

The Next Frontier:

Key Trends of Data Science and AI

South California Chapter of
American Statistical Association Meeting
December 6, 2025

Chong Ho Alex Yu, Ph.D., D. Phil., President of SCASA



We cannot fight a modern war using WWII-era weapons



Key Trends Shaping Data Science

From Coding to No-Code or Low-Code Solutions
From Programming to accessible solutions

From Binary Decision to Pattern Recognition
Beyond yes/no to nuanced insights

From Manual Processing to AI Automation
Efficient and Intelligent systems

From Local Computing to Cloud Computing
Virtually limitless storage and processing power



Key Trends Shaping Data Science

From CPU to GPU and TPU
High performance computing

From Small Samples to Big Data
Near population-level data for better predictive accuracy

From Self-Report Data to Behavioral Data
Look at what people do, not what they say

From Structured Data to Both Structured & Unstructured Data
Data are everywhere e.g., social media posts, GPS data



From Coding to Low- or No-Code Solutions

Data manipulation and analysis required programming languages like SQL, Python, or R.

User-friendly interfaces with drag-and-drop functionalities allow even non-programmers to explore and analyze data.

This democratizes data insights and empowers a wider range of people to contribute to data-driven decision making.

80%

Non-IT Users

Projected low-code users outside IT by 2026

\$82B

Market Size

Low-code platform market by 2034



The Rise of Generative AI

Prompt engineering with generative AI, such as ChatGPT, Google Gemini, and Claude is changing the landscape of data analysis.

Generative AI models can be prompted to analyze large datasets and identify patterns, trends, and insights that may not be immediately apparent to human analysts.

📺 **Watch:** Learn more about no-code data analysis at <https://youtube.com/watch?v=LZ983fjHeG4>

Coding Assistant/Vibe Programming

Welcome To Colab

File Edit View Insert Runtime Tools Help

Q Commands + Code + Text ▶ Run all Copy to Drive

Table of contents

- Welcome to Colab!
- Access popular AI models via...
- Explore the Gemini API
- Getting started
- Data science
- Machine learning
- More Resources
- Featured examples

+ Section

Welcome to Colab!

Access popular AI models via Google-Colab-AI Without an API Key

All users have access to most popular LLMs via the `google-colab-ai` Python library, and paid users have access to a wider selection of models. For more details, refer to the [getting started with google colab ai](#).

```
[ ]  
from google.colab import ai  
response = ai.generate_text("What is the capital of France?")  
print(response)
```

Explore the Gemini API

The Gemini API gives you access to Gemini models created by Google DeepMind. Gemini models are built from the ground up to be multimodal, so you can reason seamlessly across text, images, code, and audio.

How to get started?

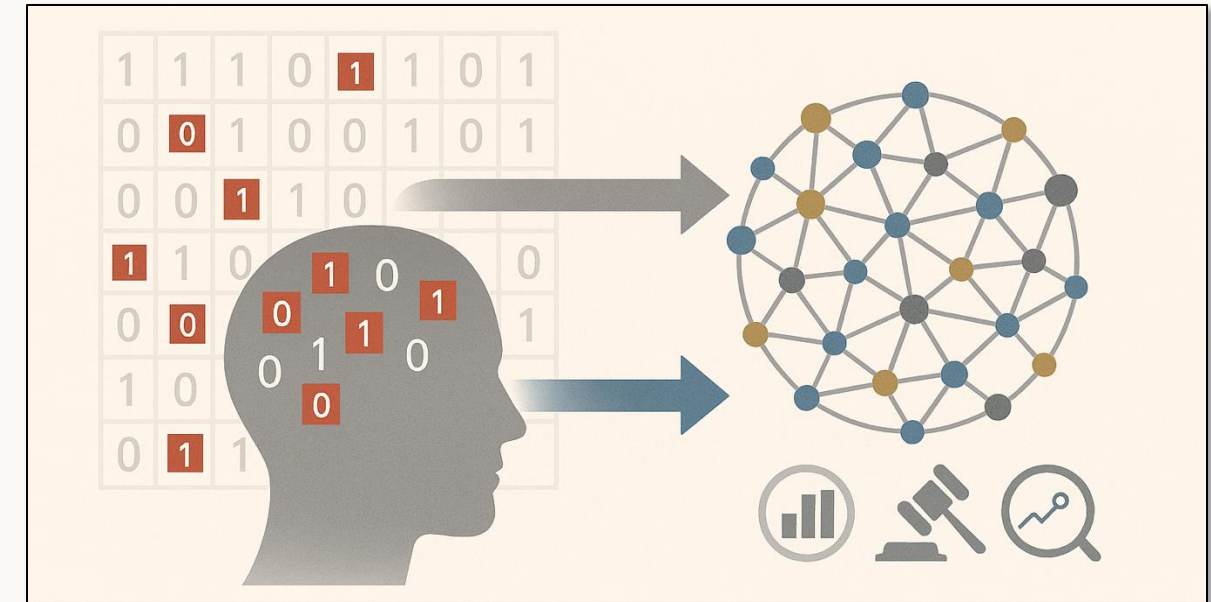
- Go to [Google AI Studio](#) and log in with your Google account.
- [Create an API key](#).
- Use a quickstart for [Python](#), or call the REST API using [curl](#).

