



TOTAL

— TI PARA CONCURSOS

TEORIA

FUNDAMENTOS DE DADOS

**Dado, informação, conhecimento,
inteligência. Dados estruturados,
semiestruturados e não estruturados.
Coleta, tratamento, armazenamento,
integração e recuperação de dados.
Dados abertos**

SUMÁRIO

1.	Dado, informação, conhecimento e inteligência.....	4
1.1	Dado.....	5
1.2	Informação	6
1.3	Conhecimento	7
1.4	Inteligência.....	8
1.5	Pirâmide Informacional	9
2.	Dados estruturados, não estruturados e semiestruturados	13
2.1	Dados estruturados.....	14
2.2	Dados não estruturados.....	15
2.3	Dados semiestruturados	16
2.4	Comparativo entre os tipos de dados	18
3.	Coleta, tratamento, armazenamento, integração e	22
	recuperação de dados.....	22
3.1	Ciclo de Vida dos Dados.....	22
3.1.1	Fases do ciclo de vida.....	22
3.1.2	Objetivos transversais do ciclo de vida	25
3.2	Ciclo de Vida da Informação.....	26

4.	Dados abertos	28
4.1	Conceito de dado aberto	28
4.2	Três leis dos dados abertos.....	29
4.3	Três normas fundamentais para os dados abertos	29
4.4	Oito princípios dos dados abertos.....	30
4.5	Cinco estrelas para os dados abertos.....	31
4.6	Cinco motivos para os dados abertos.....	32
4.7	Síntese das principais referências sobre dados abertos	33
	Glossário de Termos	36
	Referências	38

1. Dado, informação, conhecimento e inteligência

ESSENCIAL DE PROVA

Dado → **Informação** → **Conhecimento** → **Inteligência (DICI)**
representam níveis crescentes de valor e significado.

Dado: registro bruto sem significado.

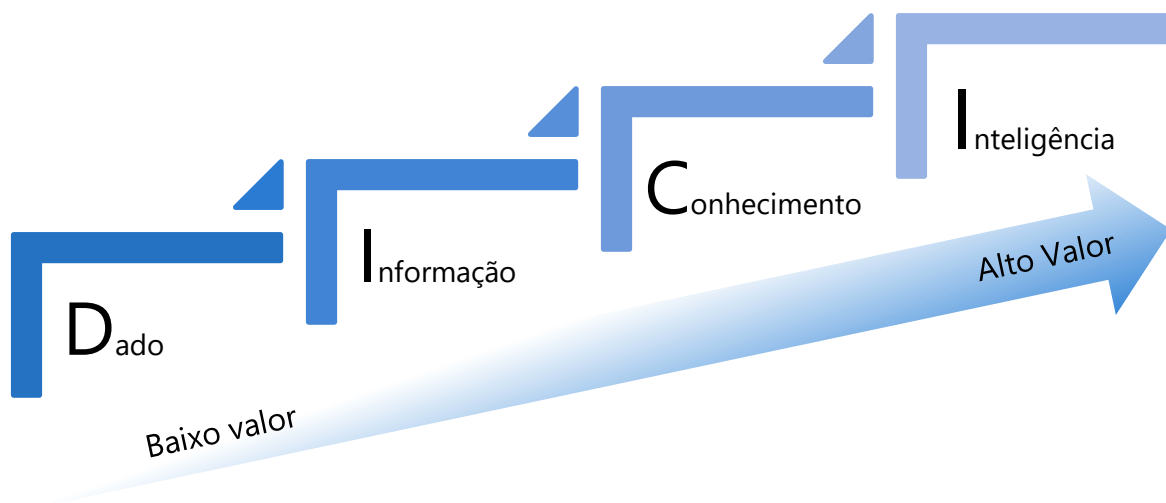
Informação: dado com significado e contexto.

Conhecimento: informação compreendida e internalizada.

Inteligência: conhecimento aplicado em decisão/ação.

Nas organizações, o valor não está nos dados isolados, mas na capacidade de transformá-los em entendimento e decisões.

Essa progressão é conhecida como **modelo DICI**.



1.1 Dado

Dados são **fatos brutos** ou **registros primários** de eventos que ocorrem nas organizações ou no ambiente físico.

Em geral, apresentam as seguintes características:

- **Não organizados**
- **Não processados**
- **Não contextualizados**

Podem ser letras, números, imagens ou símbolos armazenáveis e manipuláveis.

Isoladamente, **não possuem significado claro**.



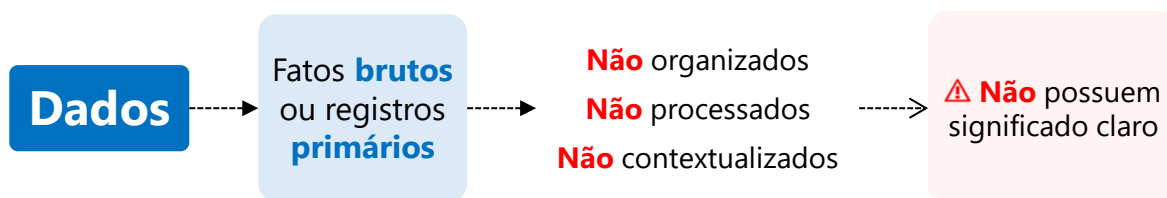
EXEMPLIFICANDO

Exemplos de dados:

- 5
- maio
- xyz
- imagem
- texto sem significado claro



ESQUEMA



Esquema 1 – Conceito de dados.

1.2 Informação

A **informação** é o **dado trabalhado e contextualizado**.

Surge quando dados são organizados, combinados ou interpretados, de modo a produzir compreensão e utilidade para o usuário.

É dotada de:

- **Contexto**
- **Significado**
- **Relevância**
- **Propósito**



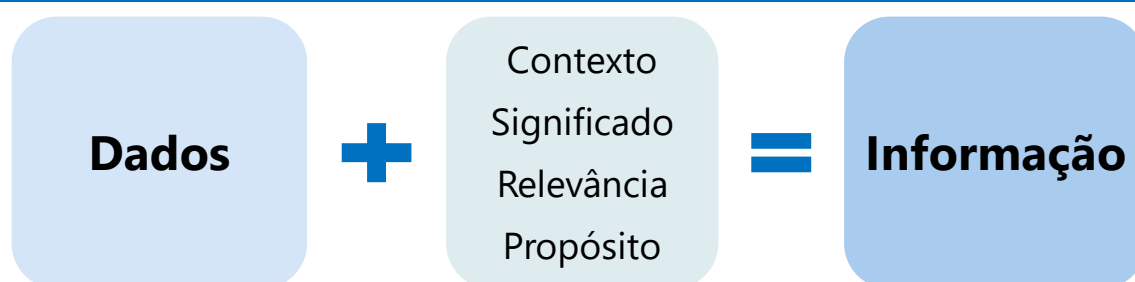
EXEMPLIFICANDO

Exemplos de informações:

- nome do cliente: João Silva
- cor do automóvel: vermelho
- valor do saldo bancário: R\$ 12.350,00



ESQUEMA



Esquema 2 – Conceito de informação.



BIZU

Dados → **sem** significado

Informação → **com** significado

1.3 Conhecimento

Conhecimento é a **informação compreendida e internalizada** utilizada para orientar decisões ou ações.

Surge quando a informação é interpretada à luz da experiência, reflexão e síntese humana.

