



## UNIDADE II

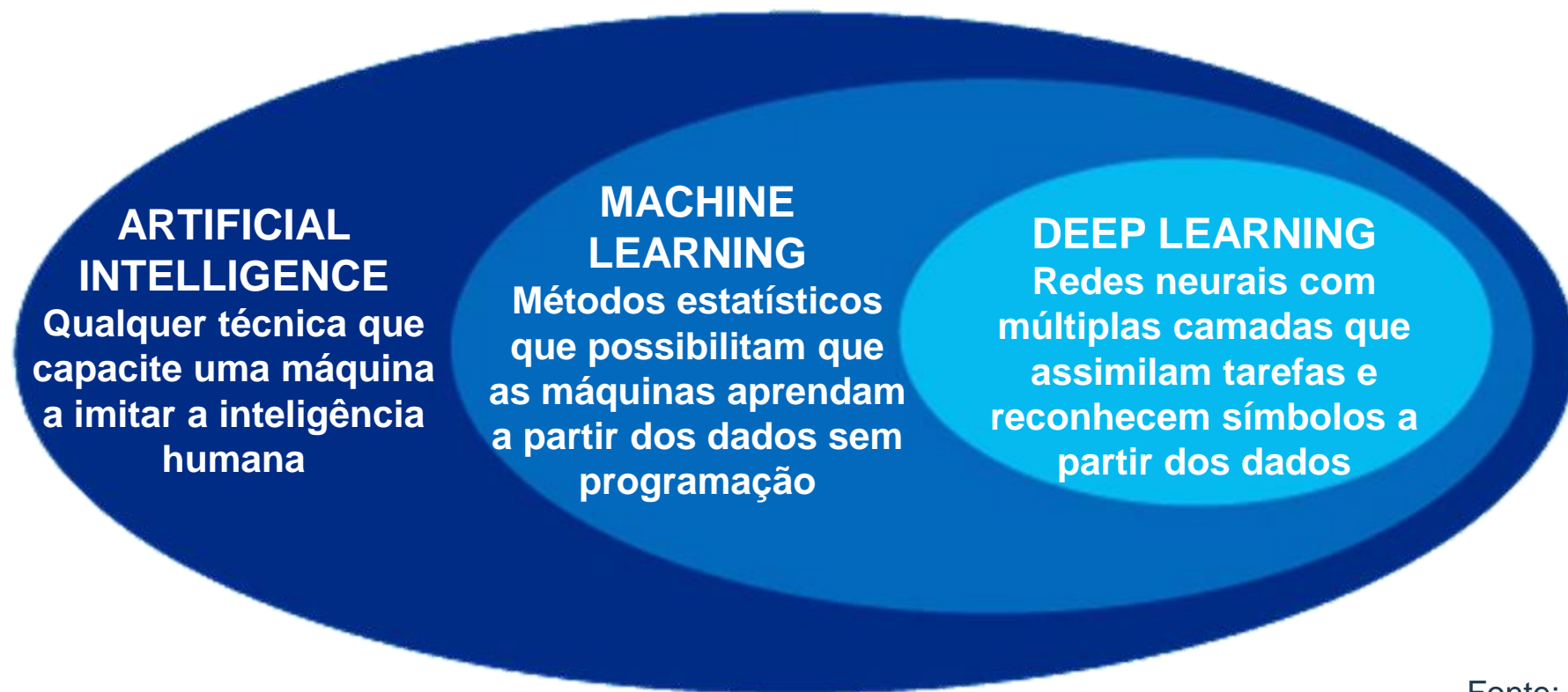
---

Estudos Disciplinares  
*Big Data Analytics*

Prof. Dr. Kauê Rosalem

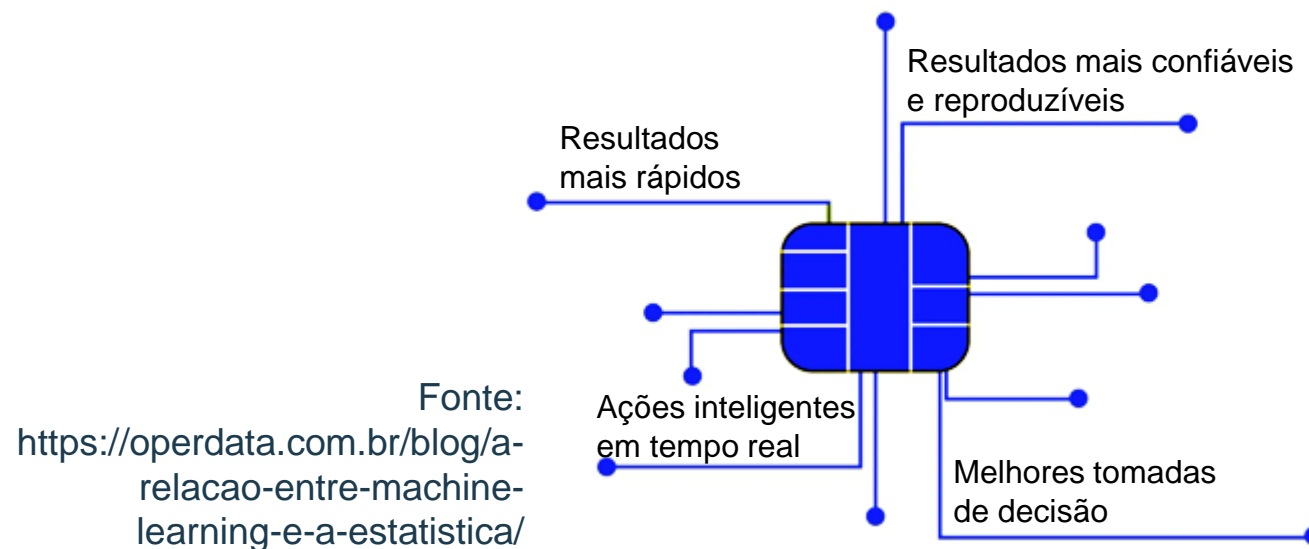
# Big Data e IA

- *Machine learning* e *deep learning* são métodos para treinar modelos, ou seja, mostrar às máquinas como elas devem aprender a classificar dados;
- A Inteligência Artificial, por sua vez, engloba os dois e pode ser definida como qualquer técnica que capacite uma máquina a imitar a inteligência humana.

