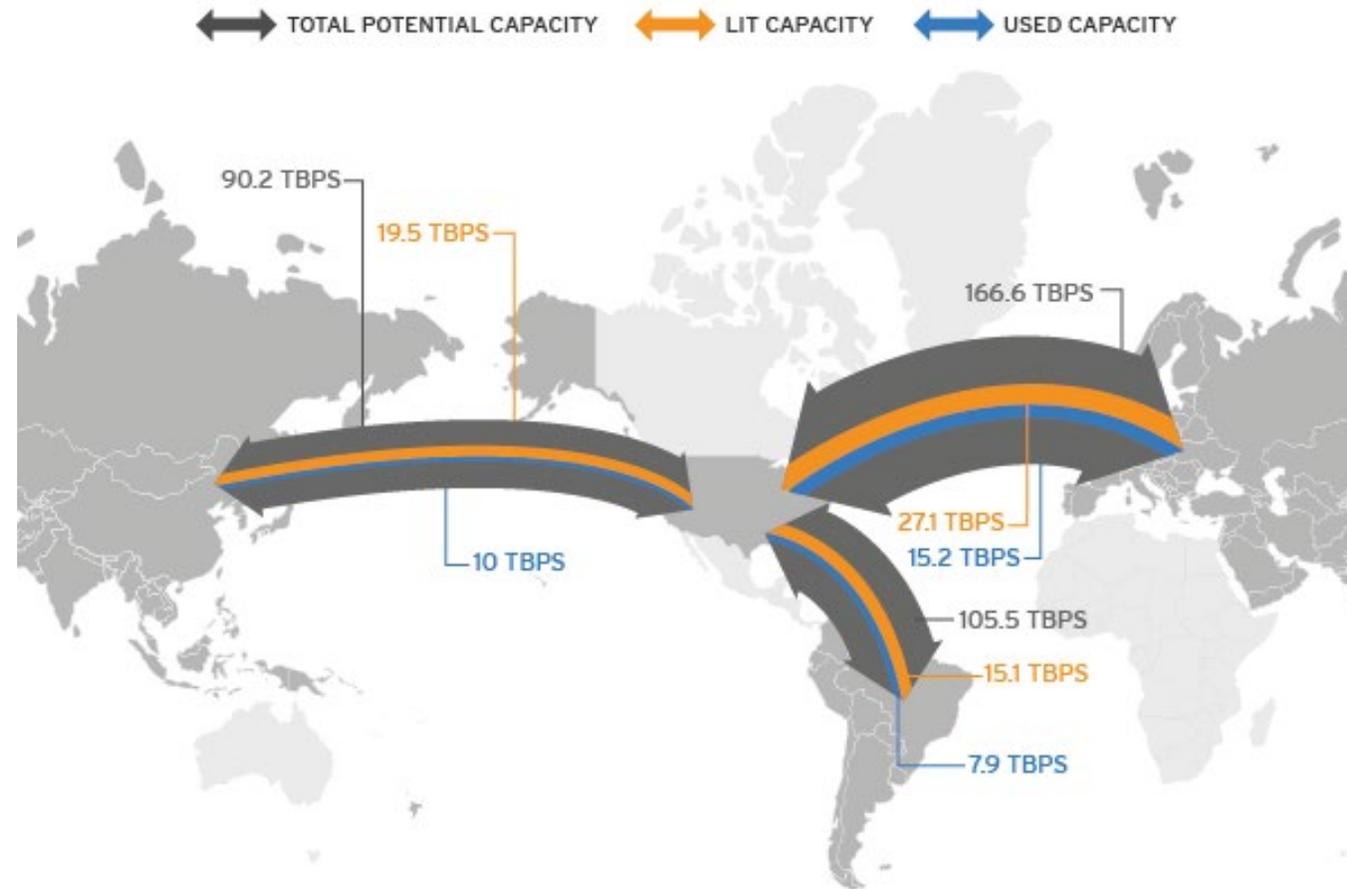


# Data Governance in TF

Sray Agarwal  
Kaushik Srinivasan

# Cross border data flows



# A Roadmap for Cross-Border Data Flows

Source: World Economic Forum

02

## Establish a level of data protection

Establish national legal frameworks that protect data of private individuals. Complement this with laws that protect proprietary rights.

04

## Hardwire accountability between nations

Establish cooperation mechanisms between national authorities to hold governments accountable for the security and confidentiality of the data they share, while making allowances for compliance.

06

## Future-proof the policy environment

Allow for the possibility of future alternative models (such as federated learning models and data trusts) that can also fulfil the spirit of cross data flows.

01

## Allow data to flow by default

Prohibit data localization requirements except in very specific circumstances in order to create regulatory certainty for businesses.

