



9th–13th November

EU Connect 2020

The Clinical Data
Science Conference

Data Integrity & Standards in the Clinical Data Pipeline

Presented By:

Mark Wheeldon
CEO



Raj Indupuri
CEO





Data Integrity & Standards in the Clinical Data Pipeline:

STANDARDS DRIVEN EFFICIENCIES POWERED BY NEW TECHNOLOGY

Mark Wheeldon | CEO Formedix
November 2020

formedix.com

OUR CREDENTIALS

Cloud based clinical MDR and study automation platform

- 2 decades of CDISC standards knowledge & expertise.
- 20+ years developing clinical trial software tools.
- Launched cloud based platform, ryze in 2020.
- Standards enabled technology partnerships.
- Trusted by pharmas, biotechs & CROs.



CDISC THEN AND NOW

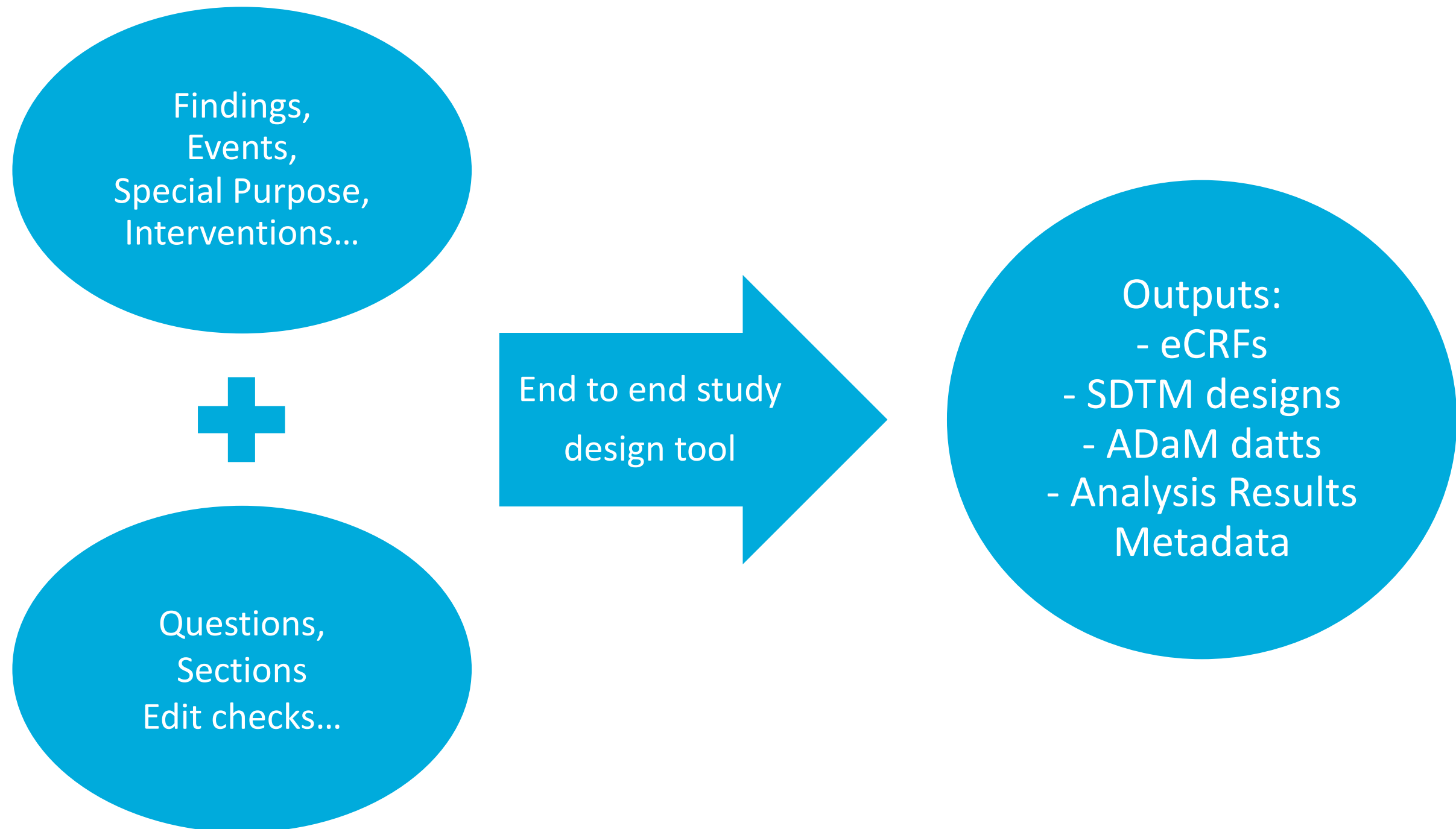
CDISC before

- No standardization = slowed down submissions & hard to compare data.
- Submission focused - only satisfying end regulation.
- Compliance can be a tick box exercise.
- You need Define – don't have it – rip it from existing studies.

