phuse·share
connect·share·advance

Single Day Event

A Look into the Near Future: A Day in the Life of a Statistical Analyst

Margarita Mersiyanova,
SAS

23 September 2025

phuse.global


Pfizer
Margarita Mersiyanova


3779

A Look into the Near Future: A Day in the Life of a Statistical Programmer

Can you generate a PROC means on a sas dataset AE.SAS7SDAT?

Looks like you mean `AE.SAS7BDAT` (the standard extension for SAS datasets).
Here's a simple `PROC MEANS` you can run on that dataset:

```sas
libname mydata "C:\your\path\to\data";  /* adjust folder path */

proc means data=mydata.AE n mean std min max;
    var _numeric_;  /* summarize all numeric variables */
run;
```

Explanation:

- `libname` tells SAS where your dataset lives. You replace the path with the folder containing `AE.sas7bdat`.
- `data=mydata.AE` points to that dataset.
- `n mean std min max` tells SAS what statistics to compute (number of observations, mean, standard deviation, min, max).
- `var _numeric_` applies it to all numeric columns. You could also list specific variables instead.

If you want a quick descriptive table of *all* variables (including categorical), you can also do:

```sas
proc means data=mydata.AE n mean std min max;
run;
```

That will summarize everything numeric without specifying variables.

Would you like me to generate a more **video-demo friendly version** (shorter code, with comments explaining each step on screen), since you mentioned recording earlier?

Thank you Chat GPT!!!!!!!!

Agenda

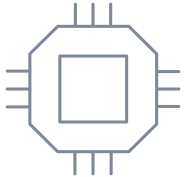
- Daily role of statistical programmers
- Automation of manual processes
- AI-augmented statistical analysts
- Responsible usage of AI
- The future of statistical programmers

Importance of Statistical Programmers

THE WHOLE JOURNEY

GUARDIANS OF TRUST

DATA DETECTIVES



EVOLVING WITH TECHNOLOGY

The Real Job of Statistical Programmers

- Wrangling study data
- Mapping datasets: SDTM → ADaM → TLFs
- Ensuring full traceability and quality control
- Operating within rules
- Managing constant change
- Facing submission pressure

Pain Points for Statistical Programmers

Traceability fatigue

Protocol amendments

Big, messy data

Shift to Open Source

Submission deadlines

Tool changes

Why Now?

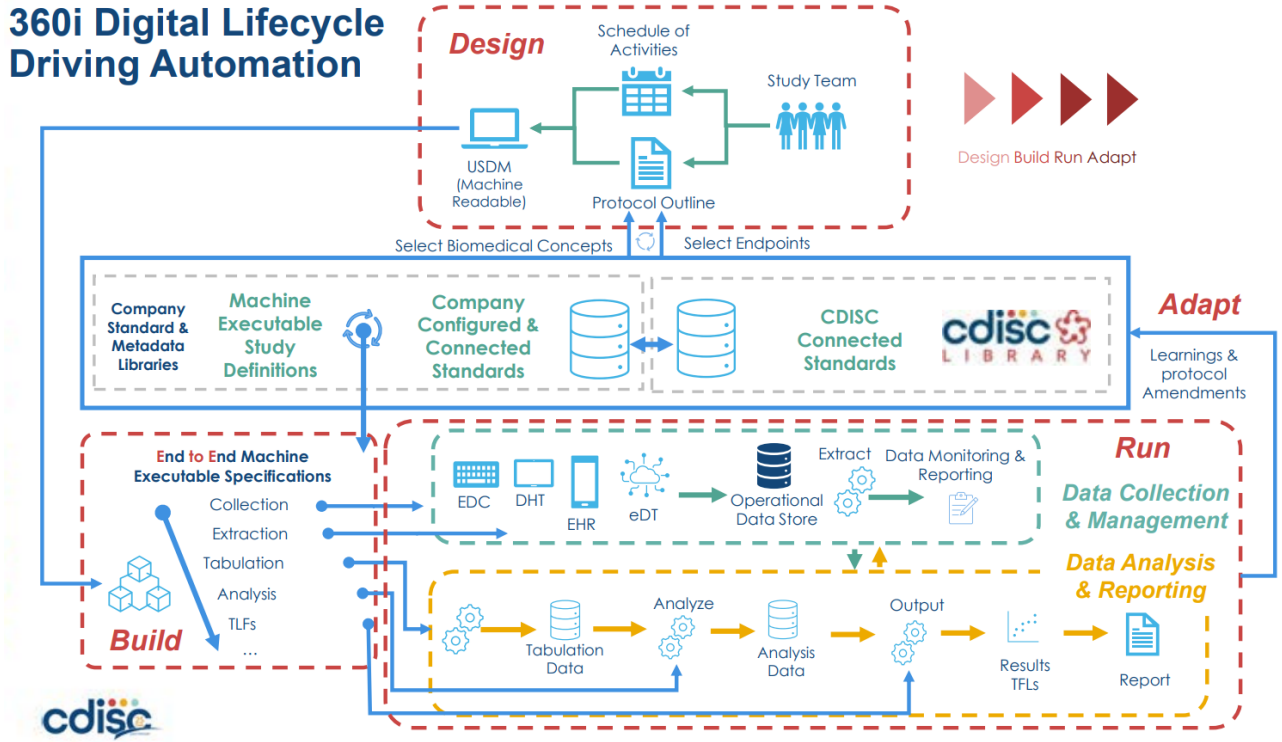
Market Drivers in Health Care and Life Sciences

