

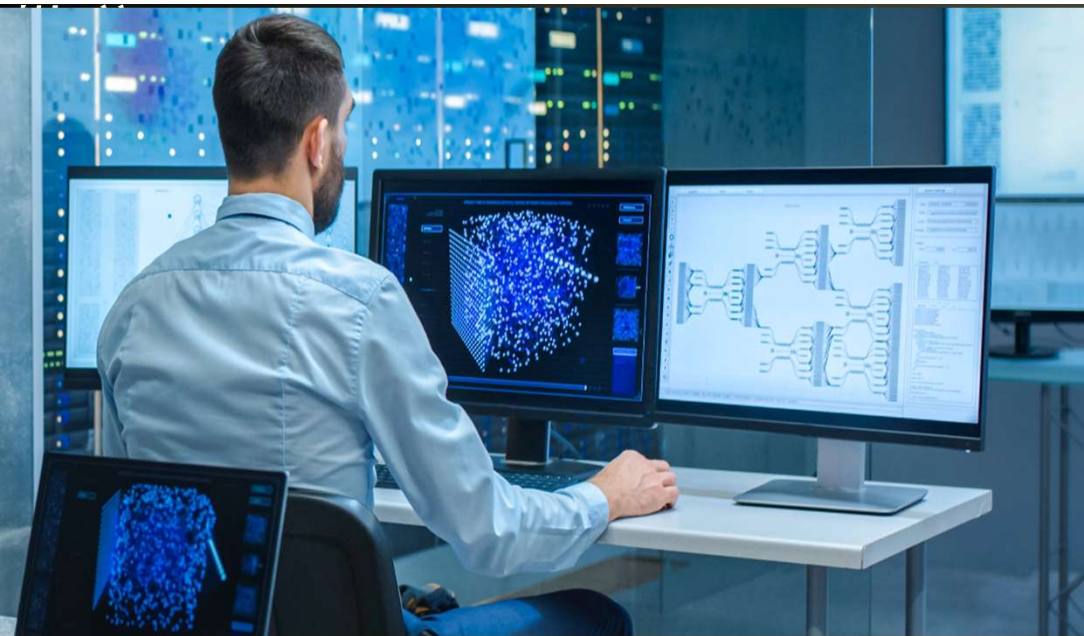
Dado e Informação

Prof. Luiz Camolesi Jr.

Dado

Dado (*Datum* em inglês) é a menor unidade de representação simbólica de um objeto (real ou abstrato), pessoa, fato, característica (propriedade ou estado) entre outras. Neste sentido, a norma **ISO 2382-4** define o Dado como “*a representation of facts, concepts, or instructions in a formalized manner, suitable for communication, interpretation, or processing by humans or by automatic means*”. Esta representação deve permitir a sua comunicação, interpretação ou processamento por seres humanos ou máquinas. Alguns exemplos são os valores: Laranja; Itália; Mercúrio; Tempo; 2000; Vela.

Data (em inglês) deve ser traduzido como **Dados** (plural).



Profissionais

Engenheiro de Dados

Administrador de Dados

Analista de Dados

Gestor de Dados



Princípios para os dados em uma organização:

1. **Dado é um ativo:** dados são recursos valiosos de uma organização e devem ser gerenciados adequadamente;
2. **Dado deve ser compartilhado:** dados devem estar distribuídos entre os indivíduos que os utilizam (usuários);
3. **Dado deve estar acessível:** dados devem estar disponíveis para os indivíduos que os utilizam (usuários) nos momentos em que são necessários;
4. **Dado deve ter um responsável pela sua qualidade:** a responsabilidade pela garantia das diversas dimensões da qualidade de um dado pode ser atribuída a um único responsável ou distribuída entre diversos atores nos processos de utilização;
5. **Dado deve ter descrição textual padronizada:** um padrão de descrição de dados e vocabulário de termos deve ser definido para facilitar o entendimento dos dados pelos seus usuários;
6. **Dado deve estar seguro:** dados devem estar protegidos quando a utilizações não autorizadas ou equivocadas, bem como, quanto a perdas acidentais ou propositais.

