NTPU
國立臺北大學
National Taipei University

Introduction to Python for Accounting Applications

1121PAA01
ACC2, NTPU (M5265) (Fall 2023)
Wed 6, 7, 8, (14:10-17:00) (9:10-12:00) (B3F10)

aws educate | Cloud Ambassador
2020 Cohort

aws academy Accredited Educator
aws certified Cloud Practitioner
aws certified Solutions Architect Associate

Min-Yuh Day, Ph.D,
Associate Professor

Institute of Information Management, National Taipei University

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

2023-09-13

國立臺北大學
National Taipei University

aws educate | Cloud Ambassador
2020 Cohort

aws academy
Accredited Educator

aws certified
Solutions Architect
Associate

aws certified
Cloud Practitioner

Min-Yuh Day, Ph.D.

aws educate | Cloud Ambassador
2020 Cohort

Associate Professor, Information Management, NTPU

Visiting Scholar, IIS, Academia Sinica

Ph.D., Information Management, NTU

Director, Intelligent Financial Innovation Technology, IFIT Lab, IM, NTPU

Associate Director, Fintech and Green Finance Center, NTPU

Artificial Intelligence, Financial Technology, Big Data Analytics,
Data Mining and Text Mining, Electronic Commerce

國立臺北大學
National Taipei University

中央研究院 ACADEMIA SINICA 1928

Course Syllabus
National Taipei University
Academic Year 112, 1st Semester (Fall 2023)

- **Course Title: Python for Accounting Applications**
- **Instructor: Min-Yuh Day**
- **Course Class: ACC2, NTPU (3 Credits, Elective)**
- **Details**
  - **EMI Course**
    (3 Credits, Elective, One Semester) (U2004)
- **Time & Place: Wed. 6, 7, 8, 14:10-17:00(B3F17)**
- **Google Meet:** https://meet.google.com/ofh-iosa-ehd

Course Objectives

1. Understand the fundamental concepts of Python for Accounting Applications.
2. Equip with Hands-on practices of Python for Accounting Applications.

Course Outline

- This course introduces the fundamental concepts and hands-on practices of Python for Accounting Applications.
- Topics include
  1. Introduction to Python for Accounting Applications,
  2. Python Programming and Data Science,
  3. Foundations of Python Programming,
  4. Data Structures,
  5. Control Logic and Loops,
  6. Functions and Modules,
  7. Files and Exception Handling,
  8. Data Analytics and Visualization with Python,
  9. Obtaining Data From the Web with Python,
  10. Statistical Analysis with Python,
  11. Machine Learning with Python,
  12. Text Analytics with Python and Large Language Models (LLMs),
  13. Applications of Accounting Data Analytics with Python, and
  14. Applications of ESG Data Analytics with Python.

Syllabus

Week	Date	Subject/Topics
1	2023/09/13	Introduction to Python for Accounting Applications
2	2023/09/20	Python Programming and Data Science
3	2023/09/27	Foundations of Python Programming
4	2023/10/04	Data Structures
5	2023/10/11	Control Logic and Loops
6	2023/10/18	Functions and Modules
7	2023/10/25	Files and Exception Handling
8	2023/11/01	Midterm Project Report

Syllabus

Week	Date	Subject/Topics
9	2023/11/08	Data Analytics and Visualization with Python
10	2023/11/15	Obtaining Data From the Web with Python
11	2023/11/22	Statistical Analysis with Python
12	2023/11/29	Machine Learning with Python
13	2023/12/06	Text Analytics with Python and Large Language Models (LLMs)
14	2023/12/13	Applications of Accounting Data Analytics with Python
15	2023/12/20	Applications of ESG Data Analytics with Python
16	2023/12/27	Final Project Report

Teaching Methods and Activities

- Lecture
- Discussion
- Practicum

Evaluation Methods

- Individual Presentation 30 %
- Group Presentation 30 %
- Case Report 20 %
- Class Participation 10 %
- Assignment 10 %

Required Texts

- **Allen B. Downey (2016), Think Python: How to Think Like a Computer Scientist, 2nd Edition, O'Reilly Media**

Reference Books

1. Frederick Kaefer and Paul Kaefer (2020), Introduction to Python Programming for Business and Social Science Applications, SAGE Publications

2. Abdullah Karasan (2021), Machine Learning for Financial Risk Management with Python: Algorithms for Modeling Risk, O'Reilly Media

3. Vic Anand, Khrystyna Bochkay, and Roman Chychyla (2020), Using Python for Text Analysis in Accounting Research, Now Publishers

4. Aurélien Géron (2022), Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 3rd Edition, O'Reilly Media.

5. Yves Hilpisch (2018), Python for Finance: Mastering Data-Driven Finance, 2nd Edition, O'Reilly Media.

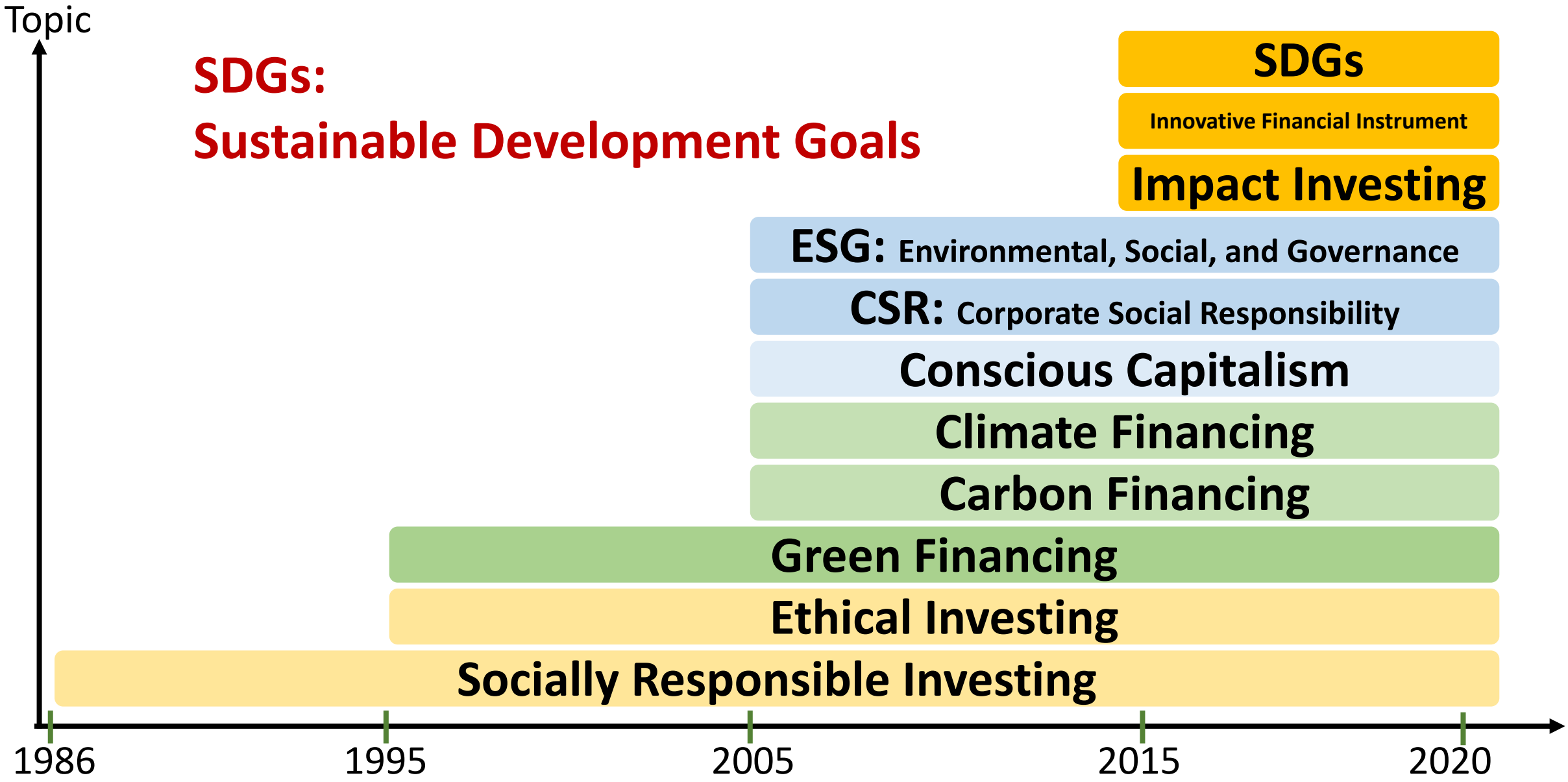
6. Yves Hilpisch (2020), Artificial Intelligence in Finance: A Python-Based Guide, O'Reilly Media.

Other References

- **Python, https://www.python.org/**

Green Finance
and
Sustainable Finance

Evolution of Sustainable Finance Research



Source: Kumar, S., Sharma, D., Rao, S., Lim, W. M., & Mangla, S. K. (2022). Past, present, and future of sustainable finance: Insights from big data analytics through machine learning of scholarly research. Annals of Operations Research, 1-44.

AI for Environmental, Social, and Governance (AI4ESG)

Source: Nenad Tomašev, Julien Cornebise, Frank Hutter, Shakir Mohamed, Angela Picciariello, Bec Connelly, Danielle Belgrave et al. (2020) "AI for social good: unlocking the opportunity for positive impact." Nature Communications 11, no. 1: 1-6.

AI for Social Good (AI4SG)

Source: Nenad Tomašev, Julien Cornebise, Frank Hutter, Shakir Mohamed, Angela Picciariello, Bec Connelly, Danielle Belgrave et al. (2020) "AI for social good: unlocking the opportunity for positive impact." Nature Communications 11, no. 1: 1-6.

Sustainability

SDGs

CSR

ESG

Sustainable Development Goals (SDGs)

Source: https://sdgs.un.org/goals

Sustainable Development Goals (SDGs) and 5P

Partnership

Peace

Prosperity

People

Planet

Source: Folke, Carl, Reinette Biggs, Albert V. Norström, Belinda Reyers, and Johan Rockström. "Social-ecological resilience and biosphere-based sustainability science."Ecology and Society 21, no. 3 (2016).

Green Finance

Generic term
implying use or diversion
of financial resources
to deploy and support projects
with long term positive impact
on the environment

Source: Bhattacharyya, Rupsha. "Green finance for energy transition, climate action and sustainable development: overview of concepts, applications, implementation and challenges." Green Finance 4, no. 1 (2022): 1-35.

Sustainable Finance

Finances

deployed in support of projects that ensure just, sustainable and inclusive growth or attainment of one or more **sustainable development goals**

Source: Bhattacharyya, Rupsha. "Green finance for energy transition, climate action and sustainable development: overview of concepts, applications, implementation and challenges." Green Finance 4, no. 1 (2022): 1-35.

Carbon Finance

Financial instruments
based on
economic value of carbon emissions
which an organization cannot avoid but which
it offsets by funding other compensatory projects
that contribute to carbon emissions reduction

Source: Bhattacharyya, Rupsha. "Green finance for energy transition, climate action and sustainable development: overview of concepts, applications, implementation and challenges." Green Finance 4, no. 1 (2022): 1-35.

Climate Finance

Finances deployed
in support of low carbon and
climate resilient projects
that help in climate change mitigation and
adaptation efforts,
particularly in the
energy and infrastructure sectors

Source: Bhattacharyya, Rupsha. "Green finance for energy transition, climate action and sustainable development: overview of concepts, applications, implementation and challenges." Green Finance 4, no. 1 (2022): 1-35.

ESG Investing

Investments considering the broad range of environmental (e.g. climate change, pollution biodiversity loss), social (e.g. working conditions, human rights, salary or compensation structures) and governance (e.g. board composition, diversity and inclusion, taxes) characteristics of the projects or companies being invested in; ethical and business sustainability considerations are integral part of financing

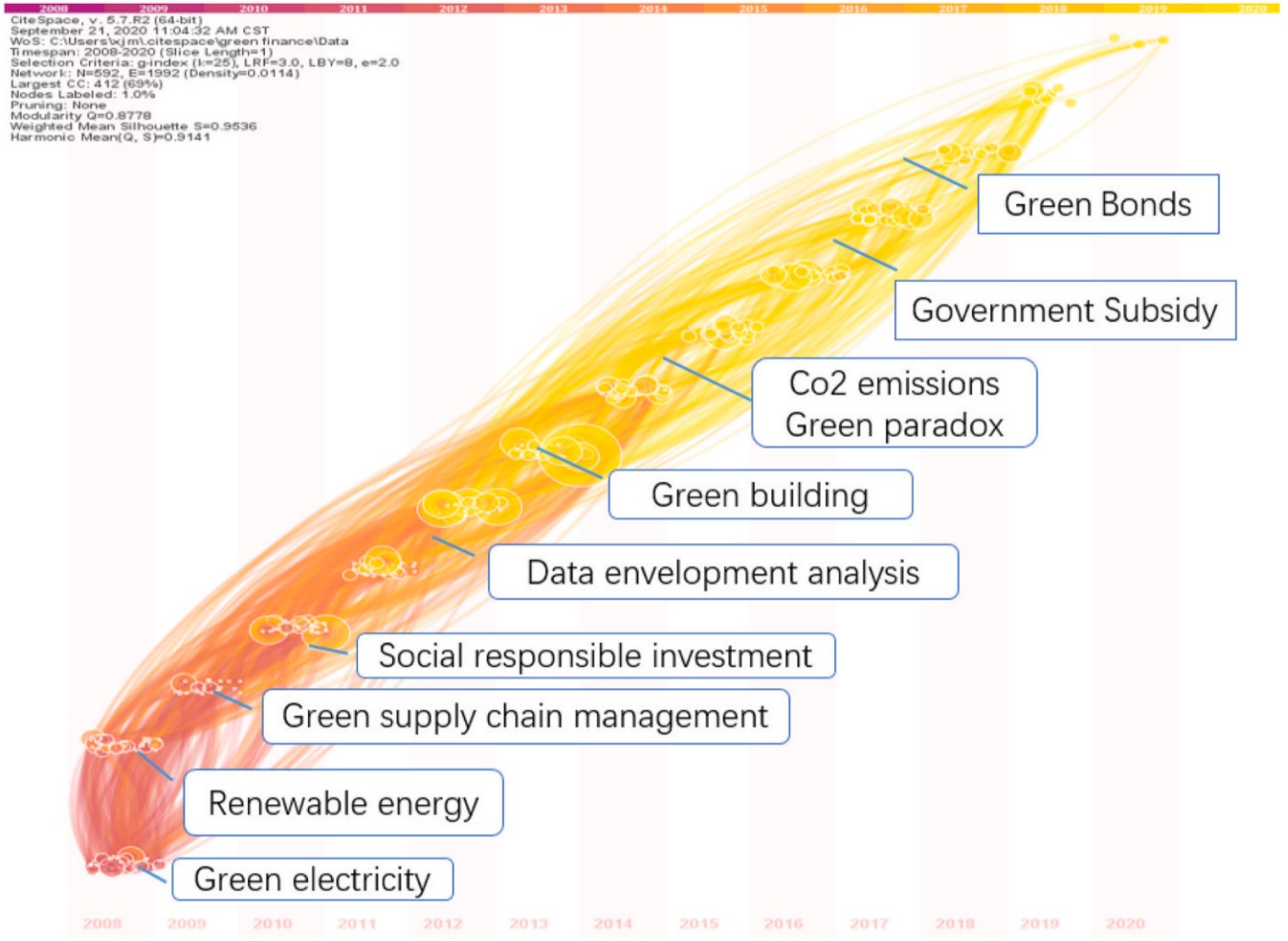
Source: Bhattacharyya, Rupsha. "Green finance for energy transition, climate action and sustainable development: overview of concepts, applications, implementation and challenges." Green Finance 4, no. 1 (2022): 1-35.

Impact Investing

Investing in projects
that solve a social or environmental problem;
the focus is on the positive impact
rather than the
means used to produce that impact

Source: Bhattacharyya, Rupsha. "Green finance for energy transition, climate action and sustainable development: overview of concepts, applications, implementation and challenges." Green Finance 4, no. 1 (2022): 1-35.

Dynamic Trends of Green Finance and Energy Policy



Source: Wang, Moran, Xuerong Li, and Shouyang Wang. (2021) "Discovering research trends and opportunities of green finance and energy policy: A data-driven scientometric analysis." Energy Policy 154 (2021): 112295.

ESG:

Environmental

Social

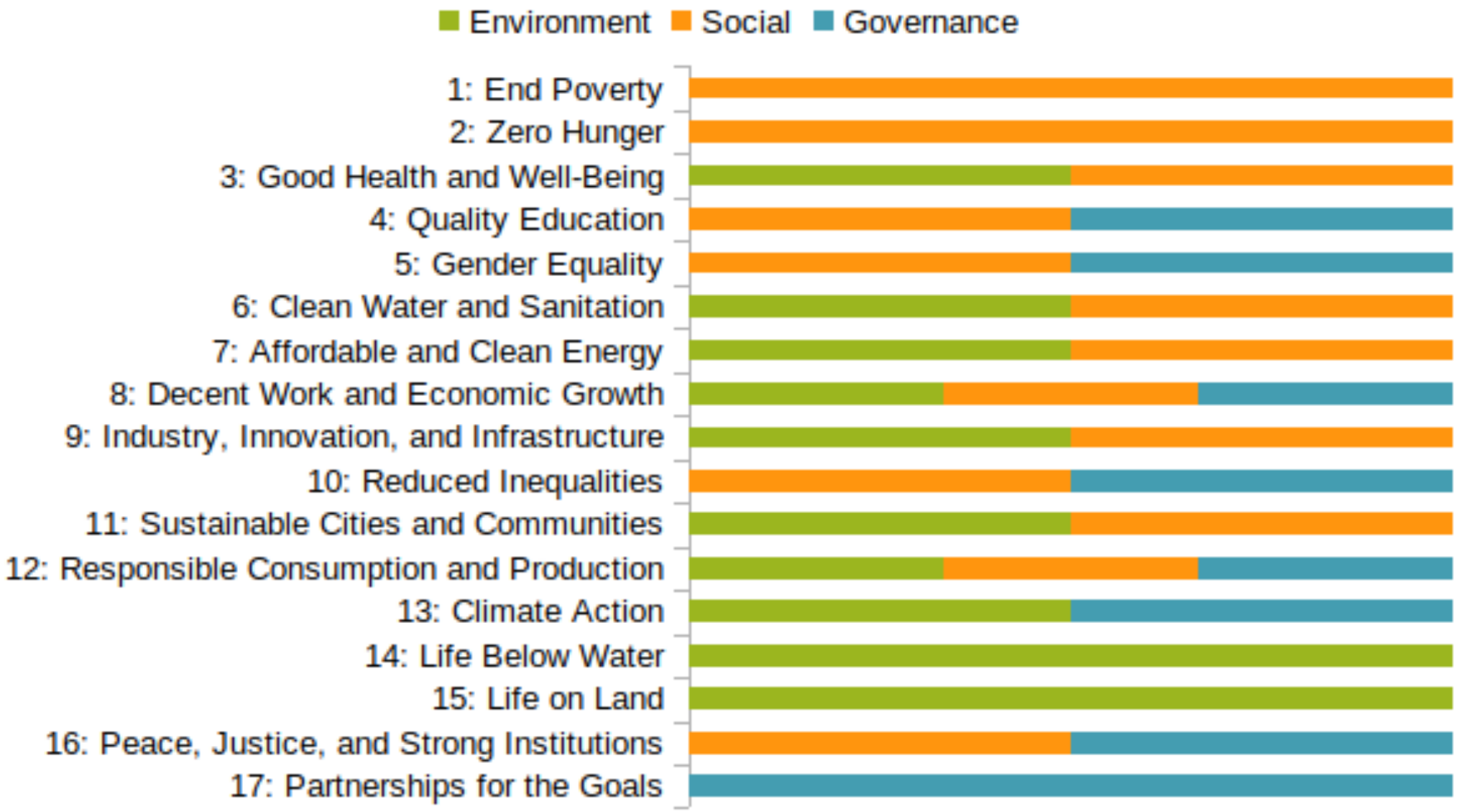
Governance

CSR: Corporate Social Responsibility

ESG to 17 SDGs

Source: Henrik Skaug Sætra (2021) "A Framework for Evaluating and Disclosing the ESG Related Impacts of AI with the SDGs." Sustainability 13, no. 15 (2021): 8503.

