



Mind the Gap: The Raw Data Model

Julie Smiley, Sr. Director, Life Sciences Product Strategy

Oracle Health Sciences Global Business Unit

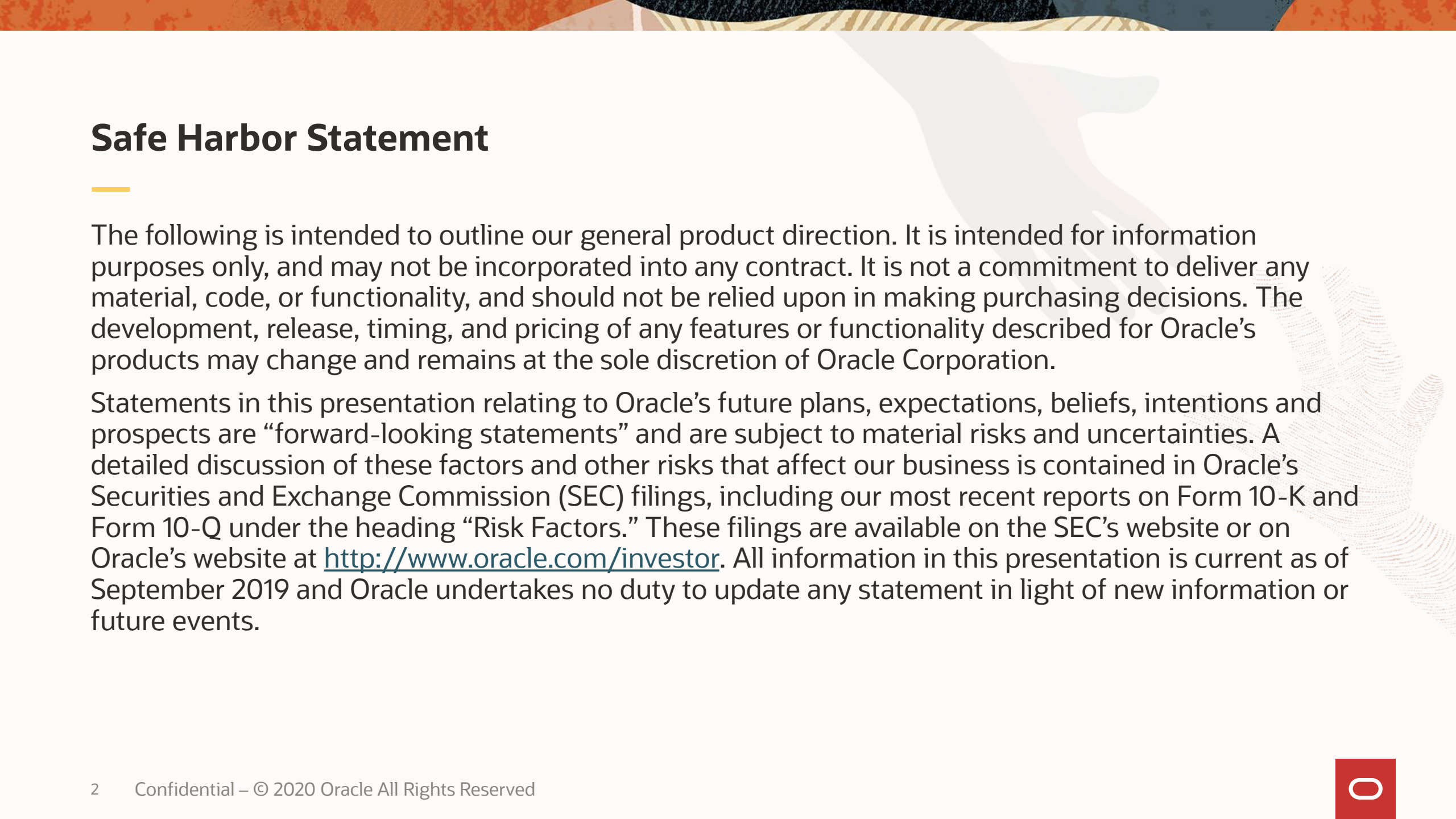
julie.s.smiley@oracle.com

PHUSE US Connect 2020 – SI02

March 9, 2020



Safe Harbor Statement



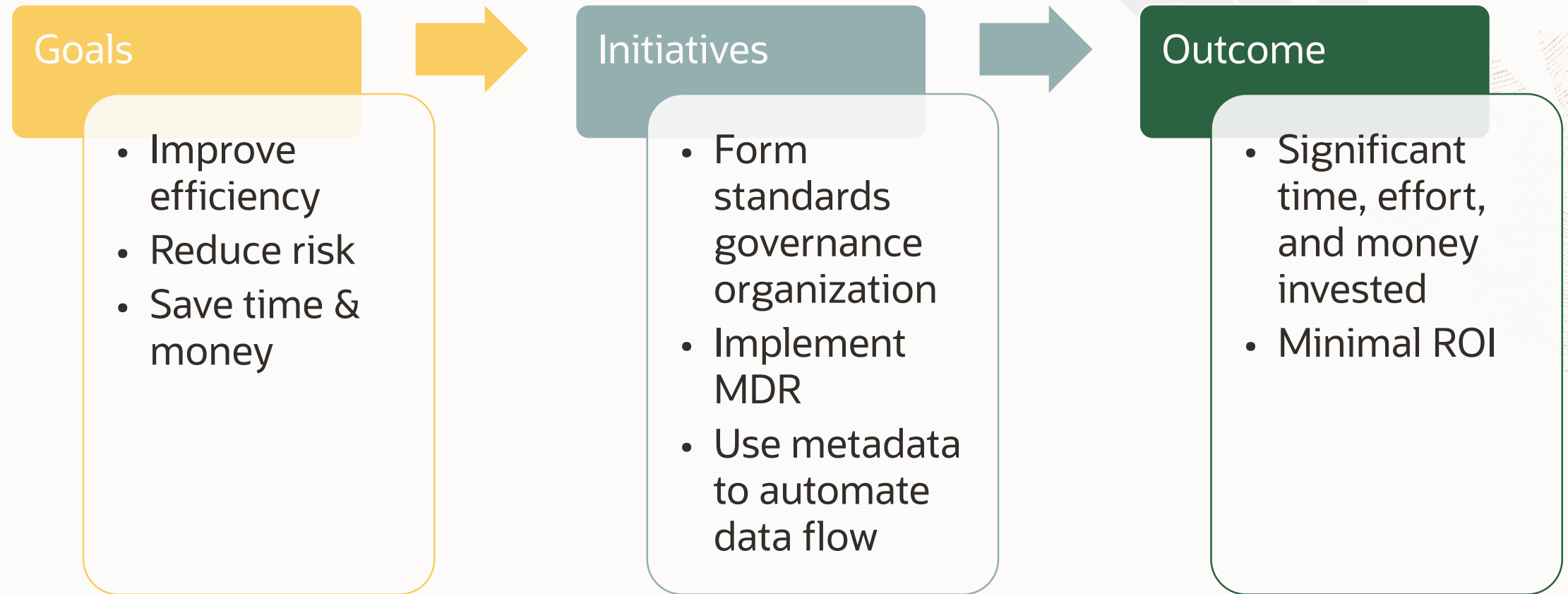
The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

Statements in this presentation relating to Oracle's future plans, expectations, beliefs, intentions and prospects are "forward-looking statements" and are subject to material risks and uncertainties. A detailed discussion of these factors and other risks that affect our business is contained in Oracle's Securities and Exchange Commission (SEC) filings, including our most recent reports on Form 10-K and Form 10-Q under the heading "Risk Factors." These filings are available on the SEC's website or on Oracle's website at <http://www.oracle.com/investor>. All information in this presentation is current as of September 2019 and Oracle undertakes no duty to update any statement in light of new information or future events.

Agenda

- 1 Defining the Gap
- 2 Minding the Gap
- 3 Conclusion

Clinical Development Initiatives



Key Barriers



People

- Internal, partners, and vendors
