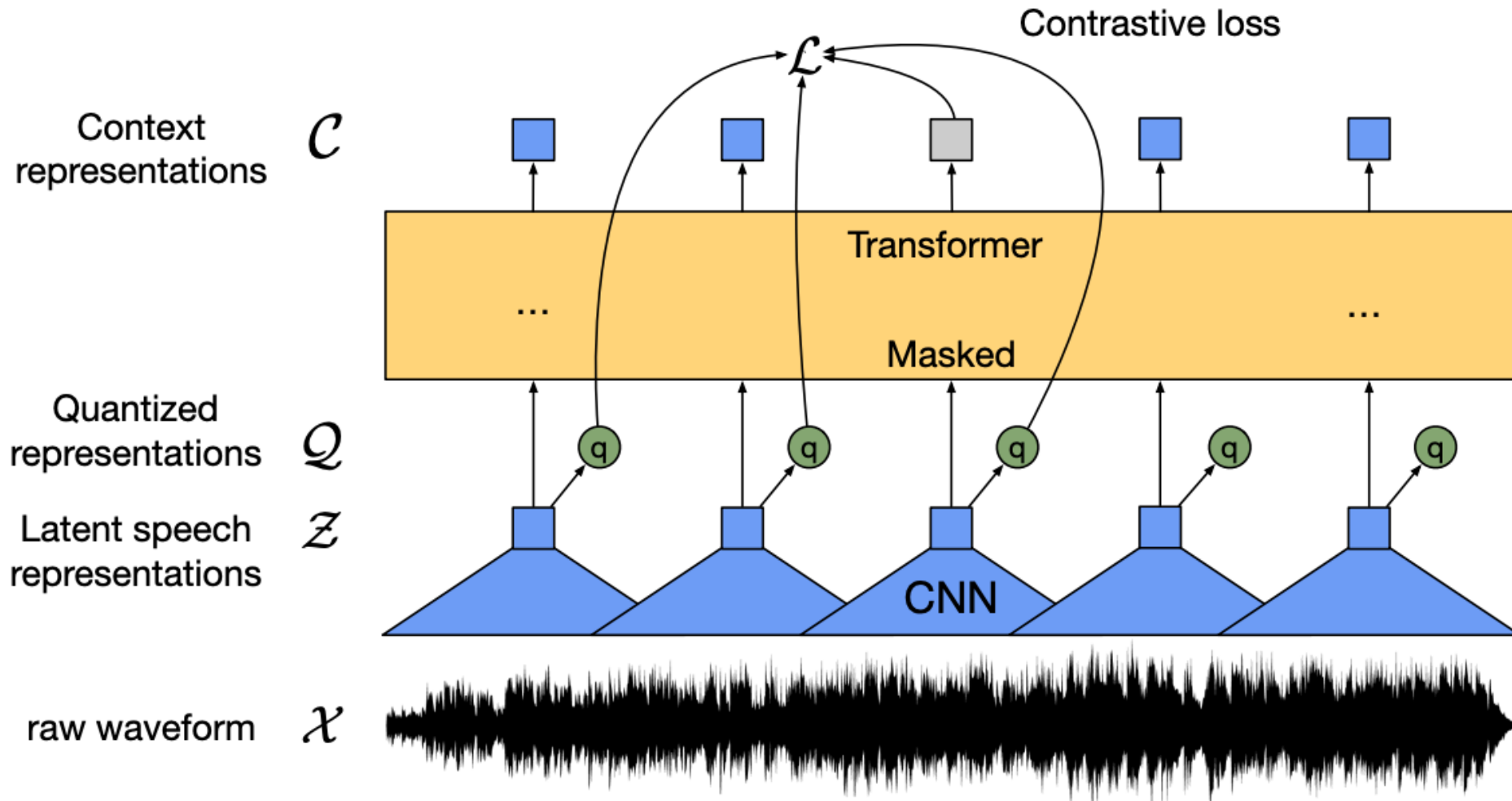
(c) $VE > MI > TE$



(d) $MI > VE = TE$

wav2vec 2.0:

A framework for self-supervised learning of speech representations

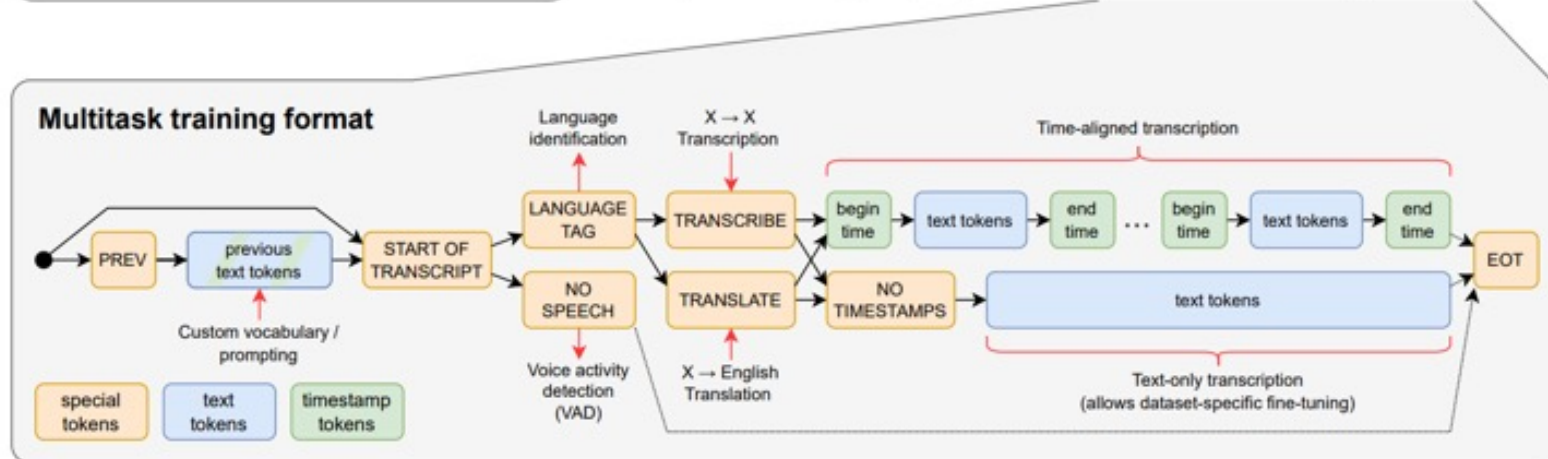
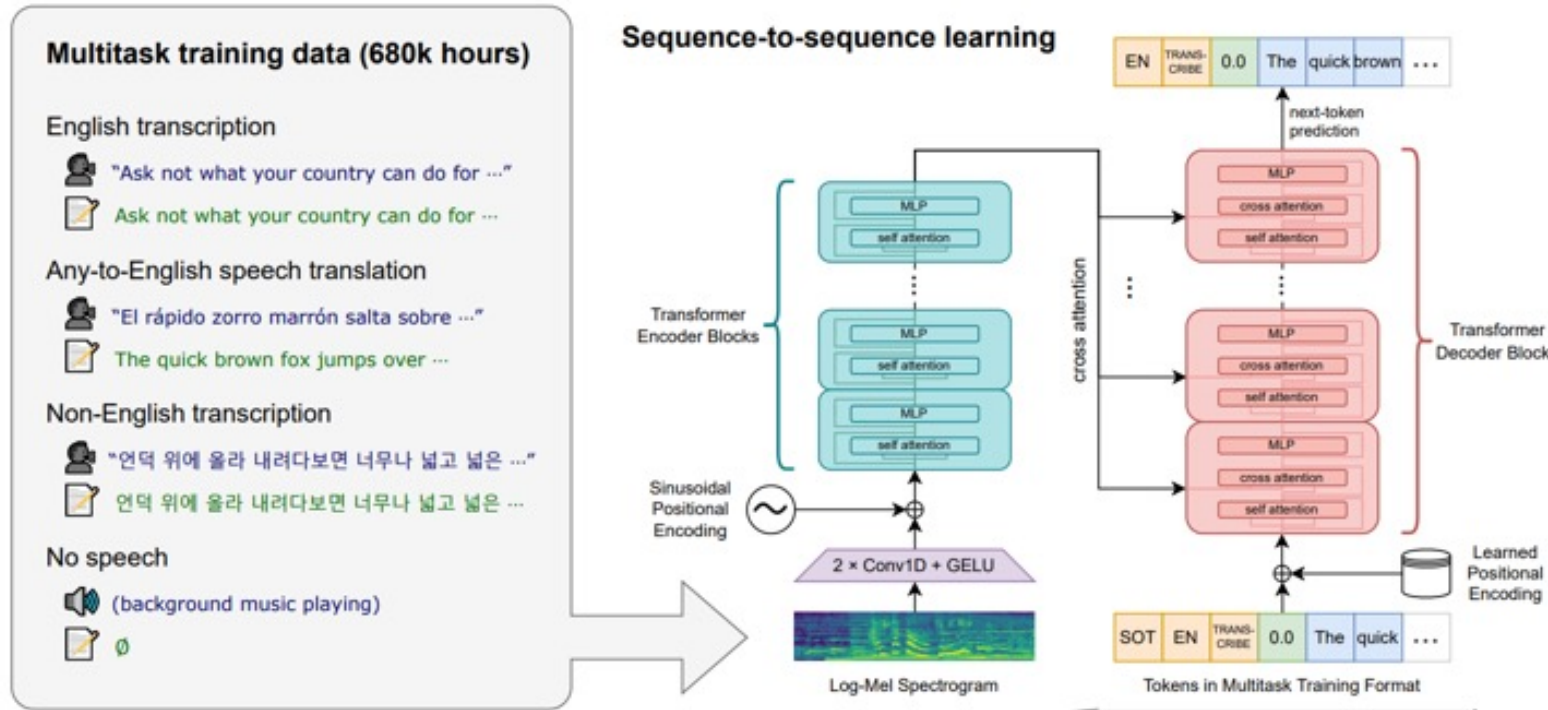


Source: Baevski, Alexei, Yuhao Zhou, Abdelrahman Mohamed, and Michael Auli.

"wav2vec 2.0: A framework for self-supervised learning of speech representations." Advances in Neural Information Processing Systems 33 (2020): 12449-12460.

Whisper:

Robust Speech Recognition via Large-Scale Weak Supervision



Microsoft Azure Text to Speech (TTS)

Text SSML

You can replace this text with any text you wish. You can either write in this text box or paste your own text here.

Try different languages and voices. Change the speed and the pitch of the voice. You can even tweak the SSML (Speech Synthesis Markup Language) to control how the different sections of the text sound. Click on SSML above to [give it a try!](#)

Enjoy using Text to Speech!

Language

English (United States) ▾

Voice

Jenny (Neural) ▾

Speaking style

General ▾

Speaking speed: 1.00



Pitch: 0.00



Play

Hugging Face



Hugging Face

Search models, datasets, spaces, docs, solutions, pricing

Models

Datasets

Spaces

Docs

Solutions

Pricing



Log In

Sign Up



The AI community building the future.

Build, train and deploy state of the art models powered by
the reference open source in machine learning.



Star

58,696

<https://huggingface.co/>

BLOOM

BigScience Large Open-science Open-access Multilingual Language Model



BigScience Large Open-science Open-access Multilingual Language Model

Version 1.3 / 6 July 2022

Current Checkpoint: **Training Iteration 95000**

Total seen tokens: **366B**

Downloads last month
12,875



⚡ Hosted inference API ⓘ

📄 Text Generation

Groups ▾

Examples ▾

I love bloom. Super simple, but so effective! I went through a similar process a couple of years ago when I

sampling greedy

ⓘ [BLOOM prompting tips](#)

Switch to "greedy" for more accurate completion e.g. math/history/translations (but which may be repetitive/less inventive)

Compute

⌘+Enter

1.3

OpenAI Whisper



Hugging Face

Models

Datasets

Spaces

Docs

Solutions

Pricing



Spaces: openai/whisper



422

Running

App

Files




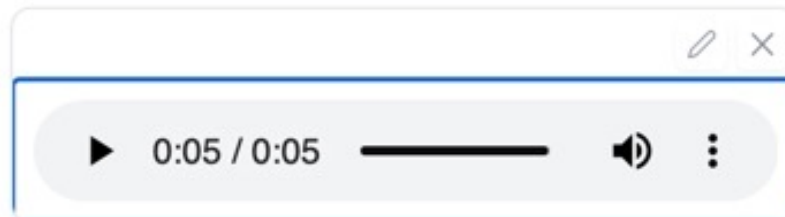
Community 49

Whisper

Whisper is a general-purpose speech recognition model. It is trained on a large dataset of diverse audio and is also a multi-task model that can perform multilingual speech recognition as well as speech translation and language identification. This demo cuts audio after around 30 secs.

You can skip the queue by using google colab for the space:

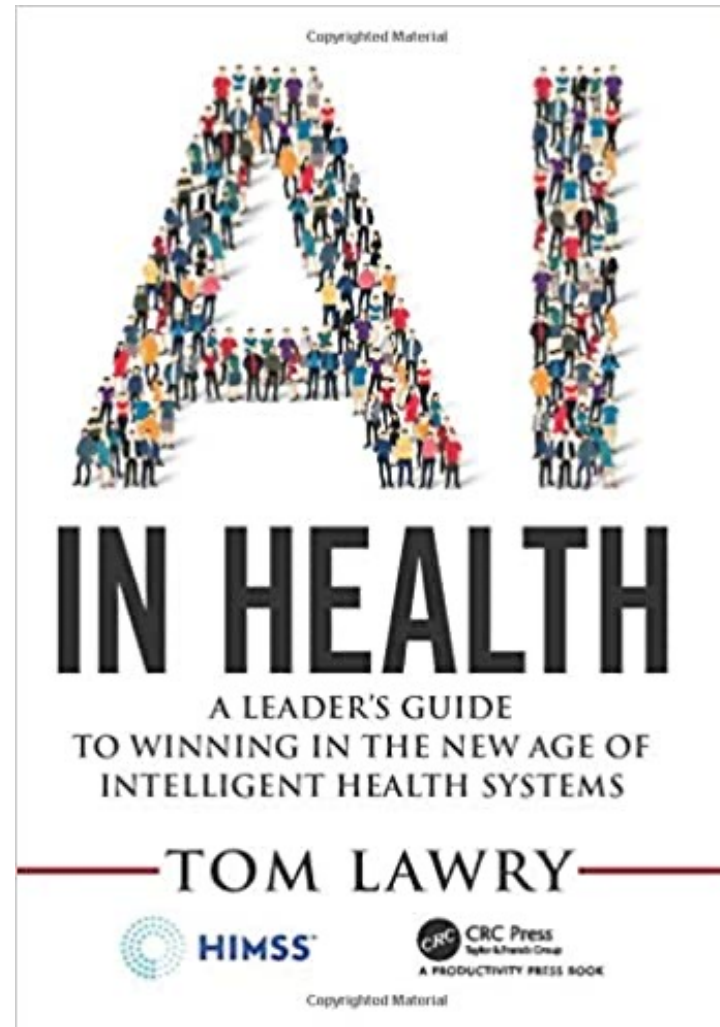
 Open in Colab



Transcribe

Source: <https://huggingface.co/spaces/openai/whisper>

Tom Lawry (2020),
AI in Health:
A Leader's Guide to Winning in the New Age of Intelligent Health Systems,
HIMSS Publishing



Source: Tom Lawry (2020), AI in Health: A Leader's Guide to Winning in the New Age of Intelligent Health Systems, HIMSS Publishing

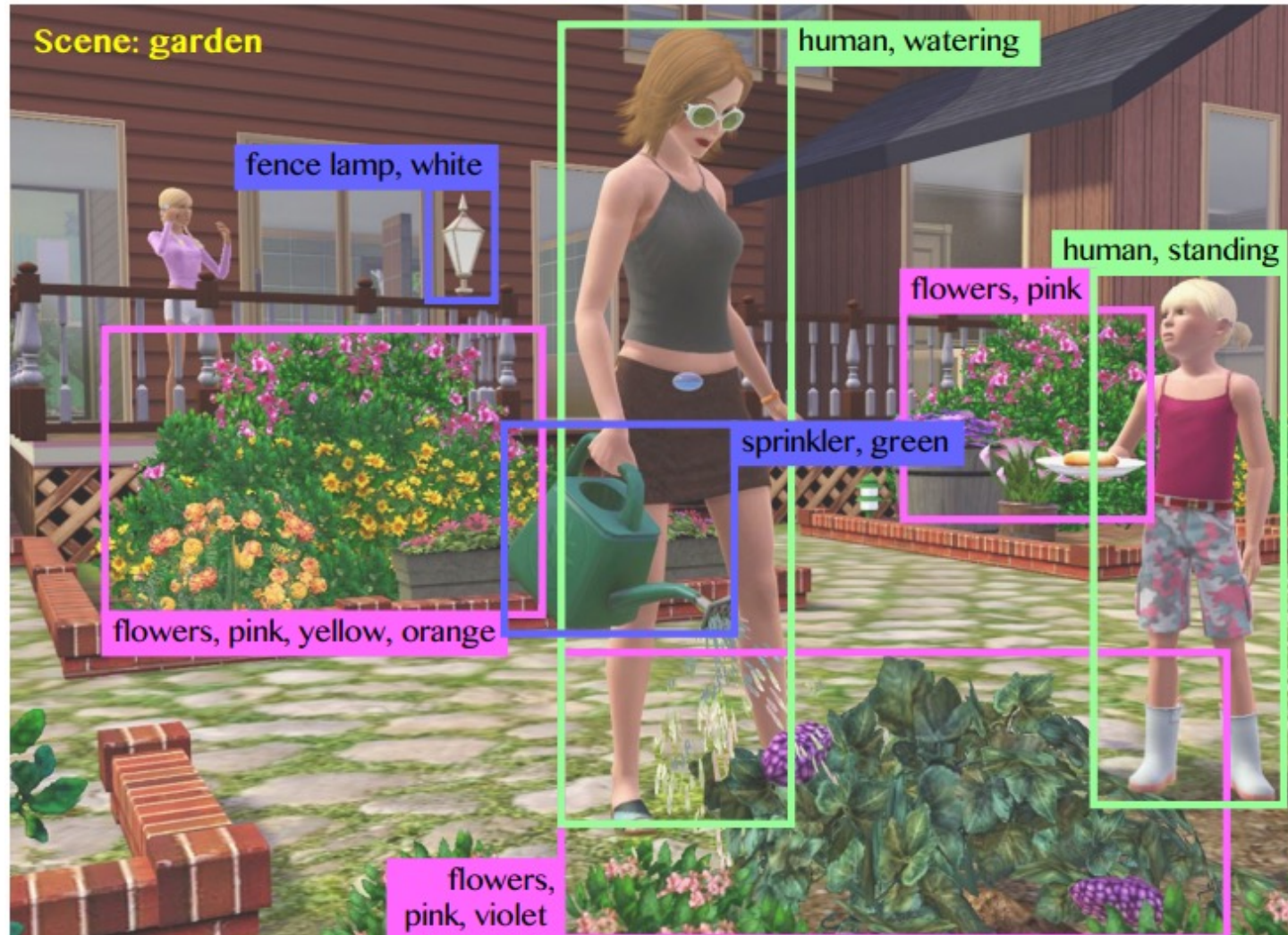
<https://www.amazon.com/Health-HIMSS-Book-Tom-Lawry/dp/0367333716/>