Tipo de Dado

Uma característica importante envolvendo a **Engenharia de Dados** envolve os Tipos de Dados que são suportados pela Plataforma Tecnológica utilizada e também pelas Aplicações usadas pelos usuários dos dados. Os Tipos de Dados podem ser entendidos como sendo a forma (ou formato) de estruturação de um dado. Esta forma está também diretamente relacionada com a quantidade de memória (bit e bytes) que é despendida em um equipamento eletrônico/digital para o armazenamento do Dado.

O Tipo de Dados também está relacionado com as operações que podem (ou devem) ser realizadas com o dado, apesar de algumas linguagens de programação flexibilizarem esta regra. Por exemplo, um dado do *Tipo inteiro* pode ser submetido a operações aritméticas como a soma ou a subtração. Um exemplo prático é um Dado do *tipo String* com o *rótulo Número de Telefone*, sobre o qual não há sentido em realizar uma operação aritmética, pois o Dado é um conjunto de caracteres que não representam um valor quantitativo.

Tipo de Dado

Os dados em muitas Aplicações são em grande parte definidos por Tipos de Dados Elementares (Primitivos). O Tipo de Dado Elementar permite definir um Dado cujo valor deve pertencer a um subconjunto a partir de um Conjunto Elementar de valores, como o Conjunto dos Números Inteiros (\mathbb{Z}), o Conjunto dos Números Naturais (\mathbb{N}), o Conjunto dos Números Reais (\mathbb{R}), Conjunto de Caracteres Unicode (unicode.org) e o Conjunto de Valores Binários (0 e 1).

Com a inserção crescente de Sistemas Computacionais no cotidiano pessoal e profissional, houve a necessidade de expandir os Tipos de Dados disponíveis nas Ferramentas de Engenharia e de Gestão para tornar o desenvolvimento mais ágil e confiável. Esta almejada agilidade foi obtida com a definição de Tipos de Dados Complexos (Estruturados ou Compostos), formados com a composição de Tipos de Dados Elementares. Alguns dos Tipos de Dados Complexos normalmente encontrados são: string (sequência de caracteres); date (data; três inteiros); time (horário; três inteiros).

Informação

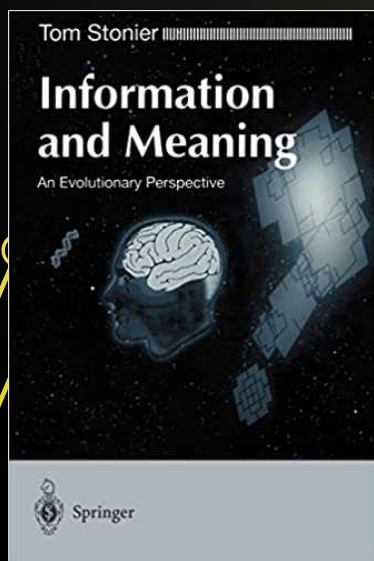
Informação é uma coleção de dados que possuem alguma correlação e que seja de alguma forma útil. Uma informação deve ser colocada em um contexto. A dificuldade em se definir o que é informação está relacionada à capacidade/condições de cada ser humano em compreender e utilizar a informação. Exemplo: “No Engenho Central em Piracicaba iniciará na próxima sexta-feira a tradicional exposição de humor”.