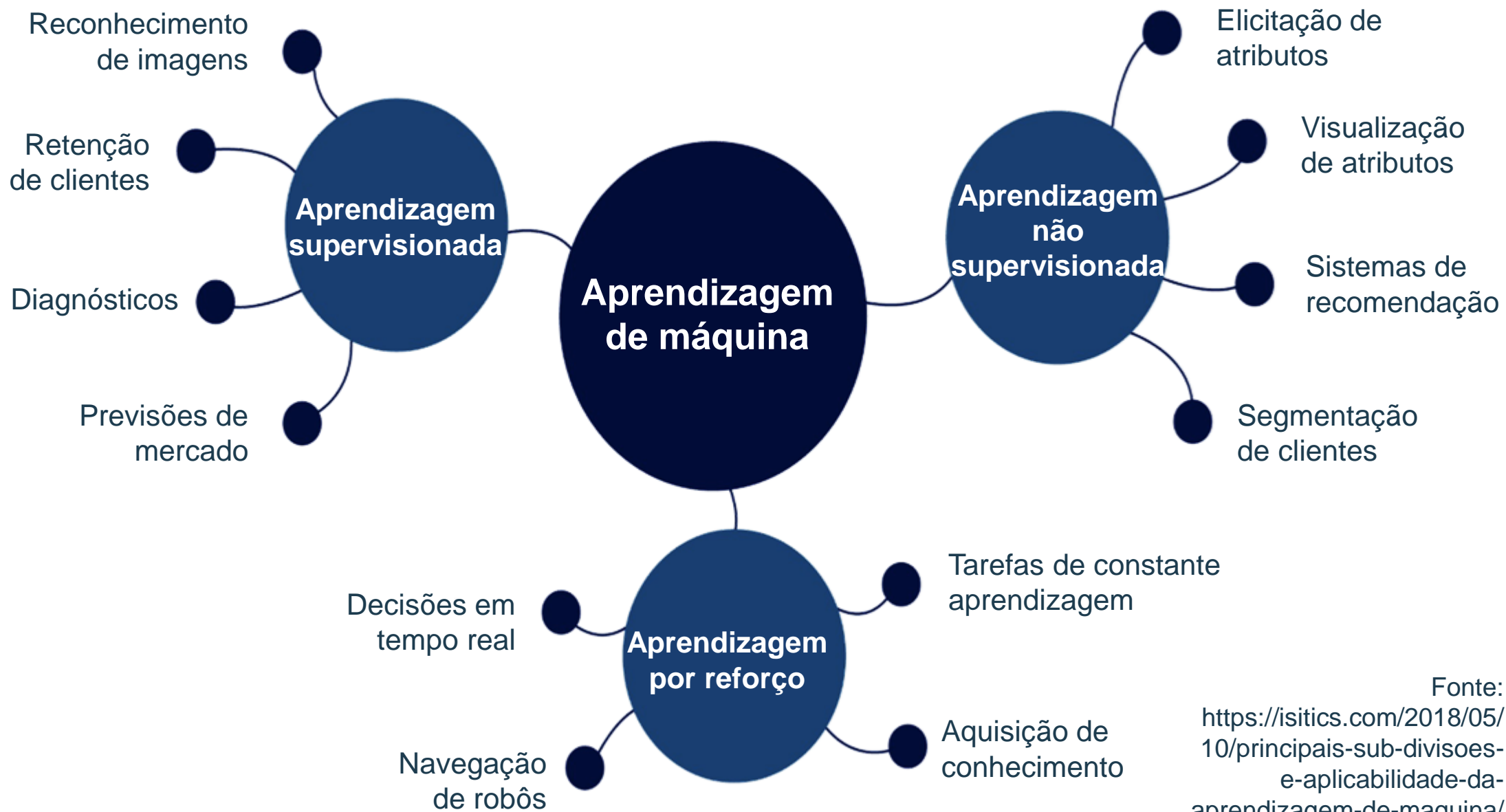


# Aprendizagem de máquina

- **Aprendizagem supervisionada:** o supervisionamento é feito através do treinamento com dados rotulados, ou seja, etiquetados com algum tipo de classe. Por exemplo, produtos podem ser classificados em duas classes: com defeito e sem defeito.
- **Aprendizagem não supervisionada:** ao contrário da supervisionada, aqui não se faz necessário o uso de etiquetas. Aqui são encontrados algoritmos de recomendação, agrupamento e segmentação.
- **Aprendizagem por reforço:** a aprendizagem por reforço é muito utilizada em sistemas de decisão em tempo real. Neste caso, não existe um grande conjunto de dados históricos para que seja gerado um conjunto de treinamento e consequentemente seja possível criar um modelo de aprendizado.



# Aplicabilidades

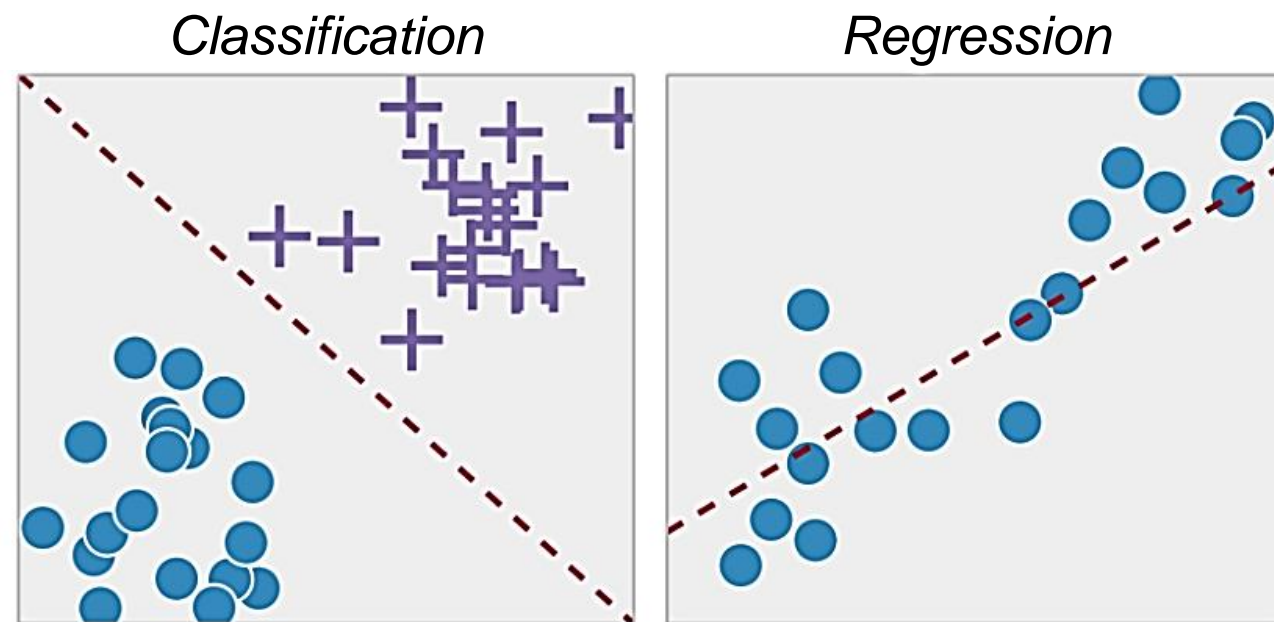


Fonte:  
<https://isitics.com/2018/05/10/principais-sub-divisoes-e-aplicabilidade-da-aprendizagem-de-maquina/>

# Aprendizagem de máquina supervisionada

Exemplos:

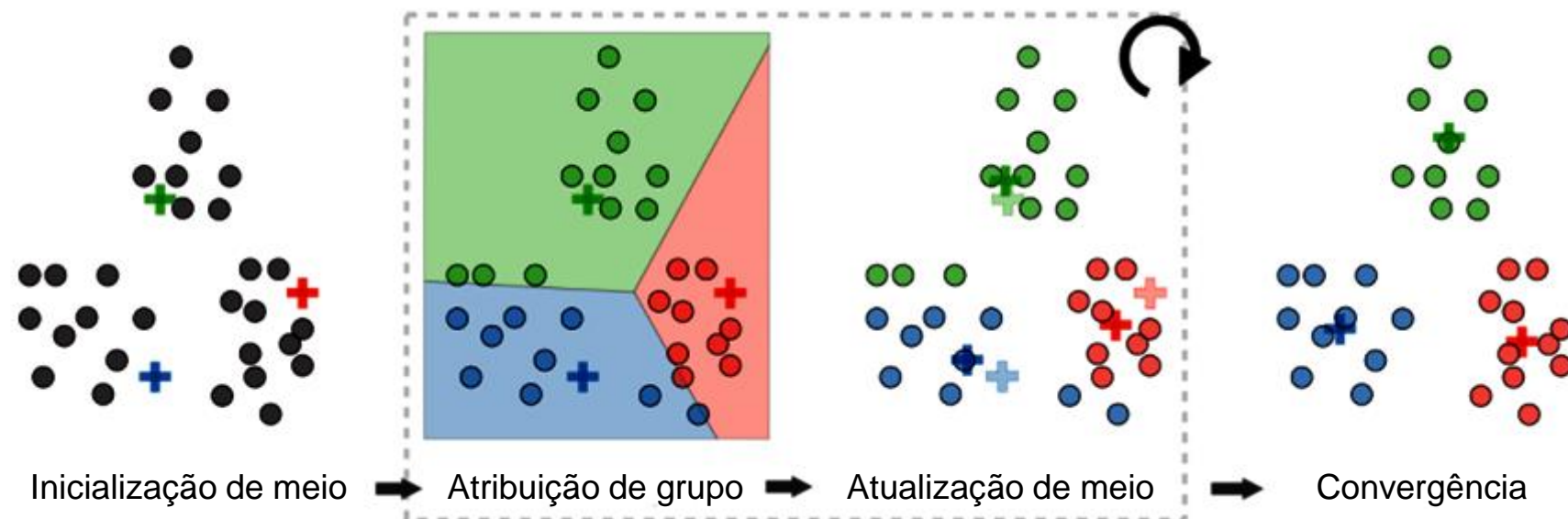
- Regressão: dada uma imagem de homem/mulher, temos de prever sua idade com base em dados da imagem.
- Classificação: dado um exemplo de tumor cancerígeno, temos de prever se ele é benigno ou maligno através do seu tamanho e idade do paciente.



# Aprendizagem de máquina não supervisionada

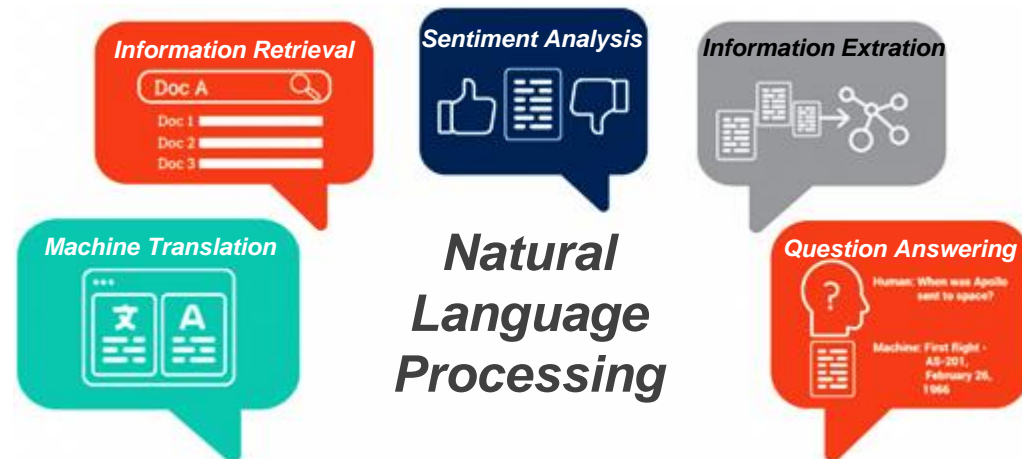
Exemplos:

- “Clustering”: dada uma coleção de 1000 pesquisas de uma universidade encontrar uma maneira de agrupar automaticamente estas pesquisas em um grupos que são de alguma forma semelhantes ou relacionados por diferentes variáveis, tais como a frequência das palavras, frases, contagem de páginas etc.;
- “Algoritmo *Cocktail Party*”, que pode encontrar em uma estrutura de dados desorganizada, como identificar as vozes individuais e música.



# Processamento de linguagem natural (PLN)

- PLN como extensão do *Machine Learning* e da Inteligência Artificial => Evolução da interação homem-máquina;
- Ao tratar línguas escritas e/ou faladas, o PLN faz com que seja possível a interação entre homem e máquina;
- Interpretar e decifrar as mensagens codificadas em linguagem natural => transmitindo-as para as máquinas;
- Para que as máquinas compreendam falas, reproduzam textos ou mesmo consigam conectar palavras de forma lógica, foi necessário aprimorar o processamento de linguagem natural;
- Capacidade de identificar sentimentos e estados de espírito em meio à análise de plataformas virtuais, a fim de conhecer as tendências de interação;
- Rastrear, comparar e combinar dados são alguns dos pontos existentes em PLN.



Fonte:  
<http://www.each.usp.br/petsi/jornal/?p=2577>



# Internet das Coisas (IoT)

- *Internet of Things* (IoT) ou Internet das Coisas: tendência que está diretamente ligada à Inteligência Artificial;
- Aparelhos eletrodomésticos, *gadgets* e máquinas entregam uma interação que otimiza processos e reduz esforços;
- Atividades que não precisam ser realizadas pelo homem sejam assimiladas por máquinas.

## Sua aplicação acontece:

- no controle inteligente do trânsito e no gerenciamento de sinais;
- na automação residencial, com o controle da temperatura do ambiente ou na iluminação;
  - com a segurança, tanto no armazenamento de dados sigilosos quanto no controle de acessos;
  - pelo monitoramento de animais, a partir de microchips de rastreamento.



Fonte: <https://www.grupomult.com.br/iot-comunicacao-de-dados-e-microservicos-uma-visao-integrada-para-suportar-industria-4-0/>



# Lei Geral de Proteção de Dados

- A Lei Geral de Proteção de Dados Pessoais é uma norma federal aprovada em 2018.
- Ela estabelece regras para o uso, coleta, armazenamento e compartilhamento de dados dos usuários por empresas públicas e privadas.
- O principal objetivo é garantir mais segurança, privacidade e transparência no uso de informações pessoais.
- Com a nova legislação, o usuário terá o direito de consultar gratuitamente quais dos seus dados as empresas têm, como armazenam e até pedir a retirada deles do sistema.
- Esses dados podem ser números de documentos como RG, CPF, PIS, endereço ou aqueles considerados pela LGPD como mais “sensíveis” — por exemplo: origem racial ou étnica, filiação à organizações políticas ou religiosas, informações genéticas e de biometria ou de orientação sexual.
  - Vale lembrar que essas características são coletadas de diversas maneiras hoje em dia.
  - É o caso dos APPs de celulares que pedem acesso às informações de usuários e os formulários preenchidos em *sites* de empresas para receber *newsletters* ou ofertas.