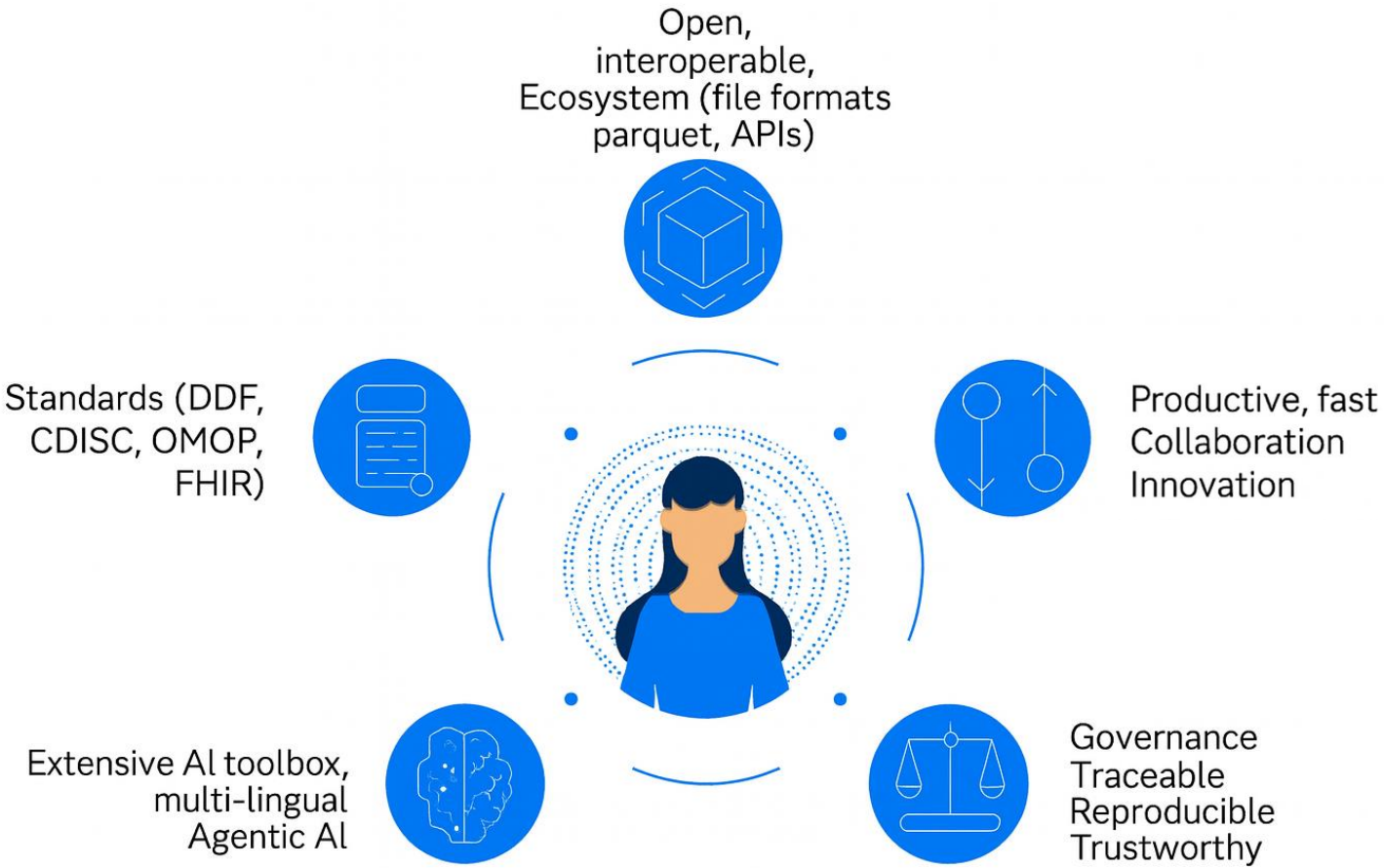
Rising complexity of scientific, regulatory and operational decisions