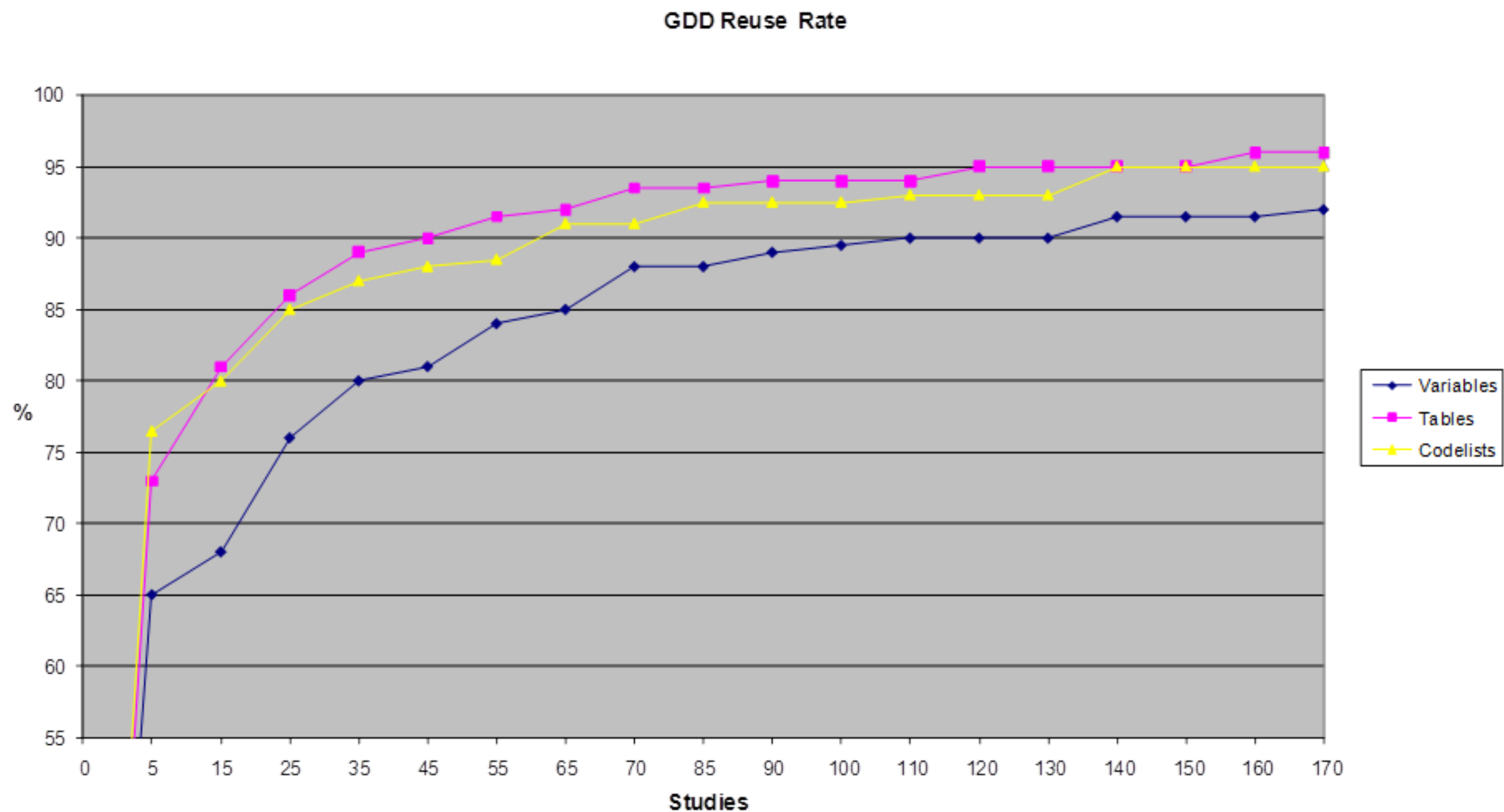
CDISC implementations should focus on:

- Content reuse.
- Efficiencies & better data quality.
- Faster study start up.
- Quicker SDTM conversions.
- Design with the end in mind.

USING CDISC TO DESIGN WITH THE END IN MIND



THE BENEFIT OF STANDARDS REUSE ACROSS STUDIES



Source: Schering Dublin May 2003

FLEXIBLE END TO END STANDARDS ORGANIZATION

Dashboard

Repository

Services

Downloads

Admin

Support

Standards

⊕ CREATE STANDARD?

AE (Adverse Events)

Updated Sep 16, 2020

Status In progress

Version 1

Asset Rules Compliance Non-Standard

View

Edit

Transition this standard

CM (Concomitant Medications)

Updated Sep 16, 2020

Status Draft

Version 2

Asset Rules Compliance Non-Standard

DM (Demographics)

Updated Sep 16, 2020

Status Draft

Version 1

Asset Rules Compliance Non-Standard

EX (Exposure)

Updated Sep 16, 2020

Status Draft

Version 3

Asset Rules Compliance Non-Standard

IE (Inclusion Exclusion Criteria)

Updated Sep 16, 2020

Status In review

Version 1

Asset Rules Compliance Non-Standard

MH (Medical History)

Updated Sep 16, 2020

Status In progress

Version 4

Asset Rules Compliance Non-Standard

VS (Vital Signs)

Updated Sep 16, 2020

Status Draft

Version 1

Asset Rules Compliance Non-Standard

COMPLETE STUDY & STANDARDS ASSET MANAGEMENT

DM (Demographics)

[Standards](#) >
 [DM \(Demographics\) \(Version 1 - Draft\)](#) >

Asset groups

ADaM (DM) (ADaM datasets)	
DM (Forms) (Data acquisition)	
SDTM (DM) (SDTM datasets) ✓	

+ Create asset group

Properties

DM (Demographics)
Version 1

Status	Draft	Updated	Sep 16, 2020 4:24:32 PM
Name	Demographics	Compliance	Compliant ✓
Identifier	STD_2	Therapeutic area	None
Sponsor	None		
Program	None		

Assets

All (279 assets)

- Visit schedules +
- Visits +
- Forms (1 asset) +
- Sections (1 asset) +
- Questions (2 assets) +
- Units +
- Conditions +
- Datasets (2 assets) +
- Variables (46 assets) +
- Values +

