

Reimagine Data Analysis and Visualization with Gen AI Agentic Technologies

Kevin Lee, Nathan Lee



Disclaimer

The views and opinions presented here represent those of the speaker and should not be considered to represent any companies or organizations.

Agenda

- ChatGPT impact on us
- Programming workflow before ChatGPT
- Programming workflow after ChatGPT
- Programming workflow with AI Agent
- Introduction of AI Agent & Agentic workflow
- AI Agents integration in Biometrics (System, Process, People)
- Reimagine with AI Agents
- Discussion



Do you use “ChatGPT” for your work?



How ChatGPT changed our work?

- Programming & Code Support
 - Code development
 - Code Debugging
 - Learning new languages (e.g., SAS, R and Python)
- Document & Specification development / interpretation
- Collaboration & Communication : Email, Report writing, Analysis, Research
- Workflow – More automation by AI
- Role of Biometrics : hard-coding to interpretation, strategy, verification



Programming workflow before ChatGPT



ChatGPT impact on programming

- Code Development
- Code Debugging
- Lowering the boundary of languages (e.g., SAS, R and Python) – vibe coding

Codes Development – R

Prompt :

Write a **R script** that groups data by "TRTA" and "AVISIT" columns in "adv.s.xpt"

and calculates the mean of "AVAL" column for each group at PARAMCD = 'WEIGHT'

Output :

```
library(haven) # For reading XPT files  
library(dplyr) # For data manipulation
```

```
# Read the ADaM dataset  
adv.s <- read_xpt("adv.s.xpt")
```

```
# Process the data
```

```
result <- adv.s %>%
```

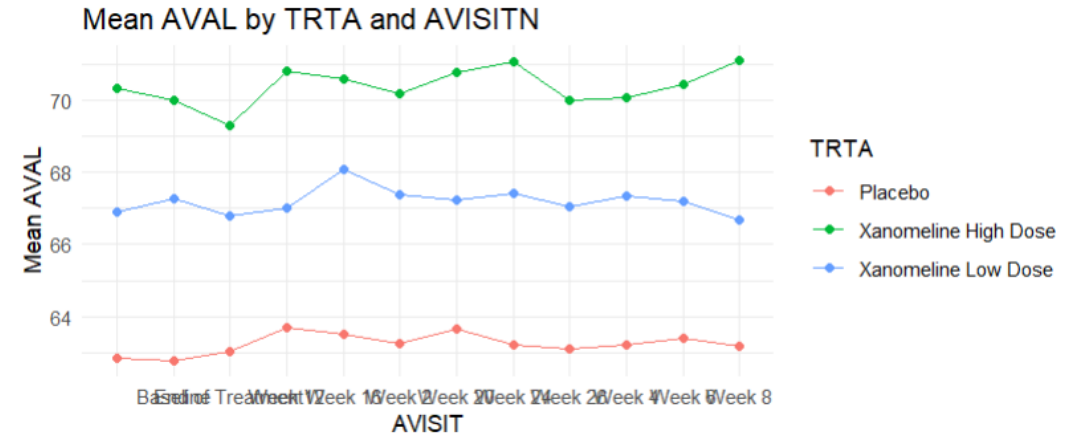
```
  filter(PARAMCD == "WEIGHT") %>% # Filter for PARAMCD = 'WEIGHT'
```

```
  group_by(TRTA, AVISIT) %>%      # Group by TRTA and AVISIT
```

```
  summarise(mean_AVAL = mean(AVAL, na.rm = TRUE)) %>% # Calculate mean of  
  AVAL  
  ungroup()
```

```
# Print the result
```

```
print(result)
```



Codes Development – SAS

Prompt :

Write a **SAS script** that groups data by “TRTA” and “AVISIT” columns in "advs.xpt"

and calculates the mean of “AVAL” column for each group at PARAMCD =‘WEIGHT’