03

## Prioritize cybersecurity

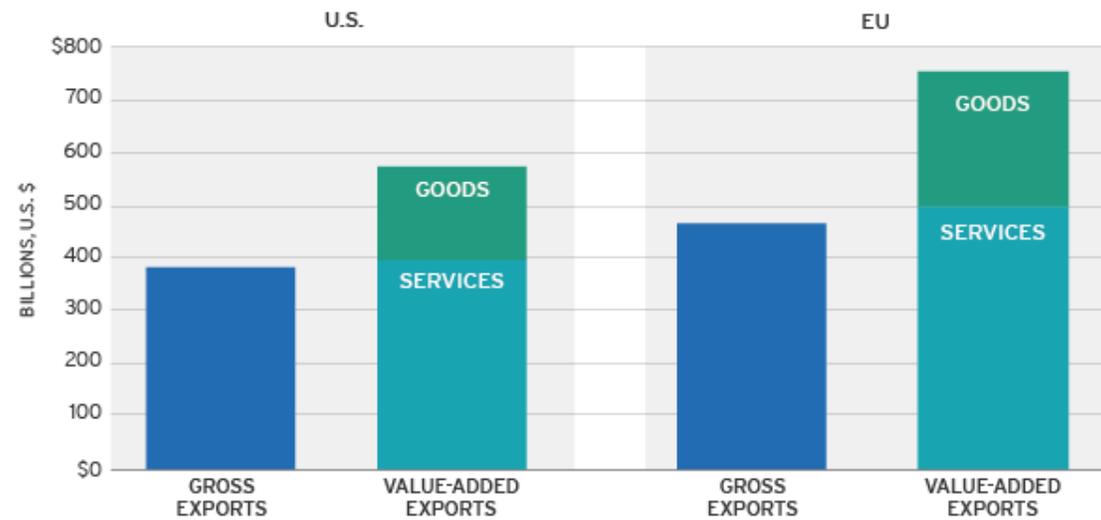
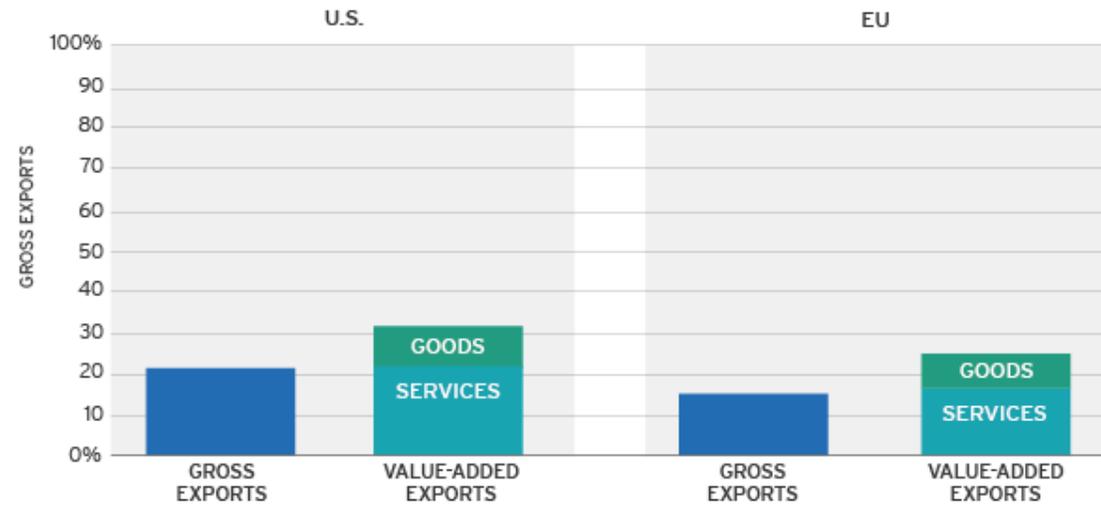
Enact transparent cybersecurity legislation in line with international norms and maintain robust data security infrastructure.

05

## Prioritize connectivity, technical interoperability, data portability and data provenance

Prioritize the development of the connectivity infrastructure as a prerequisite to building a local data economy, encourage technical standards to increase interoperability, facilitate data portability at the B2B level to support SMEs, and encourage data publishers to ensure data integrity.