Change requests

ACTIVE

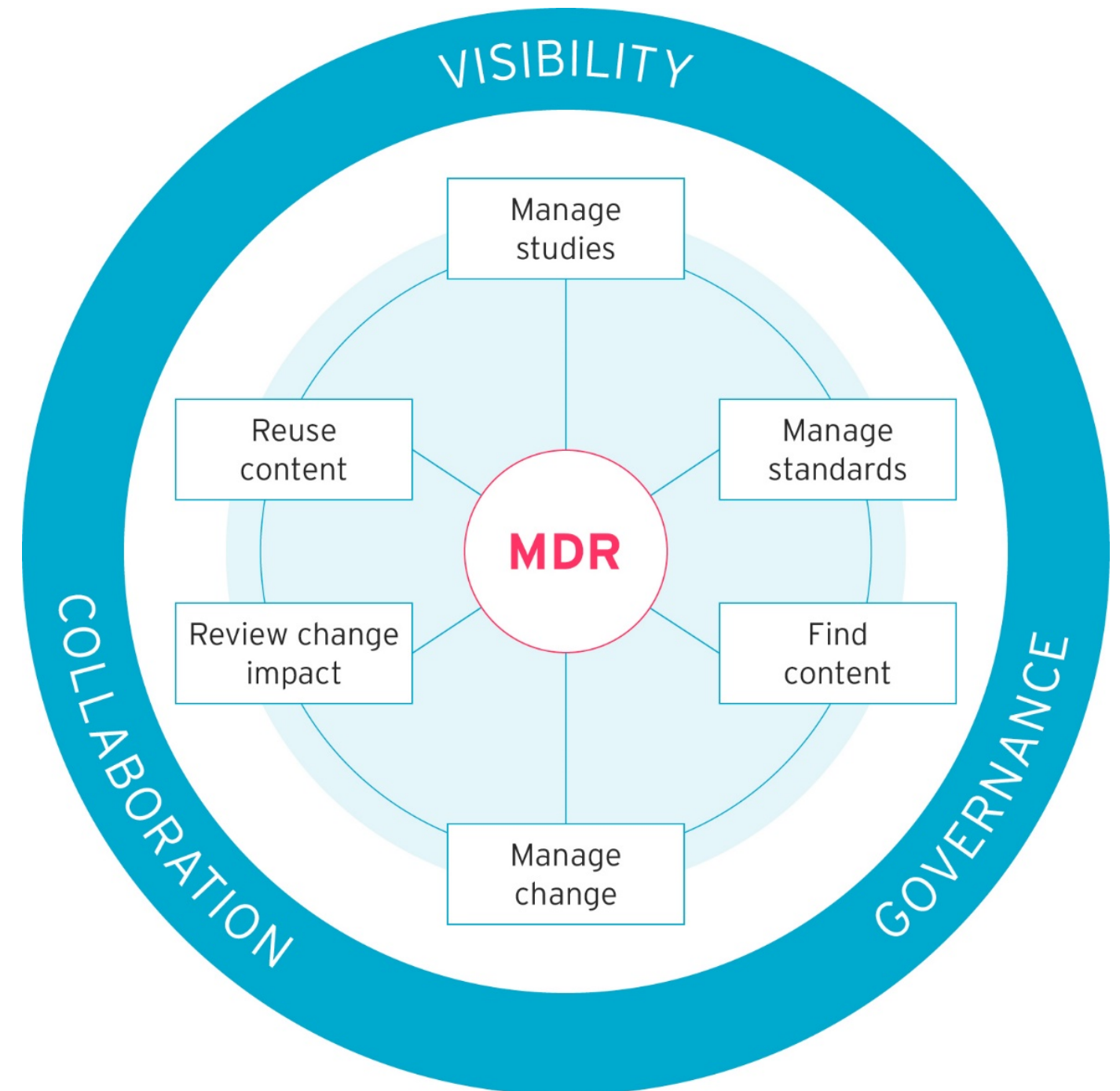
ID	Title	State	Due date
None			

+ Create change request

METADATA REPOSITORIES

End to end metadata management:

- Impact analysis.
- Change management.
- Lifecycles.
- Round tripping.
- Differencing.
- Flexibility with governance.



MANUAL PROCESSES REPLACED & NEW CAPABILITIES REALIZED

Study automation successes:

- Specifications.
- Multiple EDC design & build.
- Automated CRF annotations.
- SDTM dataset conversions.
- Validation.



EDC EXCEL SPECIFICATIONS AUTO-GENERATED

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Section															
2	OID	Name	Repeating	Description	Order No.	Mandatory	Aliases	Condition	IsReferenceData	Repeating Information	SASDatasetName	Domain	Origin	Comment	Purpose	
3	IG_DEMOGRAPHICS	Demographics	No		1	Yes	DM [SDTM\$Title]				DEMO	DM	CRF			
4	Questions															
5	OID	Name	Text	DataType	Order No.	Mandatory	Terminology	Length	Significant Digits	Units	Description	Aliases	External	Range Checks	SASFieldName	SDSVarName
6	DMDAT	Date of visit	Date of visit: Datum des Besuchs: [de] Date de la visite: [fr] 訪問の日: [jp]	date	1	Yes										
7	SUBJCTID	Subject Identifier	Enter Subject Identifier Geben Sie Subject Identifier [de] 件名識別子を入力します。 [en-US] Entrez l'identificateur Objet [fr]	text	2	Yes		20								
8	SUBJINIT	Initials	Subject Initials Subject Initials [de] Initiales Sujet [fr] 件名イニシャル [jp]	text	3	Yes		3				SUBJID [CDASH:Question] SCSTRESC, SCORRES when SCTESTCD=SU BJID [SDTM:Value]			subinit	SUBJINIT
9	CSNTSIG	Consent Signed	Informed Consent given? Informierte Einwilligung gegeben? [de] Donné son consentement éclairé? [fr] インフォームドコンセント を与えられましたか? [jp]	text	4	Yes	CL_YN	1				DSDECOD [CDASH:Question] DSDECOD [SDTM:Value]			cnsnt	DSDECODE
10	CSNTDAT	Consent Date	Date Of Informed Consent Datum der informierten Zustimmung [de] Date De Consentement Au courant [fr] インフォームドコンセント の日 [jp]	date	5	No						DSSTDAT [CDASH:Question] DSDTC [SDTM:Value]			cnsntdt	DSDTC
11	BRTHDTC	Birth Date	Date of Birth Geburtsdatum [de] Date de naissance [fr] 生年月日 [jp]	date	6	Yes						BRTHDAT [CDASH:Question] BRTHDTC [SDTM:Value]			birthdt	BRTHDTC
12	SEX	Sex	Sex Geschlecht [de] Sexe [fr] 性別 [jp]	text	7	Yes	CL_SEX	1				SEX [CDASH:Question] SEX [SDTM:Value]			sex	SEX
13	RACE	Race	Race Rasse [de] Race [fr] 種族 [jp]	text	8	Yes	CL_C74457_RACE	1				RACE [CDASH:Question] RACE [SDTM:Value]			race	RACE
			If Race is 'Other' please specify Wenn Rasse ist "Andere"									RACEOTH				