Discover Gemini's advanced capabilities

- Play with Gemini [multimodal outputs](#), mixing text and images in an iterative way.
- Discover the [multimodal Live API](#) (demo [here](#)).
- Learn how to [analyze images and detect items in your pictures](#) using Gemini (bonus, there's a [3D version](#) as well!).
- Unlock the power of [Gemini thinking model](#), capable of solving complex task with its inner thoughts.

{} Variables 📄 Terminal

Beyond Yes/No Decisions



1

Classical Statistics

Dichotomous conclusions: reject or not reject the hypothesis.
Simplistic criterion: $\text{Alpha} \leq .05$

2

Modern Analytics

Pattern recognition in complex datasets.

3

Real Applications

Predicting customer churn, identifying fraud, uncovering hidden trends in sentiment

Because reality is complicated, data analysis is moving beyond simple yes/no answers.

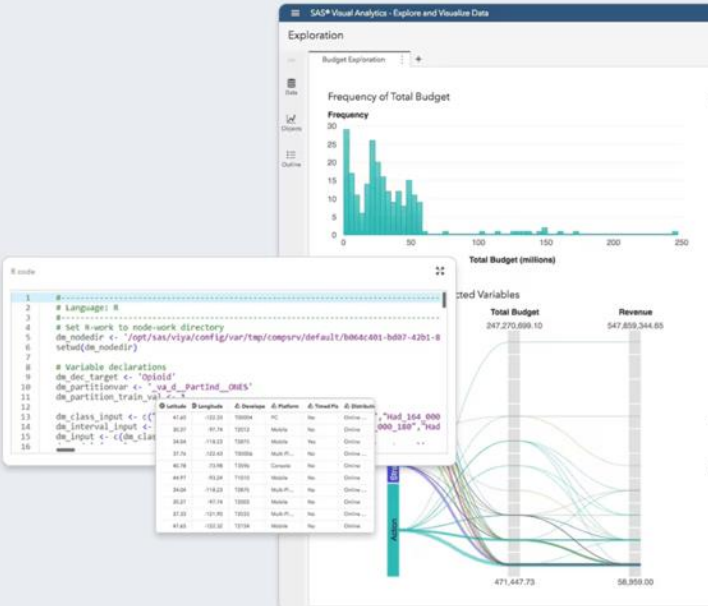
AI-Based Automation: The Shift to Intelligent Systems

SAS® Viya® for Learners

Advance your expertise in cutting-edge data science skills

To access this product, please sign in with your university email.

[Access for educators](#) [Access for students](#)



The screenshot displays the SAS Viya interface. On the left, there is a panel for R code with the following content:

```
1 # language: R
2 #-----
3 # set R-work to node-work directory
4 de_nodendir <- "/opt/sas/viya/config/var/tmp/compsrv/default/8064c801-bd07-42b1-8
5 setwd(de_nodendir)
6
7
8 # variable declarations
9 de_dec_target <- "oploid"
10 de_partitionvar <- "va_d_partind_ones"
11 de_partition_train_val <- "X"
12
13 de_class_input <- c("
14 de_interval_input <-
15 de_input <- c(de_clas
16
```