Output :

```
libname mydata xport 'path_to_advs.xpt'; /* Replace with the actual path to advs.xpt */
```

```
data weight_data;  
  set mydata.advs;  
  if PARAMCD = 'WEIGHT'; /* Filter data for PARAMCD = 'WEIGHT' */  
run;
```

```
proc sql;  
  create table weight_means as  
  select TRTA, AVISIT,  
         mean(AVAL) as mean_weight  
  from weight_data  
  group by TRTA, AVISIT;  
quit;
```

```
proc print data=weight_means;  
run;
```

Codes Development – Python

Prompt :

Write a **Python script** that groups data by “TRTA” and “AVISIT” columns in “adv.s.xpt”

and calculates the mean of “AVAL” column for each group at PARAMCD = ‘WEIGHT’

Output :

```
import pandas as pd
import xport

# Load the .xpt file using xport
with open('adv.s.xpt', 'rb') as f:
    data = xport.load(f)

# Convert the xport data to a pandas DataFrame
df = pd.DataFrame(data['ADVS'])

# Filter the data for PARAMCD = 'WEIGHT'
df_weight = df[df['PARAMCD'] == 'WEIGHT']

# Group by TRTA and AVISIT, then calculate the mean of AVAL
grouped_data = df_weight.groupby(['TRTA', 'AVISIT'])['AVAL'].mean().reset_index()

# Output the grouped data
print(grouped_data)
```

Codes Debugging – Python

Prompt :

Debug below <python codes> using <error message>.

<python codes>:

```
import pandas as pd
import xport
```

```
# Load the .xpt file using xport
with open('adv.s.xpt', 'rb') as f:
    data = xport.load(f)
```

```
# Convert the xport data to a pandas
DataFrame
```

```
df = pd.DataFrame(data['ADVS'])
```

```
# Filter the data for PARAMCD = 'WEIGHT'
```

```
df_weight = df[df['PARAMCD'] == 'WEIGHT']
```

```
# Group by TRTA and AVISIT, then calculate
the mean of AVAL
```

```
grouped_data = df_weight.groupby(['TRTA',
'AVISIT'])['AVAL'].mean().reset_index()
```

```
# Output the grouped data
print(grouped_data)
</python codes>
```

<error message>

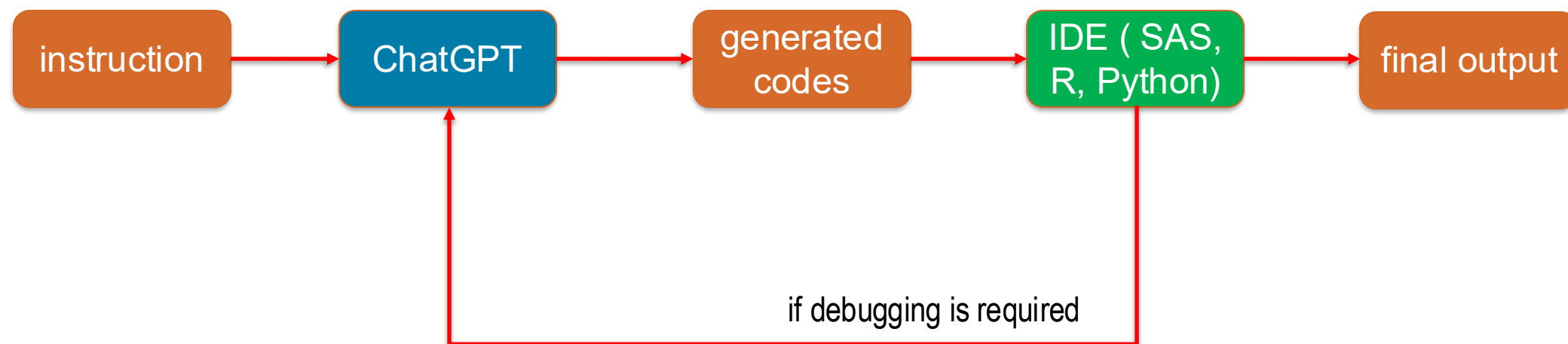
TypeError: Could not convert ['aval'] to numeric

</error message>

Programming workflow after/ using ChatGPT



Programming workflow after/using ChatGPT – Detailed



Question :

- Can we automate this workflow further?