SEE HOW FORMS LOOK IN RAVE EDC - AS YOU DESIGN

MEDIDATA RAVE

	Demographics	Additional Metadata	Annotations	Show Edit Checks/Derivations
Subject: Subject Name Page: Demographics				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Date of Birth				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Informed Consent given?				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Subject Initials				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Date Of Informed Consent				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
If Race is 'Other' please specify				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Enter Subject Identifier				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Race				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Age				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Sex				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Date of visit:				☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

INFORM – DO FORMS, EDIT CHECKS, VISIT STRUCTURES & MORE

ORACLE INFORM

Demographics
[Show Annotations](#)

Demographics 

Demographics			
1.*	Date of Birth	/ /	 
2.*	Informed Consent given?	☐ No ☐ Yes	 
3.*	Subject Initials		 
4.*	Date Of Informed Consent	/ /	 
5.*	If Race is 'Other' please specify		 
6.*	Enter Subject Identifier		 
7.*	Race		 
8.*	Age		 
9.*	Sex	☐ Male ☐ Female	 
10.*	Date of visit:	/ /	 

ANNOTATED CRFS

	<table><tr><td></td><td></td><td></td></tr><tr><td colspan="3">Site Number</td></tr></table> <table><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td colspan="5">Subject Number</td></tr></table>				Site Number								Subject Number				
Site Number																	
Subject Number																	
Protocol DM																	

Demographics		F_DM_002								
1 Demographics		IG_DEMOGRAPHICS								
1.1	Date of Birth BIRTHDT (YY)	<table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
1.2	Informed Consent given? CSNTSIG	☐ No ☐ Yes CL_YN								
1.3	Subject Initials SUBJINIT									
1.4	Date Of Informed Consent CSNTDAT (YY)	<table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
1.5	If Race is 'Other' please RACESP									
1.6	Enter Subject Identifier SUBJCTID									
1.7	Race* RACE									
1.8	Age AGE	<table><tr><td></td><td></td><td></td></tr></table>								
1.9	Sex SEX	☐ Male ☐ Fema CL_SEX								
1.10	Date of visit: DMDAT (YYYY)	<table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
*Race AMERICAN INDIAN OR ALASKA NATIVE=American Indian or Alaska Native ASIAN=Asian OTHER=Other, specify: CL_C74457_RACE BLACK OR AFRICAN AMERICAN=Black or African American WHITE=White NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER=Native Hawaiian or Other Pacific Islander										

TARGET SDTM SPECIFICATION AUTO-GENERATED IN PDF

DM (Demographics)

DM (Demographics) -- SPECIAL PURPOSE								dm.xpt
Name	Description	Key	Type	Length	Controlled Terms	Origin	Role	Comment
STUDYID	Study Identifier	1	text	7		Protocol	Identifier	
DOMAIN	Domain Abbreviation	2	text	2	["DM"]	Assigned	Identifier	
USUBJID	Unique Subject Identifier	3	text	14		Derived	Identifier	See Method (M.USUBJID)
SUBJID	Subject Identifier for the Study		text	6		CRF Page [F_DM_002]	Topic	
RFSTDTC	Subject Reference Start Date/Time		datetime			Derived	Record Qualifier	See Method (M.RFSTDTC)
RFENDTC	Subject Reference End Date/Time		datetime			Derived	Record Qualifier	See Method (M.RFENDTC)
RFXSTDTC	Date/Time of First Study Treatment		datetime			CRF Page [F_DM_002]	Record Qualifier	
RFXENDTC	Date/Time of Last Study Treatment		datetime			CRF Page [F_DM_002]	Record Qualifier	
RFICDTC	Date/Time of Informed Consent		datetime			CRF Page [F_DM_002]	Record Qualifier	
RFPENDTC	Date/Time of End of Participation		datetime			CRF Page [F_DM_002]	Record Qualifier	