# Criação da LGPD

- O aumento dos casos de vazamento de dados nos últimos anos fez com que governos, empresas e sociedade se preocupassem em criar mecanismos para evitar a invasão de privacidade.
- Outro fator relevante é a perda financeira causada por ataques cibernéticos.
- No Brasil, a perda foi de R\$ 80 bilhões de reais em 2019, como informa o levantamento mais recente da União Internacional de Telecomunicações (ITU, na sigla em inglês), órgão da Organização das Nações Unidas (ONU).
- A LGPD foi inspirada no Regulamento Geral de Proteção de Dados da União Europeia (GDPR, na sigla em inglês), criado em 2018, que trata da segurança de informação dos cidadãos europeus.
  - No Brasil, até o momento, não havia legislação específica sobre o assunto, apenas disposições gerais no Código Civil, Código de Defesa do Consumidor, na Lei de Acesso à informação e no Marco Civil da Internet.

# LGPD nas Empresas

- Todas as empresas, sejam PMEs (Pequenas e Médias Empresas) ou de grande porte, terão que atender às exigências da LGPD.
- Uma das mudanças mais importantes é que a nova lei prevê o consentimento expresso dos clientes para o uso das informações.
- Isso significa que as companhias precisarão deixar claro para quem as informações serão usadas.
- Normalmente, os formulários nas páginas de Internet e avisos eletrônicos de empresas públicas e privadas perguntam sobre o consentimento dos usuários.
- Para se adequar à LGPD será necessário mudar a cultura no que diz respeito à gestão dos arquivos, contratação de especialistas e investimento em segurança da informação.
  - É recomendável que a empresa faça um mapeamento e documentação dos dados que já possui e classifique essas informações.
  - Os profissionais de TI representam uma peça-chave no processo de adequação das empresas à LGPD porque são eles que, geralmente, já cuidam dos dados armazenados na nuvem ou em servidores das empresas.

# Interatividade

Existem duas categorias principais de algoritmos de *machine learning*. Marque as categorias corretas.

- a) Positiva e negativa.
- b) Artificial e natural.
- c) Supervisionada e não supervisionada.
- d) Neural e científica.
- e) Matemática e estatística.

# Resposta

Existem duas categorias principais de algoritmos de *machine learning*. Marque as categorias corretas.

- a) Positiva e negativa.
- b) Artificial e natural.
- c) Supervisionada e não supervisionada.
- d) Neural e científica.
- e) Matemática e estatística.

# ***Big Data e Cloud Computing***

- O aumento da computação em nuvem (*Cloud Computing*) tem sido um precursor e facilitador para o surgimento do *Big Data*.
- Embora o *Big Data* traga muitas oportunidades atraentes, as empresas também enfrentam muitos desafios.
- **Coleta dos dados, armazenamento, pesquisa, compartilhamento, análise e visualização.**
- Cada uma destas tarefas requer diferentes abordagens, diferentes níveis de segurança, infraestrutura e profissionais capacitados.
  - O *Cloud Computing* tem sido o suporte necessário para o crescimento na adoção de uma cultura *data driven* que permita às empresas extrair do *Big Data* os *insights* necessários para a tomada de decisão mais consciente.

## ***Big Data e Cloud Computing***

- Assim como o *Big Data* forneceu às organizações terabytes de dados, também apresentou uma questão de gerenciar esses dados em uma estrutura tradicional.
- Na era de conectividade de alta velocidade, mover grandes conjuntos de dados e fornecer os detalhes necessários para acessá-los também é um problema.
- Esses grandes conjuntos de dados geralmente transportam informações confidenciais, como números de cartão de crédito/débito, endereços e outros detalhes, aumentando as preocupações de segurança de dados.
- Problemas de segurança na nuvem são uma grande preocupação para empresas e provedores. Mas as pesquisas mostram que um ambiente em nuvem tende a ser mais seguro que uma rede corporativa.



# Benefícios da computação em nuvem

- A computação em nuvem é uma grande mudança na forma tradicional como as empresas pensam recursos de TI, pois **elimina o gasto capital de compra de *hardware* e *software* e instalação e execução de *datacenters* locais** – *racks* de servidores, eletricidade com disponibilidade permanente para energia e resfriamento, segurança de acesso, *backup*, manutenção e atualização de *software*, etc.
- Isso pode impulsionar os negócios da empresa, que pode então focar apenas em sua atividade-fim.
  - A maior parte dos serviços de computação em nuvem é fornecida por autosserviço e sob demanda para que até grandes quantidades de recursos de computação possam ser provisionadas em minutos, normalmente com apenas alguns cliques, fornecendo às empresas muita flexibilidade e aliviando a pressão do planejamento de capacidade.

# Benefícios da computação em nuvem

- **Os benefícios dos serviços de computação em nuvem incluem a capacidade de dimensionamento elástico.**
- Em termos de nuvem, isso significa fornecer a quantidade correta de recursos de TI, por exemplo, variações da energia de computação, armazenamento e largura de banda, quando necessário e no local geográfico correto.
- **A computação em nuvem exclui a necessidade de muitas dessas tarefas para que as equipes de TI possam usar o tempo delas para alcançar metas de negócios mais importantes (como analisar dados, por exemplo).**

Fonte:  
<https://semantix.com.br/cloud-computing-o-futuro-do-big-data/>



# Tipos de implantação em nuvem

- **Nuvem pública:** nuvens públicas são de propriedade de um provedor de serviços de nuvem e operadas por ele que, por sua vez, fornece recursos de computação, como servidores e armazenamento pela Internet.
  - O Amazon Web Service, Microsoft Azure e Google Cloud são exemplos de nuvem pública.
  - Com uma nuvem pública, todo o *hardware*, *software* e outras infraestruturas de suporte são de propriedade e gerenciadas pelo provedor de nuvem.
  - Você acessa esses serviços e gerencia sua conta usando um navegador *web*.
- **Nuvem privada:** uma nuvem privada se refere aos recursos de computação em nuvem usados exclusivamente por uma única empresa ou organização.
  - Uma nuvem privada pode estar localizada fisicamente no *datacenter* local da empresa. Algumas empresas também pagam provedores de serviço de terceiros para hospedar sua nuvem privada.
  - Uma nuvem privada é aquela em que os serviços e a infraestrutura são mantidos em uma rede privada.