Informação

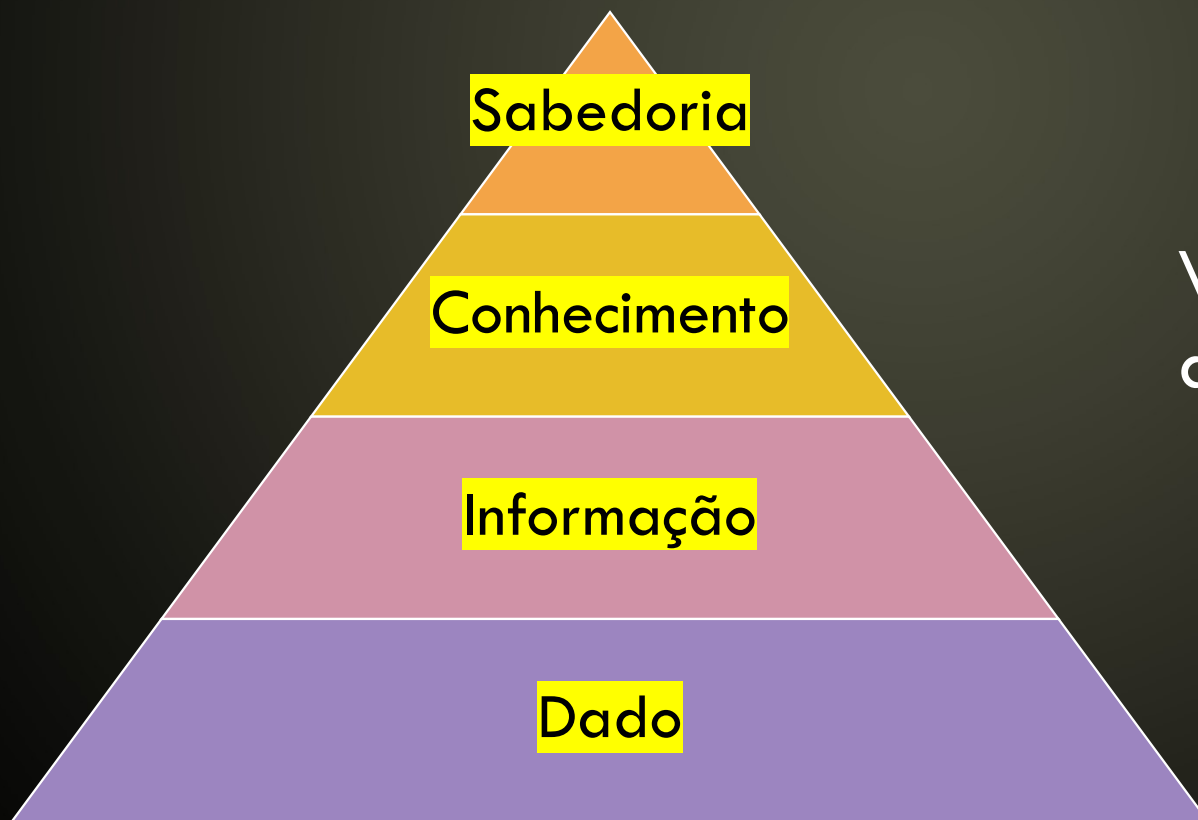
As Leis da Informação foram elaboradas por Daniel Moody e Peter Walsh e apresentadas, na sétima Conferência Europeia de Sistemas em 1999. O objetivo das leis foi ampliar a compreensão sobre a informação enquanto ativo econômico e propor uma metodologia que orientasse o planejamento de estratégias das tecnologias da informação nas organizações. O trabalho desenvolve a ideia de que a informação é elemento fundamental para que as organizações melhorem seu desempenho organizacional. As 7 leis definidas são:

- 1- A informação é (infinitamente) compartilhável
- 2- O valor da informação aumenta com o uso
- 3- A informação é perecível
- 4- O valor da informação aumenta com a precisão
- 5- O valor da informação aumenta quando há combinação de informações
- 6- Mais informação não é necessariamente melhor
- 7- A informação não é esgotável

Informação é energia e quanto mais qualidade ela tiver e mais rápido ela circular entre atores, melhor será o resultado para estes atores em ambiente em que atuam, vivem, convivem e trocam experiências




Tom Stonier (1927-1999) defende que a informação, bem como o seu armazenamento e processamento, possui uma realidade física independente do ser humano. Para Stonier, matéria e energia constituem a estrutura superficial do universo, a qual é percebida pelos nossos sentidos, enquanto que a estrutura interna do universo consistiria **não apenas de matéria e energia, mas, também, de informação**. Enquanto a energia é definida como a capacidade de realizar trabalho, a informação é definida como a capacidade de organizar um sistema.



Vamos entender
as composições

Dado
Dado
Dado
Dado
Dado
Dado
Dado
Dado
Dado
Dado
Dado
Dado

Informação

The image features a dark gray background with decorative yellow circuit-like lines in the corners. These lines consist of straight segments and small circles, resembling a stylized electronic circuit. In the center, the word 'Informação' is written in a large, bold, yellow font. To the left of this central text, there is a vertical list of the word 'Informação' repeated ten times, enclosed within a thin yellow bracket. The list is aligned to the left of the bracket, and the bracket itself is positioned to the left of the central 'Informação' text.

Informação
Informação
Informação
Informação
Informação
Informação
Informação
Informação
Informação
Informação
Informação

Informação

Informação
Informação
Informação
Informação
Informação
Informação
Informação
Informação
Informação
Informação
Informação

+



Conhecimento

Quanta informação o mundo produz ?



