

Single Day Event

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

Margarita Mersiyanova,
SAS

23 September 2025



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;
```

 Copy code

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;
```

 Copy code

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



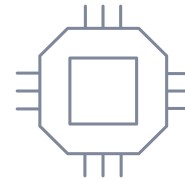
THE WHOLE JOURNEY



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?

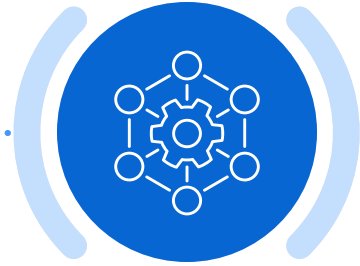
Market Drivers in Health Care and Life Sciences



Rising complexity of
scientific, regulatory
and operational
decisions



Patient and physician
expectations
for real-time
information

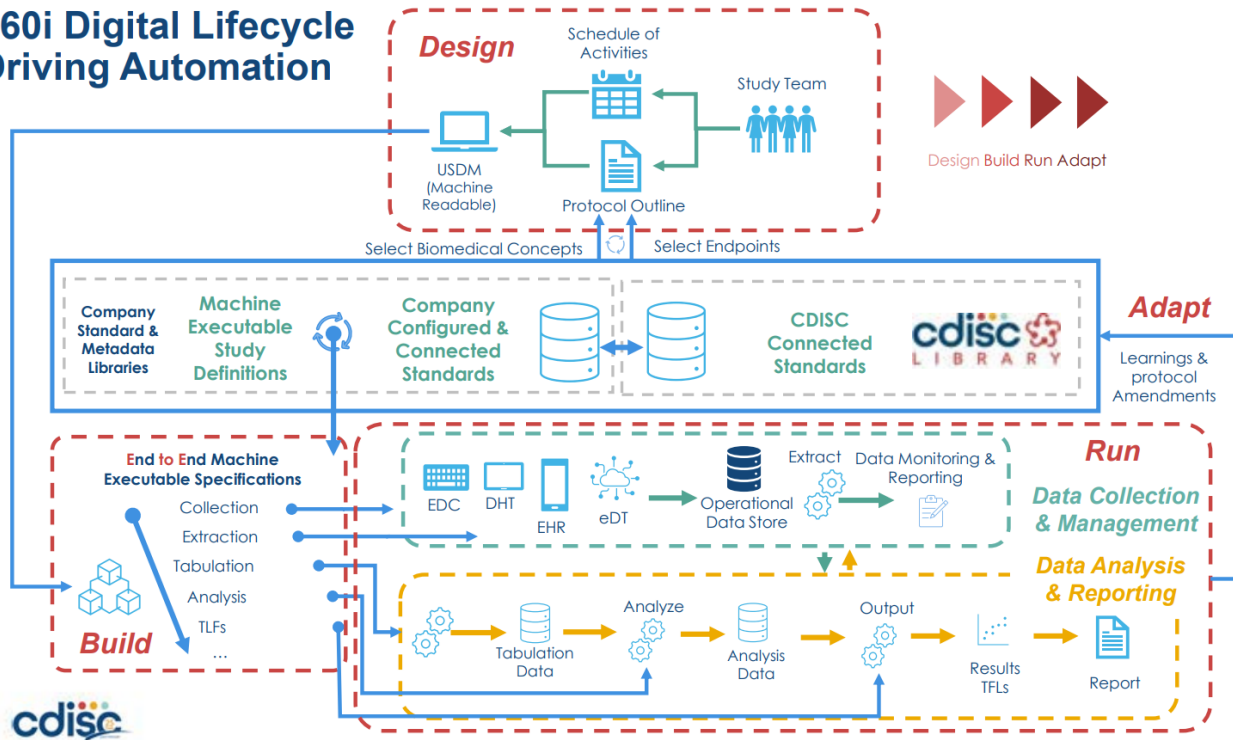


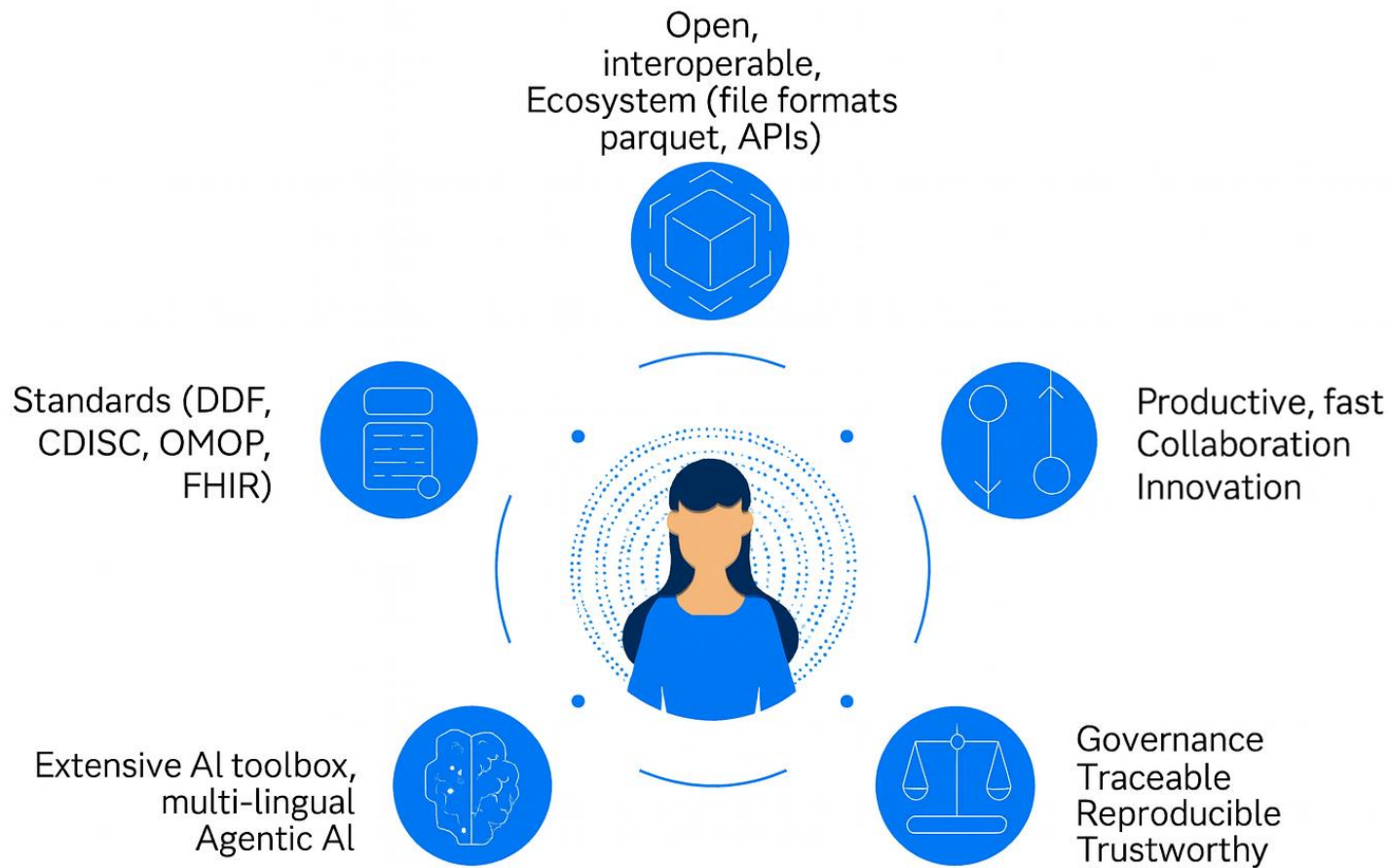
Dynamic environments
and adaptability (cost,
R&D speed, workforce
constraints)



CDISC 360i Vision

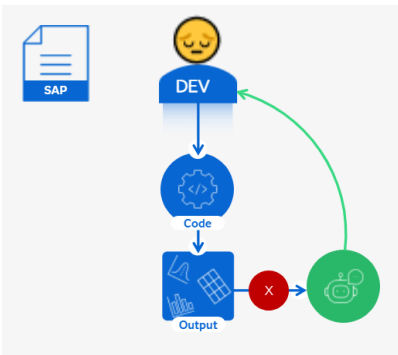
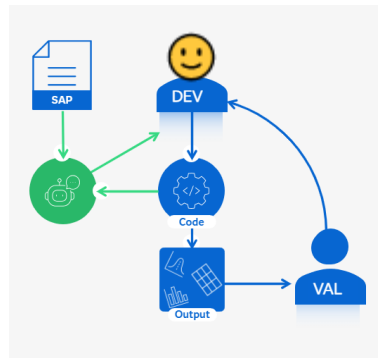
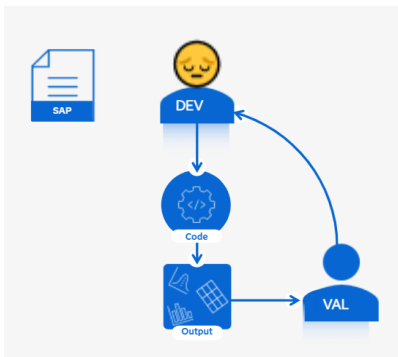
360i Digital Lifecycle Driving Automation