- Training/skillsets
- Roles & responsibilities



Processes

- Change management
- Functional silos
- Mergers & Acquisitions



Technology

- No “one size fits all” solution
- Budget constraints
- **Standards gap**

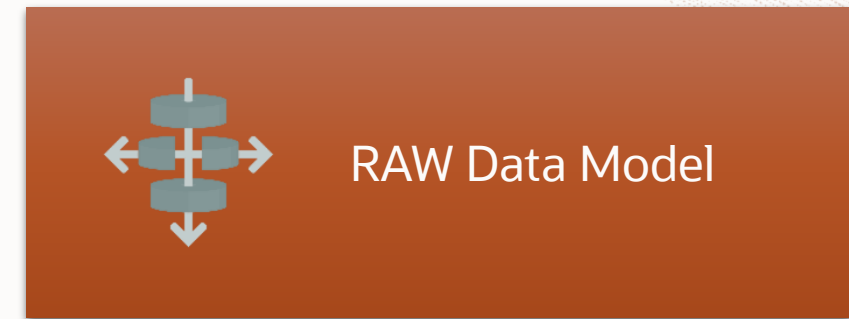
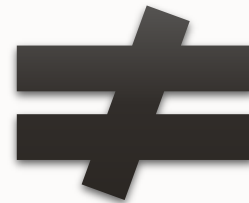


Standards Gap

- Data collection (DC) standards/specifications model how data is to be collected
- The DC model does not define how data is to be extracted/output
- DC solutions typically have proprietary physical data model for storing data with no user access
- DC system extracts (raw data model) are similar, but not the same as the DC model
- Each DC system has a set of rules that determine how the data comes out in the extract



Optimized for user entry and quality data input



Optimized for downstream usage and quality data output

Example 1: Repeating & Non-Repeating Sections

Lab Form Example:

Sample Collection Date:

Sample Collection Time:

Test	Result	Units
Hemoglobin		
Hematocrit		
WBC		
Neutrophils		

Non-Repeating

Repeating

Lab Raw Data Extract Example:

LB_NR

SUBJID	VISIT	LBDAT	LBTIM

LB

SUBJID	VISIT	LBTEST	LBRES	LBUNIT



Example 2: Multiple Form Variants

Vital Signs Form Example:

Test	Result	Units
Height		
Weight		
Temperature		
...		