AI in Healthcare



Computer Vision in the Metaverse

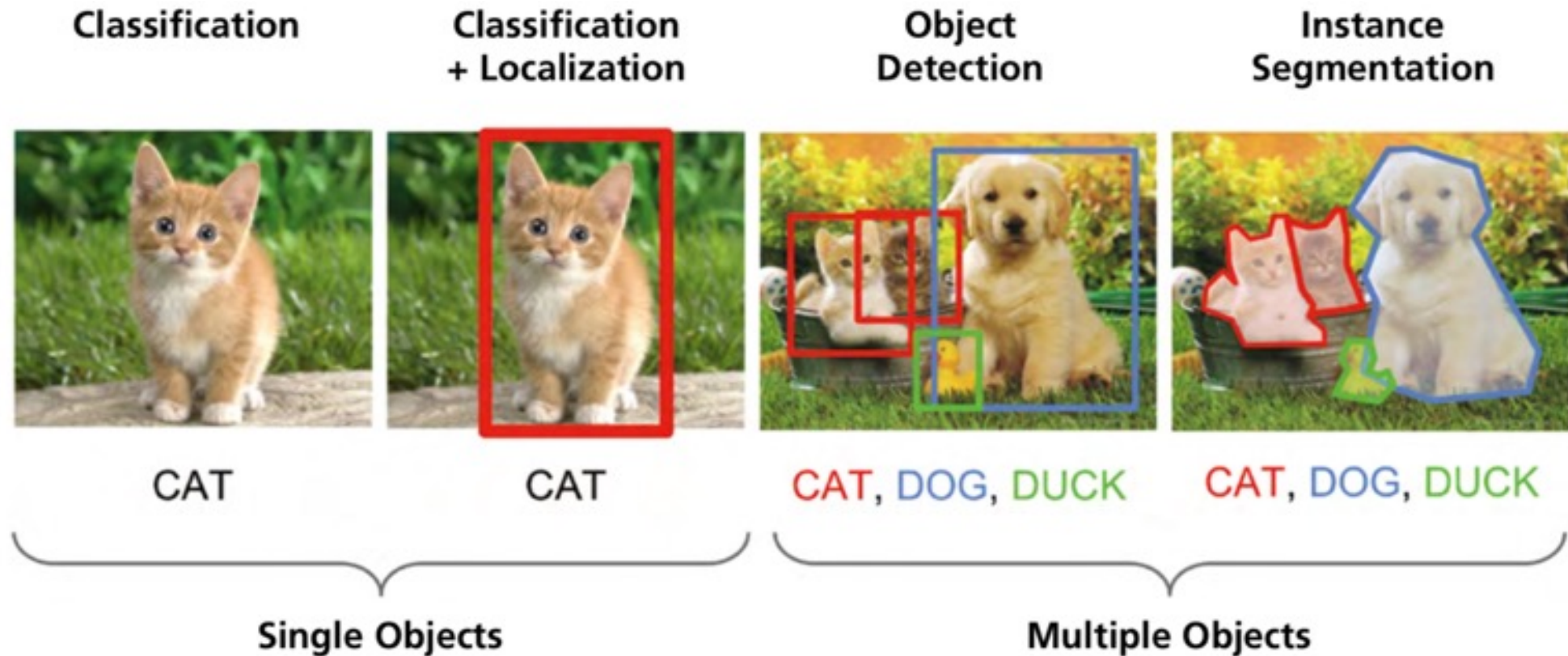
with scene understanding, object detection, and human action/activity recognition



Source: Huynh-The, Thien, Quoc-Viet Pham, Xuan-Quy Pham, Thanh Thi Nguyen, Zhu Han, and Dong-Seong Kim (2022).

"Artificial Intelligence for the Metaverse: A Survey." arXiv preprint arXiv:2202.10336.

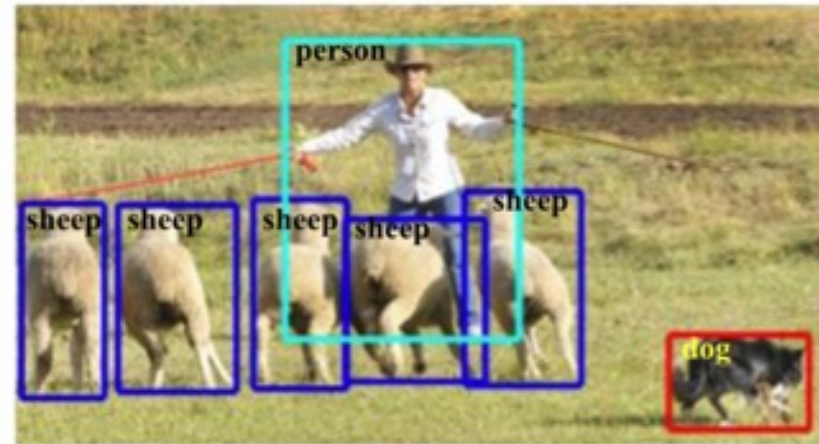
Computer Vision: Image Classification, Object Detection, Object Instance Segmentation



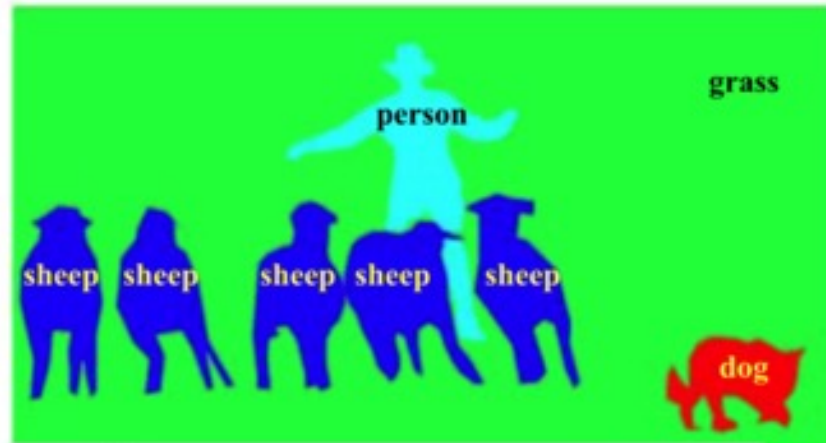
Computer Vision: Object Detection



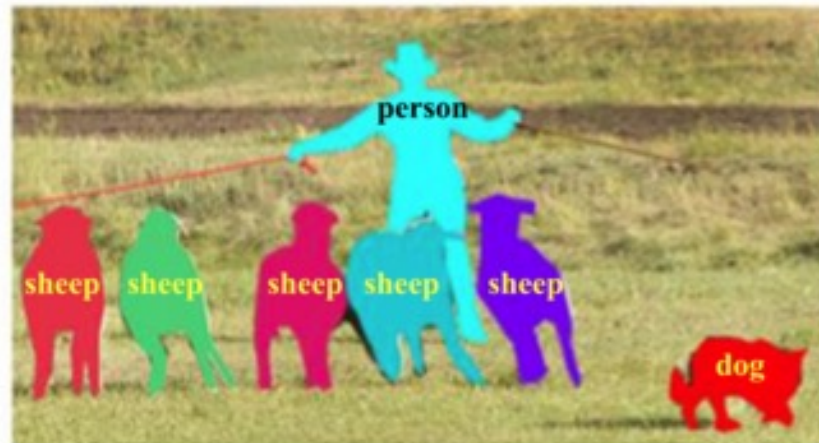
(a) Object Classification



(b) Generic Object Detection (Bounding Box)