Caracteriza-se por ser:

- **Resultante da reflexão, experiência e síntese humana.** A informação é interpretada à luz da vivência e do raciocínio humano.
- **Acionável.** O conhecimento permite orientar decisões ou ações.
- **Normalmente tácito.** Grande parte do conhecimento não está registrado formalmente.
- **Difícil de estruturar.** Nem todo conhecimento pode ser facilmente documentado ou armazenado.



EXEMPLIFICANDO

Exemplo de conhecimento:

- “clientes acima de 60 anos compram 30% mais no período da manhã”.

Esse tipo de conclusão resulta da interpretação de informações, permitindo orientar decisões como estratégias de atendimento, marketing ou organização de equipes.



ESQUEMA



Esquema 3 – Conceito de conhecimento.

1.4 Inteligência

A **inteligência** (ou **sabedoria**) é o **conhecimento sintetizado e aplicado** a situações concretas.

Permite julgar e agir de forma eficaz.

É uma capacidade essencialmente humana que se baseia na **experiência** e **intuição**.



ATENÇÃO!

Na **conceituação clássica** da gestão da informação, a **inteligência é essencialmente humana**.

Em contextos como **Inteligência Artificial (IA)**, o termo pode se referir a sistemas que imitam ou simulam comportamentos inteligentes.



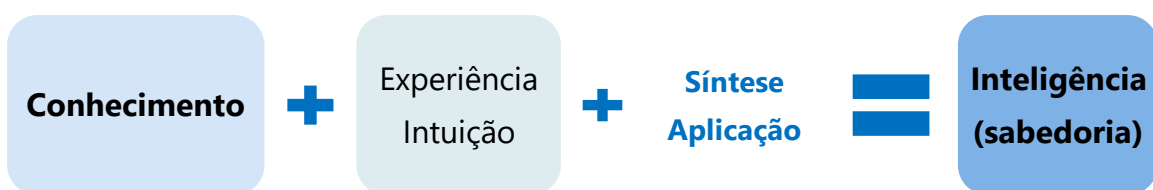
EXEMPLIFICANDO

Exemplo de inteligência:

- Direcionar campanhas do turno da manhã para clientes acima de 60 anos e ajustar o estoque para esse horário.



ESQUEMA

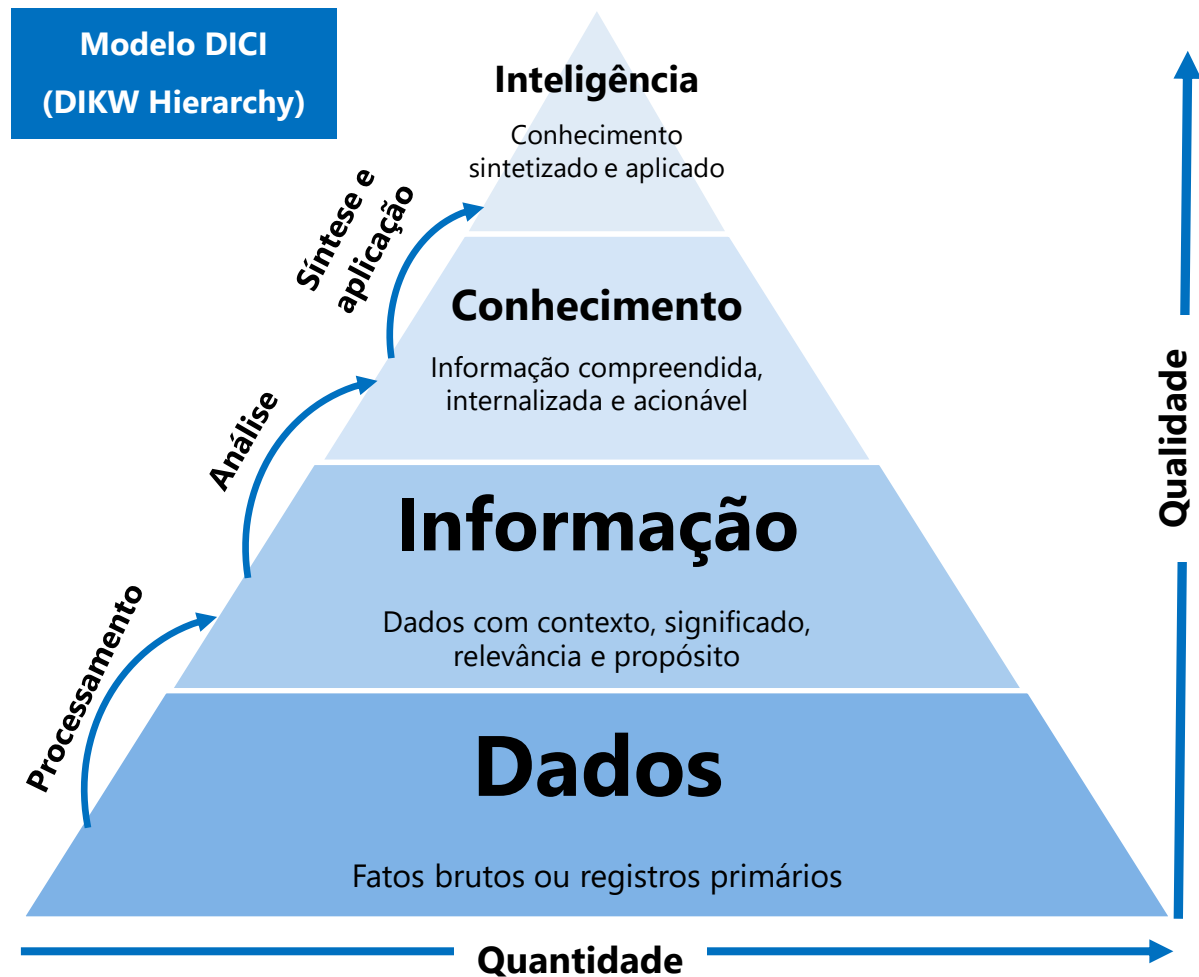


Esquema 4 – Conceito de inteligência.

1.5 Pirâmide Informacional

A relação entre Dados, Informação, Conhecimento e Inteligência pode ser representada pela chamada **pirâmide informacional**.

ESQUEMA



Esquema 5 – Pirâmide Informacional (modelo DICI ou DIKW).

ESCLARECENDO

Em inglês, temos **Data–Information–Knowledge–Wisdom (DIKW)**.

BIZU

Dados → maior **QUANTIDADE**, menor **QUALIDADE**.

Inteligência → maior **QUALIDADE**, menor **QUANTIDADE**.



ATENÇÃO!

Dado, Informação, Conhecimento e Inteligência NÃO são sinônimos.

👉 Em provas, as bancas costumam trocar as definições entre os termos.

Ex.: Dado representam registros com contexto e significado → ❌ ERRADO

Ex.: Informação é um dado desprovido de significado → ❌ ERRADO



QUESTÕES

1- (CESPE / CEBRASPE – 2025 – ANM – Especialista em Recursos Minerais)

Em relação a BI (*business intelligence*) e à modelagem de dados, julgue o item que se segue.

Em BI, a informação é o resultado da organização dos dados, ao passo que o conhecimento é a geração de valor com a informação.