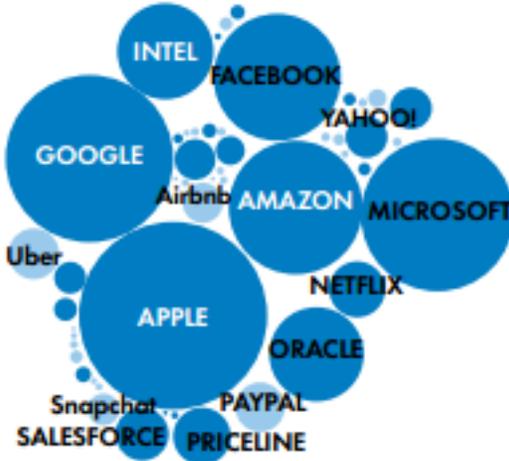
# Digitally deliverables export - 2012



# Data flow across continents

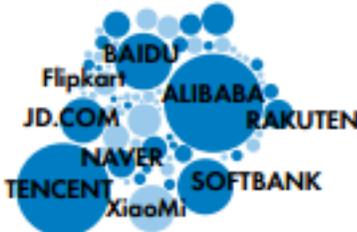
Public Private

## NORTH AMERICA



63: \$2.8 TRILLION

## ASIA



42: \$670 BILLION

## EUROPE

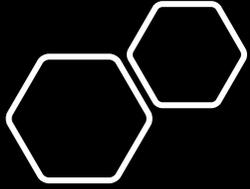


27: \$161 BILLION

## AFRICA & LATIN AMERICA



3: \$61 BILLION



# Data Governance

Centrally controlled and managed  
account

On-demand accessibility

Document control

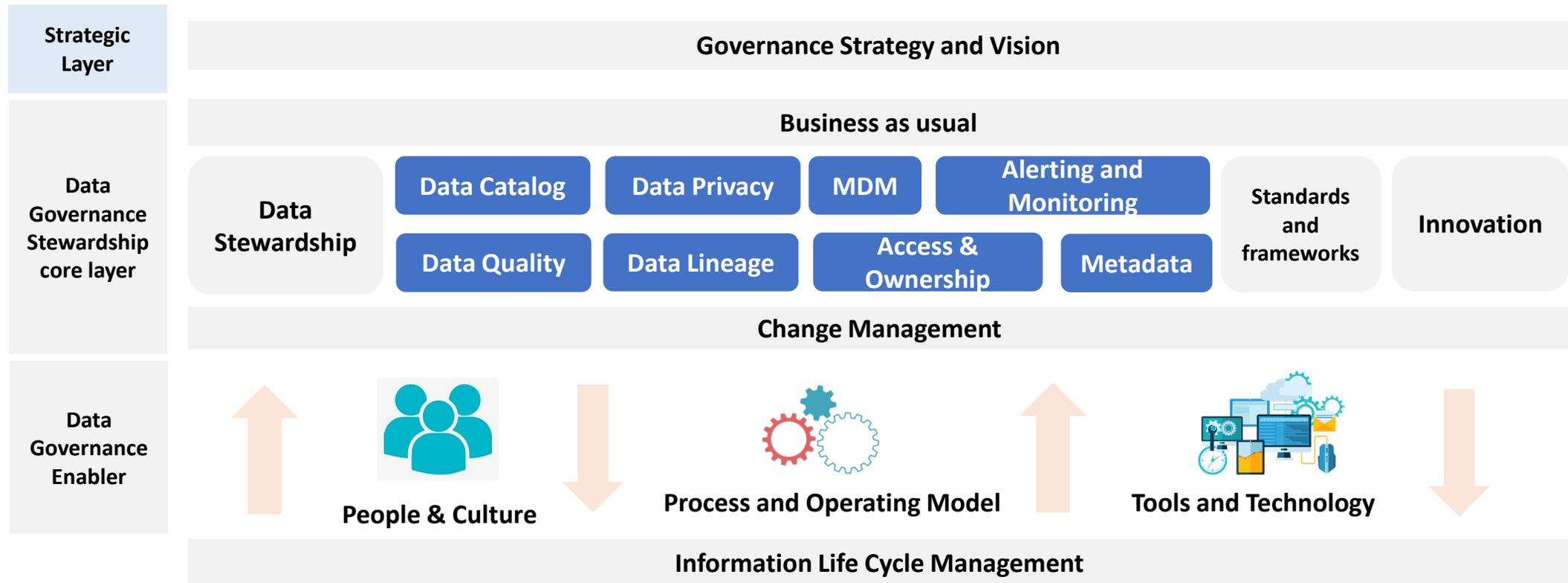
Security

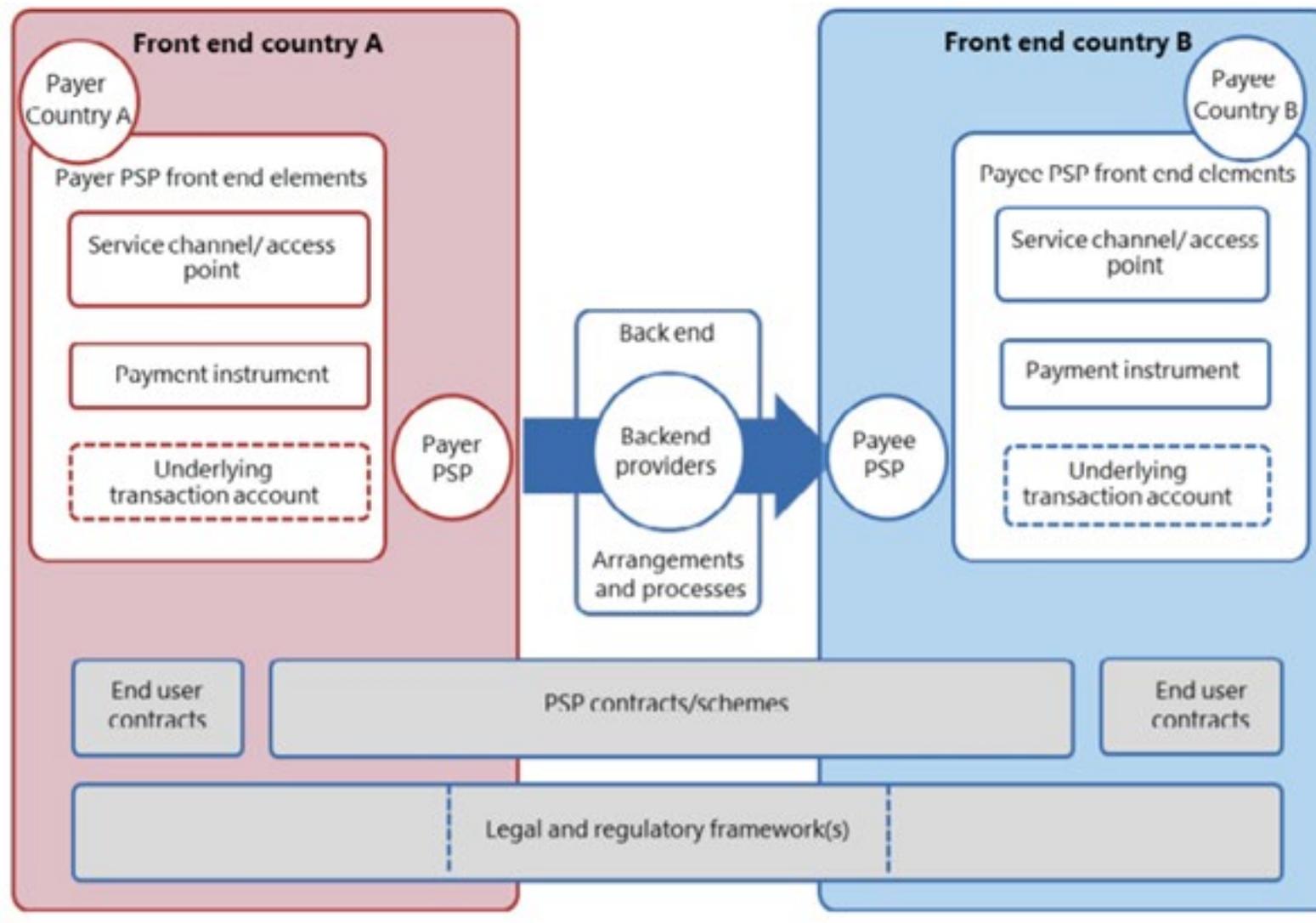
# Data Governance Framework

Data Governance Office and Policy
Data Quality Management – Automation and Operations
Data Dictionary, Business Glossary and Metadata Management
Data Catalog
Data Lineage
Data Privacy and Security
MDM and Reference Data

<b>DG Assets</b>	DG Prioritization Framework	DQ Rules Repository
	Metadata Mgmt Framework	Business Glossary Template
	DG Metric with Business KPIs	Data Dictionary Template

# Data Governance & Best Practices





*As of January 2022, the Netherlands come in hot at second place of top data breaches in Europe with 92,657 reports*

*Dutch municipality Assen, an employee had sent a file containing 530 persons' personal data to the wrong email*

*An unsecured server resulted in the **exposure of 3 terabyte of data** including airport employee records*

*In 2014, hackers gained access to databases of sensitive data via credentials of 3 employees of ebay*

# Differential Privacy (DP)

## Definition



Achieved by introducing some **noise** which is enough to protect privacy and at the same time limited enough so that the provided information is still useful.