ESG to 17 SDGs



Source: https://sustainometric.com/esg-to-sdgs-connected-paths-to-a-sustainable-future/

Generative AI for ESG Applications

AI and Sustainability Development Goals (SDGs)

SDGs	1 No poverty	2 Zero hunger	3 Good health and well-being	4 Quality education	5 Gender equality	6 Clean water and sanitation	7 Affordable and clean energy	8 Decent work and economic growth	9 Industry, innovation and infrastructure	10 Reduces inequalities	11 Sustainable cities and communities	12 Responsible consumption and production	13 Climate action	14 Life below water	15 Life on land	16 Peace, justice and strong institutions	17 Partnerships for the goals
Economic								●	●	●	○						●
Ecological		○					○				○	○	●	●	●		
Social	●	●	●	●	●	●	●				●	●				●	
Positive impact of AI*	100%	76%	69%	10%0	56%	100%	100%	92%	100%	90%	100%	82%	80%	90%	100%	58%	26%

Note: ● adopted from Vinuesa et al. (2020), ○ added based on our analysis.
*The assessment of AI's possible positive impact is based on a consensus-based expert elicitation process (Vinuesa et al., 2020).

Source: Schoormann, T., Strobel, G., Möller, F., Petrik, D., & Zschech, P. (2023).
Artificial Intelligence for Sustainability—A Systematic Review of Information Systems Literature. Communications of the Association for Information Systems, 52(1), 8.

AI for Sustainability

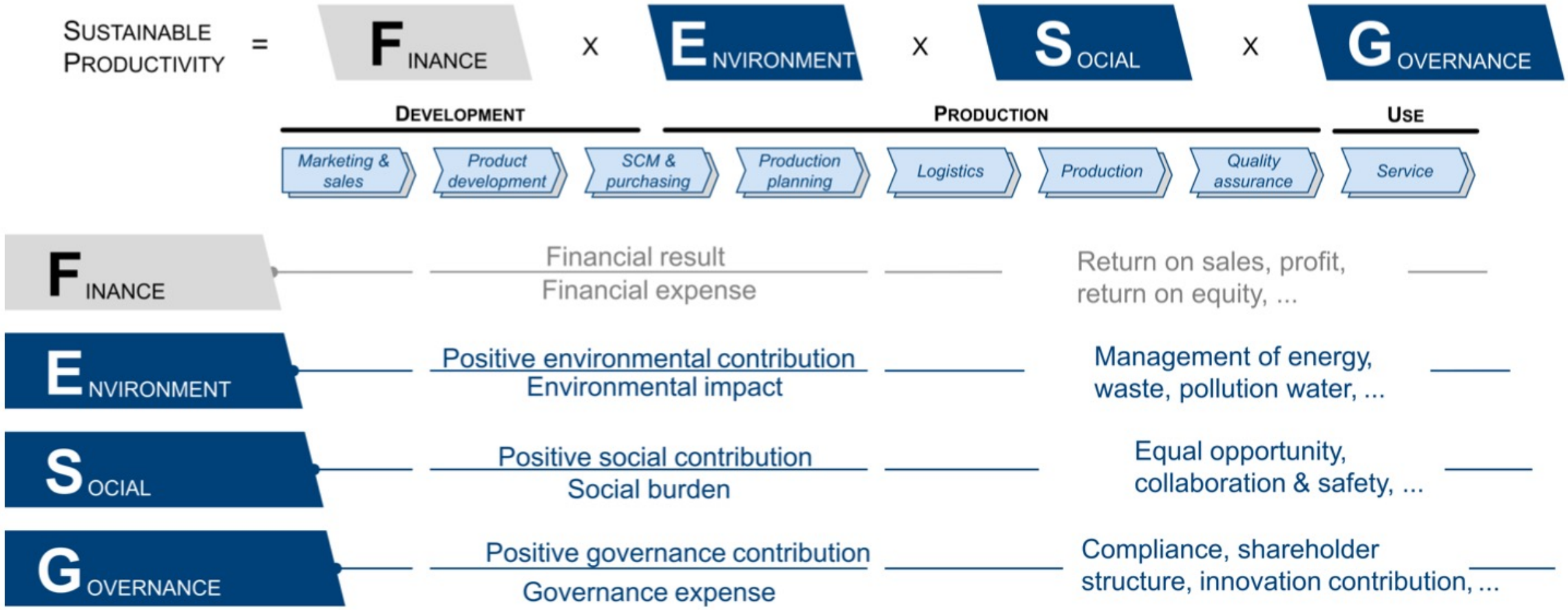
Dimension	Code characteristics					
Primary objective[1]	Develop new (AI) methods *(11/95)*	Compare (AI) methods *(39/95)*	Apply (AI) methods *(53/95)*	Develop new system *(20/95)*	Other objective *(4/95)*	
Sustainability dimension	Economic *(23/95)*		Ecological *(17/95)*		Social *(72/95)*	
Sustainable Development Goals (SDGs)	SDG 1 *(0/95)*	SDG 2 *(2/95)*	SDG 3 *(55/95)*	SDG 4 *(6/95)*	SDG 5 *(0/95)*	SDG 6 *(0/95)*
	SDG 7 *(9/95)*	SDG 8 *(7/95)*	SDG 9 *(8/95)*	SDG 10 *(1/95)*	SDG 11 *(9/95)*	SDG 12 *(8/95)*
	SDG 13 *(2/95)*	SDG 14 *(0/95)*	SDG 15 *(2/95)*	SDG 16 *(11/95)*	SDG 17 *(0/95)*	
Data source	Reviews *(12/95)*	Social media/ Online forums *(31/95)*	Health records *(21/95)*	Environment/ Weather *(10/95)*	Energy *(5/95)*	
Data source plurality	Single source *(50/95)*		Multiple sources *(44/95)*		N/A *(1/95)*	
Data sensitivity	Publicly available data *(64/95)*		Internal data *(16/95)*	Other *(11/95)*	N/A *(9/95)*	
Manual labeling	Yes *(32/95)*			No *(63/95)*		
Technology	ML *(91/95)*		NLP *(42/95)*	CV *(12/95)*	Other *(21/95)*	
Type of learning for ML approach	Supervised learning *(85/95)*			Unsupervised learning *(23/95)*		
Neural vs. non-neural	Non-neural *(45/95)*		Neural *(50/95)*		Deep learning *(38/95)*	
Evaluation	Technical evaluation *(83/95)*			Domain evaluation *(25/95)*		
Paradigm	DSR/ADR *(30/95)*			Non-DSR/ADR *(64/95)*		

	0-9	10-29	30-54	55-69	70-95

Notes: Code dimensions are not mutually exclusive; one article can be classified into one or more code characteristics; [1] 'Compare' does include 'apply'.

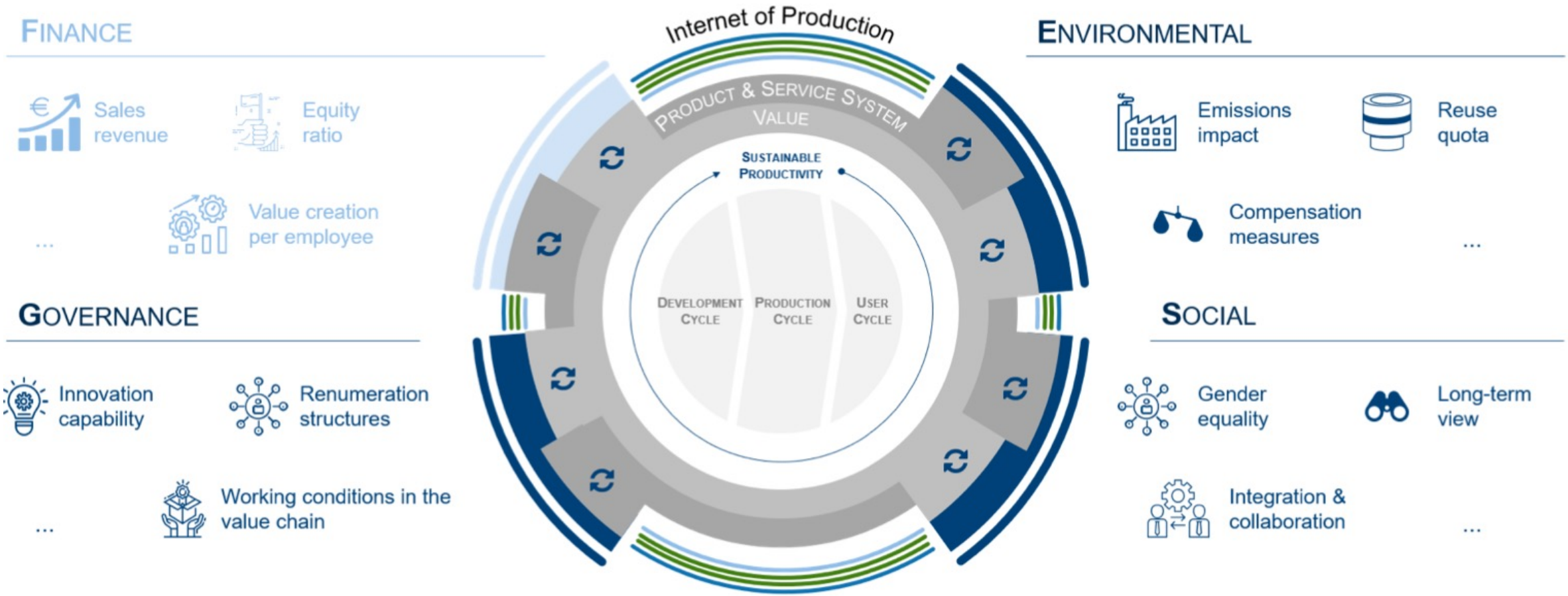
Source: Schoormann, T., Strobel, G., Möller, F., Petrik, D., & Zschech, P. (2023). Artificial Intelligence for Sustainability—A Systematic Review of Information Systems Literature. Communications of the Association for Information Systems, 52(1), 8.

Sustainable Productivity: Finance ESG



Source: Huang, Ziqi, Yang Shen, Jiayi Li, Marcel Fey, and Christian Brecher (2021). "A survey on AI-driven digital twins in Industry 4.0: Smart manufacturing and advanced robotics." Sensors 21, no. 19: 6340.

Sustainable Resilient Manufacturing
ESG

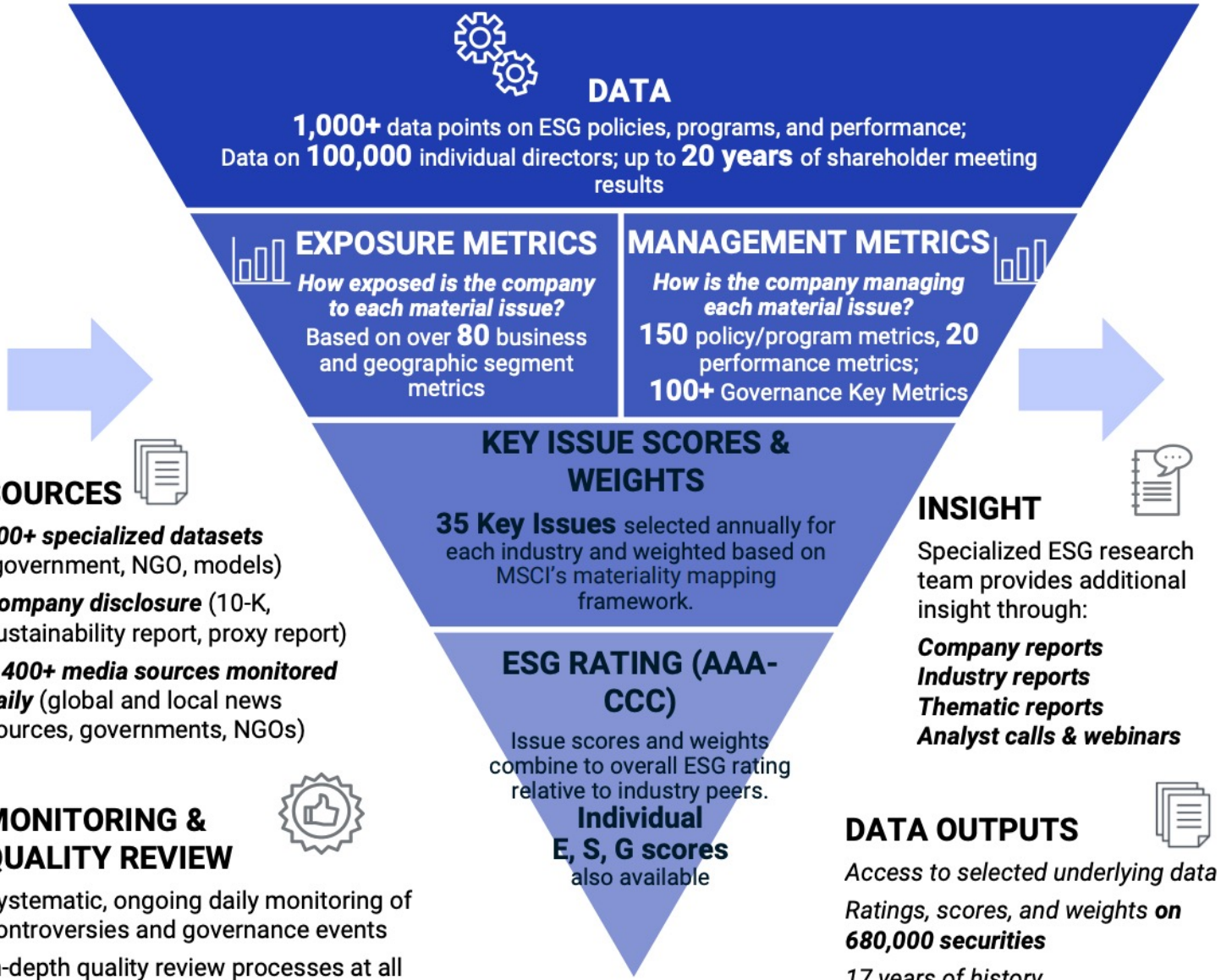


Source: Huang, Ziqi, Yang Shen, Jiayi Li, Marcel Fey, and Christian Brecher (2021). "A survey on AI-driven digital twins in Industry 4.0: Smart manufacturing and advanced robotics." Sensors 21, no. 19: 6340.

ESG Indexes

- MSCI ESG Index
- Dow Jones Sustainability Indices (DJSI)
- FTSE ESG Index

MSCI ESG Rating Framework



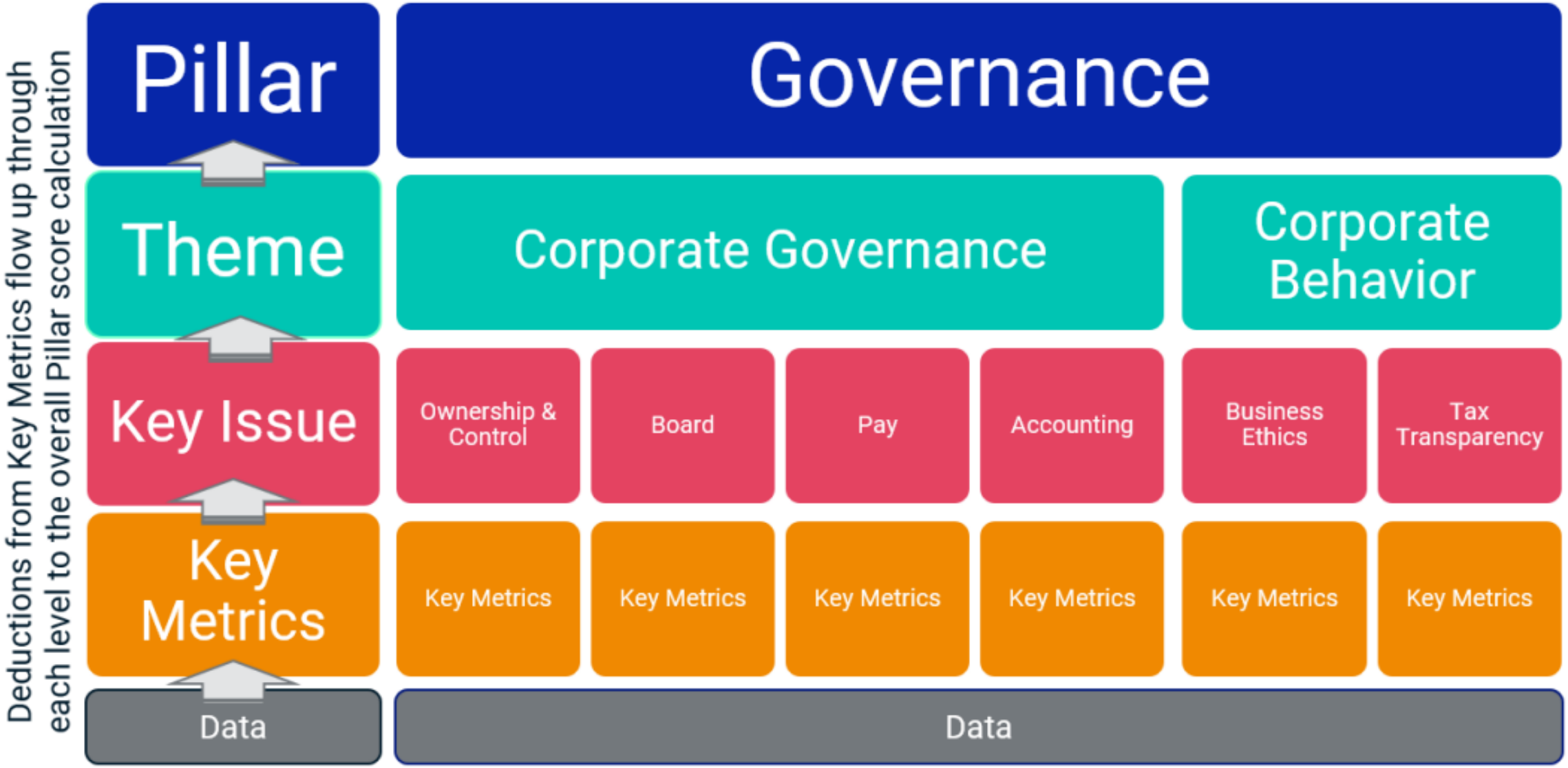
Source: https://www.msci.com/documents/1296102/21901542/ESG-Ratings-Methodology-Exec-Summary.pdf

