Curso preparatório para as certificações
Big Data Foundation, Data Science Essentials,
Data Governance Foundation, InfoSec Foundation
e Cloud Security Foundation

Big Data Foundation

Objetivo do módulo

- Este módulo tem por objetivo preparar os participantes para realizar o exame da certificação Big Data Foundation.

Introdução ao Big Data

- “Big Data faz referência ao grande **volume**, **variedade** e **velocidade** de dados que demandam formas inovadoras e rentáveis de processamento da informação, para melhor percepção e tomada de decisão.” (*Gartner)*

Os 3 Vs de Big Data

- **Os 3 Vs de Big Data**
  - Volume
  - Variedade
  - Velocidade

Os 3 Vs de Big Data

- **Volume**
  - Cerca de 2,5 bilhões de gigabytes de dados são criados diariamente;
  - De toda a quantidade de dados disponível no mundo, aproxidamente 90% foi criado nos últimos 2 anos;
  - 1,8 bilhão de usuários ativos no Facebook;
  - 1 bilhão de usuários ativos no WhatsApp;
  - 95 milhões de fotos e vídeos por dia no Instagram;
  - 44 milhões de artigos na Wikipedia;
  - 4 bilhões de visualizações por dia no YouTube;
  - 300 horas de vídeos são carregados a cada 1 minuto no Youtube;

Os 3 Vs de Big Data

- **Velocidade**
  - Processamento em tempo real e Streaming Data (dados em streaming).
    - O que acontece em 30 segundos na Internet.

Os 3 Vs de Big Data

- **Variedade**
  - Dados são criados em diferentes formatos - como e-mails, comentários no Facebook, fotos publicadas em redes sociais e transações.
  - Big Data inclui dados estruturados, semi-estruturados e não-estruturados.
    - Dados estruturados:
      - Bases de dados relacionais.
    - Dados semi-estruturados:
      - Não possuem uma estrutura pré-definida (representação estrutural heterogênea).
      - Auto-descritivos e sem esquema prévio definido.
      - Possuem esquema de representação presente (de forma explícita ou implícita).
        - Exemplo: XML (eXtensible Markup Language).
    - Dados não-estruturados:
      - Documentos, fotos, vídeos, *tweets*, comentários em redes sociais, etc.
      - Estima-se que pelo menos **80%** dos dados gerados atualmente sejam do tipo não-estruturados.

Outros Vs do Big Data

- Além dos 3 Vs, outras duas dimensões são comumente associadas à definição de Big Data. São elas:
  - **Veracidade**
    - Confiabilidade dos dados.
  - **Valor**
    - Gerar valor para o negócio;
    - Melhor entender as necessidades dos clientes;
    - Oferecer produtos e serviços que melhor atendam as necessidades dos clientes;
    - Oferecer produtos e serviços personalizados;
    - Melhorar o relacionamento com os clientes;
    - Aumentar a fidelização e satisfação dos clientes;
    - Gerar vantagem competitiva para o negócio.

Desafios do Big Data

- Onde armazenar esses dados?
- Como estruturar esses dados?
- Como consultar esses dados?
- Como extrair valor desses dados?
- Necessidade de novas tecnologias capazes de oferecer escalabilidade, disponibilidade, flexibilidade e desempenho para a manipulação de grandes volumes de dados.

Desafios do Big Data

- Big Data necessita de grande capacidade de processamento e armazenamento.
- Computação em Nuvem oferece capacidade de processamento e armazenamento conforme a necessidade do usuário.
- Computação em Nuvem é um imperativo para Big Data.
- De acordo com NIST (*National Intitute of Standards and Technology*): Computação em Nuvem é um modelo que permite um acesso sob demanda via redes de computadores a um conjunto compartilhado de recursos computacionais que podem ser rapidamente provisionado e liberado com um mínimo de esforço administrativo ou interação com o provedor de serviços.

Armazenamento

- **SGBDs mais utilizados:**

Business Intelligence (BI)

- **Business Intelligence (BI)**
  - Consolida os principais indicadores de uma empresa com base em dados obtidos no sistema de gestão integrado ERP e fornece visões precisas e analíticas, que apoiam a tomada de decisão.
  - As ferramentas de BI geram gráficos (Dashboards) e relatórios de acordo com a necessidade do usuário.
  - Por meio do BI podem ser detectadas sazonalidades, tendências e padrão de comportamento.

Business Intelligence (BI)

- Ferramentas de Business Intelligence (BI)

Armazenamento

- **Base de dados não relacionais**
  - **NoSQL (Not Only SQL)**
    - Conjunto de conceitos que permite o processamento rápido e eficiente de conjuntos de dados com foco em desempenho, confiabilidade e agilidade.
    - Diferentes formas de armazenamento
      - Orientado a documentos (o mais popular)
        - Ex.: MongoDB (https://www.mongodb.com/) e Apache CouchDB (https://couchdb.apache.org/)
      - Orientado a chave-valor (o mais simples)
        - Ex.: Amazon DynamoDB (https://aws.amazon.com/dynamodb/)
      - Orientado a grafos (o mais especializado)
        - Ex.: Neo4j (https://neo4j.com/)
      - Orientado a colunas (o mais complexo)
        - Ex.: HBase (http://hbase.apache.org/)
    - Características
      - Não-relacional
      - *Cluster-friendly*
      - Interface de consulta simples.

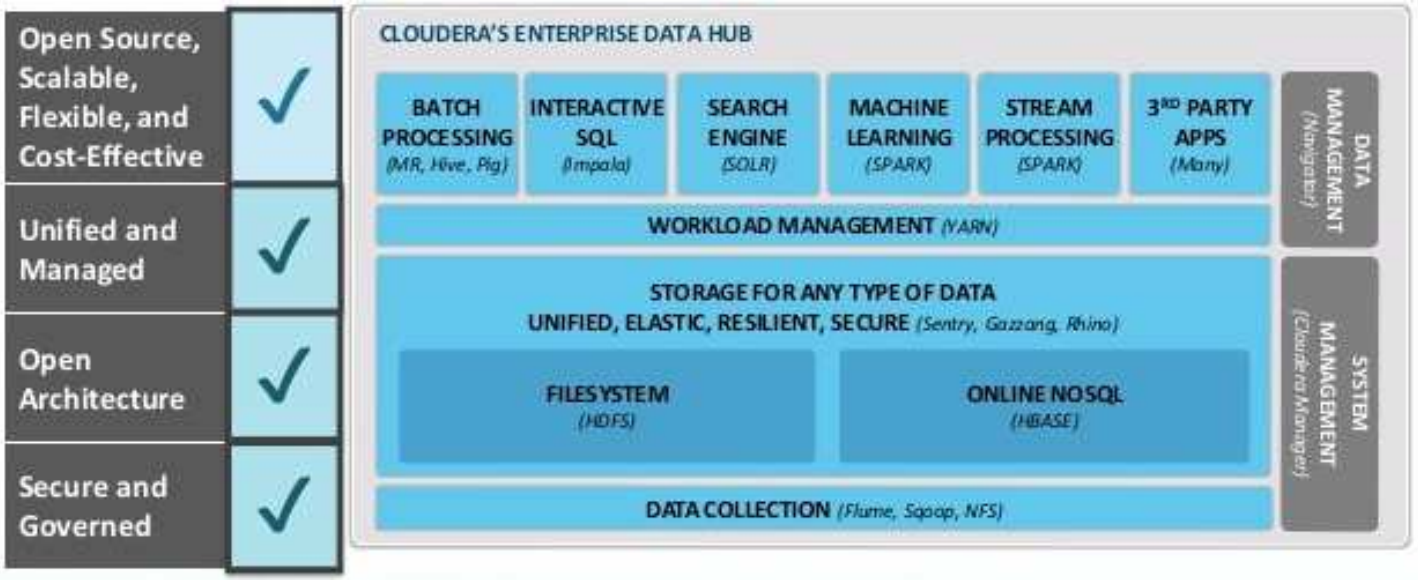
Armazenamento

- **Data Lake**
  - Repositório único no qual dados estruturados e não-estruturados, coletados de diferentes fontes, são armazenados em sua forma bruta, como foram coletadas na fonte, sem qualquer processamento.

- **Enterprise Data Hub (EDH)**
  - Permite que a empresa tenha uma fonte de dados centralizada e unificada que possa fornecer rapidamente informações a diversos usuários do negócio, apoiando a tomada de decisão.
    - Soluções:
      - Azure Data Lake
        - https://azure.microsoft.com/pt-br/solutions/data-lake/
      - Cloudera Enterprise Data Hub
        - https://www.cloudera.com/products/enterprise-data-hub.html
      - Enterprise Data Hub (MapR)
        - https://mapr.com/solutions/enterprise/enterprise-data-hub/

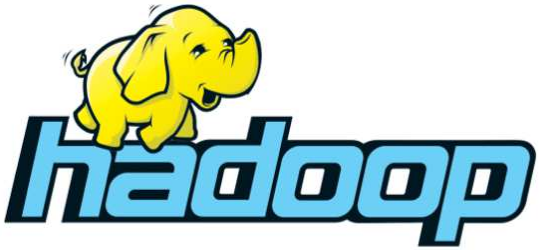
Armazenamento

- **Enterprise Data Hub (EDH)**



Processamento

- **Hadoop**
  - http://hadoop.apache.org/
  - Solução *open source* que permite a execução de aplicações de Big Data utilizando milhares de máquinas.
  - Projetado para processar grandes quantidades de dados estruturados e não-estruturados.
  - Oferece recursos de armazenamento, gerenciamento e processamento de dados distribuídos.
  - Benefícios:
    - Redução de custo
    - Flexibilidade
    - Escalabilidade
    - Desempenho



Processamento

- **Principais fornecedores de mercado**

cloudera

MAPR®

Hortonworks

Processamento

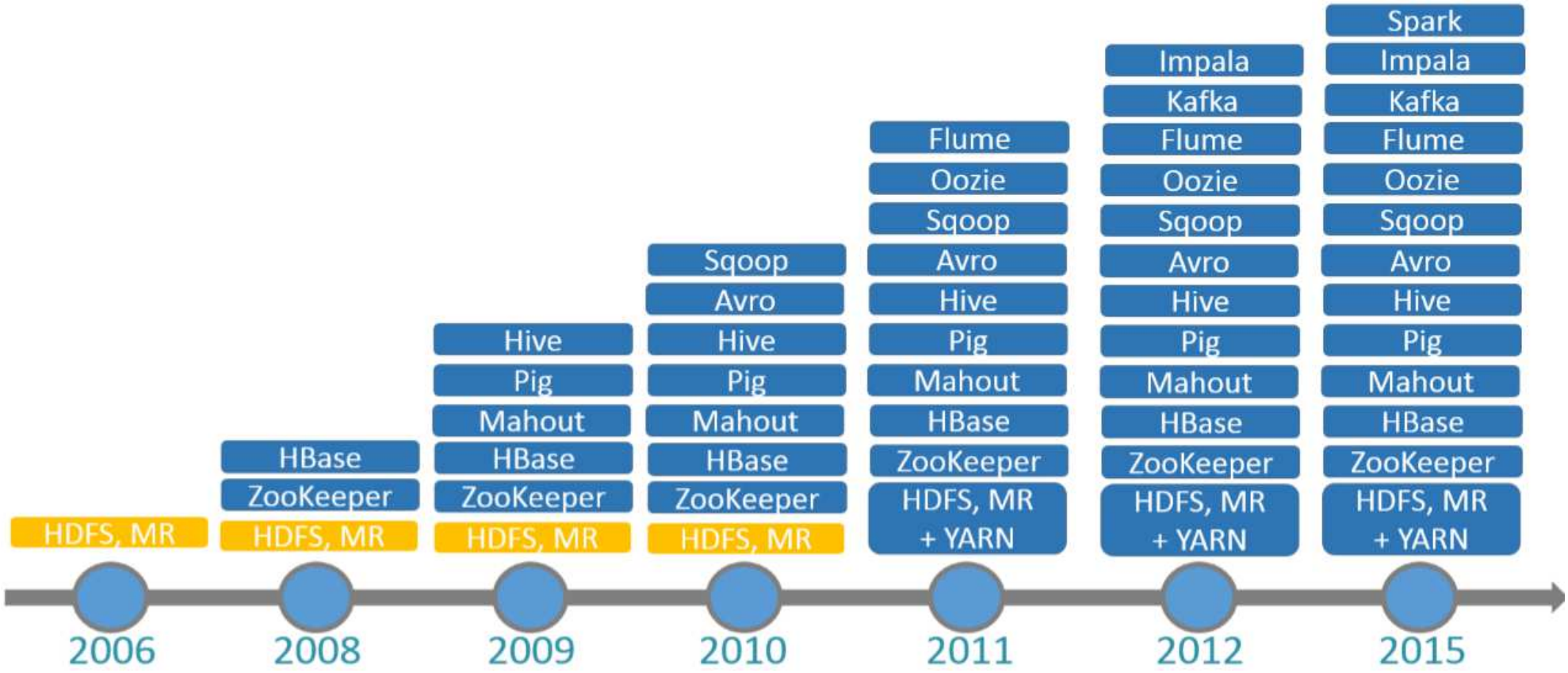
- **Ecossistema Hadoop**
  - Hadoop possui 2 componentes principais:

Processamento

- **Hadoop HDFS (Hadoop Distributed File System)**
  - Sistema de arquivos distribuídos;
  - Otimizado para processamento de **grandes volumes de dados (alta taxa de transferência)**;
  - Otimizado para ler e armazenar grandes arquivos em *clusters*;
  - Arquivos são divididos em blocos de 64 ou 128 MB (tamanho *default – pode ser configurado*);
  - Escalável e tolerante a falhas.

