verisian

# Making Your Studies AI-Ready: From Analysis to Regulatory Review

Tomás Sabat Stöfsel

CEO & Co-Founder, Verisian

# Code Traceability

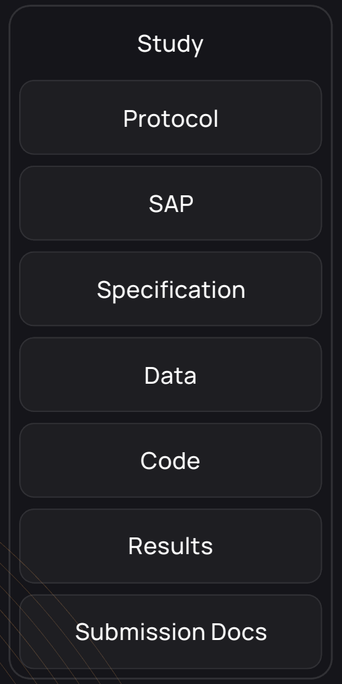
Connects your specifications, data, code, results (TLFs), and submission documentation

# Code Traceability

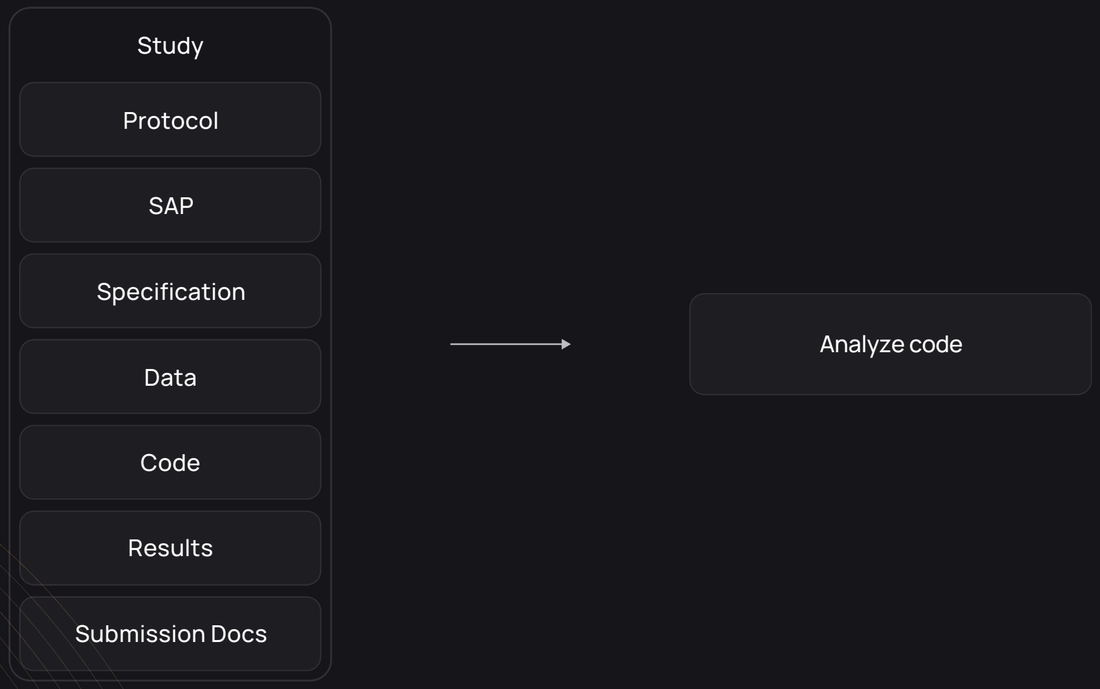
Connects your specifications, data, code, results (TLFs), and submission documentation

Independent of CDISC standards, language (R or SAS), and SCEs

# Building Traceability: From Code to Graph



# Building Traceability: From Code to Graph



# Building Traceability: From Code to Graph

Study

Protocol

SAP

Specification

Data

Code

Results

Submission Docs

Analyze code

The source of truth of a clinical trial is its code

# Variable Traceability: SDTM.DM.AGE  in a demo study

AGE is a simple concept, but can be derived in different ways

# Variable Traceability: SDTM.DM.AGE in a demo study

AGE is a simple concept, but can be derived in different ways

| dob | firstdose |
| --- | --- |
| 1974-02-05 | 2010-03-28 |
| ... | ... |
| ... | ... |
| ... | ... |
| ... | ... |

# Variable Traceability: SDTM.DM.AGE in a demo study

AGE is a simple concept, but can be derived in different ways

| dob | firstdose |
| --- | --- |
| 1974-02-05 | 2010-03-28 |
| ... | ... |
| ... | ... |
| ... | ... |
| ... | ... |

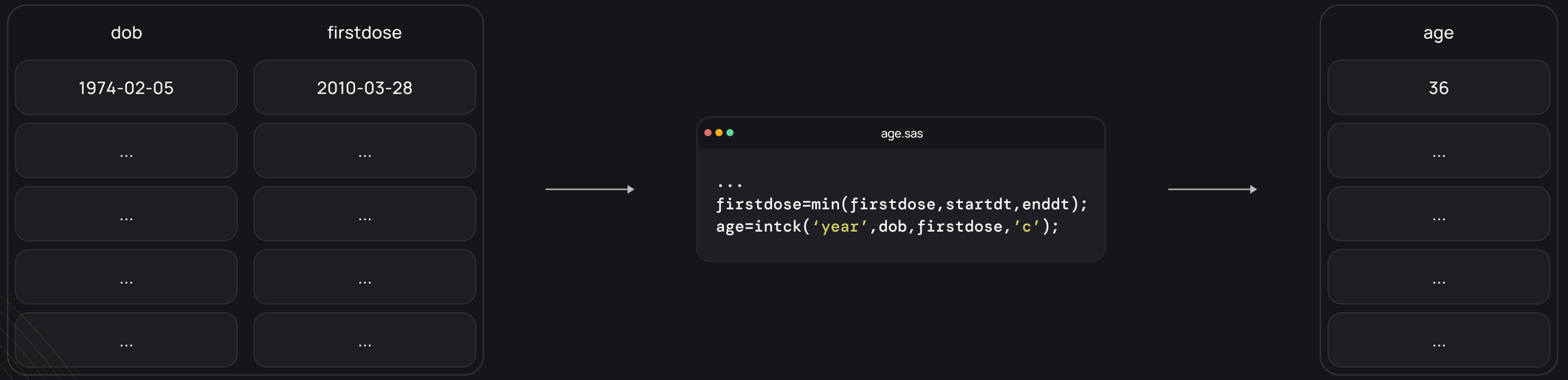


age.sas

```
...
firstdose=min(firstdose,startdt,enddt);
age=intck('year',dob,firstdose,'c');
```