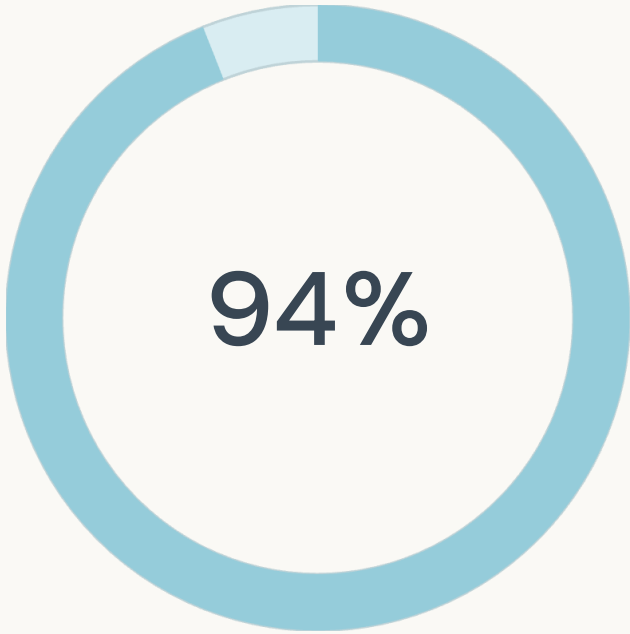
On the right, there are two main visualizations. The top one is a histogram titled 'Frequency of Total Budget' showing the distribution of 'Total Budget (millions)'. The bottom one is a line chart titled 'Revenue' showing 'Revenue' over time, with 'Total Budget' also plotted as a reference line.

Manual data cleaning, analysis, and reporting are becoming increasingly automated with the help of AI.

AI can handle repetitive tasks, identify anomalies, and even generate reports. Examples: SAS Viya, Microsoft Power BI with Copilot, Azure Auto ML, IBM Watson...etc.

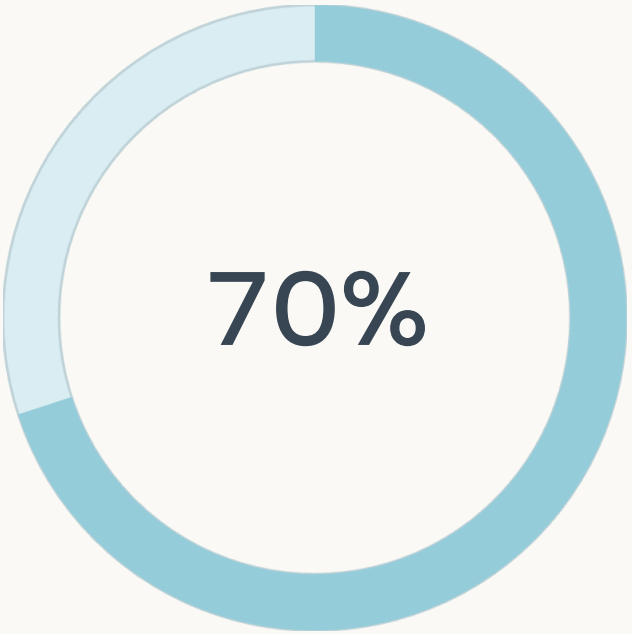
This transformation frees up human analysts to focus on higher-level tasks like interpreting results and strategic decision making, amplifying human intelligence rather than replacing it.

The Cloud Computing Revolution



Corporate Adoption

Corporations currently using cloud computing



Industry Platforms

Enterprises using cloud platforms by 2027



Market Growth

Global cloud market value by 2030

Cloud computing offers virtually limitless storage capacity and processing power compared to traditional laptops. Cloud platforms also enable collaboration and real-time data access for teams working remotely.

GPU-and TPU-Based High-Performance Computing

Central Processing Units (CPUs) are good for general-purpose tasks, but data analysis often involves complex calculations. Graphics Processing Units (GPUs), originally designed for video processing, are much faster for specific computations required in machine learning.