- **Hadoop MapReduce**
  - É uma implementação do Hadoop;
  - Ferramenta utilizada para facilitar o processamento de grandes volumes de dados (multi-terabyte data-sets) de forma distribuída;
  - Tolerante a falhas;
  - Funcões *Map* e *Reduce*.

Ecossistema Hadoop



Processamento

- **Algumas empresas que utilizam o Hadoop**

Aplicações

- **Varejo**
  - Melhor segmentação de clientes;
  - Propaganda personalizada;
  - Melhor oferta de produtos e serviços com base no perfil, rastro digital e no histórico de compras do cliente;
  - Previsão e prevenção de *Customer Churn* (identificação de clientes com alta propensão a cancelar produtos e serviços);
  - *Chatbots*.

- **Setor financeiro**
  - Detecção de transações fraudulentas envolvendo utililização de Internet Banking e cartões de crédito;
  - Análise de crédito;
  - Melhor relacionamento com os clientes.

Aplicações

- **People Analytics (HR Analytics)**
  - Processo de coleta, armazenamento e análise de dados sobre o comportamento dos colaboradores em uma organização.
  - Utilização de análise de dados em Gestão de Pessoas.
  - Utilização de informações disponíveis em redes sociais.
  - Utilização de leitores biométricos e crachás inteligentes.
  - Análise de currículo utilizando *Text Analytics* (Análise de Texto).
  - Utilização de rastro digital e informações de redes sociais para ajudar na identificação do perfil mais adequado para cada vaga.
  - Principais benefícios:
    - Otimização do processo de Recrutamento e Seleção;
    - Avaliação de Desempenho;
    - Aumento da produtividade;
    - Desenvolvimento de programas de treinamento e capacitação;
    - Retenção de talentos;
    - Redução da Rotatividade.

Aplicações

- **Internet of Things (Internet das Coisas)**
  - Rede formada por milhares de dispositivos (objetos) inteligentes conectados a Internet.
  - Dispositivos capazes de capturarem grandes quantidades de dados por meio de sensores.
  - Exemplos:
    - Veículos autônomos e conectados;
    - Casas inteligentes;
    - Eletrodomésticos inteligentes;
    - *Wearables* (Relógios e pulseiras inteligentes).

Exemplos de empresas com negócios centrados em dados

Simulados

1. Selecione nas alternativas abaixo os 3 Vs do Big Data.

a. Volume, Velocidade e Variedade
b. Velocidade, Vazão e Volume
c. Volume, Visibilidade e Virtude
d. Visão, Valores e Variedade

Simulados

2. Selecione nas alternativas abaixo o sistema de arquivos distribuído otimizado para processamento de grande volume de dados.

a. Ext4

b. NTFS

c. FAT

d. HDFS

Simulados

3. Selecione nas alternativas abaixo os dois componentes centrais do ecossistema Hadoop.

a. MySQL e NTFS

b. Hadoop MapReduce e NTFS

c. Hadoop MapReduce e HDFS

d. PostgreSQL e NTFS

Simulados

4. Além das três dimensões (3 Vs - Volume, Velocidade e Variedade), quais são as outras duas dimensões comumente associadas à definição de Big Data?

a. Garantia e Utilidade

b. Utilidade e Volatilidade

c. Autenticidade e Validade

d. Valor e Veracidade

Simulados

5. Selecione nas alternativas abaixo um exemplo de Banco Dados NoSQL orientado a documentos.

a. MongoDB

b. Microsoft Access

c. SQL Server

d. Oracle Database

Simulados

6. Considere as seguintes informações sobre Data Lake:

I. O Data Lake não pode ser criado na nuvem.

II. No Data Lake os dados são armazenados em seu formato bruto como foram coletados na fonte de dados.

III. O Data Lake armazena apenas dados estruturados.

É correto o que se afirma em:

a. II, apenas

b. I e II, apenas

c. II e III, apenas

d. III, apenas

Simulados

7. Selecione nas alternativas abaixo três ferramentas largamente utilizadas em BI:

a. Microsoft Access, IBM Db2 e SQLite

b. Microsoft Access, MySQL e MariaDB

c. Qlik, Tableau e Power BI

d. SQLite, Microsoft Access e PostgreSQL

Data Science Essentials

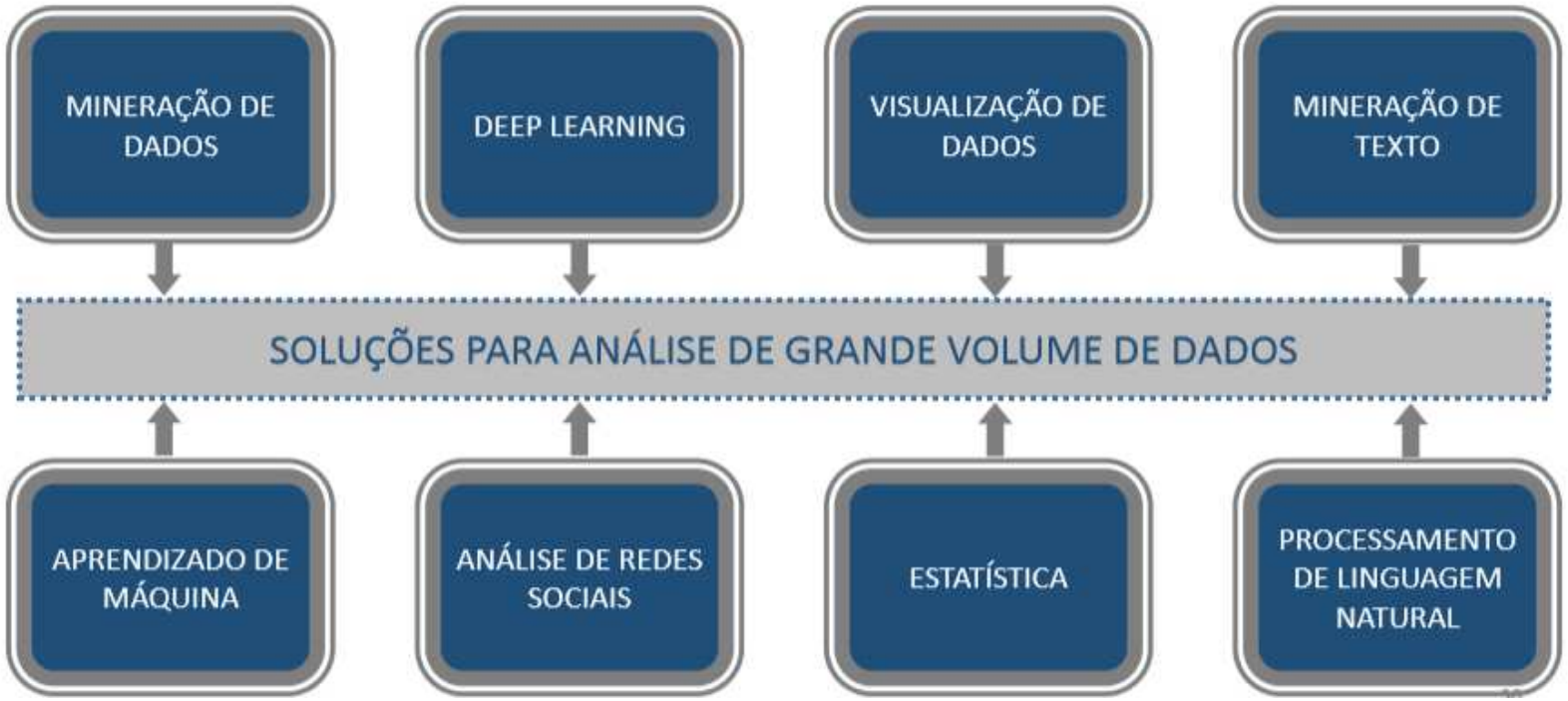
Objetivo do Módulo

- Este módulo tem por objetivo preparar os participantes para realizar o exame da certificação Data Science Essentials.

Introdução a Data Science

- **Data Science (Ciência dos Dados)**
  - Termo utilizado para descrever o processo de extração, análise e interpretação de grandes volumes de dados, gerados à partir de diversas fontes, a fim de extrair insights e informações valiosas para auxiliar na tomada de decisões.
  - Incorpora conhecimentos das áreas de Matemática e Estatística e utiliza diversas técnicas, como Modelagem Preditiva, Mineração de Dados (*Data Mining*), Análise de Texto (*Text Analysis*), Aprendizado de Máquina (*Machine Learning*) e Visualização de Dados (*Data Visualization*).

Inovação na análise de dados



Técnicas utilizadas em Análise de Dados

- ***Data Mining* (Mineração de Dados)**
  - Processo que tem por objetivo analisar grandes quantidades de dados a fim de identificar padrões e extrair informações.
    - Exemplo: Weka (https://www.cs.waikato.ac.nz/ml/weka/)

- ***Machine Learning* (Aprendizado de Máquina)**
  - Subcampo da Inteligência Artificial dedicado ao desenvolvimento de algoritmos e técnicas que permitem ao computador aprender à partir do reconhecimento de padrões.
    - Exemplo: IBM Watson Machine Learning (https://www.ibm.com/cloud/machine-learning).

- ***Deep Learning* (Aprendizado Profundo)**
  - Subárea de Maching Learning.
  - Aprendizado de máquina utilizando redes neurais artificais.
    - Processamento de Linguagem Natural;
    - Reconhecimento de Fala;
    - Visão Computacional;
    - Processamento de Imagens.