AI Agent

- System that can interact with its environment, collect data, and **autonomously** perform tasks and make decisions to achieve the goals.
- Main Components
 - LLM
 - Prompts
 - Tools
 - Memory (short-term /long-term memory, RAG)
- Junior stat programmer who
 - has knowledge of coding
 - with the instruction,
 - access to tools (SAS, R, Python, functions & macros) and data
 - understand our workflow



AI Agentic workflow

AI Agentic Workflow is a workflow or process where an **AI agent** not only generates outputs (like code or text) but also **plans, executes, monitors, and adapts its actions** step by step to achieve a goal.

Ex. For Junior Stat Programmer,

- Develop codes
- Execute codes
- Find the errors
- Debug codes
- Re-execute codes
- Do debugging and executing the codes until the codes are error-free or He or She can't resolve the errors.
- Then, deliver the final results.



Programming workflow using AI Agents



Question :

- What is the problem in this picture?

ChatGPT or AI Agent is like a very smart, capable, recently graduated intern/new hire.



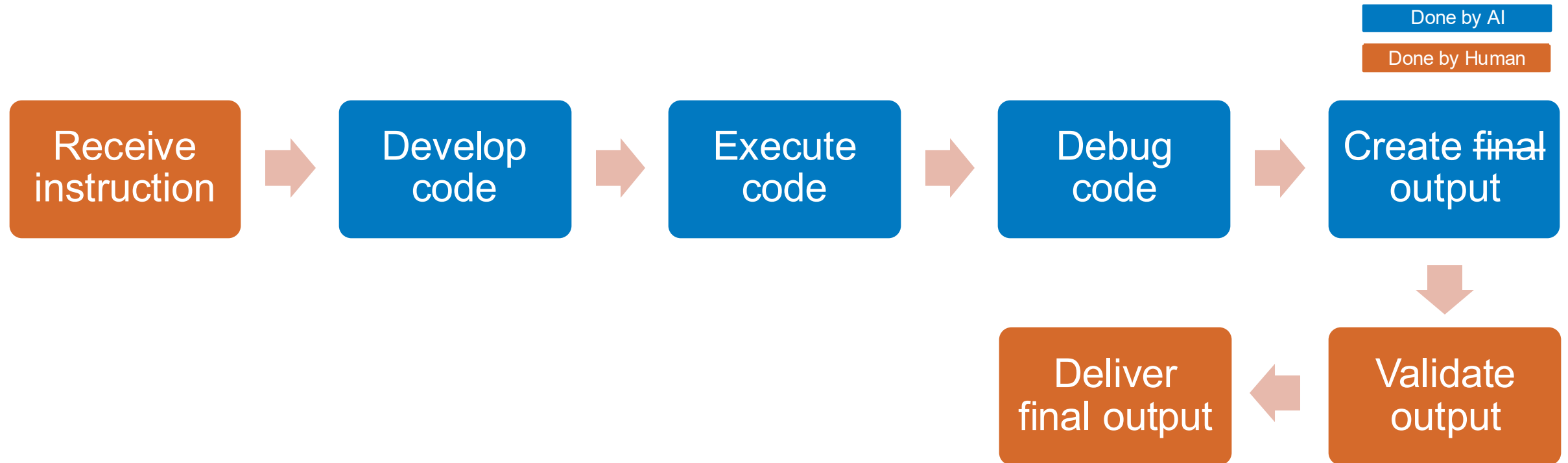
=



What will we do with the results from an intern?