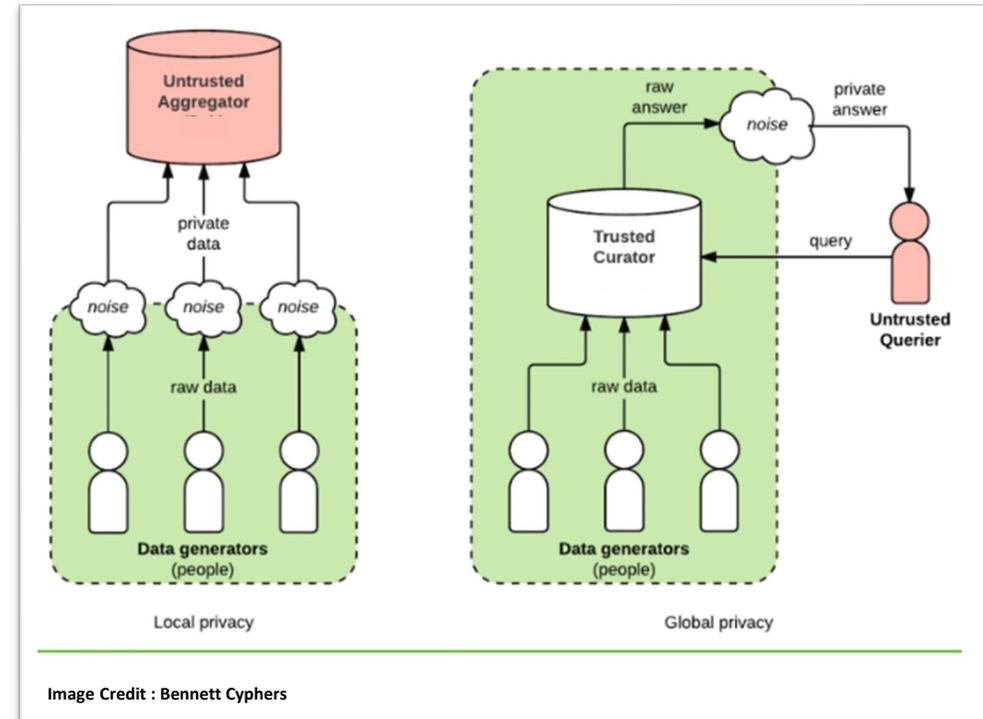
## Goal



Allow data analysts to build accurate models without sacrificing the privacy of the individual data points.

# How DP Works

- Introduces privacy loss parameter ( $\epsilon$ ) to the dataset. This adds randomness to data.
- A high value of  $\epsilon$  means more accurate but less private data
- Noise can be added to **data** and/or **algorithm**