# Variable Traceability: SDTM.DM.AGE in a demo study

AGE is a simple concept, but can be derived in different ways



# How is AGE derived and what data does it depend on?

- macros
  - %derive_baseline.sas
  - %dtc2dt.sas
  - %derive_age.sas
  - %flag_last.sas
  - %km_plot.sas
  - %visit_window.sas
  - +14 more
- source
  - demographic.sas
  - dosing.sas
  - labs.sas
  - adverse.sas
  - exposure.sas
  - vital_signs.sas
  - +18 more
- sdtm
  - dm.sas
  - ae.sas
  - cm.sas
  - ex.sas
  - lb.sas
  - vs.sas
  - +30 more
- adam
  - adae.sas
  - adef.sas
  - adsl.sas
  - adef.sas
  - adtte.sas
  - advs.sas
  - +16 more
- tlf
  - t14_1_1.sas
  - t14_2_1.sas
  - t14_3_1.sas
  - l162_1.sas
  - l16_2_2.sas
  - f14_2_1.sas
  - +174 more

dm.sas

```
proc sort data=source.dosing(keep=subject startdt enddt) out=dosing;
by subject startdt;
run;

data dose_sum;
set dose_raw;
by subject;

retain firstdose lastdose;

if first.subject then do;
firstdose=.;
lastdose=.;
end;

firstdose=min(firstdose,startdt,enddt);
lastdose=max(lastdose,startdt,enddt);

if last.subject;
run;

proc sort data=source.demographic out=demographic;
by subject;
run;

data demog_dose;
merge demographic dosing;
by subject;
run;

options missing=' ';

data sdtm.dm;
set EMPTY_DM demog_dose(rename=(race=_race));

studyid='XYZ123';
domain='DM';
usubjid=left(uniqueid);
subjid=put(subject,3.);
rfstdtc=put(firstdose,yymmdd10.);
rfendtc=put(lastdose,yymmdd10.);
rfxstdtc=put(firstdose,yymmdd10.);
rfxendtc=put(lastdose,yymmdd10.);
rficdtc=put(icdate,yymmdd10.);
rfpendtc=put(lastdoc,yymmdd10.);
dthfl='N';
siteid=substr(subjid,1,1) || "00";
brthdtc=put(dob,yymmdd10.);
age=intck('year', dob, firstdose, 'c');
if age ne . then ageu='YEARS';
sex=put(gender,sex_demographic_gender.);
race=put(_race,race_demographic_race.);
armcd=put(trt,armcd_demographic_trt.);
arm=put(trt,arm_demographic_trt.);
actarmcd=put(trt,armcd_demographic_trt.);
actarm=put(trt,arm_demographic_trt.);
country="USA";
run;
```

dm.sas

```
proc sort data=source.dosing(keep=subject startdt enddt) out=dosing;
by subject startdt;
run;

data dose_sum;
set dose_raw;
by subject;

retain firstdose lastdose;

if first.subject then do;
firstdose=.;
lastdose=.;
end;

firstdose=min(firstdose,startdt,enddt);
lastdose=max(lastdose,startdt,enddt);

if last.subject;
run;

proc sort data=source.demographic out=demographic;
by subject;
run;

data demog_dose;
merge demographic dosing;
by subject;
run;

options missing=' ';

data sdtm.dm;
set EMPTY_DM demog_dose(rename=(race=_race));

studyid='XYZ123';
domain='DM';
usubjid=left(uniqueid);
subjid=put(subject,3.);
rfstdtc=put(firstdose,yymmdd10.);
rfendtc=put(lastdose,yymmdd10.);
rfxstdtc=put(firstdose,yymmdd10.);
rfxendtc=put(lastdose,yymmdd10.);
rficdtc=put(icdate,yymmdd10.);
rfpendtc=put(lastdoc,yymmdd10.);
dthfl='N';
siteid=substr(subjid,1,1) || "00";
age=intck('year', dob, firstdose, 'c');
sex=put(gender,sex_demographic_gender.);
race=put(_race,race_demographic_race.);
armcd=put(trt,armcd_demographic_trt.);
arm=put(trt,arm_demographic_trt.);
actarmcd=put(trt,armcd_demographic_trt.);
actarm=put(trt,arm_demographic_trt.);
country="USA";
run;
```

dm.sas

```
proc sort data=source.dosing(keep=subject startdt enddt) out=dosing;
by subject startdt;
run;

data dose_sum;
set dose_raw;
by subject;

retain firstdose lastdose;

if first.subject then do;
firstdose=.;
lastdose=.;
end;

firstdose=min(firstdose,startdt,enddt);
lastdose=max(lastdose,startdt,enddt);

if last.subject;
run;

proc sort data=source.demographic out=demographic;
by subject;
run;

data demog_dose;
merge demographic dosing;
by subject;
run;

options missing=' ';

data sdtm.dm;
set EMPTY_DM demog_dose(rename=(race=_race));

studyid='XYZ123';
domain='DM';
usubjid=left(uniquesubj);
subjid=put(subject,3.);
rfstdtc=put(firstdose,yymmdd10.);
rfendtc=put(lastdose,yymmdd10.);
rfxstdtc=put(firstdose,yymmdd10.);
rfxendtc=put(lastdose,yymmdd10.);
rficdtc=put(icdate,yymmdd10.);
rfpendtc=put(lastdoc,yymmdd10.);
dthfl='N';
siteid=substr(subjid,1,1) || "00";
brthdtc=put(dob,yymmdd10.);
age=intck('year', dob, firstdose, 'c');
if age ne . then ageu='YEARS';
sex=put(gender,sex_demographic_gender.);
race=put(_race,race_demographic_race.);
armcd=put(trt,armcd_demographic_trt.);
arm=put(trt,arm_demographic_trt.);
actarmcd=put(trt,armcd_demographic_trt.);
actarm=put(trt,arm_demographic_trt.);
country="USA";
run;
```

# How is AGE derived and what data does it depend on?

- macros
  - %derive_baseline.sas
  - %dtc2dt.sas
  - %derive_age.sas
  - %flag_last.sas
  - %km_plot.sas
  - %visit_window.sas
  - +14 more
- source
  - demographic.sas
  - dosing.sas
  - labs.sas
  - adverse.sas
  - exposure.sas
  - vital_signs.sas
  - +18 more
- sdtm
  - dm.sas
  - ae.sas
  - cm.sas
  - ex.sas
  - lb.sas
  - vs.sas
  - +30 more
- adam
  - adae.sas
  - adef.sas
  - adsl.sas
  - adef.sas
  - adtte.sas
  - advs.sas
  - +16 more
- tlf
  - t14_1_1.sas
  - t14_2_1.sas
  - t14_3_1.sas
  - l162_1.sas
  - l16_2_2.sas
  - f14_2_1.sas
  - +174 more

demographics.sas

```
proc format;
value trt 1="Active" 0="Placebo";
value gender 1="Male" 2="Female";
value race 1="White" 2="Black" 3="Other";
run;

data source.demographic;
set raw.demographic;
label subject="Subject Number" trt="Treatment" gender="Gender"
race="Race"
orace="Other Race Specify" dob="Date of Birth" uniqueid=
"Company Wide Subject ID" randdt="Randomization Date" icdate=
"Informed Consent Date" lastdoc="Last date of contact";
input subject 1-3 trt 5 gender 7 race 9 orace $ 11-20 +1 dob
mmddyy10. +1
randdt mmddyy10. +1 icdate mmddyy10. +1 lastdoc mmddyy10.;
uniqueid='UNI' || put(subject,3.);
format dob randdt icdate lastdoc mmddyy10.;
run;

data _null_;
rc = filename('xptdir', "&xptdir");
if fexist('xptdir') = 0 then rc = dcreate("xpt",
"/mnt/viya-share/homes/demo-study/data/xpt");
run;

libname xpt_demo xport "&xptdir./demographic.xpt";
data xpt_demo.demograp;
set source.demographic;
run;
libname xpt_demo clear;
```