# Tipos de implantação em nuvem

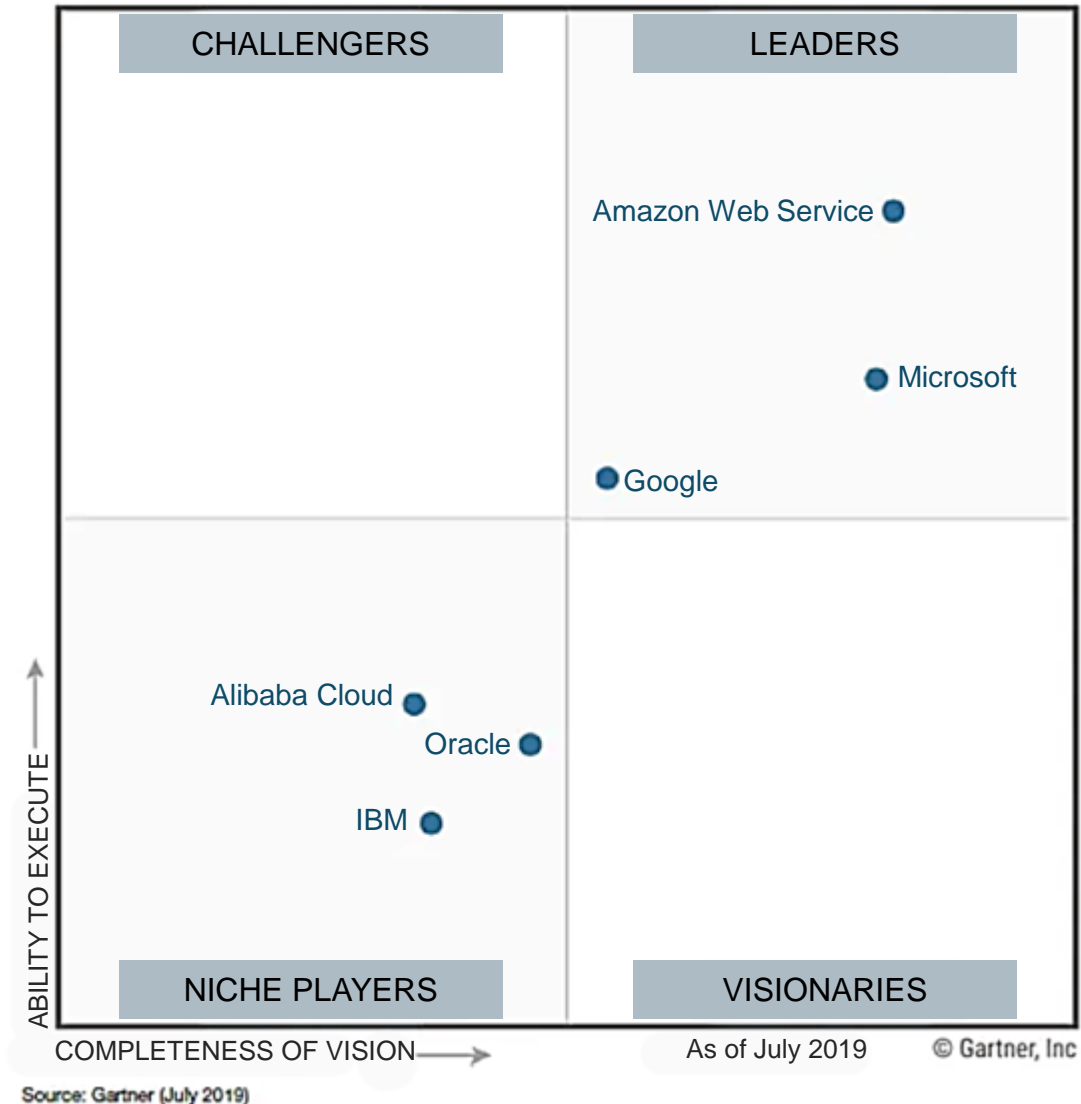
- **Nuvem híbrida:** nuvens híbridas combinam nuvens públicas e privadas ligadas por uma tecnologia que permite que dados e aplicativos sejam compartilhados entre elas.
  - Ao permitir que dados e aplicativos sejam movidos entre nuvens públicas e privadas, a nuvem híbrida dá aos negócios mais flexibilidade e mais opções de implantação.



# Desafios para *Big Data* na computação em nuvem

Os principais desafios na adoção de *Cloud Computing*, especialmente para projetos de *Big Data*, incluem:

- Cultura da empresa que prefere manter os dados “em casa”.
- **Falta de conhecimento ou profissionais capacitados para trabalhar com projetos em nuvem.**
- Resistência da área de TI em “perder importância” dentro da organização (o que é um erro, pois a computação em nuvem libera a TI para focar em outras atividades mais críticas para o negócio).
  - Problemas com largura de banda para transferência de dados entre a rede da empresa e o *Cloud Provider*.



# Oportunidades para *Big Data* na computação em nuvem

- **Redução de custo com infraestrutura de TI para armazenamento de grandes conjuntos de dados.**
- Redução no custo com licença e manutenção de *software*.
- Dados acessíveis de qualquer localidade.
- Acesso a infraestrutura de ponta usada pelos *Cloud Providers*.
- Segurança dos dados normalmente maior do que na rede corporativa.
- Não é necessário se preocupar com detalhes de Engenharia de um *datacenter*, como luz, energia, resfriamento, etc. Isso fica a cargo do *Cloud Provider*.
- **Escalabilidade, com rápida expansão da infraestrutura para acomodar um novo projeto.**
  - Capacidade de dimensionamento elástico.
  - **Foco maior na análise de dados, ao invés de foco no armazenamento.**
  - Redução nos custos de *backup* de dados.
  - **Recuperação de desastre e continuidade dos negócios.**

# Interatividade

A habilidade de pagar pelo uso dos recursos à medida que eles são utilizados é característica da computação em nuvem. Selecione a alternativa que justifique esta vantagem dos serviços em nuvem.

- a) Apresentar uma interface padrão que disponibiliza uma infinidade de serviços.
- b) Esconder as características físicas da plataforma computacional.
- c) Deslocar a localização de toda a infraestrutura computacional.
- d) Sugerir que o usuário tem em suas mãos toda a Internet e os seus serviços.
- e) Permitir que os recursos sejam liberados caso não sejam utilizados, evitando um consumo desnecessário.



## Resposta

A habilidade de pagar pelo uso dos recursos à medida que eles são utilizados é característica da computação em nuvem. Selecione a alternativa que justifique esta vantagem dos serviços em nuvem.

- a) Apresentar uma interface padrão que disponibiliza uma infinidade de serviços.
- b) Esconder as características físicas da plataforma computacional.
- c) Deslocar a localização de toda a infraestrutura computacional.
- d) Sugerir que o usuário tem em suas mãos toda a Internet e os seus serviços.
- e) Permitir que os recursos sejam liberados caso não sejam utilizados, evitando um consumo desnecessário.

