

# FIAP GRADUAÇÃO

# DATA SCIENCE

## DATA GOVERNANCE & DATA SECURITY MANAGEMENT

Prof. Dr. Renê de Ávila Mendes

# Objetivos da disciplina

**DISCIPLINA:** Data Governance & Data Security Management

**OBJETIVOS:** Descubra como funciona um **projeto de banco de dados** dentro de um ambiente corporativo, aplicando **técnicas de levantamento e documentação de requisitos**, aderente aos projetos de bancos de dados e aprenda a representar esses requisitos em arquiteturas de solução tecnológica para Data distribution e Data integration, modelos de estruturas de dados e dicionários de dados buscando **Data quality**. Garanta a qualidade dos dados de uma empresa para prover os melhores subsídios à tomada de decisão de negócio, praticando **Data cleaning** para limpar, harmonizar, complementar e corrigir dados inconsistentes, incompletos ou incorretos. Compreenda como funciona o **ciclo de vida da informação** e as responsabilidades administrativas sobre os dados de negócio, buscando qualidade, segurança e compatibilidade com políticas de administração de informação corporativas auditáveis, aplicando práticas atuais de **Data profiling** e conhecendo os princípios de **Data auditing**, de forma a atender a **Lei Geral de Proteção de Dados (LGPD)**.

# Assuntos – 2º Semestre

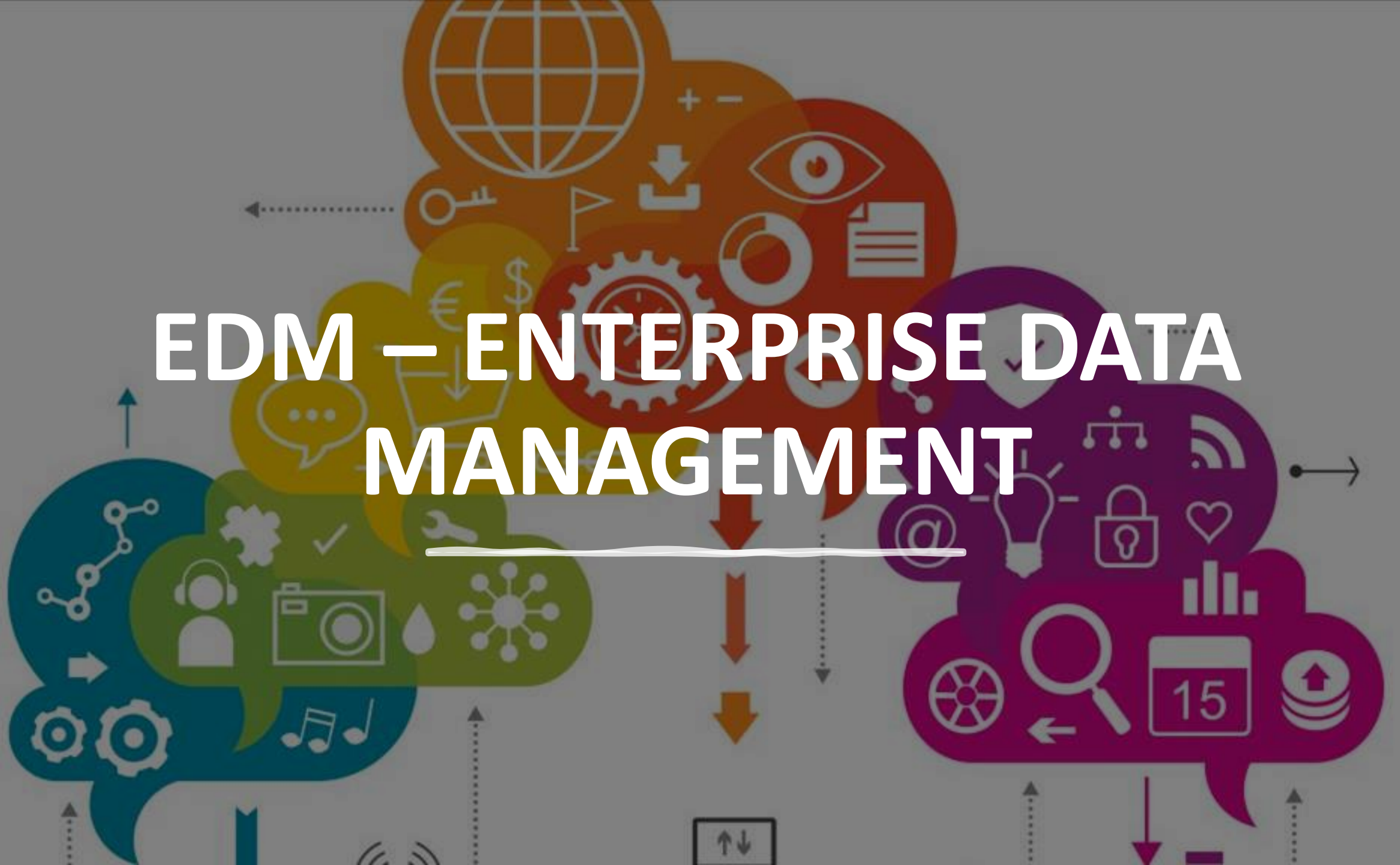
- Qualidade em metadados
- Arquiteturas de integração e distribuição física de banco de dados
- Master Data Management e Data Hub
- **Qualidade de dados**
- Enterprise Data Management
- LGPD