Patient and physician expectations for real-time information

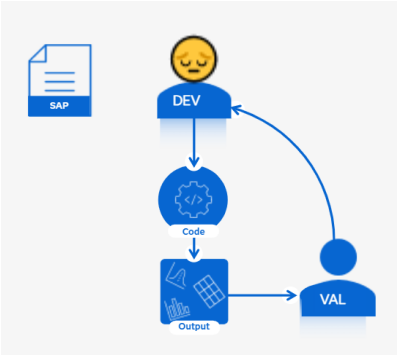
Dynamic environments and adaptability (cost, R&D speed, workforce constraints)

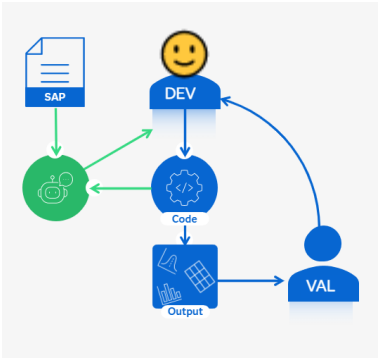
CDISC 360i Vision

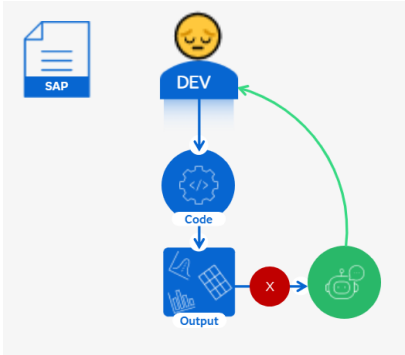



Open,
interoperable,
Ecosystem (file formats
parquet, APIs)
Standards (DDF,
CDISC, OMOP,
FHIR)
Productive, fast
Collaboration
Innovation
Extensive AI toolbox,
multi-lingual
Agentic AI
Governance
Traceable
Reproducible
Trustworthy