} Height collected

Test	Result	Units
Height		
Weight		
Temperature		
...		

} Height not collected

Vital Signs Raw Data Extract Example:

VS1

SUBJID	VISIT	VSTEST	VSORRES	VSORRESU

VS2

SUBJID	VISIT	VSTEST	VSORRES	VSORRESU



Example 3: Repeating & Non-Repeating Sections Across Form Variants

Collection Date:

Test	Result	Units
Height		
Weight		
Temperature		
...		

Non-Repeating

Repeating, height collected

Collection Date:

Test	Result	Units
Weight		
Temperature		
...		
...		

Non-Repeating

Repeating, height not collected

VS1_NR

SUBJID	VISIT	VSDAT

VS1

SUBJID	VISIT	VSTEST	VSORRES	VSORRESU

VS2_NR

SUBJID	VISIT	VSDAT

VS2

SUBJID	VISIT	VSTEST	VSORRES	VSORRESU



Example 4: Multiple Variables Per Item

Demographics Form Example:

Birth Date:

What is the sex of the subject:

☐ Male

☐ Female

☐ Unknown

☐ Undifferentiated

What is the subjects ethnicity:

☐ Hispanic or Latino

☐ Not Hispanic or Latino

☐ Not Reported

What is the subject's ...:

Demographics Raw Data Extract Example:

DM

SUBJID	VISIT	BRTHTDC	SEX	SEX_C	ETHNIC	ETHNIC_C

Code

Value

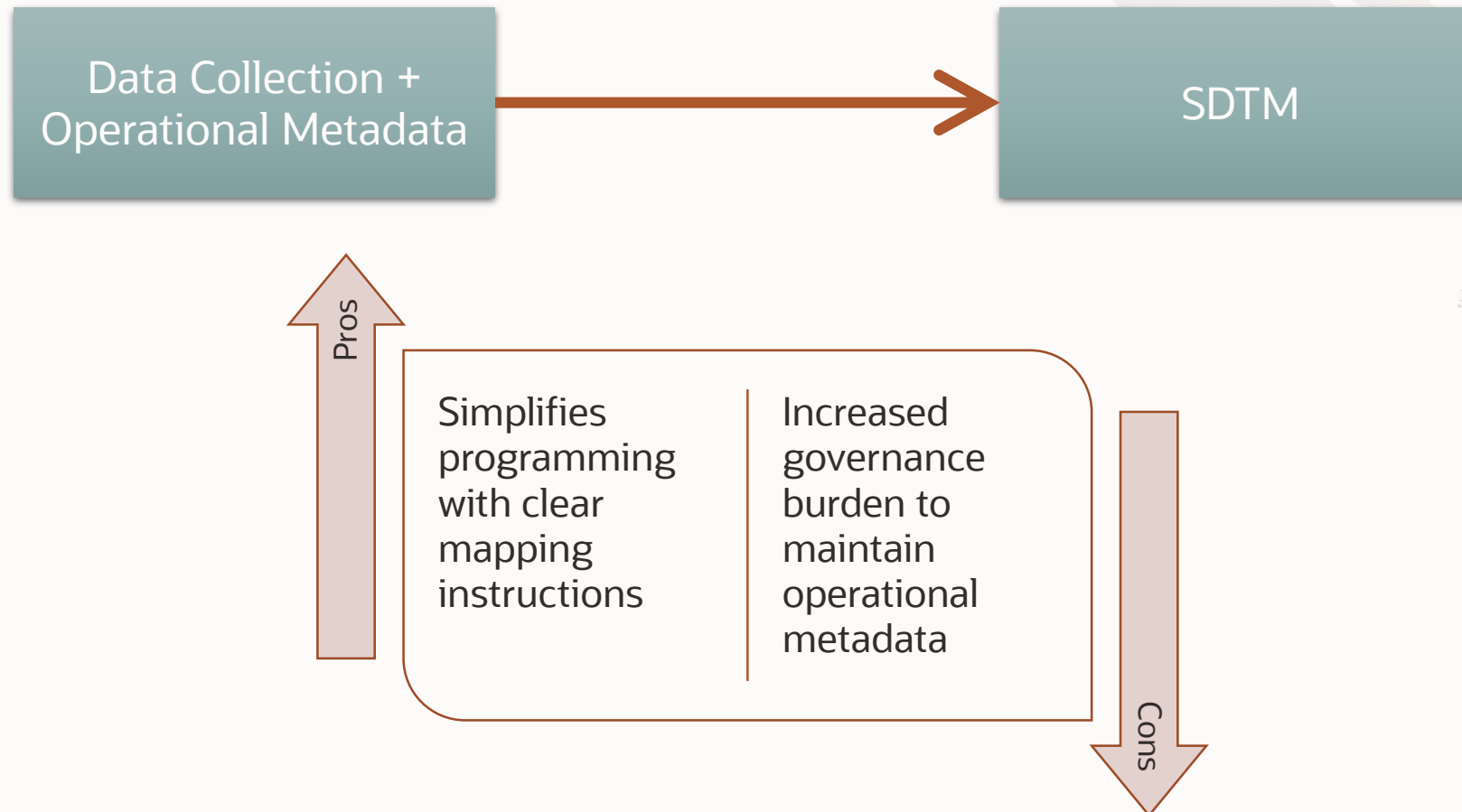


Minding the Gap



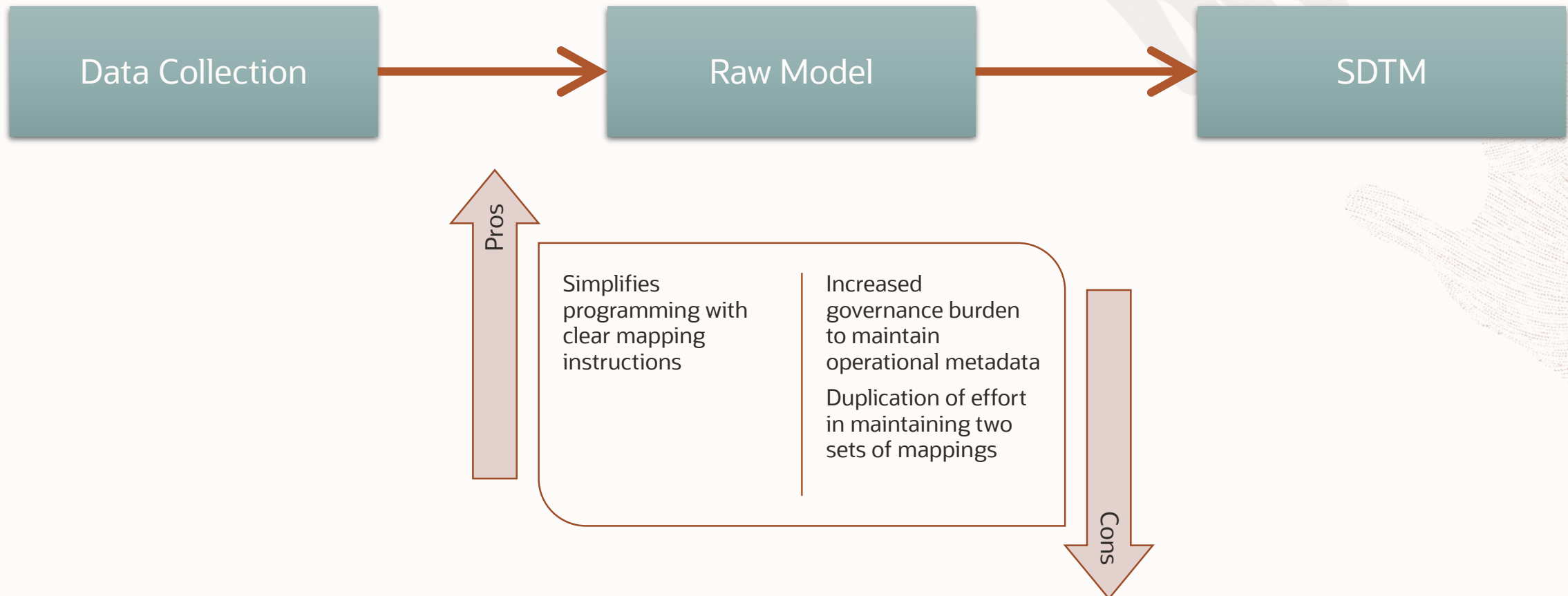
Method 1

Manage Operational Metadata as Part of Data Collection Standard



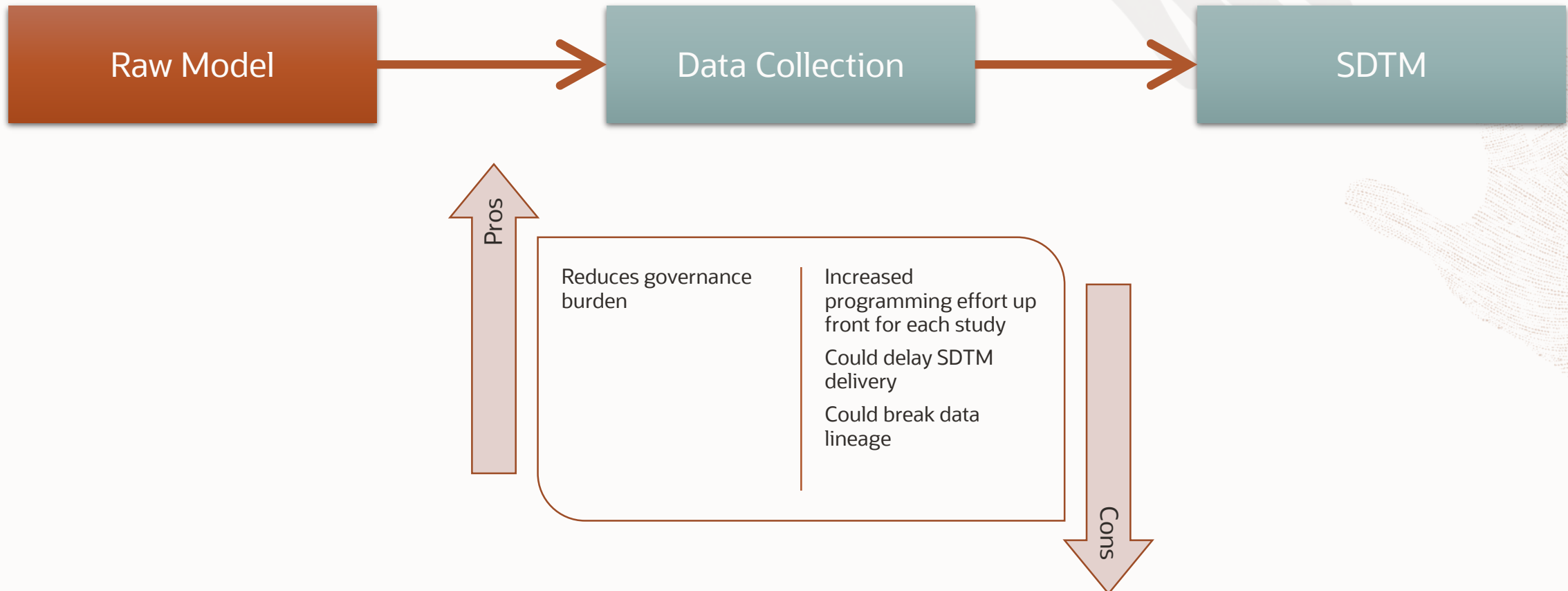
Method 2

Manage Raw Data Model as a Standard



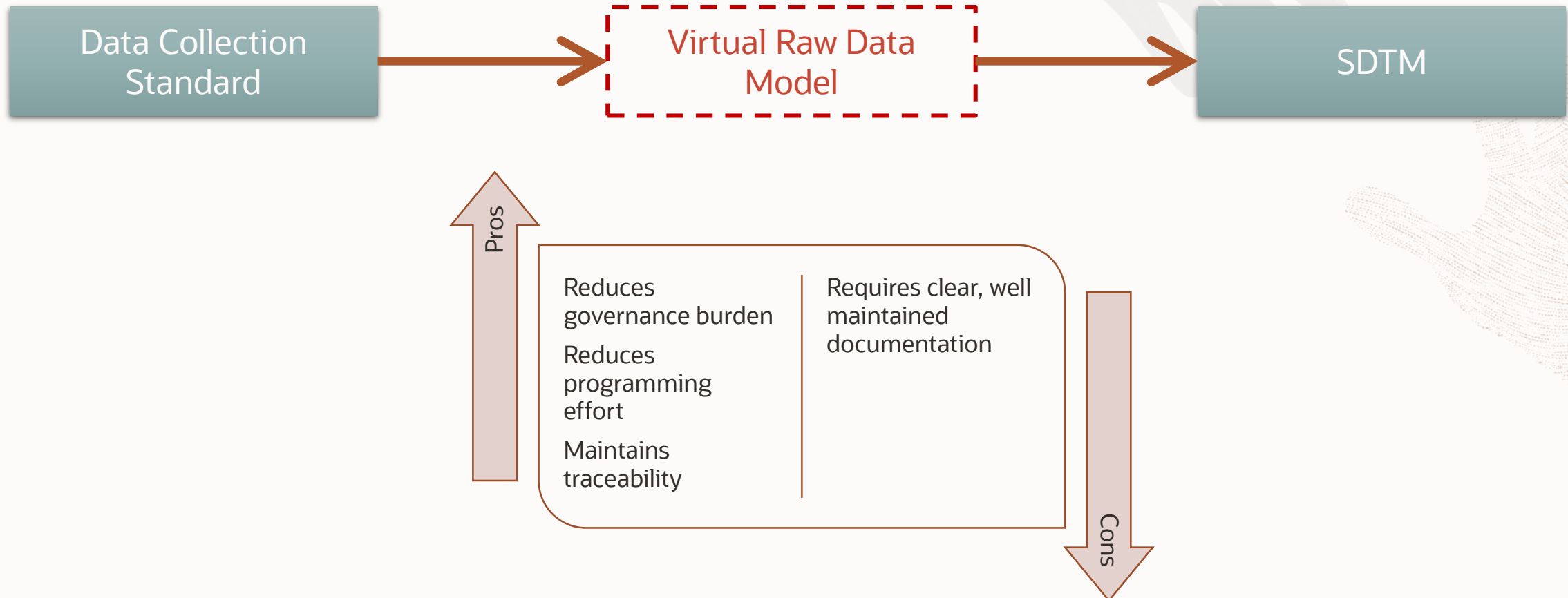
Method 3

Programmatically Convert Raw to Data Collection Standard

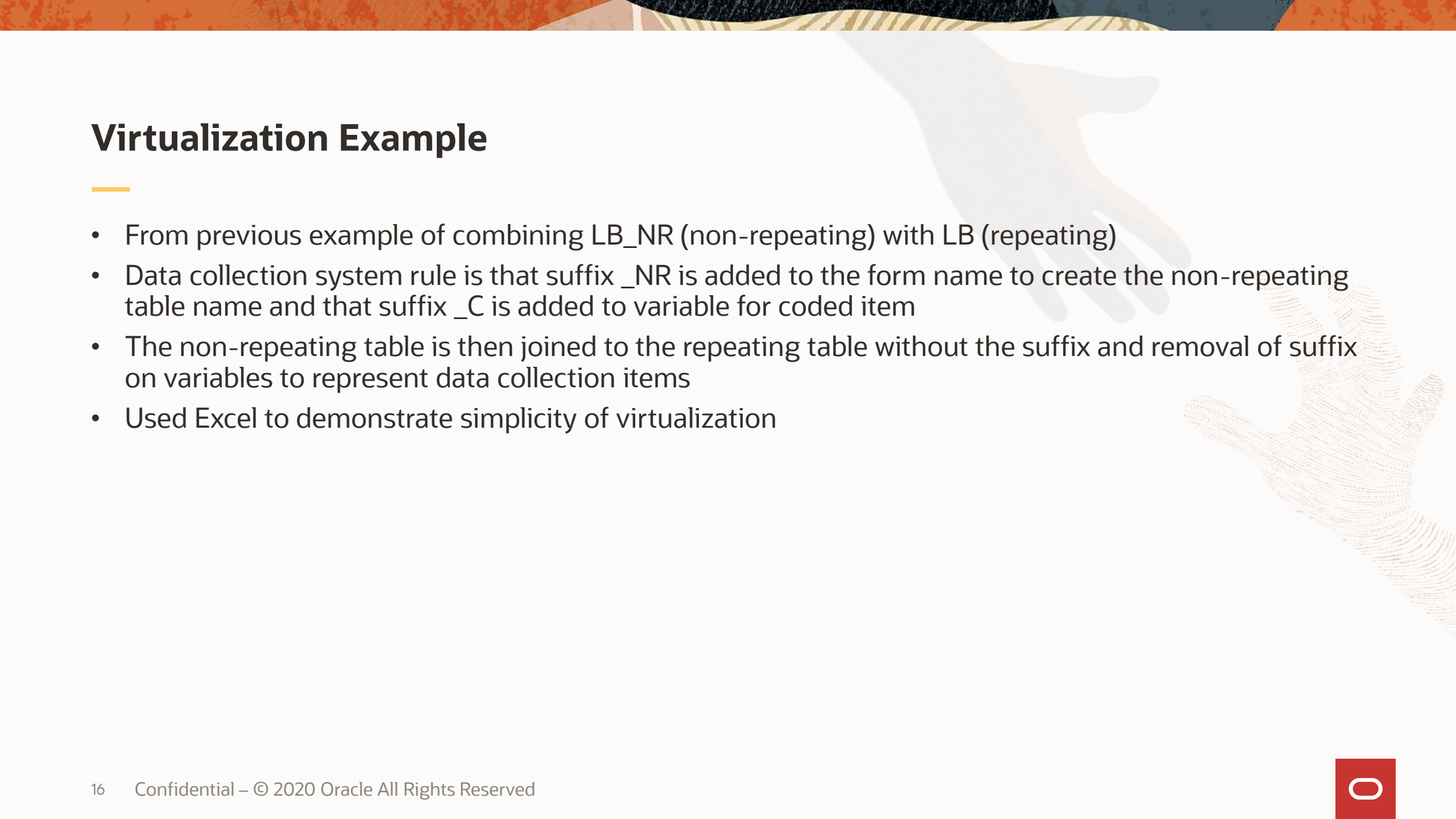


Method 4

Virtually Map to SDTM from Raw Data Model



Virtualization Example

The background of the slide features a decorative design. At the top, there is a horizontal band with an orange and brown pattern. Below this, a large, faint, light-gray hand is shown with fingers spread, reaching towards the right. In the lower right area, there is a fingerprint graphic composed of many small, overlapping text characters in a light gray color.

- From previous example of combining LB_NR (non-repeating) with LB (repeating)
- Data collection system rule is that suffix _NR is added to the form name to create the non-repeating table name and that suffix _C is added to variable for coded item
- The non-repeating table is then joined to the repeating table without the suffix and removal of suffix on variables to represent data collection items
- Used Excel to demonstrate simplicity of virtualization

Virtualization Example: Step 1

From Lab Data Collection Specifications, added in a new column and concatenated the form and item name, then sorted by this column.