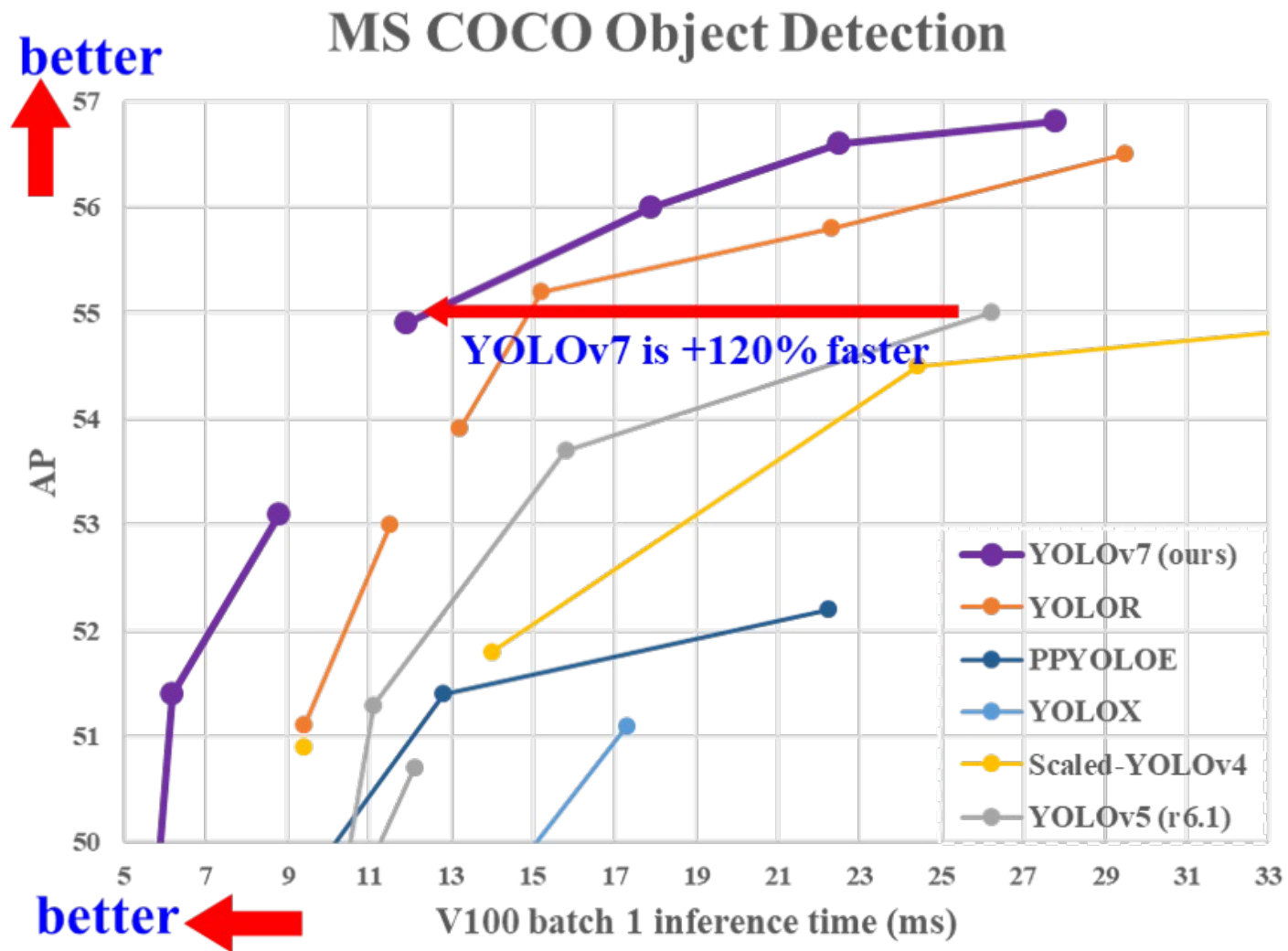
(c) Semantic Segmentation



(d) Object Instance Segmentation

YOLOv7:

Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors



Multimodal Fall Detection

18398

IEEE SENSORS JOURNAL, VOL. 21, NO. 17, SEPTEMBER 1, 2021



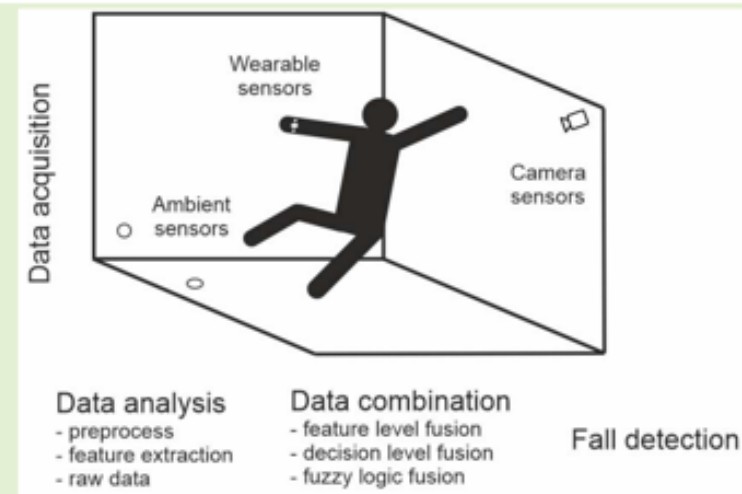
Performance, Challenges, and Limitations in Multimodal Fall Detection Systems: A Review

Vasileios-Rafail Xeferis^{ID}, Athina Tsanousa, Georgios Meditskos^{ID}, Stefanos Vrochidis^{ID},
and Ioannis Kompatsiaris

Ambient Assisted Living (AAL)

Abstract—Fall events among older adults are a serious concern, having an impact on their health and well-being. The development of the Internet of Things (IoT) over the last years has led to the emergence of systems able to track abnormal body movements and falls, thus facilitating fall detection and in some cases prevention. Fusing information from multiple unrelated sources is one of the recent trends in healthcare systems. This work aims to provide a survey of recent methods and trends of multisensor data fusion in fall detection systems and discuss their performance, challenges, and limitations. The paper highlights the benefits of developing multimodal systems for fall detection compared to single-sensor approaches, categorizes the different methods applied to this field, and discusses issues and trends for future work.

Index Terms—Data fusion, fall detection, multisensor fusion, non-wearable sensors, wearable sensors.



Multimodal Fall Detection

Ambient Assisted Living (AAL)

Sensor modalities	Intrusion	ROI specific	Accuracy	Power needs	Computational needs	Environment affected
Wearable	Obtrusive	No	Scenario dependent	High	Low/dependent	No
Ambient	No	Yes	Scenario dependent	Low	Low/dependent	Yes
Camera	Privacy	Yes	High	Low	High	Yes

Challenges of Multimodal Fall Detection

Modalities combined	Performance	Response time	Power consumption	Unaddressed issues	Other advantages
Wearable	Reasonable accuracy.	Reasonably low time.	Up to 62 days.	Obtrusiveness.	Offer to other healthcare applications, continuous monitoring.
Non-wearable	High accuracy.	Reasonably low response time.	No action needed.	ROI restriction.	No recharge power needs.
Wearable and non-wearable	High accuracy.	Low response time.	No evidence.	Complexity.	Takes advantage of both modalities, no ROI restriction.