Resolução:

A assertiva está correta. **Informação** corresponde ao dado organizado e contextualizado, enquanto **conhecimento** se refere ao uso da informação para gerar valor ou orientar decisões.

Gabarito: Certo.

2- (CESPE / CEBRASPE – 2025 – FUB – Técnico) Julgue o item a seguir, relativo aos conceitos de administração de dados.

Informação é uma unidade básica de dado que pode ser medida, registrada ou armazenada sem que esteja associada a um significado ou conhecimento, ao passo que dado é uma informação dotada de significado que pode ser usada para a geração de conhecimento ou a tomada de decisões.

Resolução:

A assertiva está incorreta, pois troca os conceitos de dado e informação.

🔗 **Informação** ~~Dado~~ é uma unidade básica de dado que pode ser medida, registrada ou armazenada sem que esteja associada a um significado ou conhecimento, ao passo que ~~dado~~ **informação** é um ~~informação~~ **dado** dotado de significado que pode ser usado para a geração de conhecimento ou a tomada de decisões.

Gabarito: Errado.

3- (FGV - 2024 – TCE-PA – Auditor de Controle Externo) Com relação aos conceitos de dado, informação e conhecimento, avalie as afirmativas a seguir.

I. Os dados são itens elementares, são cadeias de símbolos e não possuem significado.

II. São exemplos de conhecimento: tendência de vendas de um produto A em uma região B; relação entre o aumento ou queda do preço de uma ação X, na bolsa de valores, e a variação do câmbio.

III. As informações correspondem ao dado processado, com significado e um contexto indefinido.

Está correto o que se afirma em

- a) I, apenas.
- b) I e III, apenas.
- c) I e II, apenas.
- d) II, apenas.
- e) I, II e III.

Resolução:

Vamos analisar cada um dos itens:

I. **Correto:**

Dados são elementares e não possuem significado, incluindo símbolos.

II. **Correto:**

Esses exemplos são de conhecimento, pois envolvem a reflexão humana e avaliação contextual.

III. **Incorreto:**

O contexto da informação é definido e não indefinido.

Gabarito: Letra C.

4- (FCC - 2018 - TCE-RS - Auditor Público Externo - Administração Pública ou de Empresas) Os conceitos de dados, informação e conhecimento são de grande importância no contexto de sistemas de informação. Sobre eles, é correto afirmar que

- a) não são necessários os dados para que se obtenha o conhecimento.
- b) a informação é obtida acrescentando-se significado aos dados.
- c) a informação é obtida a partir do conceito de conhecimento.
- d) o processo de tomada de decisão em um sistema de informação tem por base apenas os dados brutos.
- e) os dados consistem do conhecimento analisado sob diferentes pontos de vista.

Resolução:

Vamos analisar cada um dos itens:

- a) **Incorreto:**

Dados são a base da informação e conhecimento. Portanto, não é possível obter conhecimento sem a existência prévia de dados.

- b) **Correto:**

A informação surge quando se acrescenta significado aos dados, isto é, quando os dados são organizados, interpretados ou contextualizados.

- c) **Incorreto:**

O conhecimento é que obtido da informação, não o contrário.

- d) **Incorreto:**

A tomada de decisão baseia-se em informações e conhecimentos, e não apenas em dados brutos

- e) **Incorreto:**

Os dados não derivam do conhecimento. Pelo contrário, o conhecimento é construído a partir da interpretação das informações, que por sua vez derivam dos dados.

Gabarito: Letra B.

2. Dados estruturados, não estruturados e semiestruturados



ESSENCIAL DE PROVA

Quanto a estrutura os dados podem ser:

Dados estruturados: estrutura rígida e previamente definida.

Dados semiestruturados: alguma estrutura, mas flexível.

Dados não estruturados: sem organização pré-definida.

Nas organizações, os dados são produzidos em diferentes formatos e níveis de organização. Alguns seguem modelos rígidos e padronizados, enquanto outros são livres e heterogêneos.

Essa diferença permite classificá-los em **três categorias** principais:



2.1 Dados estruturados

Os **dados estruturados (ou tipados)** são aqueles que possuem uma **estrutura rígida e previamente definida**, chamada de esquema.

O **esquema** determina como os dados devem ser registrados, armazenados e consultados, facilitando sua organização, recuperação e análise.

Além disso, os dados estruturados possuem **restrições de integridade**, que são regras responsáveis por garantir a consistência dos dados armazenados.

ESCLARECENDO

Geralmente, o **esquema** organiza os dados em blocos semânticos (relações) e define as mesmas descrições para dados de um mesmo grupo (atributos).

Os dados estruturados são amplamente utilizados em:

- **Bancos de dados relacionais.**
- **Planilhas eletrônicas** (como Excel).
- **Cubos multidimensionais** (Data Warehouse).

EXEMPLIFICANDO

A tabela a seguir representa o cadastro dos clientes de uma loja.

Código do cliente	Nome do cliente	CPF	Endereço
134512	João Aprovado dos Santos	123.123.123-12	Rua A, nº 123
251218	Maria Concursada Fernandes	456.456.456-45	Rua B, nº 456

Nesse exemplo:

- **Esquema:** tabela previamente definida com quatro atributos.
- **Atributos:** código do cliente, nome do cliente, CPF e endereço.
- **Restrições de integridade:** exigir CPF no formato XXX.XXX.XXX-XX.

2.2 Dados não estruturados

Os **dados não estruturados (não tipados)** são aqueles **sem organização pré-definida**, ou seja, **não possuem estrutura de atributos pré-definida**.

A maioria dos dados corporativos são não estruturados (cerca de 90%).

Os dados não estruturados podem conter dados de diferentes tipos e em qualquer ordem, sem restrições ou regras específicas.

Essa heterogeneidade dos dados dificulta a consulta ou pesquisa.

Os dados não estruturados são amplamente utilizados em:

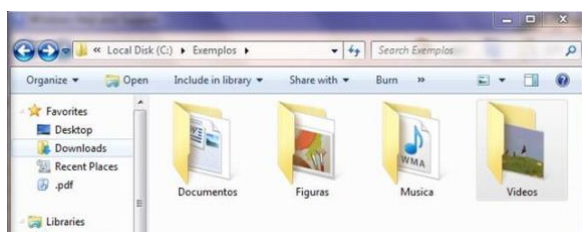
- **Sistemas de arquivos**
- **Bancos de dados NoSQL**



EXEMPLIFICANDO

Exemplos de dados não estruturados:

- Documentos
- Imagens
- Gravações de áudio
- Vídeos



Você pode criar um documento de texto e escrever o que quiser, na ordem que quiser. Pode adicionar imagens, vídeos, formas e um monte de elementos da forma que bem entender. Ou seja, não há uma estrutura para ser respeitada.



ATENÇÃO!

Dados não estruturados podem possuir organização visual (títulos, parágrafos, seções, etc.), mas não possuem estrutura de atributos pré-definida.

2.3 Dados semiestruturados

Os dados estruturados e não estruturados representam os extremos da classificação quanto à estrutura.

Entre eles, encontram-se os **dados semiestruturados**, que apresentam uma **estrutura intermediária e heterogênea**, não sendo **nem completamente não-estruturados e nem estritamente estruturados**.

Nos dados semiestruturados, o esquema não é rigidamente separado dos dados, podendo aparecer de forma explícita ou implícita no próprio arquivo. Por isso, dizemos que esses dados são **auto-descritivos**.



