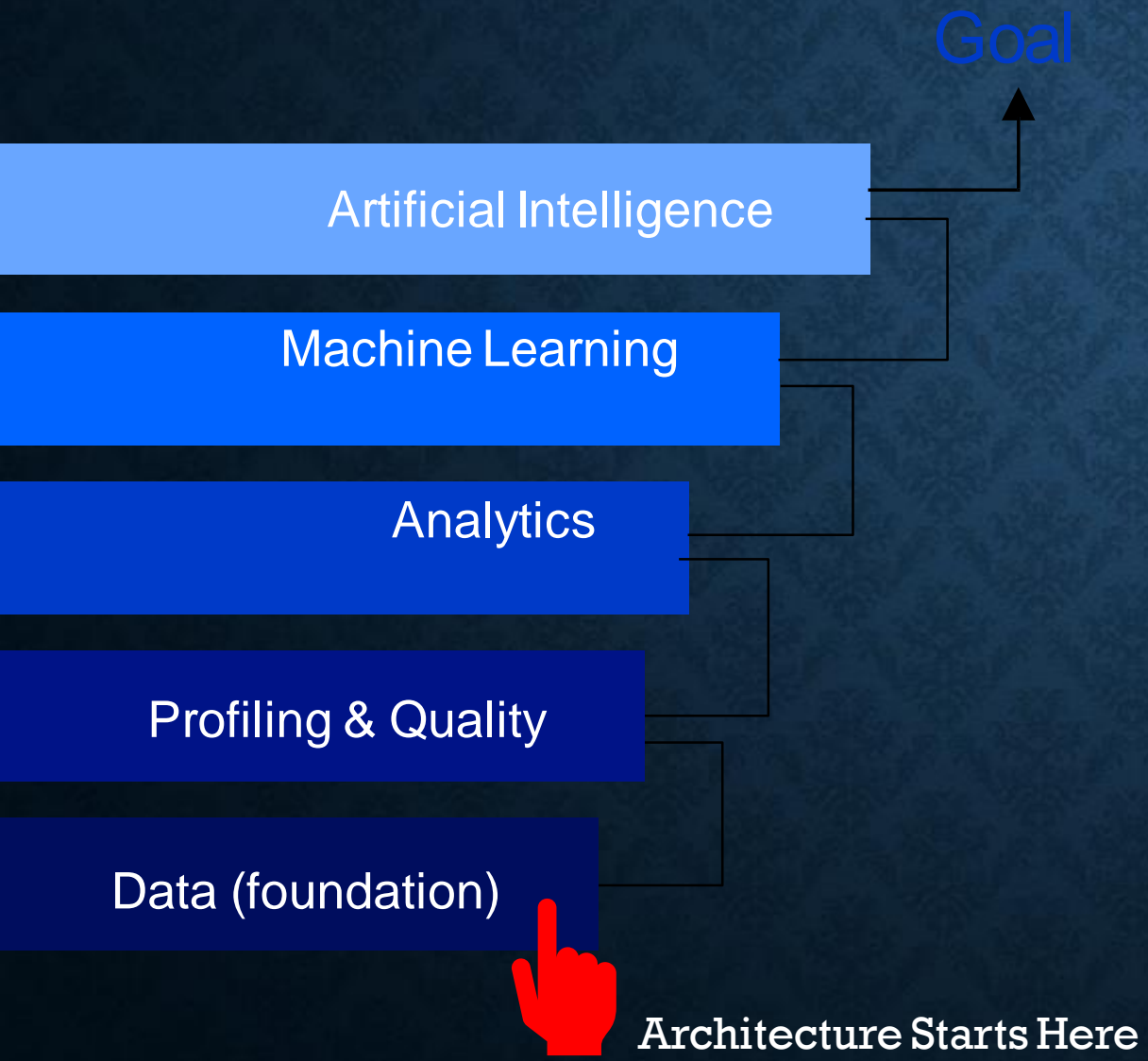


THE BUILDING BLOCKS OF AI



....IT'S ALL ABOUT THE DATA

AI Model



Structured data

Unstructured and exogenous data

^ ...New approaches are required

Familiar structured data approaches to build trust

Discovery Tools

Lineage

Data Quality Profiling

DQ Rules

Business Glossary

Active Stewardship

Governance

There is no AI without IA

Information architecture and process matter!