# Value generation with synthetic data

*Any data algorithmically generated approximating original data.*

## Motivation:



Safety and Privacy



Industry Collaboration



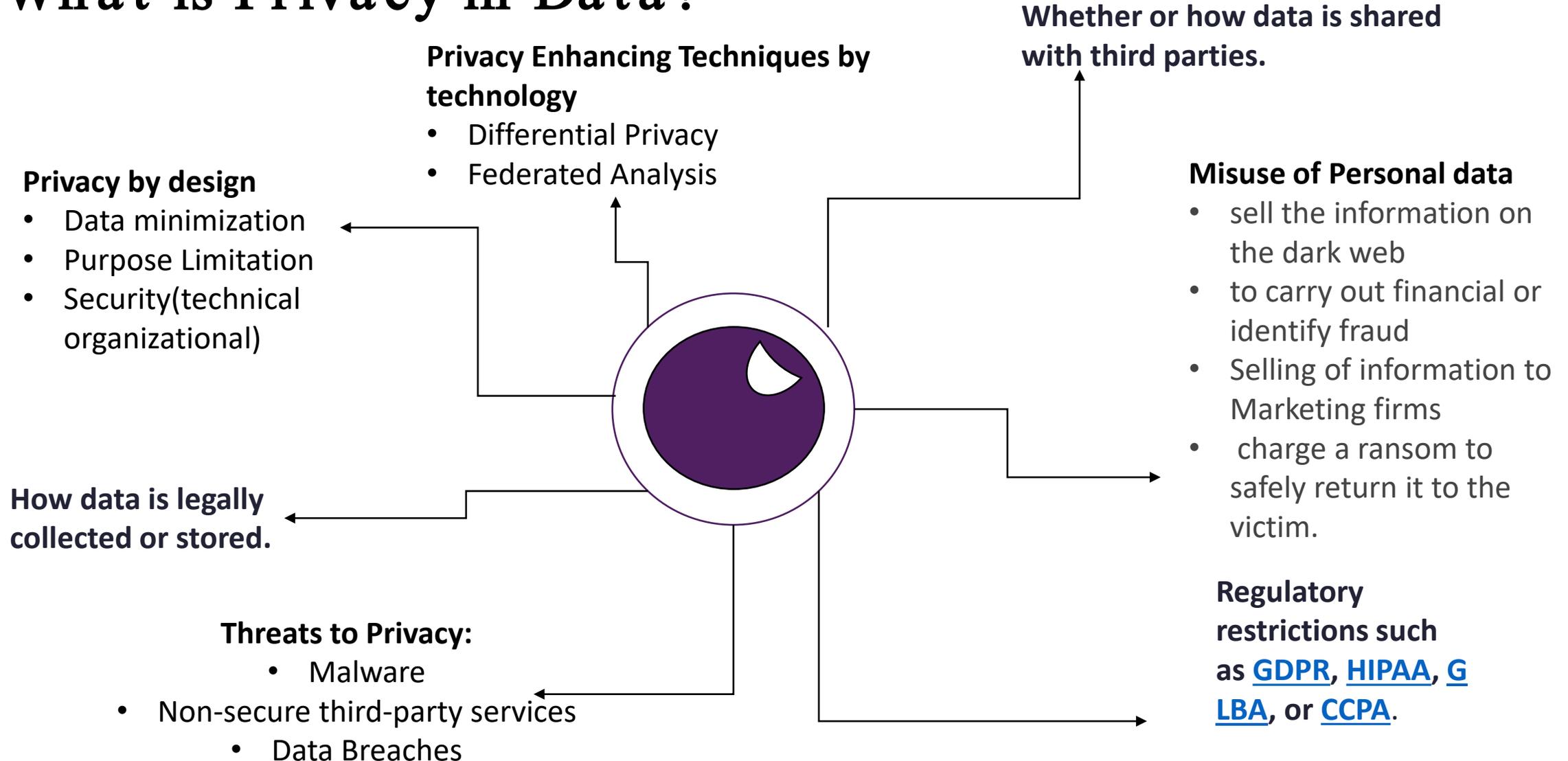
Responsible AI

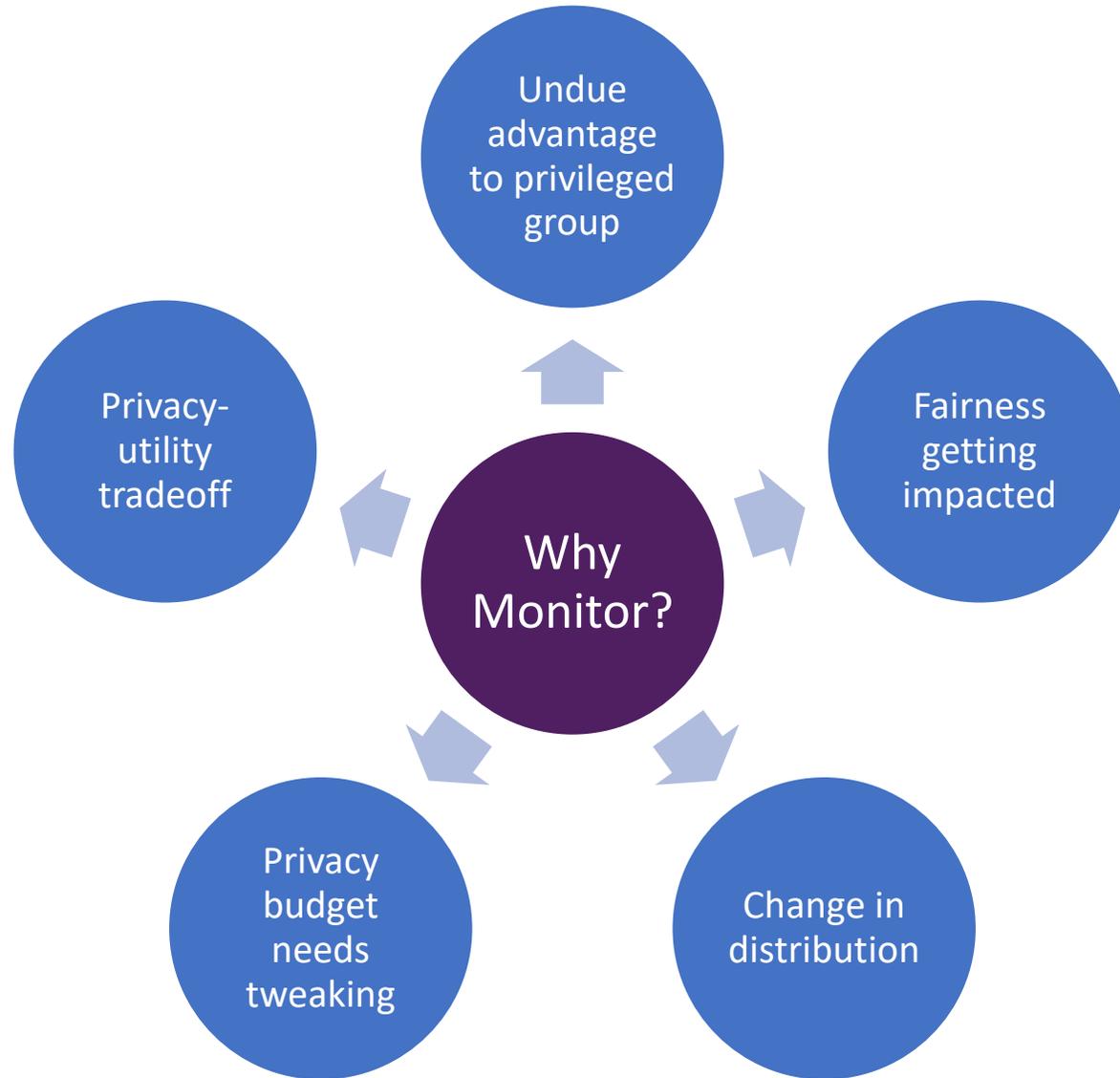


Secure

- Create safe datasets that retain the same insights and statistical integrity equivalent to original data source
- Develop cross-domain AI use cases to drive industry collaboration
- Drive responsible AI practices through balanced synthetic datasets or de-bias datasets for ML/AI model testing
- Defend against re-identification and joinability attacks.