EXEMPLIFICANDO

Exemplos de formatos de dados semiestruturados:

- XML
- JSON
- CSV

Observe um exemplo simplificado de arquivo XML:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<receita nome="pão" tempo_de_preparo="5 minutos" tempo_de_cozimento="1 hora">
  <titulo>Pão simples</titulo>
  <ingredientes>
    <ingrediente quantidade="3" unidade="xícaras">Farinha de Trigo</ingrediente>
    <ingrediente quantidade="7" unidade="gramas">Fermento</ingrediente>
    <ingrediente quantidade="1.5" unidade="xícaras" estado="morna">Água</ingrediente>
    <ingrediente quantidade="1" unidade="colheres de chá">Sal</ingrediente>
  </ingredientes>
  <instrucoes>
    <passo>Misture todos os ingredientes, e dissolva bem.</passo>
    <passo>Cubra com um pano e deixe por uma hora em um local morno.</passo>
    <passo>Misture novamente, coloque numa bandeja e asse num forno.</passo>
  </instrucoes>
</receita>
```

Além dos valores (dados), o arquivo também contém **informações sobre sua própria estrutura**, representadas por tags.

- **Estrutura:** título, ingredientes e instruções.
- **Dados em si:** Pão Simples; Farinha de Trigo, Fermento, Água e Sal; três passos para produção.

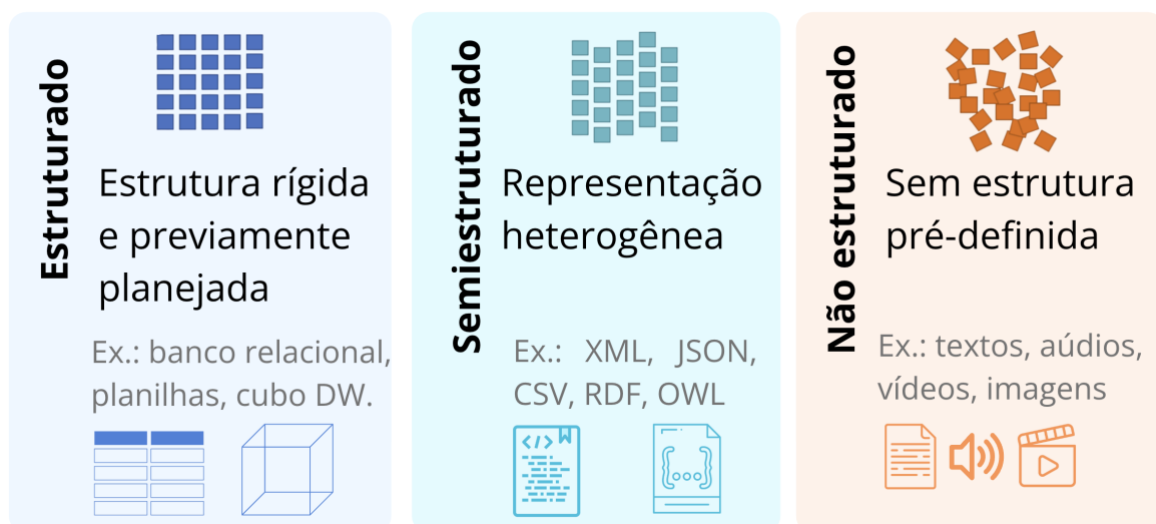
A literatura técnica descreve algumas características adicionais que podem ajudar a compreender melhor os dados semiestruturados.

- **Definição à posteriori:** os esquemas costumam ser definidos após a existência dos dados, a partir da análise de suas estruturas e similaridades.
- **Estrutura irregular:** dados semanticamente semelhantes podem estar organizados de maneiras diferentes, com campos ausentes ou adicionais.
- **Estrutura implícita:** a estrutura pode não estar explicitamente definida, sendo identificada ou extraída por processamento computacional.
- **Estrutura parcial:** apenas parte dos dados possui estrutura, de modo que o esquema pode ser incompleto e nem todas as informações esperadas estão presentes.
- **Estrutura extensa:** a estrutura pode ser ampla e heterogênea, devido à grande variedade de formas de organização dos dados.
- **Estrutura evolucionária:** a estrutura pode mudar com frequência, evoluindo juntamente com os próprios dados.
- **Estrutura descritiva e não prescritiva:** a estrutura tende a descrever os dados existentes, sem impor esquemas rígidos ou fortes restrições de integridade.
- **Distinção entre estrutura e dados não é clara:** como a estrutura está embutida nos próprios dados, nem sempre é fácil separar o que é estrutura e o que é valor.

2.4 Comparativo entre os tipos de dados

O esquema a seguir resume os tipos de dados quanto a estruturação.

ESQUEMA



Esquema 6 - Classificação dos dados quanto à estruturação.

QUESTÕES

- 5- (CESPE / CEBRASPE – 2025 – SEFAZ-SE – Auditor Fiscal Tributário)** Dados estruturados são tipos de dados que
- a) não podem ser indexados, mesmo com organização tabular, devendo ser percorridos integralmente em buscas sequenciais.
 - b) dependem de metadados embutidos em arquivos binários para permitirem qualquer forma de busca ou filtragem.
 - c) podem ser armazenados somente em sistemas de gerenciamento de banco de dados relacionais.
 - d) não podem conter valores nulos, pois a não atribuição de valores descaracterizaria a consistência estrutural exigida.
 - e) têm formato e organização que, definidos por esquemas rígidos que estabelecem tipos e restrições para cada campo, permitem consultas por meio de linguagens como SQL.

Resolução:

Vamos analisar cada um dos itens:

a) **Incorreto:**

Dados estruturados podem ser indexados e consultados de forma eficiente.

b) **Incorreto:**

Dados estruturados não dependem de metadados em arquivos binários; sua estrutura é definida pelo esquema do banco ou da tabela.

c) **Incorreto:**

Dados estruturados são comuns em bancos relacionais, mas também podem existir em planilhas ou outros sistemas estruturados.

d) **Incorreto:**

Dados estruturados podem conter valores nulos, dependendo das restrições definidas no esquema.

e) **Correto:**

Item descreve exatamente a característica central dos dados estruturados: esquema rígido com tipos e restrições, permitindo consultas estruturadas.

Gabarito: Letra E.

6- (CESPE / CEBRASPE - 2024 – TCDF – Auditor de Controle Externo) Julgue o próximo item, relativos a dado, informação, conhecimento e inteligência.

É possível que grandes dados numéricos ou textuais não sejam estruturados em razão da ineficiência em modelá-los como uma tabela.

Resolução:

A assertiva está correta porque mesmo que os dados sejam numéricos ou textuais, eles podem ser considerados não estruturados quando não se encaixam adequadamente em um modelo tabular.

Isso ocorre porque, em muitos casos, o volume, a variedade ou a complexidade dos dados tornam ineficiente ou impraticável representá-los em linhas e colunas — como acontece com grandes volumes de textos, mensagens, imagens ou registros de redes sociais.

Gabarito: Certo.

7- (FCC – 2026 – SCGE PE – Controle Interno) Durante o processo de análise de grandes volumes de informações, um gestor precisa identificar corretamente o tipo de dado com o qual está trabalhando. Em um mesmo projeto, ele utiliza uma planilha contendo:

- I. arquivos JSON com informações de navegação.
- II. postagens em redes sociais contendo textos e imagens.
- III. registros tabulares de clientes.

Nessa situação, os dados são classificados como

- a) I. semiestruturados; II. estruturados; III. não estruturados.
- b) I. estruturados; II. não estruturados; III. semiestruturados.
- c) I. semiestruturados; II. não estruturados; III. estruturados.
