#### Accelerate Data Product Delivery from Model to Marketplace

Presenters: Yetkin Ozkucur & Nadeem Siddique



#### **Gartner's recommendation**

"As analytics artifacts are delivered to consuming users, more must be taken into consideration than just the object itself. Organizations taking a product approach to analytics delivery will find increased trust across domains and reduced redundancies in development."



- Gartner Reference Architecture for Federated Analytics, July 2023



#### Data product

Noun. [ dey-tuh prod-uhkt]

1. A product that facilitates an end goal through the use of data – DJ Patil, former U.S. Data Scientist

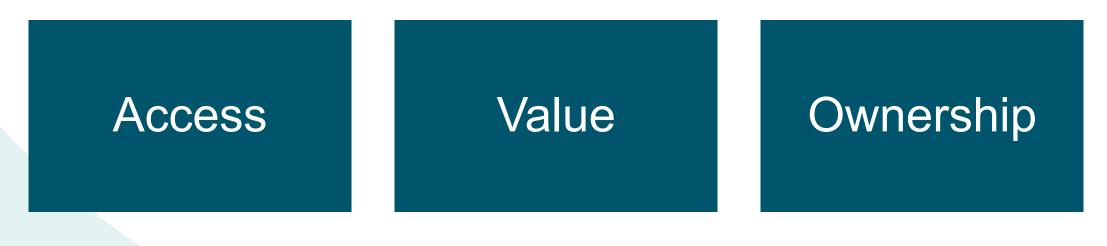
2. Primary Stakeholders include data modelers, data scientists, data analysts, end data consumers and decision makers

3. IDC Prediction: "By 2026, 60% of leading enterprise intelligence companies will have **identified data products**, and 15% will have **attributed business value** to the products with a data valuation methodology



erwin.com | confidential

#### Three key attributes of data products



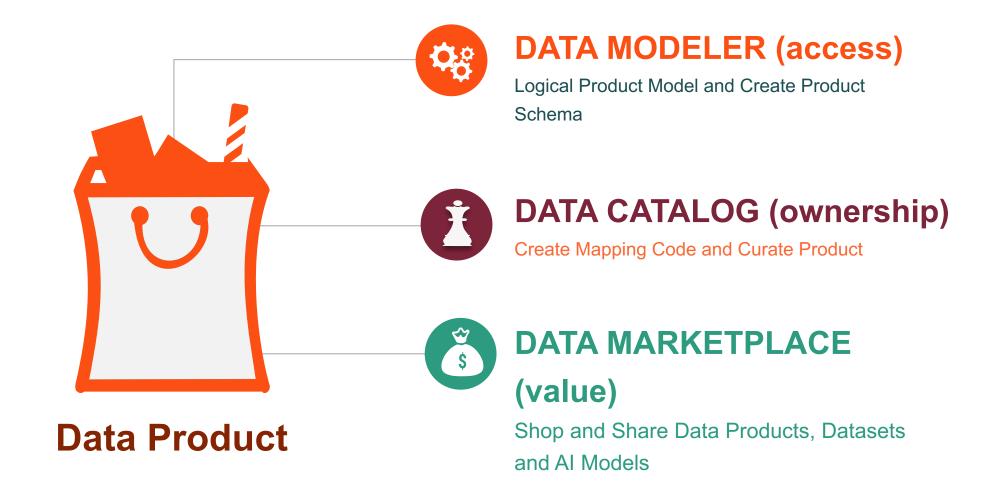
Making data products available, discoverable, and reusable The business value derived from the data products to drive data valuation Who is responsible to maintain, monetize and nurture the data product through its life