# What is Privacy in Data?





## A New Approach to Governing Cross-border Data Flows

1 Encourage states to develop plans for the regulation and exchange of different types of data

2 Give people greater voice and greater control over their data

3 Clarify the rules and exceptions to the rules so nations do not restrict cross-border data flows more frequently or broadly than necessary, especially in the name of national security or cyber security

4 Provide clarity on what types of practices should be banned because they are trade-distorting

5 Delineate how nations should or should not respond to state actions that distort cross-border data flows

# Data Privacy and localization

## Data Residency

- Data stored in geographical location of choice

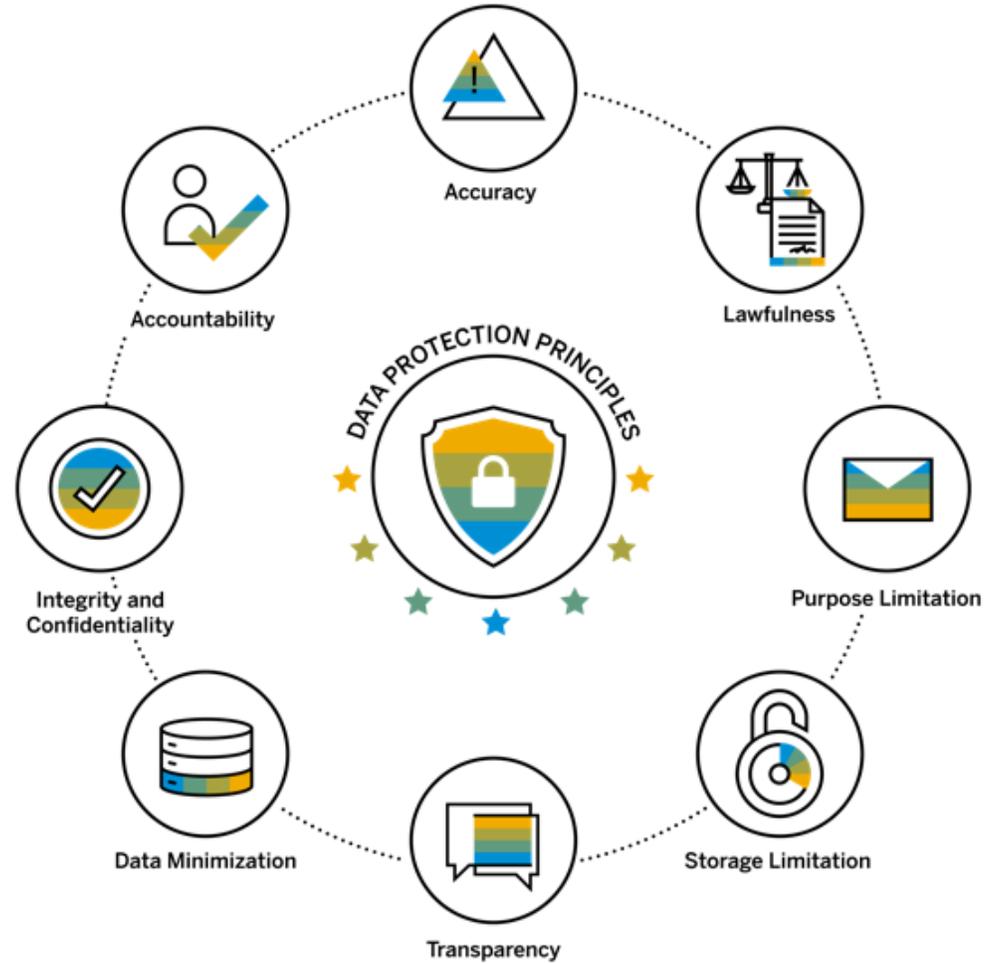
## Data Sovereignty

- Data is subject to law of the country where it is stored

## Data Localization

- Copy of data to be held within country's border for ease of auditing

# Data Protection



# Need for data governance and standards

- **Integrity:** All actors within the program should act honestly and be forthcoming about things like constraints, challenges, and other impacts of data governance decisions.
- **Transparency:** Processes should be clear and transparent to both participants and auditors in how practices and controls will be introduced and implemented.
- **Auditability:** Data Governance activities should be auditable and accompanied by documentation to support compliance-based and operational auditing requirements.
- **Accountability:** You need to define accountabilities for cross-functional and data-related decisions, processes, and controls. Know where the buck stops.
- **Stewardship:** Know, assign, and delegate governance stewardship activities that are the responsibilities of both individual contributors and data stewardship groups.
- **Checks & balances:** Introduce checks and balances between business and technology teams, creators and collectors of data, and anyone who uses or manages information.
- **Standardization:** Your Data Governance program's focus should be on introducing and supporting the standardization of enterprise data.
- **Change management:** Support proactive and reactive change management activities throughout the processes, from working with data to personnel best practices.