dosing.sas

```
* dosing.sas;
data source.dosing;
label subject="Subject Number" startdt="Dosing start date" enddt=
"Dosing end date" ddose="Daily dose taken (pills)" uniqueid=
"Company Wide Subject ID" startmm="Month of Start Dose" startdd=
"Day of Start Dose" startyy="Year of Start Dose" endmm=
"Month of End Dose" enddd="Day of End Dose" endyy=
"Year of End Dose";
input subject 1-3 startmm 5-6 startdd 8-9 startyy 11-14 endmm 16-17
enddd
19-20 endyy 22-25 ddose 27;
uniqueid='UNI' || put(subject,3.);
startdt=mdy(startmm, startdd , startyy);
enddt=mdy(endmm, enddd, endyy);
format startdt enddt mmddyy10.;
datalines;
;
run;
```

dm.sas

```
proc sort data=source.dosing(keep=subject startdt enddt) out=dosing;
by subject startdt;
run;

data dose_sum;
set dose_raw;
by subject;

retain firstdose lastdose;

if first.subject then do;
firstdose=.;
lastdose=.;
end;

firstdose=min(firstdose,startdt,enddt);
lastdose=max(lastdose,startdt,enddt);

if last.subject;
run;

proc sort data=source.demographic out=demographic;
by subject;
run;

data demog_dose;
merge demographic dosing;
by subject;
run;

options missing=' ';

data sdtm.dm;
set EMPTY_DM demog_dose(rename=(race=_race));

studyid='XYZ123';
domain='DM';
usubjid=left(uniqueid);
subjid=put(subject,3.);
rfstdtc=put(firstdose,yymmdd10.);
rfendtc=put(lastdose,yymmdd10.);
rfxstdtc=put(firstdose,yymmdd10.);
rfxendtc=put(lastdose,yymmdd10.);
rficdtc=put(icdate,yymmdd10.);
rfpendtc=put(lastdoc,yymmdd10.);
dthfl='N';
siteid=substr(subjid,1,1) || "00";
brthdtc=put(dob,yymmdd10.);
age=intck('year', dob, firstdose, 'c');
if age ne . then ageu='YEARS';
sex=put(gender,sex_demographic_gender.);
race=put(_race,race_demographic_race.);
armcd=put(trt,armcd_demographic_trt.);
arm=put(trt,arm_demographic_trt.);
actarmcd=put(trt,armcd_demographic_trt.);
actarm=put(trt,arm_demographic_trt.);
country="USA";
run;
```

demographics.sas

```
proc format;
value trt 1="Active" 0="Placebo";
value gender 1="Male" 2="Female";
value race 1="White" 2="Black" 3="Other";
run;

data source.demographic;
set raw.demographic;
label subject="Subject Number" trt="Treatment" gender="Gender"
race="Race"
orace="Other Race Specify" dob="Date of Birth" uniqueid=
"Company Wide Subject ID" randdt="Randomization Date" icdate=
"Informed Consent Date" lastdoc="Last date of contact";
input subject 1-3 trt 5 gender 7 race 9 orace $ 11-20 +1 dob
mmddyy10. +1
randdt mmddyy10. +1 icdate mmddyy10. +1 lastdoc mmddyy10.;
uniqueid='UNI' || put(subject,3.);
format dob randdt icdate lastdoc mmddyy10.;
run;

data _null_;
rc = filename('xptdir', "&xptdir");
if fexist('xptdir') = 0 then rc = dcreate("xpt",
"/mnt/viya-share/homes/demo-study/data/xpt");
run;

libname xpt_demo xport "&xptdir./demographic.xpt";
data xpt_demo.demograp;
set source.demographic;
run;
libname xpt_demo clear;
```

dosing.sas

```
* dosing.sas;
data source.dosing;
label subject="Subject Number" startdt="Dosing start date" enddt=
"Dosing end date" ddose="Daily dose taken (pills)" uniqueid=
"Company Wide Subject ID" startmm="Month of Start Dose" startdd=
"Day of Start Dose" startyy="Year of Start Dose" endmm=
"Month of End Dose" enddd="Day of End Dose" endyy=
"Year of End Dose";
input subject 1-3 startmm 5-6 startdd 8-9 startyy 11-14 endmm 16-17
enddd
19-20 endyy 22-25 ddose 27;
uniqueid='UNI' || put(subject,3.);
startdt=mdy(startmm, startdd , startyy);
enddt=mdy(endmm, enddd, endyy);
format startdt enddt mmddyy10.;
datalines;
;
run;
```

dm.sas

```
proc sort data=source.dosing(keep=subject startdt enddt) out=dosing;
by subject startdt;
run;

data dose_sum;
set dose_raw;
by subject;

retain firstdose lastdose;

if first.subject then do;
firstdose=.;
lastdose=.;
end;

firstdose=min(firstdose,startdt,enddt);
lastdose=max(lastdose,startdt,enddt);

if last.subject;
run;

proc sort data=source.demographic out=demographic;
by subject;
run;

data demog_dose;
merge demographic dosing;
by subject;
run;

options missing=' ';

data sdtm.dm;
set EMPTY_DM demog_dose(rename=(race=_race));

studyid='XYZ123';
domain='DM';
usubjid=left(uniqueid);
subjid=put(subject,3.);
rfstdtc=put(firstdose,yymmdd10.);
rfendtc=put(lastdose,yymmdd10.);
rfxstdtc=put(firstdose,yymmdd10.);
rfxendtc=put(lastdose,yymmdd10.);
rficdtc=put(icdate,yymmdd10.);
rfpendtc=put(lastdoc,yymmdd10.);
dthfl='N';
siteid=substr(subjid,1,1) || "00";
brthdtc=put(dob,yymmdd10.);
age=intck('year'. dob. firstdose.'c');
if age ne . then ageu='YEARS';
sex=put(gender,sex_demographic_gender.);
race=put(_race,race_demographic_race.);
armcd=put(trt,armcd_demographic_trt.);
arm=put(trt,arm_demographic_trt.);
actarmcd=put(trt,armcd_demographic_trt.);
actarm=put(trt,arm_demographic_trt.);
country="USA";
run;
```

demographics.sas

```
proc format;
value trt 1="Active" 0="Placebo";
value gender 1="Male" 2="Female";
value race 1="White" 2="Black" 3="Other";
run;

data source.demographic;
set raw.demographic;
label subject="Subject Number" trt="Treatment" gender="Gender"
race="Race"
orace="Other Race Specify" dob="Date of Birth" uniqueid=
"Company Wide Subject ID" randdt="Randomization Date" icdate=
"Informed Consent Date" lastdoc="Last date of contact";
input subject 1-3 trt 5 gender 7 race 9 orace $ 11-20 +1 dob
mmddyy10. +1
randdt mmddyy10. +1 icdate mmddyy10. +1 lastdoc mmddyy10.;
uniqueid='UNI' || put(subject,3.);
format dob randdt icdate lastdoc mmddyy10.;
run;

data _null_;
rc = filename('xptdir', "&xptdir");
if fexist('xptdir') = 0 then rc = dcreate("xpt",
"/mnt/viya-share/homes/demo-study/data/xpt");
run;

libname xpt_demo xport "&xptdir./demographic.xpt";
data xpt_demo.demograp;
set source.demographic;
run;
libname xpt_demo clear;
```