Fall Detection

Non-Wearable Sensors Fusion

Reference	Year	Sensors	Method	Evaluation	Performance
[46]	2013	PIR and PM sensors.	Graph-theoretical concepts to track user and rule-based algorithm to detect falls.	Falls and ADLs from 5 healthy young subjects.	Accuracy: 82.86%
[47]	2014	Doppler radar sensor and PIR motion sensors.	SVM classifier on Doppler radar features, rule-based algorithm to correct false alarms using PIR data.	A week of continuous data monitoring of a volunteer.	Reduced false alarms by 63% with 100% detection rate.
[48]	2018	IR sensor and an ultrasonic distance sensor.	Thermal IR and ultrasonic features, SVM classifier.	180 falls and ADLs from 3 healthy young subjects, 6 continuous recordings.	Accuracy: 96.7% (discrete test), 90.3% (continuous test).
[52]	2018	Doppler radar sensor and RGB camera.	Multiple CNN, movement classification from radar, aspect ratio sequence from camera, max voting fusion.	1 type of fall and 3 types of ADLs from 3 subjects.	Accuracy: 99.85%
[53]	2019	Doppler radar and depth camera.	Joints' coordinates from depth camera, feature extraction from joints' coordinates and radar data, Linear Discriminant Classifier.	3 different datasets.	Sensitivity: 100% (FD).

Fall Detection Datasets

Datasets	Posture samples	Subject					Type sensor	year
		Number	Height(cm)	Weight(kg)	Age(year)	Gender(M/F)		
Fall detection ⁴	380	4	159-182	48-85	24-31	3M-1F	RGB camera	2007
Fall detection ⁵	72	2	N/A	N/A	N/A	2M	RGB camera	2008
Multicam Fall ⁶	24	1	N/A	N/A	N/A	M	8 RGB camera	2010
Le2i ⁷	249	10	N/A	N/A	N/A	N/A	RGB camera	2013
Thermal simulated fall [8]	35	10	N/A	N/A	N/A	N/A	Thermal camera	2016
SisFall[9]	154	45	149-183	42-102	19-75	23M-21F	RGB camera, 2 accelerometers, 1 gyroscope	2016
UR Fall Detection[10]	70	5	N/A	N/A	N/A	5M	2 Kinect camera, accelerometer	2016
NTU RGB+D Action Recognition [11]	56880	302	N/A	N/A	N/A	N/A	Kinect camera v2	2016
UMA Fall [12]	531	17	155-195	50-93	18-55	10M-7F	Mobility sensors (smartphone)	2017
CMD Fall [13]	20	50	N/A	N/A	21-40	30M-20F	Kinect camera, accelerometer	2018
TST Fall Detection Dataset V2 ⁸	264	11	N/A	N/A	N/A	N/A	Microsoft Kinect v2, accelerometer	2018
UP-Fall[14]	561	17	N/A	N/A	22-58	N/A	Infrared ,inertial measurement	2019

Note: N/A_ Not Available; M_Male; F_Femal

Source: Oumaima, Guendoul, Ait Abdelali Hamd, Tabii Youness, Oulad Haj Thami Rachid, and Bourja Omar.

"Vision-based fall detection and prevention for the elderly people: A review & ongoing research." In 2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS), pp. 1-6. IEEE, 2021.

Human Action Recognition (HAR)

Human Action Recognition from Various Data Modalities: A Review







Zehua Sun, Qihong Ke, Hossein Rahmani, Mohammed Bennamoun, Gang Wang, and Jun Liu

Abstract—Human Action Recognition (HAR) aims to understand human behavior and assign a label to each action. It has a wide range of applications, and therefore has been attracting increasing attention in the field of computer vision. Human actions can be represented using various data modalities, such as RGB, skeleton, depth, infrared, point cloud, event stream, audio, acceleration, radar, and WiFi signal, which encode different sources of useful yet distinct information and have various advantages depending on the application scenarios. Consequently, lots of existing works have attempted to investigate different types of approaches for HAR using various modalities. In this paper, we present a comprehensive survey of recent progress in deep learning methods for HAR based on the type of input data modality. Specifically, we review the current mainstream deep learning methods for single data modalities and multiple data modalities, including the fusion-based and the co-learning-based frameworks. We also present comparative results on several benchmark datasets for HAR, together with insightful observations and inspiring future research directions.

Index Terms—Human Action Recognition, Deep Learning, Data Modality, Single Modality, Multi-modality.

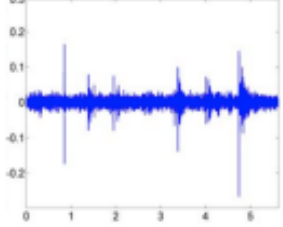
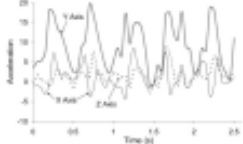
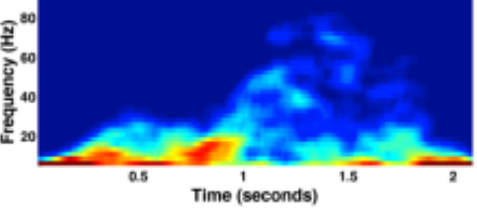
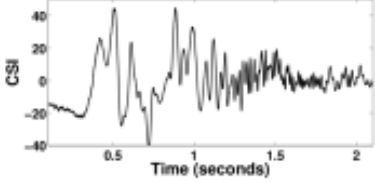
Human Action Recognition (HAR)

Modality

	Modality	Example	Pros	Cons
Visual Modality	RGB	 <p>Hand-waving [27]</p>	<ul style="list-style-type: none"> · Provide rich appearance information · Easy to obtain and operate · Wide range of applications 	<ul style="list-style-type: none"> · Sensitive to viewpoint · Sensitive to background · Sensitive to illumination
	3D Skeleton	 <p>Looking at watch [28]</p>	<ul style="list-style-type: none"> · Provide 3D structural information of subject pose · Simple yet informative · Insensitive to viewpoint · Insensitive to background 	<ul style="list-style-type: none"> · Lack of appearance information · Lack of detailed shape information · Noisy
	Depth	 <p>Mopping floor [29]</p>	<ul style="list-style-type: none"> · Provide 3D structural information · Provide geometric shape information 	<ul style="list-style-type: none"> · Lack of color and texture information · Limited workable distance
	Infrared Sequence	 <p>Pushing [30]</p>	<ul style="list-style-type: none"> · Workable in dark environments 	<ul style="list-style-type: none"> · Lack of color and texture information · Susceptible to sunlight
	Point Cloud	 <p>Bending over [31]</p>	<ul style="list-style-type: none"> · Provide 3D information · Provide geometric shape information · Insensitive to viewpoint 	<ul style="list-style-type: none"> · Lack of color and texture information · High computational complexity
	Event Stream	 <p>Running [32]</p>	<ul style="list-style-type: none"> · Avoid much visual redundancy · High dynamic range · No motion blur 	<ul style="list-style-type: none"> · Asynchronous output · Spatio-temporally sparse · Capturing device is relatively expensive

Human Action Recognition (HAR)