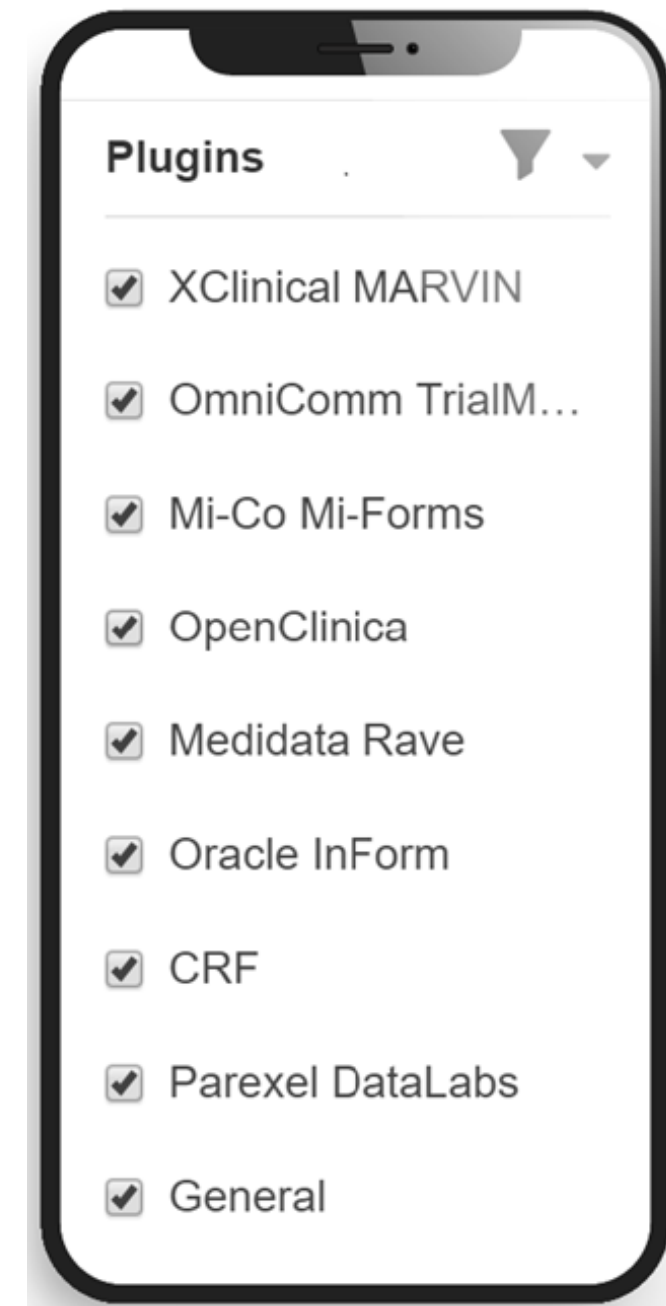
MAPPING SPECIFICATIONS AUTO-GENERATED

[illegible]

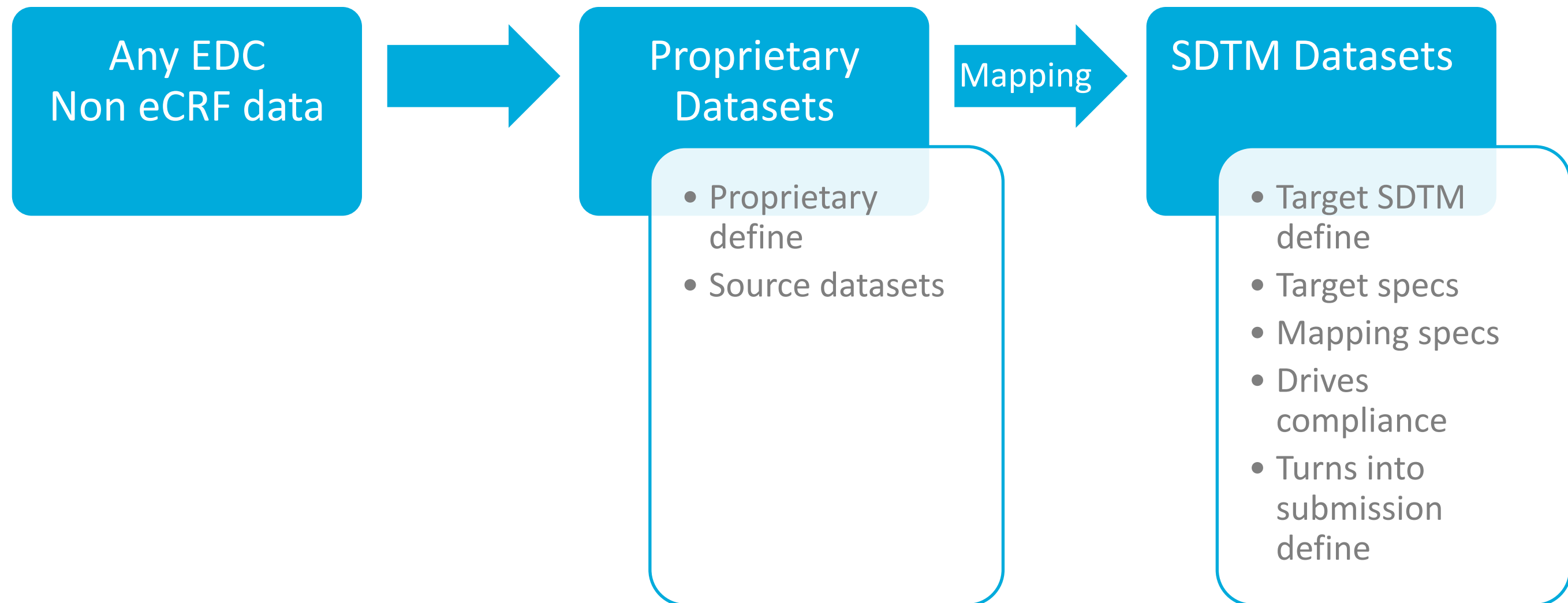
EDC DESIGN & BUILD STUDY AUTOMATION

Automated EDC build – at the push of a button!

- Design CRFs for 7 EDCs
- See what eCRFs look like *as you design forms – before EDC build.*
- Instant EDC build – no delays.
- Fewer author, review & approval cycles.
- 50-68% decrease in study set-up time.



METADATA DRIVEN DATA TRANSFORMATIONS



9 TYPICAL SDTM MAPPING SCENARIOS

Why do I need a mapping language?