81%

do not understand
the data required
for AI

8X

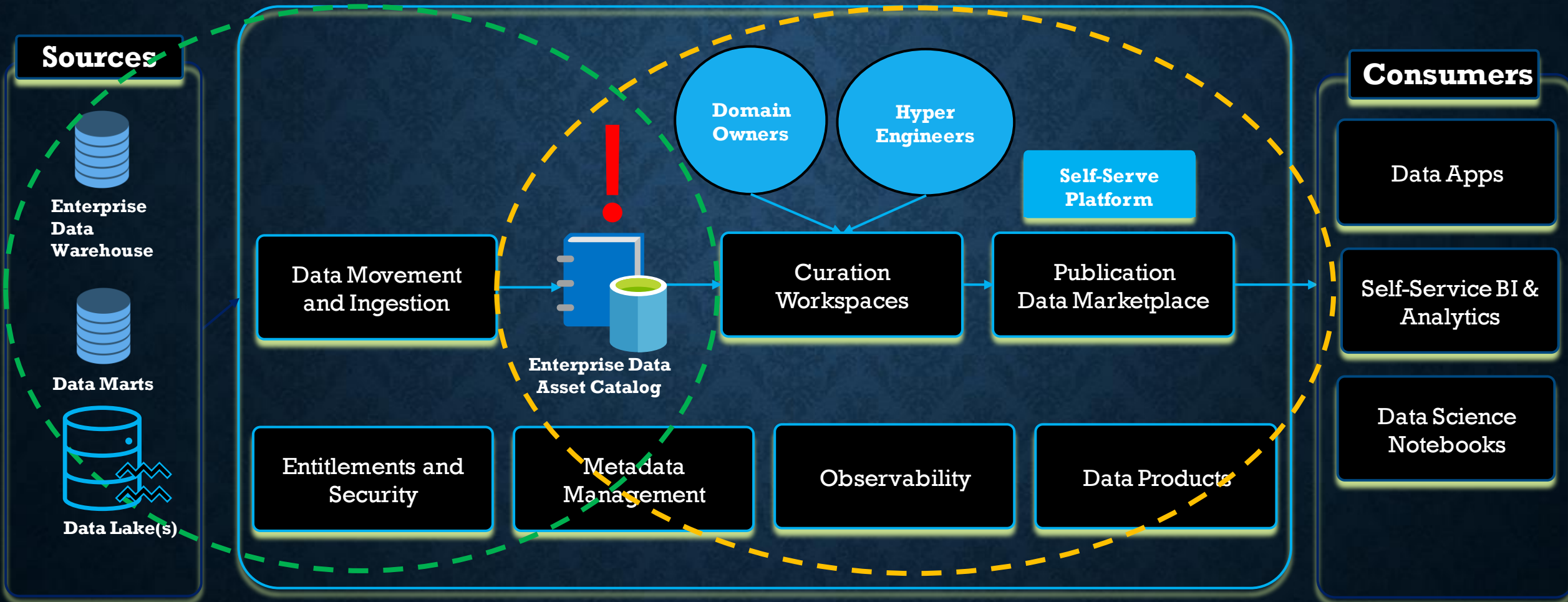
AI pioneers are
8X more likely to
have a robust
data architecture

*“No amount of AI algorithmic sophistication
will overcome a lack of data (architecture)...”*

Data collection and preparation is the most
time consuming and difficult part of AI.



