Estratégia Concursos

Aula 00

Análise de Informações p/ TCU (Auditor de TI) Com videoaulas - 2020

Autor:
Diego Carvalho, Equipe Informática e TI, Thiago Rodrigues Cavalcanti

20 de Janeiro de 2020

Sumário

1 – Análise de Informação 2
1.1 – Dado, Informação, Conhecimento e Inteligência 2
1.1.1 – Dado 3
1.1.2 – Informação 4
1.1.3 – Conhecimento 6
1.1.4 – Inteligência 9
1.2 – Dados Estruturados e Não-Estruturados 11
1.2.1 – Dados Estruturados 11
1.2.2 – Dados Não-Estruturados 12
1.2.3 – Dados Semi-Estruturados 12
1.3 – Dados Abertos 15
1.4 – Coleta, Tratamento, Armazenamento, Integração e Recuperação 18

1 – ANÁLISE DE INFORMAÇÃO

1.1 – DADO, INFORMAÇÃO, CONHECIMENTO E INTELIGÊNCIA

Pessoal, vamos começar falando de um assunto que não cai tanto em prova, mas que é importante para compreender o que veremos mais à frente – estou falando da Hierarquia DIKW. Essa é a sigla em inglês para *Data*, *Information*, *Knowledge* e *Wisdom* – Dados, Informação, Conhecimento e Inteligência (ou Sabedoria). Esses são elementos fundamentais para a comunicação e para a tomada de decisão nas organizações, mas têm significados diferentes.

Essas palavras formam um sistema hierárquico em que um dado para um indivíduo pode ser uma informação e/ou conhecimento para outro. Galera, em nosso dia-a-dia, nós costumamos usar alguns desses termos indiscriminadamente – como se fossem sinônimos. No entanto, no contexto de Sistemas de Informações, esses termos são – na verdade – complementares. Para entender melhor essas diferenças, vamos analisar como são definidos cada um desses conceitos.



1.1.1 – Dado

DADO
Dados são correspondências de um atributo, característica ou propriedade que, sozinho, não tem significado.
Dados são elementos brutos, sem significado, desvinculados da realidade.
Dados são simples observações sobre o estado do mundo.
Dados são um conjunto de fatos objetivos e discretos sobre eventos.
Dados são a menor partícula estruturada que compõe uma informação.

O que podemos concluir? Podemos concluir que dados são fatos ou estatísticas coletadas para servir de referência ou análise. O grande lance quando se pensa em dados é que eles não têm significado ou semântica quando vistos isoladamente – eles são brutos, simples e desvinculados de contexto/realidade. Professor, ainda não entendi! Cara... imagine que sem nenhum contexto, eu te mostro um relatório com uma tabela escrito a palavra abaixo em todas as linhas e colunas.

manga



Você consegue tirar alguma informação desse relatório? Não, você pode até inferir que se trata – por exemplo – de algum relatório sobre frutas. No entanto, você não pode afirmar com certeza! *Por que?* Porque isso pode ser tanto um relatório de um hortifruti quanto um relatório de um camiseteria. *E aí, sem nenhum contexto, tem como saber?* Não, é impossível! *Por que?* Porque isso é apenas um dado: bruto, simples, objetivo e desvinculado da realidade.

Dados são sucessões de fatos brutos, que não foram organizados, processados, relacionados, contextualizados, avaliados ou interpretados, representando apenas partes isoladas de eventos, situações ou ocorrências. Eles se constituem de unidades básicas a partir das quais informações poderão ser elaboradas ou obtidas. Quando os dados são organizados, processados, relacionados, contextualizados, avaliados ou interpretados, ele passa a ser uma informação.

(FADESP – COREN/PA – Administrador) Nos Sistemas de Informação, dados são:

a) informações que possuem significado isoladamente.
b) representações do conhecimento contextualizado.
c) definições do real sentido das coisas.
d) correspondências de um atributo, característica ou propriedade que, sozinho, não tem significado.

Comentários: conforme vimos em aula, dados são correspondências de um atributo, característica ou propriedade que, sozinho, não tem significado (Letra D).

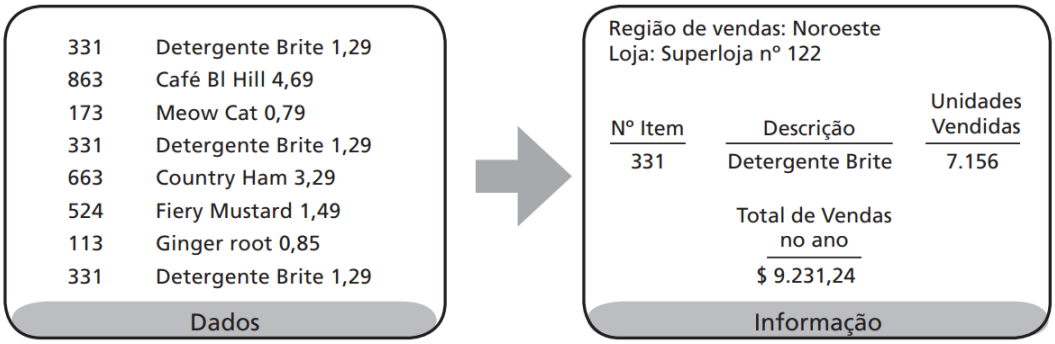
1.1.2 – Informação

INFORMAÇÃO
Conjunto de dados com significado que reduza a incerteza ou que permita o conhecimento a respeito de algo.
Conjunto dos dados presentes em um contexto, carregado de significados e entregue à pessoa adequada.
Conhecimento inscrito (gravado) sob a forma escrita (impressa ou numérica), oral ou audiovisual.
Conjunto de dados contextualizados que visam fornecer uma solução para determinada situação de decisão.
Fatos e/ou dados que encontramos nas publicações, na internet ou mesmo aquilo que as pessoas trocam entre si.
Resultado do processo de acrescentar significado aos dados.
Dados sobre determinado assunto de interesse, que possam ser interpretados ou tenham significado para o receptor.