IDC FutureScape Worldwide Future of Enterprise Intelligence 2024 Predictions



erwin.com | confidential

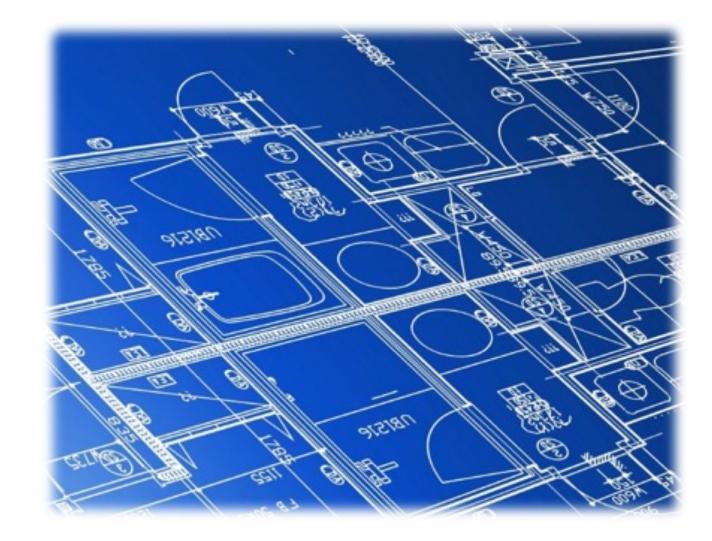
#### **Composition of a Data Product**





#### The role of data modeling

- Sets the table by structurally collecting business requirements for data
- Conceptual, logical and physical modeling
- Uniform definitions, descriptions and sensitive data information to be used in data product curation





#### The benefits of data catalog automation

- Generation of source-to-target data mappings and ETL code to gather data
- Automated generation of data lineage for data consumers
- Automated data profiling and data quality scoring to build data trust
- Integrated business context and governance guidance for consumers





		oto m	orkotr						
The valu			arkeip	nace	22	ASSET EXTENDED			
A one-stop sho	p for data pro	oducts	ADD TO COMPARE 🧪 ♀	<pre>x</pre>		Customer And Date Enterprise Date. 4	lytics Dottaset ent Punctional Area.	Compliance Dataset Dataset Discontectional Area,	Personal Loan Dotaset Dotaset B Pringelina Data, + B Functional Area,
Interprise Datasets Function.     Customer Analytics Dataset     Customer Analytics Dataset     Customer Analytics Dataset	ota Value core OLD Workflow Status No Status	Roting ****	t Togs Comptonce (COR) (Nicco)			Lorem (psum is simply dummy te industry: Lorem (psum has been tent ever since the 1500, when an type and scrambled it to my	the industry's standard dummy unknown printer took a gatey a		Leven tourn is circuly duriney text of the prioriting and toperatil Industry, Leven tourns tourns to level. The Industry is control of domini- ties over since the ISOs, when an universe priorities tour is galaxy type and iscontrolled 15 months is type spectrum book. The survival not our hybrid and and the transmission of the transmission toperating remaining essentially unchanged 15 most prior toperating remaining essentially unchanged 15 most prior to the ISOs with the network of unchan these controls galaxy symmetry possible. Registriater including variations of some types of uname to the set of the transmission of unchanged to the transmission.
Bad Par	Asset Information  sfinition rem (prum is simply dummy text of the printing and typesetting		Governance Responsibilitie	Covernance Responsibilities Dataset Owner		★★★★★		****	*****
Coney     C		nen	Concession however		with specific individuals subject to Approvals Roboro Ronning		Instactore is contained and access should any be given subject to approvals	This Dataset contains restricted and confidential information Rationa Ranning	
Control Woother Dataset     Contempose Datasets or Punction.     Contempose Datasets or Punction.	Iord Ismen 😧 Classification	thould be shored		Customer And	o 😋	Compliance Dataset	Personal Loan Dataset		
Sho	Additional Properties	industry.	Share		D) Classification	Here days		Compar	(Conserve) (2)
Datasets, Al moc products all in o			Idresses dat oncerns thro ernance and	ugh data		A 2 200		et reviews w o a consume experien	r shopping
		Asset Type	Enterprise Sotoests Functional Area Datasets Compliance Dataset Usine spours is simply during test of the printing and types		Comptionce No Sta	•	0	ACC TABLE (C) CONTRACT (C) CONT	
		Data Value Score	Ind Party Credit Bureau Data Mil Credit Bureau data rates to the information callected and n				0	202100 0 100000	
			Emerginas Datasetis Functional Area Datasetis Customer Analytics Dataset Loren spoum is simply durrens text of the printing and types			····	0	ps and Reviews X fee Seasons — Auxiliarial level bacavait war Analytics Doctower 4,5	
		Dotanei Ourier A	Enterprise Datasets Functional Area Datasets Personal Loan Detaset etc. Lorum ipsum is simply durreny test of the printing and types			···	۰.	Na. 3	
		Enclore leaving	Consumer Banking ATM and branch location data with The data set includes information on ATM and branch location			tura e e e e e e e e e e e e e e e e e e e	0		
		I M A	Consumer Iterating Coll center data His data set includes information on customer interactions				0	Antoniadadar 2 de como de activitador de la como d	
		100	Analytics Card Customer Data ets: A Customer Credit Card Information Dataset which can be a			tan Distance	0	It is the most accurate subtrane analytics dataset for 8 hepothing	
erwin		Workflow Boh.s         V           Ratings         A           ★★★★☆ 5.800%         B	Consumer Credit Cord Credit Cord Dataset for Chatering this data set is used to develop a customer segmentation to				0	endow     e	
by Quest			erwin.com   confidentia	1					