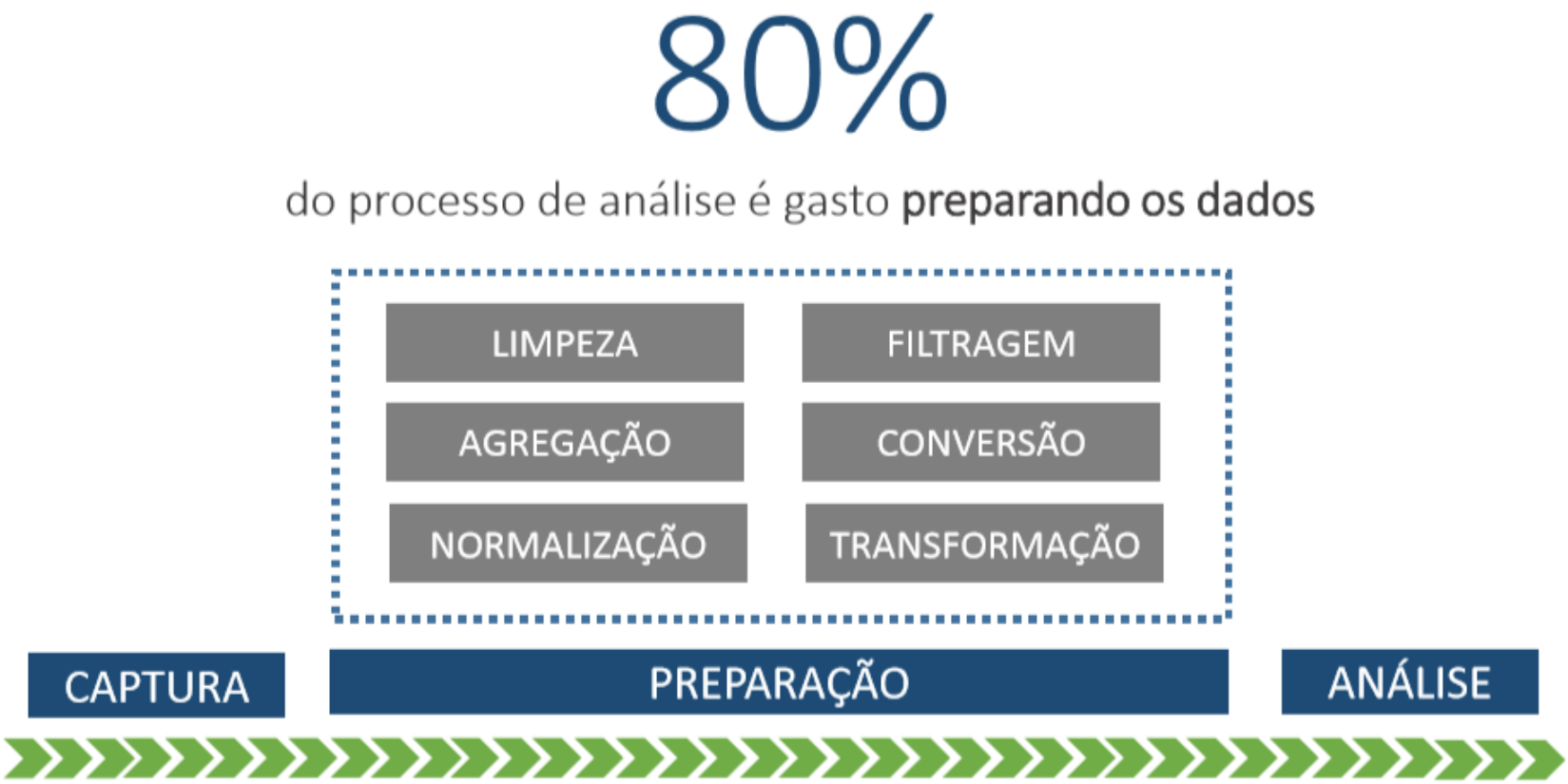
Técnicas utilizadas em Análise de Dados

- **Estatística**
  - Conjunto de métodos largamente utilizados na coleta e interpretação de dados em *Data Science*.

- **Análise de Texto (*Text Analysis*)**
  - Análise de dados não estruturados (textos).
  - Utiliza Processamento de Linguagem Natural.
    - Subcampo da Inteligência Artificial que estuda a compreensão da linguagem natural (Análise Semântica e Análise sintática).
  - Técnica utilizada para realizar análise de conteúdo em mídias sociais.

- ***Data Storytelling***
  - Técnica de construção de narrativas por meio da utilização de visualização de dados.

Inovação na análise de dados



Oportunidades

- Big Data Analytics
- Serviços orientados a dados
- Monetização de dados
- Ferramentas para análise de dados
- Serviços de visualização de dados

Big Data Analytics

- Processo de coletar, organizar e analisar enormes conjuntos de dados a fim de se descobrir padrões, tendências e fazer correlações.
- Faz uso de dados históricos, mineração de dados, aprendizado de máquina, entre outros métodos.
- Quatro abordagens:
  - Descritiva
  - Diagnóstica
  - Preditiva
  - Prescritiva

Big Data Analytics

- **Descritiva**
  - Entender os eventos ocorridos.
  - Faz uso de ferramentas de BI (*Business Intelligence*);
  - Baseado em métricas;
  - *Dashboards*, relatórios e alertas;
  - Estima-se que mais de 80% das análises de negócios realizadas sejam descritivas.

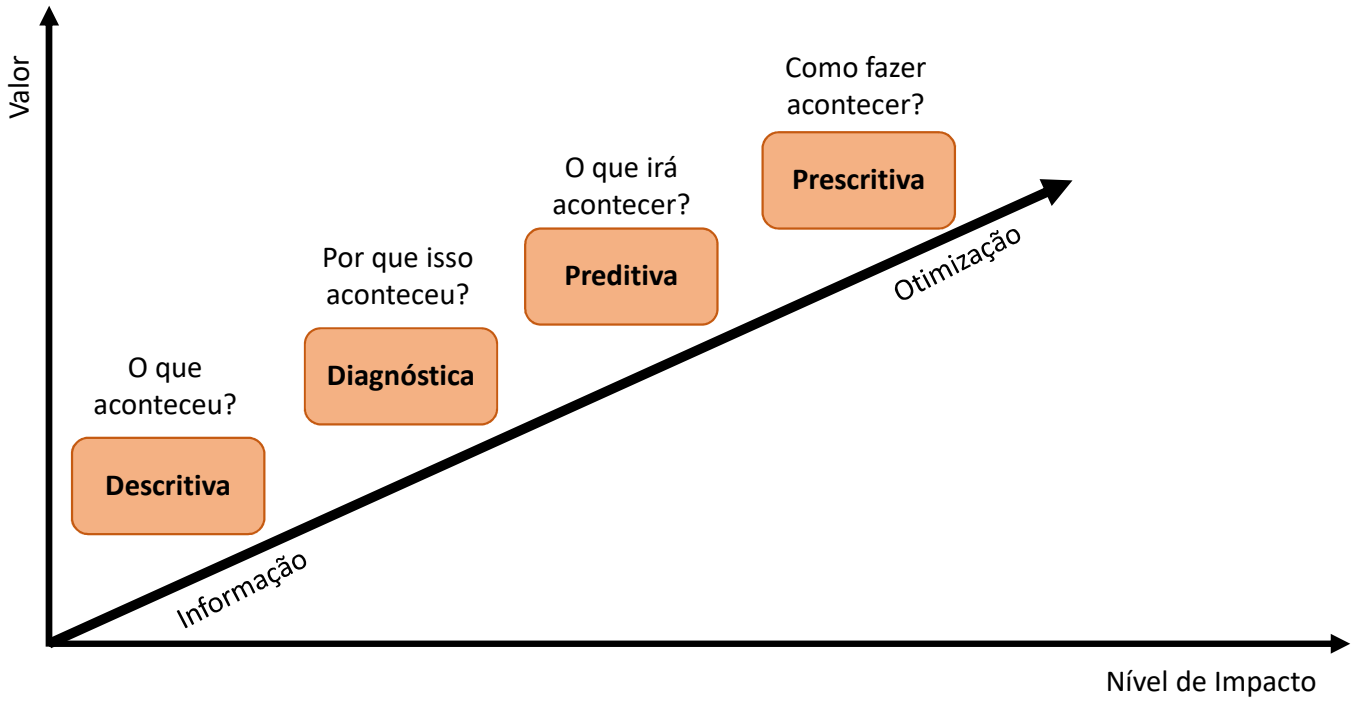
- **Diagnóstica**
  - Tem por objetivo determinar a causa de um evento;
  - Uso de análise de correlação, análise de variância e testes de hipótese.

Big Data Analytics

- **Preditiva**
  - Prever tendências baseadas em dados;
  - Predição de eventos futuros;
  - Faz uso de diferentes métodos e ferramentas (análises estatísticas, técnicas de simulação, mineração de dados e aprendizado de máquina);
  - Utilizada para diferentes aplicações, como por exemplo, prever quais clientes de uma operadora de telefonia estariam mais propensos a cancelarem um determinando serviço.

- **Prescritiva**
  - Predizer as possíveis consequências para as diferentes escolhas que forem feitas;
  - Sugere ações baseadas no conhecimento extraído dos dados.

Big Data Analytics



Serviços orientados a dados

- **Web sites**
  - Análise de experiência do usuário
  - Personalização de conteúdo
- **Geolocalização**
  - Recomendação de serviços
- **Comércio eletrônico**
  - Recomendação de produtos
  - Segmentação de clientes
- **E-mails e mensagens**
  - Publicidade personalizada
- **Redes sociais**
  - Análise de influência
  - Análise de sentimento

Monetização de dados

- Transformar ativos de informação em dinheiro, direta ou indiretamente, por meio de troca, comercialização ou venda direta.
  - Dados genéticos para pesquisadores;
  - Dados de comportamento de usuários para campanhas de marketing;
  - Dados de sensores para seguradoras.

Ferramentas para análise de dados

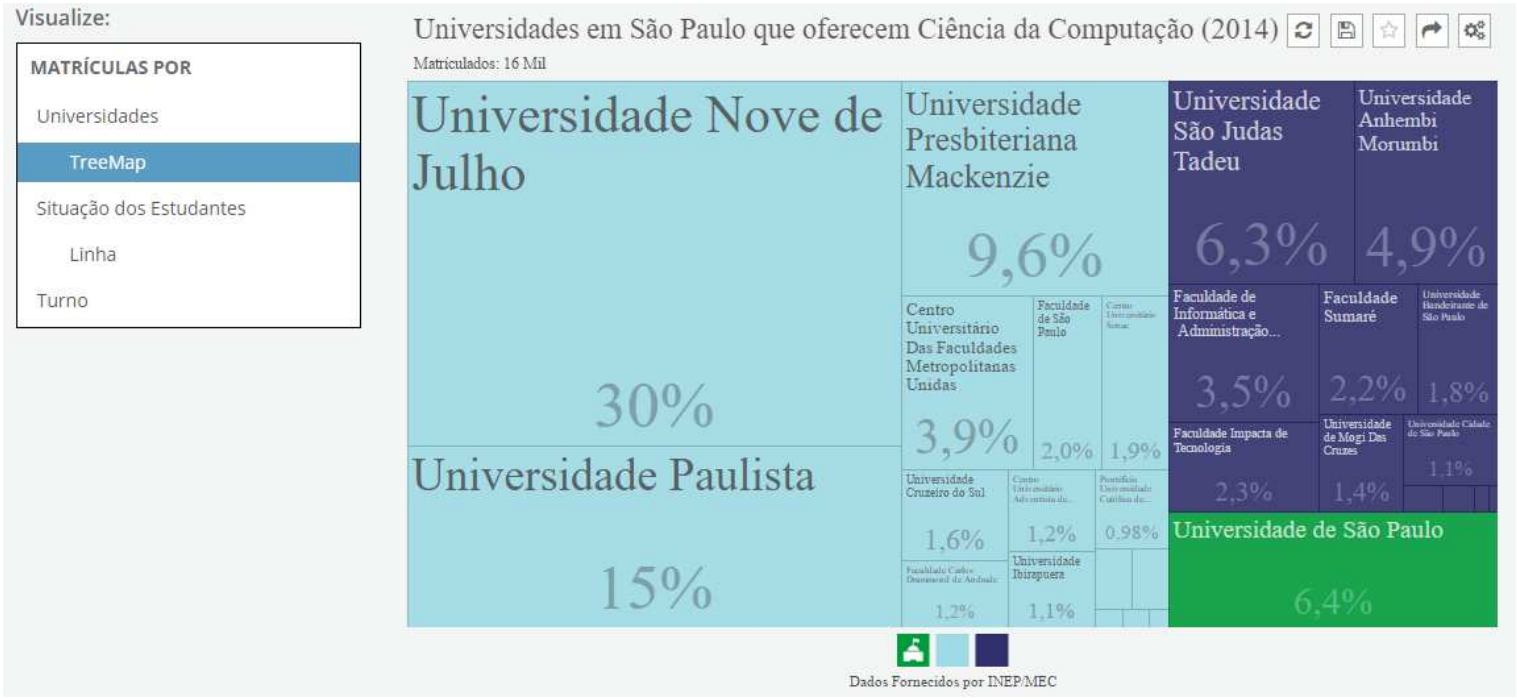
- Processamento distribuído dos dados;
- Processamento em tempo real;
- Algoritimos de aprendizado de máquina (*Machine Learning*).