MSCI ESG Key Issue Hierarchy

3 Pillars	10 Themes	35 ESG Key Issues	
Environment	**Climate Change**	Carbon Emissions	Financing Environmental Impact
		Product Carbon Footprint	Climate Change Vulnerability
	Natural Capital	Water Stress	Raw Material Sourcing
		Biodiversity & Land Use	
	Pollution & Waste	Toxic Emissions & Waste	Electronic Waste
		Packaging Material & Waste	
	Environmental Opportunities	Opportunities in Clean Tech	Opportunities in Renewable Energy
		Opportunities in Green Building	
Social	**Human Capital**	Labor Management	Human Capital Development
		Health & Safety	Supply Chain Labor Standards
	Product Liability	Product Safety & Quality	Privacy & Data Security
		Chemical Safety	Responsible Investment
		Consumer Financial Protection	Health & Demographic Risk
	Stakeholder Opposition	Controversial Sourcing	
		Community Relations	
	Social Opportunities	Access to Communications	Access to Health Care
		Access to Finance	Opportunities in Nutrition & Health
Governance	**Corporate Governance**	Ownership & Control	Pay
		Board	Accounting
	Corporate Behavior	Business Ethics	
		Tax Transparency	

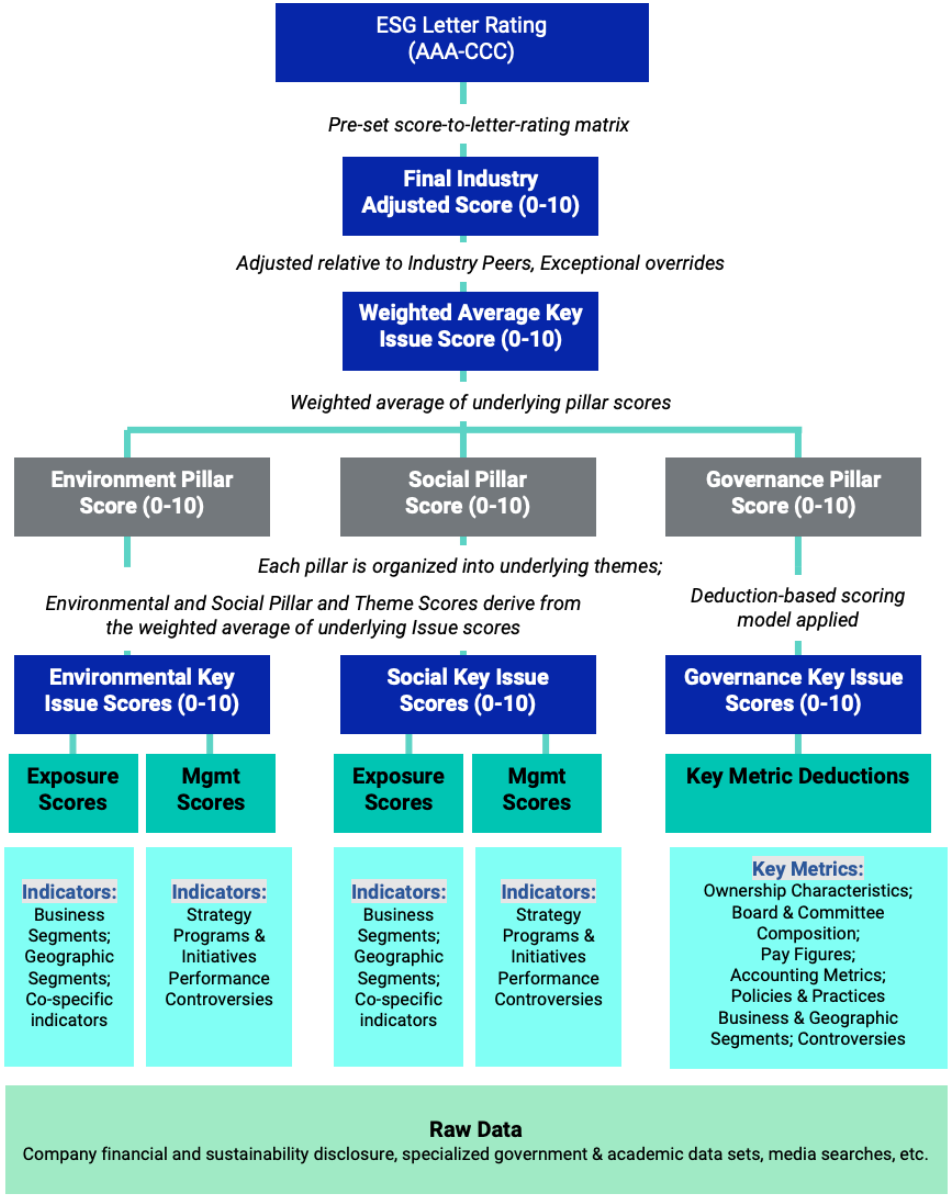
Source: https://www.msci.com/documents/1296102/21901542/ESG-Ratings-Methodology-Exec-Summary.pdf

MSCI Governance Model Structure



Source: https://www.msci.com/documents/1296102/21901542/ESG-Ratings-Methodology-Exec-Summary.pdf

MSCI Hierarchy of ESG Scores



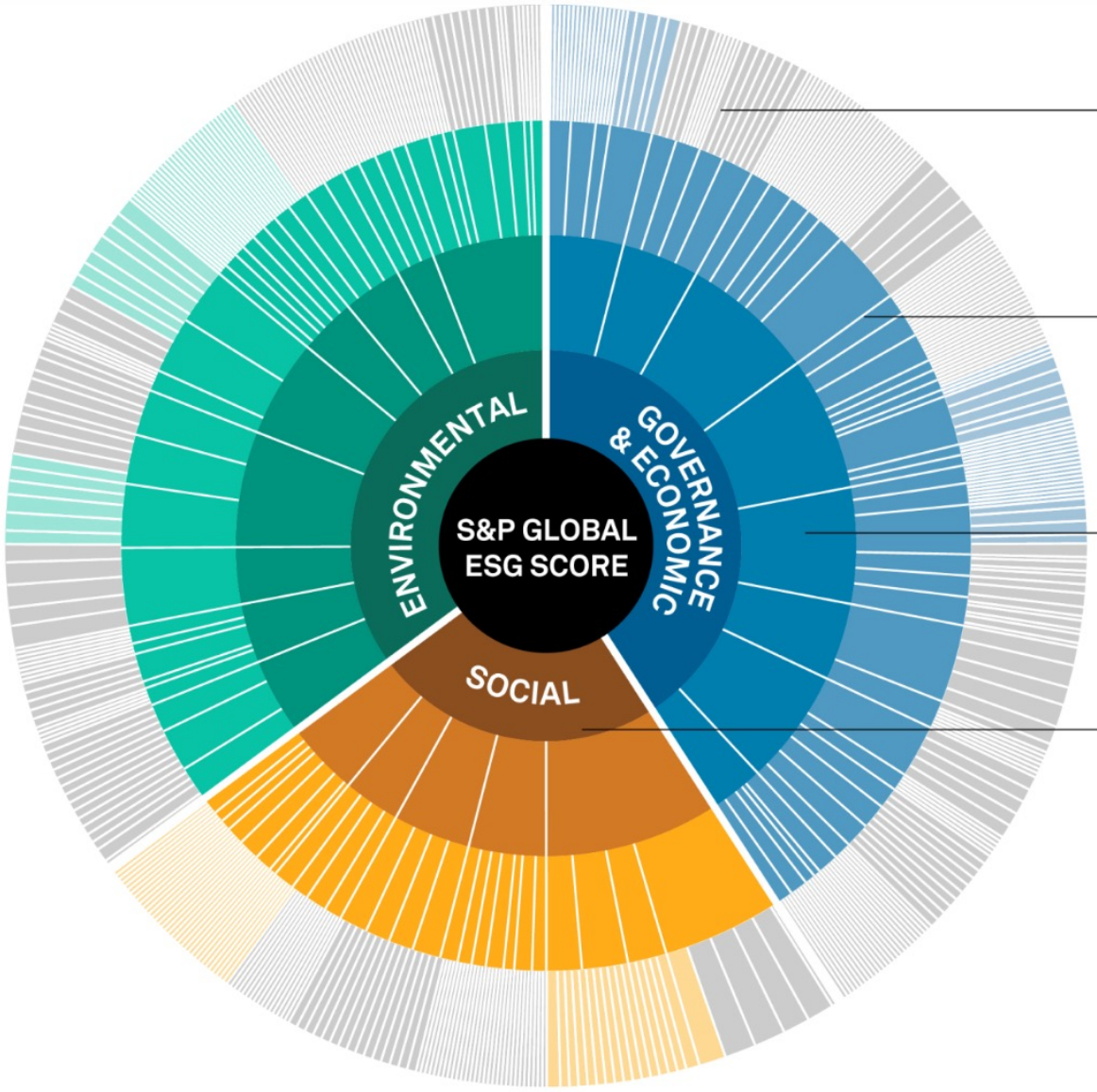
Source: https://www.msci.com/documents/1296102/21901542/ESG-Ratings-Methodology-Exec-Summary.pdf

DJSI S&P Global ESG Score

8,000
Companies

90%
Global market capitalization

340,000+
Current Research Universe and Active Securities



Approx.
1,000
Datapoints

Assessed values, text, checkboxes, documents
Sources: Web-based questionnaire and company documents

130+
Questions

Weighted data point scores
Up to 50% industry-specific

Ave.
30+
Criteria scores

Weighted question scores
61 industry specific approaches, with tailored questions, criteria and related weightings

3
Dimension scores

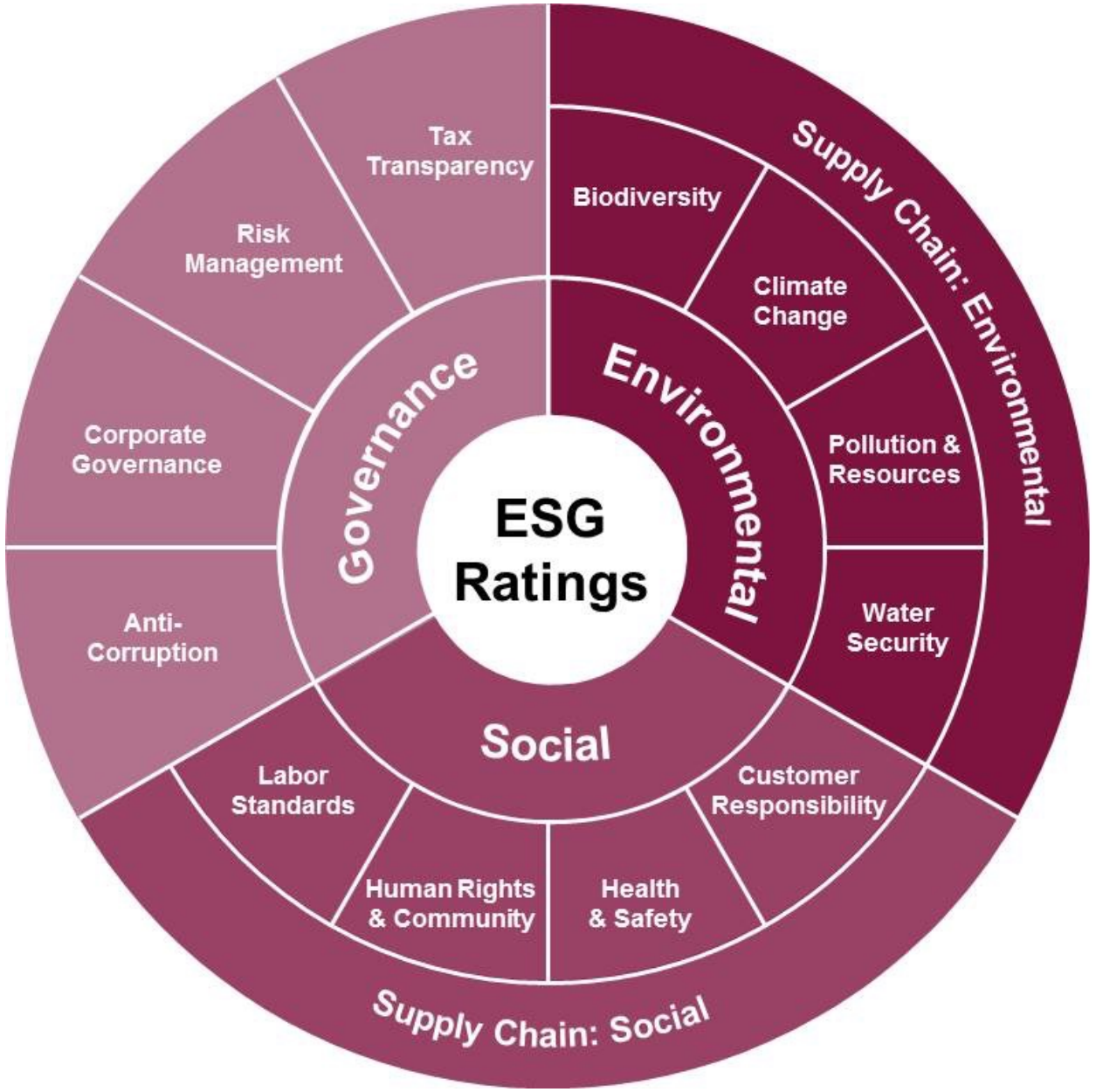
Weighted criteria scores
Adjusted for corporate ESG controversies where applicable

1
S&P Global ESG Score

Sum of weighted dimension scores

Source: DJSI, S&P Global Corporate Sustainability Assessment (CSA), https://www.spglobal.com/esg/solutions/data-intelligence-esg-scores

FTSE Russell ESG Ratings



Source: https://www.ftserussell.com/data/sustainability-and-esg-data/esg-ratings

SUSTAINALYTICS
a Morningstar company

Sustainalytics ESG Risk Ratings

Analyst-based approach

Sustainalytics' ESG Risk Ratings measure a company's exposure to industry-specific material ESG risks and how well a company is managing those risks.

Negligible	Low	Medium	High	Severe
0 - 10	10 - 20	20 - 30	30 - 40	40+

Source: https://www.sustainalytics.com/esg-data

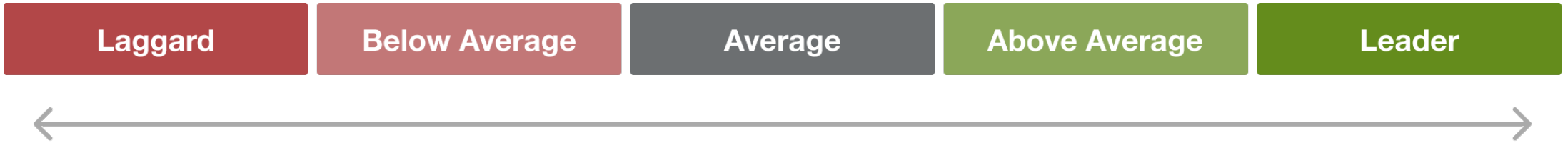
TruValue Labs

FACTSET

Truvalue ESG Ranks

Machine-based approach

- **Truvalue Labs applies AI to analyze over 100,000 sources and uncover ESG risks and opportunities hidden in unstructured text.**
- **The ESG Ranks data service produces an overall company rank based on industry percentile leveraging the 26 ESG categories defined by the Sustainability Accounting Standards Board (SASB).**
- **The data feed covers 20,000+ companies with more than 13 years of history.**



Source: :https://developer.truvaluelabs.com/data/esg-ranks

Analyst-driven vs. AI-driven ESG

Analyst-driven ESG research

Derives ratings in a structured data model

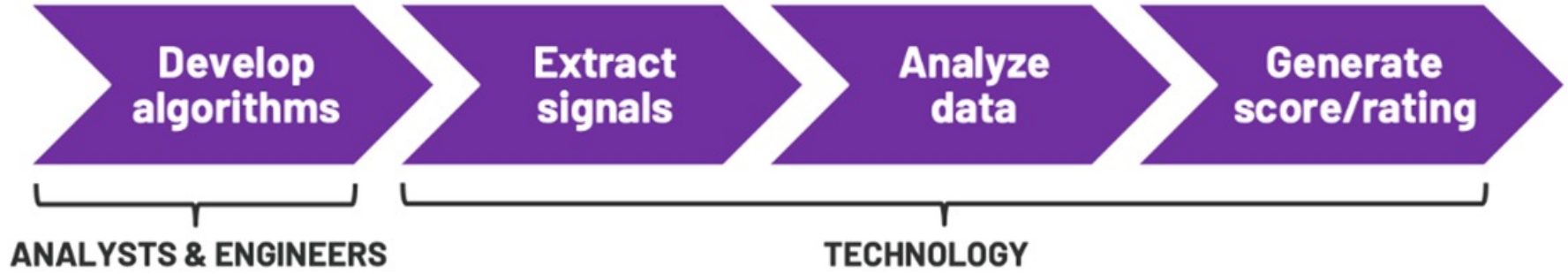
Sustainalytics

Analyst role at the end of the process allows subjectivity to color results

AI-driven ESG research

Derives signals from unstructured data

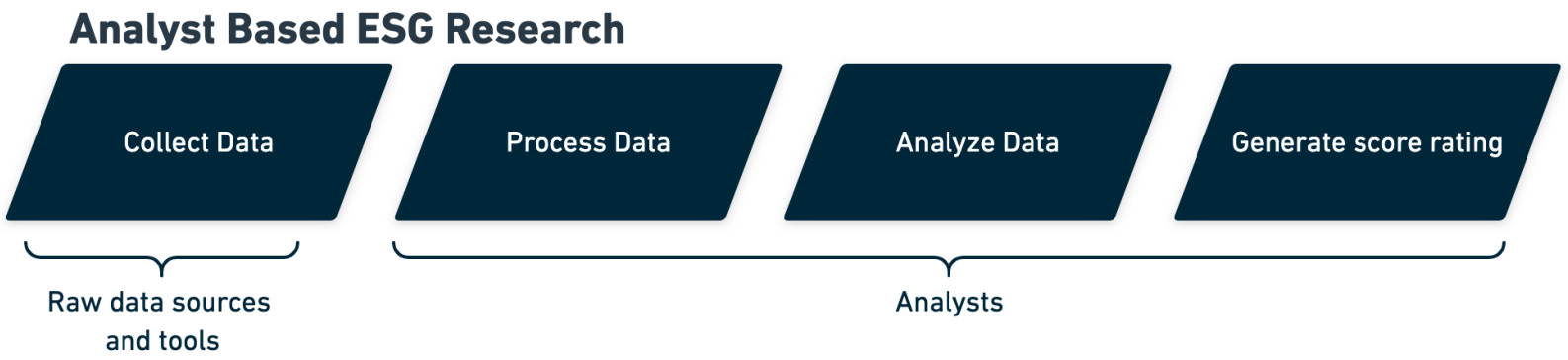
Truvalue Labs



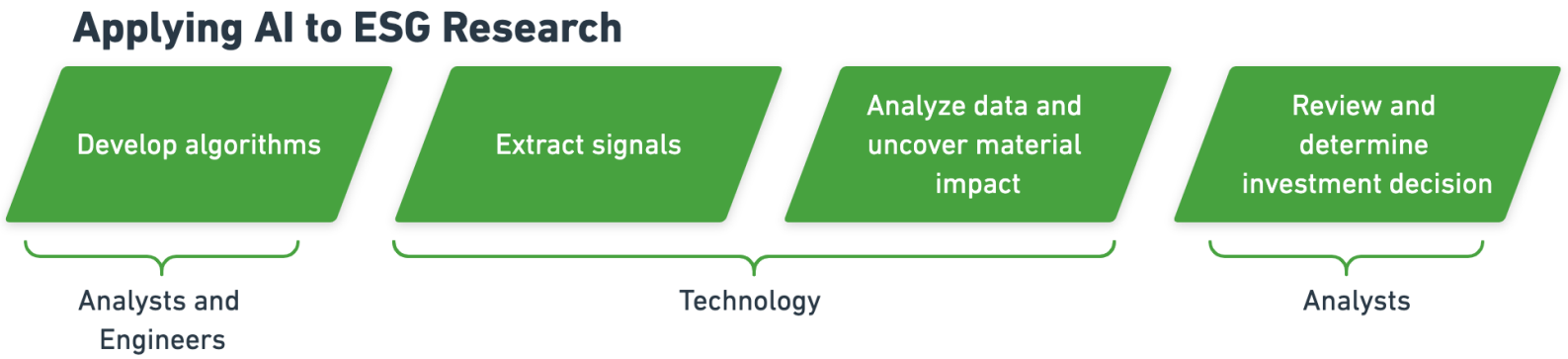
Analyst expertise at the beginning of the process produces consistent results