SAS AI Agent for Code Validation







DEV
Original

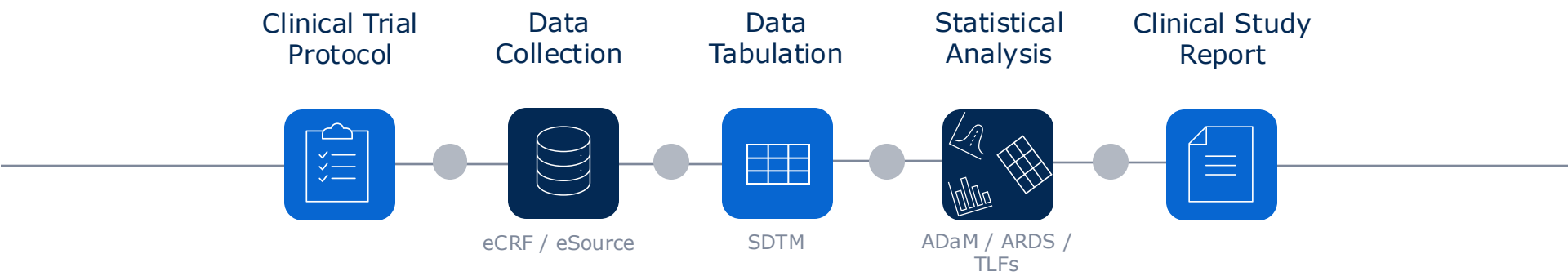
```
else if kupcase(vstestcd) eq 'DIABP' then do;

   ATOXGRHN = 0 + (aval ge 141) + (aval ge 151) + (aval gt 155);
   if not missing(atoxgrhn) then ATOXDSCH = 'HYPERTENSION (DIASTOLIC)';
end;
```

LLM Update

```
else if kupcase(vstestcd) eq 'DIABP' then do;
   /* Corrected the ranges for Hypertension (Diastolic Blood Pressure) */
   ATOXGRHN = 0 + (aval ge 91) + (aval ge 96) + (aval gt 100);
   if not missing(atoxgrhn) then ATOXDSCH = 'HYPERTENSION (DIASTOLIC)';
end;
```