Serviços de visualização de dados

- Representação gráfica de informações (utilização de gráficos de barra, gráficos de linha, gráficos de área, imagens, mapas, entre outros).
- Tem por objetivo facilitar a compreensão dos dados.
- **Benefícios**
  - Tomada de decisão aperfeiçoada;
  - Monitoramento e aumento da produtividade;
  - Melhoria na análise de dados;
  - Melhor experiência ao usuário.
- Exemplos:
  - DataViva – Portal de visualização de dados da economia brasileira. Disponibiliza dados socioeconômicos de mais de 5 mil municípios brasileiros.
    - http://dataviva.info/pt/
  - http://circos.ca/
  - Tweetping - https://tweetping.net/
  - Infogr.am https://infogr.am/

Serviços de visualização de dados

- DataViva – Portal de visualização de dados da economia brasileira.
- Disponibiliza dados socioeconômicos de mais de 5 mil municípios brasileiros.



Linguagens de programação utilizadas em Data Science

- **Linguagem R**
  - Linguagem de programação estatística largamente utilizada em Big Data.
  - https://www.r-project.org/

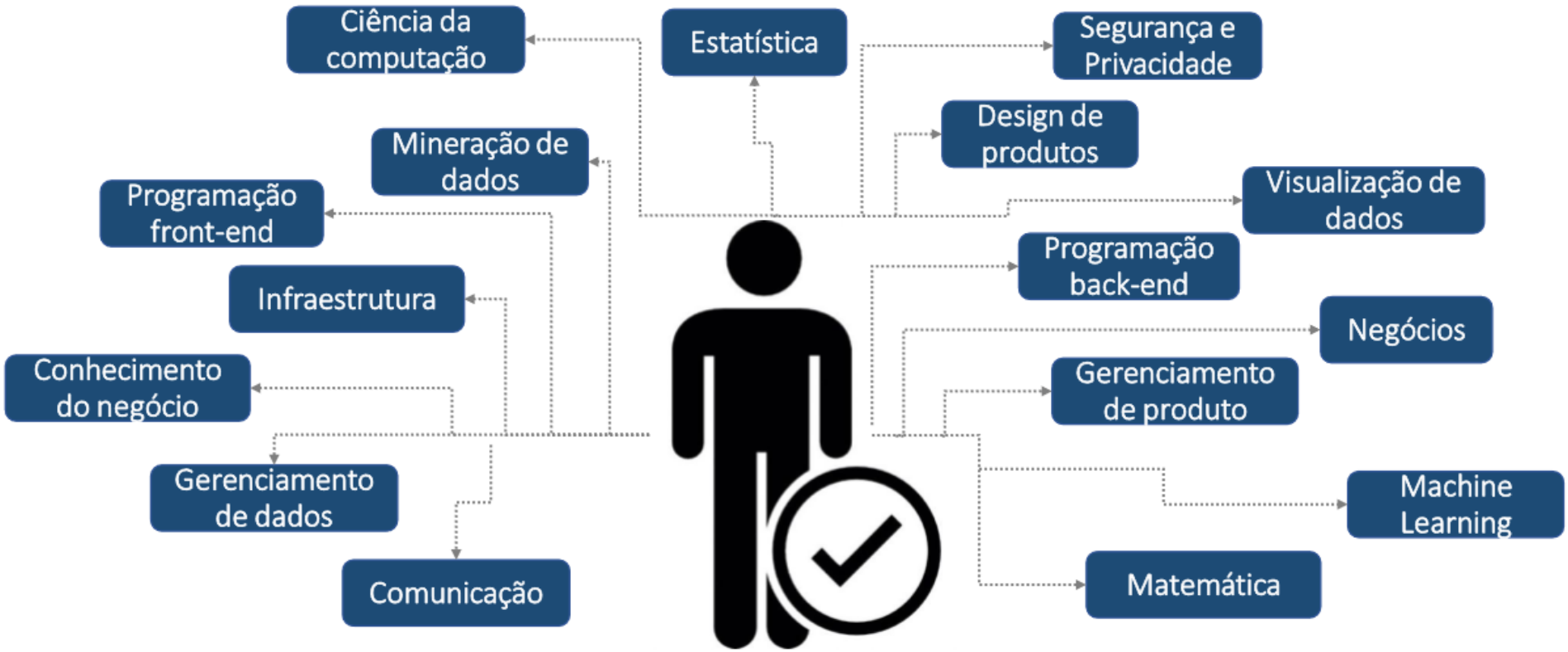
- **Python**
  - Python é uma das linguagens de programação orientada a objetos mais versáteis, fáceis e rápidas de serem aprendidas.
  - https://www.python.org/

- **Scala (Scalable Language)**
  - Scala é uma linguagem de programação que combina os paradigmas de programação orientada a objetos e funcional;
  - Utiliza uma sintaxe concisa que é totalmente compatível com Java e é executado na JVM (*Java Virtual Machine*);
  - Tem ganhado cada vez mais adeptos por conta da sua capacidades de lidar com grandes quantidades de dados de maneira escalável e confiável;
  - http://www.scala-lang.org/

Cientista de dados

- Profissional responsável por extrair informações de grandes volumes de dados estruturados e não estruturados;
- Profissão em alta;
- Salários podem chegar a 22 mil reais (Robert Half/Computerworld);
- Carência de profissionais qualificados;
- Profissionais capacitados são altamente disputados;
- Carreira altamente promissora.

Cientista de Dados



Simulados

1. Selecione nas alternativas abaixo a linguagem de programação estatística largamente utilizada em Big Data.

a. Linguagem C

b. Linguagem de Máquina

c. Linguagem R

d. Linguagem A

Simulados

2. Qual tipo de análise poderia ser utilizada por uma operadora de telefonia para prever quais clientes estariam mais propensos a cancelarem um serviço?

a. Análise Preditiva

b. Análise Diagnóstica

c. Análise Descritiva

d. Análise Prescritiva

Simulados

3. Selecione nas alternativas abaixo o processo que tem por objetivo analisar grandes quantidades de dados a fim de identificar padrões e extrair informações.

a. Data Sets

b. Data Mining

c. Data Mart

d. Data Management

Simulados

4. Selecione nas alternativas abaixo o tipo de análise que tem por objetivo compreender a causa de um evento.

a. Análise Diagnóstica

b. Análise Preditiva

c. Análise Prescritiva

d. Análise Descritiva

Simulados

5. Selecione nas alternativas abaixo como é chamado o principal executivo de dados em uma organização.

a. Chief Executive Officer (CEO)

b. Chief Information Officer (CIO)

c. Chief Information Security Officer (CISO)

d. Chief Data Officer (CDO)

Simulados

6. Selecione nas alternativas abaixo o tipo de análise que tem por objetivo ajudar a entender os eventos ocorridos e faz uso de ferramentas de BI (Business Intelligence), dashboards, relatórios e alertas.

a. Análise Descritiva

b. Análise Preditiva

c. Análise Diagnóstica

d. Análise Prescritiva

Data Governance Foundation

Objetivo do Módulo

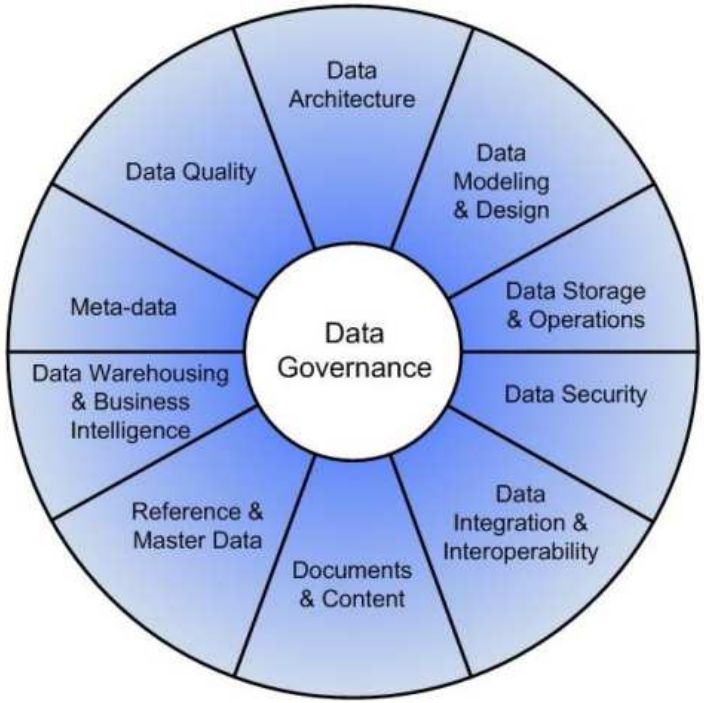
- Este módulo tem por objetivo preparar os participantes para realizar o exame da certificação Data Governance Foundation.

Introdução a Gestão de Dados

- Gestão de Dados é a disciplina responsável por definir, planejar, implantar e executar estratégias, procedimentos e práticas necessárias para gerenciar de forma efetiva os dados das organizações.
- Cuida do planejamento, controle e entrega de ativos de dados e de informação.

Guia DAMA-DMBOK

- O Guia DAMA-DMBOK® é um conjunto de boas práticas de Gestão de Dados.
- Corpo de conhecimento de Gestão de Dados utilizado internacionalmente.
- O DAMA-DMBOK® estabelece 11 (onze) áreas de conhecimento de Gestão de Dados (Versão 2 - 2017).



Governança de Dados

- Função central do guia DAMA-DMBOK® e que influencia todas as demais funções.
- Governança de Dados é o exercício de autoridade e controle (planejamento, monitoramento e engajamento) sobre o gerenciamento de ativos de dados.
- É o exercício de tomada de decisão e autoridade para as questões relacionadas a dados.

Arquitetura de Dados

- Responsável por definir as necessidades de dados da empresa.
- **Arquitetura de Dados**
  - Conjunto de componentes de dados e suas relações dentro de uma organização.
  - A Arquitetura de Dados descreve como os dados são processados, armazenados e utilizados pelas aplicações.
  - A Arquietura de Dados inclui:
    - Arquitetura de Dados Corporativos
    - Modelos de Dados Corporativos
      - Modelos conceituais de dados
      - Modelos lógicos de dados
    - Arquitetura de Integração de Dados
    - Infraestrutura

Modelagem de Dados e Projeto

- O objetivo da Modelagem de Dados e Projeto é projetar, implementar e manter soluções que satisfaçam as necessidades de dados da empresa, ao longo de todo o ciclo de desenvolvimento do sistema.
- Inclui atividades de Modelagem de Dados, Análise de Requisitos, Projeto, Implementação e Manutenção de Banco de Dados.

Armazenamento de Dados e Operações

- O objetivo de Armazenamento de Dados e Operações é planejar, controlar a suportar os ativos de dados ao longo do seu ciclo de vida, desde a criação e aquisição (obtenção) até o arquivamento final e eliminação dos dados.
- Inclui atividades relacionadas a planejamento de backup e recuperação de dados, definição de tipo de SGDB, obtenção de dados de fontes externas, performance, políticas de armazenamento, retenção e eliminação de dados, entre outras.

Dados Mestres e Referência

- Responsável por definir e controlar atividades para garantir a consistência e disponibilização de visões únicas dos dados mestres e de referência da empresa.
  - **Dados mestres**
    - São os dados críticos do negócio geralmente utilizados por diversas aplicações dentro de uma organização.
  - **Dados de referência**
    - São utilizados para categorizar outros dados.

Qualidade de Dados

- Qualidade de Dados tem por objetivo garantir que os dados armazenados estejam corretos, precisos, consistentes, completos, integrados, aderentes às regras de negócio e aos domínios estabelecidos.
- Responsável por medir, avaliar, melhorar e garantir a qualidade dos dados da organização.
- **Requisitos de Qualidade de Dados:**
  - **Atualidade**
    - Determina se os dados estão atualizados e representam verdadeiramente o estado atual das informações.
  - **Completude**
    - O requisito de Qualidade de Dados que determina se os dados estão completos de acordo com as informações exigidas na execução dos processos de negócios.
  - **Consistência**
    - Determina se os dados estão integros e coerentes.
  - **Unicidade**
    - Indica que o dado é único e exclusivo dentro da organização.
  - **Precisão**
    - Indica se os dados são precisos.

