

# data thinking

RICARDO  
CAPPRA  
+

CAPPRA  
INSTITUTE  
FOR DATA  
SCIENCE

20

20





CAPPRA  
INSTITUTE  
FOR DATA  
SCIENCE

# Cappra Institute

A transformação digital está rapidamente mudando o mundo, e se o presente é digital, o futuro inevitavelmente será analítico.

Para enfrentar essa nova realidade, o **Cappra Institute** está ajudando a moldar uma nova geração de pessoas e organizações, através de instrumentos que potencializam a adoção da cultura analítica. Em nossos laboratórios, pesquisamos sobre *Big Data* e criamos formas mais eficientes para pessoas e organizações lidarem com dados, através da tecnologia da informação. Os programas do Cappra Institute ajudam a acelerar negócios data-driven, a aprimorar habilidades analíticas e a democratizar a ciência de dados.

**A missão do Cappra Institute for Data Science é preparar pessoas e organizações para um futuro mais analítico.**

Entre as organizações que já se beneficiaram dos métodos desenvolvidos em nossos laboratórios estão: World Bank, Ambev/ABInbev, Banco Itaú, Rede Globo, Centauro, Gerdau, UOL, Grupo Fleury, Gol Linhas Aéreas, Unilever, Kroton, B3, Turner, Abbott, entre tantas outras espalhadas pelo mundo.

Mais informações em [cappralab.com](http://cappralab.com)







# Ricardo Cappra

Ricardo Cappra é um cientista de dados que pesquisa o impacto dos dados na sociedade e nos negócios, ele lidera uma comunidade global de pesquisadores e cientistas especializados em métodos analíticos.

*Cappra trabalha na criação de métodos científicos, consultorias estratégicas data-driven, aconselhamento na governança da cultura analítica, concepção de laboratórios de dados, participando ativamente do desenvolvimento analítico de algumas das maiores organizações do mundo. Os métodos criados em seus laboratórios são utilizados por Governo dos EUA, Coca-Cola, Banco Santander, Ambev, Whirlpool, Banco Mundial, Rede Globo, Banco Itaú, Unilever, Ambev,...*

Ricardo Cappra fundou em 2017 o *Cappra Institute for Data Science* nos EUA, onde atua como líder de pesquisas e na criação de metodologias analíticas. Em 2015 planejou e lançou o laboratório de big data *Mission Control*, com bases em São Paulo, Porto Alegre e Miami. Fundou em 1998 a *Cappra Consultoria*, que desde lá atua no mercado brasileiro com projetos em tecnologia da informação. Frequentemente é professor convidado de Pós-Graduações, MBAs Internacionais e cursos livres. Autor de artigos, teorias e pesquisas relacionadas a ciência de dados, big data, estratégia, data-driven, temas que realiza Palestras no mundo todo.

Veja o perfil completo, trabalho e projetos em [ricardocappra.com](http://ricardocappra.com)



CAPPRA  
INSTITUTE  
FOR DATA  
SCIENCE

# Casa da Cultura Analítica

Em 2019 inauguramos um laboratório público para as pessoas vivenciarem - e aprenderem - sobre o mundo dos dados. Um espaço que abriga atividades para quem quer compreender, experimentar e aperfeiçoar o conhecimento sobre Big Data.

A Casa tem a missão de democratizar a Cultura Analítica, simplificando o acesso para não técnicos e transferindo conhecimento avançado para especialistas. É um local de encontro para cientistas que utilizam técnicas avançadas, profissionais de negócio que trabalham com dados, executivos que estão transformando seus negócios, estudantes que estão fazendo análises, e tantos outros que buscam o desenvolvimento data-driven.

Mais informações em [cappralab.com/casa](http://cappralab.com/casa)

Endereço:

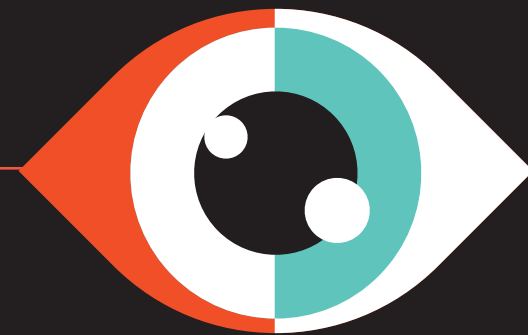
Rua Dom Pedro II, 210, Porto Alegre, RS

Horário de funcionamento:

De segunda à sexta, das 9:00 às 17:00.