Agora ficou fácil saber o que é informação: é simplesmente um dado contextualizado. No momento em que um dado é tratado, ele passa a transmitir uma mensagem e, então, temos uma informação. A informação é a ordenação e organização dos dados de forma que passa a transmitir uma mensagem compreensiva dentro de um determinado contexto. Seguindo nosso exemplo anterior, se eu digo: "*Que manga deliciosa!",* nós já conseguimos inferir que se trata da fruta.



Dentro os métodos utilizados para transformar dados em informação, temos: contextualização; categorização; cálculos; correção; e sumarização. Em tempo, as principais características de uma informação são: precisão; completude; confiabilidade; relevância; verificabilidade; acessibilidade; e segurança. *Galera, preciso dizer que a informação é importantíssima e que pode afetar um comportamento, uma decisão ou um resultado?* Vamos ver alguns exercícios...



(CESPE – TCM/BA – Auditor Estadual de Controle Externo) O diretor de uma montadora de veículos necessita tomar uma decisão acerca da continuidade ou não de um dos produtos vendidos no Brasil. Para tanto, solicitou um relatório sobre as vendas de carros da marca do último trimestre de 2018, por faixa de preço, região, modelo e cor. Nessa situação, no contexto de análise da informação, o relatório representa:

a) conhecimento.

b) inteligência.
c) dados.
d) informação.
e) sabedoria.

_____

Comentários: conforme vimos em aula, o relatório em si representa a informação. Trata-se de um conjunto de dados contextualizados – veremos conhecimento e inteligência mais à frente (Letra D).

(IESES – CRA/SC – Agente Administrativo) É um conjunto de dados com determinado significado, ou seja, que reduz a incerteza a respeito de algo ou que permite o conhecimento a respeito de algo:

a) Emissor. b) Dado. c) Comunicação. d) Informação.

_____

Comentários: conforme vimos em aula, trata-se da definição de informação, isto é, dados com significado (Letra D).

(FADESP – COREN/PA – Administrador) Nos Sistemas de Informação, informação:

a) tem o mesmo significado que dados.
b) não tem conexão com geração de conhecimento.
c) é qualquer dado coletado e mantido por uma pessoa ou organização, de forma não contextualizada.
d) é o conjunto dos dados presentes em um contexto, carregado de significados e entregue à pessoa adequada.

_____

Comentários: (a) Errado, informação é dado processado e com significado; (b) Errado, tem conexão forte com a geração de conhecimento; (c) Errado, a informação é dado contextualizado; (d) Correto, trata-se realmente de um conjunto de dados presentes em um contexto, carregado de significados e entregue à pessoa adequada – excelente definição (Letra D).

(FCC – TCE/RS – Auditor Público Externo) 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.

_____

Comentários: (a) Errado, dados e informações são necessários para que se obtenha conhecimento; (b) Correto, a informação realmente é obtida acrescentando-se significado aos dados; (c) Errado, a informação é obtida a partir do conceito de dados; (d) Errado, não são apenas dados brutos, são também a partir de informações; (e) Errado, o conhecimento consiste dos dados e informações analisados sob diferentes pontos de vista (Letra B).

1.1.3 – Conhecimento

Legal, mas como uma informação se torna conhecimento? Essa é uma diferença mais sutil! Basicamente, o conhecimento acontece quando a informação é aplicada. *Vamos mudar um pouco o exemplo?* Vejam só: vocês decidem aprender inglês. Palavras em um dicionário de inglês são apenas dados. Quando essas palavras estão em um livro, por exemplo, elas estão processadas e contextualizadas, logo o livro contém informações.

Agora imaginem que vocês fazem cinco anos de escola de inglês – do nível básico ao avançado. Vocês aprendem várias palavras, aprendem a utilizá-las em frases, conseguem entender textos, conseguem formar frases e conseguem pronunciar tudo perfeitamente. Ora, vocês adquiriram conhecimento de inglês. Se um curso de inglês ensinar vocabulário, morfologia e sintaxe, mas você não aprender a se comunicar em inglês, você terá informação, mas não terá conhecimento.

CONHECIMENTO
Propriedade subjetiva, inerente a quem analisa os dados ou as informações.
Conhecer é o processo de compreender e interiorizar as informações recebidas, possivelmente combinando-as de forma a gerar mais conhecimento.
É o que cada indivíduo constrói como produto do processamento, da interpretação, da compreensão dos dados e fatos – é o significado que atribuímos e representamos em nossas mentes sobre a realidade.
O que cada indivíduo constrói como produto do processamento, da interpretação, da compreensão da informação.

Notem as palavras-chave das definições acima: propriedade **subjetiva**; **interiorizar** informações; construção **individual**. É preciso entender que o conhecimento é o recurso de integração da informação. Isso quer dizer que conhecer algo significa saber onde aquilo se integra como parte de algo maior, pois o conhecimento tem a propriedade de um sistema: compõe um todo; tem sua lógica; estabelece um meio. Segundo Turban, conhecimento é a informação aplicada (em ação).

O conhecimento vai além de informações, pois ele além de ter um significado tem uma aplicação. As informações são valiosas, mas o conhecimento constitui um saber. Produz ideias e experiências que as informações por si só não serão capazes de mostrar. Se informação é dado processado, então conhecimento e informação processada. *Bacana?* Então vamos fazer agora alguns exercícios para saber se vocês adquiriram informações ou conhecimento!

(FADESP – COREN/PA – Administrador) Nos Sistemas de Informação, conhecimento é uma propriedade:

a) subjetiva, inerente a quem analisa os dados ou as informações.
b) decorrente de uma posição objetiva, sem qualquer correlação com dados ou informações.
c) não atrelada ao ser humano, uma vez que não se tem condição de verificar de fato o que é.

d) que não consegue atribuir significado aos dados e informações presentes no contexto.