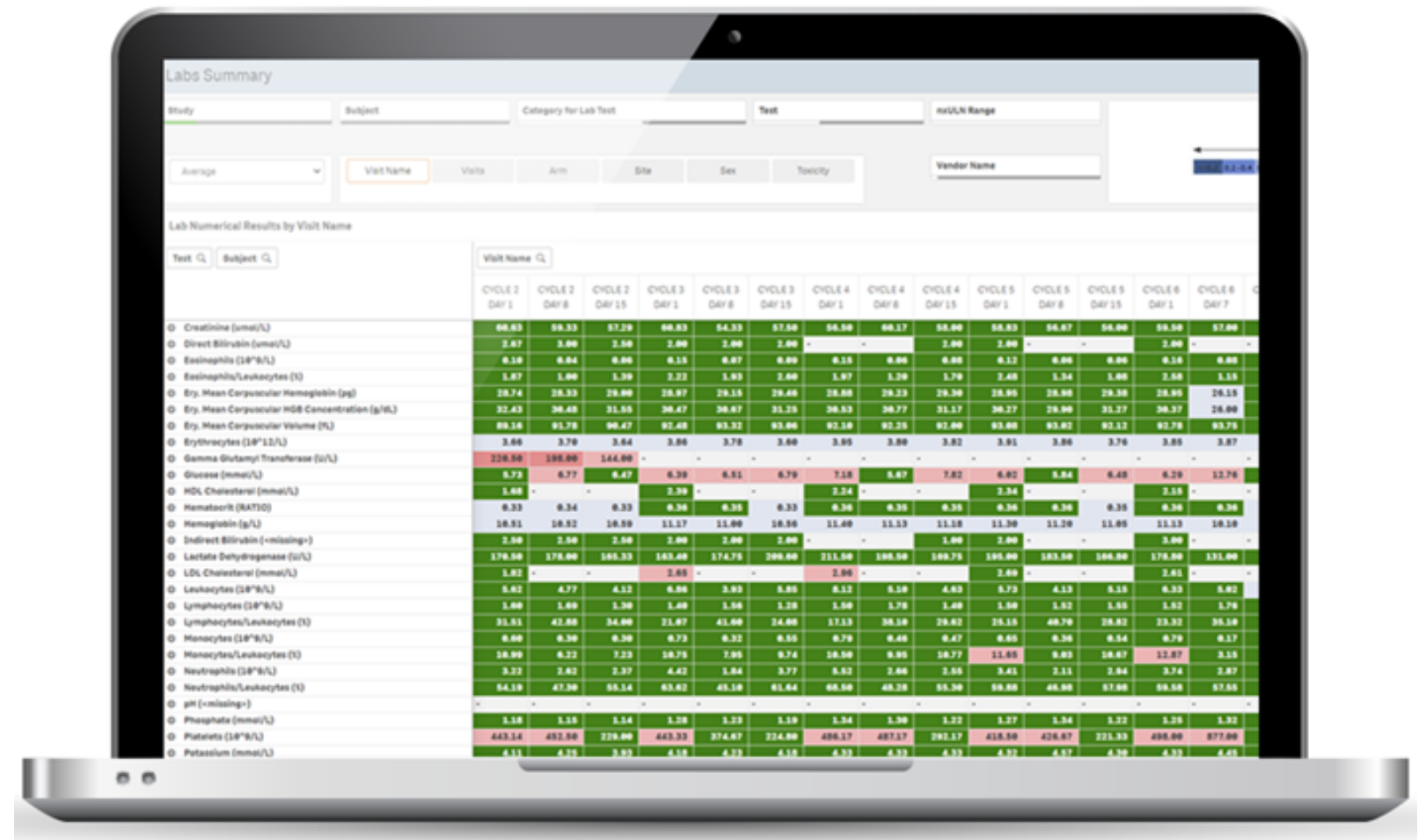
1. The direct carry forward.
2. The variable rename.
3. The variable attribute change.
4. The reformat.
5. The combine.
6. The split.
7. The derivation.
8. The variable value map & new code list application.
9. The horizontal-to-vertical data structure transpose.



SDTM AUTOMATION

It's all about time to data insights!

- Auto-generation of consistent mapping specs.
- Metadata driven automated conversions.
- Eliminates double programming!
- 70% reduction in dataset conversion time and effort.
- Reduces lag time – see your data quicker.



Labs Summary

Study: Subject: Category for Lab Test: Test: n/N Range:

Average: Visit Name: Visits: Arm: Site: Sex: Toxicity: Vendor Name:

Lab Numerical Results by Visit Name

Test	CYCLE 2 DAY 1	CYCLE 2 DAY 8	CYCLE 2 DAY 15	CYCLE 3 DAY 1	CYCLE 3 DAY 8	CYCLE 3 DAY 15	CYCLE 4 DAY 1	CYCLE 4 DAY 8	CYCLE 4 DAY 15	CYCLE 5 DAY 1	CYCLE 5 DAY 8	CYCLE 5 DAY 15	CYCLE 6 DAY 1	CYCLE 6 DAY 7
① Creatinine (umol/L)	66.83	66.33	67.29	66.83	64.33	67.58	66.88	66.17	68.88	68.83	66.67	66.88	69.58	67.88
② Direct Bilirubin (umol/L)	2.67	3.88	2.58	2.88	2.88	2.88	-	-	2.88	2.88	-	-	2.88	-
③ Eosinophils (10 ⁹ /L)	0.18	0.84	0.88	0.15	0.87	0.88	0.15	0.88	0.88	0.12	0.88	0.88	0.18	0.88
④ Eosinophils/Leukocytes (1)	1.87	1.88	1.38	1.22	1.83	2.88	1.87	1.28	1.78	2.48	1.54	1.88	2.58	1.55
⑤ Ery. Mean Corpuscular Hemoglobin (pg)	28.74	28.33	28.88	28.87	28.15	28.48	28.88	28.23	28.38	28.85	28.88	28.38	28.85	28.15
⑥ Ery. Mean Corpuscular HGB Concentration (g/dL)	32.43	36.48	31.55	36.47	36.67	31.15	36.53	36.77	31.17	36.17	28.88	31.17	36.17	28.88
⑦ Ery. Mean Corpuscular Volume (fL)	89.18	91.78	98.47	92.48	93.32	93.88	92.18	92.15	92.88	93.88	93.82	92.12	92.78	93.75
⑧ Erythrocytes (10 ¹² /L)	3.88	3.78	3.84	3.88	3.78	3.88	3.88	3.88	3.82	3.81	3.88	3.78	3.88	3.87
⑨ Gamma Glutamyl Transferase (u/L)	128.88	188.88	188.88	-	-	-	-	-	-	-	-	-	-	-
⑩ Glucose (mmol/L)	5.73	6.77	6.47	6.38	6.11	6.78	7.18	5.87	7.82	6.82	6.84	6.48	6.28	12.76
⑪ HDL Cholesterol (mmol/L)	1.88	-	-	2.38	-	-	2.34	-	-	2.34	-	-	2.15	-
⑫ Hematocrit (KAT10)	0.33	0.34	0.33	0.34	0.35	0.33	0.36	0.35	0.35	0.36	0.36	0.35	0.36	0.36
⑬ Hemoglobin (g/L)	18.81	18.82	18.88	11.17	11.88	18.88	11.48	11.13	11.18	11.38	11.38	11.88	11.13	18.18
⑭ Indirect Bilirubin (-missing-)	2.58	2.58	2.58	2.88	2.88	2.88	-	-	1.88	2.88	-	-	1.88	-
⑮ Lactate Dehydrogenase (u/L)	178.88	178.88	188.33	183.48	174.75	288.88	211.88	188.88	188.75	188.88	183.88	188.88	178.88	131.88
⑯ LDL Cholesterol (mmol/L)	1.82	-	-	2.88	-	-	2.88	-	-	2.88	-	-	2.81	-
⑰ Leukocytes (10 ⁹ /L)	5.82	4.77	4.12	6.88	3.83	5.88	8.12	5.18	4.83	5.73	4.13	5.18	6.33	5.82
⑱ Lymphocytes (10 ⁹ /L)	1.88	1.88	1.38	1.48	1.58	1.28	1.88	1.78	1.48	1.82	1.55	1.82	1.82	1.78
⑲ Lymphocytes/Leukocytes (1)	31.51	42.88	34.88	21.87	41.48	24.88	17.13	38.18	28.82	25.15	48.78	28.82	23.32	35.18
⑳ Monocytes (10 ⁹ /L)	0.88	0.38	0.38	0.73	0.12	0.15	0.78	0.48	0.47	0.45	0.18	0.54	0.78	0.17
㉑ Monocytes/Leukocytes (1)	18.88	6.22	7.23	18.78	7.88	0.74	18.88	0.88	18.77	11.88	0.83	18.87	12.87	3.15
㉒ Neutrophils (10 ⁹ /L)	3.22	2.42	2.37	4.42	1.84	3.77	5.82	2.88	2.15	3.41	2.11	2.84	3.74	2.87
㉓ Neutrophils/Leukocytes (1)	84.18	47.38	58.14	83.82	45.18	81.84	88.88	48.18	55.38	58.88	48.88	57.88	58.88	87.88
㉔ pH (-missing-)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
㉕ Phosphate (mmol/L)	1.18	1.18	1.14	1.18	1.13	1.18	1.14	1.18	1.22	1.27	1.14	1.22	1.15	1.32
㉖ Platelets (10 ⁹ /L)	443.14	482.88	228.88	443.33	374.47	224.88	488.17	487.17	287.17	418.88	428.87	221.33	488.88	877.88
㉗ Potassium (mmol/L)	4.11	4.18	3.93	4.18	4.13	4.18	4.13	4.13	4.13	4.13	4.17	4.18	4.13	4.18

BUSINESS BENEFITS OF A METADATA REPOSITORY

ROI from reuse and automation

- 70% reuse of eCRFs, edit checks, proprietary datasets, SDTM, mappings, ADaM....
- 6th trial onwards you're at break even.
- Benefits in submission:
 - Cost avoidance of US\$1.5million.
 - Savings in conversions, quality checks, data analysis & pooling.

Source Innovion



BUSINESS BENEFITS OF STUDY AUTOMATION

Deliver proven ROI – time and cost savings.

- ~~1,484~~ 476 EDC labor hours.
- ~~20 weeks~~ 4 weeks to EDC go-live.
- ~~486~~ 140 SDTM labor hours.
- ~~8 weeks~~ 4 weeks to see SDTM data.





Data Integrity & Standards in the Clinical Data Pipeline: **Building a Robust Clinical Data Pipeline**

Raj Indupuri
CEO, eClinical Solutions

PhUSE EU Connect
November 2020



The imperative to modernize and manage the clinical data deluge

- "External" or non-EDC data sources make up **60% of trial data** according to a recent Tufts-eClinical Solutions study
- Last Patient Last Visit (LPLV) to database lock cycle times increasing, especially for sponsors managing **5-6 data sources**
- On an average, in 2019-2020 **six data sources per trial** were integrated and leveraged for analytics
- Biomarkers are increasingly common data sources from specialty labs, often with varying formats from different vendors



Wearables



EHR



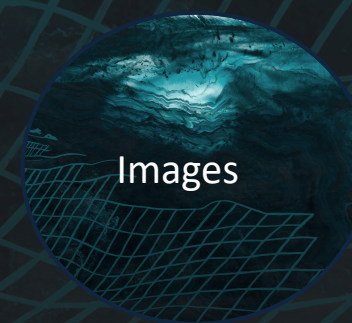
Safety



ePRO



CTMS



Images



Biomarkers



Labs

Tufts-eClinical Solutions data strategies and transformation study

A transformational shift is required to modernize technology infrastructure



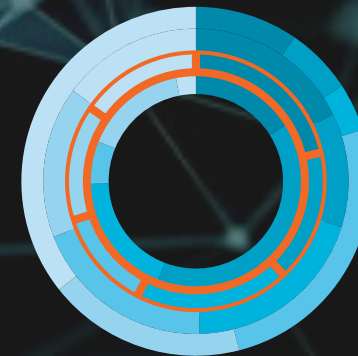
Increased cycle time LPLV to DBL with >5 data sources – no data strategy