# Próximas aulas

- 11/09 - Revisão de conteúdo
- 18/09 - Checkpoint 6
- 25/09 - Aula para preparação para Challenge Sprint 4
- **02/10 - EDM aula 1**
- 09/10 - EDM aula 2
- 16/10 - Segurança e Auditoria aula 1
- 23/10 - Segurança e Auditoria aula 2
- 30/10 - LGPD aula 1
- 06/11 - LGPD aula 2
- 13/11 - Kick-off Global Solutions
- 20/11 - Feriado
- 27/11 - Global Solutions
- 04/12 - Substitutiva

# EDM – ENTERPRISE DATA MANAGEMENT

---



# EDM – O que é ?

- Gerenciamento dos dados corporativos:
  - Identificação e definição dos elementos de dados
  - Integração dos dados
  - Recuperação simplificada dos dados
- Para atender às demandas por dados
  - Na hora certa
  - Na qualidade necessária
  - Com uma visão do todo

# EDM – Quem é responsável ?

- Tecnologia da Informação
  - Custodiante dos dados
  - Entra com as ferramentas e com os recursos ligados às ferramentas
- Stakeholders de negócio
  - Data stewards/data owners
  - Definem o elemento de dado



# EDM – Por que é importante ?

- Continuidade operacional
- Minimização do impacto operacional nas mudanças nos processos e tecnologia
- Evolução dos negócios
- Melhor direcionamento das decisões internas: empresa data-driven