Deep Learning Acceleration

Training deep convolutional neural networks: from days to hours with GPUs

Diverse Applications

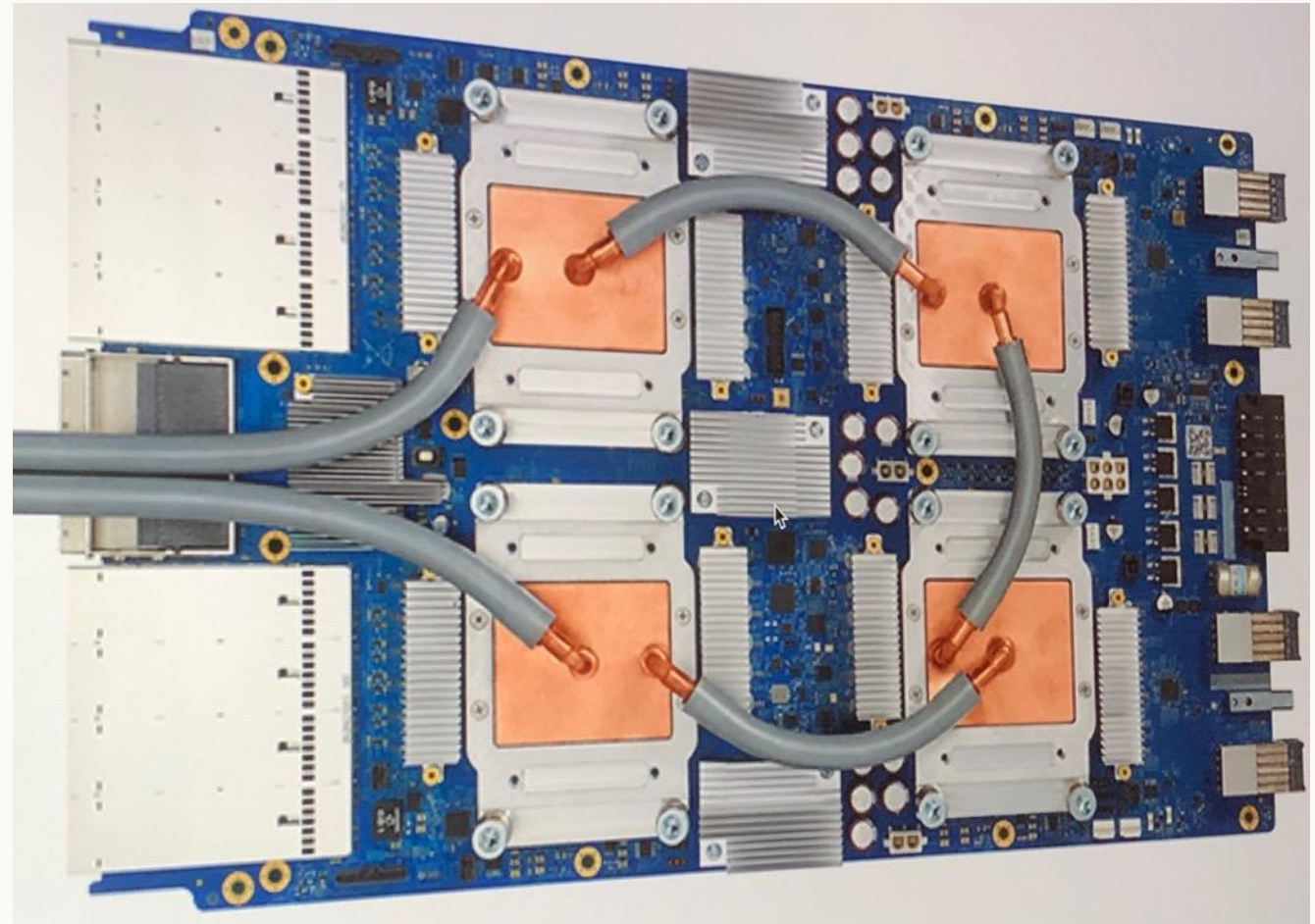
Image recognition & natural language processing

📄 **Learn More:** Explore GPU computing at <https://youtube.com/watch?v=VWShTgP5KEk>

Google's Tensor Processing Units

Google has developed **Tensor Processing Units (TPUs)** – specialized hardware designed specifically for accelerating machine learning and deep learning tasks.

Accessible via Google Colab



❏ **Learn More:** Explore TPU computing at <https://www.youtube.com/watch?v=OalzyQj3B68>

The Big Data Revolution

From Samples to Populations

Traditionally, data analysis relied on smaller, manageable datasets. Today, we have access to vast amounts of data from various sources.

Real-World Impact

Manufacturers leverage big data to anticipate equipment failures and refine maintenance timetables.

AI Bias?

Weapon of Math Destruction?

Contrary to popular belief, machine learning is less biased than classical statistics and humans. Rather, it can overcome our bias!



