


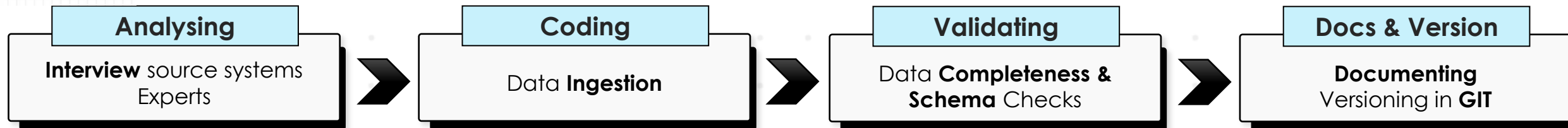
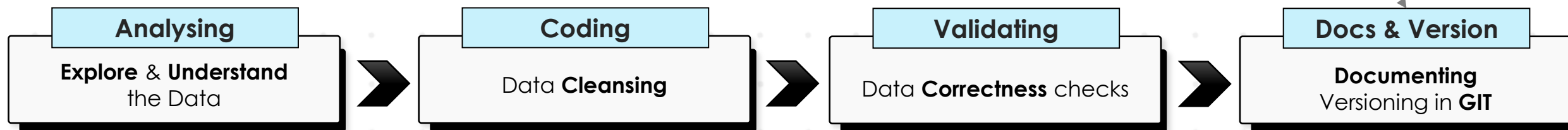
 ASHESH SAMRAJYA	 <div>Bronze Layer</div>	 <div>Silver Layer</div>	 <div>Gold Layer</div>
Definition	Raw, unprocessed data as-is from sources	Clean & standardized data	Business-Ready data
Objective	Traceability & Debugging	(Intermediate Layer) Prepare Data for Analysis	Provide data to be consumed for reporting & Analytics
Object Type	Tables	Tables	Views
Load Method	Full Load (Truncate & Insert)	Full Load (Truncate & Insert)	None
Data Transformation	None (as-is)	<ul style="list-style-type: none"> - Data Cleaning - Data Standardization - Data Normalization - Derived Columns - Data Enrichment 	<ul style="list-style-type: none"> - Data Integration - Data Aggregation - Business Logic & Rules
Data Modeling	None (as-is)	None (as-is)	<ul style="list-style-type: none"> - Star Schema - Aggregated Objects - Flat Tables
Target Audience	- Data Engineers	<ul style="list-style-type: none"> - Data Analysts - Data Engineers 	<ul style="list-style-type: none"> - Data Analysts - Business Users



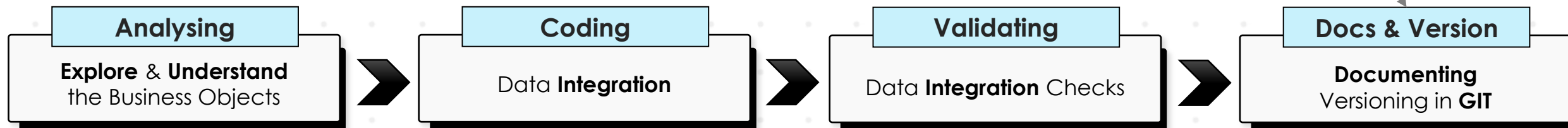
Bronze Layer



Silver Layer



Gold Layer





Source System Interview

Business Context & Ownership

Who **owns** the **data**?

What **Business Process** it supports?

System & Data **documentation**

Data **Model** & Data **Catalog**

Architecture & Technology Stack

How is data **stored**?
(SQL Server, Oracle, AWS, Azure, ...)

What are the **integration capabilities**?
(API, Kafka, File Extract, Direct DB, ...)

Extract & Load

Incremental vs. **Full Loads** ?

Data **Scope** & **Historical** Needs

What is the expected **size** of the **extracts**?

Are there any data **volume limitations**?

How to **avoid impacting** the source
system's **performance**?

authentication and **authorization**
(tokens, SSH keys, VPN, IP whitelisting, ...)