<u>Parâmetros</u>	<u>2 elevado a:</u>	<u>Unidades em bytes</u>
KiloByte.....(KB)...	10 bits	1024
MegaByte.....(MB)...	20 bits	1048576
GigaByte.....(GB)....	30 bits	107341824
TeraByte.....(TB)....	40 bits	1099511627776
PetaByte.....(PB)....	50 bits	1125899906852624
ExaByte.....(EB)....	60 bits	1152921504606846976
ZetaByte.....(ZB)....	70 bits	1180591620717411303424
YottaByte.....(YB)....	80 bits	1208925819614629174706176
BrontoByte.....(BB)....	90 bits	1237940039295380274899124224
GeopByte.....(GeB).	100 bits	1267650600228229401496703205376
SaganByte.....(SB)...	110 bits.....é o mesmo que multiplicar 1 “GeB” por 1024 Bytes	
JotaByte.....(JB)...	120 bits.....é o mesmo que multiplicar 1 “SB” por 1024 Bytes	
JotaKByte.....(JkB).	130 bits.....é o mesmo que multiplicar 1 “JB” por 1024 Bytes	

1 How much data is generated every minute?

Source: Domo



41,666,667

messages shared
by WhatsApp users



1,388,889

video / voice calls made
by people worldwide



404,444

hours of video streamed
by Netflix users



347,222

stories posted by Instagram users



150,000

messages shared by Facebook users



147,000

photos shared by Facebook users

2 Estimated Data Consumption from 2021 to 2024

Source: IDC / Statista



3 Data Growth in 2021

Sources: TechJury, Internet Live Stats, Cisco, PurpleSec



2 TRILLION

searches on Google by the end of 2021



1.134 TRILLION MB

volume of data created every day



3,026,626

emails sent every second, 67% of which are spam



278,108 PETABYTES

global IP data per month by the end of 2021



230,000

new malware versions created every day



82%

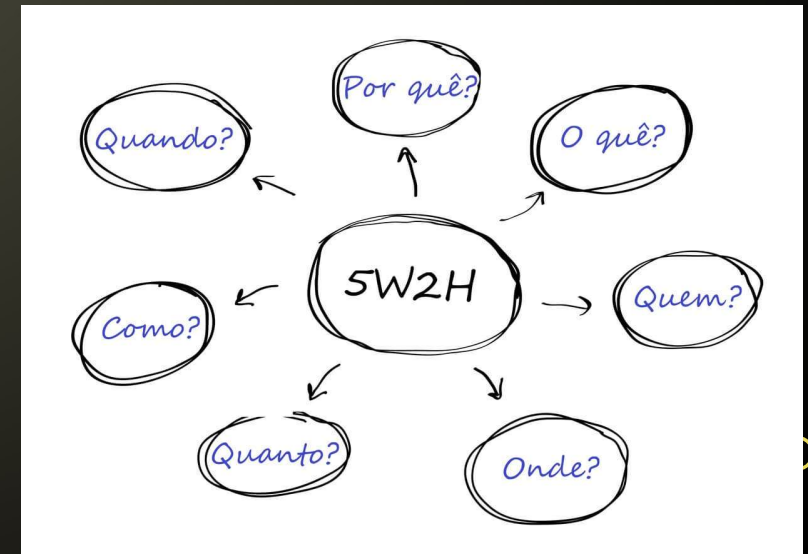
share of video in total global internet
traffic at the end of 2021

então O que faz o

Engenheiro de Dados

?

- Quais dados são importantes ?
- Como estes dados serão armazenados ?
- Como estes dados serão disponibilizados ?
- Quem terá acessos ?
- Quais os tipos de acessos (segurança) ?
- Até quando estes dados são importantes ?



então O que faz o

Engenheiro de Dados

?

- Quais dados são importantes ?
- Como estes dados serão armazenados ?
- Como estes dados serão disponibilizados ?
- Quem terá acessos ?
- Quais os tipos de acessos (segurança) ?
- Até quando estes dados são importantes ?



Estruturas de Dados

Ou

Estruturas de Armazenamento

Por que “como estes dados serão armazenados” é importante ?

Velocidade de Acesso