dosing.sas

```
* dosing.sas;
data source.dosing;
label subject="Subject Number" startdt="Dosing start date" enddt=
"Dosing end date" ddose="Daily dose taken (pills)" uniqueid=
"Company Wide Subject ID" startmm="Month of Start Dose" startdd=
"Day of Start Dose" startyy="Year of Start Dose" endmm=
"Month of End Dose" enddd="Day of End Dose" endyy=
"Year of End Dose";
input subject 1-3 startmm 5-6 startdd 8-9 startyy 11-14 endmm 16-17
enddd
19-20 endyy 22-25 ddose 27;
uniqueid='UNI' || put(subject,3.);
startdt=mdy(startmm, startdd , startyy);
enddt=mdy(endmm, enddd, endyy);
format startdt enddt mmddyy10.;
datalines;
;
run;
```

dm.sas

```
proc sort data=source.dosing(keep=subject startdt enddt) out=dosing;
by subject startdt;
run;

data dose_sum;
set dose_raw;
by subject;

retain firstdose lastdose;

if first.subject then do;
firstdose=.;
lastdose=.;
end;

firstdose=min(firstdose,startdt,enddt);
lastdose=max(lastdose,startdt,enddt);

if last.subject;
run;

proc sort data=source.demographic out=demographic;
by subject;
run;

data demog_dose;
merge demographic dosing;
by subject;
run;

options missing=' ';

data sdtm.dm;
set EMPTY_DM demog_dose(rename=(race=_race));

studyid='XYZ123';
domain='DM';
usubjid=left(uniqueid);
subjid=put(subject,3.);
rfstdtc=put(firstdose,yymmdd10.);
rfendtc=put(lastdose,yymmdd10.);
rfxstdtc=put(firstdose,yymmdd10.);
rfxendtc=put(lastdose,yymmdd10.);
rficdtc=put(icdate,yymmdd10.);
rfpendtc=put(lastdoc,yymmdd10.);
dthfl='N';
siteid=substr(subjid,1,1) || "00";
brthdtc=put(dob,yymmdd10.);
age=intck('year'. dob. firstdose.'c');
if age ne . then ageu='YEARS';
sex=put(gender,sex_demographic_gender.);
race=put(_race,race_demographic_race.);
armcd=put(trt,armcd_demographic_trt.);
arm=put(trt,arm_demographic_trt.);
actarmcd=put(trt,armcd_demographic_trt.);
actarm=put(trt,arm_demographic_trt.);
country="USA";
run;
```

demographics.sas

```
proc format;
value trt 1="Active" 0="Placebo";
value gender 1="Male" 2="Female";
value race 1="White" 2="Black" 3="Other";
run;

data source.demographic;
set raw.demographic;
label subject="Subject Number" trt="Treatment" gender="Gender"
race="Race"
orace="Other Race Specify" dob="Date of Birth" uniqueid=
"Company Wide Subject ID" randdt="Randomization Date" icdate=
"Informed Consent Date" lastdoc="Last date of contact";
input subject 1-3 trt 5 gender 7 race 9 orace $ 11-20 +1 dob
mmddyy10. +1
randdt mmddyy10. +1 icdate mmddyy10. +1 lastdoc mmddyy10.;
uniqueid='UNI' || put(subject,3.);
format dob randdt icdate lastdoc mmddyy10.;
run;

data _null_;
rc = filename('xptdir', "&xptdir");
if fexist('xptdir') = 0 then rc = dcreate('xpt',
"/mnt/viya-share/homes/demo-study/data/xpt");
run;

libname xpt_demo xport "&xptdir./demographic.xpt";
data xpt_demo.demograp;
set source.demographic;
run;
libname xpt_demo clear;
```

dosing.sas

```
* dosing.sas;
data source.dosing;
label subject="Subject Number" startdt="Dosing start date" enddt=
"Dosing end date" ddose="Daily dose taken (pills)" uniqueid=
"Company Wide Subject ID" startmm="Month of Start Dose" startdd=
"Day of Start Dose" startyy="Year of Start Dose" endmm=
"Month of End Dose" enddd="Day of End Dose" endyy=
"Year of End Dose";
input subject 1-3 startmm 5-6 startdd 8-9 startyy 11-14 endmm 16-17
enddd
19-20 endyy 22-25 ddose 27;
uniqueid='UNI' || put(subject,3.);
startdt=mdy(startmm, startdd , startyy);
enddt=mdy(endmm, enddd, endyy);
format startdt enddt mmddyy10.;
datalines;
;
run;
```

dm.sas

```
proc sort data=source.dosing(keep=subject startdt enddt) out=dosing;
by subject startdt;
run;

data dose_sum;
set dose_raw;
by subject;

retain firstdose lastdose;

if first.subject then do;
firstdose=.;
lastdose=.;
end;

firstdose=min(firstdose,startdt,enddt);
lastdose=max(lastdose,startdt,enddt);

if last.subject;
run;

proc sort data=source.demographic out=demographic;
by subject;
run;

data demog_dose;
merge demographic dosing;
by subject;
run;

options missing=' ';

data sdtm.dm;
set EMPTY_DM demog_dose(rename=(race=_race));

studyid='XYZ123';
domain='DM';
usubjid=left(uniqueid);
subjid=put(subject,3.);
rfstdtc=put(firstdose,yymmdd10.);
rfendtc=put(lastdose,yymmdd10.);
rfxstdtc=put(firstdose,yymmdd10.);
rfxendtc=put(lastdose,yymmdd10.);
rficdtc=put(icdate,yymmdd10.);
rfpendtc=put(lastdoc,yymmdd10.);
dthfl='N';
siteid=substr(subjid,1,1) || "00";
brthdtc=put(dob,yymmdd10.);
age=intck('year'. dob. firstdose.'c');
if age ne . then ageu='YEARS';
sex=put(gender,sex_demographic_gender.);
race=put(_race,race_demographic_race.);
armcd=put(trt,armcd_demographic_trt.);
arm=put(trt,arm_demographic_trt.);
actarmcd=put(trt,armcd_demographic_trt.);
actarm=put(trt,arm_demographic_trt.);
country="USA";
run;
```