Modality

Non-visual Modality	Audio	 <p>Audio wave of jumping [33]</p>	<ul style="list-style-type: none"> · Easy to locate actions in temporal sequence 	<ul style="list-style-type: none"> · Lack of appearance information
	Acceleration	 <p>Acceleration measurements of walking [34]</p>	<ul style="list-style-type: none"> · Can be used for fine-grained HAR · Privacy protecting · Low cost 	<ul style="list-style-type: none"> · Lack of appearance information · Capturing device needs to be carried by subject
	Radar	 <p>Spectrogram of falling [35]</p>	<ul style="list-style-type: none"> · Can be used for through-wall HAR · Insensitive to illumination · Insensitive to weather · Privacy protecting 	<ul style="list-style-type: none"> · Lack of appearance information · Capturing device is relatively expensive
	WiFi	 <p>CSI waveform of falling [35]</p>	<ul style="list-style-type: none"> · Simple and convenient · Privacy protecting · Low cost 	<ul style="list-style-type: none"> · Lack of appearance information · Sensitive to environments · Noisy

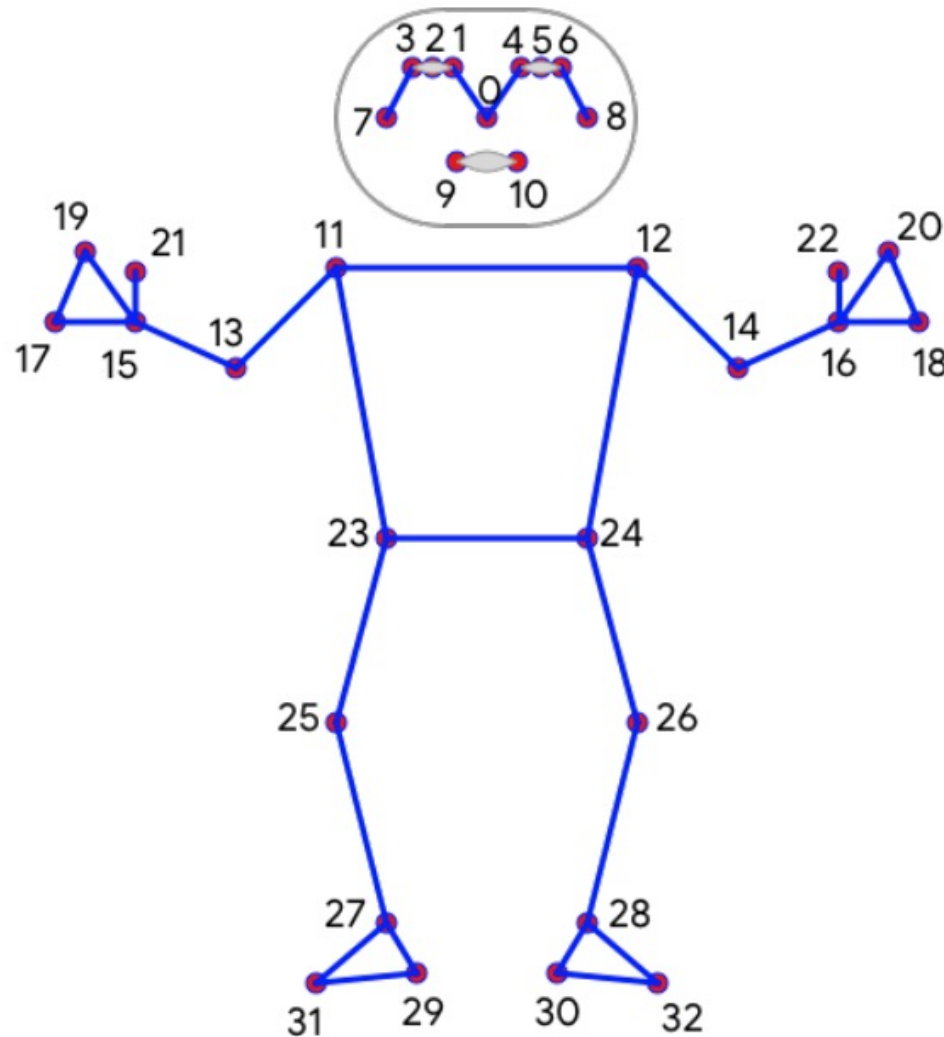
Fall Detection



BlazePose:

On-device Real-time Body Pose tracking

BlazePose 33 Keypoint topology



0. Nose

1. Left eye inner

2. Left eye

3. Left eye outer

4. Right eye inner

5. Right eye

6. Right eye outer

7. Left ear

8. Right ear

9. Mouth left

10. Mouth right

11. Left shoulder

12. Right shoulder

13. Left elbow

14. Right elbow

15. Left wrist

16. Right wrist

17. Left pinky #1 knuckle

18. Right pinky #1 knuckle

19. Left index #1 knuckle

20. Right index #1 knuckle

21. Left thumb #2 knuckle

22. Right thumb #2 knuckle

23. Left hip

24. Right hip

25. Left knee

26. Right knee

27. Left ankle

28. Right ankle

29. Left heel

30. Right heel

31. Left foot index

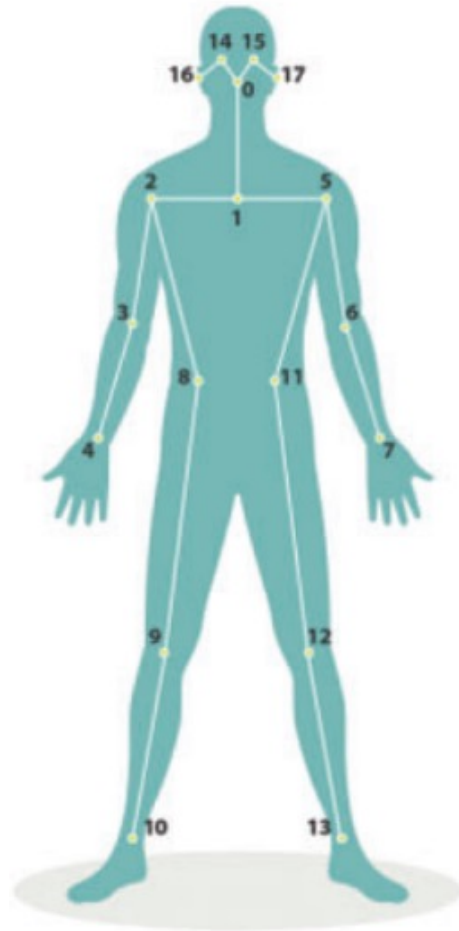
32. Right foot index

BlazePose results on yoga and fitness poses

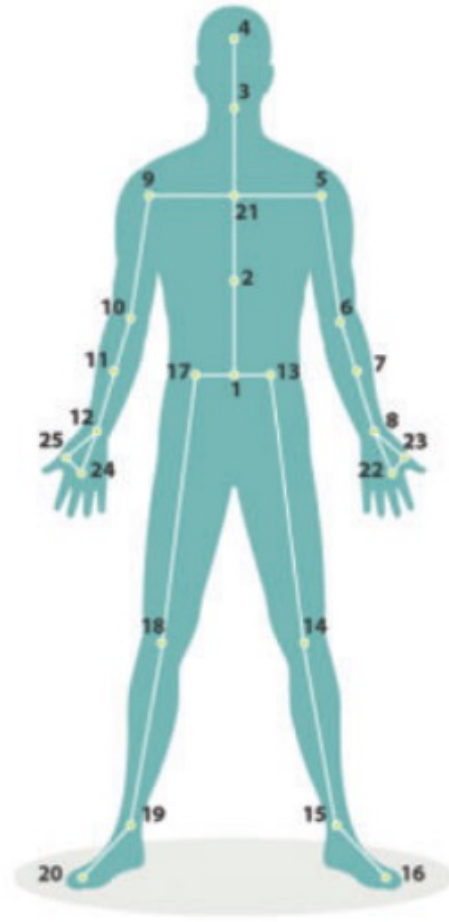


SourceBazarevsky, Valentin, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. "Blazepose: On-device real-time body pose tracking." arXiv preprint arXiv:2006.10204 (2020).