- d) I. estruturados; II. semiestruturados; III. não estruturados.
- e) I. não estruturados; II. estruturados; III. semiestruturados.

Resolução:

Vamos classificar cada item:

- I. JSON → possui estrutura flexível e auto descritiva → semiestruturado
- II. Postagens em redes sociais → conteúdo livre (texto, imagem) → não estruturado
- III. Registros tabulares → esquema rígido (linhas/colunas) → estruturado

Gabarito: Letra C.

8- (FGV - 2025 – TCE-PE – Analista de Gestão) Trabalhar com dados não estruturados representa um grande desafio para as empresas, pois esses dados:

- a) são planejados desde a elaboração da estrutura para um propósito estrito.
- b) possuem estrutura rígida e previamente definida.
- c) correspondem a estruturas que não limitam a quantidade de informações coletadas e hierarquizam seus documentos por meio de elementos semânticos, como o código HTML.
- d) não requerem estruturas bem definidas ou padronizadas e podem ser compostos por elementos diversos, comuns ao cotidiano das pessoas.
- e) representam cerca de 80% deste universo, podendo ser fotos, áudios, vídeos, textos, formulários online e tabelas de bancos de dados relacionais.

Resolução:

Vamos analisar cada um dos itens:

a) **Incorreto:**

Item descreve dados estruturados.

b) **Incorreto:**

Item descreve dados estruturados.

c) **Incorreto:**

Item descreve dados semiestruturados.

d) **Correto:**

Item descreve corretamente os dados semiestruturados, que apresentam estruturas flexíveis.

e) **Incorreto:**

Tabelas de bancos relacionais são exemplos de dados estruturados. Os demais são realmente exemplos de dados não estruturados.

Gabarito: Letra D.

3. Coleta, tratamento, armazenamento, integração e recuperação de dados

ESSENCIAL DE PROVA

O **Ciclo de Vida dos Dados (CVD)** descreve as etapas pelas quais os dados passam ao longo de seu uso nos sistemas.

Costuma incluir quatro fases principais:

Coleta → Armazenamento → Recuperação → Descarte

O CVD não se confunde com o **Ciclo de Vida da Informação (CVI)**, que descreve o fluxo organizacional da informação.

Costuma incluir cinco fases principais:

Produção → Manuseio → Armazenamento → Transporte → Descarte

3.1 Ciclo de Vida dos Dados

Os dados em sistemas e processos organizacionais não surgem prontos para uso.

Eles passam por um conjunto de etapas que vão desde a obtenção inicial até sua eventual eliminação, caracterizando o chamado **Ciclo de Vida dos Dados (CVD)**.

3.1.1 Fases do ciclo de vida

O **ciclo de vida dos dados** representa um **fluxo contínuo de utilização dos dados**, desde sua obtenção inicial até sua eliminação.



ATENÇÃO!

Não há padronização única das fases do ciclo de vida dos dados.

Diferentes modelos apresentam etapas e denominações distintas.

O modelo proposto por **Sant'Ana** apresenta quatro fases principais.



A fase de **coleta** envolve a **obtenção inicial dos dados**.

Inclui atividades como:

- Planejamento da coleta
- Identificação das fontes de dados
- Definição do formato dos dados
- Descrição por meio de metadados
- Seleção e validação de registros

A fase **armazenamento** envolve a **persistência dos dados** em repositórios digitais.

Inclui atividades como:

- Processamento
- Transformação
- Inserção
- Modificação
- Migração
- Transmissão



ESCLARECENDO

Tratamento de dados e **processamento de dados** são frequentemente usados como sinônimos. Referem-se a execução sistemática de operações com objetivo de organizar, classificar, transformar e preparar os dados para análise.



ATENÇÃO!

Em alguns modelos, atividades como **processamento ou transformação** aparecem como fases próprias.

A **recuperação** corresponde ao **acesso aos dados** armazenados.

Inclui atividades como:

- Consulta
- Filtragem
- Visualização
- Exploração analítica

ATENÇÃO!

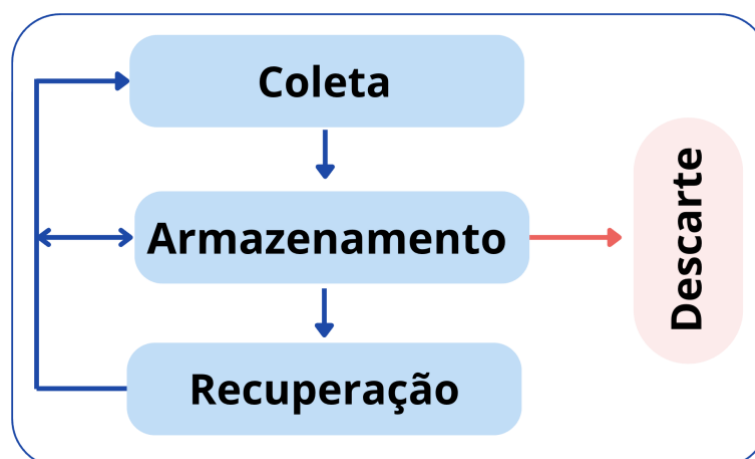
Em contexto técnico, **recuperação de dados (data recovery)** pode significar a restauração de dados perdidos ou corrompidos.

A fase de **descarte** envolve a **eliminação de dados** que não possuem mais utilidade.

Pode ocorrer quando:

- Os dados perderam valor analítico
- Existem restrições legais de retenção
- O volume compromete a eficiência do sistema

ESQUEMA



Esquema 7 – Fases do ciclo de vida dos dados.

3.1.2 Objetivos transversais do ciclo de vida

Além das fases operacionais, o modelo de Sant'Ana identifica um conjunto de **objetivos transversais** que devem **permeiar todas as etapas** do ciclo de vida dos dados.

Entre eles destacam-se:

- **Privacidade**
- **Qualidade**
- **Direitos autorais**
- **★ Integração**
- **Disseminação**
- **Preservação**



DICA

Dentre esses objetivos, a **integração de dados** merece destaque. Além de aparecer em alguns editais, ela fundamenta temas recorrentes em provas, como ETL, Data Warehouse e Business Intelligence.

Por essa razão, a integração será detalhada a seguir.

A **integração** refere-se à **articulação de dados provenientes de diferentes fontes ou sistemas**, permitindo a construção de uma visão unificada das informações.

No ciclo de vida dos dados, a integração pode ocorrer em diferentes fases:

- **Coleta:** identificação de atributos que permitam relacionar registros entre bases (ex.: chaves primárias).
- **Armazenamento:** definição de estruturas e mecanismos que permitam o acesso integrado aos dados.
- **Recuperação:** análise conjunta de dados provenientes de diferentes sistemas.
- **Descarte:** remoção de registros pode afetar relacionamentos entre bases.

3.2 Ciclo de Vida da Informação