Source: Mark Tulay (2020), Man vs. machine: A tale of two sustainability ratings systems, GreenBiz, https://www.greenbiz.com/article/man-vs-machine-tale-two-sustainability-ratings-systems

Analyst based ESG Research

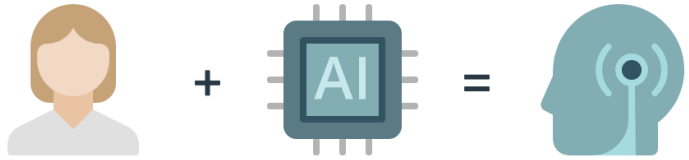


AI based ESG Research



It would take an analyst over 5 years to do what our AI can in 1 week

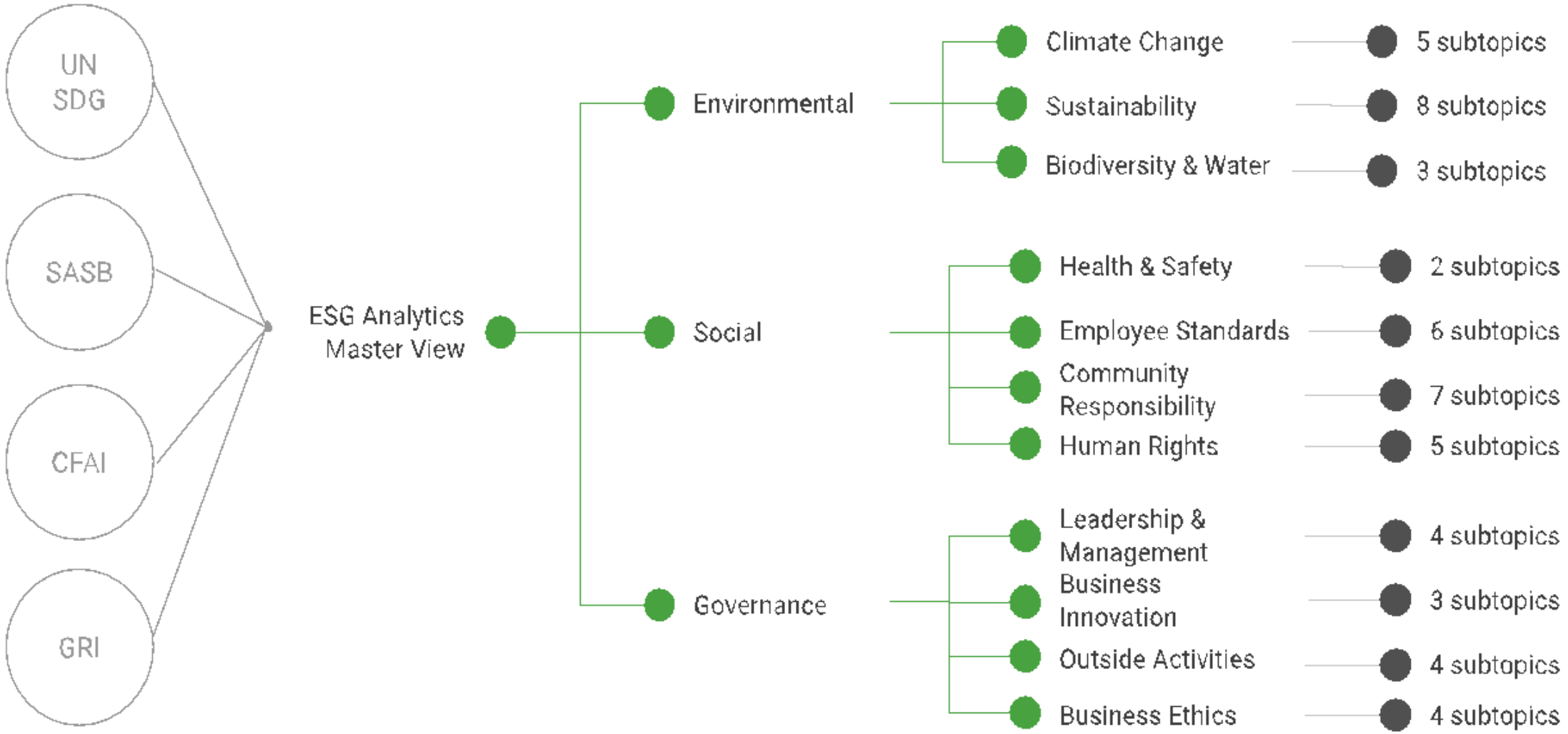
Combining analysts with AI creates gives you the full picture



ESG ANALYTICS
Invest where it matters.

Source: https://www.esganalytics.io/insights/how-data-is-accelerating-esg

ESG Analytics: NLP Taxonomy

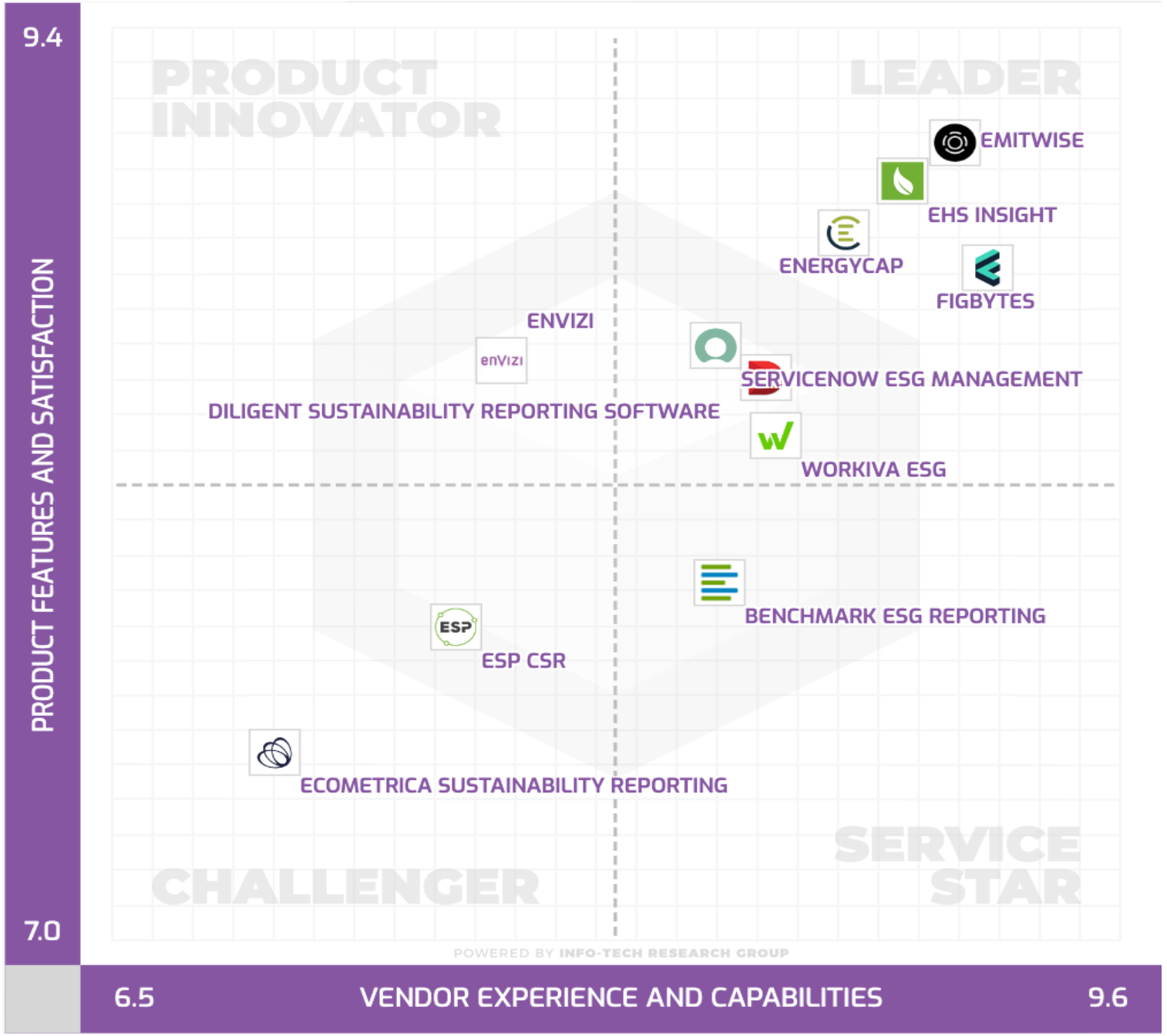


ESG ANALYTICS

Source: https://www.esganalytics.io/insights/how-data-is-accelerating-esg

Top ESG Reporting Software

Environmental, Social and Governance (ESG) Reporting software or **Sustainability software** helps organizations manage their operational data, evaluate their impact on the environment and provide reporting to perform audits.



Source: https://www.softwarereviews.com/categories/environmental-social-and-governance-reporting

ESG Reporting Software: Emitwise

- **Emitwise is the carbon management platform for companies with complex manufacturing supply chains to confidently understand, track and reduce their complete carbon footprint.**
- **Combining 100 years of carbon accounting experience and machine learning technology, we accelerate climate action by increasing the accuracy of scope 3 emissions.**
- **The platform empowers manufacturers and their supply chains to make carbon-led business decisions that lower risk, increase profitability and deliver ambitious climate action.**

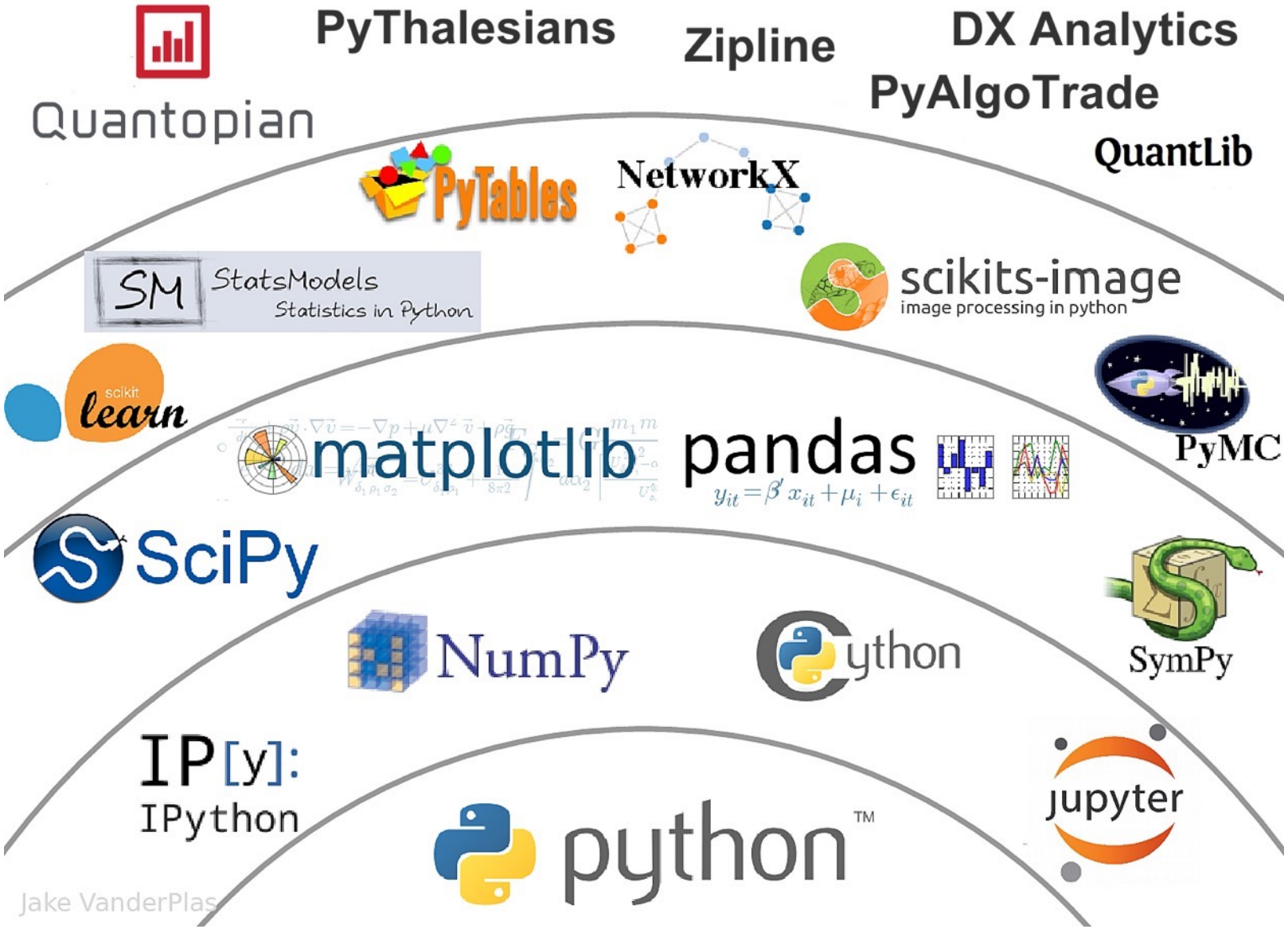
Source: https://www.softwarereviews.com/categories/environmental-social-and-governance-reporting

ESG Reporting Software: Workiva ESG

- **Workiva is a cloud native platform that simplifies the complexities of reporting and compliance.**
- **Workiva ESG is the end-to-end platform that allows you to integrate financial data, nonfinancial data, and XBRL.**
- **Workiva, the platform that streamlines your entire ESG process.**
- **Automate data collection, utilize frameworks, and directly connect to all your ESG reports. in meaningful glossy reports, accurate survey responses, and regulatory filings with integrated XBRL tagging.**

Source: https://www.softwarereviews.com/categories/environmental-social-and-governance-reporting

The Quant Finance PyData Stack



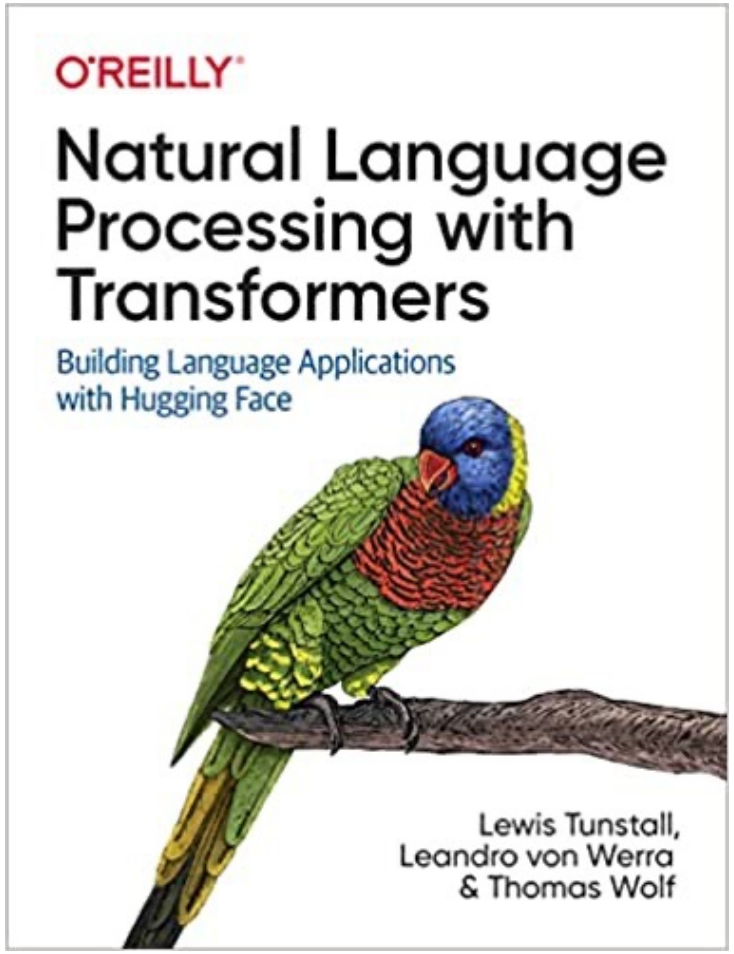
Source: http://nbviewer.jupyter.org/format/slides/github/quantopian/pyfolio/blob/master/pyfolio/examples/overview_slides.ipynb#/5

Lewis Tunstall, Leandro von Werra, and Thomas Wolf (2022),

Natural Language Processing with Transformers:

Building Language Applications with Hugging Face,

O'Reilly Media.



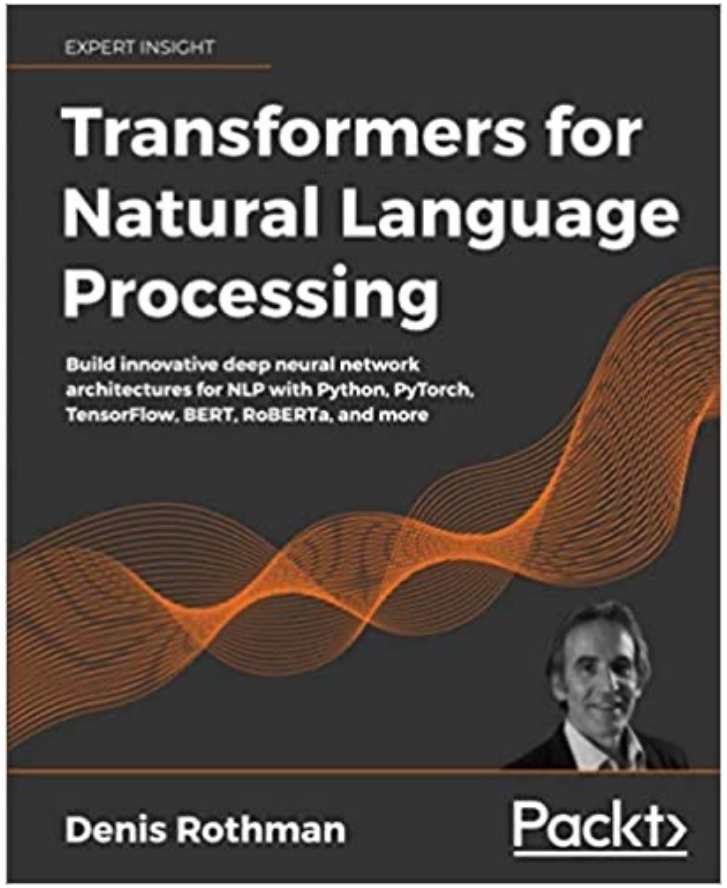
Source: https://www.amazon.com/Natural-Language-Processing-Transformers-Applications/dp/1098103246

Denis Rothman (2021),

Transformers for Natural Language Processing:

Build innovative deep neural network architectures for NLP with Python, PyTorch, TensorFlow, BERT, RoBERTa, and more,

Packt Publishing.