Lab Data Collection Form Specifications:

FORM.ITEM	FORM	ITEM	DATA TYPE	CODELIST	SDTM Mapping	SDTM Mapping Instructions
LB.LBDAT	LB	LBDAT	CHAR		LBDTC	Concatenate LBDAT and LBTIM to an ISO 8601 format
LB.LBRES	LB	LBRES	CHAR		LBRES	
LB.LBTEST	LB	LBTEST	CHAR	TESTCD	LBTEST	
LB.LBTIM	LB	LBTIM	CHAR		LBDTC	Concatenate LBDAT and LBTIM to an ISO 8601 format
LB.LBUNIT	LB	LBUNIT	CHAR	UNIT	LBUNIT	
LB.SUBJID	LB	SUBJID	CHAR		DM.SUBJID	
LB.VISIT	LB	VISIT	CHAR		VISIT	

=[@FORM] & "." & [@ITEM]



Virtualization Example: Step 2

From the raw data extract metadata (data dictionary), programmatically generated the DC Form.Item Name

New Virtual Raw Model Mapping Specifications to SDTM:

TABLE	VARIABLE	DATA TYPE	DC Form.Item Name	SDTM Mapping	SDTM Mapping Instructions	
LB_NR	SUBJID	CHAR	LB.SUBJID	DM.SUBJID		0
LB_NR	VISIT	CHAR	LB.VISIT	VISIT		0
LB_NR	LBDAT	CHAR	LB.LBDAT	LBDTC	Concatenate LBDAT and LBTIM to an ISO 8601 format	
LB_NR	LBTIM	CHAR	LB.LBTIM	LBDTC	Concatenate LBDAT and LBTIM to an ISO 8601 format	
LB	SUBJID	CHAR	LB.SUBJID	DM.SUBJID		0
LB	VISIT	CHAR	LB.VISIT	VISIT		0
LB	LBTEST	CHAR	LB.LBTEST	LBTEST		0
LB	LBTEST_C	CHAR	LB.LBTEST	LBTEST		
LB	LBRES	CHAR	LB.LBRES	LBRES		
LB	LBUNIT	CHAR	LB.LBUNIT	LBUNIT		
LB	LBUNIT_C	CHAR	LB.LBUNIT	LBUNIT		

=IFERROR(LEFT([@TABLE],FIND("-",[@TABLE])-1),[@TABLE]) & "." & IFERROR(LEFT([@VARIABLE],FIND("-",[@VARIABLE])-1),[@VARIABLE])

Data Dictionary Output



Virtualization Example: Step 3

Used DC Form.Item Name to programmatically look up and add the SDTM Mappings and Mapping Instructions from the Lab Data Collection Specifications table

New Virtual Raw Model Mapping Specifications to SDTM:

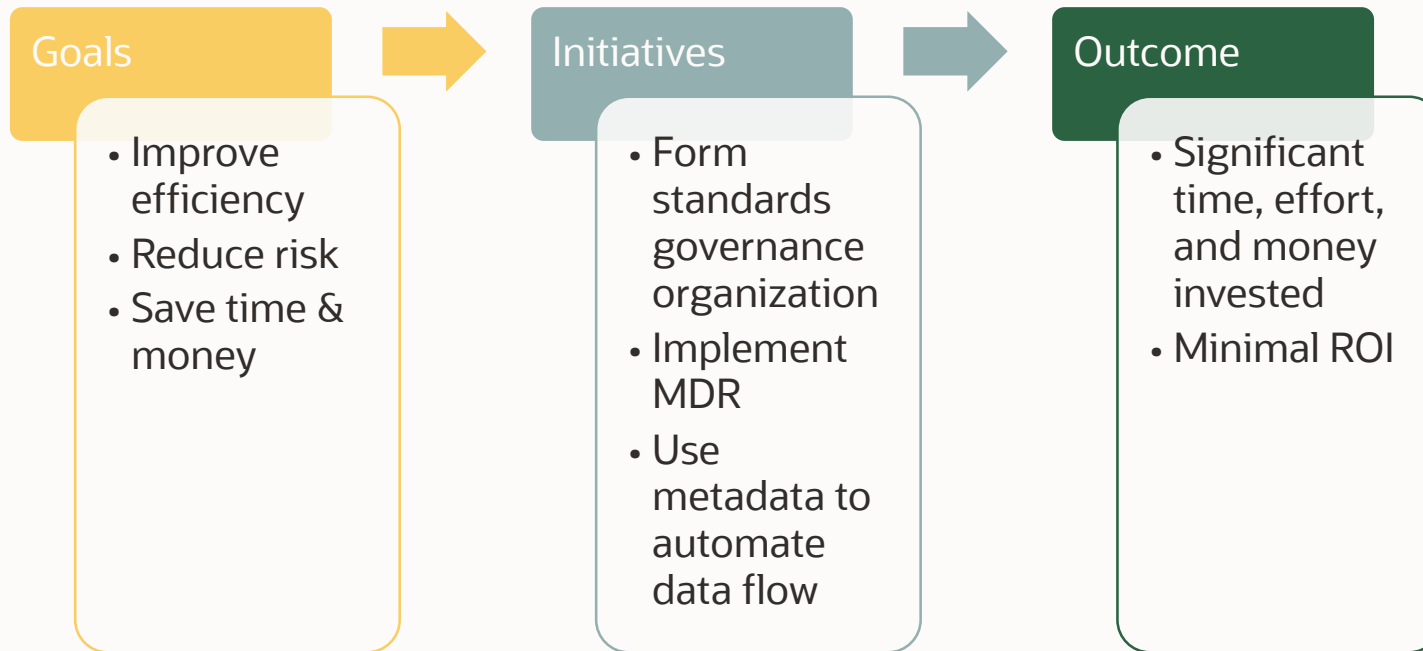
TABLE	VARIABLE	DATA TYPE	DC Form.Item Name	SDTM Mapping	SDTM Mapping Instructions	
LB_NR	SUBJID	CHAR	LB.SUBJID	DM.SUBJID		0
LB_NR	VISIT	CHAR	LB.VISIT	VISIT		0
LB_NR	LBDAT	CHAR	LB.LBDAT	LBDTC	Concatenate LB.LDAT and LBTIM to an ISO 8601 format	
LB_NR	LBTIM	CHAR	LB.LBTIM	LBDTC	Concatenate LBDAT and LBTIM to an ISO 8601 format	
LB	SUBJID	CHAR	LB.SUBJID	DM.SUBJID		0
LB	VISIT	CHAR	LB.VISIT	VISIT		0
LB	LBTEST	CHAR	LB.LBTEST	LBTEST		
LB	LBTEST	CHAR	LB.LBTEST	LBTEST		
LB	LBRES	CHAR	LB.LBRES	LBRES		
LB	LBUNIT	CHAR	LB.LBUNIT	LBUNIT		
LB	LBUNIT_C	CHAR	LB.LBUNIT	LBUNIT		0