Programming workflow using AI Agents with human involvement – **AI Agentic workflow**



AI Agents in demo

TFL Validation Tool



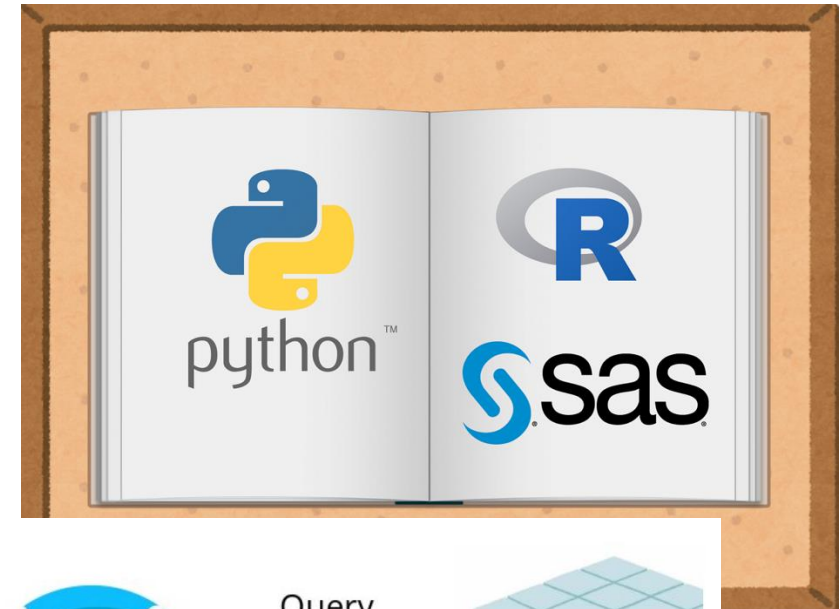
AI Agents integration in Biometrics

- Identifying / Developing AI agents
- Integrating with existing infrastructure
- Determine its usage
- Validation framework
- Compliant AI Agentic workflow integration