# Data value scoring speeds discovery and aids in valuation and monetization

- User-defined, weighting of data quality scores, data consumer rankings, and governance curation completeness
- Automated assessment applying gold, bronze and silver tiering of value
- Based on *Infonomics* concepts developed by Doug Laney

≣ ♠		Datasets 🔻			Q		Sort by : Relevance 👻
Applied Filters	CLEAR ALL	1 - 20 of 32 result	ts for Browsing "Datasets"				CARD VIEW  < < 1/2
Datasets 🐼		Datasets	Consumer Banking ATM and branch location data This data set includes information on ATM and branch location	PII Confidential +1	Analytics Customer +2 Pending Appro	2 Ratings	
Explore By Category	/ Attribute	Datasets	Consumer Banking Call center data	Restricted Confidential	customer service Pending Appro	2 Ratings	
Asset Type	~		This data set includes information on customer interactions wit		Concenter	V -	¥ Mile
Data Score Value	~	岽	Analytics Card Customer Data	Pil Internal Only +1	No Tags Publish	Z Ratings	
Data Set Owner	~	Datasets	A Customer Credit Card Information Dataset which can be use			· ·	Q MIN
Data Steward Sensitivity Classification	~	) Datasets	3rd Party Credit Bureau Data	Pll Restricted	Analytics customer +2	2 Ratings Data Value Score 0 DQ Score: 40.37 0 Data Value Score	
lags	~	Dutusets	Credit bureau data refers to the information collected and mai			Oataset Ratings: 16     Dataset Complete	
Workflow Status	~	岸	Consumer Credit Card Credit Card Dataset for Clustering	Pil Internal Only	Critical Data Ele Draft	2 Ratings	
Ratings	^	Datasets	This data set is used to develop a customer segmentation to d		card	· ·	
★ ★ ★ ☆ & Above ★ ★ ★ ☆ ☆ & Above ★ ★ ☆ ☆ ☆ & Above ★ ☆ ☆ ☆ ☆ & Above ★ ☆ ☆ ☆ ☆ & Above ☆ ☆ ☆ ☆ ☆ None		Datasets	Consumer Credit Card Credit Card Fraud Detection Data This data set includes information on credit card transactions, L.	PII Restricted +1	Analytics Publish customer +1	*1Ratings ★	발 REQUEST ACCESS 문과 CON
Show My Favourites	^	र्ट्सू Datasets	Lending → Credit Card Lending <b>Credit Card Marketing Data</b> This data set includes information on marketing campaigns for	PII	customer Pending Appro	2 Ratings	알 REQUEST ACCESS 환 con
		Datasets	Customer Operations Customer Customer data sets are essential for businesses as they help to	Pll Confidential +1	Critical Data Ele Publish	* 2 Ratings ♥	
		Datasets	Customer Operations Customer Success Customer Churn Focused customer retention programs	PII Restricted +2	Critical Data Ele Customer +8	* å at ings * €	



### Solving hard problems with data products

Where should a bank build their next physical location?

The Problem(s)

Should it be an ATM or a full branch?

How should it be staffed?