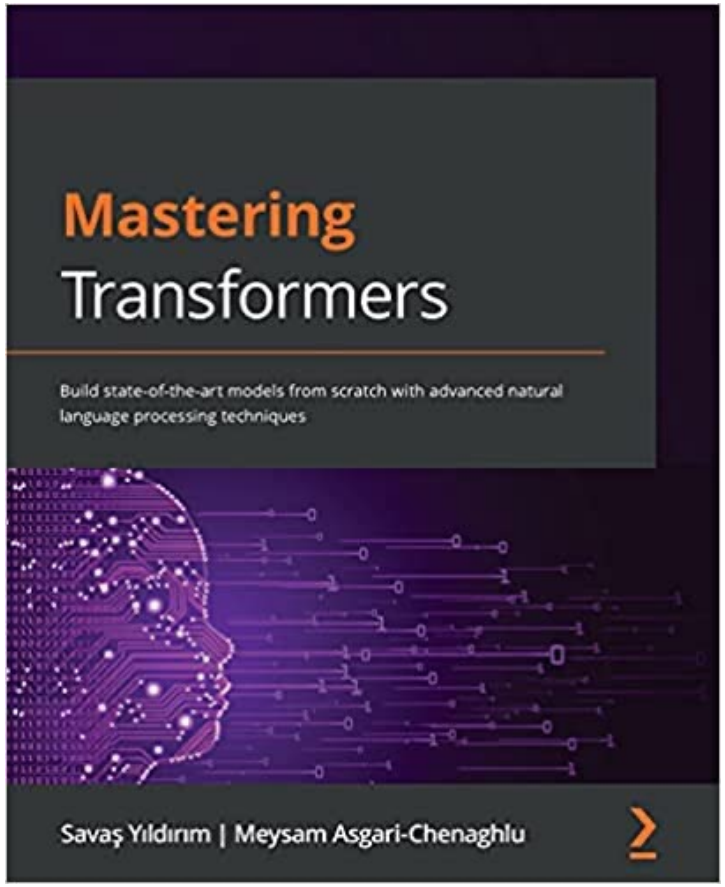
Source: https://www.amazon.com/Transformers-Natural-Language-Processing-architectures/dp/1800565798

Savaş Yıldırım and Meysam Asgari-Chenaghlu (2021),

Mastering Transformers:

Build state-of-the-art models from scratch with advanced natural language processing techniques,

Packt Publishing.



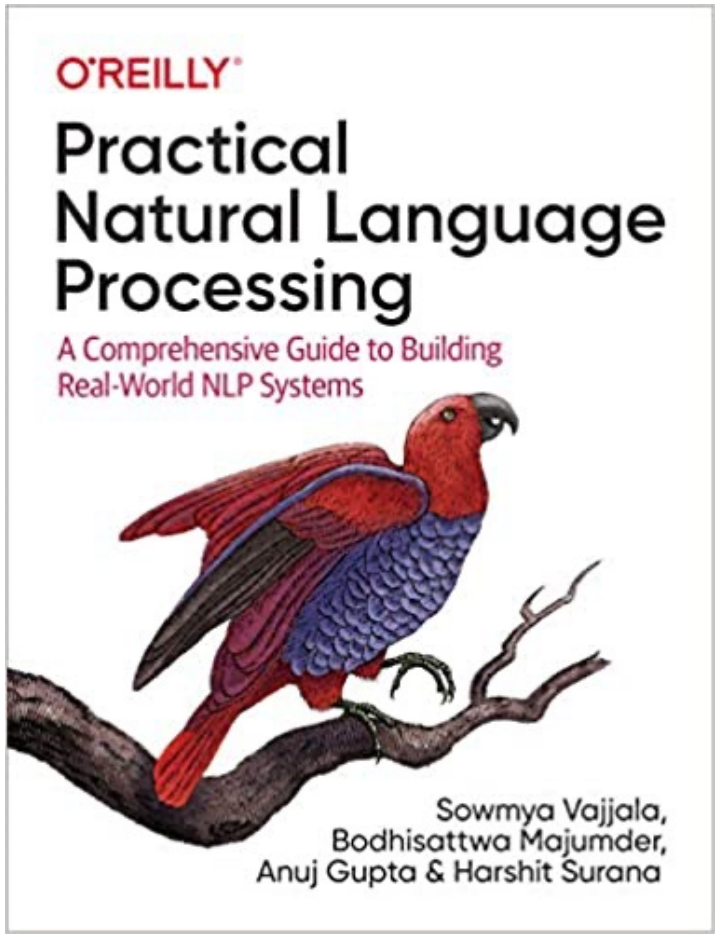
Source: https://www.amazon.com/Mastering-Transformers-state-art-processing/dp/1801077657/

Sowmya Vajjala, Bodhisattwa Majumder, Anuj Gupta (2020),

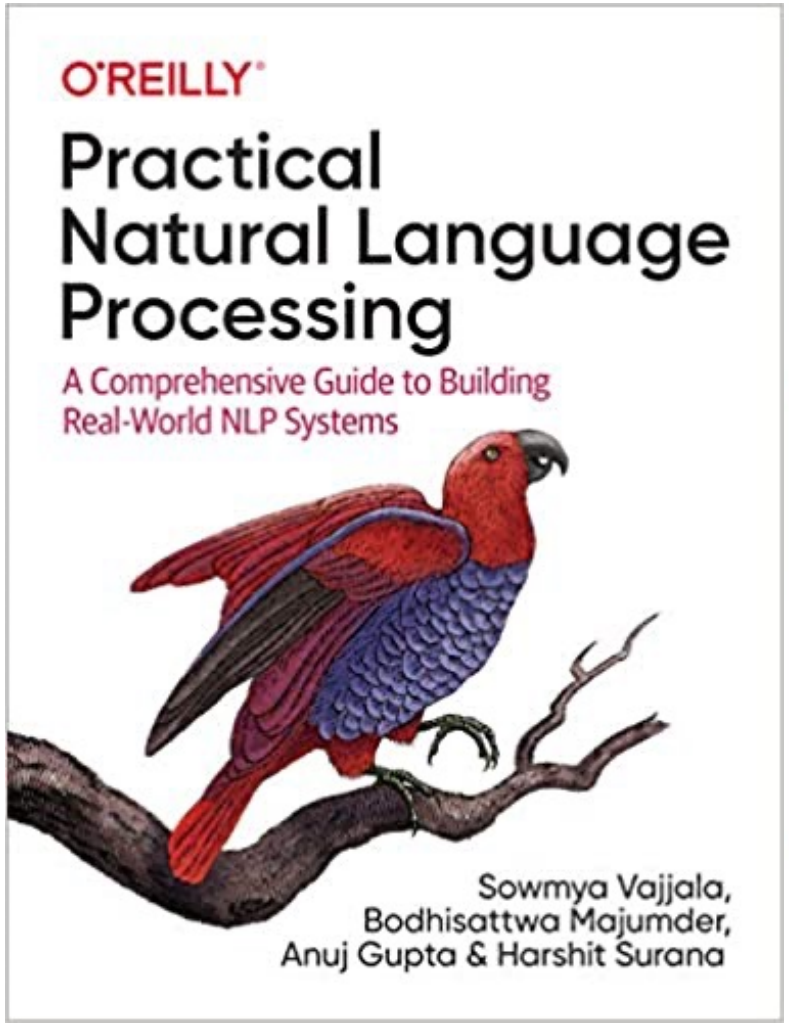
Practical Natural Language Processing:

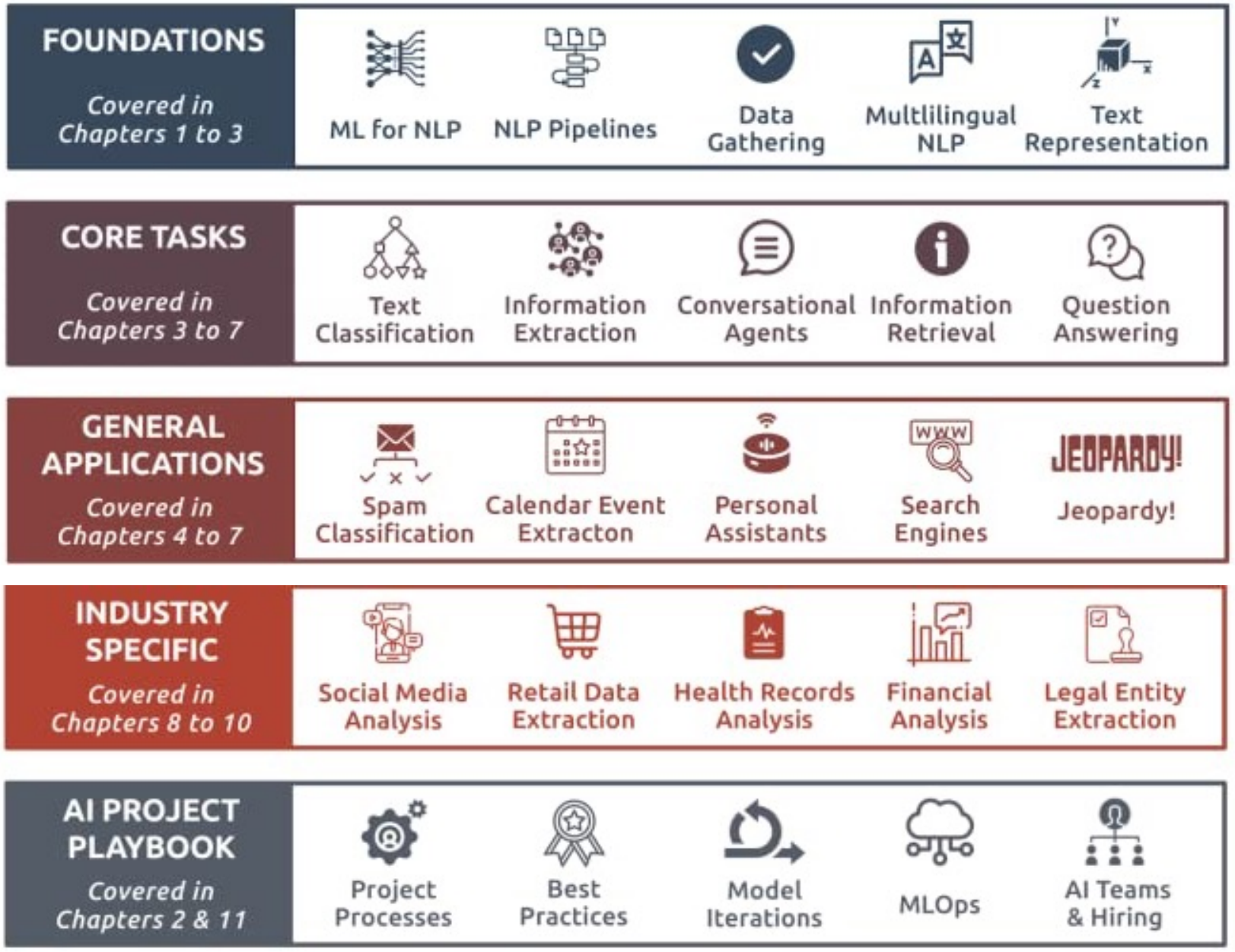
A Comprehensive Guide to Building Real-World NLP Systems,

O'Reilly Media.



Source: https://www.amazon.com/Practical-Natural-Language-Processing-Pragmatic/dp/1492054054

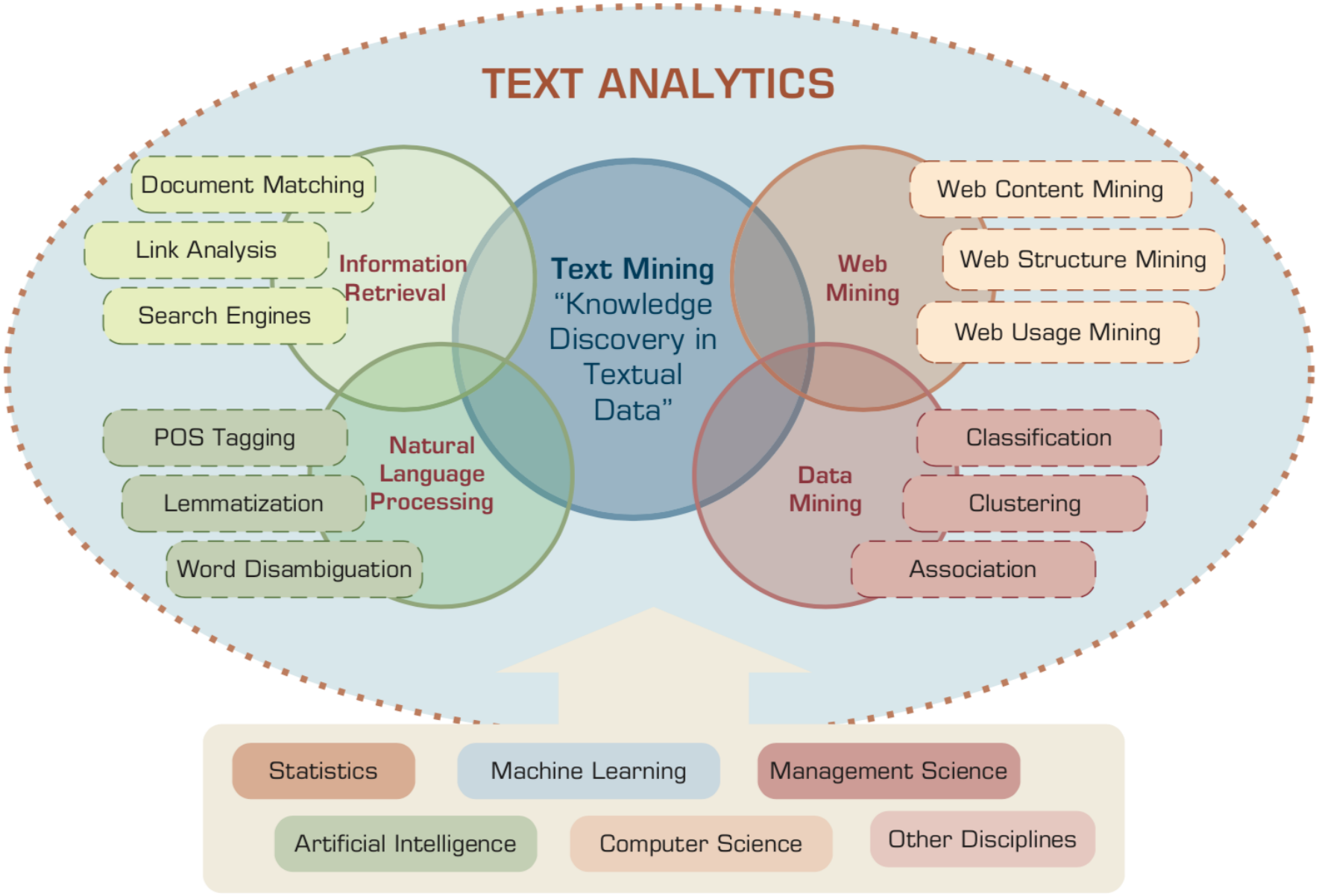




Source: Sowmya Vajjala, Bodhisattwa Majumder, Anuj Gupta (2020), Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems, O'Reilly Media.

Source: https://www.amazon.com/Practical-Natural-Language-Processing-Pragmatic/dp/1492054054

Text Analytics and Text Mining



Source: Ramesh Sharda, Dursun Delen, and Efraim Turban (2017), Business Intelligence, Analytics, and Data Science: A Managerial Perspective, 4th Edition, Pearson

Generative AI

Text, Image, Video, Audio Applications

Generative AI (Gen AI)

AI Generated Content (AIGC)

Image Generation

Instruction 1:
An astronaut riding a horse in a photorealistic style.

Instruction 2:
Teddy bears working on new AI research on the moon in the 1980s.

Figure 1

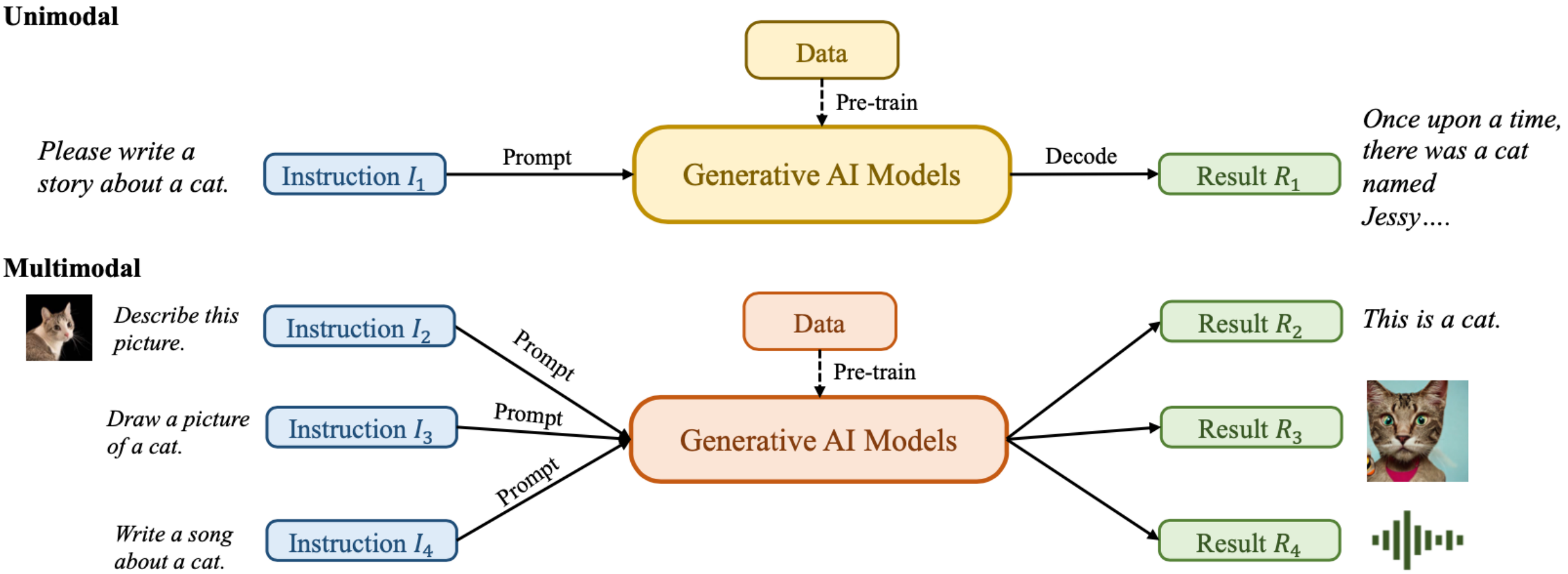
Figure 2

OpenAI DALL·E 2

Source: Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.