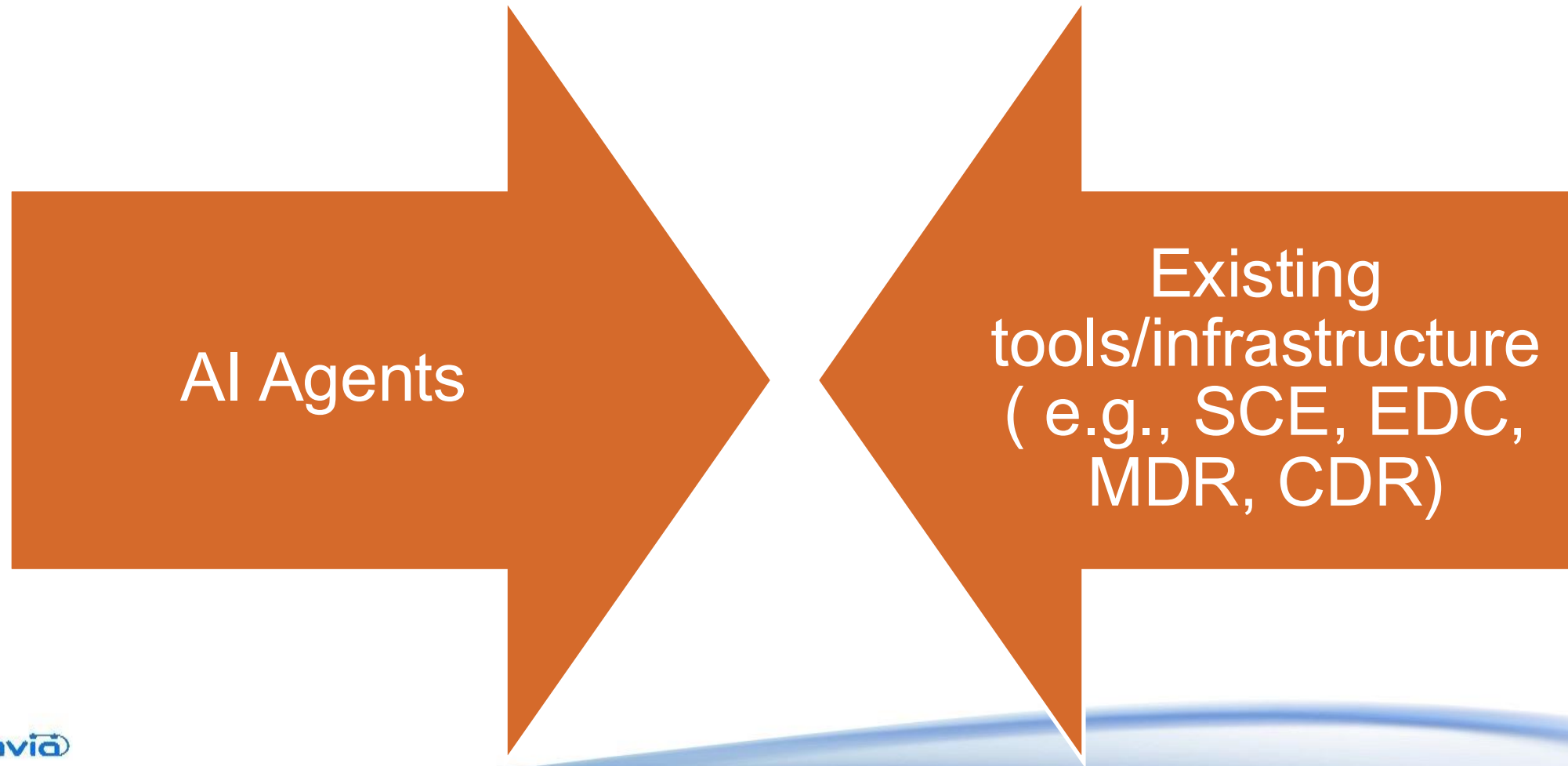


Where could we use AI Agent?