Data Product Components

Internal datasets

- customer addresses
- Loan histories

External datasets

- home values per zip code
- School demographics

AI models

Associations to a Customer Segmentation data product

Business terms, rules and policies

#### **Potential Insights**

Identify which locations are underutilized or overburdened to adjust staffing accordingly

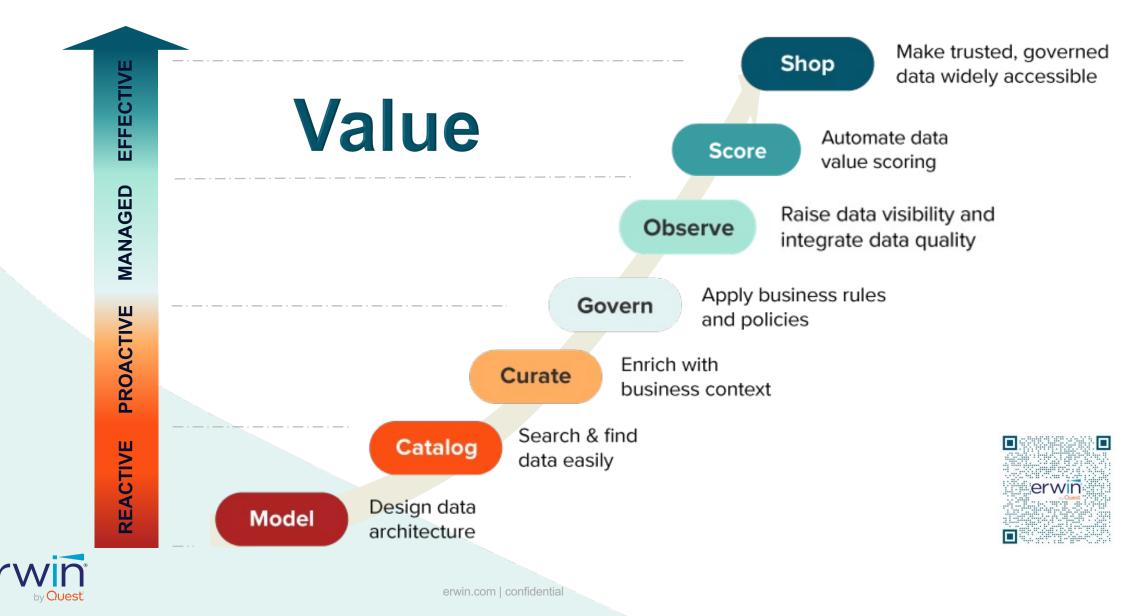
Improve the customer experience by staffing bi-lingual employees

Plan strategically, i.e., evaluate the potential ROI in opening new location branches or ATMs

### Data Product Delivery A data marketplace view



#### 7 steps from model to marketplace

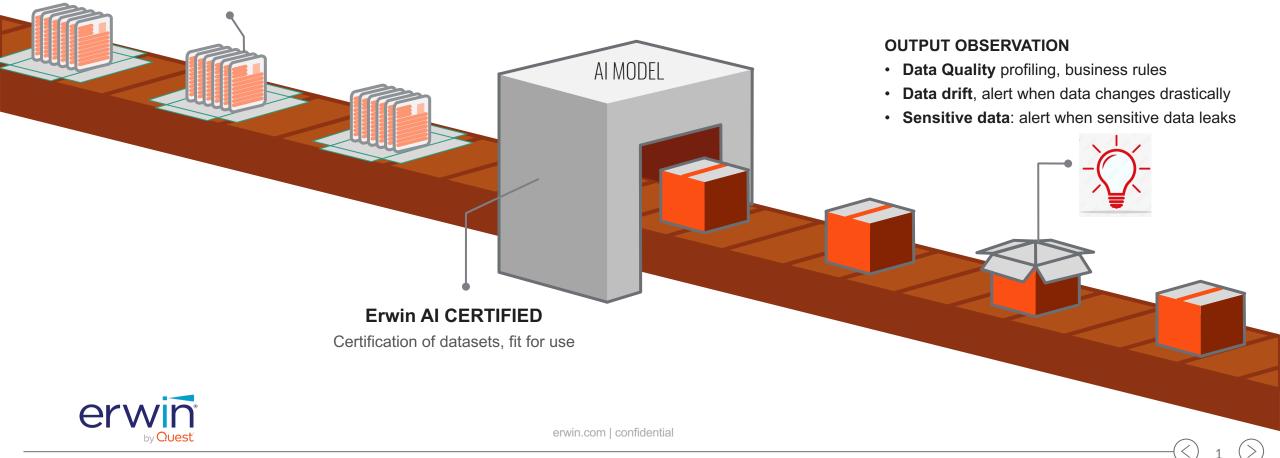


#### **INPUT DATASETS (7 steps)**

- Model, the blueprint
- Catalog, where the data resides
- Context, rules, policies, definitions
- Owner
- Sensitive data: PII, PHI, Confidential
- Data Lineage, where the data coming from
- Data Quality profiling, business rules
- Data Score, popularity, value, processes

#### erwin Al Model Governance

#### Govern Input/ Stamp /Monitor Output



## Thank You

Want to learn how erwin can help you?