_____

Comentários: conforme vimos em aula, conhecimento é uma construção individual de quem analisa dados ou informações (Letra A).

(FADESP – COREN/PA – Administrador) Dado, informação e conhecimento são elementos fundamentais para a comunicação e a tomada de decisão nas organizações, mas seus significados não são tão evidentes. Eles formam um sistema hierárquico de difícil delimitação. O que é um dado para um indivíduo pode ser informação e/ou conhecimento para outro. Davenport corrobora esse ponto de vista colocando resistência em fazer essa distinção e a considera nitidamente imprecisa. Com isso podemos afirmar que o grande desafio dos tomadores de decisão:

a) é o de minimizar e transformar as interferências individuais em dados e dados em informação, nesse processo de transformação do conhecimento;

b) é o de transformar conhecimento em informação e informação em dados, minimizando as interferências individuais nesse processo de transformação;

c) é o de transformar dados em informação e informação em conhecimento, minimizando as interferências individuais nesse processo de transformação;

d) é o de transformar informação em dados e conhecimento em informação, minimizando as interferências individuais nesse processo de transformação;

e) na transformação da informação e dos dados é interferir o mínimo nesse processo de conhecimento.

_____

Comentários: conforme vimos em aula, em nossa hierarquia temos a transformação de dados em informação e informação em conhecimento (Letra C).

(CESPE – Polícia Federal – Agente de Polícia Federal) O conceito de conhecimento é mais complexo que o de informação, pois conhecimento pressupõe um processo de compreensão e internalização das informações recebidas, possivelmente combinando-as.

_____

Comentários: galera... que questão linda! Conhecimento é realmente mais complexo que informação porque pressupõe que haja um processo de compreensão e internalização das informações recebidas (Correto).

(FUMARC – PC/MG – Analista da Polícia Civil) Sobre conceitos e definições de dados e informações no contexto de Sistemas de Informação, analise as seguintes afirmativas:

I. Dados se referem a uma descrição elementar de coisas, eventos, atividades e transações que são registrados, classificados e armazenados, mas não são organizados para transmitir significados específicos.

II. Informação se refere a dados que foram organizados para terem significado e valor para o receptor, que interpreta o significado e elabora conclusões sobre a informação.

III. Conhecimento consiste em informações organizadas para transmitir experiência, aprendizagem e prática aplicadas a um problema ou atividade empresarial.

Estão CORRETAS as afirmativas:

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

Comentários: (I) Correto. Dados, ou dados brutos, são as descrições básicas de produtos, clientes, eventos, atividades e transações que são registradas, classificadas, armazenadas. Os dados são a matéria prima a partir da qual as informações são produzidas. Exemplos disso são o número de horas que um funcionário trabalhou em determinada semana ou o número de novos veículos que a Toyota vendeu no primeiro trimestre de 2010; (II) Correto. Informação é um dado que já foi processado, organizado ou colocado em contexto de forma a ter significado para a pessoa que recebe. Por exemplo: o número de vendas trimestrais de novos veículos da Toyota de 2008 até 2010 é uma informação, porque dará a ideia de como os recalls dos veículos de 2009 e 2010 causaram impacto nas vendas; (III) Correto. Conhecimento é um dado e/ou uma informação que foi processado, organizado e colocado em contexto para ter significado, e transmite um conhecimento, experiência, aprendizado acumulado e especialização, já que se aplica a um problema ou atividade atual. Saber como gerir o recall de veículos para minimizar os impactos negativos nas vendas é um exemplo de conhecimento (Letra A).

1.1.4 – Inteligência

Bacana! Já sabemos dado é um fato ou estado não processado; informação é um quando você processa um dado; conhecimento é quando você aplica a informação; e a inteligência (ou sabedoria) é quando você utiliza o seu conhecimento para algum propósito. *Como assim, professor?* Galera, vejam a tabela a seguir - ela resumirá tudo que nós vimos nos tópicos anteriores. Vem comigo...

NOME	SEXO	IDADE	SALÁRIO	NATURALIDADE
Alice	Feminino	18	R$ 8.500,00	DF
Bernardo	Masculino	26	R$ 2.000,00	GO
Carolina	Feminino	30	R$ 9.500,00	TO
Davi	Masculino	22	R$ 1.500,00	RR
Eduarda	Feminino	32	R$ 7.500,00	SP
Felipe	Masculino	24	R$ 3.000,00	RJ
Gabriela	Feminino	28	R$ 6.500,00	PA
Hugo	Masculino	20	R$ 1.000,00	RS

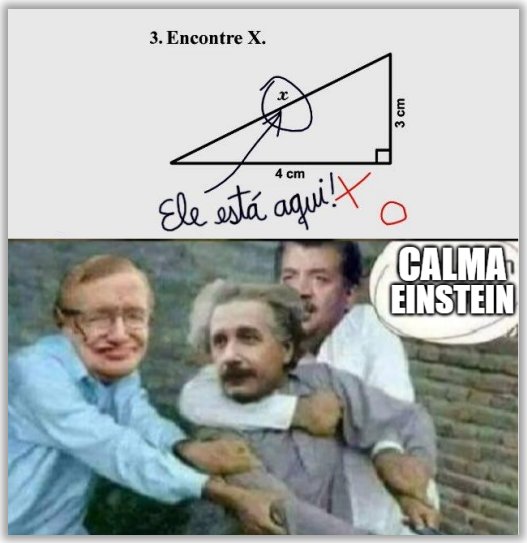
Se vocês entenderam o que nós vimos anteriormente, vocês serão capazes de me responder algumas perguntas. *Se nós olhássemos isoladamente apenas o número 28, nós teríamos um...?* Dado! *Por que?* Porque esse número poderia ser o peso de um animal, o número de gols do Gabigol, as polegadas de um monitor, enfim... sem um contexto, uma análise, um processamento seria impossível identificar do que se trata porque é apenas um dado bruto e simples.