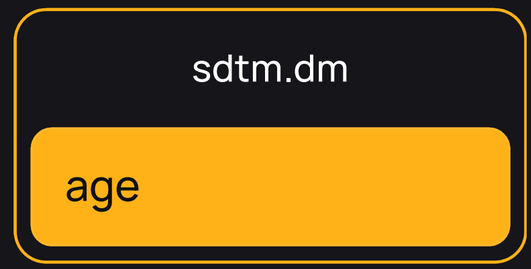
# Trace code in files

Trace code in files

Build a mental representation of the logic

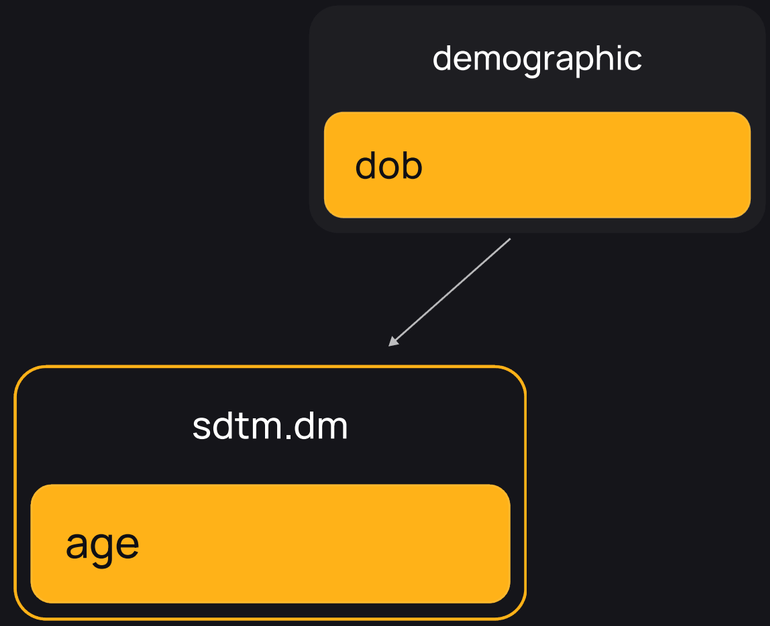
# What variables does AGE depend on?

```
data sdtm.dm;
age=intck('year',dob,firstdose,'c');
```



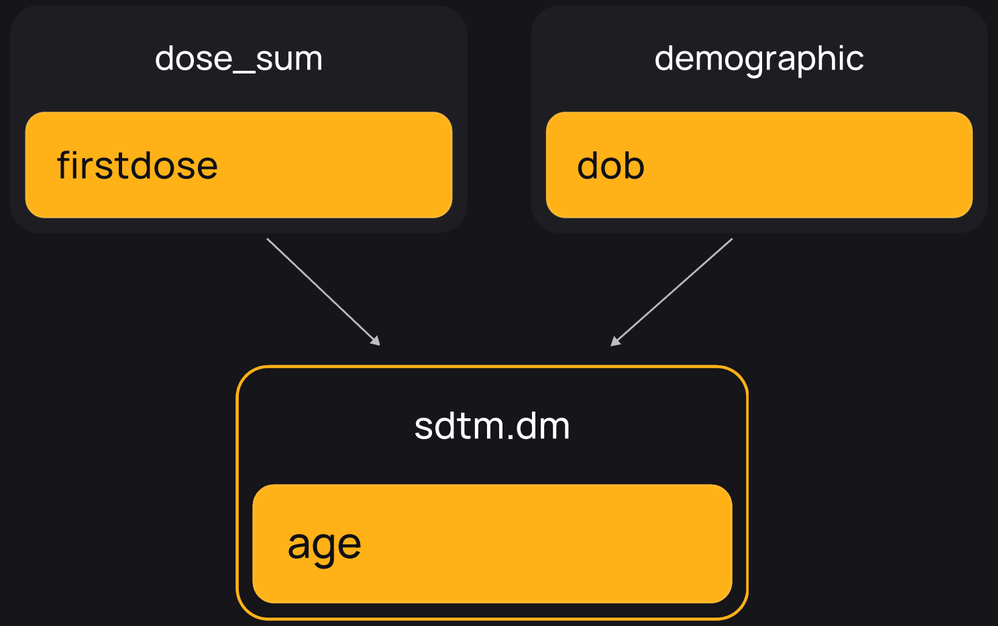
# What variables does AGE depend on?

```
data demographic;
set raw_demographics;

data sdtm.dm;
set demographic dose_sum;
age=intck('year', dob, firstdose, 'c');
```



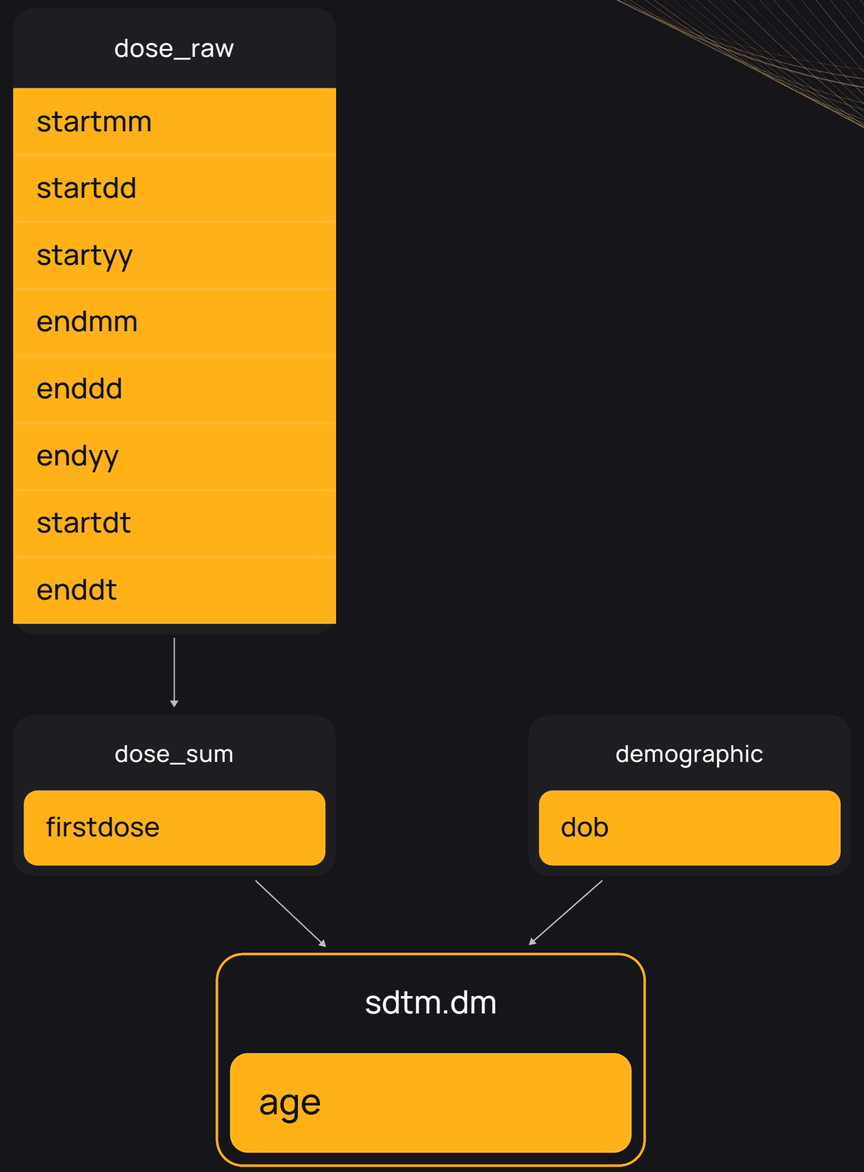
# What variables does AGE depend on?

```
data dose_sum;
if first.subject then do;
firstdose=.;
firstdose=min(firstdose,startdt,enddt);

data demographic;
set raw_demographics;

data sdtm.dm;
set demographic dose_sum;
age=intck('year', dob, firstdose, 'c');
```