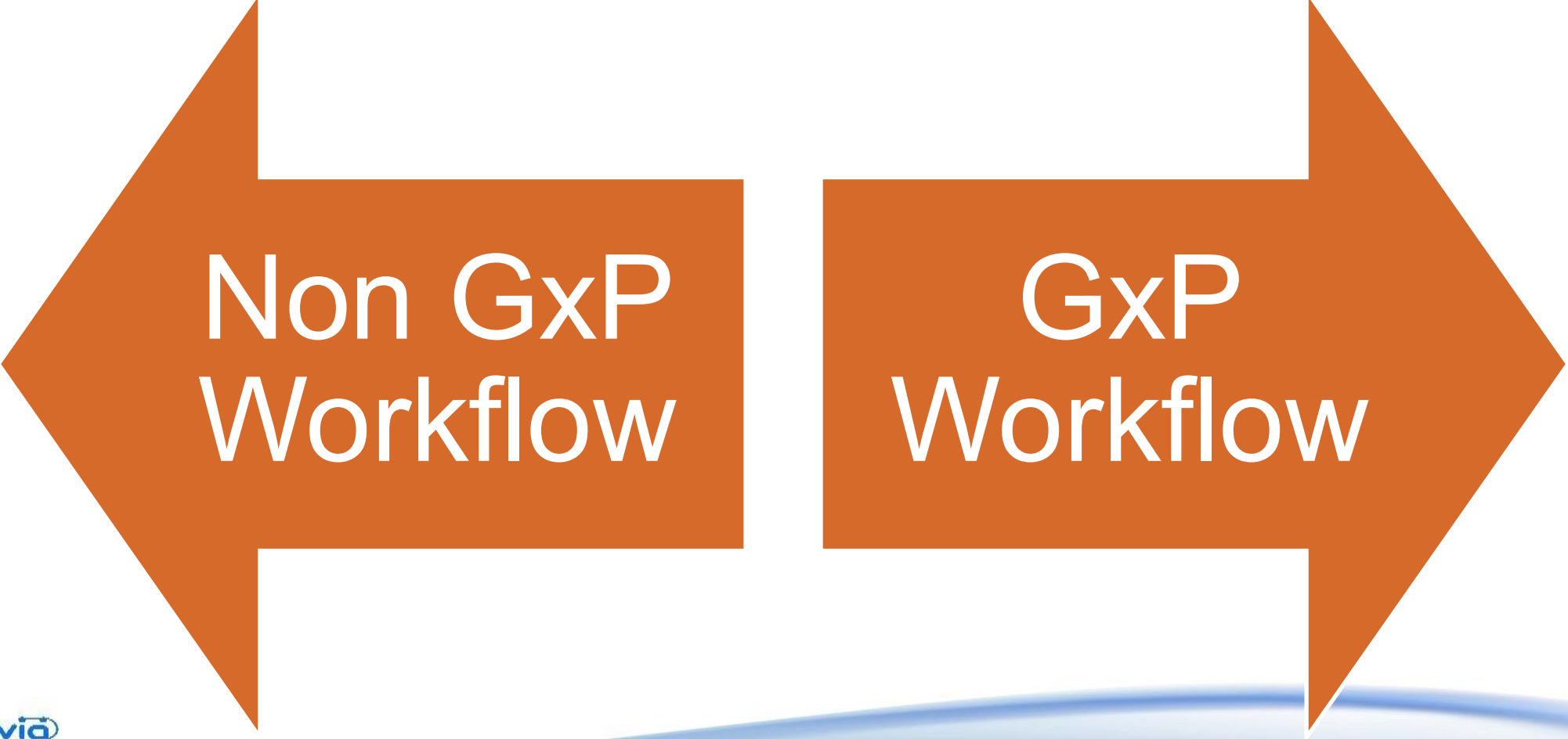
- Code Development (SAS, R, Python) for SDTM, ADaM, TFL
- Validation of the deliverables – SDTM, ADaM, TFL
- Generation of clinical trial artifacts – Protocol, SAP, Mock up Shell, Data Specification,,,
- Chatbot
 - Clinical trial study
 - Internal documents
- Project management
- Data Reconciliation



Integration with existing infrastructure



Where AI Agentic workflow???



Non GxP
Workflow

GxP
Workflow

Validation framework

- **Regulatory Importance:** AI Agentic workflow outputs must meet the **same quality standards** as traditionally produced results – traceability, reproducibility, accuracy, compliance.
- **Biometrics Leadership Role:** Must establish rigorous, **risk-based validation protocols/systems** tailored for AI Agentic workflow implementation.
- **Framework Components:**
 - Risk-based validation planning
 - Verification of AI Agentic workflow output accuracy
 - Application of GxP / GAMP 5 principles
 - Comprehensive documentation (plans, protocols, test scripts, reports)
 - Version control and audit trails
 - Package validation