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

Clinical Trial
Protocol



Data
Collection



eCRF / eSource

Data
Tabulation



SDTM

Statistical
Analysis



ADaM / ARDS /
TLFs

Clinical Study
Report



Raw data to SDTM format

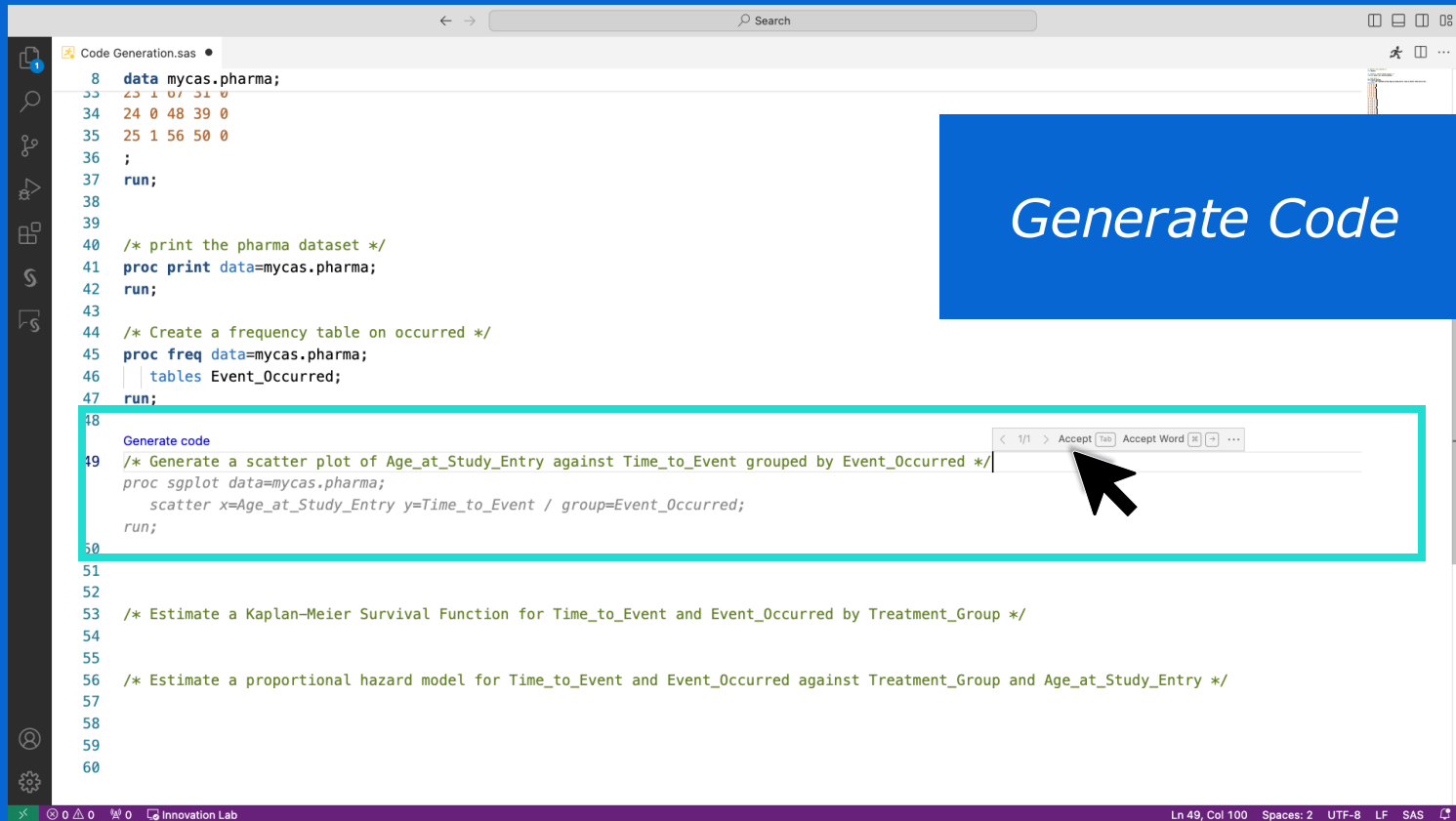
SDTM to ADaM

ADaM Validation

Creation of define.xml

Creation of TLFs

Generate Code



```
Code Generation.sas •
8  data mycas.pharma;
33  23 1 07 31 0
34  24 0 48 39 0
35  25 1 56 50 0
36  ;
37  run;
38
39
40  /* print the pharma dataset */
41  proc print data=mycas.pharma;
42  run;
43
44  /* Create a frequency table on occurred */
45  proc freq data=mycas.pharma;
46  | tables Event_Occurred;
47  run;
48
49  /* Generate a scatter plot of Age_at_Study_Entry against Time_to_Event grouped by Event_Occurred */
50  proc sgplot data=mycas.pharma;
51  | scatter x=Age_at_Study_Entry y=Time_to_Event / group=Event_Occurred;
52  run;
53
54  /* Estimate a Kaplan-Meier Survival Function for Time_to_Event and Event_Occurred by Treatment_Group */
55
56  /* Estimate a proportional hazard model for Time_to_Event and Event_Occurred against Treatment_Group and Age_at_Study_Entry */
57
58
59
60
```

Generate code

< 1/1 > Accept Tab Accept Word [X] ...

Ln 49, Col 100 Spaces: 2 UTF-8 LF SAS