# Como o *Big Data* está transformando o recrutamento e o RH

- O recrutamento está passando por muitas mudanças, graças às tecnologias de análise de dados e à automação está se tornando um fator maior no recrutamento inteligente.
- Novas ferramentas estão surgindo para ajudar as equipes de RH a identificar e avaliar os melhores candidatos, plataformas como o LinkedIn e o Glassdoor dão a todos os empregadores, não importa quão grandes ou pequenos, a oportunidade de acessarem dados valiosos.
- Acreditamos firmemente que as equipes de RH que trabalham com dados são as que mais irão recrutar com sucesso nos próximos anos.

Fonte: <https://www.gruposelpe.com.br/>



# Compreender e impulsionar a marca de empregador

- Uma forte “Marca Empregadora” (*Employer Brand*) fará toda a diferença nas tentativas da empresa de manter os funcionários felizes e atrair os melhores talentos.
- Um estudo da Risesmart descobriu que 84% dos funcionários considerariam deixar seu emprego atual e se mudar para um empregador com uma reputação fantástica – mesmo que o aumento salarial não fosse tão grande.
- Então, como dados e análises podem ajudá-lo a desenvolver uma forte marca de empregador?
  - O que você quer que sua empresa represente?
  - Como você deseja que os funcionários se sintam trabalhando para a empresa?
  - O que te faz diferente dos outros empregadores?

# Sucesso da marca de empregador

- A empresa pode, por exemplo, contratar um Cientista de Dados e conduzir uma análise de sentimentos em entrevistas e respostas de pesquisas e postagens de mídia social para determinar o sucesso de sua marca de empregador.
- Ou, se sua empresa passa por grandes mudanças, como uma grande reestruturação, você pode avaliar o sentimento do funcionário antes e depois das alterações para avaliar o impacto na sua marca.
- Pesquisas curtas e anônimas do tipo “*Pulse Surveys*”, podem informar a probabilidade de os funcionários recomendarem a empresa a outras pessoas.
  - Crucialmente, em vez de realizar “pesquisa de clima organizacional” uma vez por ano em uma única e grande pesquisa de pessoal ou fazer a pergunta em entrevistas de saída, “*Pulse Surveys*” permitem que você pergunte aos funcionários uma vez por semana, uma vez por mês ou uma vez por trimestre para ter uma noção mais forte de como eles estão se sentindo durante todo o ano.

# Sucesso da marca de empregador

- No entanto, a marca do empregador não se limita a manter seus funcionários atuais satisfeitos; também é importante ser atraente para pessoas de fora, incluindo ex-funcionários.
- Faz sentido, então, que um Estudo de Transição de Severidade e Mão de Obra tenha mostrado que mais e mais empresas estão extraíndo dados de mídias sociais e *sítes* de avaliação do empregador, como a Glassdoor, após demitir um funcionário.
- Além disso, o *feedback* de qualquer pessoa que tenha deixado a empresa voluntariamente também fornecerá informações úteis sobre a percepção das pessoas sobre a marca da empresa.

# Focando nos melhores canais de recrutamento

- A maioria das empresas usa uma combinação de canais de recrutamento, geralmente incluindo jornais, *headhunters*, campanhas de mídia social, *sites* de emprego *online* e pesquisas no LinkedIn.
- Com uma variedade de canais de recrutamento para escolher, é importante ter uma compreensão clara de quais canais geram o maior retorno do investimento, para que você possa concentrar seu tempo, energia e orçamento de acordo.
- A beleza dos dados é que eles permitem que você teste seus canais de recrutamento e meça sua taxa de sucesso de maneira muito mais precisa.
- Portanto, em vez de focar em indicadores óbvios como o número de currículos obtidos em diferentes canais (que só lhe dizem quantidade, não qualidade), você pode procurar indicadores mais valiosos, como quantas ofertas foram feitas a candidatos de canais específicos.
  - Ou você pode avaliar seus funcionários mais bem-sucedidos em funções específicas e identificar de quais canais eles vieram.
  - O objetivo é direcionar seu recrutamento para que você alcance exatamente os tipos de pessoas que deseja atrair.

## Exemplo de recrutamento utilizando *Big Data*

- Um bom exemplo é a Marriott Hotels e sua impressionante estratégia de recrutamento social.
- A Marriott Hotels possui a maior página de recrutamento no Facebook, com mais de 1,2 milhão de curtidas e milhares de pessoas interagindo com a página a cada semana.
- Obviamente, a página lista as vagas disponíveis, mas também mostra lindamente, por meio de fotos e vídeos, como é trabalhar na empresa.
- A empresa encoraja ativamente o engajamento constante por meio de curtidas e comentários – e essa é uma via de mão dupla, com a Marriott respondendo a comentários.
- Tudo foi projetado para atrair a “pessoa das pessoas” para a marca do empregador da Marriott e mostrar a empresa como um local desejável para trabalhar.



# Como será o futuro do trabalho?

1

## **Flexibilidade permanente:**

Arranjos de trabalho cada vez mais flexíveis.

3

## **Chatbots:**

Seu uso irá crescer, especialmente como uma maneira de ajudar os funcionários no processo de *onboarding*.

5

## **Data driven:**

Um RH cada vez mais orientado por dados, com decisões de contratação mais assertivas.

2

## **Trabalhos temporários:**

Menos empregos de tempo integral e mais profissionais livres, com atuações remotas.

4

## **Gamificação:**

Será usado no recrutamento e seleção, transformando os testes e medindo a personalidade e a habilidade dos candidatos.

6

## **Avaliação de desempenho contínua:**

As empresas estão percebendo que as reuniões individuais entre o gerente e o funcionário ajudam a movimentar os projetos mais rapidamente.

Fonte: <https://www.gupy.io/blog/recursos-humanos>

# Identificando talentos

- Muitos profissionais de RH ou gerentes de contratação provavelmente admitiriam que eles fazem consultas com base no instinto.
- Mas dados e análises estão ajudando os empregadores a eliminar o trabalho de adivinhação do recrutamento e a encontrar pessoas mais adequadas que permanecerão felizes e no cargo por mais tempo.
- Os empregadores de todos os setores estão se voltando para os dados e cada dia precisamos mais de ferramentas para processar os dados de várias maneiras.
- Ferramentas que permitem que os empregadores encontrem a melhor pessoa para qualquer trabalho, com base em suas habilidades, interesses e ações.
  - Além disso, ferramentas de *Big Data* e Inteligência Artificial estão sendo cada vez mais oferecidas por fornecedores como o LinkedIn para filtrar os perfis dos candidatos e identificar as pessoas mais adequadas para uma posição, o que também é bom, considerando que 52% dos líderes de aquisição de talentos afirmam que a parte mais difícil do recrutamento é identificar as pessoas certas de um grande grupo de candidatos.

## Exemplos de identificação de talentos

- A JetBlue Airlines nos oferece um ótimo exemplo de análise de dados sendo usada para encontrar os candidatos mais adequados.
- Anteriormente, a empresa tinha se concentrado em “*niceness*” (em uma tradução livre, algo como “gentileza” ou “ser legal”) como o atributo mais importante para comissários de bordo.
- Então, depois de realizar uma análise de dados de clientes com a Wharton Business School, a JetBlue descobriu que, aos olhos de seus clientes, ser útil é mais importante do que “ser legal” – e até compensa o fato de pessoas que não são tão legais.
  - A empresa foi então capaz de usar essas informações para restringir os candidatos de forma mais eficaz.