Documentos e Conteúdo

- Define como planejar, implementar e controlar atividades para armazenar, proteger e acessar dados não estruturados (textos, imagens, gráficos, imagens, áudio e vídeo) da empresa.

Metadados

- Área de conhecimento responsável por gerir e armazenar os metadados da empresa, bem como viabilizar formas de acesso.

Data Warehousing e Business Intelligence

- Área de conhecimento responsável por definir e controlar processos para prover dados de suporte à tomada de decisão, geralmente disponibilizados em aplicações analíticas.

Segurança de Dados

- Área de conhecimento responsável por definir as politicas, os controles e procedimentos necessários para garantir a adequada autenticação, controle de acesso e auditoria de dados.

Integração de Dados e Interoperabilidade

- Área de conhecimento responsável pela aquisição, extração, transformação, replicação, federação e virtualização de dados.

Simulados

1. Selecione nas alternativas abaixo o número de áreas de conhecimento de Gestão de Dados de acordo com o DAMA-DMBOK® (Versão 2).

a. 11

b. 2

c. 20

d. 31

Simulados

2. Selecione nas alternativas abaixo a área de conhecimento (função central) do guia DAMA-DMBOK® e que influencia todas as demais áreas.

a. Segurança dos Dados

b. Governança de Dados

c. Arquitetura de Dados

d. Gestão de Metadados

Simulados

3. Selecione nas alternativas abaixo a área de conhecimento da Gestão de Dados que tem por objetivo garantir que os dados armazenados estejam corretos, precisos, consistentes, completos, íntegros e aderentes às regras de negócio e aos domínios estabelecidos.

a. Qualidade de Dados

b. Gestão de Pessoas

c. Segurança de Dados

d. Armazenamento de Dados e Operações

Simulados

4. Selecione nas alternativas abaixo a área de conhecimento da Gestão de Dados que define como planejar, implementar e controlar atividades para armazenar, proteger e acessar dados não estruturados (textos, imagens, gráficos, imagens, áudio e vídeo).

a. Gestão de Risco
b. Segurança de Dados
c. Gestão de Projetos
d. Documentos e Conteúdo

Simulados

5. Qual área de conhecimento inclui atividades relacionadas a planejamento de backup e recuperação de dados, definição de tipo de SGDB, obtenção de dados de fontes externas, performance, políticas de armazenamento, retenção e eliminação de dados, entre outras.

a. Gestão de Projetos

b. Segurança de Dados

c. Armazenamento de Dados e Operações

d. Gestão de Risco

Simulados

6. Como é chamado o conjunto de componentes de dados e suas relações dentro de uma organização (descreve como os dados são processados, armazenados e utilizados)?

a. Segurança de Dados

b. Qualidade de Dados

c. Arquitetura de Dados

d. Arquitetura de Computadores

Simulados

7. Selecione nas alternativas abaixo a função responsável por definir e controlar processos para prover dados de suporte à tomada de decisão, geralmente disponibilizados em aplicações analíticas.

a. Data Warehousing e Business Intelligence
b. Segurança de Dados
c. Gestão de Riscos
d. Gestão de Pessoas

InfoSec Foundation

Objetivo do módulo

- Este módulo tem por objetivo preparar os participantes para realizar o exame da certificação InfoSec Foundation.

Normas

- **Normas**
  - Normas são documentos estabelecidos por consenso e aprovado por um organismo reconhecido, que fornece, para uso comum e repetitivo, regras, diretrizes ou características para atividades ou seus resultados, visando a obtenção de um grau ótimo de ordenação em um dado contexto.

- **Normas da família ISO IEC 27000**
  - As normas da família ISO/IEC 27000 são normas internacionais que apresentam os requisitos necessários para a implementação de um Sistema de Gestão da Segurança da Informação (SGSI) em qualquer organização por meio do estabelecimento de políticas de segurança, controles e gerenciamento de risco.

Principais normas de segurança

- ISO/IEC 27000
  - Overview and vocabulary (Termos e definições aplicáveis a todas as normas da família 27000)
- ABNT NBR ISO/IEC 27001
  - Sistemas de gestão da segurança da informação — Requisitos
- ABNT NBR ISO/IEC 27002
  - Código de prática para controles de segurança da informação
- ABNT NBR ISO/IEC 27003
  - Diretrizes para implantação de um sistema de gestão da segurança da informação
- ABNT NBR ISO/IEC 27004
  - Gestão da segurança da informação — Medição
- ABNT NBR ISO/IEC 27005
  - Gestão de riscos de segurança da informação

Outras normas importantes

- ABNT NBR ISO 31000:2018
  - Gestão de riscos - Diretrizes
- ABNT NBR ISO/IEC 20000-1
  - Tecnologia da informação — Gestão de serviços
    Parte 1: Requisitos do sistema de gestão de serviços
- ABNT NBR ISO/IEC 20000-2
  - Tecnologia da informação — Gerenciamento de serviços
    Parte 2: Guia de aplicação do sistema de gestão de serviços
- ABNT NBR ISO/IEC 38500
  - Governança corporativa de tecnologia da informação
- ABNT NBR ISO 22301
  - Requisitos para um sistema de gestão da continuidade de negócios

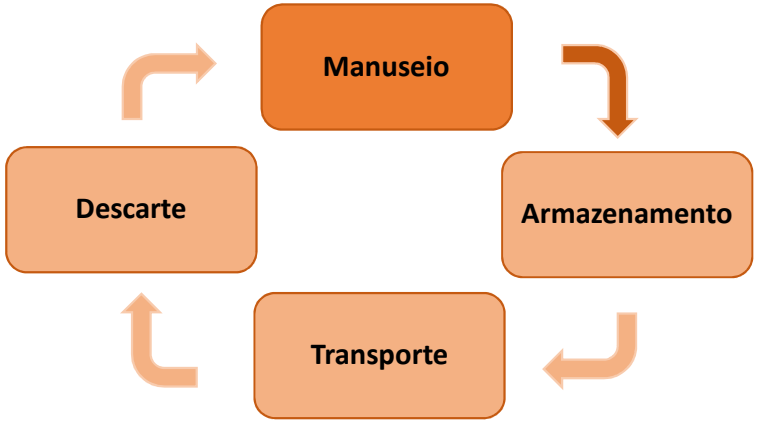
Conceitos e definições

- **Informação**
  - A informação é um ativo que, como qualquer outro ativo importante, é essencial para os negócios de uma organização e, consequentemente, necessita ser adequadamente protegida.

  - A informação pode existir em diversas formas:
    - Impressa ou escrita em papel;
    - Armazenada eletronicamente;
    - Transmitida pelo correio ou por meios eletrônicos ;
    - Arquivos de imagem, áudio ou vídeo.

Conceitos e definições

- **Ciclo de vida da informação**
  - Manuseio: trata-se do início do ciclo, onde a informação é gerada e manipulada;
  - Armazenamento: momento em que a informação é armazenada;
  - Transporte: momento em que a informação é enviada e/ou transportada;
  - Descarte: parte final do ciclo, onde a informação é descartada, eliminada, apagada, destruída de forma definitiva.



Conceitos e definições

- **Requisitos de Segurança da Informação**
  - **Confiabilidade**
    - **Aspectos da Confiabilidade da Informação**
      - **C**onfidencialidade: proteger uma informação contra acesso não autorizado.
      - **I**ntegridade: proteger a informação contra alteração não autorizada.
      - **D**isponibilidade: garantir que um recurso esteja disponível sempre que necessário.
  - **Autenticação**: verificar se a entidade é realmente quem ela diz ser.
  - **Autorização**: determinar as ações que a entidade pode executar.
  - **Identificação**: permitir que uma entidade se identifique, ou seja, diga quem ela é.
  - **Não-repúdio (Irretratabilidade):** evitar que uma entidade possa negar que foi ela quem executou uma ação.

Conceitos e definições

- **Ativo**
  - Qualquer coisa que tenha valor para a organização.

- **Classes de ativos**
  - Ativo tangível – produto, bem, equipamento, imóvel, informação em papel;
  - Ativo intangível – marca, reputação e catálogo intelectual.

- **Ativo de informação**
  - Bases de dados, arquivos, documentação de sistema, manuais de usuário, planos de continuidade do negócio, contratos, etc...

Conceitos e definições

- **Vulnerabilidade**
  - Fragilidade de um ativo ou grupo de ativos que pode ser explorada por uma ou mais ameaças.
  - É uma condição que, quando explorada por um atacante, pode resultar em uma violação de segurança.
- **Patch**
  - Termo atribuído à correção desenvolvida para eliminar falhas de segurança em um programa ou sistema operacional.

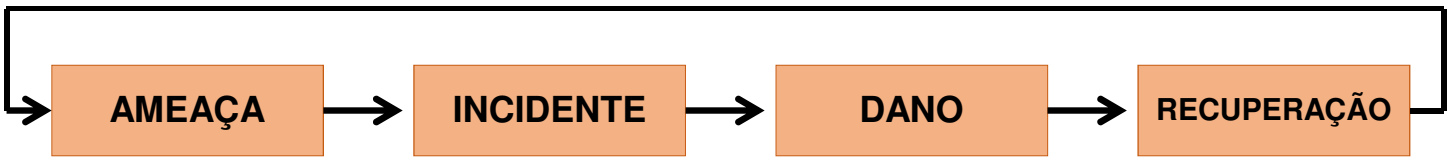
Conceitos e definições

- **Ameaça**
  - Causa potencial de um incidente indesejado, que pode resultar em dano para um sistema ou organização.
  - **Tipos de ameaça**
    - **Humana intencional** – danos causados de forma proposital.
      - Hackers;
      - Engenharia social;
      - Vandalismo;
      - Roubo e furto;
      - Sabotagem;
      - Incêndio culposo.
    - **Humana não intencional** – danos causados de forma involuntária.
      - Pen-drive com vírus;
      - Uso inadequado de extintor de incêndio.
    - **Não humana**
      - Incêndio;
      - Relâmpagos;
      - Inundação;
      - Enchente.

Conceitos e definições

- **Incidente de segurança da informação**
  - Qualquer evento adverso, confirmado ou sob suspeita, relacionado à segurança de sistemas de computação ou de redes de computadores.
  - Alguns exemplos de incidentes de segurança são: tentativa de uso ou acesso não autorizado a sistemas ou dados, tentativa de tornar serviços indisponíveis, modificação em sistemas (sem o conhecimento ou consentimento prévio dos donos) e o desrespeito à política de segurança ou à política de uso aceitável de uma organização.

- **Ciclo de vida do incidente**



Conceitos e definições

- **Gestão de incidentes de segurança da informação**
  - Convém que responsabilidades e procedimentos de gestão sejam estabelecidos para assegurar respostas rápidas, efetivas a incidentes de segurança da informação.

- **Notificação de fragilidades e incidentes de segurança da informação**
  - Assegurar que fragilidades e eventos de segurança da informação associados com sistemas de informação sejam comunicados, permitindo a tomada de ação corretiva em tempo hábil.

- **Time de respostas a incidentes de segurança da informação (CSIRT)**
  - É o nome dado à organização responsável por receber, analisar e responder a notificações e atividades relacionadas à incidentes de segurança da informação.

Conceitos e definições

- **Danos**
  - São as consequências de um incidente. Os danos podem ser diretos ou indiretos.
    - **Danos diretos** – são consequências diretas do incidente.
      - Exemplo: furto de um veículo.
    - **Danos indiretos** – são consequências indiretas do incidente.
      - Exemplo: após o furto do veículo, a pessoa perder compromissos.

- **Impacto**
  - Mudança adversa no nível obtido dos objetivos do negócio.