EMERGING IA & DATA STRATEGY - DATA FABRIC AND DATA MESH

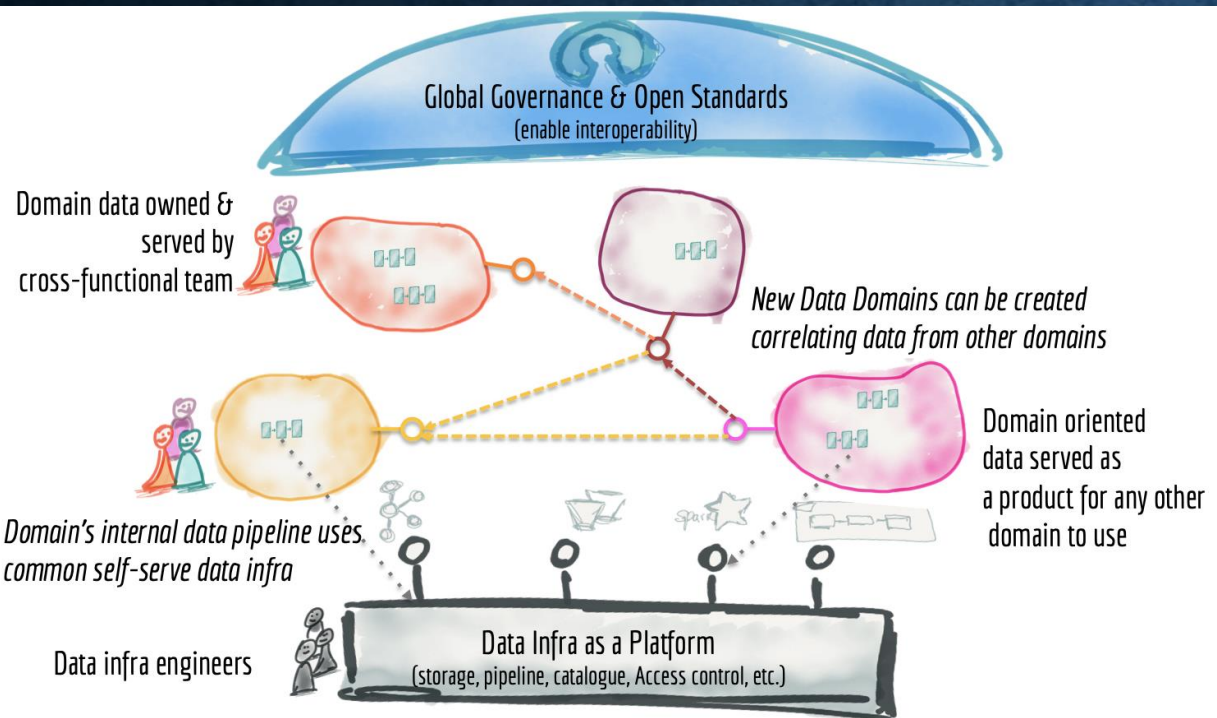


Data Fabric – Infrastructure, Technology, Architecture driven

Data Mesh - Business Centric domain ownership of the data

DATA CATALOG

THE NUCLEI OF A DATA MESH*



- A data product must be easily discoverable especially with a data catalogue, with their meta information such as their owners, source of origin, lineage, sample datasets, etc. This centralized discoverability service allows data consumers, engineers and scientists in an organization, to find a dataset of their interest easily. Each domain data product must register itself with this centralized data catalogue for easy discoverability.
- Note the perspective shift here is from a single platform extracting and owning the data for its use, to each domain providing its data as a product in a discoverable fashion.
- Data catalog platforms provide central discoverability, access control and governance of distributed domain datasets.

HOW TO BUILD FOR THE FUTURE – 5 KEY AREAS TO FOCUS

1. **Align your data and business strategies.** Act as a change agent to champion data literacy across the business.
2. **Build the right teams.** Ensure all roles—data architects, data engineers, data analysts and data scientists—work together to establish a standard of data observability and trust.
3. **Empower your teams with self-service data usage.** Provide the right tools for real-time, self-service data ingestion, access, and analysis so that your team can focus on operationalizing AI.
4. **Ensure data governance.** Adopt metadata, data catalogs to help standardize data usage across your organization and ensure that data is fit for purpose.
5. **Adopt the right technology.** Look for a solution that simplifies access to data, such as a data fabric architecture approach.

QUESTIONS?



<http://www.linkedin.com/in/srinisankar>

<https://twitter.com/srinisankar>