PROJECT IMPLEMENTATION

Questions to ask to client:

1. What is the Source of the Data? (API, SharePoint, Databases, third-party providers, Files, DropZone, etc.…)
2. How often should data be retrieved or synchronized?
3. Are there any privacy or compliance concerns we need to address?
4. What is the expected volume of data?
5. How should the data be integrated with existing systems or workflows?
6. Who is the point of contact for troubleshooting or further questions?
7. What are the SharePoint site URL and access credentials?
8. Are there any specific file naming conventions or metadata to be aware of?
9. Is there available API documentation or an API key for accessing it?
10. What are the API endpoint URLs, and what authentication method is required?
11. What is the format of the data returned (e.g., JSON, XML)?
12. What is the directory structure and where is the data located?
13. What is the expiry date for the API access key/ token?
14. Do you have any specific primary key for the tables?
15. How are changes to the data sources communicated to us?
16. Are there different versions of the API, and which version should we use?
17. Are there backup sources or alternative methods for data retrieval in case of primary source failure?
18. Are there any specific data formats or standards that need to be followed?
19. Are there any constraints or preferences regarding cloud providers (e.g., AWS, Azure, GCP) or on-premise solutions?
20. How should the data be integrated (e.g., ETL, ELT)?
21. What transformations or aggregations are required on the data?
22. Are there any specific tools or technologies you prefer for data processing?
23. What is the expected timeline for the project?
24. Are there any critical milestones or deadlines that we need to be aware of?
25. How often do you expect progress updates and reports?
26. Who are the primary stakeholders for this project?
27. Who will be the main point of contact, and how should we communicate (e.g., email, meetings, collaboration tools)?

Motivation:

This page is the landing zone for anything you need to know about the data source. It does highlight the data scope in terms of regional span as well as account or tenant structure. In addition, you will find all integrated data assets as well as an entity relationship diagram to understand the context of each asset.

Last Doc Update	05-06-2024
Doc Status	In-Progress
Development Status	In-Progress

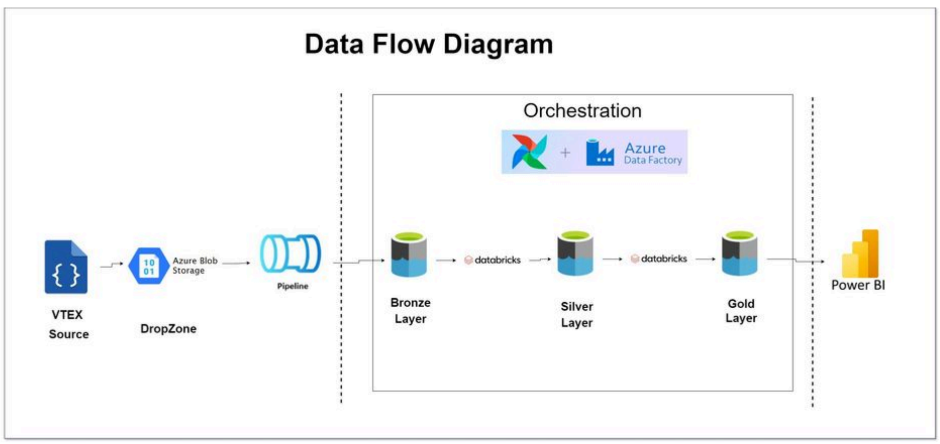
Contributor	Basheer Ahmed Manvi Khandelwal Sathya Priya

Solution Brief:

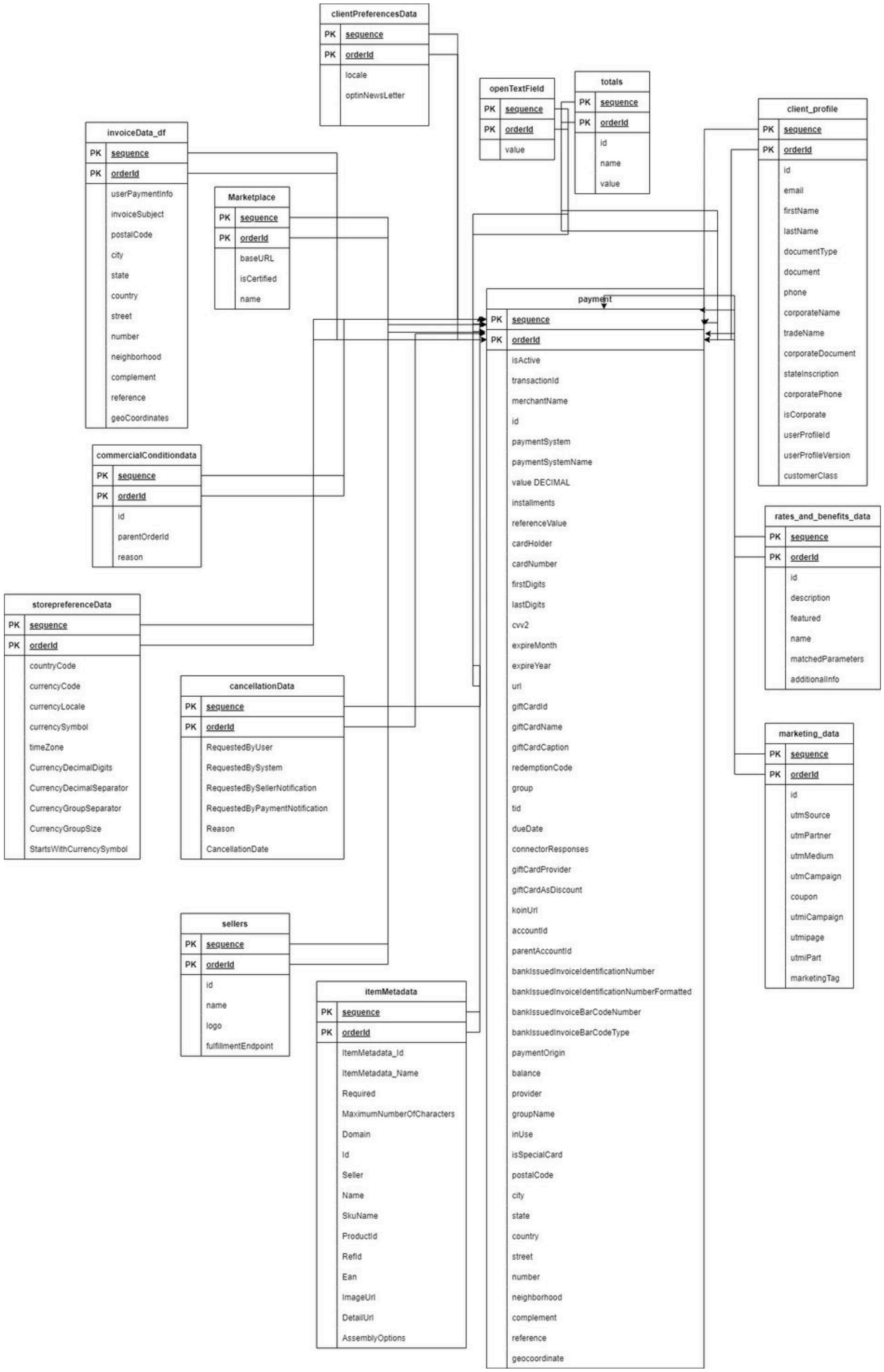
Stakeholder:

Role	Data Owner/s	Source System Contact Person	Source System Resolver Group	Business Sponsor	GDP Product Owner	GDP Data Architect	Engaged developer

Data Flow Diagram:



Entity Relationship Diagram:



Data Dictionary:

Bronze

Location:	/Workspace/Users/rahul.siwal2022@vitstudent.ac.in/vtex_project/soruce_to_bronze
Format:	json
Filename:	
	Data profiling file link:project_table.xlsx
Silver	
Location	/Workspace/Users/rahul.siwal2022@vitstudent.ac.in/vtex_project/bronze_to_silver
Format:	Parquet
Table name	vtex_db
	Data profiling file link: project_table.xlsx

Tables to be created from the given data:

Parent	**Child**	**Sub Child**	**Sub Sub Child**	**Sub Sub Sub Child**
orderId				
sequence				
marketplaceOrderId				
marketplaceServicesEndpoint				
sellerOrderId				
origin				
affiliateId				
salesChannel				
merchantName				
status				
workflowIsInError				
statusDescription				
value				
creationDate				
lastChange				
orderGroup				
totals				
items	itemAttachment			

	offerings			
	attachmentOfferings			
	additionalInfo	categories		
		dimension		
	priceDefinition	sellingPrices		
marketplaceItems				
clientProfileData				
giftRegistryData				
marketingData				
ratesAndBenefitsData				
shippingData	address			
	logisticsInfo	slas	pickupStoreInfo	
		deliveryIds		
		deliveryChannels		
		pickupStoreInfo		
	selectedAddresses			
transactions	payments			
packageAttachment				
packages				
callCenterOperatorData				
followUpEmail				
lastMessage				
hostname				
invoiceData				
changesAttachment				
openTextField				
roundingError				
orderFormId				
commercialConditionData				
isCompleted				
customData				
storePreferencesData	currencyFormatInfo			

allowCancellation				
allowEdition				
isCheckedIn				
marketplace				
authorizedDate				
invoicedDate				
cancelReason				
itemMetadata	Items	AssemblyOptions	InputValues	takeback
subscriptionData				
taxData				
checkedInPickupPointId				
cancellationData				
clientPreferencesData				

Primary and Foreign Keys:

- **Order**: PK - orderId
- **Totals**: PK - totalId, FK - orderId
- **Items**: PK - itemId, FK - orderId
- **ItemAttachment**: PK - itemAttachmentId, FK - itemId
- **Offerings**: PK - offeringId, FK - itemId
- **AttachmentOfferings**: PK - attachmentOfferingId, FK - itemId
- **AdditionalInfo**: PK - additionalInfoId, FK - itemId
- **Assemblies**: PK - assemblyId, FK - itemId
- **MarketplaceItems**: PK - marketplaceItemId, FK - orderId
- **ClientProfileData**: PK - clientProfileId, FK - orderId
- **MarketingData**: PK - marketingDataId, FK - orderId
- **RatesAndBenefitsData**: PK - ratesAndBenefitsDataId, FK - orderId
- **ShippingData**: PK - shippingDataId, FK - orderId
- **Address**: PK - addressId, FK - shippingDataId
- **LogisticsInfo**: PK - logisticsInfoId, FK - shippingDataId, addressId
- **PaymentData**: PK - paymentDataId, FK - orderId
- **Transactions**: PK - transactionId, FK - paymentDataId
- **Payments**: PK - paymentId, FK - transactionId
- **Sellers**: PK - sellerId, FK - orderId
- **StorePreferencesData**: PK - storePreferencesDataId, FK - orderId

Tables to be created:

1. clientProfileData
2. cancellationData
3. itemData

4. paymentData
5. orderDetails
6. marketingData
7. marketplaceData
8. itemsData
9. sellersData
10. ratesandbenefitData
11. shippingData
12. totalsData