Medidas de segurança

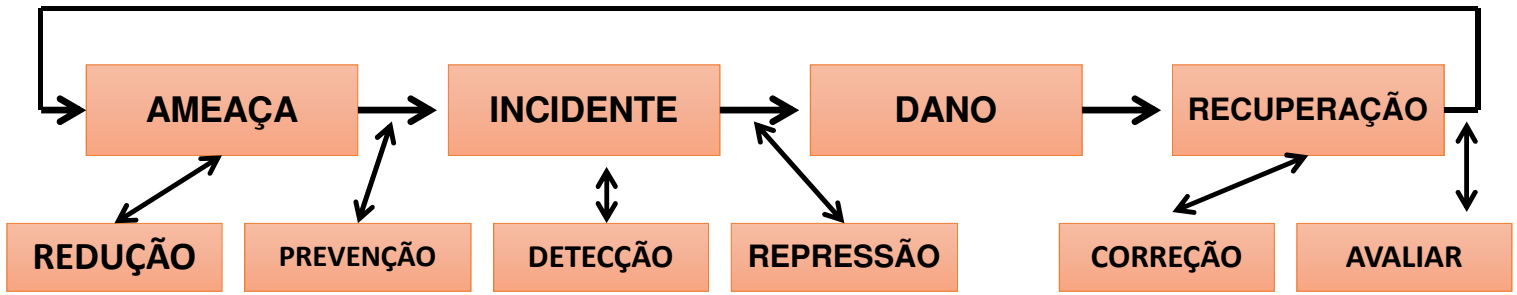
- **Funções das medidas de segurança**
- **Redutivas**
  - Medidas redutivas são aquelas destinadas a reduzir a probabilidade de que um incidente ocorra.
    - Exemplo: Instalação de um antivírus.
- **Preventivas**
  - Medidas preventivas são aquelas cujo objetivo é evitar a exploração de uma vulnerabilidade. Visam evitar o risco e reduzir a zero a probabilidade de ocorrência de um incidente, eliminando também a atividade geradora do risco.
    - Exemplo: uso de um sistema de chave de acesso (crachá) ou o armazenamento de informações sigilosas em um cofre.
- **Detectivas**
  - As medidas detectivas são aquelas que procuram identificar um incidente no momento em que ele ocorre.
    - Exemplo: sistema de detecção de instrusão.

Medidas de segurança

- **Funções das medidas de segurança (cont.)**
- **Repressivas**
  - As medidas repressivas são aquelas que combatem o dano causado pelo incidente.
    - Exemplo: combate a um incêndio.
- **Corretivas**
  - As medidas corretivas, ou de recuperação, visam a restauração do ambiente após um incidente de segurança. As medidas corretivas são importantes para que as operações da organização voltem à normalidade após um incidente.
    - Exemplo: restaurar o banco de dados usando backup.

Medidas de segurança

- **Medidas de segurança x ciclo de vida do incidente**



Medidas de segurança

- **Segurança física e do ambiente**
  - A segurança física e do ambiente tem por objetivo previnir o acesso físico não autorizado, danos e interferências com as instalações e informações da organização.

- **Perímetro de segurança física**
  - Proteção contra acesso físico não autorizado;
  - Barreiras como paredes , portões de entrada, portas, fechaduras, etc;
  - Alarmes;
  - Sistemas de deteccão de intrusos.

- **Controles de entrada**
  - Assegurar que somente pessoas autorizadas tenham acesso;
  - Identificações, registro de entrada e saída, crachás, etc.

- **Segurança em escritórios, salas e instalações**

Medidas de segurança

- **Segurança física e do ambiente (cont.)**
- **Proteção contra ameaças externas e do meio ambiente**
  - Convém que sejam levadas em consideração todas ameças à segurança representadas por instalações vizinhas.

- **Segurança de equipamentos**
  - Impedir perdas, danos, furto ou roubo, ou comprotimento de ativos e interrupção das atividades da organização;
  - Monitoramento das condições ambientais, como temperatura e umidade;
  - Proteção contra raios;
  - Proteção contra falta de energia (no-breaks, UPS, geradores de emergência, etc.).

- **Segurança do cabeamento**
  - Evitar interferências;
  - Cabos de energia segregados dos cabos de comunicações;
  - Blindagem eletromagnética para proteção dos cabos;
  - Piso elevado.

Medidas de segurança

- **Segurança em recursos humanos**
  - Tem por objetivo estabelecer diretrizes e controles para a implementação de uma efetiva gestão de segurança em recursos humanos.
- **Antes da contratação**
  - Seleção
    - Verificações do histórico de todos os candidatos a emprego (Referências, informações do currículo, confirmação das qualificações acadêmicas e profissionais, verificação independente da identidade e atestado de Antecedentes Criminais).
  - Termos e condições de contratação
    - Assinatura de um termo de confidencialidade.
- **Durante a contratação**
  - Conscientização, educação e treinamento em segurança da informação;
  - Processo disciplinar.
- **Encerramento da contratação**
  - Devolução de ativos;
  - Retirada de direito de acesso.

Medidas de segurança

- **Proteção contra códigos maliciosos**
  - Visa proteger a integridade do software e da informação por meio da implantação de controles de detecção, prevenção e recuperação contra códigos maliciosos e conscientização de usuários.

- **Diretrizes para implantação**
  - Proibir o uso de softwares não autorizados;
  - Instalar e atualizar regularmente softwares de detecção e remoção de códigos maliciosos;
  - Estabelecer planos de continuidade do negócio para recuperação em casos de ataques por códigos maliciosos;
  - Conscientização dos usuários.

Medidas de segurança

- **Descarte de mídias**
  - Garantir que as mídias sejam descartadas de forma segura e protegida quando não forem mais necessárias para minimizar o risco de vazamento de informacões sensíveis para pessoas não autorizadas.

Medidas de segurança

- **Cópias de segurança**
  - Convém que as cópias de segurança sejam efetuadas e testadas regularmente conforme a política de backup.
  - Convém que as cópias de segurança sejam armazenadas em uma localidade remota.
  - Convém que os procedimentos de recuperação sejam testados e verificados regularmente.

Medidas de segurança

- **Autenticação segura**
  - A autenticação, apesar de ser também utilizada para controle de acesso lógico, é um instrumento indispensável na segurança física. A autenticação se dá através de um ou mais fatores dentre os três a seguir:
    - O que você sabe? – Nome de usuário e senha;
    - O que você tem? – Cartão, crachá, *smartcards* e *tokens*;
    - Quem você é? – Dispositivo biométrico.
  - Para termos uma autenticação considerada segura, aconselha-se a utilização de, no mínimo, dois requisitos de autenticação agregados. Exemplos:
    - ID + Senha + Crachá;
    - Crachá + Biometria.

Medidas de segurança

- **Segregação de funções**
  - Prega a divisão de tarefas e permissões na organização, não concentrando o conhecimento em apenas uma pessoa, reduzindo, consequentemente, o risco de fraudes, uma vez que seriam necessários dois ou mais colaboradores para que essa se consumasse.

- **Gerenciamento de acesso**
  - Manter um controle efetivo sobre os direitos de acesso necessários para que os colaboradores exerçam suas atribuições, sem que lhes seja concedido nenhum direito além do necessário.

Medidas de segurança

- **Gestão de mudanças**
  - Modificações em equipamentos, sistemas operacionais e aplicativos devem ser devidamente controladas. Em particular, devem ser considerados os seguintes items:
    - Identificação e registro das mudanças significativas;
    - Planejamento e testes de mudanças;
    - Avaliação de impactos;
    - Comunicação dos detalhes das mudanças para todas as pessoas envolvidas;
    - Procedimento formal de aprovação das mudanças;
    - Procedimentos de recuperação.

Medidas de segurança

- **Monitoramento**
  - Tem por objetivo detectar atividades não autorizadas de processamento da informação. Inclui itens como:
    - Registros de auditoria;
    - Monitoramento do uso dos sistemas;
    - Proteção das informações de registro (*log*);
    - Registro (*log*) de falhas;
    - Sincronização dos relógios.

Medidas de segurança

- **Gerenciamento de acesso do usuário**
  - Assegurar o acesso de usuário autorizado e previnir acesso não autorizado a sistemas de informação.
  - Convém que procedimentos formais sejam implementados para controlar a distribuição de direitos de acesso a sistemas de informação e serviços.
  - Inclui itens como:
    - Registro de usuário;
    - Gerenciamento de privilégios;
    - Gerenciamento de senha do usuário.

Medidas de segurança

- **Responsabilidades dos usuários**
  - Garantir que os usuários estejam conscientes de suas responsabilidades para manter efetivo controle de acesso, principalmente em relação ao uso de senhas e equipamentos.
    - Uso de Senhas;
    - Política de mesa limpa e tela limpa.

- **Controle de acesso a rede**
  - Prevenir acesso não autorizado aos serviços da rede.
    - Política de uso dos recursos da rede;
    - Autenticação para conexão externa do usuário;
    - Segregação de redes.

Medidas de segurança

- **Vulnerability Management Foundation**
- **Gestão de vulnerabilidades técnicas**
  - A Gestão de vulnerabilidades técnicas tem por objetivo reduzir riscos resultantes da exploração de vulnerabilidades técnicas conhecidas.
  - Convém que seja obtida informação em tempo hábil sobre vulnerabilidades técnicas dos sistemas de informação em uso, avaliada a exposição da organização a estas vulnerabilidades e tomadas as medidas apropriadas para lidar com os riscos associados.
  - A norma ISO/IEC 27002 estabelece que a gestão de vulnerabilidades técnicas seja implementada de forma efetiva, sistemática e de forma repetível com medições de confirmação de efetividade.

Medidas de segurança

- **Gestão de vulnerabilidades técnicas (cont.)**
  - Diretrizes para implementação:
    - Um inventário completo e atualizado dos ativos de informação é um pré-requisito para uma gestão efetiva de vulnerabilidade técnica.
    - Utilização de ferramentas para identificação de vulnerabilidades técnicas.
      - Exemplo: OpenVAS - http://www.openvas.org/ (Ferramenta de varreduras e gerenciamento de vulnerabilidades open-source largamente utilizada).
    - Uma vez que uma vulnerabilidade técnica potencial tenha sido identificada, convém que a organização avalie os riscos associados e as ações a serem tomadas.
    - Se um *patch* for disponibilizado, convém que sejam avaliados os riscos associados a sua instalação (*patches* devem ser testados e avaliados antes de serem instalados).
    - O processo de gestão de vulnerabilidades técnicas deve ser regularmente monitorado e avaliado.

Medidas de segurança

- **Gestão de vulnerabilidades técnicas (cont.)**
  - **Exemplos de vulnerabilidades técnicas**
    - Vulnerabilidade de software
      - Procedimentos de teste de software insuficientes ou inexistentes.
    - Vulnerabilidade de hardware
      - Sensibilidade à variação de temperatura.
    - Vulnerabilidade de rede
      - Conexão de redes públicas desprotegidas.
    - Vulnerabilidade do local ou das instalacões
      - Inexistência de mecanismos de proteção física no prédio, portas e janelas.
    - Vulnerabilidade em recursos humanos
      - Procedimentos de recrutamento e seleção inadequados.

Medidas de segurança

- **Sistema de gestão da segurança da informação (SGSI)**
  - A organização deve estabelecer, implementar, operar, analisar criticamente, manter e melhorar um SGSI;
  - O SGSI deve ser baseado no modelo “Plan-Do-Check-Act” (PDCA).

Política de Segurança da Informação

- **Information Security Policy Foundation**
- **Introdução a Política de Segurança da Informação (PSI)**
  - Uma política de segurança não é apenas um documento contendo instruções de uso de senhas, mas, sim, um documento estruturado que estabelece um conjunto de regras, normas e procedimentos que define as obrigações e as responsabilidades referentes à segurança da informação e deve ser observado e seguido pelos colaboradores da organização, sob pena de advertência e até desligamento por justa causa, no caso do não cumprimento.
  - É considerada como um importante mecanismo de segurança, tanto para as instituições como para os usuários.