Agora se víssemos os títulos das colunas dessa tabela, poderíamos inferir que 28 é a idade – em anos – de Gabriela. *O que nós acabamos de fazer?* Transformamos dado em informação, isto é, contextualizamos, analisamos, processamos esse dado em uma informação. Em geral, computadores trabalham com dados, quem é capaz de processá-los somos nós – humanos! *Tudo legal até aqui?* Então vamos seguir...

Nesse caso, o que seria um exemplo de conhecimento? O salário médio do sexo feminino é mais que duas vezes maior que o salário médio do sexo masculino. *Galera, isso estava explícito na tabela?* Não, eu precisei cruzar informações diferentes da tabela para chegar a essa conclusão. Logo, eu processei as informações que eu coletei de forma a gerar novos conhecimentos. *Tudo legal?* Agora vamos para a parte final: inteligência.

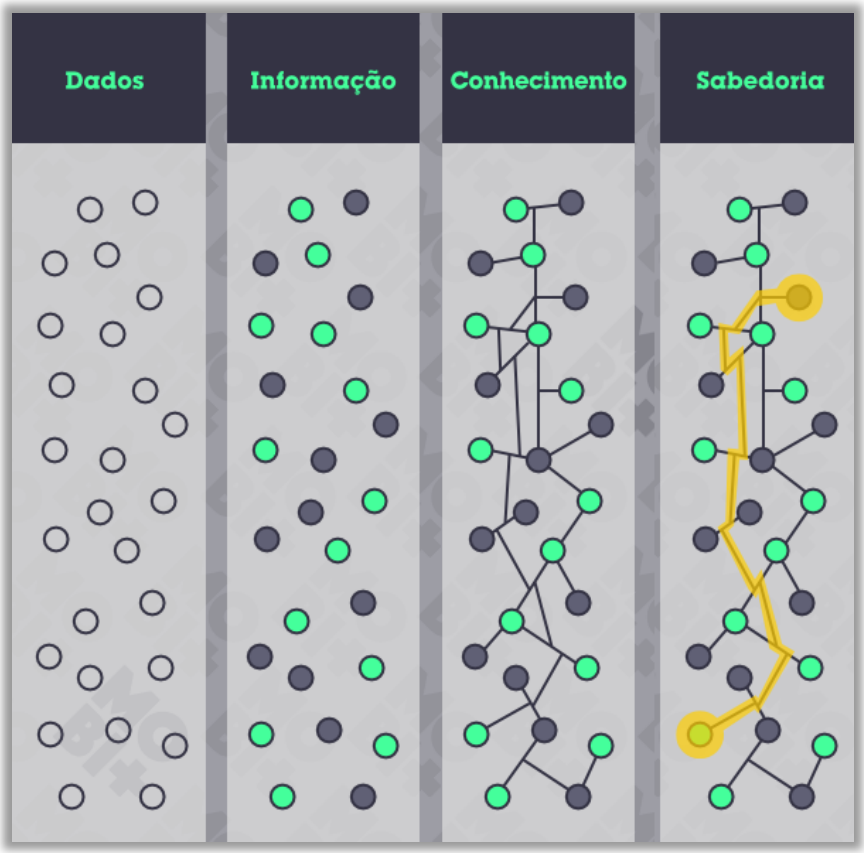
A inteligência tem uma premissa interessante: deve ser utilizada com algum propósito, i.e., para resolver algum problema. *Como assim?* Imagine que uma empresa de pesquisa esteja curiosa para entender porque, contrariando estatísticas nacionais, essa amostra de pessoas apresenta uma desigualdade salarial em favor das mulheres. Ela chama todos os seus pesquisadores – todos com formações e níveis de conhecimento semelhantes – e apresenta esse problema a cada um.

Os pesquisadores isoladamente analisam o problema, mas apenas alguns conseguem resolvê-lo. *Por que isso ocorre uma vez que os pesquisadores têm acesso aos mesmos dados, coletam as mesmas informações e possuem os mesmos conhecimentos estatísticos?* Porque é isso que difere



a inteligência de conhecimento. Cada pesquisador utiliza seus conhecimentos de uma forma particular, utilizando suas conexões cerebrais, memória e experiência para resolver o problema.

Ora... alguns pesquisadores são mais sábios! *Vocês já viram aqueles problemas de lógica que são utilizados para avaliar o QI (Quociente de Inteligência) de um indivíduo?* Dizem que o Einstein tinha uma QI entre 160 e 190! *Por que?* Porque dado um conjunto de dados, informações e conhecimentos, ele conseguia resolver problemas que ninguém antes dele sequer chegou perto de resolver. Einstein era um gênio: sábio e inteligente! *Entendido?*



(CESPE – Polícia Federal – Escrivão de Polícia Federal) O conhecimento é embasado na inteligência das informações que são coletadas e analisadas para uma organização.

Comentários: conforme vimos em aula, a questão inverteu os conceitos – a inteligência é embasada no conhecimento das informações que são coletadas e analisadas para uma organização (Errado).

1.2 – Dados Estruturados e Não-Estruturados

Nós já sabemos o que é um dado e sabemos que computadores adoram manipulá-los. No entanto, é preciso entender que nem todo dado é criado igual. Em outras palavras, os dados gerados por um aplicativo de rede social (Instagram, Twitter, Facebook, etc) é completamente diferente dos dados gerados por sistema de estoque de mercadorias de um supermercado. *Como é, Diego?* É isso mesmo, alguns dados são estruturados, mas a maioria é não-estruturado.