The diagram illustrates the SAS Viya ecosystem and its integration with various development and deployment environments. At the center is a large blue circle representing the SAS Viya cloud platform, with the text "SAS Viya" and a cloud icon. Above this central hub is the Git logo and the text "Share and cooperate", with a double-headed arrow indicating integration. To the right of the central hub is a green hexagonal shape representing "SAS Clinical Acceleration", with a double-headed arrow indicating integration. Below the central hub is a blue rectangular shape representing "Low Code", with a double-headed arrow indicating integration. To the left of the central hub is a row of five development environments, each represented by a blue rectangular shape with a person icon and a screen showing the environment's interface: "VS Code Plugin", "Jupyter LAB", "R Studio*", and "SAS Studio/EG". A large blue arrow points from the central hub towards these development environments, with the text "CONSUME DATA FROM DATABASES, OPEN DATA FORMATS" above it. The entire diagram is set against a light blue background with a grid pattern.

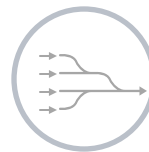
Introducing SAS® Clinical Acceleration



Integration



Automation

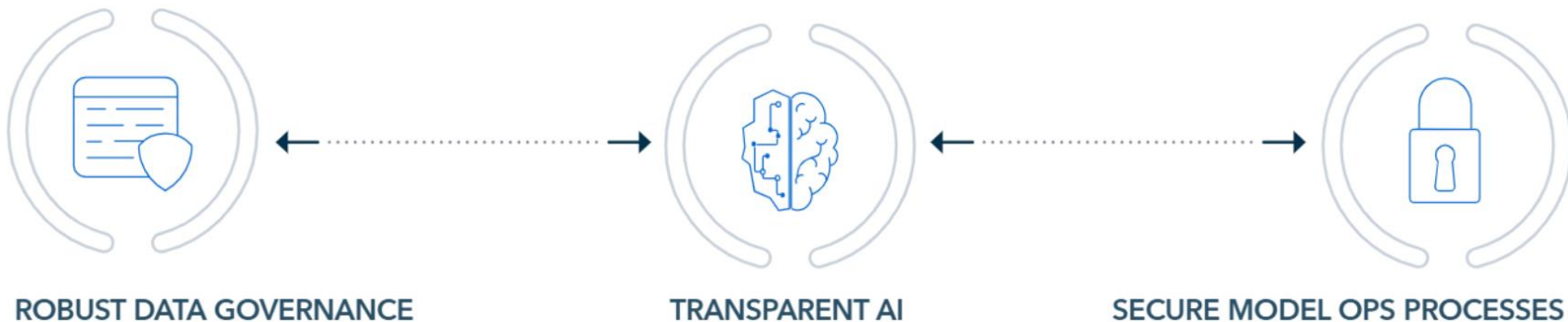


Streamline Processes

**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](https://twitter.com/phusetwitta)
📘 [/phusebook](https://facebook.com/phusebook)
▶ [/phusetube](https://youtube.com/phusetube)
🌐 [/company/phuse](https://linkedin.com/company/phuse)