=IFERROR(VLOOKUP([@[DC Form.Item Name]],Table1[#All],6,FALSE),""))

=IFERROR(VLOOKUP([@[DC Form.Item Name]],Table1[#All],7,FALSE),""))

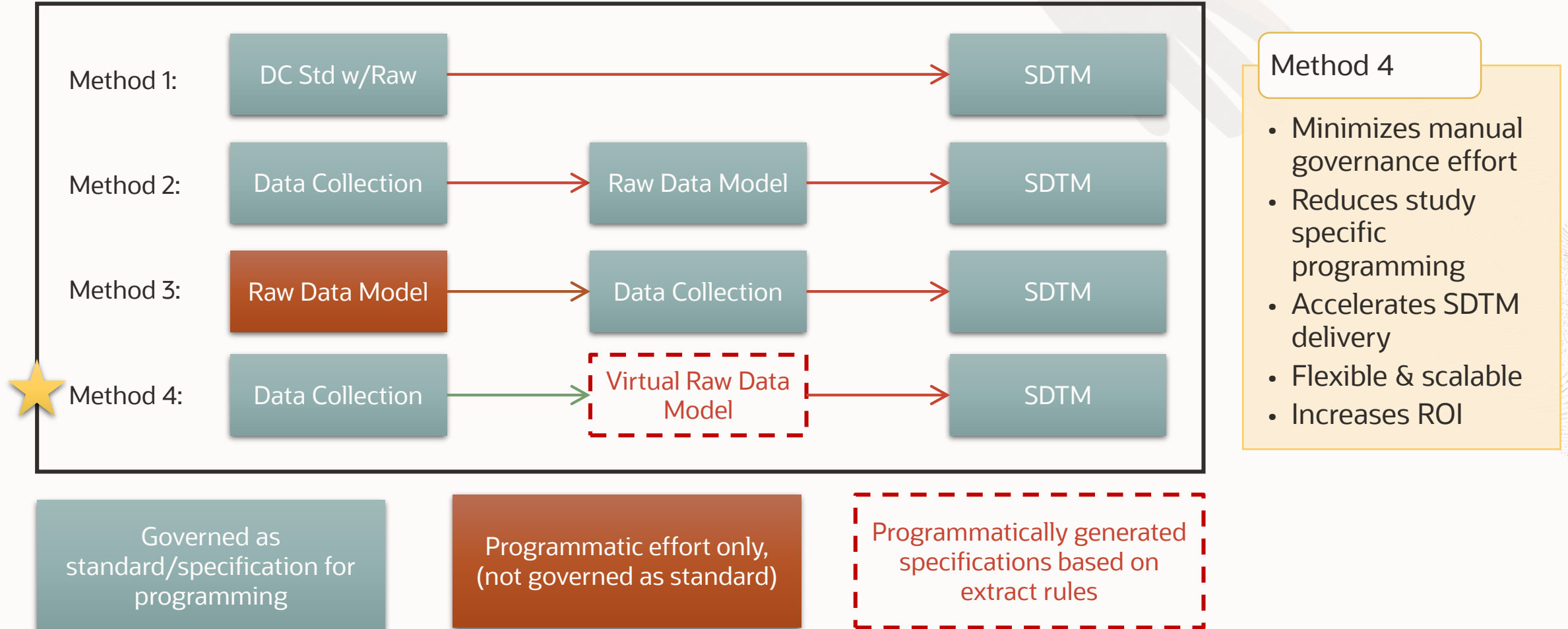


Summary



**Raw Data
Model Gap**

Addressing the Gap



An illustration of three hands reaching upwards towards a white rectangular area. The hands are stylized with different colors and patterns: a light blue hand with fine lines, a dark blue hand with orange and yellow patterns, and a pink hand with dark blue patterns. The background is a dark blue textured surface with large, abstract shapes in orange, black, and yellow.

Questions?



ORACLE