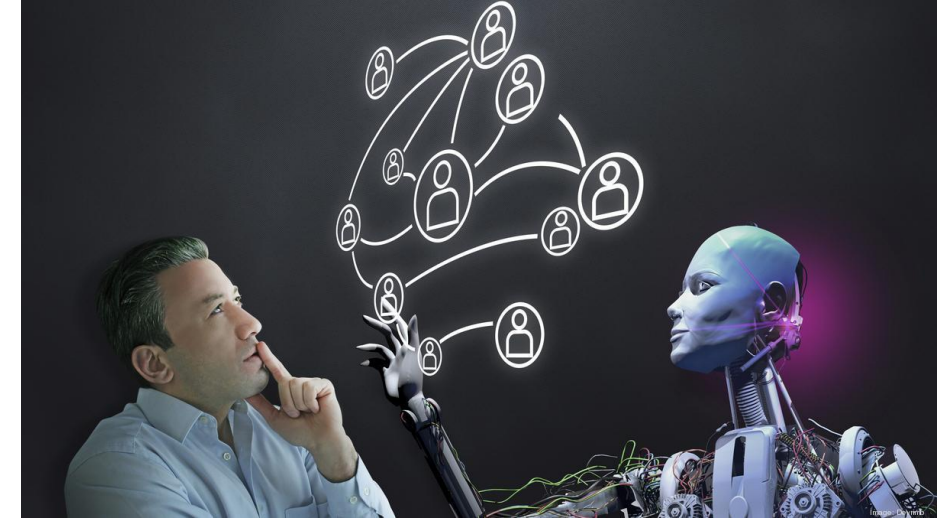


Change Management – How we could lead this AI Agents era

- Define Human vs AI Agents Collaboration Policy
- Capability Building
- Culture Transformation

Human vs AI Agents Collaboration Policy

- Clear delineation
 - AI-driven activities
 - Human-driven activities
- AI roles
 - *Assistant*: e.g., generating initial code or report templates
 - *Autonomous Operator*: e.g., handling routine quality checks or standardized visualizations
- Human intervention- required to ensure **oversight** and **accountability**.



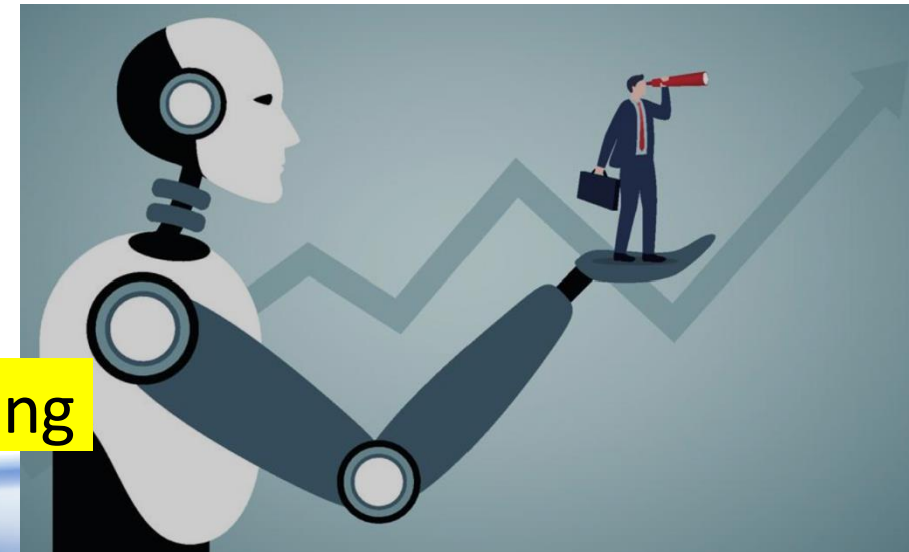
Comprehensive Capability Building

- Assessment
 - Existing skill gaps
 - Implement targeted training initiatives.
- Focus areas for capability building include:
 - Technical proficiency with AI Agents & Agentic workflow
 - Prompt engineering skills
 - Validation skills for AI agentic workflow outputs
- A systematic approach ensures department-wide readiness for AI Agentic workflow integration.



Culture Transformation

- Foster psychological safety to encourage building AI agents and experimentation with AI agents.
- Promote a mindset of **responsible innovation**—where creativity thrives without compromising compliance.
- Go **beyond submission** works.
- Goals : **To ensure continued commitment to regulatory compliance while embracing and leading AI Agents integration.**