# Avaliando talentos

- Ao recrutar um novo candidato, a personalidade é tão importante quanto a qualificação.
- Assim, além de pensar nas habilidades, qualificações e experiência que são ideais para uma posição específica, você também deve, sem dúvida, pensar sobre atributos de cultura, adequação e personalidade.
- Tudo isso pode ser avaliado com precisão nos dias de hoje. É relativamente fácil usar o *software* de análise para filtrar candidatos em potencial e encontrar aqueles com pontos de dados que melhor se ajustam à sua “lista de compras” de atributos ideais – em questão de minutos.
  - É claro, a decisão final da contratação sempre será de um humano, mas dados e análises podem economizar muito tempo, reduzindo o campo de talvez centenas de candidatos para os 10 ou 20 mais adequados. Essa automação de certos processos libera a equipe de RH para se concentrar em outras atividades.
  - Estamos apenas no início desta revolução, em que *Big Data*, Análise de Dados e IA estarão cada vez mais presentes em nossas vidas, pessoais e profissionais.

# Interatividade

Quais os três perfis básicos de profissionais engajados em Ciência de Dados e *Big Data*?

- a) Cientista de requisitos, analista de sistemas, profissionais de tecnologia.
- b) Cientista de dados, analista de sistemas e estatísticos de inferência.
- c) Administradores, gerentes comerciais, publicitários.
- d) Cientista de dados, analista de negócios, profissionais de tecnologia.
- e) Analista de dados, analista de negócios, analista de requisitos.

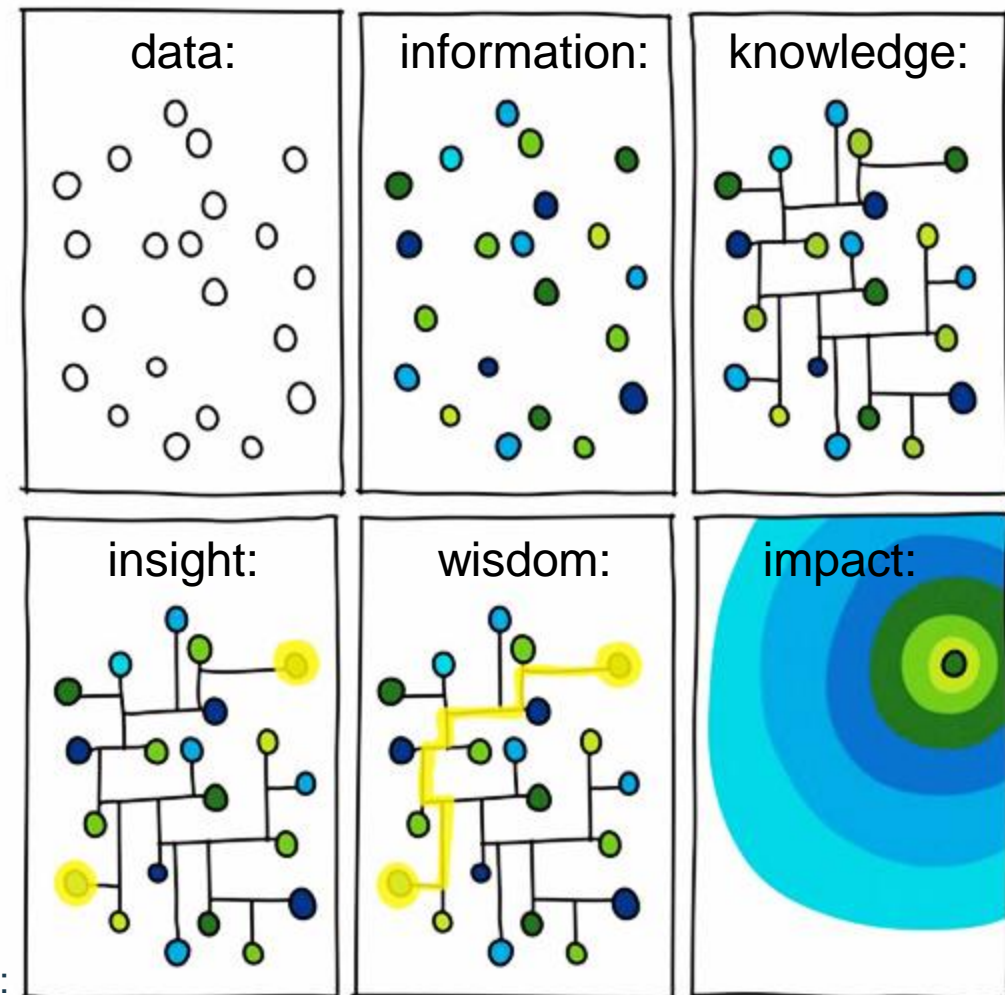
# Resposta

Quais os três perfis básicos de profissionais engajados em Ciência de Dados e *Big Data*?

- a) Cientista de requisitos, analista de sistemas, profissionais de tecnologia.
- b) Cientista de dados, analista de sistemas e estatísticos de inferência.
- c) Administradores, gerentes comerciais, publicitários.
- d) Cientista de dados, analista de negócios, profissionais de tecnologia.
- e) Analista de dados, analista de negócios, analista de requisitos.

# Produtividade das equipes

- As métricas de desempenho têm sido bastante utilizadas nas grandes empresas aliadas ao *big data* para RH. Essas estratégias são aplicadas para melhorar as condições do local de trabalho e impulsionar o desempenho das equipes.
- Por exemplo, ao utilizar um *software* de RH com *big data* é possível estudar os movimentos e as interações entre os funcionários e saber quais são as equipes com melhores vínculos e também quais são as mais eficientes.
- Esses *insights* são valiosos para uma empresa e resultam no aumento da produtividade dos funcionários.



Fonte:

<https://www.gapingvoid.com/>

gapingvoid  
Culture Design Group

@gapingvoid



# Identificação de falhas de processos

- Quanto mais rápido uma falha é descoberta, mais fácil é a sua resolução. Em um mundo ideal, esse cenário seria o perfeito para qualquer empresa.
- Porém, sabemos que detectar problemas não é tão simples assim. Se a organização não contar com a tecnologia em seus processos, esse trabalho fica ainda mais difícil.
- O relacionamento entre funcionários e empresa pode ser bastante delicado, por isso é preciso todo cuidado para que não haja nenhum tipo de ruído.
  - Com a utilização do *big data* no RH, o gestor consegue identificar possíveis falhas e trazer soluções definitivas e adequadas para diferentes casos. Além disso, o conhecimento dos dados permite que seja implementada uma estratégia para que as mesmas divergências não aconteçam no futuro.