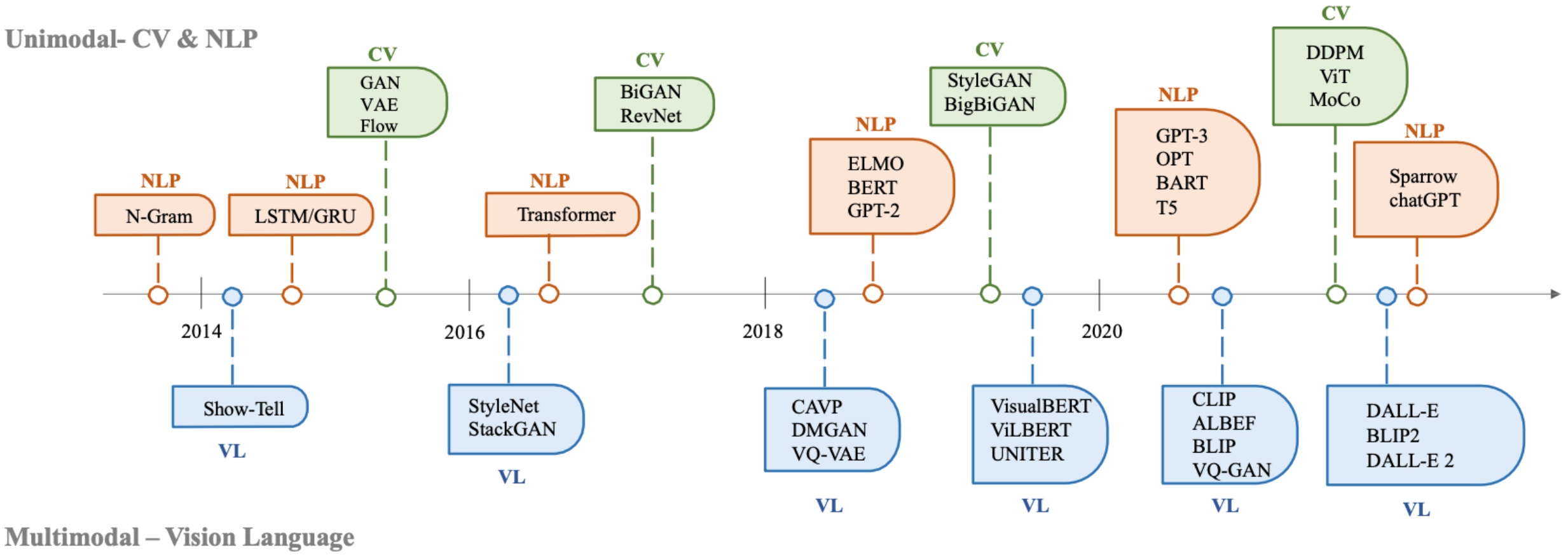
Generative AI (Gen AI)

AI Generated Content (AIGC)



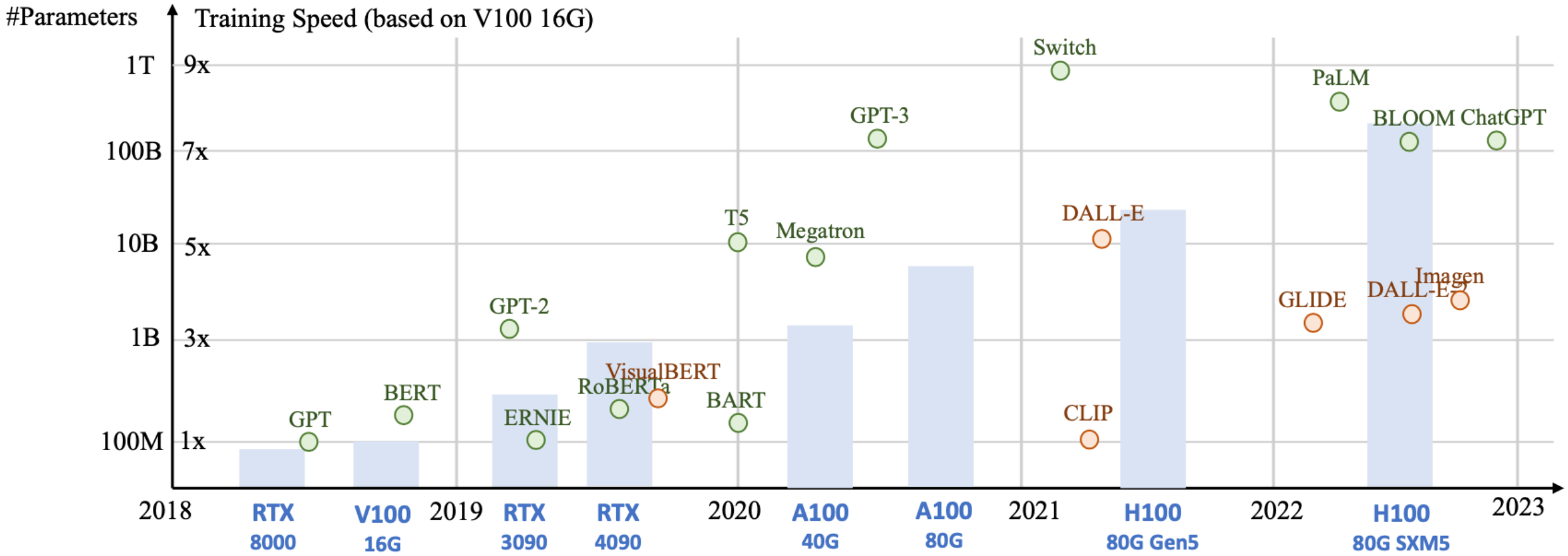
Source: Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.

The history of Generative AI in CV, NLP and VL



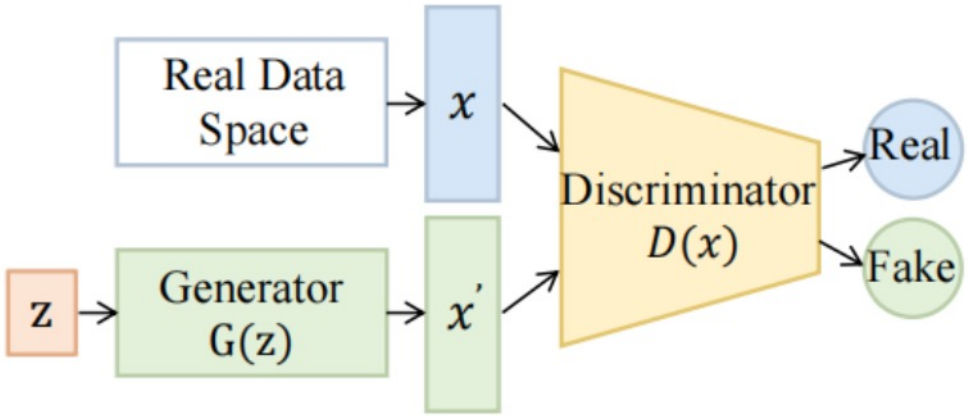
Source: Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.

Generative AI Foundation Models

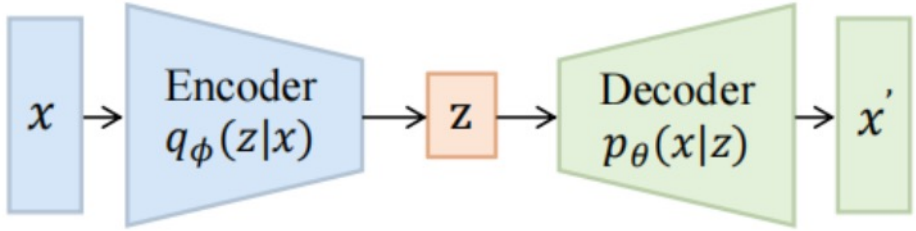


Source: Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.

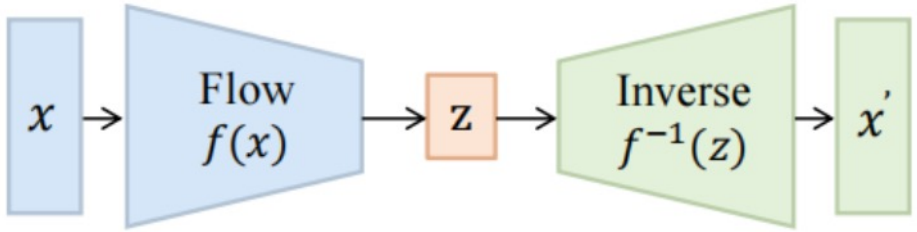
Categories of Vision Generative Models



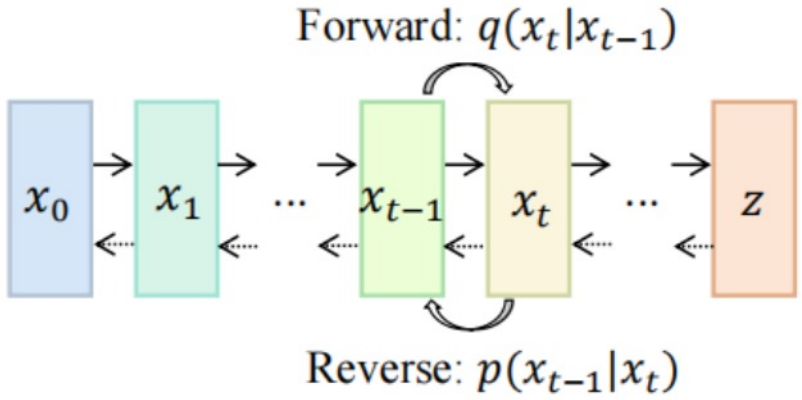
(1) Generative adversarial networks



(2) Variational autoencoders



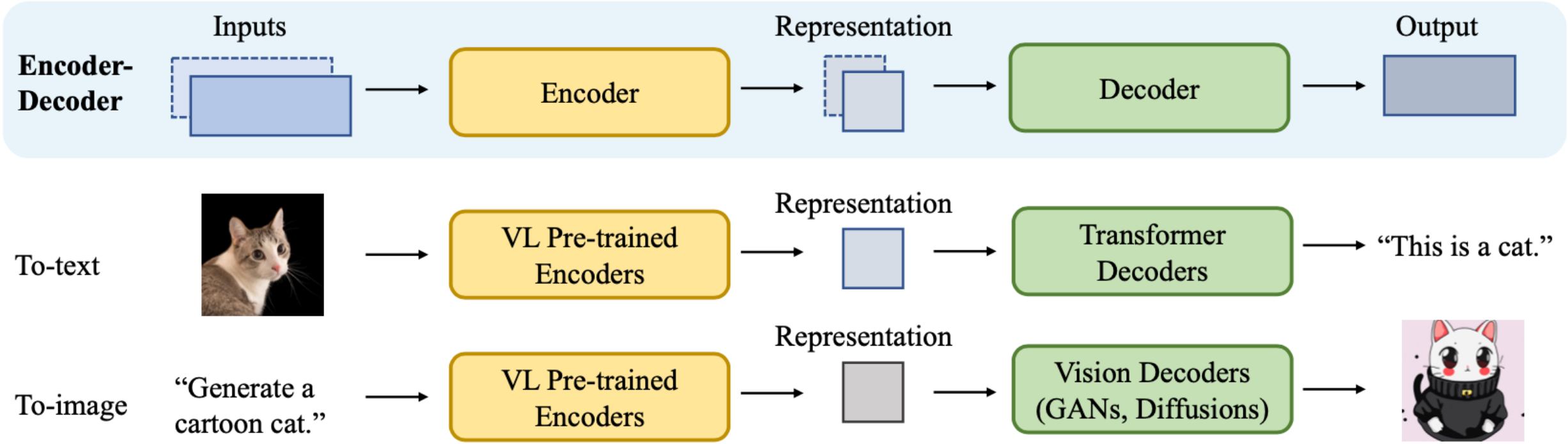
(3) Normalizing flows



(4) Diffusion models

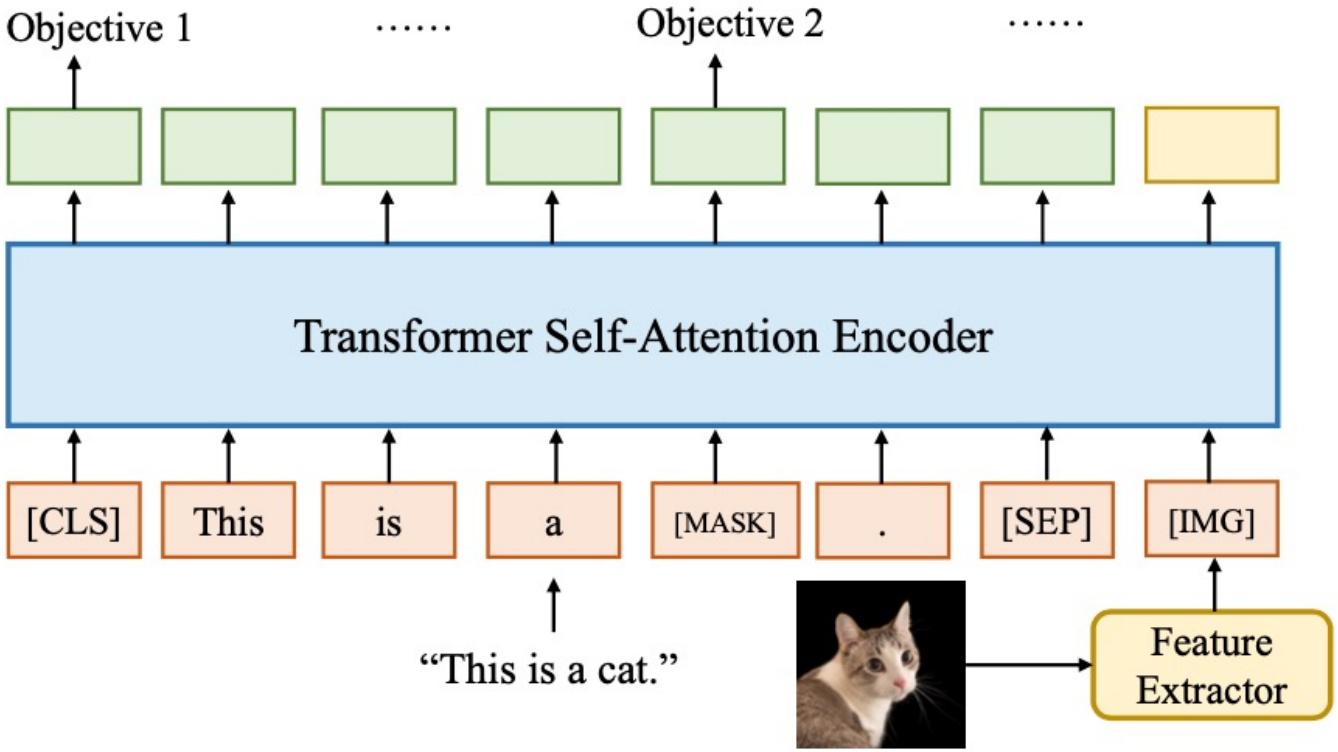
Source: Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.

The General Structure of Generative Vision Language

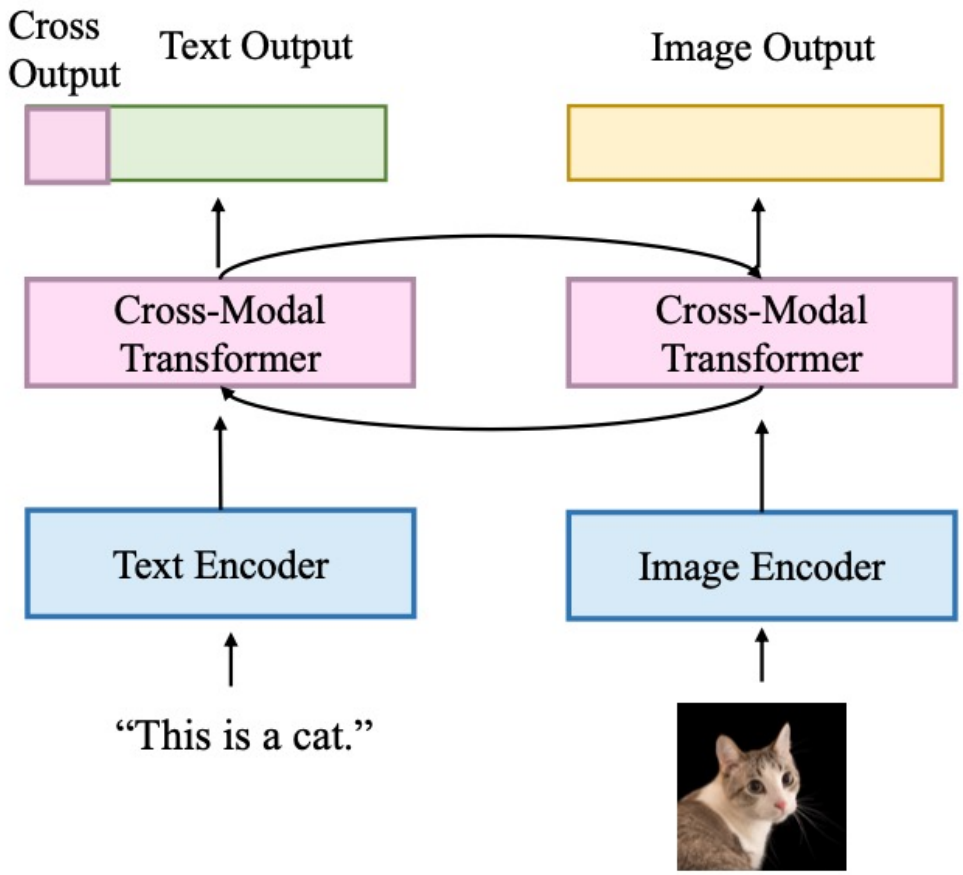


Source: Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.

Two Types of Vision Language Encoders:
Concatenated Encoders and Cross-aligned Encoders



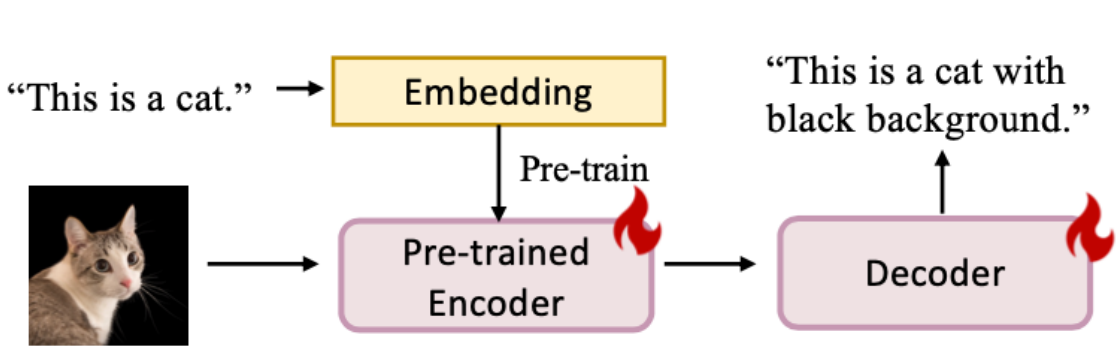
(a) Concatenated Encoder



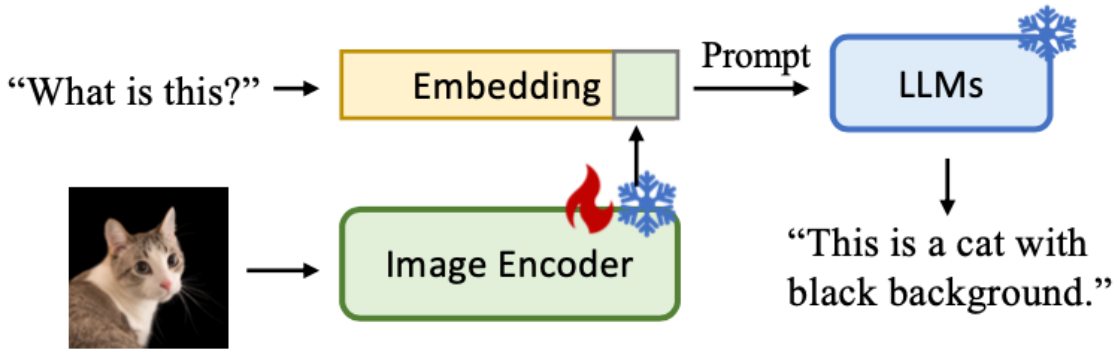
(b) Cross-aligned Encoder

Source: Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.