O **Ciclo de Vida da Informação (CVI)** descreve as **etapas pelas quais a informação passa dentro das organizações**, desde sua produção até seu descarte.

As fases costumam incluir:

- **Produção:** criação ou geração da informação a partir de dados ou atividades organizacionais.
- **Manuseio:** utilização e manipulação da informação durante atividades operacionais ou analíticas.
- **Armazenamento:** guarda da informação em suportes físicos ou digitais para uso futuro.
- **Transporte:** circulação ou compartilhamento da informação entre sistemas, pessoas ou áreas da organização.
- **Descarte:** eliminação da informação quando ela perde utilidade ou ultrapassa seu prazo de retenção.



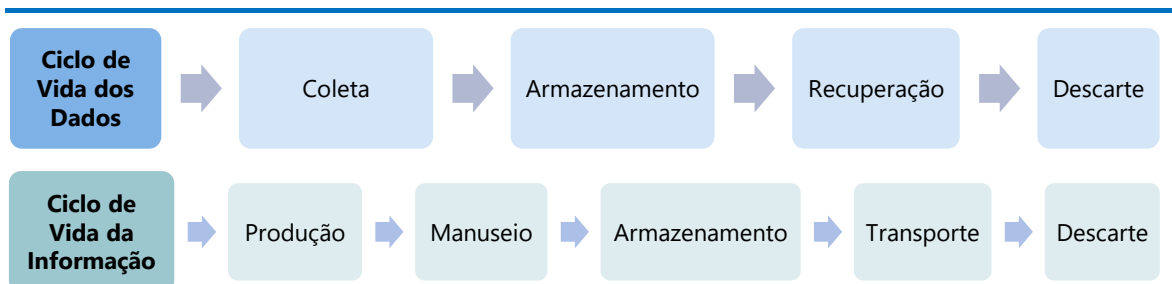
ESCLARECENDO

O **Ciclo de Vida dos Dados (CVD)** trata do fluxo dos registros em sistemas.

O **Ciclo de Vida da Informação (CVI)** enfatiza o fluxo da informação nas organizações.



ESQUEMA



Esquema 8 - Ciclo de vida dos dados x Ciclo de Vida da Informação.



QUESTÕES

9- (CESPE / CEBRASPE - 2024 – ANAC – Especialista em Regulação da Aviação Civil) A respeito de dado, informação, conhecimento e inteligência, julgue o item a seguir.

Na fase de armazenamento do ciclo de vida dos dados, o foco principal é prover meios que ampliem os níveis de utilização desses dados, seja por ampliação das possibilidades de acesso via cópia, por obtenção de conjuntos para análise, ou por meio da disponibilização de recursos de visualização desses dados.

Resolução:

A assertiva está incorreta. A descrição corresponde à fase de recuperação/uso dos dados, não à fase de armazenamento.

↯ Na fase de ~~armazenamento~~ **recuperação** do ciclo de vida dos dados, o foco principal é prover meios que ampliem os níveis de utilização desses dados, seja por ampliação das possibilidades de acesso via cópia, por obtenção de conjuntos para análise, ou por meio da disponibilização de recursos de visualização desses dados.

Gabarito: Errado.

10- (CESPE / CEBRASPE - 2020 - Ministério da Economia - Tecnologia da Informação - Ciência de Dados) A respeito de dados, informação, conhecimento e inteligência, julgue o próximo item.

O ciclo de vida da informação compreende as fases de produção, manuseio, armazenamento, transporte e descarte.

Resolução:

A assertiva está correta. O ciclo de vida da informação envolve etapas relacionadas à criação, uso, armazenamento, movimentação e descarte das informações ao longo de sua existência.



Gabarito: Certo.

4. Dados abertos

ESSENCIAL DE PROVA

Dados abertos são dados acessíveis ao público, em meio digital, em formato aberto, processáveis por máquina, referenciados na internet e disponibilizados sob licença que permita uso, reutilização e redistribuição.

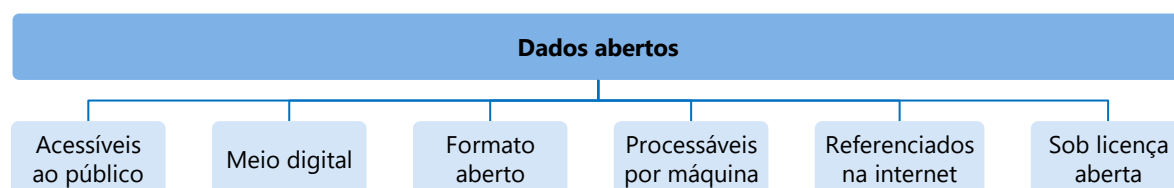
O conceito de dados abertos está associado à disponibilização de dados de forma acessível, reutilizável e sem restrições desnecessárias de uso. Essa abordagem busca ampliar a transparência, o acesso à informação e a reutilização de dados por cidadãos, pesquisadores e organizações.

4.1 Conceito de dado aberto

Dados abertos são dados que possuem as seguintes características:

- **Acessíveis ao público**
- Representados em **meio digital**
- Disponibilizados em **formato aberto**
- **Processáveis por máquina**
- **Referenciados na internet**
- **Disponibilizados sob licença aberta** que permita sua livre utilização, consumo ou cruzamento, limitando-se a creditar a autoria ou a fonte.

ESQUEMA



Esquema 9 – Conceito de dados abertos.



DICA

Em provas, o tema “dados abertos” costuma ser cobrado por meio de **listas e classificações**, como leis, normas, princípios, modelos de maturidade e motivos.

Por isso, é importante reconhecer essas listas, pois muitas questões exploram diretamente seus elementos.

4.2 Três leis dos dados abertos

O ativista de dados abertos David Eaves propôs o que se convencionou chamar de **três leis dos dados abertos**. São elas:

1. Se o dado não pode ser encontrado e indexado na Web, ele não existe;
2. Se não estiver aberto e processável por máquina, ele não pode ser reaproveitado; e
3. Se algum dispositivo legal impedir sua replicação, ele não é útil.

4.3 Três normas fundamentais para os dados abertos

Segundo o Open Group, os dados abertos carregam **três normas fundamentais**:

- **Disponibilidade e acesso:** os dados devem estar disponíveis integralmente e a custo razoável de reprodução, preferencialmente para download pela Internet. Devem também estar em formato conveniente e modificável.
- **Reuso e redistribuição:** os dados devem ser disponibilizados sob termos que permitam sua reutilização e a redistribuição, inclusive em combinação com outros conjuntos de dados.
- **Participação universal:** qualquer pessoa deve poder usar, reutilizar e redistribuir os dados, sem discriminação. Restrições como uso “não comercial” ou “limitado a certos fins” excluem dados do conceito de dados abertos.



ATENÇÃO!

As chamadas “**leis**” e “**normas**” dos dados abertos **não correspondem a leis ou normas jurídicas formais**.

São formulações conceituais propostas por especialistas e organizações para avaliar ou orientar a abertura de dados.

4.4 Oito princípios dos dados abertos

Em 2007, um grupo de especialistas reuniu-se para definir os princípios dos Dados Abertos Governamentais.

Desse encontro resultaram os chamados **oito princípios** dos dados abertos.

1. **Completo.** Todos os dados públicos devem ser disponibilizados, exceto os sujeitos a restrições de privacidade, segurança ou controle de acesso.
2. **Primários.** Publicados na forma coletada na fonte, com o maior nível de granularidade possível, sem agregação ou transformação.