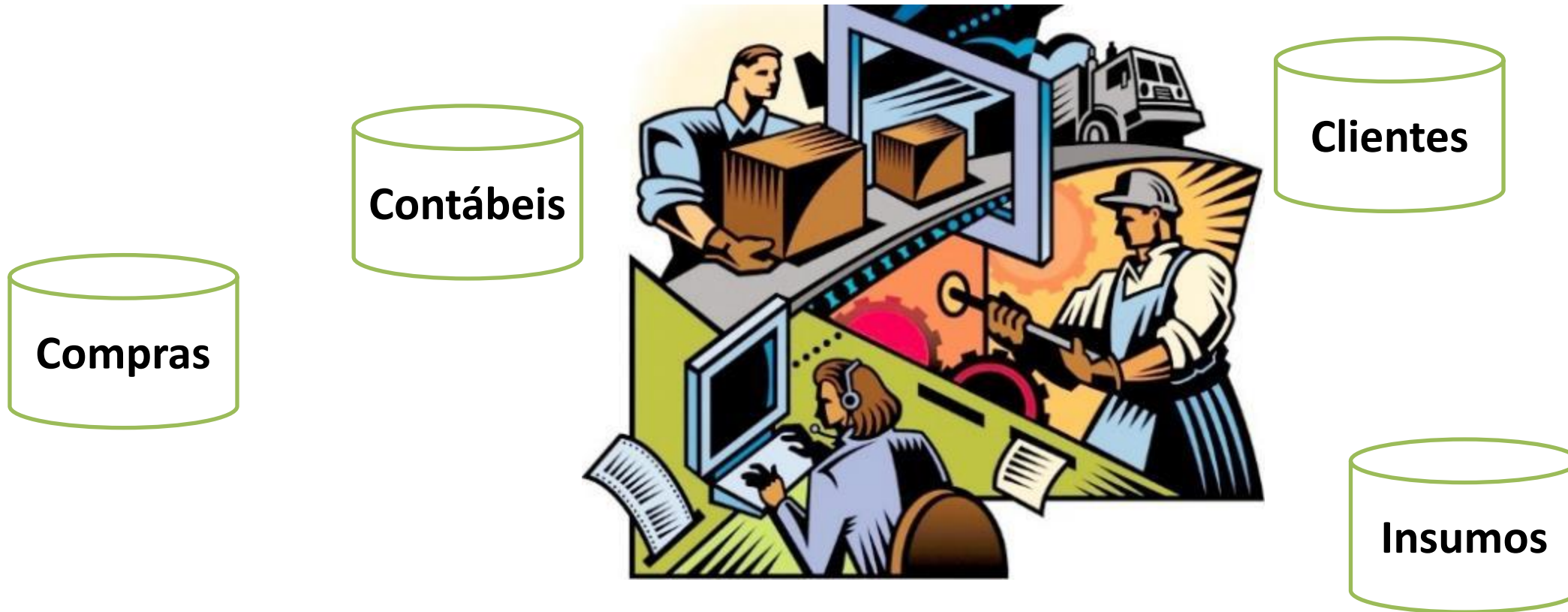


# O PROBLEMA

# Dados no ambiente corporativo



# Dados no ambiente corporativo

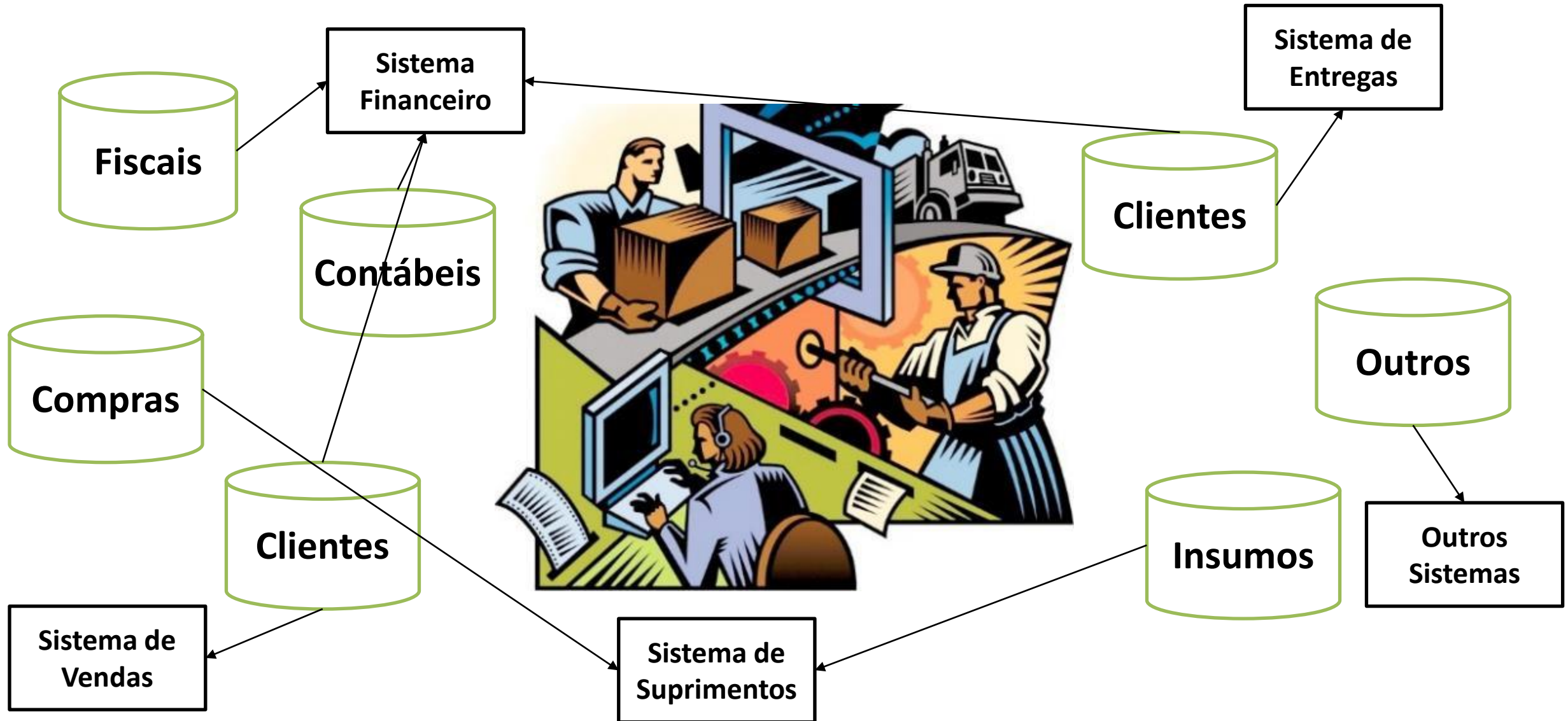


# Dados no ambiente corporativo

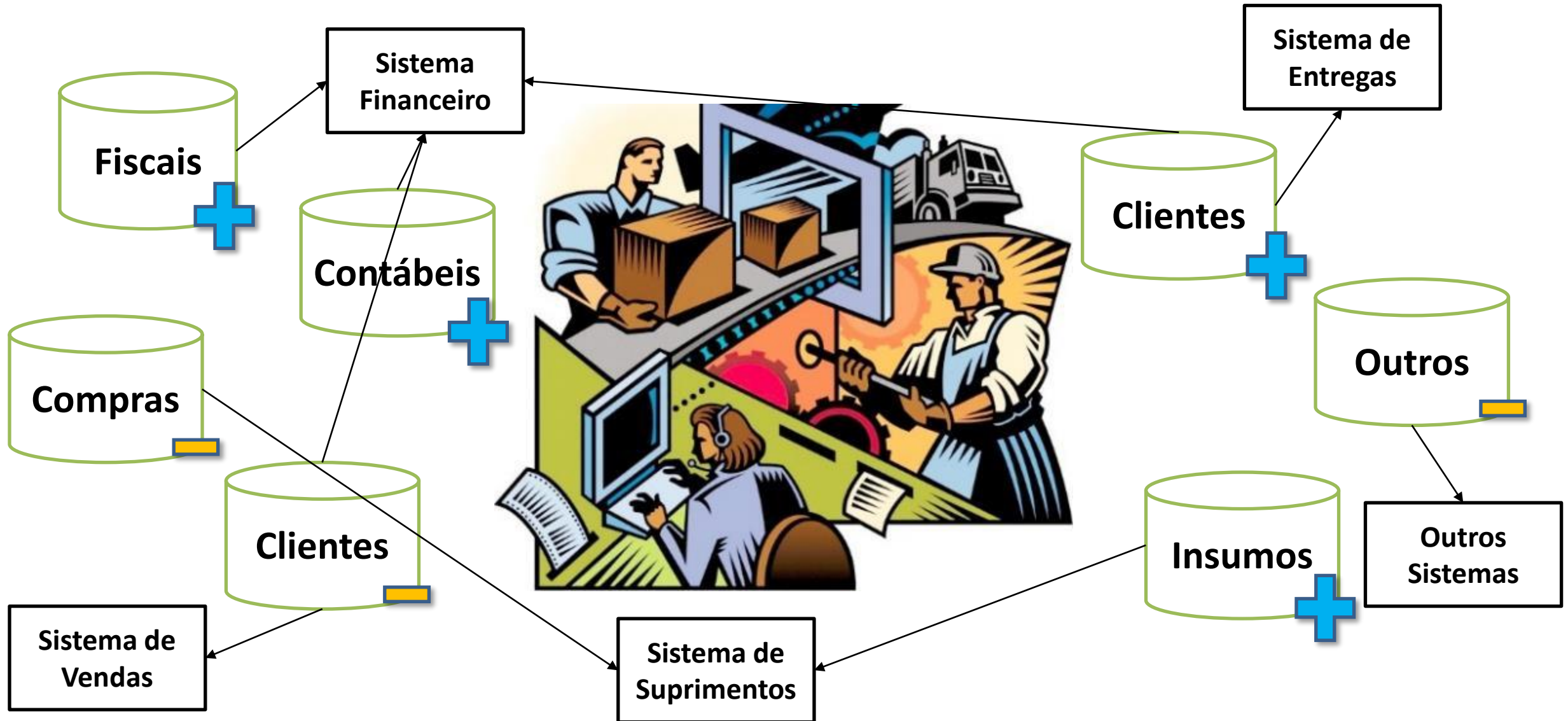




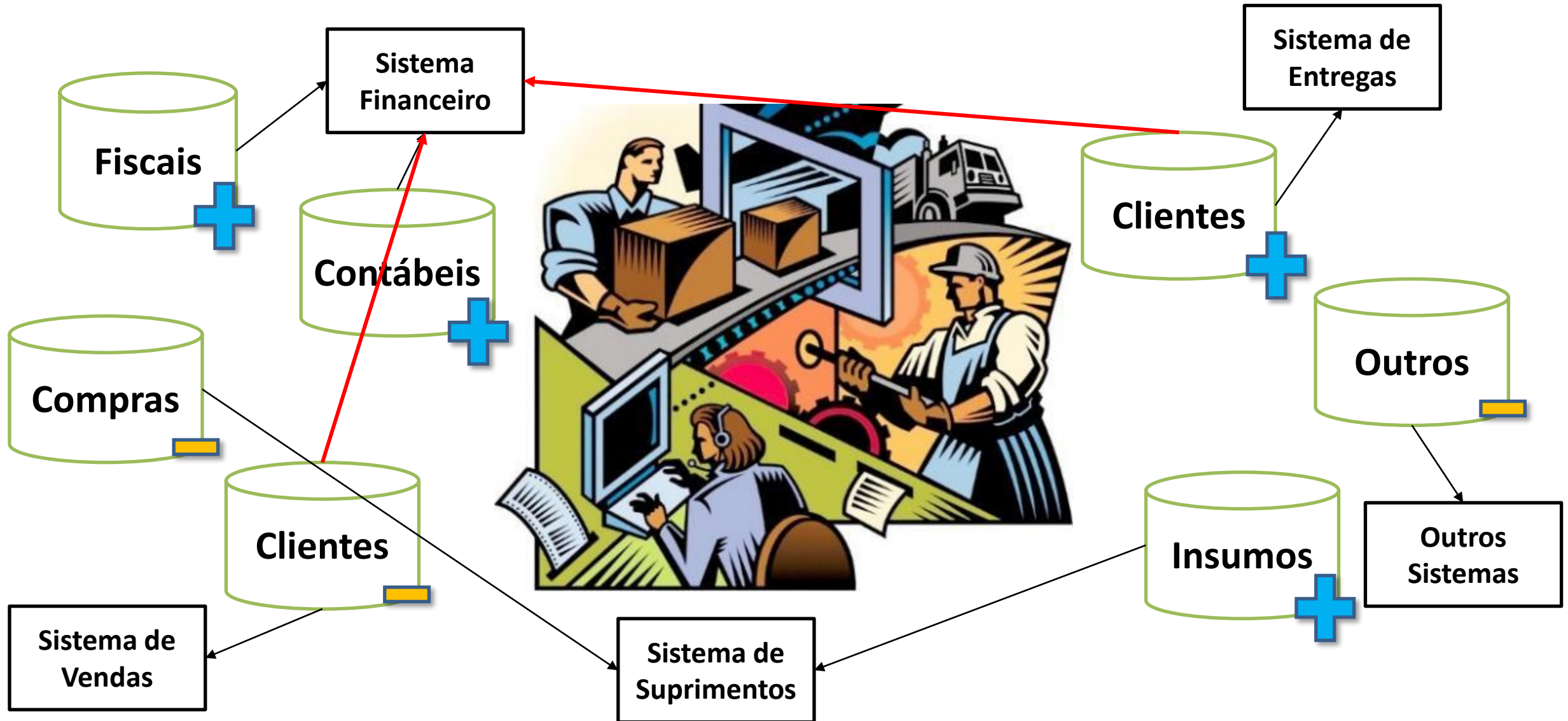
# Dados no ambiente corporativo



# Dados no ambiente