Two Types of to-language Decoder Models: Jointly-trained Models and Frozen Models



(a) Jointly-trained Models



(b) Frozen Models

Source: Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.

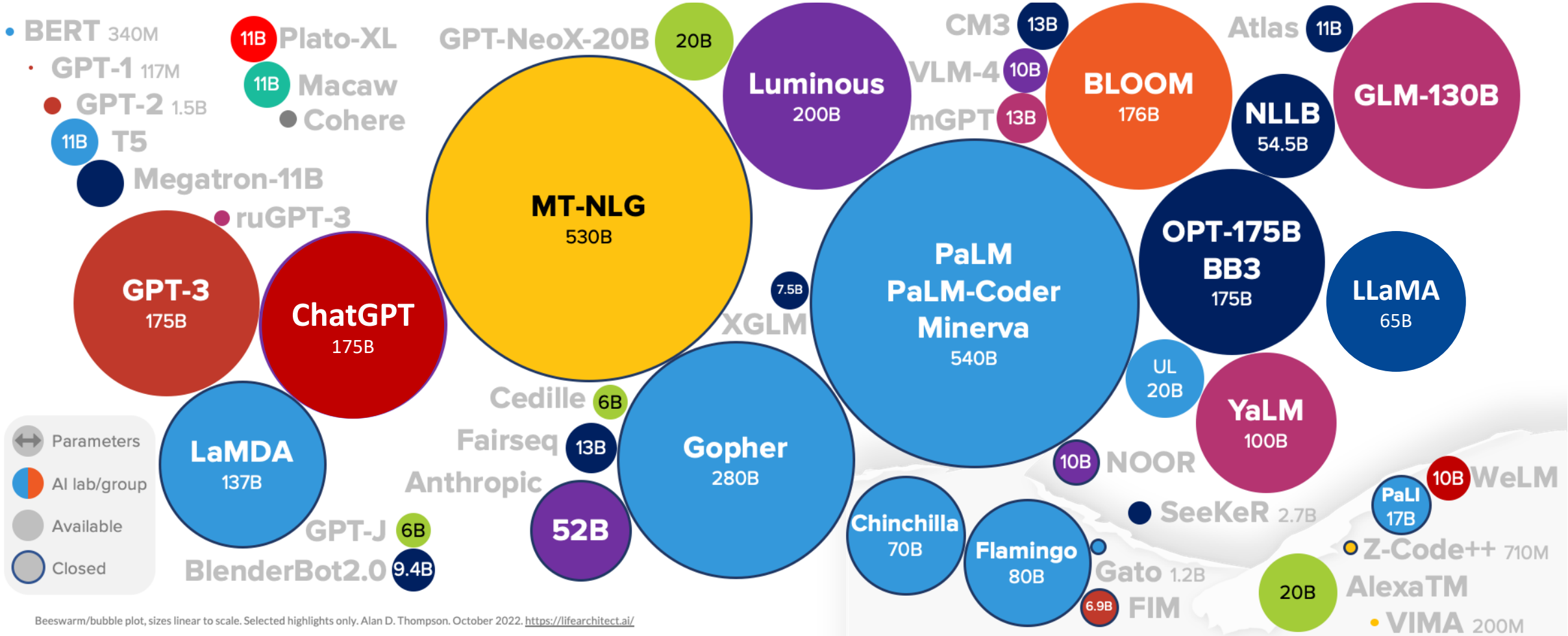
ChatGPT

Large Language Models (LLMs)

Foundation Models

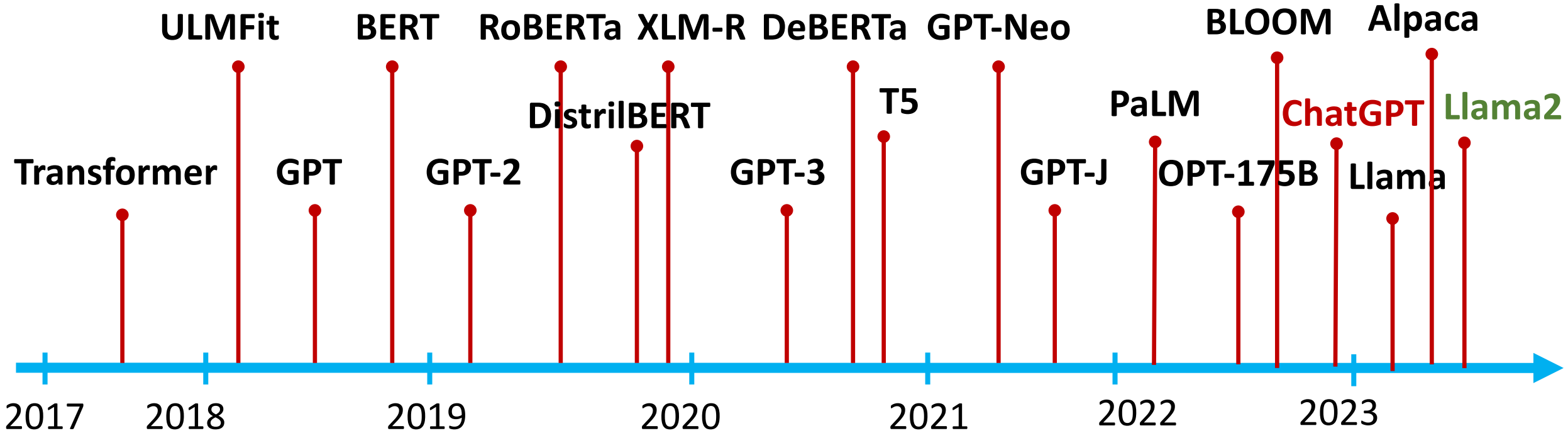
Large Language Models (LLM)

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



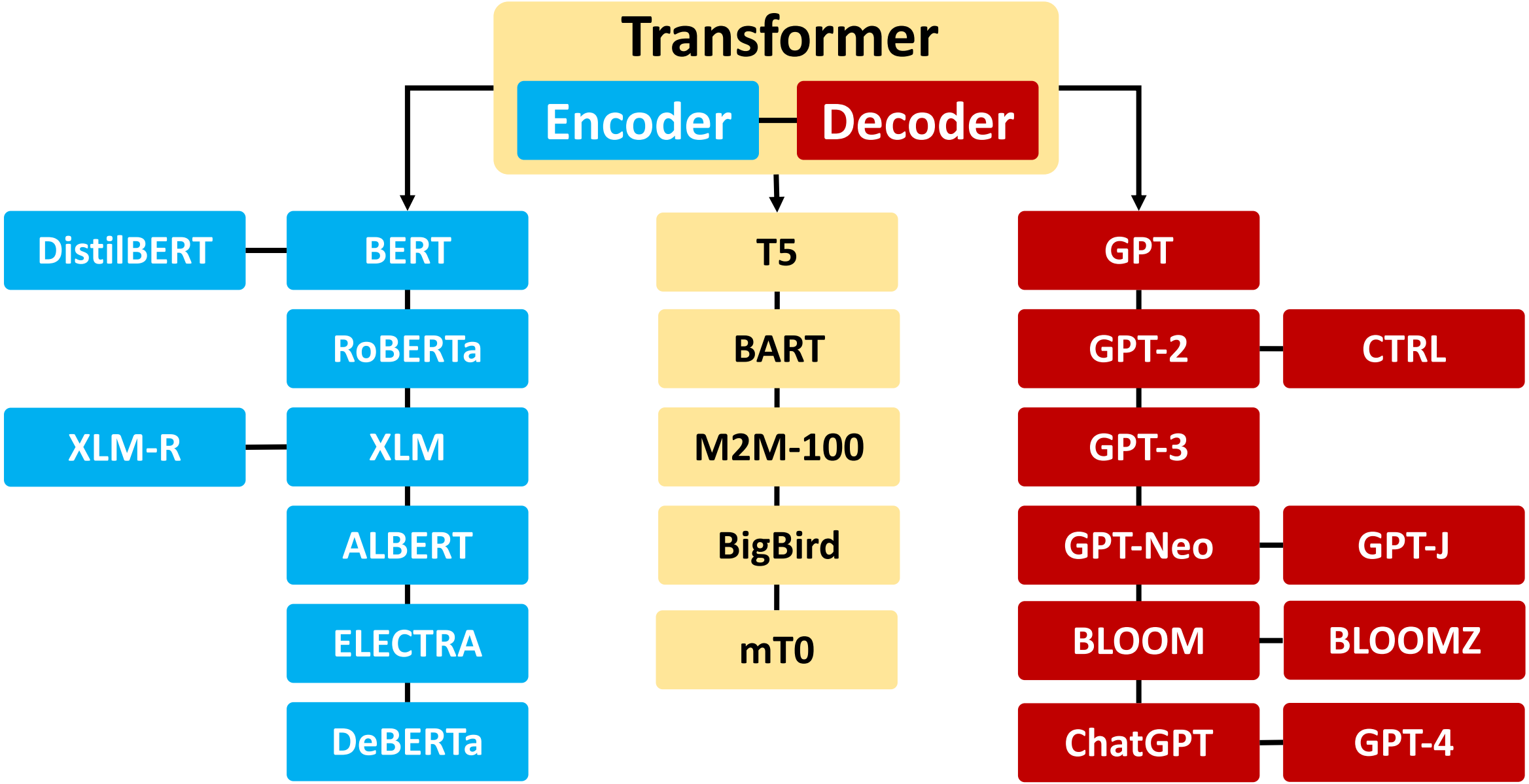
Source: https://lifearchitect.ai/models/

The Transformers Timeline



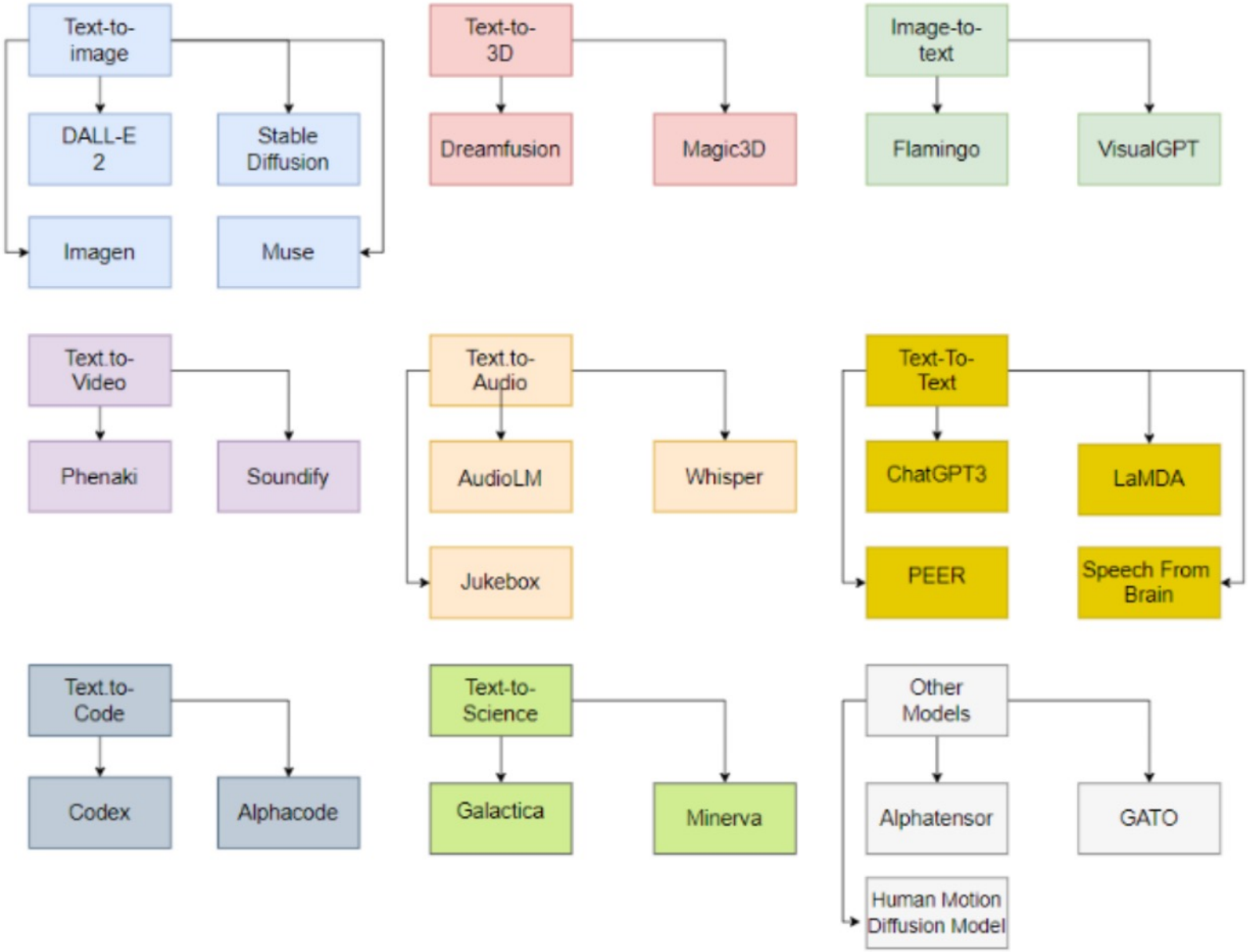
Source: Lewis Tunstall, Leandro von Werra, and Thomas Wolf (2022), Natural Language Processing with Transformers: Building Language Applications with Hugging Face, O'Reilly Media.

Transformer Models



Source: Lewis Tunstall, Leandro von Werra, and Thomas Wolf (2022), Natural Language Processing with Transformers: Building Language Applications with Hugging Face, O'Reilly Media.

Generative AI Models

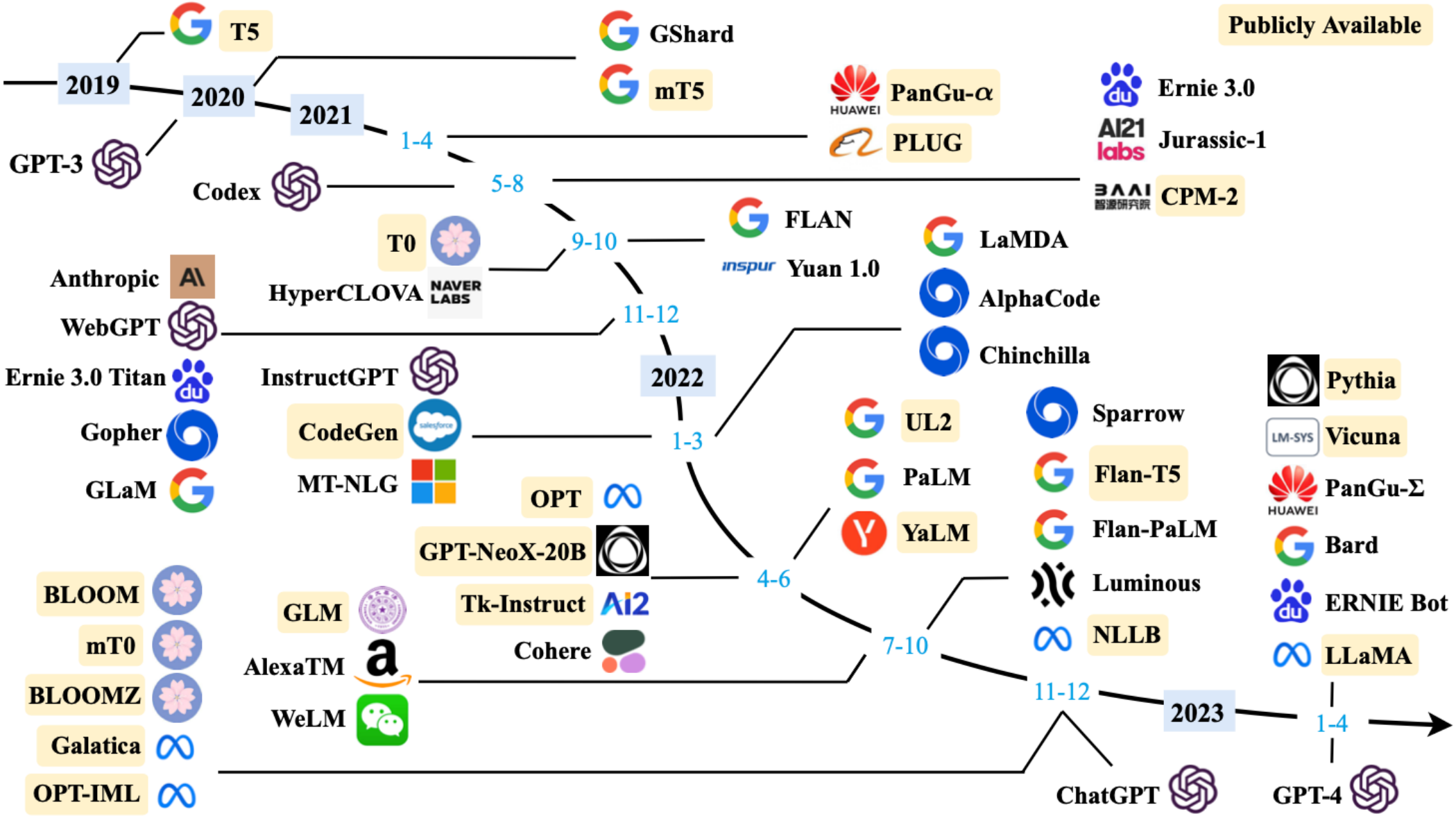


ChatGPT is not all you need

Attention is all you need

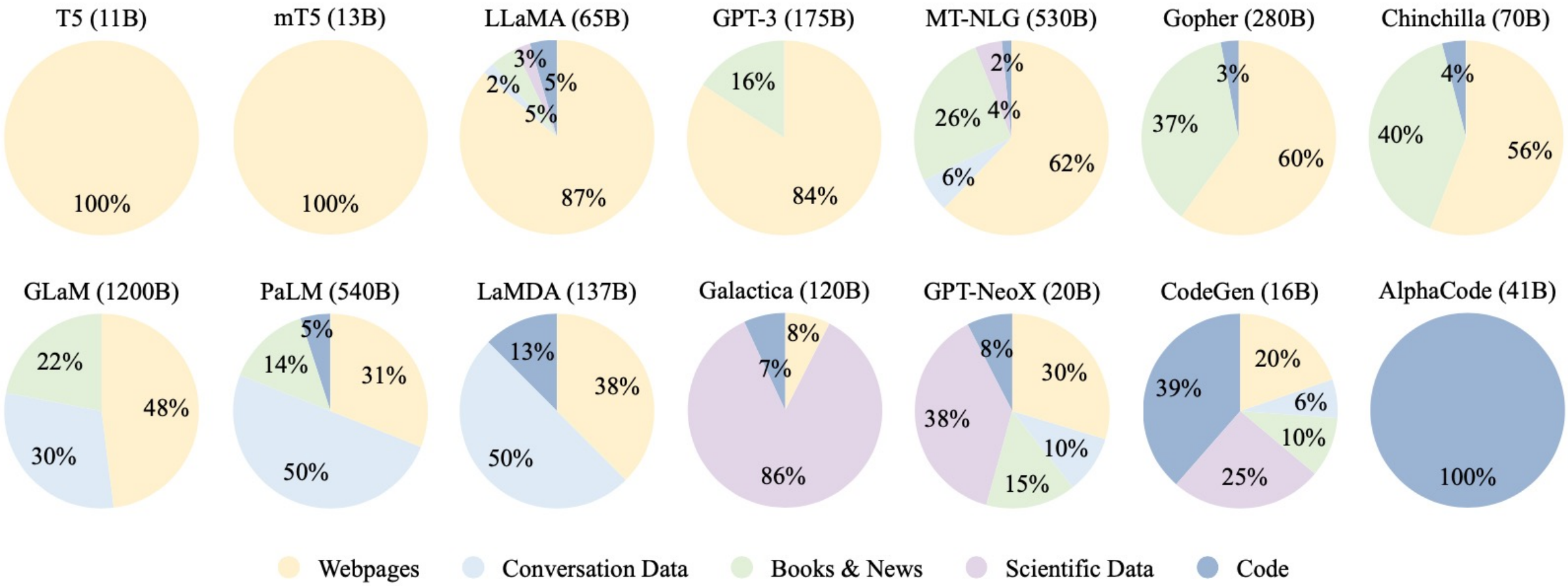
Source: Gozalo-Brizuela, Roberto, and Eduardo C. Garrido-Merchan (2023). "ChatGPT is not all you need. A State of the Art Review of large Generative AI models." arXiv preprint arXiv:2301.04655 (2023).