3. **Atuais.** Disponibilizados com a maior rapidez possível para preservar o valor.
4. **Acessíveis.** Disponíveis ao público mais amplo possível e para os mais diversos usos.
5. **Processáveis por máquina.** Estruturados de forma que permitam processamento automatizado.
6. **Acesso não discriminatório.** Disponíveis a todos, sem necessidade de identificação ou registro.
7. **Formatos não proprietários.** Disponibilizados em formatos sobre os quais nenhum ente possua controle exclusivo.
8. **Licenças livres.** Não estão sujeitos a restrições por regulações de direitos autorais, marcas, patentes ou segredo industrial.



ATENÇÃO!

Dados abertos não significa ausência total de restrições.

Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

4.5 Cinco estrelas para os dados abertos

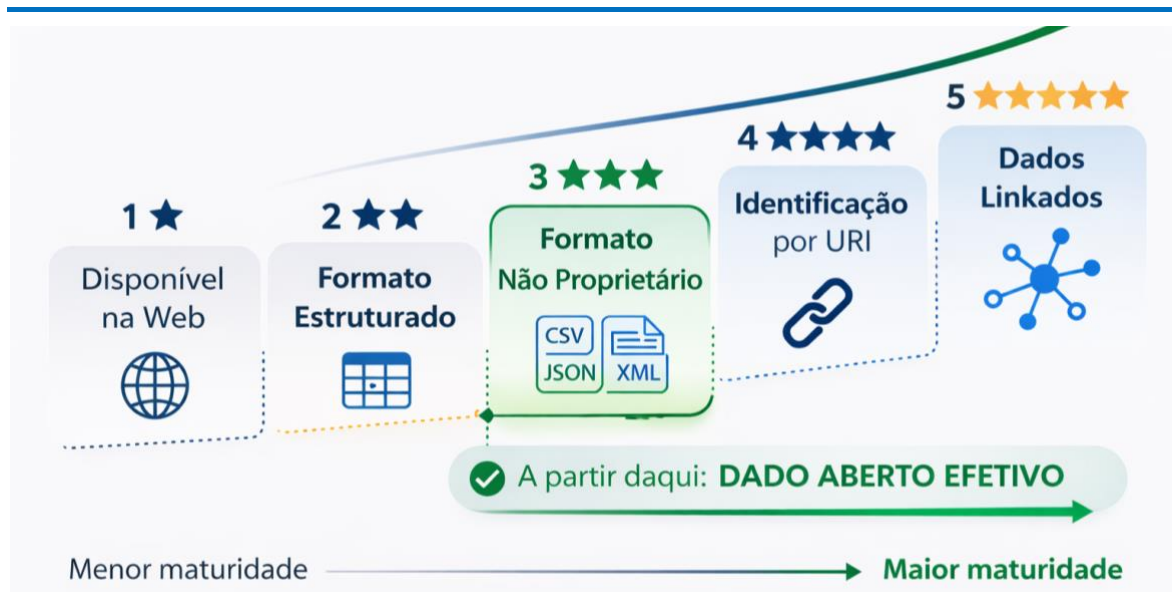
Tim Berners-Lee sugeriu um esquema de implementação das **5 estrelas** para dados abertos. Essas estrelas funcionam como uma medida de intensidade da abertura de dados. São elas:

- ★ Torne seus recursos disponíveis na Web (tanto faz o formato) sob uma licença aberta.
- ★★ Torne seus recursos disponíveis como dados estruturados (ex. excel no lugar de imagem escaneada).
- ★★★ Utilize formatos não-proprietários (ex. CSV e não excel).
- ★★★★ Utilize URIs para identificar recursos. Isso vai ajudar as pessoas a apontarem para eles.
- ★★★★★ Conecte seus dados com dados de outras pessoas para prover contexto (dados linkados).



Dado aberto efetivo → no mínimo, **três estrelas**.

ESQUEMA



Esquema 10 – Modelo de 5 estrelas para maturidade de dados abertos.

4.6 Cinco motivos para os dados abertos

A publicação 5 motivos para a abertura de dados na Administração Pública elaborada pelo Tribunal de Contas da União, apresenta razões para que as organizações públicas invistam em iniciativas de abertura de dados governamentais. Os **cinco motivos para a abertura dos dados** são:

1. **Transparência na gestão pública.** A publicação de dados governamentais em formato aberto permite que a sociedade acompanhe, avalie e fiscalize as ações do governo.
2. **Contribuição da sociedade com serviços inovadores ao cidadão.** Dados públicos podem ser utilizados por cidadãos, empresas e instituições para desenvolver novos conhecimentos, aplicações e serviços.
3. **Aprimoramento na qualidade dos dados governamentais.** A abertura dos dados permite que a sociedade identifique erros e contribua para o aprimoramento das bases governamentais.
4. **Viabilização de novos negócios.** Empresas podem utilizar dados abertos governamentais para desenvolver produtos e serviços comercializados à sociedade.
5. **Obrigatoriedade por lei.** A legislação brasileira, como a Lei de Acesso à Informação (Lei 12.527/2011), incentiva a transparência e a abertura de dados governamentais.

4.7 Síntese das principais referências sobre dados abertos

As principais referências sobre dados abertos são resumidas a seguir:

ESQUEMA



Esquema 11 – Leis, princípios, motivos, normas e estrelas para dados abertos.

ESCLARECENDO

Quando um conjunto de dados não apresenta as características discutidas nas listas anteriores, ele **não pode ser considerado um dado aberto**. Por exemplo:

- Dados não disponíveis na internet;
- Dados disponíveis na internet, mas em formatos proprietários;
- Dados disponíveis apenas em PDF ou imagem, que não são facilmente processáveis por máquina.
- Dados que exigem a identificação do interessado;
- Dados desatualizados;
- Dados com restrições de licença.



QUESTÕES

11- (CESPE / CEBRASPE – 2025 – SEFAZ-SE – Auditor Fiscal Tributário)

Assinale a opção **correta** a respeito de dados abertos.

- a) Os dados abertos são aqueles cujo acesso exige cadastro prévio e aceite de termos de uso restritivos que proíbem a modificação das informações.
- b) Os dados abertos são aqueles que somente podem ser compartilhados com órgãos governamentais mediante acordos de confidencialidade para se evitar mau uso por terceiros.
- c) Dados abertos são disponibilizados de forma gratuita, em formatos abertos e legíveis por máquina, podendo ser livremente utilizados, reutilizados e redistribuídos por qualquer pessoa, desde que citada a fonte.
- d) Dados abertos podem ser publicados em qualquer formato eletrônico, independentemente de estarem protegidos por direitos autorais ou licenças restritivas.
- e) Para que mantenham o *status* de abertos, os dados devem ser obrigatoriamente atualizados em tempo real.

Resolução:

Vamos analisar cada um dos itens:

- a) **Incorreto:**

Dados abertos não exigem restrições que impeçam modificação.

- b) **Incorreto:**

Dados abertos não são limitados a órgãos governamentais.

- c) **Correto:**

Dados abertos são livremente acessíveis, reutilizáveis e redistribuíveis, normalmente em formatos abertos e legíveis por máquina, podendo exigir atribuição da fonte

- d) **Incorreto:**

Dados abertos não podem estar sujeitos a licenças restritivas.

- e) **Incorreto:**

Os dados abertos precisam ser disponibilizados com a maior rapidez possível, mas não necessariamente em tempo real.

Gabarito: Letra C.

12- (CESPE / CEBRASPE - 2024 – ANAC – Especialista em Regulação de Aviação Civil) A respeito de dado, informação, conhecimento e inteligência, julgue o item a seguir.

Dados abertos são disponibilizados sob uma modalidade de licença que permite sua livre utilização, consumo ou cruzamento, por qualquer pessoa, e estão sujeitos, no máximo, ao crédito da autoria ou da fonte.

Resolução:

A assertiva está correta. Dados abertos podem ser livremente utilizados, reutilizados e compartilhados, podendo exigir apenas atribuição da autoria ou da fonte.