Política de Segurança da Informação

- **Objetivos de uma PSI**
  - O objetivo da política de segurança da informação é prover uma orientação e apoio da direção para a segurança da informação de acordo com os requisitos do negócio e com as leis e regulamentações pertinentes.
  - "*Convém que a direção estabeleceça uma clara orientação da política, alinhada com os objetivos do negócio e demonstre apoio e comprometimento com a segurança da informação por meio da publicação e manutenção de uma política de segurança da informação para toda a organização*." (ISO/IEC 27002).
  - Os principais objetivos de uma política de segurança da informação são:
    - Proteger os negócios da organização frente ao impacto de incidentes;
    - Padronizar a segurança da informação dentro da organização;
    - Orientar colaboradores, prestadores de serviço e terceiros a respeito de suas obrigações quanto à segurança da informação.
  - Uma política de segurança atribui direitos e responsabilidades às pessoas em relação à segurança dos recursos computacionais com os quais trabalham.
  - Uma política de segurança também estipula as penalidades às quais estão sujeitos aqueles que a descumprem.

Política de Segurança da Informação

- **Benefícios da adoção de uma PSI**
  - O principal benefício da adoção de uma PSI é o estabelecimento de um padrão de conduta que seja amplamente difundido na organização e que sirva como referência para tomada de decisões da alta direção em assuntos relacionados à segurança da informação.
  - A Política de Segurança da Informação tem a importante função de fornecer as diretrizes para proteger ativos de informação contra ameaças ou incidentes, assegurando:
    - Redução da probabilidade da ocorrência de incidentes por meio da adoção de controles de segurança e diminuição dos riscos;
    - Tratamento imediato de quaisquer violações de segurança da informação detectadas e minimização dos danos provocados;
    - Efetividade dos planos de continuidade de negócios por meio de avaliações, manutenções e testes periódicos;
    - Comprometimento e responsabilidade de todos os funcionários com a PSI, observando as normas de conduta e ética da empresa;
    - Treinamentos e conscientização regulares disponíveis para todos os usuários com acesso ao sistema de informações.

Política de Segurança da Informação

- **Diretrizes para implementação de uma PSI**
  - **Conjunto de regras efetivas e atuais**
    - Uma PSI deve especificar um conjunto de regras efetivas e atuais:
      - Efetivas: as regras precisam ser tangíveis e aplicáveis dentro da realidade da organização no momento de sua efetivação e publicação;
      - Atuais: as regras devem cobrir todos os elementos relativos às novas tecnologias.
  - **Extratificação**
    - A política de segurança da informação deve especificar processos e controles de segurança da informação, em diferentes níveis de detalhamento, com a finalidade de proporcionar a devida segurança da informação.
    - As regras devem ser organizadas de forma hierárquica.
    - Extratificação da PSI em diretrizes, regras e procedimentos.
  - **Responsabilização**
    - Indica as sansões cabíveis em casos de violação ou não observância à PSI.

Política de Segurança da Informação

- **Diretrizes para implementação de uma PSI (cont.)**
  - **Viabilidade**
    - O custo de implantação de todas as exigências da PSI deve ser justificado pelo valor do ativo e do negócio a ser protegido.
    - A política deve estar alinhada com os objetivos do negócio e ser aderente à realidade da organização.
  - **Aplicabilidade**
    - As regras estabelecidas em uma PSI devem ser aplicáveis e implementáveis.
  - **Clareza e objetividade**
    - A linguagem utilizada na redação da PSI deve ser clara, objetiva e concisa, facilitando a leitura e a compreensão. Textos longos podem desestimular a leitura ou suscitar dubiedade na interpretação.
    - "*A política de segurança da informação deve ser um documento simples e de fácil entendimento, pois será lida por todos os colaboradores da organização, de todos os níveis hierárquicos.*" (Campos, 2007)

Política de Segurança da Informação

- **Diretrizes para implementação de uma PSI (cont.)**
  - **Respaldo**
    - O comprometimento da alta direção é indispensável para o sucesso da implementação da PSI.
    - "*A aprovação da direção para as iniciativas de segurança da informação é essencial.*" (Campos, 2007)
  - **Conhecimento**
    - A PSI deve ser amplamente divulgada. As pessoas precisam estar devidamente informadas e conscientizadas sobre a importância do cumprimento das regras, normas e procedimentos estabelecidos na política de segurança da informação.
    - "*Convém que um documento da política da segurança da informação seja aprovado pela direção, publicado e comunicado para todos os funcionários e partes externas relevantes.*" (ISO/IEC 27002)
  - **Obrigatoriedade**
    - O cumprimento da PSI deve ser obrigatório com consequências efetivas em caso de descumprimento.
    - Recomenda-se que os usuários estejam cientes de que existam ou possam existir meios de identificação do descumprimento da PSI.

Política de Segurança da Informação

- **Análise crítica da PSI**
  - *"Convém que a política de segurança da informação seja analisada criticamente, atualizada e aprimorada dentro de intervalos preestabelecidos ou quando ocorrem mudanças significativas que possam comprometer a sua pertinência, adequação e eficácia."* (ISO/IEC 27002)

Política de Segurança da Informação

- **Políticas específicas**
  - A política de segurança pode conter outras políticas específicas, como:
    - **Política de senhas**: define as regras sobre o uso de senhas nos recursos computacionais, como tamanho mínimo e máximo, regra de formação e periodicidade de troca.
    - **Política de backup**: define as regras sobre a realização de cópias de segurança, como tipo de mídia utilizada, período de retenção e frequência de execução.
    - **Política de privacidade**: define como são tratadas as informações pessoais, sejam elas de clientes, usuários ou funcionários.
    - **Política de confidencialidade**: define como são tratadas as informações institucionais, ou seja, se elas podem ser repassadas a terceiros.

Política de Segurança da Informação

- **Políticas específicas (cont.)**
  - **Política de mesa limpa e tela limpa**
    - Política que tem por objetivo evitar que papéis e mídias removíveis fiquem acessíveis a terceiros.
    - Informações sensíveis ou críticas, por exemplo, em papel ou em mídia de armazenamento eletrônico devem ser guardadas em lugar seguro (cofre ou armário) quando não em uso.
    - Documentos que contenham informação sensível devem ser removidos da impresssora imediatamente.
    - Computadores e terminais quando não utilizados devem ser mantidos desligados ou protegidos por mecanismos de travamento de tela e teclado.
    - Controle do uso de copiadoras.
    - Proteção de correspondências e fax.

Política de Segurança da Informação

- **Política de uso aceitável (PUA) ou Acceptable Use Policy (AUP)**: também chamada de "Termo de Uso" ou "Termo de Serviço", define as regras de uso dos recursos computacionais, os direitos e as responsabilidades de quem os utiliza e as situações que são consideradas abusivas.
- A política de uso aceitável costuma ser disponibilizada na página *Web* e/ou ser apresentada no momento em que a pessoa passa a ter acesso aos recursos. Algumas situações que geralmente são consideradas de uso abusivo (não aceitável) são:
  - compartilhamento de senhas;
  - divulgação de informações confidenciais;
  - envio de boatos e mensagens contendo *spam* e códigos maliciosos;
  - envio de mensagens com objetivo de difamar, caluniar ou ameaçar alguém;
  - cópia e distribuição não autorizada de material protegido por direitos autorais;
  - ataques a outros computadores;
  - comprometimento de computadores ou redes.

Política de Segurança da Informação

- **Código de conduta**
  - Dentro das organizações, utiliza-se o código de conduta como uma forma de direcionar as atitudes dos colaboradores para que estejam em conformidade com a conduta esperada pela alta gestão. Para que seja efetivo, o código precisa ser publicado e divulgado constantemente, desde o momento da contratação até o desligamento.
  - Para que a política de segurança seja igualmente inserida no cotidiano dos colaboradores, ela deve ser inserida no código de conduta da empresa, tornando-se parte do conjunto de diretrizes que todo colaborador deve seguir para atender aos requisitos da organização.

Política de Segurança da Informação

- **Considerações finais**
  - O desrespeito à política de segurança ou à política de uso aceitável de uma instituição pode ser considerado como um incidente de segurança e, dependendo das circunstâncias, ser motivo para encerramento de contrato (de trabalho, de prestação de serviços, etc.).
  - Se a política de segurança da informação for distribuída fora da organização, convém que sejam tomados cuidados para não revelar informações sensíveis.

Classificação da informação

- A classificação da informação tem por objetivo assegurar que a informação receba um nível adequado de proteção.
- Convém que a informação seja classificada em termos do seu valor, requisitos legais, sensibilidade e criticidade para a organização.
- A classificação dada a informação é a maneira de determinar como a informação vai ser tratada e protegida.
  - A classificação da informação não é necessariamente fixa;
  - Documentos de outras organizações devem ser reclassificados;
  - Muitos níveis de classificação podem deixar o processo complexo e economicamente inviável;
  - O proprietário da informação deve ser o responsável pela sua classificação e análise crítica;
  - Rótulos e tratamento da informação: definir e implementar um conjunto de procedimentos para rotulação e tratamento da informação segundo o esquema de classificação adotado pela empresa.
- **Níveis de classificação**
  - Pública;
  - Interna;
  - Restrita;
  - Confidencial.

Gestão da continuidade do negócio

- **Gestão da continuidade do negócio**
  - A gestão da continuidade do negócio tem por objetivo não permitir a interrupção das atividades do negócio, proteger os processos críticos contra efeitos de falhas ou desastres significativos e assegurar a sua retomada em tempo hábil.

- **Planos de continuidade do negócio**
  - Os planos de continuidade do negócio devem ser testados e atualizados regularmente, de forma a assegurar a sua permanente atualização e efetividade.
  - Os planos de continuidade do negócio devem incluir o plano de recuperação de desastres e o plano de gerenciamento de crise.

Conformidade com requisitos legais

- **Direitos de propriedade intelectual**
  - Convém que sejam adotados procedimentos apropriados para assegurar a proteção aos direitos de propriedade intelectual.

- **Proteção de registros organizacionais**
  - Registros organizacionais devem ser protegidos contra perda, destruição e falsificação.

- **Proteção de dados e privacidade de informações pessoais**
  - Convém que privacidade e a proteção de dados sejam asseguradas conforme exigido nas legislações, regulamentações e, se aplicável, nas cláusulas contratuais pertinentes.

Legislações e regulamentações

- **PCI DSS**
  - Conjunto de requisitos de segurança desenvolvido para proteger os dados de portadores de cartão de crédito.

- **Lei Sarbanes-Oxley**
  - Lei promulgada pelo governo norte-americano que tem por objetivo estabelecer maior responsalibidade e transparência na divulgação de informações financeiras por parte dos executivos.

Legislações e regulamentações

- **GDPR (*General Data Protection Regulation*) - Regulamentação Geral de Proteção de Dados – Europa**
  - Entrou em vigor no dia 25/05/2018 em todos os países da União Europeia.
  - A GDPR aplica-se a toda e qualquer organização que ofereça bens ou serviços que coletem dados pessoais de residentes da União Europeia. Não importa onde a empresa esteja situada.
  - Empresas de outros países, como o Brasil, que possuem negócios - e que armazenem dados - na Europa, também devem estar em conformidade com a nova regulamentação.
  - Estabelece a exigência de consentimento do usuário antes que qualquer dado pessoal seja coletado. Esse consentimento deverá ser claro, específico e facilmente revogado a qualquer momento.
  - Qualquer serviço de internet que direta ou indiretamente colete dados pessoais de usuários deverá informar sobre isso de forma clara, acessível e transparente.
  - Além disso qualquer pessoa terá direito a acessar, corrigir ou apagar suas informações pessoais, bem como solicitar a interrupção da coleta de dados a qualquer momento.
  - Em caso de violação, as multas podem chegar a € 20 milhões ou 4% do faturamento global da companhia - o que for maior.

Legislações e regulamentações

- **Lei 12.737, de 30 de novembro de 2012 (Lei "Carolina Dieckmann")**
  - Lei que torna crime no Brasil invadir dispositivo informático alheio, conectado ou não à rede de computadores, mediante violação indevida de mecanismo de segurança e com o fim de obter, adulterar ou destruir dados ou informações sem autorização expressa ou tácita do titular do dispositivo ou instalar vulnerabilidades para obter vantagem ilícita.

- **LEI Nº 12.965, DE 23 DE ABRIL DE 2014 (Marco Civil da Internet)**
  - Estabelece princípios, garantias, direitos e deveres para o uso da internet no Brasil.