#### **Experience** Data Intelligence

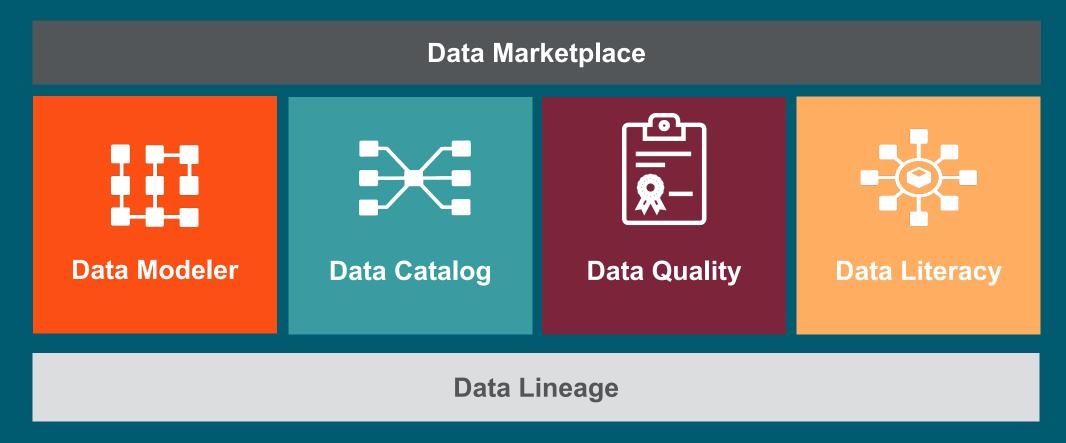
See first-hand the impact erwin can make in a hands-on "day-in-the-life" session for your company.

> To get started, visit <u>www.erwin.com</u> and request a demo here



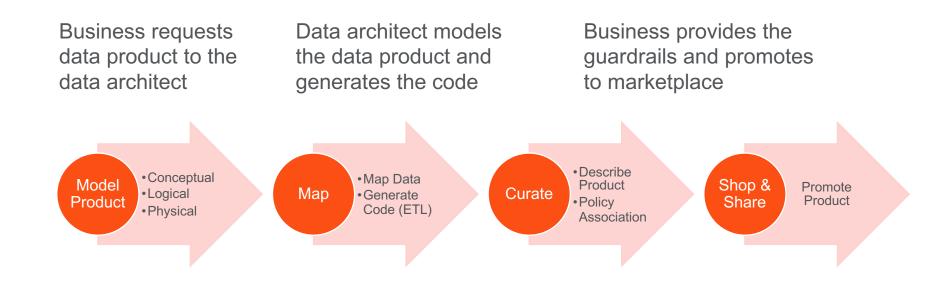


#### Today's erwin portfolio





### Accelerating model to marketplace delivery





### Where does data preparation fit in?

- Combine data marketplace capabilities with data prep tools like Quest Toad Data Point to deliver blended data from multiple sources easily
- Use scheduling capabilities to keep provided datasets refreshed

	{JS <b>O</b> N}	denodo <sup>:‡</sup>	RECORPT	cassandr	TERA	DATA		
	Power Bl	aws	& kafka	() ma	ongoDB	CouchDB	New Source	
MuleS	DB2	<b>Sho</b>	•	Ş	Spark	IBM	cloudera	
SQLS	Server 🕥	DATA VAULT <sup>24</sup>	SAP	python	@ per	ntaho 👖	AicroStrategy	
	Informa	tica	ableau		ACLE'	0	COBOL	
(State	ave 🦻	talend		(IP)	IBM Cog	GitHub	Java	
(	erw	/10						

by Quest