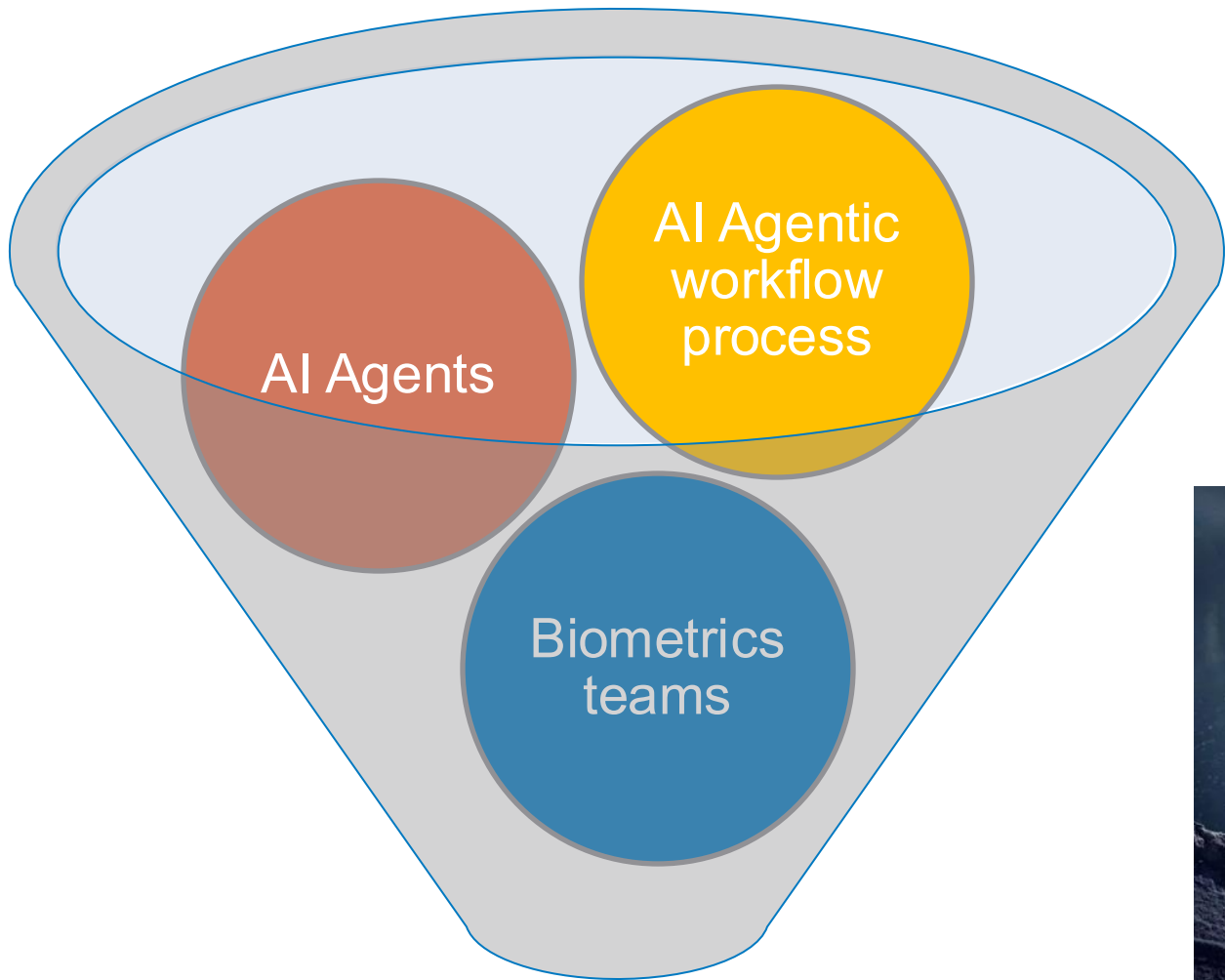


ChatGPT Users vs Non-Users Comparison in Boston Consulting

- Results of ChatGPT Users
 - finished **12.2%** more tasks on average
 - completed tasks **25.1%** more quickly
 - produced **40%** higher quality results

Reimagine : How about with AI Agents????





AI Agents-powered “Super Biometrics Team (Avenger)”

Reimagine into “Super Biometric Team”



ANY QUESTIONS



Contact : kevin@clinvia.com

LinkedIn:

<https://www.linkedin.com/in/HelloKevinLee/>