Simulados

1. Identifique nas alternativas abaixo a etapa na qual se inicia o ciclo de vida da informação.

a. Descarte

b. Manuseio

c. Transporte

d. Armazenamento

Simulados

2. Selecione nas alternativas abaixo o termo atribuído a fragilidade de um ativo ou grupo de ativos que pode ser explorada por uma ou mais ameaças.

a. Vulnerabilidade

b. Impacto

c. Ameaça

d. Dano

Simulados

3. Selecione nas alternativas abaixo o conjunto de requisitos de segurança desenvolvido para proteger os dados de portadores de cartão de crédito.

a. PCI DSS

b. PGP

c. SSH

d. WPA

Cloud Security Foundation

Objetivo do módulo

- Este módulo tem por objetivo preparar os participantes para realizar o exame da certificação Cloud Security Foundation.

Introdução à Computação em Nuvem

- Computação em Nuvem é um modelo que provê acesso sob demanda via rede de computadores a um conjunto compartilhado de recursos computacionais que pode ser rapidamente provisionado e liberado com um mínimo de esforço administrativo ou interação com o provedor de serviços (NIST).

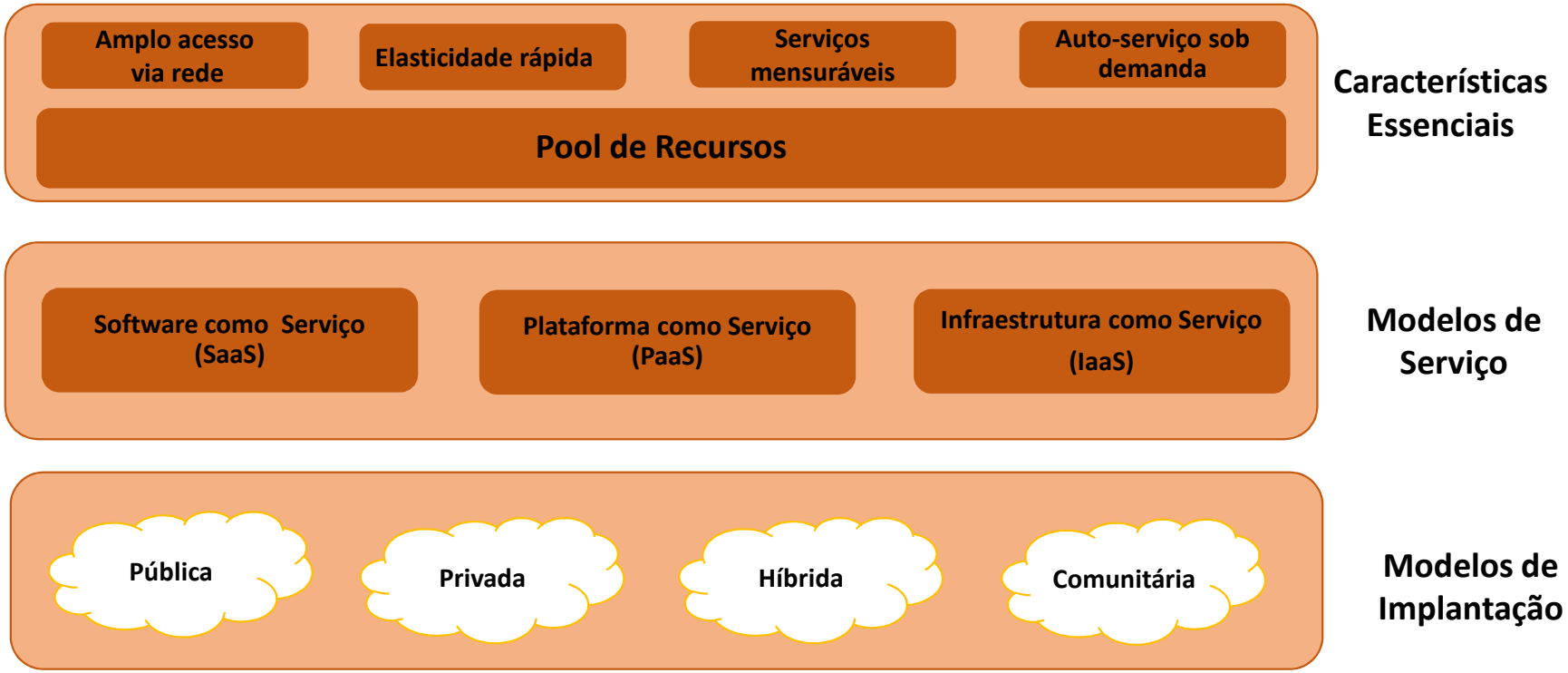
Características essenciais de acordo com o NIST

- **Auto-serviço sob demanda**
  - Usuário pode provisionar recursos computacionais conforme necessidade, sem demandar interação manual do provedor, podendo ser de forma automática ou não.
- **Amplo acesso via rede**
  - Recursos devem estar disponíveis via rede (tipicamente via Internet).
  - Acesso via múltiplas plataformas.
- **Pool de recursos**
  - Recursos físicos e virtuais dinamicamente alocados de acordo com a demanda do usuário. Exemplo: armazenamento, processamento, memória e banda.
- **Elasticidade rápida**
  - Recursos podem ser rapidamente provisionados para atender o aumento da demanda. De forma análoga, recursos podem ser rapidamente desalocados caso não haja demanda.
- **Serviços mensuráveis**
  - Métricas de uso e tarifação.

Modelos de implantação

- **Nuvem pública**
  - Provedor fornece os recursos de computação, como servidores e armazenamento pela Internet.
- **Nuvem privada**
  - Modelo no qual os recursos são utilizados exclusivamente por uma única empresa ou organização.
- **Nuvem híbrida**
  - Modelo que combina nuvens públicas e privadas.
- **Nuvem comunitária**
  - Modelo no qual a infraestrutura da nuvem é compartilhada por grupos de organizações com interesses em comum.

Modelo visual da definição de Computação em Nuvem do NIST



Modelos de serviço

- **Infraestrutura como um serviço (*Infrastructure as a Service* - IaaS)**
  - Fornece infraestrutura computacional (servidores, máquinas virtuais, armazenamento, redes e sistemas operacionais) como serviço de forma provisionada e gerenciada pela Internet.
    - Amazon EC2, Microsoft Azure, etc.
- **Plataforma como um serviço (*Platform as a Service* - PaaS)**
  - Fornece um ambiente sob demanda para desenvolvimento, teste e gerenciamento de aplicativos de software e que permite aos desenvolvedores criarem aplicativos, sem se preocupar com a configuração ou o gerenciamento de infraestrutura de servidores, armazenamento, rede e bancos de dados necessários para desenvolvimento.
  - Salesforce.com
- **Software como um serviço (*Software as a Service* - SaaS)**
  - Permite às empresas fornecer aplicativos de software pela Internet, sob demanda, geralmente, em forma de assinaturas:
    - Gmail, Flickr, Google Apps, etc.

Provedores de Computação em Nuvem

- **Amazon Elastic Compute Cloud**
  - Serviço de Computação em Nuvem da Amazon.
  - https://aws.amazon.com/
- **Microsoft Azure**
  - Plataforma de Computação em Nuvem da Microsoft.
  - https://azure.microsoft.com/
- **Google Cloud Plataform**
  - Plataforma de Computação em Nuvem do Google.
  - https://cloud.google.com/
- **OpenStack**
  - Software *open-source* largamente utilizado para criação de nuvens privadas.
  - Fundado pela Rackspace e NASA.
  - https://www.openstack.org/

Normas, padrões e frameworks

- **ISO/IEC 27017** (disponível em Português)
  - Código de prática para controles de segurança da informação para serviços em nuvem.
  - Fornece diretrizes para os controles de segurança da informação aplicáveis à prestação e utilização de serviços em nuvem.

- **Cloud Security Alliance Guidance v3** (disponível em Português)
  - Guia de boas práticas voltadas para segurança em computação em nuvem.
  - Estabelece diretrizes para operações em computação em nuvem.

- **NIST Special publication 800-144** (Guidelines on Security and Privacy in Public Cloud Computing)
  - Conjunto de diretrizes sobre segurança e privacidade em computação em nuvem.

Aspectos relacionados a segurança na nuvem

- **Conformidade**
  - Leis, regulamentações e normas (ISO/IEC 27001, ISO/IEC 27017, PCI, etc...)
- **Localização geográfica do provedor**
- **Interoperabilidade**
- **Gestão de ativos**
- **Segurança física**
- **Privacidade e proteção dos dados**
- **Criptografia e gerenciamento de chaves**
  - Proteção para dados em trânsito ou armazenados.

Aspectos relacionados a segurança na nuvem

- **Isolamento dos dados**
  - Importante requisito de segurança em relação ao armazenamento dos dados a fim de evitar que dados de diferentes clientes se misturem, especialmente em ambientes de recursos compartilhados em nuvem.
- **Solução de Prevenção de Perdas de Dados (Data Loss Prevention)**
- **Replicação de dados, redundância e tolerância a falhas**
- **Virtualização**
- **Identidade e controle de acesso**
- **Autenticação e autorização**
  - Mecanismo de segurança que tem por objetivo garantir que apenas usuários autorizados tenham acesso a recursos em ambiente de computação em nuvem.

Aspectos relacionados a segurança na nuvem

- **Auditoria**
- **Monitoramento**
- **Resposta a incidentes de segurança**
- **Continuidade dos negócios e plano de recuperação de desastres (*Disaster Recovery Plan*)**
- **Acordo de Nível de Serviço (*SLA – Service Level Agreement*)**
  - Acordo formal, realizado entre o provedor de serviçso em nuvem e o cliente e que estabelece políticas relacionadas a requisitos de disponibilidade de serviço e dos dados e procedimentos de segurança a serem adotados.

Simulados

1. Selecione nas alternativas abaixo a norma ISO/IEC que fornece um código de prática para controles de segurança da informação para serviços em nuvem.

a. ISO/IEC 27017

b. ISO/IEC 38500

c. ISO/IEC 9000

d. ISO/IEC 20000

Simulados

2. Selecione nas alternativas abaixo a característica essencial de Computação em Nuvem que permite que recursos possam ser rapidamente provisionados para atender o aumento da demanda.

a. Elasticidade rápida

b. Disponibilidade

c. Conformidade

d. Auto-serviço sob demanda

Simulados

3. Selecione nas alternativas abaixo o modelo de serviço de Computação em Nuvem que permite as empresas fornecer aplicativos de software pela Internet, sob demanda, geralmente, em forma de assinaturas.

a. SaaS

b. CaaS

c. PaaS

d. IaaS

Referências bibliográficas

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR ISO/IEC 27002. Rio de Janeiro, 2005.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR ISO/IEC 27017. Rio de Janeiro, 2016.
- CAMPOS, André. Sistema de segurança da informação: controlando os riscos. 2. ed. São Paulo: Visual Books, 2007.
- CARTILHA DE SEGURANÇA PARA INTERNET – Cert.br. Disponível em http://cartilha.cert.br/
- Cloud Security Alliance Guidance v3. Disponível em https://chapters.cloudsecurityalliance.org/brazil/files/2017/02/Guia-CSA-v-3.0.1-PT-BR-Final.pdf
- CUKIER, Kenneth; MAYER-SCHONBERGER, Viktor. Big Data - Como Extrair Volume, Variedade, Velocidade e Valor da Avalanche de Informação Cotidiana. 1. ed. Rio de Janeiro: Elsevier, 2013.
- HURWITZ, Judith; NUGENT, Alan; KAUFMAN, Marcia. Big Data para Leigos. 1. ed. Rio de Janeiro: Alta Books, 2016.
- NIST Special publication 800-144 (Guidelines on Security and Privacy in Public Cloud Computing). Disponível em http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-144.pdf
- PROVOST, Foster; FAWCETT, Tom. Data Science para Negócios. 1. ed. Rio de Janeiro: Alta Books, 2016.
- RÊGO, Bergson Lopes. Gestão e Governança de Dados. 1. ed. Rio de Janeiro: Brasport, 2013.
- SÊMOLA, Marcos. Gestão de Segurança da Informação - Uma visão executiva. 2. ed. Rio de Janeiro: Elsevier, 2013.