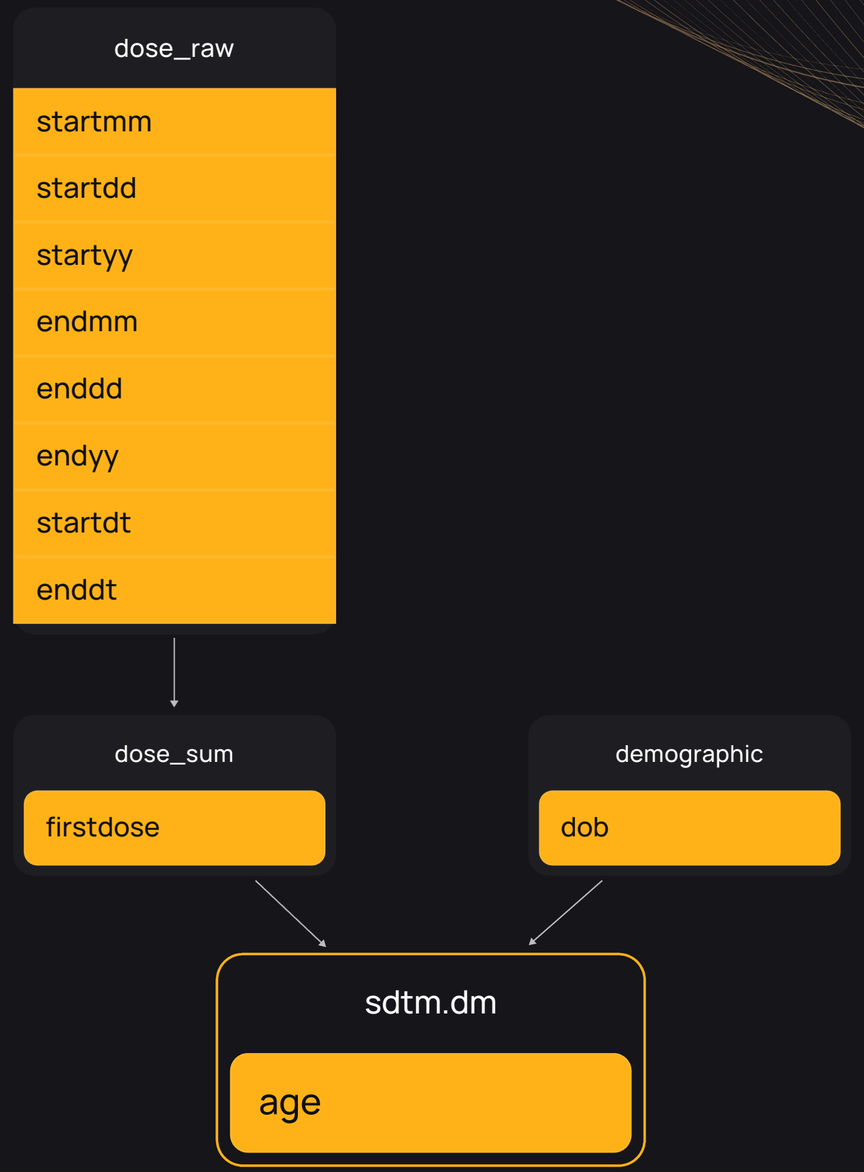
# What variables does AGE depend on?

```
data dose_raw;
startdt=mdy(startmm, startdd , startyy);
enddt=mdy(endmm, enddd, endyy);
run;

data dose_sum;
if first.subject then do;
firstdose=.;
firstdose=min(firstdose,startdt,enddt);

data demographic;
set raw_demographics;

data sdtm.dm;
set demographic dose_sum;
age=intck('year', dob, firstdose, 'c');
```



# What variables does AGE depend on?

| Dataset | Variable |
|---|---|
| dose_raw | startmm |
| dose_raw | startdd |
| dose_raw | startyy |
| dose_raw | endmm |
| dose_raw | enddd |
| dose_raw | endyy |
| dose_raw | startdt |
| dose_raw | enddt |
| dose_raw | firstdose |
| demographic | dob |



# What is age?

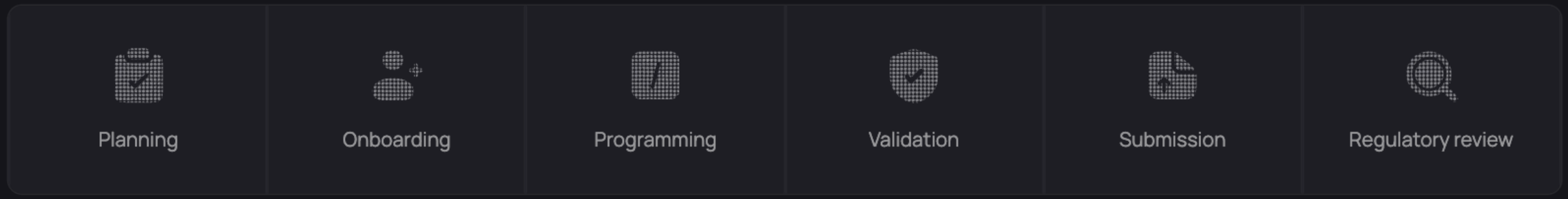

dob
firstdose
1974-02-05
2010-03-28
...
...
...
...
...
...
...
...
...
age.sas
...
firstdose=min(firstdose,startdt,enddt);
age=intck('year',dob,firstdose,'c');
age
36
...
...
...
...

Perfect transparency, trust and confidence in submissions

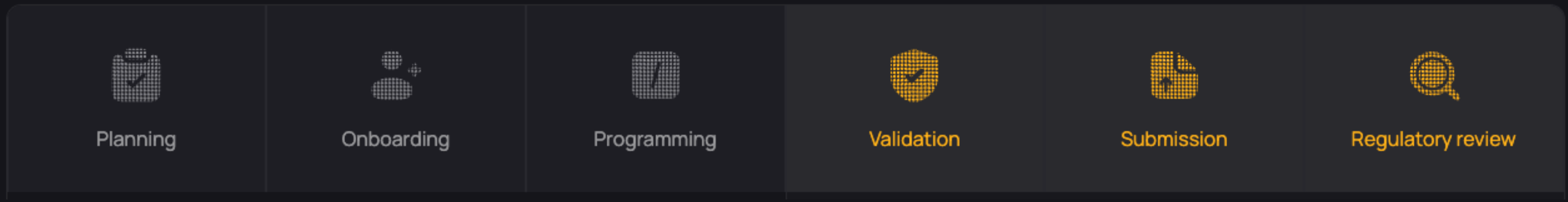
Handle more studies, faster, with fewer resources at higher quality