corporativo

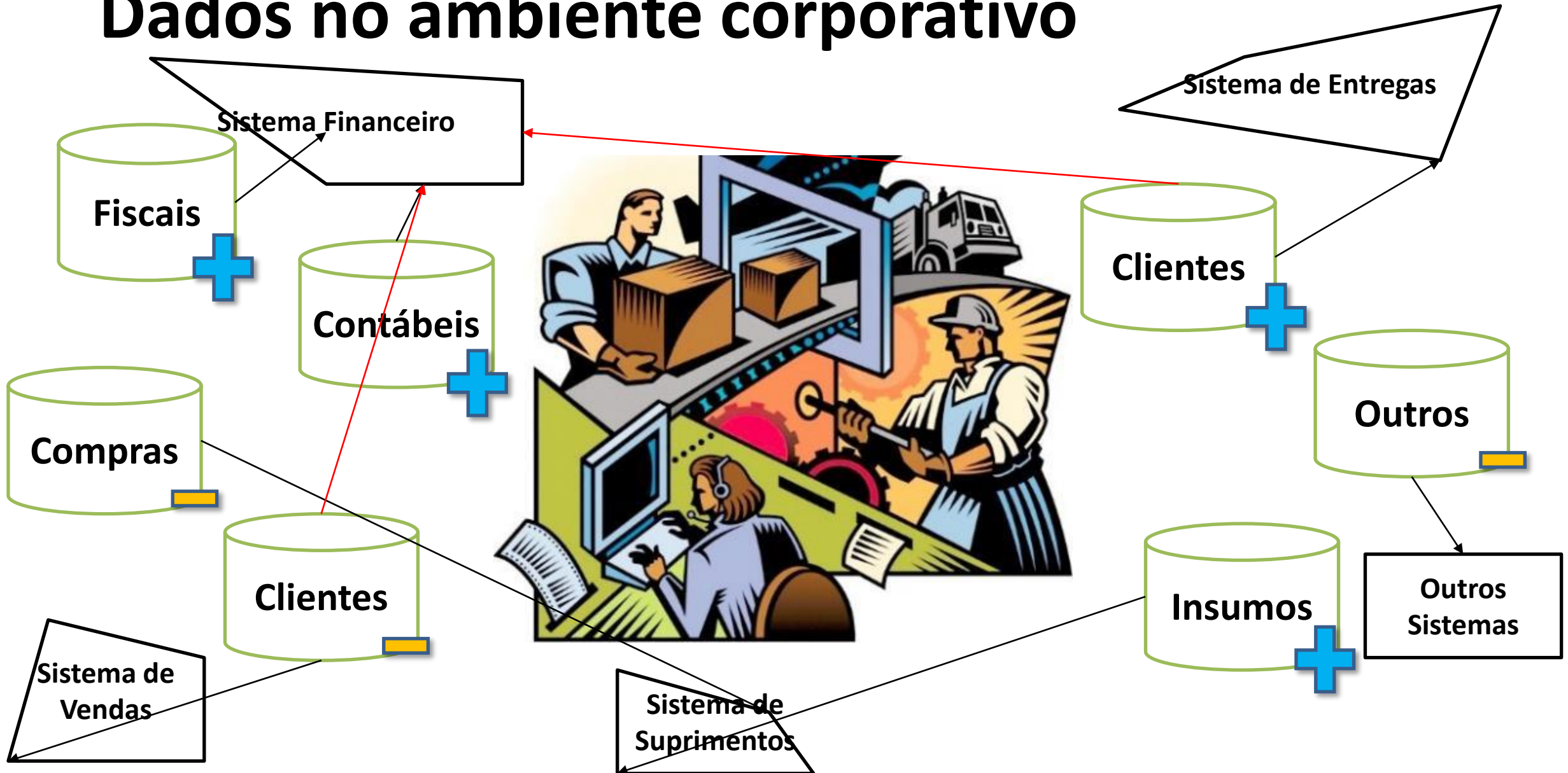


# Dados no ambiente corporativo

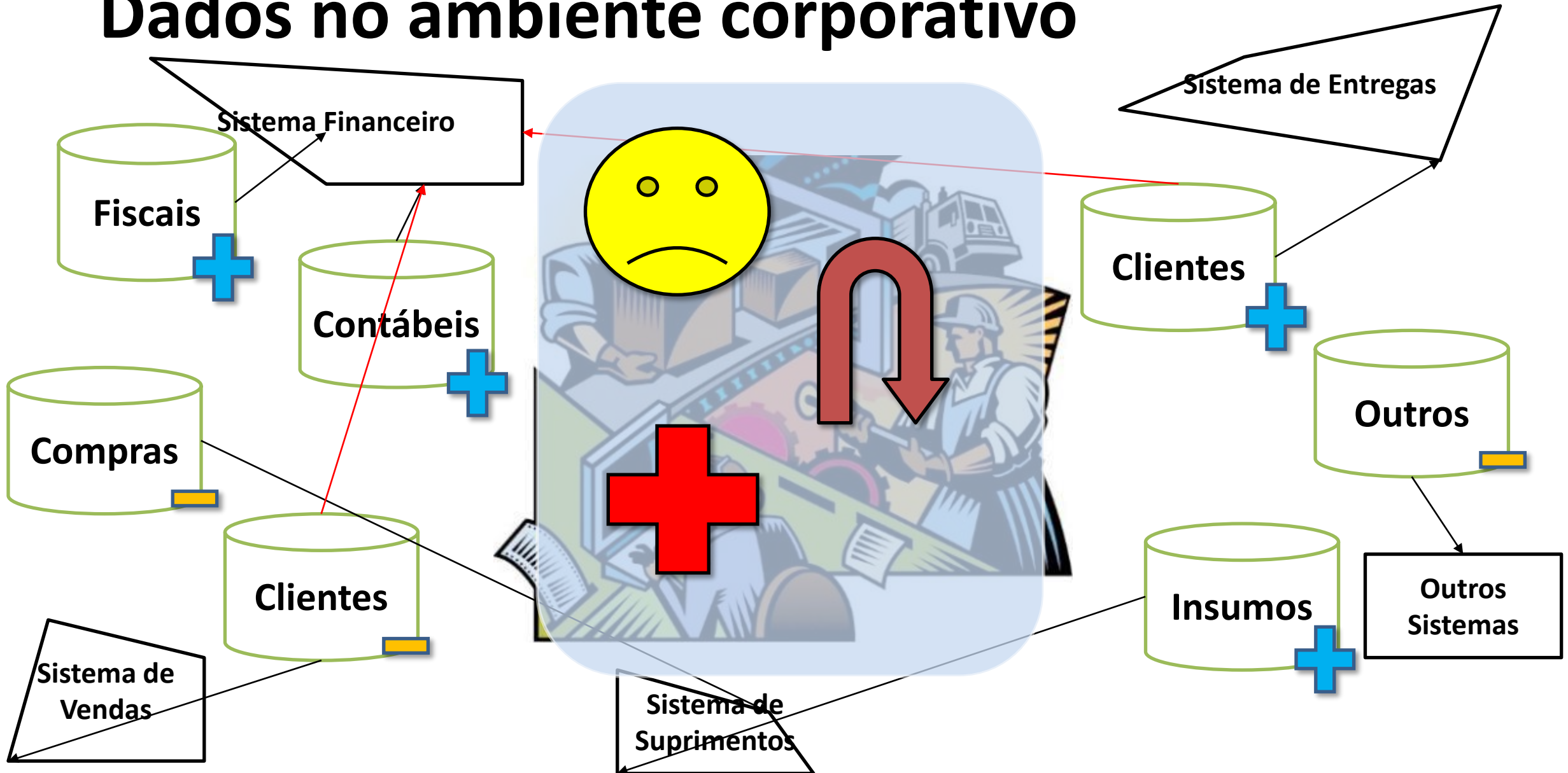




# Dados no ambiente corporativo



# Dados no ambiente corporativo





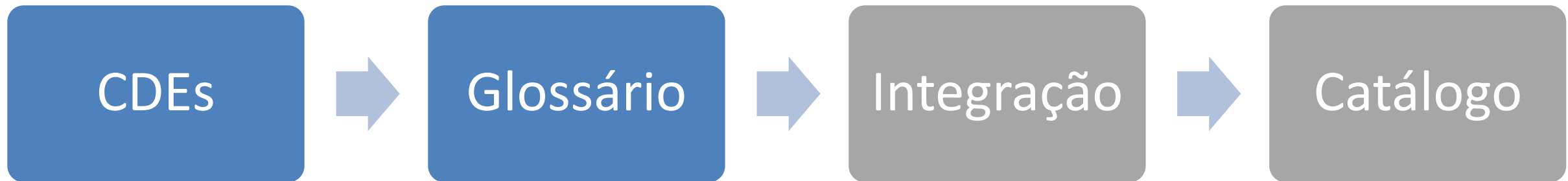
# ABORDAGEM POSSÍVEL

# EDM – Uma abordagem

- As abordagens para lidar com o problema de descentralização dos dados:
  - Concordam que é necessário **identificar, definir e disponibilizar** os elementos de dados corporativos;
  - Oferecem diferentes soluções para unificar os dados descentralizados.

# EDM – Uma abordagem

- A abordagem que seguiremos:





# CDE – CRITICAL DATA ELEMENTS

# Elemento de dados

- Um simples fato usado como base para a argumentação, para a discussão ou para o cálculo;
- Mais básica e elementar unidade de dado que pode ser identificada e descrita em um repositório de dados;
- Dado que não pode ser subdividido mas que pode ser agregado a outras entidades de dados, como um relatório, um grupo ou um registro.
- Exemplos: CPF, logradouro, cidade, CEP

# CDE – Critical Data Elements

- Ou KDE – Key Data Elements
- **Dados que são críticos para o sucesso da organização, em uma área específica de negócio**
- Dados que são necessários para que o trabalho seja feito
- Variam de negócio para negócio
- Sua identificação envolve os data owners e especialistas de negócio