E qual é a diferença entre dados estruturados e não-estruturados? Basicamente, dados estruturados são rigidamente organizados e formatados de uma maneira que seja facilmente pesquisável em bases de dados relacionais. *Professor, você falou grego agora!* Calma, tudo isso será detalhado futuramente, mas – simplificando – bases de dados relacionais nada mais são que tabelas (linhas e colunas) que armazenam dados.

Logo, se eu consigo organizar um conjunto de dados em um formato que possa ser exibido em uma tabela, então esse é um conjunto de dados estruturados. *Certinho?* Já os dados não-estruturados não possuem um formato ou uma organização predefinida, tornando muito mais difícil sua coleta, processamento e análise. Sim, a maneira como os dados são coletados, processados e analisados depende completamente do formato em que eles estão.

1.2.1 – Dados Estruturados

Dados estruturados são frequentemente categorizados como dados quantitativos. No seu trabalho, provavelmente você está acostumado a trabalhar com eles. Tente se lembrar! *Os dados que você trabalha podem ser organizados em linhas e colunas de bases de dados relacionais?* Se sim, você trabalha com dados estruturados! Exemplos: nomes, datas, endereços, números de cartão de crédito, tabelas do Excel, entre outros.

Galera, o bacana dos dados estruturados é que eles podem ser facilmente compreendidos e manipulados por linguagens de máquinas – pode-se manipular dados estruturados com relativa rapidez e essa é uma de suas maiores vantagens: você pode ordenar, agregar, separar, juntar, inserir, atualizar, deletar e consultar dados e tabelas com facilidade. A tabela que vimos no exemplo do tópico anterior é um exemplo de dado estruturado.

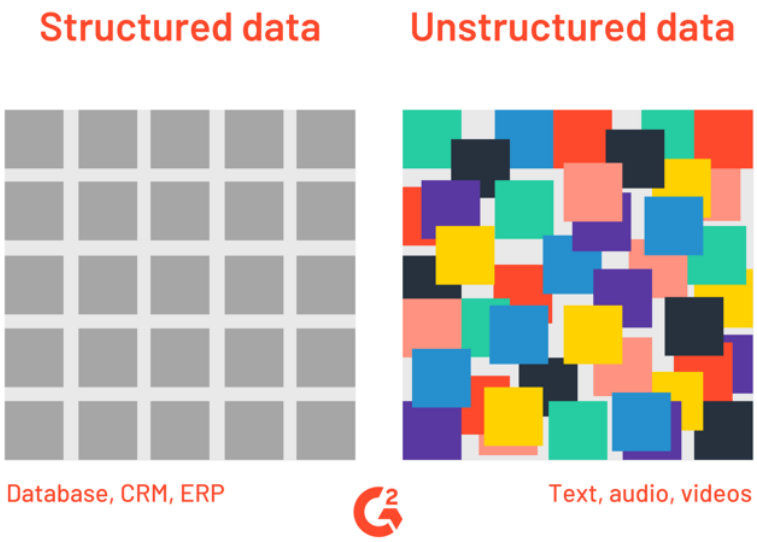
NOME	SEXO	IDADE	SALÁRIO	NATURALIDADE
Alice	Feminino	18	R$ 8.500,00	DF
Bernardo	Masculino	26	R$ 2.000,00	GO
Carolina	Feminino	30	R$ 9.500,00	TO
Davi	Masculino	22	R$ 1.500,00	RR
Eduarda	Feminino	32	R$ 7.500,00	SP
Felipe	Masculino	24	R$ 3.000,00	RJ
Gabriela	Feminino	28	R$ 6.500,00	PA
Hugo	Masculino	20	R$ 1.000,00	RS

Note que, para cada linha dessa tabela, nós temos sempre as mesmas cinco colunas, com os mesmos atributos e os atributos são sempre do mesmo tipo para cada coluna (Ex: Salário é sempre um Número; Sexo é sempre uma Palavra; Naturalidade são sempre duas letras; entre outras). *Viram como é rígido?* E tem mais: eu não consigo inserir um novo atributo na tabela sem antes modificá-la (Ex: CPF) – caso eu queira inserir essa informação, devo adicionar uma coluna!

1.2.2 – Dados Não-Estruturados

Dados não-estruturados são dados que não possuem um formato ou organização predefinida. Assim como os dados estruturados são frequentemente categorizados como dados quantitativos, os dados não-estruturados são frequentemente categorizados como qualitativos, e não podem ser processados e analisados utilizando ferramentas e métodos convencionais. Exemplos: texto, vídeo, áudio, atividades de mídias sociais, entre outros.

Dados não estruturados são difíceis de desconstruir porque não têm um formato ou modelo predefinido, significando que não podem ser organizados em uma base de dados relacional. Mais de oitenta por cento de todos os dados gerados atualmente são considerados não-estruturados, e esse número só tende a continuar crescente com o surgimento da Internet das Coisas – tecnologia que permite que vários objetos estejam conectados à internet.



1.2.3 – Dados Semi-Estruturados

Dados semi-estruturados é uma mistura dos dados estruturados com os dados não-estruturados. Eles não estão de acordo com a estrutura formal dos modelos de dados associados com bancos de dados relacionais ou outras formas de tabelas de dados, mas contêm tags ou outros marcadores para separar elementos semânticos e impor hierarquias de registros e campos dentro dos dados. *Como é, Diego?*