Lower cost

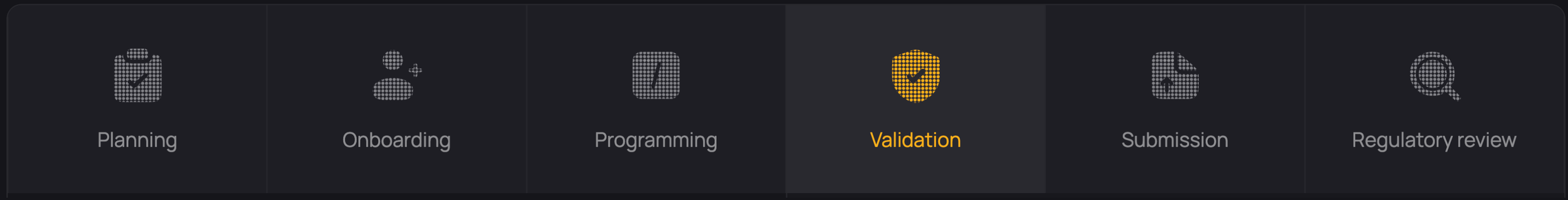
# From planning to regulatory review



# Validation, submission & regulatory review



# Automate validation of study analyses



# Automate validation of study analyses

1. Validating that code matches specs

# Automate validation of study analyses

1. Validating that code matches specs
2. Validating that data is correct

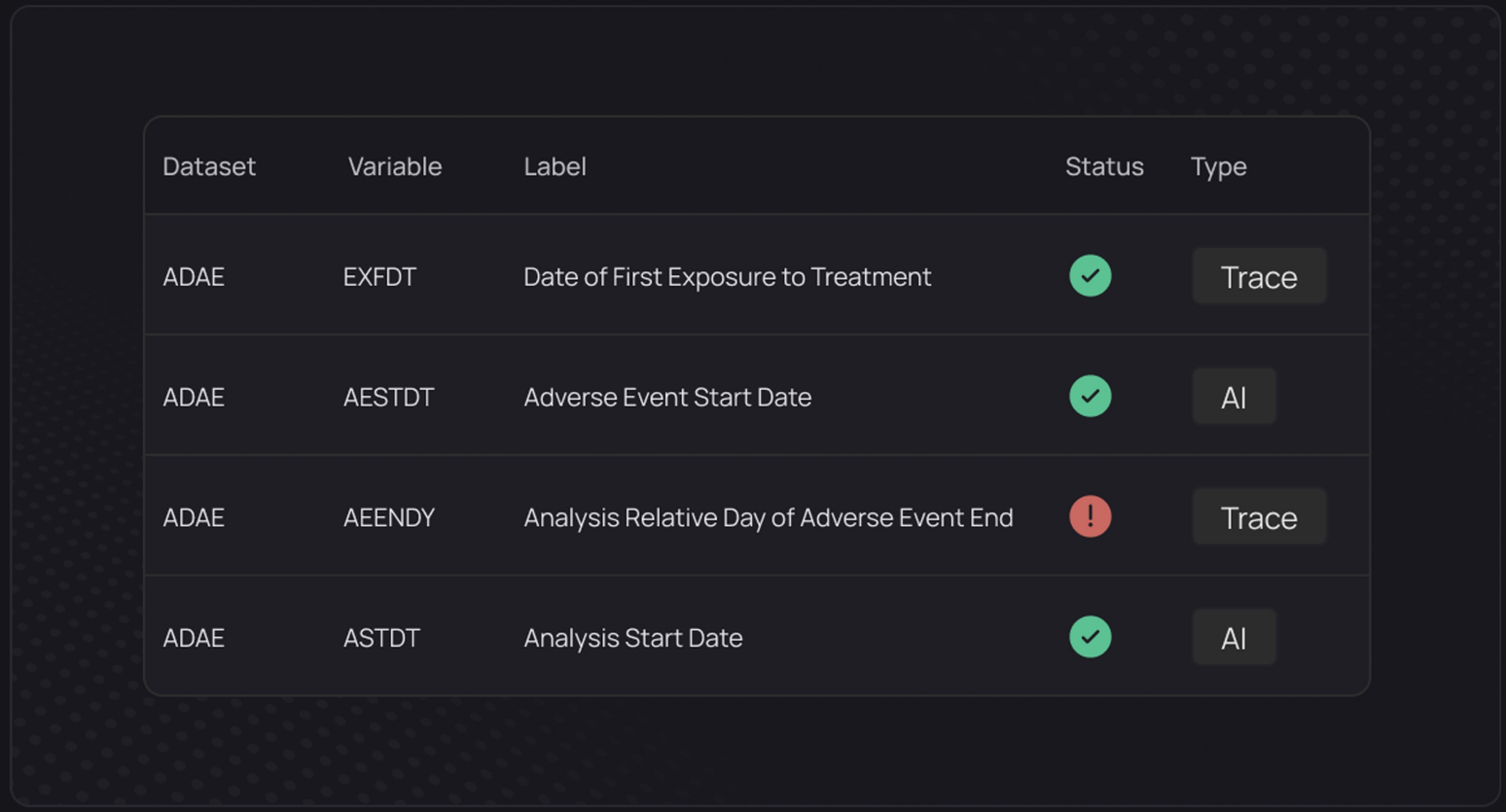
# Automate validation of study analyses

1. Validating that code matches specs
2. Validating that data is correct
3. Validating outputs

# Automate validation of study analyses

1. Validating that code matches specs
2. Validating that data is correct
3. Validating outputs

Automated by 95%+



| Dataset | Variable | Label | Status | Type |
|---|---|---|---|---|
| ADAE | EXFDT | Date of First Exposure to Treatment | ✓ | Trace |
| ADAE | AESTDT | Adverse Event Start Date | ✓ | AI |
| ADAE | AEENDY | Analysis Relative Day of Adverse Event End | ! | Trace |
| ADAE | ASTDT | Analysis Start Date | ✓ | AI |

# Validating that code matches specs

1. Constants

# Validating that code matches specs

1. Constants
2. Mappings

# Validating that code matches specs

1. Constants
2. Mappings
3. Standard macros

# Validating that code matches specs

1. Constants
2. Mappings
3. Standard macros
4. Study-specific derivations

# Validating that code matches specs

1. Constants
2. Mappings
3. Standard macros
4. Study-specific derivations

Goal is to fully automate validation of SDTM and ADaM analyses

# Validating that data is correct

1. CDISC CORE

# Validating that data is correct

1. CDISC CORE
2. Missing values

# Validating that data is correct

1. CDISC CORE
2. Missing values
3. Outliers

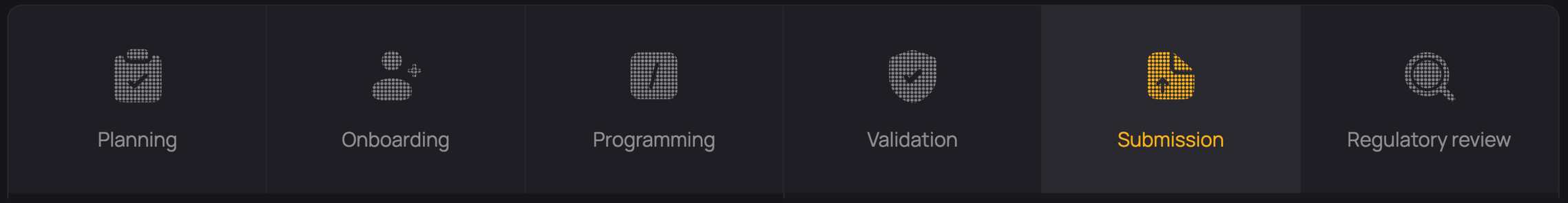
# Validating that data is correct

1. CDISC CORE
2. Missing values
3. Outliers
4. Codelist mismatches

# Validating outputs

Tracing from output to raw data

# Automate submission documentation



# Submission documentation is a bottleneck

Reviewer's Guides and Define-XMLs are time-consuming and error-prone to produce

- Documentation is often inconsistent with the actual analysis and data
- Documentation is started after analyses are completed, delaying submission readiness

Inconsistencies confuse regulators, slow review, and delay time-to-market

# Verisian AI to generate submission documentation

Bayer applied Verisian to build full study traceability and automate submission documentation

- Documentation was generated from the code that was actually run, not the specifications
- Guaranteed consistency between code, data, results, and documentation
- Powered by deterministic, fully reproducible graph theory

"The Verisian AI infrastructure is enabling Bayer to take big steps towards delivering our data caterer vision, where 80% of the busywork is automated by trustworthy and evidence-based AI."