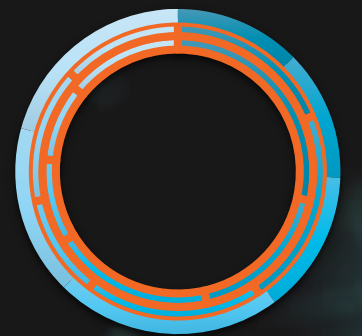
Labor-intensive or time-consuming to manage the entire data lifecycle



Labor intensive to initiate relationship with data providers and curate data



Dependent on more traditional methods like SAS to integrate and organize data



Dependent on SAS and Excel to consume, review, and analyze data

Building a reliable, robust and automated clinical data pipeline to reduce cycle times



Data Ingestion Automation

Automating data ingestion from a myriad of sources using event-based framework and end-to-end automation

Persisting and storing various formats of data for efficient processing



No-code Transformations

Streamlining integration and transformation workflows to better manage data assets

Tools to combine, curate and standardize data with built-in governance



Automated Insights and Actions

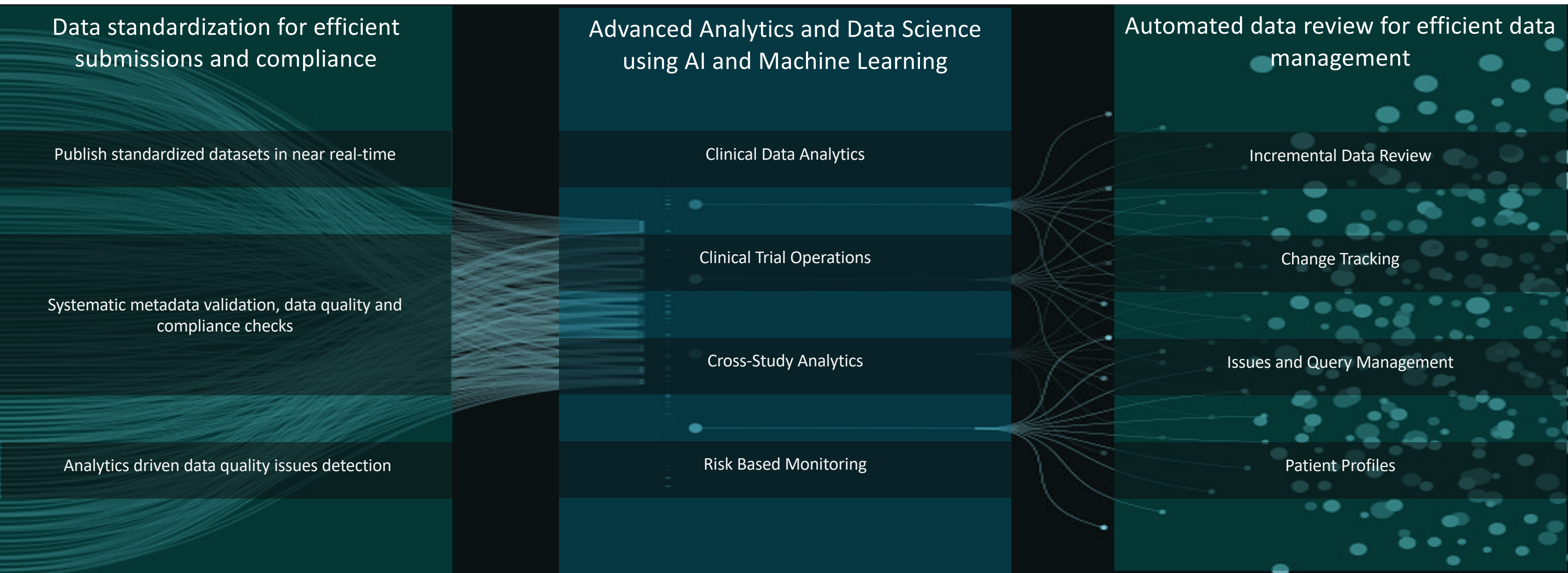
Publish data into more analysis-ready formats for downstream analytics

Analytics-ready data for quick and easy consumption across variety of use cases and external systems

Scaling clinical data pipeline automation for faster standardization and governed access to data for analytics



Automating a clinical data pipeline enables data-driven actions and decisions for better outcomes



Key Takeaways: The Industry is rapidly evolving to modernize clinical trials and adopt new technologies and with it....

Novel approaches to modernize clinical trials and data management

Leverage of technology to manage the increasing volumes and varieties of data

Invest into interoperable ecosystem of tools, technology and clinical systems

1

Build a robust data and technology strategy with defined business outcomes for pervasive transformation and automation

2

Taking a more end-to-end approach to build fully automated data pipelines leveraging a Clinical Data Platform or Data Hub rather than point solutions

3

Ensuring interoperability between clinical systems and flexibility to fit with existing infrastructure to manage shifts in industry trends



FIND OUT HOW FORMEDIX CAN HELP:

Find out more...

<https://www.formedix.com/>

See it in action, book a demo:

<https://www.formedix.com/book-a-demo/>

Contact Mark with any questions or to
request white papers:

Mark Wheeldon | CEO | Formedix
markwheeldon@formedix.com



LEARN MORE ABOUT eCLINICAL SOLUTIONS:

Find out more...

eclinicalsol.com/

Contact Us:

eclinicalsol.com/contact-us/

Contact Raj with any questions or to
request more information:

Raj Indupuri | CEO | eClinical Solutions
rindupuri@eclinicalsol.com