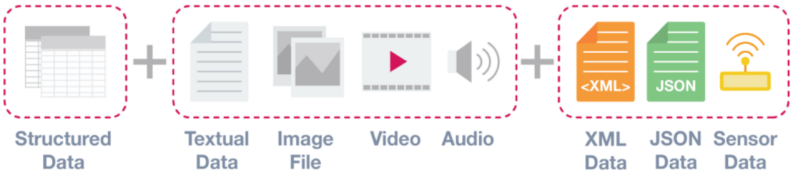
Pensem em um dado que não pode ser perfeitamente encaixado em uma base de dados relacional, mas ele também não é completamente sem formatos. *É difícil imaginá-lo na prática,*

mas vocês já ouviram falar em XML? XML é uma linguagem que permite autodescrever um dado! Em outras palavras, ela apresenta o dado e, junto com ele, apresenta uma autodescrição. Eu sei que está difícil de imaginar, mas eu vou dar um exemplo...

```
<xml version="1.0" encoding="UTF-8" standalone="yes" ?>

<MUSICAS xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
    <MUSICA>
        <NOME> THE SCRIPT FOR MY REQUIEM </NOME>
        <BANDA> BLIND GUARDIAN </BANDA>
        <ALBUM> IMAGINATIONS FROM THE OTHER SIDE </ALBUM>
    </MUSICA>
    <MUSICA>
        <NOME> STARLESS </NOME>
        <BANDA> KING CRIMSON </BANDA>
        <ALBUM> MELTDOWN </ALBUM>
    </MUSICA>
    <MUSICA>
        <NOME> CALADVWCH </NOME>
        <BANDA> DARK AVENGER </BANDA>
        <ALBUM> ALIVE IN THE DARK </ALBUM>
    </MUSICA>
</MUSICAS>
```

Esse é um exemplo de dado escrito em linguagem XML! Notem que os dados em si são os que estão com a cor preta – todo o resto são dados que auxiliam a descrever os dados. Apesar de, nesse exemplo, termos estruturas idênticas para cada música/banda/álbum, nada impede que haja estruturas diferentes para cada um. Por essa razão, não é possível dizer que se trata de dados completamente estruturados ou dados completamente não-estruturados.



(CESPE – TCE/PB – Auditor de Contas Públicas) Com relação a dados estruturados e não estruturados, assinale a opção correta.

a) Dados não estruturados podem ser caracterizados por arquivos de diversos tipos — textos, imagens, vídeos, entre outros —, cujas estruturas não são descritas implicitamente.

b) Por padrão, documentos do tipo XML (eXtensible Markup Language) são estruturados.

c) Dados não estruturados de um mesmo grupo possuem as mesmas descrições e, consequentemente, os mesmos atributos.

d) Por padrão, dados não estruturados são organizados em blocos semânticos.

e) A alta heterogeneidade facilita as consultas aos dados não estruturados, desde que estes estejam ligados por ponteiros.

_____

Comentários: (a) Correto, vimos essa definição insistentemente; (b) Errado, são semi-estruturados; (c) Errado, são os dados estruturados de um mesmo grupo que possuem as mesmas descrições e atributos; (d) Errado, dados estruturados são organizados em blocos semânticos iguais; (e) Errado, ela dificulta as consultas aos dados não-estruturados (Letra A).

(FEPESE – SJC/SC – Agente Penitenciário) Qual software, dentre os listados abaixo, permite a importação (e atualização) de dados estruturados de bancos de dados relacionais para posterior manipulação pelo aplicativo?

a) Word
b) Excel
c) Powerpoint
d) Internet Explorer
e) Outlook

_____

Comentários: dados estruturados são aqueles que podem ser armazenados em... tabelas. Logo, é o MS-Excel (Letra B).

(FAURGS – TJ/RS – Administrador de Banco de Dados) Os dados _____ são armazenados em _____. Os dados _____ misturam nomes de tipos de dados e valores de dados, mas nem todos eles precisam seguir uma estrutura predefinida fixa. Os dados _____ referem-se à informação exibida na Web, caso em que a informação sobre os tipos dos itens de dados não existe.

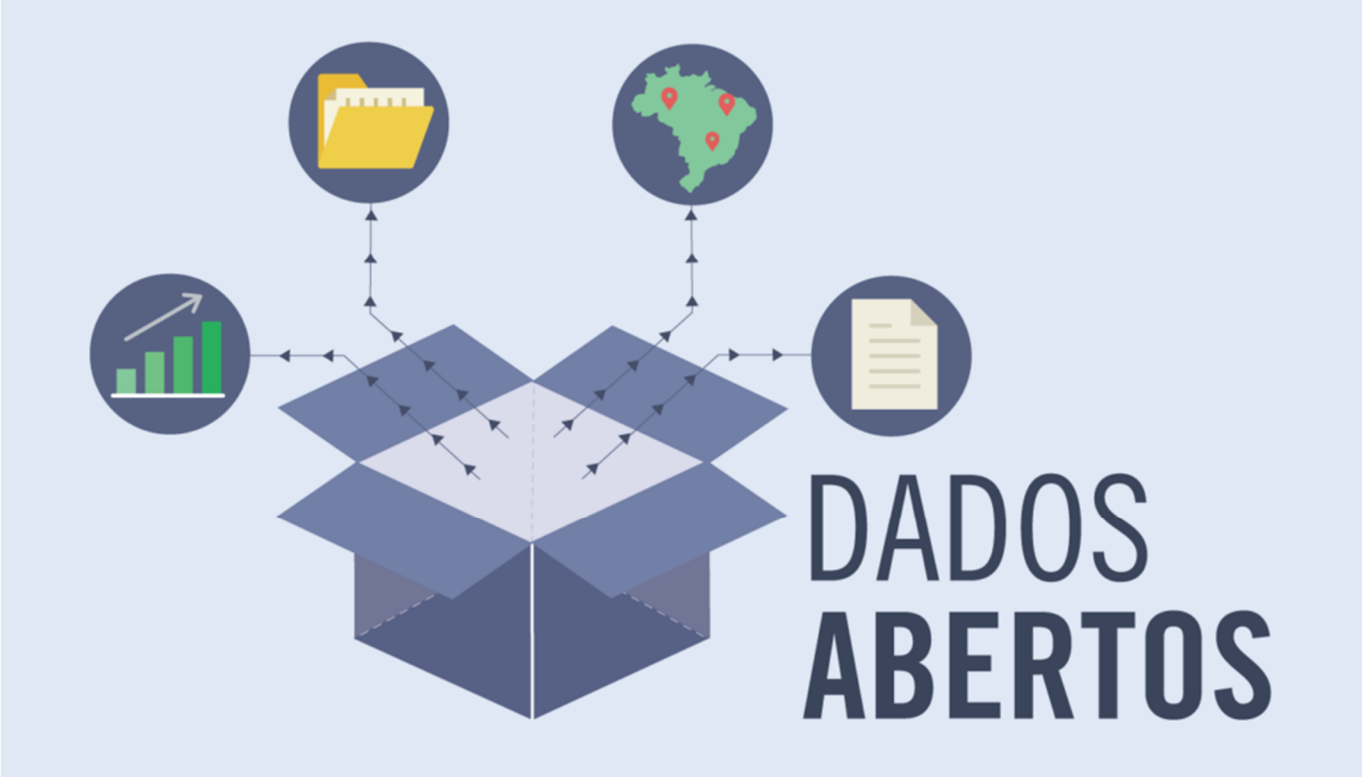
Assinale a alternativa que completa, correta e respectivamente, as lacunas do texto acima.