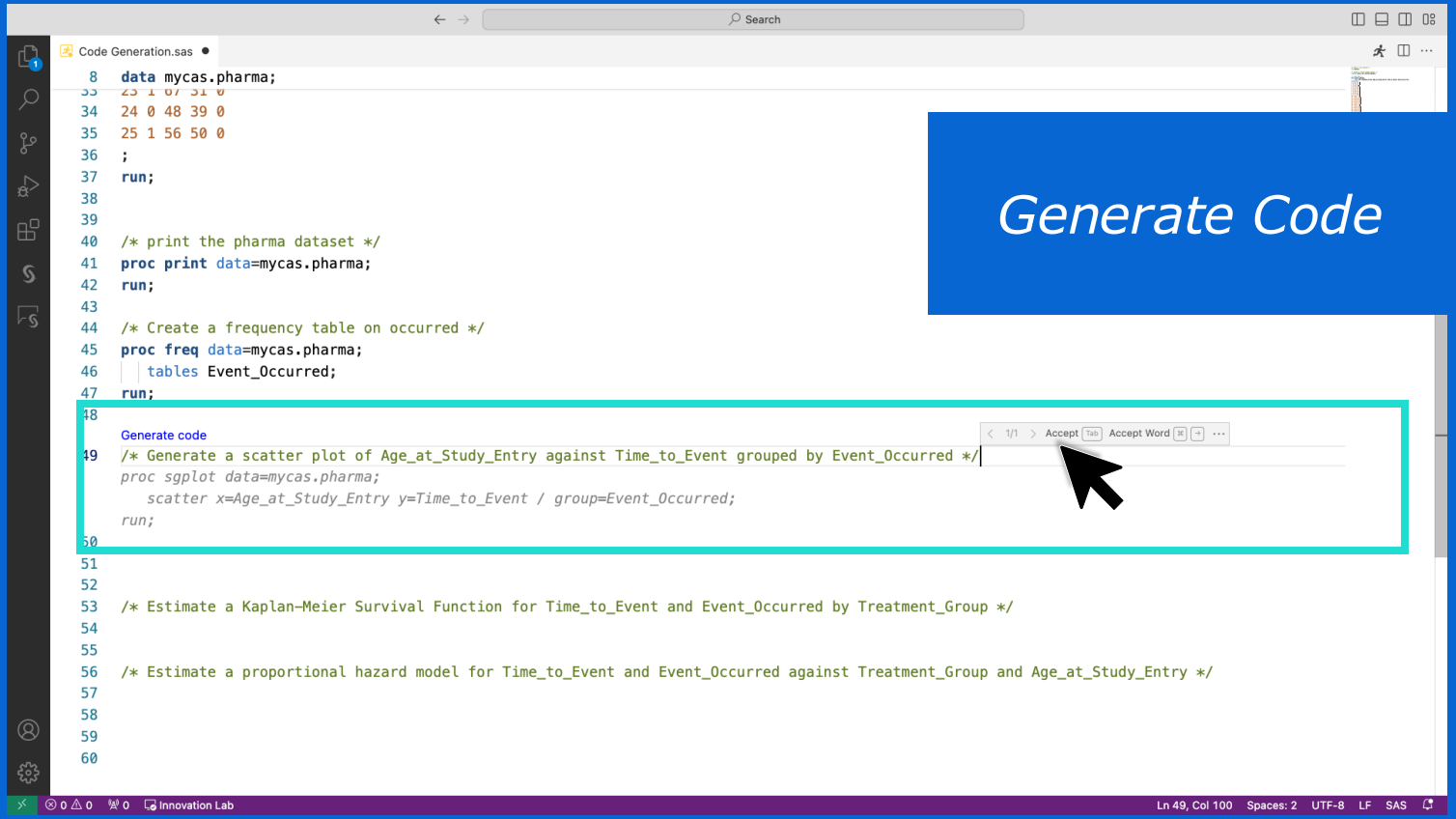
SAS Clinical Trials Workflow Vision



- Raw data to SDTM format
- SDTM to ADaM
- ADaM Validation
- Creation of define.xml
- Creation of TLFs