# Valorização da história da empresa

- Há uma valiosa oportunidade de usar o *big data* no RH para contar histórias mais ricas e impactantes, o que pode ser um componente essencial para futuros parceiros de negócio, líderes ou funcionários.
- O uso de um *software* de RH para esse fim pode aumentar a transparência e melhorar a cultura e a marca geral da organização.
- Enfim, o *big data* no RH pode oferecer ao departamento a capacidade de tomar decisões baseadas em dados e de forma objetiva para os funcionários e para a própria empresa.
  - Como explicado, a ferramenta ajuda na pesquisa, no armazenamento e na análise de grandes dados, mas cabe à própria organização interpretar cuidadosamente os resultados usando programas apropriados e parâmetros de projeto.

## Outras aplicações

- **Saúde:** em parceria com um consórcio da área de saúde, a IBM está usando seu supercomputador Watson para escanear mutações genéticas e descobrir o melhor tratamento para cada tipo de câncer.
- **Inflação:** em São Francisco, a *startup* Premise paga 700 colaboradores que tiram fotos do preço e disponibilidade de alguns produtos nas prateleiras de 25 cidades na Ásia, América Latina (incluindo o Brasil) e Estados Unidos. Reunidos, os dados são usados para medir inflação.
- **Acabar com engarrafamentos:** a Prefeitura de Dublin fez um acordo com a IBM e usou câmeras e GPS para monitorar o trânsito da cidade afim de evitar congestionamentos e a lotação de transportes públicos.
  - **Rastrear o lixo:** o MIT (Instituto de Tecnologia de Massachusetts, em inglês) convidou 500 moradores de Seattle a etiquetarem o lixo para estudo de logística que visava melhorar o fluxo dos detritos pela cidade.

## Outras aplicações

- **Evitar suicídios:** o Dyrkheim é um projeto que acompanha redes sociais para detectar palavras e frases que caracterizem o autor como um potencial suicida.
- **Substituir currículos:** a Votorantim e outras empresas já estão substituindo a velha análise de currículos pela coleta e observação de dados publicados pelos candidatos sobre si em redes sociais e outros espaços da internet.
- **Estipular preços:** nos EUA, *sites* como o Netflix, lojas, empresas aéreas já usam *big data* para estipular variações no preço dos produtos e serviços que oferecem (Case App Amazon).
  - **Desenvolver vacinas:** a Merck usa *big data* para colher dados que, reunidos e analisados, são usados para identificar padrões que ajudam no desenvolvimento de vacinas pela empresa.

## Outras aplicações

- **Superar a dengue:** a IBM desenvolveu um modelo que articula dados como chuvas, temperatura e acidez do solo para prever surtos de dengue e malária.
- **Eleger candidatos:** nas eleições de 2012, Barack Obama recolheu por 18 meses dados sobre seus eleitores na internet. Isso permitiu uma melhor compreensão dos interesses e preocupações do eleitorado e, provavelmente, o ajudou a ser reeleito presidente dos EUA.
- **Esportes:** nos EUA, a NBA fez um acordo com a SAP e a Stats LLC para oferecer aos telespectadores de jogos de basquete estatísticas detalhadas dos times e jogadores do esporte.
- **Gerir cidades:** no Rio de Janeiro, o Centro de Operações reúne dados recolhidos por câmeras e outros suportes afim de mapear a cidade e apontar locais com problemas – que são repassados a 30 órgãos públicos que podem solucioná-los.
  - **Espionar pessoas:** em 2013, Edward Snowden revelou a existência na NSA do X-Keyscore, que captura todo o conteúdo que trafega na conexão interceptada para, depois, analisá-lo e extrair os dados desejados. Só em 2012, agências federais dos EUA gastaram 5 bilhões de dólares com pesquisas em *big data* – o investimento deve chegar a 8 bilhões de dólares em 2017.

## Outras aplicações

- **Evitar fraudes:** em parceria com a Accenture, uma grande operadora de celular brasileira (que a Accenture não conta qual é por razões contratuais) desenvolveu um sistema baseado em *big data* voltado para a indústria financeira. Ao informar dados relativos à localização de *smartphones*, a novidade pretende inibir as fraudes nas compras com cartão.
- **Atender melhor:** o Bank of America registrou um ganho de produtividade e economizou cerca 15 milhões de dólares após descobrir que funcionários que conversavam mais atendiam mais rapidamente (e, assim, deixá-los conversar).
  - **Ganhar dinheiro:** por meio do programa BankAmeriDeals, o Bank of America devolve a seus clientes parte do dinheiro gasto em compras feitas com cartões de crédito e débito do banco. A vantagem é oferecida pelo banco de acordo com a análise de dados de compras passadas de seus clientes.
  - **Estudar o universo:** o centro de dados do CERN, laboratório suíço que estuda a origem do universo e outros temas, conta com 65 mil processadores que analisam 30 petabytes de dados por ano.

# Newsletters para se manter atualizado sobre *Big Data*

1. Data Science Central
2. Data Science Plus
3. Data Science 101
4. Dataversity
5. Inside Big Data
6. Digital Trends
7. IT World
8. CB Insights
9. Dataflog
10. Data Robot Blog
11. Open AI Blog
12. MIT Technology Review
13. Emerj AI
14. KD Nuggets Blogs
15. Google News Data Science
16. Google News RPA (Robotic Process Automation)
17. Google News Blockchain
18. Kaggle Blog
19. Data Science Dojo
20. Futurism

Fonte: <https://sertms.com/blog/como-funciona-o-big-data-no-rh-entenda-de-uma-vez-por-todas/>





# Interatividade

Considere as seguintes afirmações:

- I. Nos próximos anos viveremos uma escassez de profissionais de Ciência de Dados. Esta escassez ao mesmo tempo em que abre muitas perspectivas profissionais para os que abraçarem a função, também atuará como um entrave, pois dificultará às empresas usarem *Big Data* com eficiência.
- II. Há uma grande procura pelos Cientistas de Dados – remuneração média nos EUA já é superior a 100 mil dólares por ano.
- III. Cientistas de dados fazem parte de uma nova geração de especialistas analíticos que têm as habilidades técnicas para resolver problemas complexos.

Estão corretas:

- a) As afirmações I e II.
- b) As afirmações II e III.
- c) As afirmações I e III.
- d) As afirmações I, II e III.
- e) Apenas a afirmação I.



# Resposta

Considere as seguintes afirmações:

- I. Nos próximos anos viveremos uma escassez de profissionais de Ciência de Dados. Esta escassez ao mesmo tempo em que abre muitas perspectivas profissionais para os que abraçarem a função, também atuará como um entrave, pois dificultará às empresas usarem *Big Data* com eficiência.
- II. Há uma grande procura pelos Cientistas de Dados – remuneração média nos EUA já é superior a 100 mil dólares por ano.
- III. Cientistas de dados fazem parte de uma nova geração de especialistas analíticos que têm as habilidades técnicas para resolver problemas complexos.

Estão corretas:

- a) As afirmações I e II.
- b) As afirmações II e III.
- c) As afirmações I e III.
- d) **As afirmações I, II e III.**
- e) Apenas a afirmação I.

**ATÉ A PRÓXIMA!**