a) semiestruturados / bancos de dados tradicionais / estruturados / não estruturados
b) não estruturados / XML / semiestruturados /estruturados
c) estruturados / HTML / não estruturados / semiestruturados
d) estruturados / bancos de dados tradicionais / semiestruturados / não estruturados
e) semiestruturados / Web / estruturados / não estruturados

_____

Comentários: os dados **estruturados** são armazenados em **bancos de dados tradicionais**. Os dados **semiestruturados** misturam nomes de tipos de dados e valores de dados, mas nem todos eles precisam seguir uma estrutura predefinida fixa. Os dados **não-estruturados** referem-se à informação exibida na Web, caso em que a informação sobre os tipos dos itens de dados não existe (Letra D).

1.3 – Dados Abertos



Vamos lá... está acabando! Ânimo e altivez! Segundo a Open Knowledge Foundation - OKFn, "dados são abertos quando qualquer pessoa pode livremente usá-los, reutilizá-los e redistribuí-los, estando sujeito a, no máximo, a exigência de creditar a sua autoria e compartilhar pela mesma licença". Quando os dados são produzidos, coletados ou custodiados por autoridades públicas e disponibilizados em formato aberto, diz-se que são dados abertos governamentais.

Ainda segundo a OKFn, dados abertos também são pautados por três leis e oito princípios. Já o Tribunal de Contas da União (TCU) recomenda cinco motivos para utilizá-los.

LEIS DOS DADOS ABERTOS
I - Se o dado não pode ser encontrado e indexado na Web, ele não existe;
II - Se não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado; e
III - Se algum dispositivo legal não permitir sua replicação, ele não é útil.

As chamadas três "leis" dos dados abertos não são leis no sentido literal, promulgadas por algum Estado. São, em suma, um conjunto de testes para avaliar se um dado pode, de fato, ser considerado aberto. As leis foram propostas para os dados abertos governamentais, mas pode-se dizer que elas se aplicam aos dados abertos de forma geral, mesmo fora de ambientes governamentais (empresas privadas, organizações da sociedade civil, etc).

O Banco Mundial, por exemplo, disponibiliza dados abertos. Dados também podem ser abertos voluntariamente por organizações privadas, por diversos motivos. Nos últimos anos, especialistas

têm discutido a abertura de dados pelo setor privado para ações que beneficiam o interesse público, os chamados "colaborativos de dados". Vamos agora conhecer quais são os princípios que norteiam os dados abertos:

PRINCÍPIOS	DESCRIÇÃO
Completos	Todos os dados públicos são disponibilizados. Dados são informações eletronicamente gravadas, incluindo, mas não se limitando a documentos, bancos de dados, transcrições e gravações audiovisuais. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou controle de acesso, reguladas por estatutos.
Primários	Os dados são publicados na forma coletada na fonte, com a mais fina granularidade possível, e não de forma agregada ou transformada.
Atuais	Os dados são disponibilizados o quão rapidamente seja necessário para preservar o seu valor.
Acessíveis	Os dados são disponibilizados para o público mais amplo possível e para os propósitos mais variados possíveis.
Processáveis por máquina	Os dados são razoavelmente estruturados para possibilitar o seu processamento automatizado.
Acesso não discriminatório	Os dados estão disponíveis a todos, sem que seja necessária identificação ou registro.
Formatos não proprietários	Os dados estão disponíveis em um formato sobre o qual nenhum ente tenha controle exclusivo.
Livres de licenças	Os dados não estão sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

MNEMÔNICO							
CONTA	PARA	AQUELAS	ANTES	PRUDENTES	ACEITAREM	FLAMENGO	LIVREMENTE

Em 2007, um grupo de trabalho de 30 pessoas reuniu-se na Califórnia, Estados Unidos da América, para definir os princípios dos Dados Abertos Governamentais. Além disso, o grupo afirmou que a conformidade com esses princípios precisa ser verificável e uma pessoa deve ser designada como contato responsável pelos dados. Por fim, o TCU fez uma publicação com cinco motivos para a abertura de dados na Administração Pública que vocês podem ver a seguir:

MOTIVOS PARA ABERTURA DOS DADOS
Transparência na gestão pública;
Contribuição da sociedade com serviços inovadores ao cidadão;

Aprimoramento na qualidade dos dados governamentais;
Viabilização de novos negócios;
Obrigatoriedade por lei.

Para a administração pública, abrir dados é uma obrigação legal. A Lei de Acesso à Informação – LAI se aplica a os órgãos públicos da administração direta e entes da administração indireta dos Poderes Executivo, Legislativo, incluindo as Cortes de Contas, e Judiciário e do Ministério Público, em todas as esferas (art. 1ª, parágrafo único). No que diz respeito à transparência ativa, a LAI traz consigo conceitos de dados abertos, em especial em seu art. 8º:

Art. 8º É dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas.