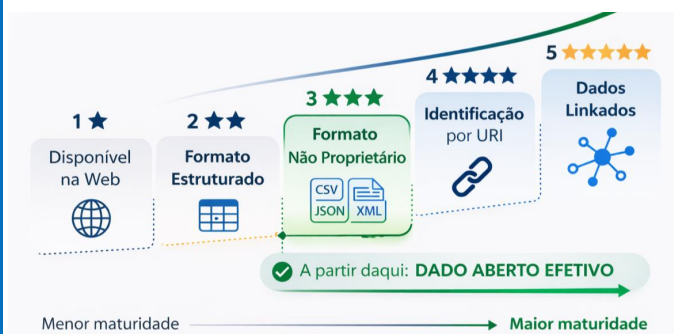
Gabarito: Certo.

13- (FGV - 2024 – TCE-GO – Auditor de Controle Externo) Dados abertos são informações digitais disponibilizadas ao público sem restrições de uso, geralmente sob formatos acessíveis e licenças que permitem sua livre utilização, redistribuição e modificação. O chamado Sistema de 5 Estrelas classifica por meio de estrelas o grau de abertura dos dados. O nível três estrelas é caracterizado por sua disponibilidade na internet

- a) em qualquer formato.
- b) em formato RDF.
- c) de maneira estruturada e formato não proprietário.
- d) em formato SPARQL.
- e) em formato RDF Conectado.

Resolução:

A alternativa correta é a letra C, pois o nível 3 estrelas corresponde a disponibilização em formato não proprietário.



Gabarito: Letra C.

GLOSSÁRIO DE TERMOS

Armazenamento de dados: fase do ciclo de vida dos dados responsável pela persistência dos dados em repositórios digitais, incluindo inserção, organização e manutenção.

Ciclo de Vida da Informação: fluxo organizacional da informação ao longo das fases de produção, manuseio, armazenamento, transporte e descarte.

Ciclo de Vida dos Dados (CVD): conjunto de etapas pelas quais os dados passam desde a coleta até o descarte, incluindo armazenamento e recuperação.

Coleta de dados: fase inicial do ciclo de vida dos dados que envolve obtenção, seleção, descrição e preparação dos dados para uso.

Dados abertos: dados acessíveis ao público, em formato aberto e processável por máquina, disponibilizados sob licença que permite uso, reutilização e redistribuição.

Dados estruturados: dados organizados segundo esquema rígido e previamente definido, normalmente armazenados em tabelas ou bancos relacionais.

Dados não estruturados: dados que não possuem estrutura de atributos pré-definida, como textos livres, imagens, vídeos e documentos.

Dados semiestruturados: dados que possuem alguma estrutura flexível ou auto-descritiva, na qual o esquema está embutido nos próprios dados (ex.: XML, JSON).

Descarte de dados: fase do ciclo de vida dos dados que trata da eliminação de dados que não possuem mais utilidade ou valor.

Formato aberto: formato de dados não proprietário, livre de restrições tecnológicas ou legais, que permite acesso e processamento por qualquer ferramenta (ex.: CSV).

Integração de dados: processo de combinação e articulação de dados provenientes de diferentes fontes ou sistemas, permitindo visão unificada.

Metadados: dados que descrevem outros dados, incluindo estrutura, origem, significado e características.

Processamento de dados: conjunto de operações computacionais aplicadas aos dados para organizá-los, transformá-los ou analisá-los (sinônimo de tratamento de dados).

Recuperação de dados: fase do ciclo de vida dos dados que envolve acesso, consulta e obtenção de dados armazenados.

Recuperação de dados (data recovery): processo técnico de restauração de dados perdidos, corrompidos ou inacessíveis em dispositivos de armazenamento.

Repositório de dados: ambiente ou sistema destinado ao armazenamento e gerenciamento de dados (ex.: banco de dados, data warehouse).

Tratamento de dados: execução sistemática de operações sobre dados para organizá-los, classificá-los ou transformá-los em informação útil (sinônimo de processamento).

REFERÊNCIAS

5STARDATA. 5 ★ DOS DADOS ABERTOS. Disponível em: <<https://5stardata.info/pt-BR/>>. Acesso em: 17. jan. 2021.

BEAL, Adriana. **Gestão estratégica da informação**: como transformar a informação e a tecnologia da informação em fatores de crescimento e de alto desempenho nas organizações. Atlas, 2012.

BIO, S. R. **Sistemas de informação**: um enfoque gerencial. São Paulo: Atlas.

CLARO, Daniela B. **DadosEstruturados x SemiEstruturados x NaoEstruturados**. Disponível em: <<http://homes.dcc.ufba.br/~dclaro/download/mate04-20121/DadosEstruturadosxSemiEstruturadosxNaoEstruturados.pdf>> Acesso em: 15 jan. 2018.

DADOS. O que são dados abertos? Disponível em: <<https://dados.gov.br/pagina/dados-abertos>>. Acesso em: 23 nov. 2021.

D'IGNAZIO, Catherine; BHARGAVA, Rahul. Approaches to building big data literacy. In: **Proceedings of the Bloomberg data for good exchange conference**. 2015. p. 6.

HOSCO. **Recuperação de dados**. Disponível em: <recuperacao-de-dados.eti.br>. Acesso em: 30 set. 2017.

LOPES, Adriano; VILLA, Cláudia Regina. A Gestão da Qualidade como facilitadora da Gestão da Informação e do Conhecimento em ambiente jurídico. **XI Seminário Latino-Iberoamericano de Gestión Tecnológica**. 2005.

MELLO, Ronaldo dos Santos et al. Dados semi-estruturados. **XV Simpósio Brasileiro de Banco de Dados**, 2000.

REZENDE, Denis Alcides. **Planejamento de sistemas de informação e informática**: guia prático para planejar a tecnologia da informação integrada ao planejamento estratégico das organizações. 5ed. São Paulo: Atlas, 2016.

SANT'ANA, Ricardo César Gonçalves. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, v. 21, n. 2, p. 116-142, 2016.

SETZER, Valdemar W. Dado, informação, conhecimento e competência. **DataGramZero Revista de Ciência da Informação**, n. 0, 1999.

TAKEUCHI, Hirotaka; NONAKA, Ikujiro. **Gestão do conhecimento**. Bookman, 2009.

TCU. **5 Motivos para a Abertura de Dados na Administração Pública**. Disponível em: <<https://portal.tcu.gov.br/5-motivos-para-a-abertura-de-dados-na-administracao-publica.htm>>. Acesso em: 23 nov. 2021.

UFPB. **Hardware (Generalidades)**. Disponível em: <<http://www.di.ufpb.br/raimundo/Generalidades/ProcDados.html>> Acesso em: 30 set. 2017.