Learn More:

<https://www.youtube.com/watch?v=3YflKEyEQe8>

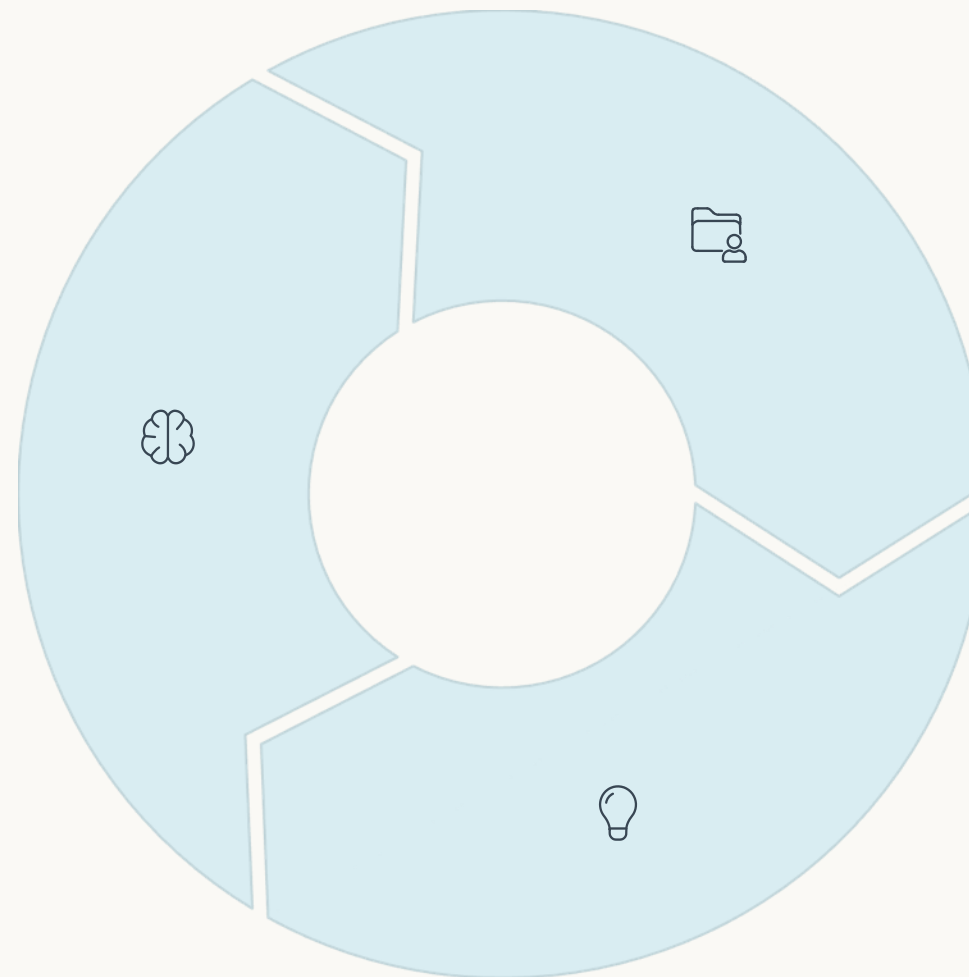
<https://www.youtube.com/watch?v=FlkE3tkalT0>



AI and Big Data: Inseparable Forces

The true power of big data is unlocked only when paired with artificial intelligence.

AI Provides
Analytical brain for big data



Big Data Provides
Memory and experience for AI

Together Create
Deeper insights and robust models

📄 **Explore Further:** Watch big data in action at <https://youtube.com/watch?v=-SSQKpxNrl8>