> *(...)*
> *§ 2º Para cumprimento do disposto no caput, os órgãos e entidades públicas deverão utilizar todos os meios e instrumentos legítimos de que dispuserem, sendo obrigatória a divulgação em sítios oficiais da rede mundial de computadores (internet).*
>
> *§ 3º Os sítios de que trata o § 2º deverão, na forma de regulamento, atender, entre outros, aos seguintes requisitos:*
>
> *(...)*
> *II – possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações;*
>
> *III – possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina;*

A lei também define as hipóteses de sigilo e de informações pessoais, que são consideradas exceções à regra geral de que os dados devem ser abertos. O objetivo, além da promoção da transparência governamental, é fornecer insumos aos cidadãos e à sociedade civil para que sejam desenvolvidas ferramentas que proporcionem uma percepção mais efetiva das atividades e dos gastos do governo e contribuam para a melhoria da gestão pública, o estímulo ao controle e participação social, a geração de emprego e renda e o fomento à inovação tecnológica.

Eu trabalho na Secretaria do Tesouro Nacional! Caso você queira ter acesso aos dados abertos disponibilizados por esse órgão, poderá acessá-los em:

HTTP://DADOS.GOV.BR/HARVEST/SECRETARIA-DO-TESOURO-NACIONAL-STN

(CESPE – TCE/SC – Auditor Fiscal de Controle Externo) Dados abertos são os dados de livre utilização, reutilização e redistribuição, exigindo-se, no máximo, créditos à autoria e compartilhamento pela mesma licença.

Comentários: conforme vimos em aula, está perfeito – literalmente como nós acabamos de ver (Correto).

1.4 – Coleta, Tratamento, Armazenamento, Integração e Recuperação

Galera, os dados possuem um ciclo de vida! *O que é um ciclo de vida, professor?* Em suma, são as fases pelas quais os dados passam por toda sua vida. Diversos autores publicam suas versões para esse ciclo, nós vamos nos ater à teoria de Ricardo César Gonçalves Sant'Ana. De acordo com ele, o ciclo de vida de um dado é contemplado por quatro fases: coleta, armazenamento, recuperação e descarte – envolvendo todas essas, nós temos a integração. *Qual é a ideia aqui?* Vejamos...

- **Coleta:**

Essa fase busca obter os dados que podem ser utilizados para atender a alguma necessidade específica ou uma demanda prevista de informações sobre um determinado contexto. Definem-se os dados que serão utilizados para algum propósito, identifica-se uma estrutura compatível, e finalmente os dados são coletados.

- **Armazenamento:**

Essa fase busca processar, transformar, inserir, modificar, migrar, transmitir ou qualquer ação que vise persistir os dados em um dispositivo. Persistir – no jargão de computação – é o mesmo que armazenar em algum dispositivo de memória não-volátil. Em outras palavras, é armazenar em alguma base de dados.

- **Recuperação:**

Essa fase busca encontrar, acessar, consultar e interpretar dados armazenados pelos usuários. Uma vez que os dados tenham sido coletados e estejam armazenados pode-se proporcionar uma nova fase que seria aquela em que, tomando como foco os dados, passa-se a tornar estes dados disponíveis para acesso e uso.

- **Descarte:**

Essa fase busca basicamente descartar os dados armazenados. Uma vez concluídas as reflexões sobre as fases de coleta, armazenamento e recuperação, poderia se supor que o ciclo de vida dos dados está completo, principalmente em um momento em que o limite para o volume de dados parece cada vez mais alto, mas não é o que ocorre – os dados ainda podem ser descartados.

Galera, a integração envolve todas essas fases! Na coleta, busca identificar e validar os atributos que serão responsáveis pela identificação unívoca de cada registro; no armazenamento, define a forma de acesso com proteção e interação; na recuperação, faz análises de entidades distintas, mas integradas; e no descarte, foca-se na degeneração à base de dados que um registro excluído poderá causar. *Legal?* Já o tratamento nada mais é que o processamento dos dados...

ESSA LEI TODO MUNDO CONHECE:
PIRATARIA É CRIME.

Mas é sempre bom revisar o porquê e como você pode ser prejudicado com essa prática.

1. 

Professor investe seu tempo para elaborar os cursos e o site os coloca à venda.

2. 

Pirata divulga ilicitamente (grupos de rateio), utilizando-se do anonimato, nomes falsos ou laranjas (geralmente o pirata se anuncia como formador de "grupos solidários" de rateio que não visam lucro).

3. 

Pirata cria alunos fake praticando falsidade ideológica, comprando cursos do site em nome de pessoas aleatórias (usando nome, CPF, endereço e telefone de terceiros sem autorização).

4. 

Pirata compra, muitas vezes, clonando cartões de crédito (por vezes o sistema anti-fraude não consegue identificar o golpe a tempo).

5. 

Pirata fere os Termos de Uso, adultera as aulas e retira a identificação dos arquivos PDF (justamente porque a atividade é ilegal e ele não quer que seus fakes sejam identificados).

6. 

Pirata revende as aulas protegidas por direitos autorais, praticando concorrência desleal e em flagrante desrespeito à Lei de Direitos Autorais (Lei 9.610/98).

7. 

Concurseiro(a) desinformado participa de rateio, achando que nada disso está acontecendo e esperando se tornar servidor público para exigir o cumprimento das leis.

8. 

O professor que elaborou o curso não ganha nada, o site não recebe nada, e a pessoa que praticou todos os ilícitos anteriores (pirata) fica com o lucro.



Deixando de lado esse mar de sujeira, aproveitamos para agradecer a todos que adquirem os cursos honestamente e permitem que o site continue existindo.