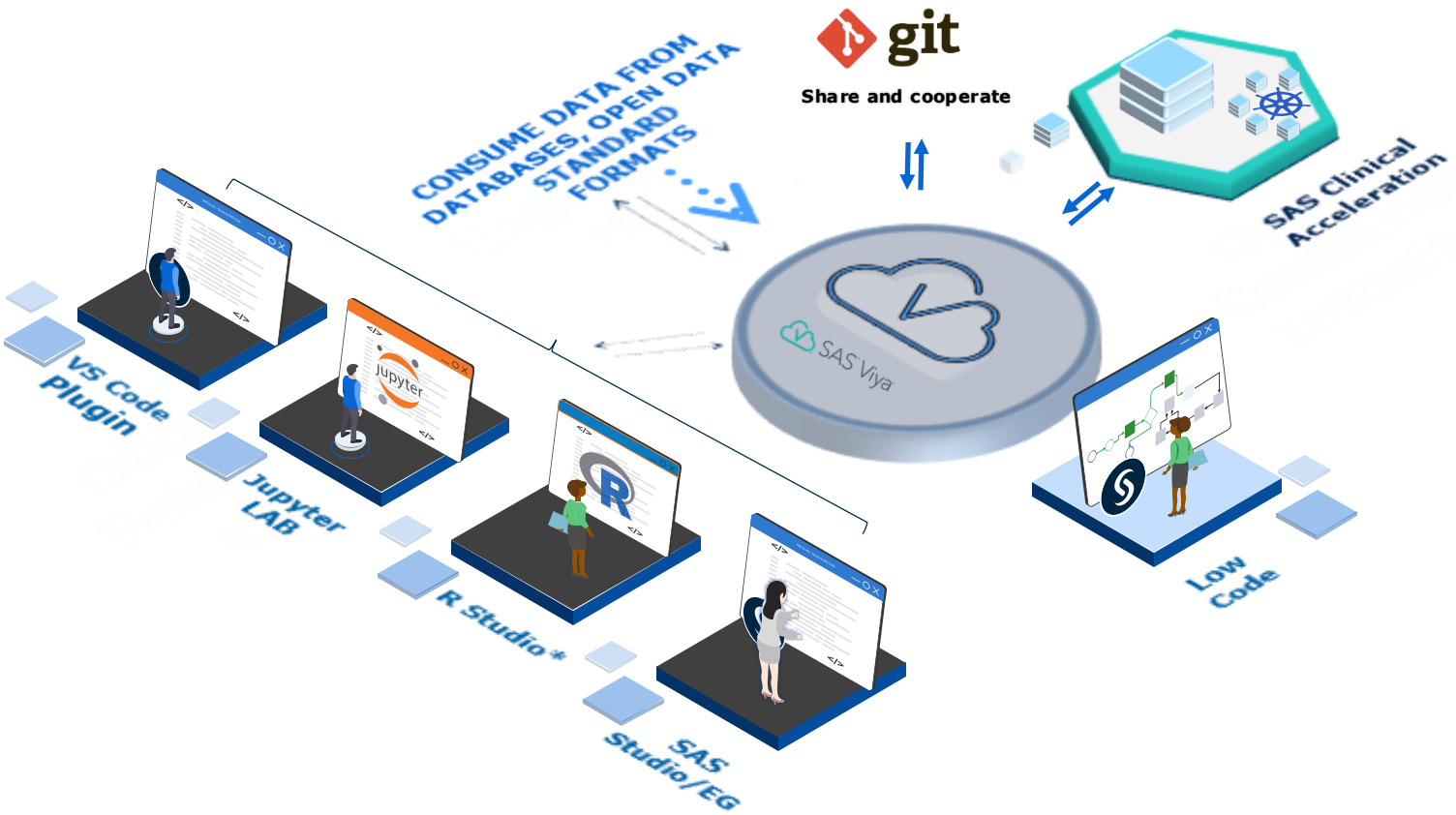
SAS Code Generation

SAS® Viya® Copilot



Generate Code

SAS Viya



Introducing SAS® Clinical Acceleration

Integration

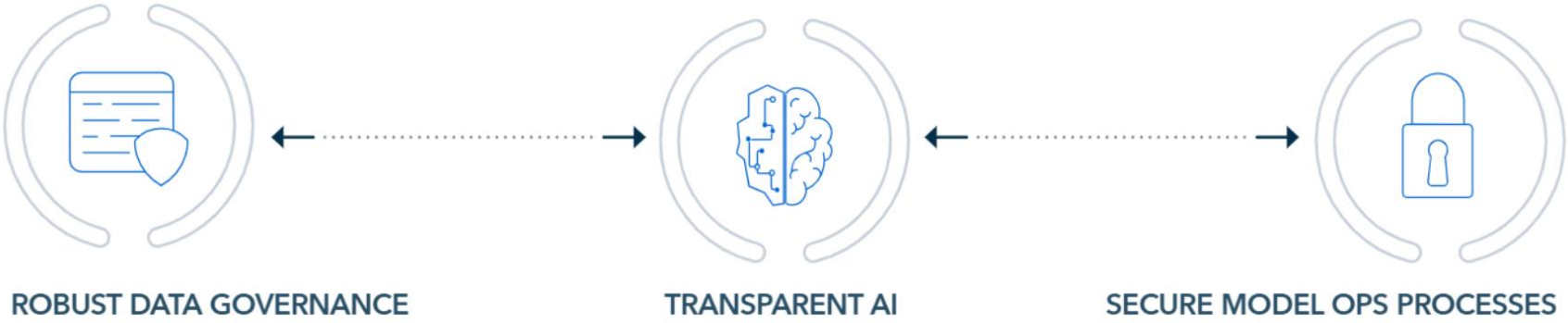
Automation

Streamline Processes

Key technical areas for clinical trial management

1. Clinical data repository
2. Statistical computing environment
3. Collaborative workflows
4. Regulatory reporting
5. Management insight

SAS Viya Delivers Trustworthy AI



Why fully autonomous AI agents fail enterprises?

In the enterprise world, you need more than full autonomy:

- **Decisioning Traceability and Explainability**
- **Risk Mitigation and Control**
- **Governance and Auditability**

Fully autonomous agents, as they're often conceptualized today, struggle in the enterprise environment.

Statistical Programmer of Tomorrow

- From code writer → *clinical data scientist*
- Skills: domain expertise, communication
- Focus of strategy, trust, traceability & interpretation
- Oversight of AI-generated outputs
- Fewer in numbers, greater in influence

The Global Healthcare Data Science Community

Contact Channels

UK +44 1843 609600
office@phuse.global
phuse.global

Social Media

@phusetwitta
/phusebook
/phusetube
/company/phuse