Sascha Ahrweiler, VP, Head Statistical Programming & Analysis - Strategy & Operations

# What Verisian AI delivered to Bayer

On a non-CDISC clinical trial for a NMPA Full Approval Submission:

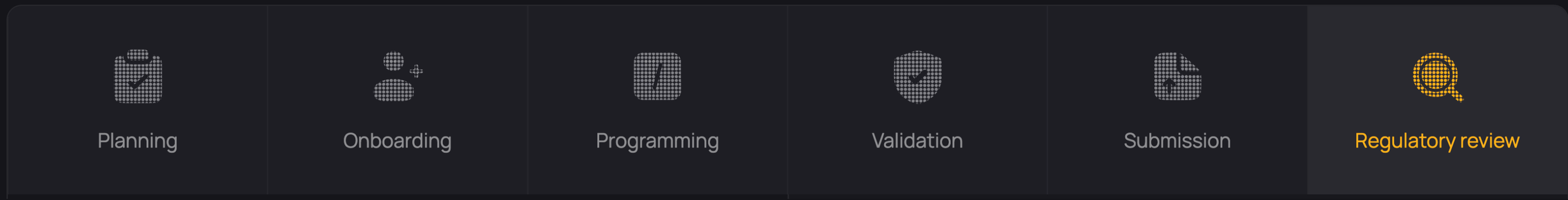
- Complete ADaM-level Define-XML generated, 90-95% at submission quality
- Crucial Reviewer's Guide sections generated at submission quality
- Generated in minutes

“While the computational methods for complicated efficacy endpoint variables are not perfect and require a review, overall about 90-95% are good enough to submit as-is. The important sections in the ADRG that needed an update from the last NMPA submission, mainly those concerned with explaining dataset lineage and summarizing their inputs, content, and ultimate use, were generated at submission quality as well.”

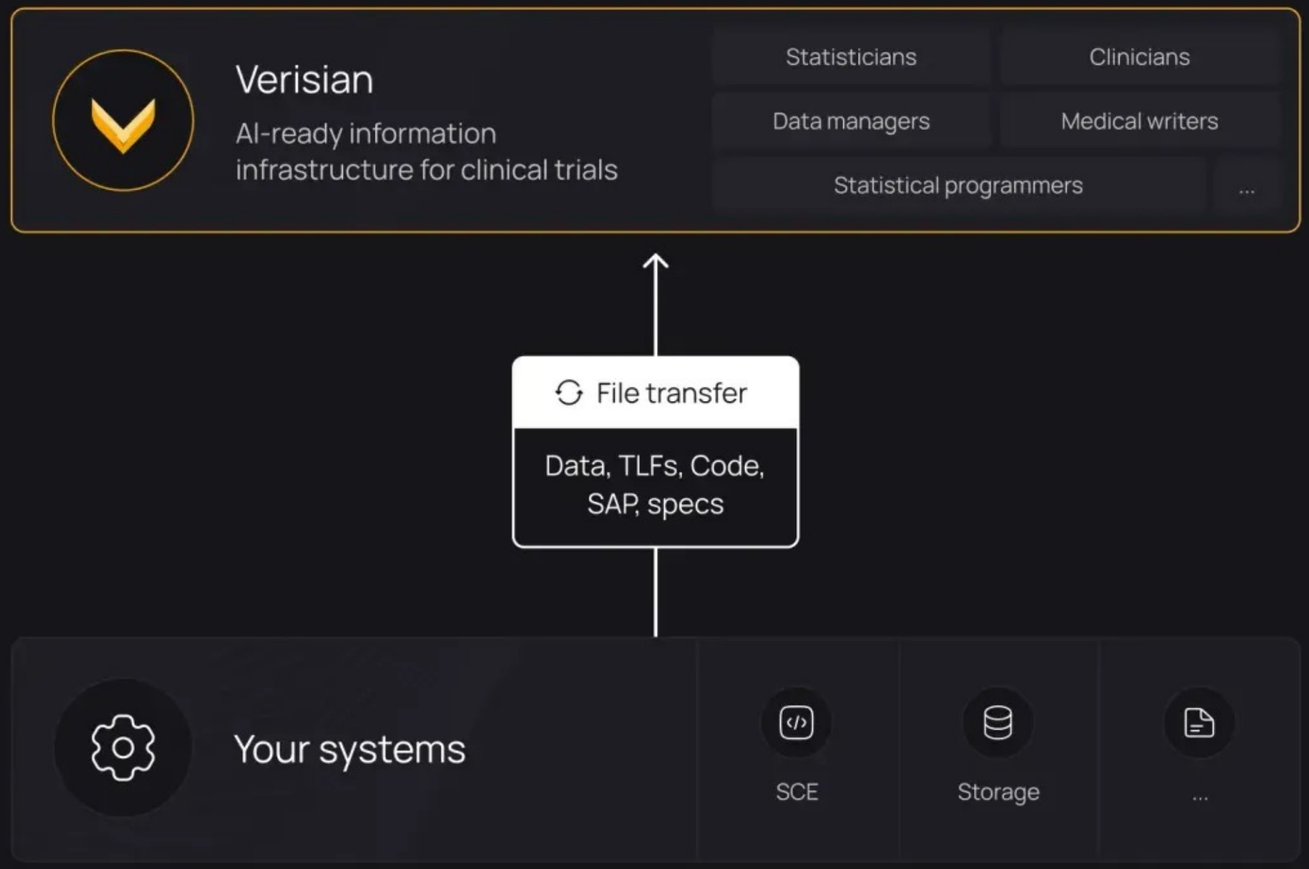
"Given the short timeframe and low resource availability, this would not have been possible without Verisian.”
Cornelia Fulgenzi, Compound Lead, Bayer

# Augmented submissions

# Faster approvals



Make your studies AI-ready in one day


Verisian
AI-ready information infrastructure for clinical trials
Statisticians
Clinicians
Data managers
Medical writers
Statistical programmers
...
File transfer
Data, TLFs, Code, SAP, specs
Your systems
SCE
Storage
...

# Across any operating model

In-house teams

CRO outsourcing

# ...and your entire study team

Perfect transparency, trust and confidence in submissions

Perfect transparency, trust and confidence in submissions

Handle more studies, faster, with fewer resources at higher quality

Perfect transparency, trust and confidence in submissions

Handle more studies, faster, with fewer resources at higher quality

Lower cost