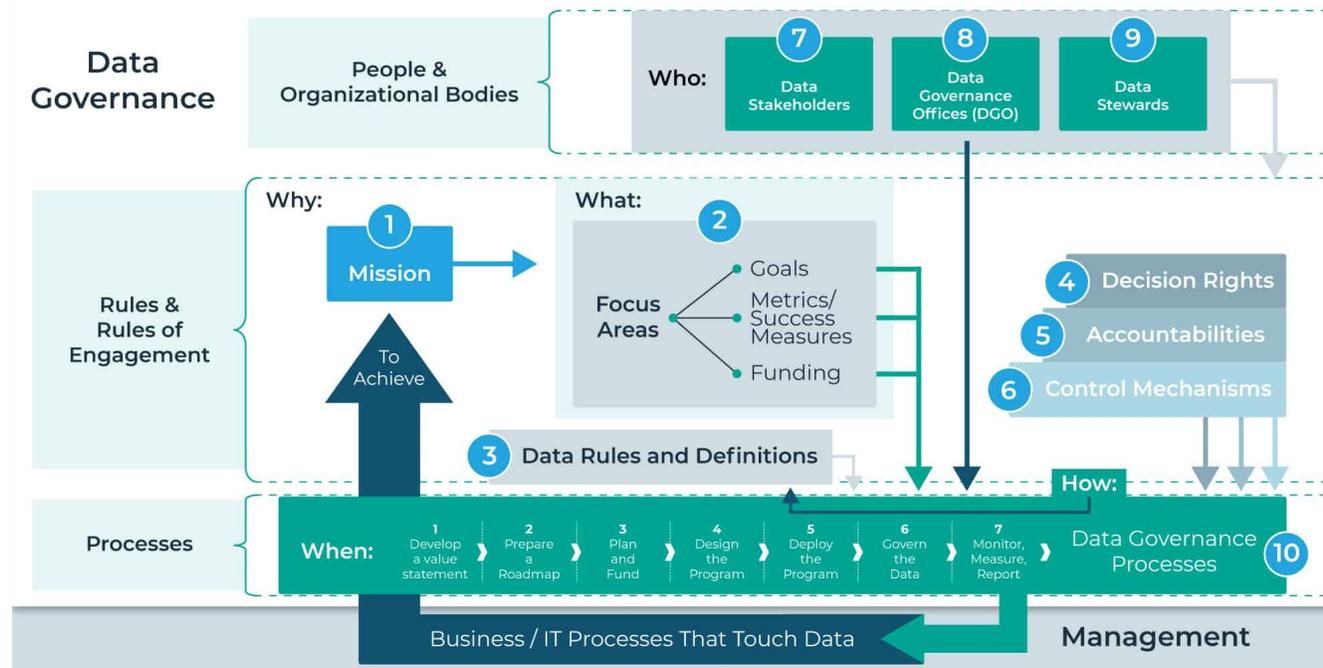
# Benefits of data governance

- **Centralized policies and systems** reduce IT costs related to data governance
  - **Data standards** allow for better cross-functional decision making and communication
  - **Compliance audits** are easier to manage, and compliance standards are easier to maintain
  - **Business intelligence** for short and long-term planning, including mergers and acquisitions, gets fuel from data
  - **Data growth** is controlled and organized
  - **Stable data** makes adapting to new data and privacy legislation easier
- 



# Why data governance

- **Management:** For top management this will ensure the oversight of corporate data assets, their value and their impact on the changing business operations and market opportunities
  - **Finance:** For finance, this will safeguard consistent and accurate reporting
  - **Sales:** For sales and marketing, this will enable trustworthy insight into customer preferences and behavior
  - **Procurement:** For procurement and supply chain management, this will fortify cost reduction and operational efficiency initiatives based on exploiting data and business ecosystem collaboration
  - **Production:** For production, this will be essential in deploying automation
  - **Legal:** For legal and compliance, this will be the only way to meet increasing regulation requirements
- 



**Definition:**

Data Governance is the exercise of decision making and authority for data-related matters.

It's a system of decision rights and accountabilities for information-related processes, executed according to agreed upon models which describe who can take what actions with what information and under what circumstances, using what methods.

**Processes for governing how data is used, and when, and by whom**

- |   |   |
|---|---|
| 1. Aligning Policies, Requirements & Controls | 7. Issue Resolution                     |
| 2. Establishing Decision Rights               | 8. Specifying Data Quality Requirements |
| 3. Establishing Accountability                | 9. Building Governance into Technology  |
| 4. Performing Stewardship                     | 10. Stakeholder Care and Support        |
| 5. Managing Change                            | 11. Stakeholder Communications          |
| 6. Defining Data                              | 12. Measuring and Reporting Value       |



Thank You