Large Language Models (LLMs) (larger than 10B)



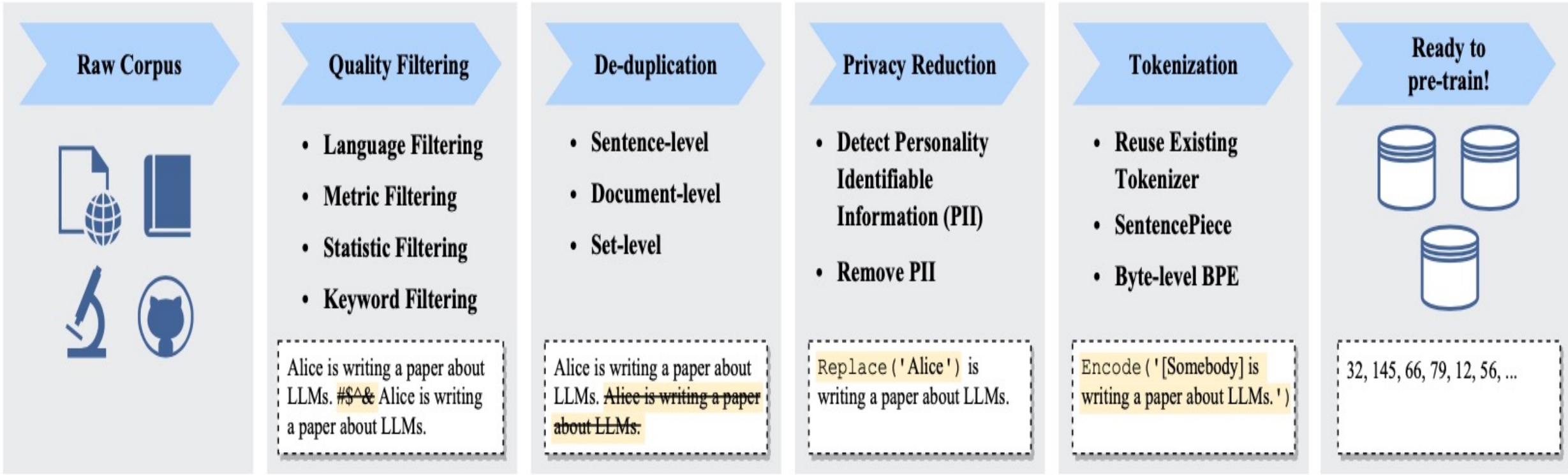
Source: Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min et al. (2023) "A Survey of Large Language Models." arXiv preprint arXiv:2303.18223.

Ratios of various data sources in the pre-training data for existing LLMs



Source: Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min et al. (2023) "A Survey of Large Language Models." arXiv preprint arXiv:2303.18223.

Typical Data Preprocessing Pipeline for Pre-training Large Language Models (LLMs)



Source: Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min et al. (2023) "A Survey of Large Language Models." arXiv preprint arXiv:2303.18223.

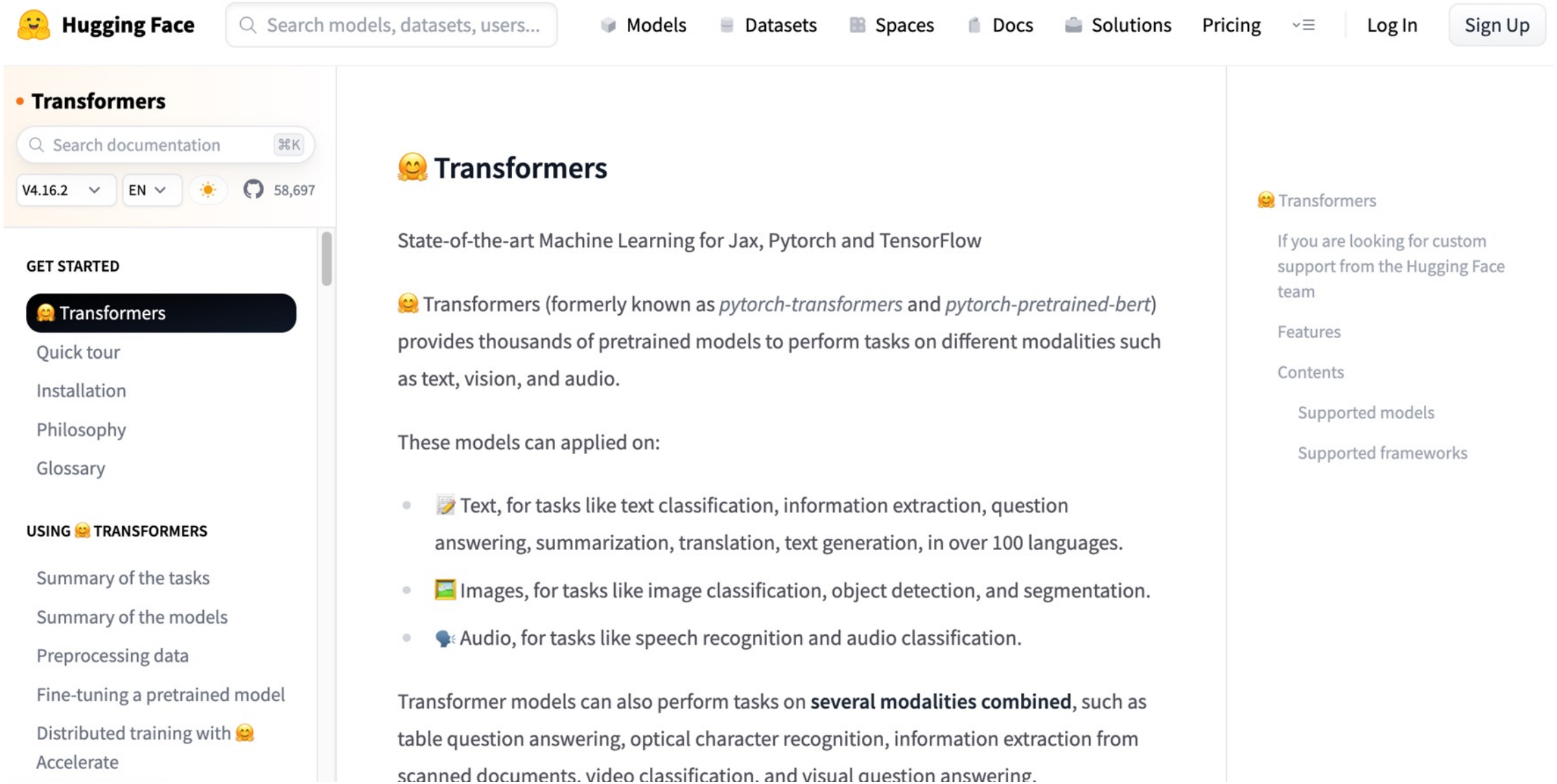
Hugging Face

The AI community building the future.

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

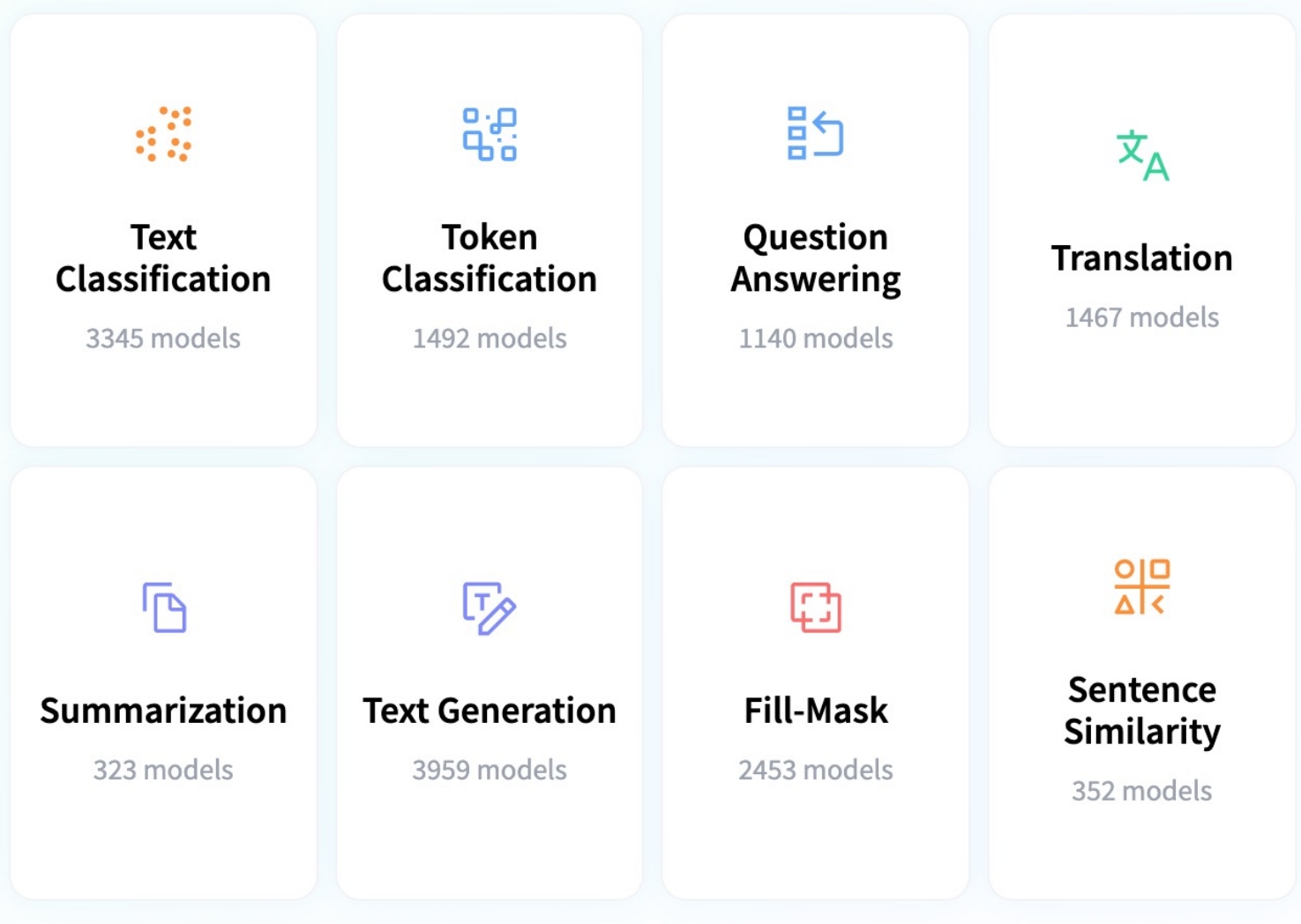
https://huggingface.co/

Hugging Face Transformers



https://huggingface.co/docs/transformers/index

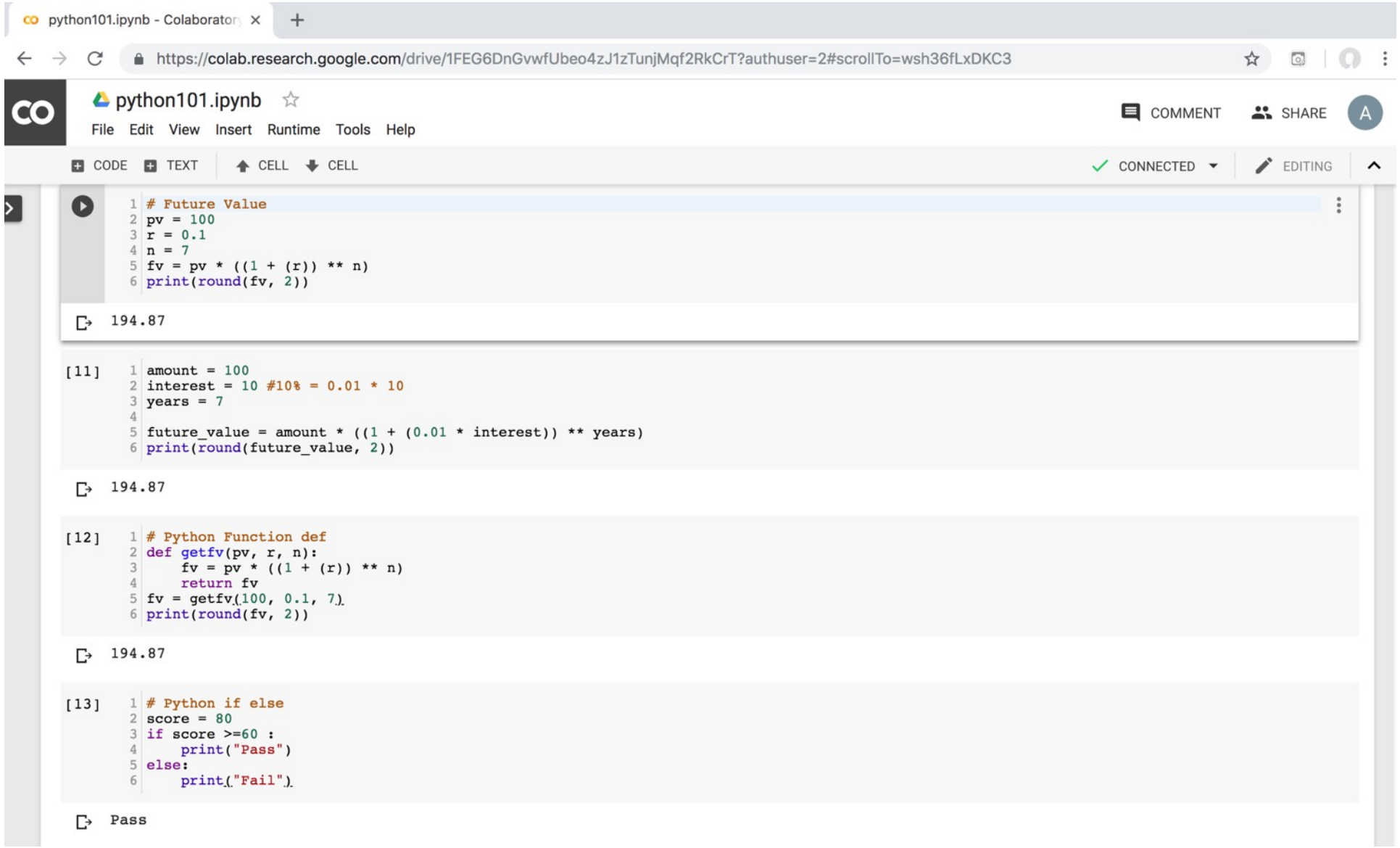
Hugging Face Tasks
Natural Language Processing



https://huggingface.co/tasks

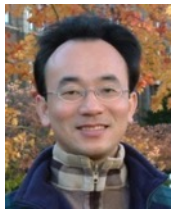
Python in Google Colab (Python101)

https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT



https://tinyurl.com/aintpupython101

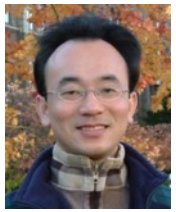
Teaching



- Artificial Intelligence in Finance and Quantitative
  - Fall 2021, Fall 2022, Fall 2023
- Artificial Intelligence for Text Analytics
  - Spring 2022, Fall 2023
- Big Data Analytics
  - Fall 2020, Spring 2023
- Software Engineering
  - Fall 2020, Fall, 2021, Spring 2022, Spring 2023
- Artificial Intelligence
  - Spring 2021, Fall 2022
- Data Mining
  - Spring 2021
- Foundation of Business Cloud Computing
  - Spring 2021, Spring 2022, Spring 2023
- Python for Accounting Applications
  - Fall 2023

https://web.ntpu.edu.tw/~myday/teaching.htm

Research Projects



1. **Applying AI technology to construct knowledge graphs of cryptocurrency anti-money laundering: a few-shot learning model**
   - MOST, 110-2410-H-305-013-MY2, 2021/08/01~2023/07/31
2. **Fintech Green Finance for Carbon Market Index, Corporate Finance, and Environmental Policies.** Carbon Emission Sentiment Index with AI Text Analytics
   - NTPU, 112-NTPU_ORDA-F-003，2023/01/01~2024/12/31
3. **Digital Support, Unimpeded Communication: The Development, Support and Promotion of AI-assisted Communication Assistive Devices for Speech Impairment.** Multimodal Cross-lingual Task-Oriented Dialogue System for Inclusive Communication Support
   - NSTC 112-2425-H-305-002-, 2023/05/01-2026/04/30
4. **Establishment and Implement of Smart Assistive Technology for Dementia Care and Its Socio-Economic Impacts.** Intelligent, individualized and precise care with smart AT and system integration
   - NSTC, NSTC, 112-2627-M-038-001-, 2023/08/01~2024/07/31
5. **Use deep learning to identify commercially dental implant systems - observational study**
   - USTP-NTPU-TMU, USTP-NTPU-TMU-112-01, 2023/01/01~2023/12/31
6. **Metaverse AI Multimodal Cross-Language Task-Oriented Dialogue System**
   - ATEC Group x NTPU, NTPU-112A413E01, 2023/05/01~2026/04/30
7. **Metaverse Avatar Automatic Metadata Generation Module**
   - FormosaVerse x NTPU, NTPU-111A413E01, 2022/12/01~2023/11/30
8. **Pilot Study on Universal Data Processing for Code Generation Engine**
   - III x NTPU, NTPU-112A513E01, 2023/08/01~2023/12/22

https://web.ntpu.edu.tw/~myday/cindex.htm#projects

Summary

- This course introduces the fundamental concepts and hands-on practices of Python for Accounting Applications.
- Topics include
  - Introduction to Python for Accounting Applications,
  - Python Programming and Data Science,
  - Foundations of Python Programming,
  - Data Structures, Control Logic and Loops,
  - Functions and Modules,
  - Files and Exception Handling,
  - Data Analytics and Visualization with Python,
  - Obtaining Data From the Web with Python,
  - Statistical Analysis with Python, Machine Learning with Python,
  - Text Analytics with Python with LLMs,
  - Applications of Accounting Data Analytics with Python, and
  - Applications of ESG Data Analytics with Python.

Artificial Intelligence for Text Analytics

國立臺北大學
National Taipei University

aws educate | Cloud Ambassador
2020 Cohort

aws academy
Accredited Educator

aws certified
Solutions Architect
Associate

aws certified
Cloud Practitioner

Contact Information

Min-Yuh Day, Ph.D.

Associate Professor

Institute of Information Management, National Taipei University

Tel: 02-86741111 ext. 66873

Office: B8F12

Address: 151, University Rd., San Shia District, New Taipei City, 23741 Taiwan

Email: myday@gm.ntpu.edu.tw

Web: http://web.ntpu.edu.tw/~myday/