From Self-Report to Behavioral Data

1 Traditional Approach

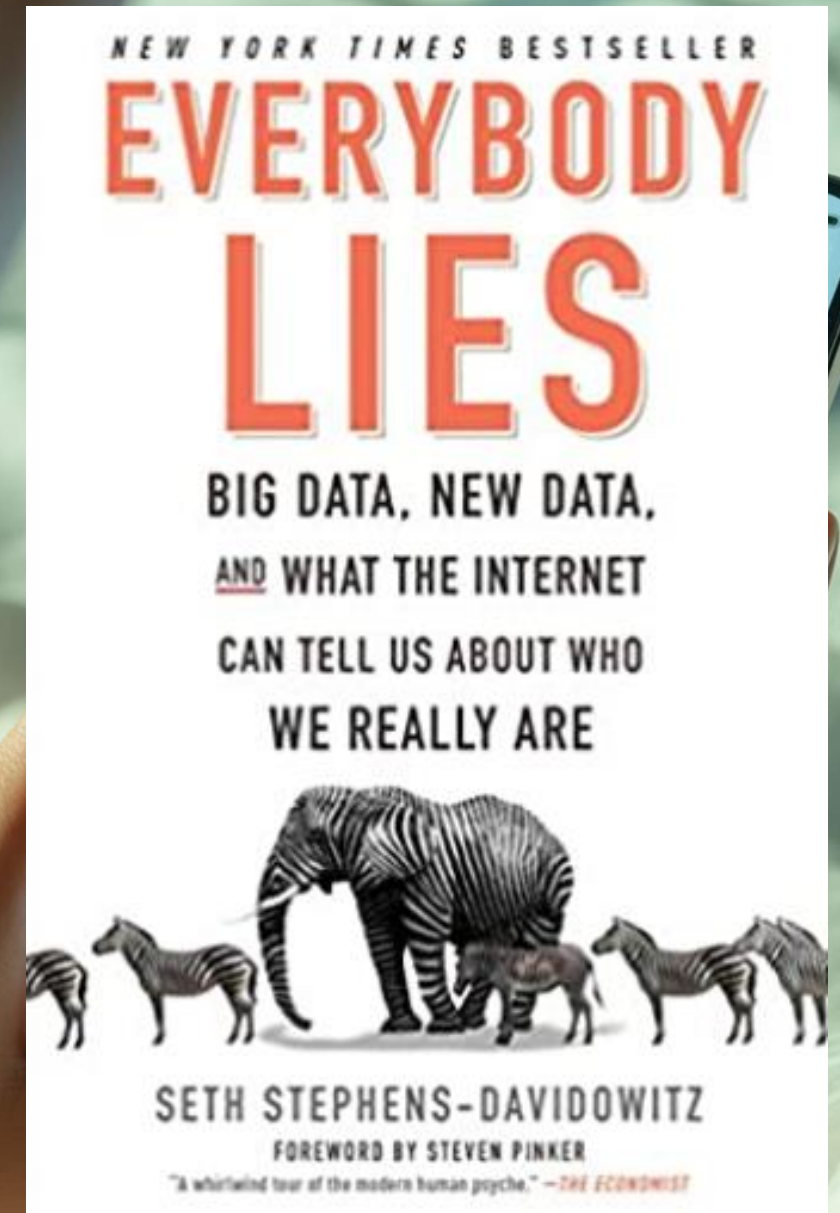
Surveys and questionnaires as primary data sources about preferences and behaviors

2 Modern Approach

Behavioral data through website interactions, app usage, sensors, and social media activity

3 The Advantage

More objective and richer understanding of how people truly behave



Real-World Behavioral Analytics



Employs predictive analytics to deliver personalized product recommendations by analyzing users' browsing and purchase history



Utilizes predictive analytics to suggest personalized content by analyzing viewing history and preferences, crafting tailored recommendations



Structured vs. Unstructured Data

Structured Data

Information stored in predefined formats such as databases. Organized, easily searchable, and traditionally the focus of data analysis.

- Database tables
- Spreadsheets
- Organized records

Unstructured Data

A substantial portion of modern data originates from unstructured sources. Emerging techniques can now effectively manage this data.

- Social media posts
- Emails and text documents
- Customer call recordings
- Videos and images

Unlocking Unstructured Data Insights



Airlines

Utilize sentiment analysis to measure passenger satisfaction through social media messages



Mental Health Apps

Evaluate users' emotional states based on voice patterns and speech analysis



Retailers

Forecast foot traffic patterns in stores by leveraging video surveillance data

In the past it was tedious to read through open-ended responses, but today AI-powered natural language processing can perform text mining in just a few seconds, extracting valuable insights from vast amounts of unstructured text.

📄 **See It in Action:** Learn about unstructured data analysis at <https://youtube.com/watch?v=RADRU9zySi0>

The future is now!

Resistance is futile



Contact me

Chong Ho Alex Yu, Ph.D., D. Phil.

chonghoyu@gmail.com

[YouTube Channel on Data Science and AI](#)
[Profile on Google Scholar](#)