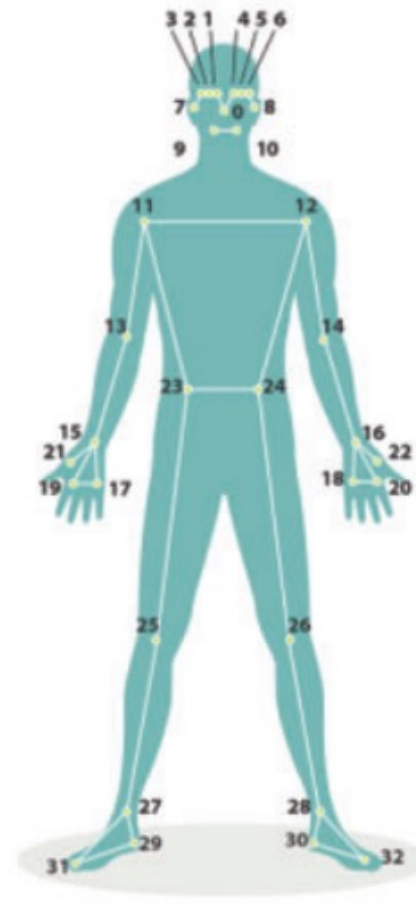
OpenPose vs. BlazePose



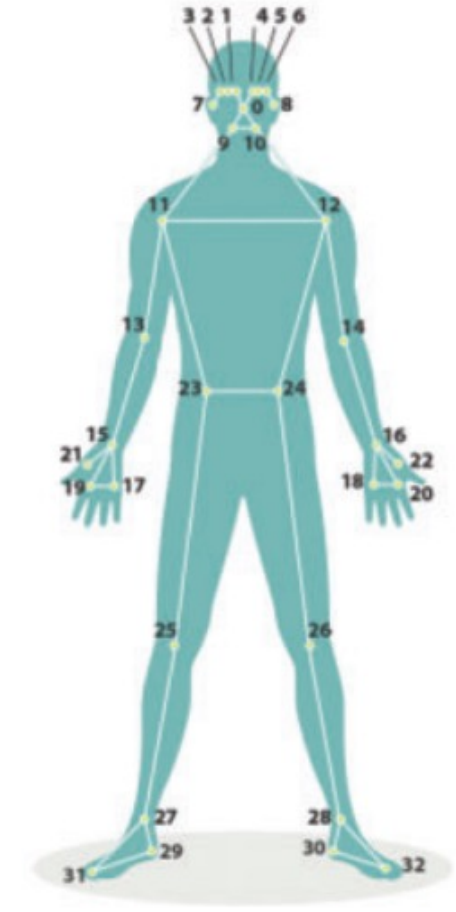
a) OpenPose COCO



b) NTU-RGB+D



c) BlazePose



d) Enhanced-BlazePose

Papers with Code State-of-the-Art (SOTA)

Computer Vision



▶ [See all 1415 tasks](#)

Natural Language Processing



▶ [See all 664 tasks](#)

Summary

- **AI**
- **Data Science**
- **Big Data Analysis**

References

- Stuart Russell and Peter Norvig (2020), *Artificial Intelligence: A Modern Approach*, 4th Edition, Pearson.
- Aurélien Géron (2019), *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd Edition, O'Reilly Media.
- Steven D'Ascoli (2022), *Artificial Intelligence and Deep Learning with Python: Every Line of Code Explained For Readers New to AI and New to Python*, Independently published.
- Varun Grover, Roger HL Chiang, Ting-Peng Liang, and Dongsong Zhang (2018), "Creating Strategic Business Value from Big Data Analytics: A Research Framework", *Journal of Management Information Systems*, 35, no. 2, pp. 388-423.
- Junliang Wang, Chuqiao Xu, Jie Zhang, and Ray Zhong (2022). "Big data analytics for intelligent manufacturing systems: A review." *Journal of Manufacturing Systems* 62 (2022): 738-752.
- Ramesh Sharda, Dursun Delen, and Efraim Turban (2017), *Business Intelligence, Analytics, and Data Science: A Managerial Perspective*, 4th Edition, Pearson
- Arun Amaithi Rajan and Vetriselvi V (2023). "Systematic Survey: Secure and Privacy-Preserving Big Data Analytics in Cloud." *Journal of Computer Information Systems* (2023): 1-21
- Lewis Tunstall, Leandro von Werra, and Thomas Wolf (2022), *Natural Language Processing with Transformers: Building Language Applications with Hugging Face*, O'Reilly Media.
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. *arXiv preprint arXiv:2203.02155*.
- Gozalo-Brizuela, Roberto, and Eduardo C. Garrido-Merchan (2023). "ChatGPT is not all you need. A State of the Art Review of large Generative AI models." *arXiv preprint arXiv:2301.04655* (2023).
- Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. (2022) "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." *arXiv preprint arXiv:2207.02696*.
- Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry et al. (2021) "Learning transferable visual models from natural language supervision." In *International Conference on Machine Learning*, pp. 8748-8763. PMLR.
- Wonjae Kim, Bokyung Son, and Ildoo Kim. (2021) "Vilt: Vision-and-language transformer without convolution or region supervision." In *International Conference on Machine Learning*, pp. 5583-5594. PMLR.
- Meng-Hao Guo, Tian-Xing Xu, Jiang-Jiang Liu, Zheng-Ning Liu, Peng-Tao Jiang, Tai-Jiang Mu, Song-Hai Zhang, Ralph R. Martin, Ming-Ming Cheng, and Shi-Min Hu. (2022) "Attention mechanisms in computer vision: A survey." *Computational Visual Media* ,:1-38.
- Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann.(2020) "Blazepose: On-device real-time body pose tracking." *arXiv preprint arXiv:2006.10204*.
- Huynh-The, Thien, Quoc-Viet Pham, Xuan-Quy Pham, Thanh Thi Nguyen, Zhu Han, and Dong-Seong Kim (2022). "Artificial Intelligence for the Metaverse: A Survey." *arXiv preprint arXiv:2202.10336*.
- Min-Yuh Day (2023), *Python 101*, <https://tinyurl.com/aintpupython101>