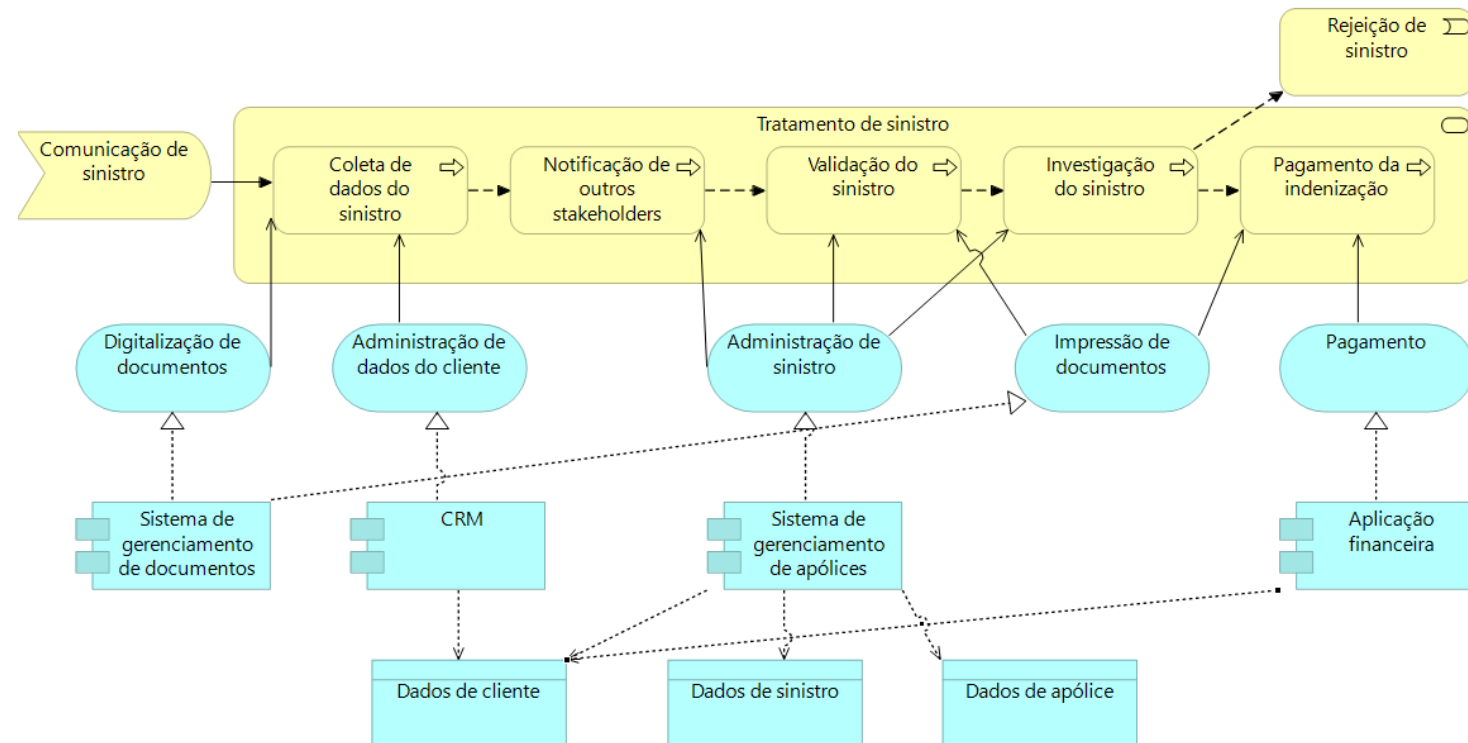
# CDE – Um pouco mais de perto

## CDE:

- São elementos de dados com impacto real nas operações da corporação, em suas decisões e em outras definições de dados
- São dados dos quais a empresa depende para:
  - Sobreviver
  - Crescer
  - Manter sua credibilidade

# CDE – Relação com o TOGAF

- CDE estão diretamente relacionados aos **objetos de dados** identificados na Arquitetura de Aplicação e Dados do TOGAF



# CDE - Exemplo

## Digital Equipment Corporation (DEC) Example

- American international company in the computer industry from the 1960s to the 1990s, VAX systems, VMS operating system...
- The “Order Number” – they were running out of them...
  - It took 2 years and \$2M in 1977 dollars to expand the Order Number from 5-9 characters
- 1983 – 100K Employees – No data standards
- 12 Critical Data Elements – Identifiers for critical corporate Entities
  - Employee (Corporate vs. Local)
  - Customer (Consolidation, and Who owns Who?)
  - Product (xx-xxxxx-xx , BOM structure, would be expanding to 13 characters)
  - Cost Center (Stockroom...)
  - Company Locations (Real Estate)
  - Orders (Customer vs. Manufacturing vs. Shipping)



# GLOSSÁRIO

# Glossário de dados do negócio

- Aplicação utilizada para comunicar e governar **os conceitos e as terminologias de negócio**, as **definições associadas** e os **relacionamentos entre os termos**.

# Glossário - Objetivos

- **Maximizar o entendimento** dos conceitos e das terminologias principais do negócio;
- **Minimizar o uso indevido** de dados devido a um entendimento equivocado de conceitos e termos;
- **Melhorar o alinhamento das áreas** de negócio com recursos tecnológicos oferecidos pela organização;
- **Maximizar a acurácia dos resultados** das pesquisas por conceitos de negócio e pelo conhecimento associado.

# Glossário – Por que é necessário

- Cria clareza e melhora a **comunicação** resolvendo o problema dos usos diversos dos termos;
- Essencial para a **integração dos dados**, permitindo uma melhor **compreensão** dos dados;
- É o ponto de partida para a **linhagem dos dados** e a fundação para a definição de **metadados**.

## Data Glossary

The Data Glossary ([pdf](#) or [excel](#)) documents institutional definitions for common data elements and terms.

This tool provides:

- University wide standardization and maintenance of the definitions of data.
- A common definition of data structures and the data content represented by data

	A	B	C	D	E
1	Category	Subcategory	Term	Definition	Source
2	Academic Program	Academic Program	Academic Program	A cohesive arrangement of college level credit courses and experiences designed to accomplish predetermined objectives leading to the awarding of a degree, diploma, or certificate.	System
3	Academic Program	Academic Program	Academic Program Emphasis	A focused component of an academic program.	System
4	Academic Program	Academic Program	Academic Program Inventory	The official list of academic programs offered by system colleges and universities.	System
5	Academic Program	Academic Program	Classification of Instructional Program (CIP) Code	A six-digit code in the form xx.xxxx that identifies instructional program specialties within educational institutions. A taxonomic coding scheme for secondary and postsecondary instructional programs. It is intended to facilitate the organization, collection, and reporting of program data using classifications that capture the majority of reportable data. The CIP is the accepted federal government statistical standard on instructional program classifications and is used in a variety of education information surveys and databases (IPEDS).	IPEDS
6	Academic Program	Degree	Degree	An award conferred by a college, university, or other postsecondary education institution as official recognition for the successful completion of a program of studies. Degree types include: Associate Degree; Baccalaureate Degree; Certificate; Graduate Certificate; Master's Degree and Doctoral Degree.	IPEDS
7	Academic Program	Degree Type	Associate Degree	An award that normally requires at least 2 but less than 4 years of full-time equivalent college work.	IPEDS



# Glossário – Ferramentas

- As ferramentas para gestão de termos de negócio vão de uma simples planilha a ferramentas integradas com mais atividades de governança de dados.
- <https://dbmstools.com/categories/business-glossary-tools>

# **Exercício – CDE e glossário**

- Identificar e definir os CDE do projeto apresentado no Challenge.

# Na aula que vem.....