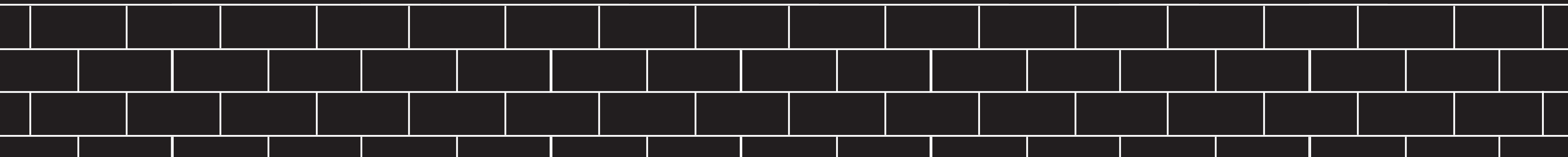






“A tecnologia da informação está aí para todos, aqueles que melhor se utilizarem desses recursos serão os mais preparados para o futuro.”

CAPPRA, Ricardo.







 **data**  
**thinking**  
**2020**

*Os pesquisadores e cientistas do Cappra Institute for Data Science investigam ininterruptamente como os dados estão fazendo parte da vida das pessoas e transformando negócios ao redor do mundo todo. São análises de cenários que nascem em nossos laboratórios, e após muitas horas de pesquisa, dão vida para teorias, metodologias e projetos.*

*Anualmente selecionamos e publicamos aquelas descobertas que terão um forte impacto em um curto espaço de tempo (1 ano), para que empresas e pessoas se preparem de uma forma adequada para essas mudanças.*

*Nossa intenção com essa publicação é levar textos simples e amigáveis para os mais diferentes tipos de públicos - na forma de ensaio - que estimule a reflexão desses temas relacionados com Cultura Analítica. Os ensaios dessa edição 2020 abrangem: Big Data, Ciência de Dados, Data-Driven, Crimes Cibernéticos, Evolução da Cognição, Novos Modelos Organizacionais, Robôs, Times de Dados, ou seja, tudo em torno da Tecnologia da Informação.*

*Entendemos que o futuro inevitavelmente será mais analítico, afinal a transformação digital - que está impactando toda a sociedade e organizações - está deixando um volume absurdo de dados. Não transformar tudo isso em informação é um desperdício, existem muitas oportunidades escondidas nesses amontoados de dados, mas também ameaças que precisamos nos preparar para enfrentar. A tecnologia da informação está aí para todos, aqueles que melhor se utilizarem desses recursos serão os mais preparados para o futuro.*

*Aproveite essas reflexões e divirta-se.*



**Ricardo Cappra** @cappra

Chief Scientist, **Cappra Institute for Data Science** @cappraLAB



# data thinking 2020

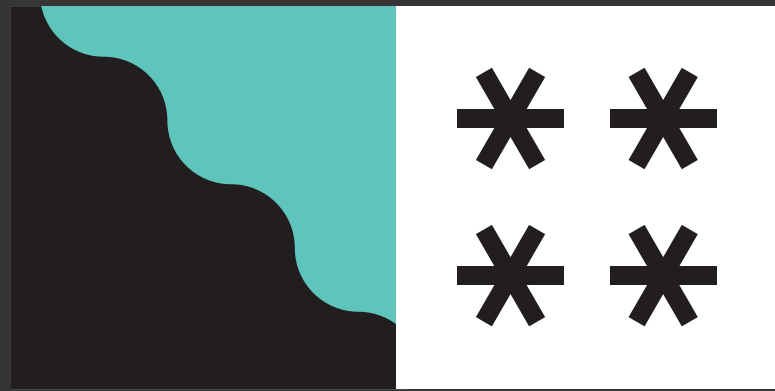
Nossos ensaios sobre o impacto dos dados nos negócios e na sociedades exploraram os seguintes cenários para 2020:

### criminosos analíticos



O LADO SOMBRIO DO  
BIG DATA

### detox de informação



ELIMINANDO O EXCESSO  
DE INFORMAÇÃO

### superpoderes cognitivos



CÉREBROS ARTIFICIAIS PARA  
AJUDAR CÉREBROS ORGÂNICOS

### inforganization




ORGANISMOS ORIENTADOS  
POR INFORMAÇÃO

### times data-driven

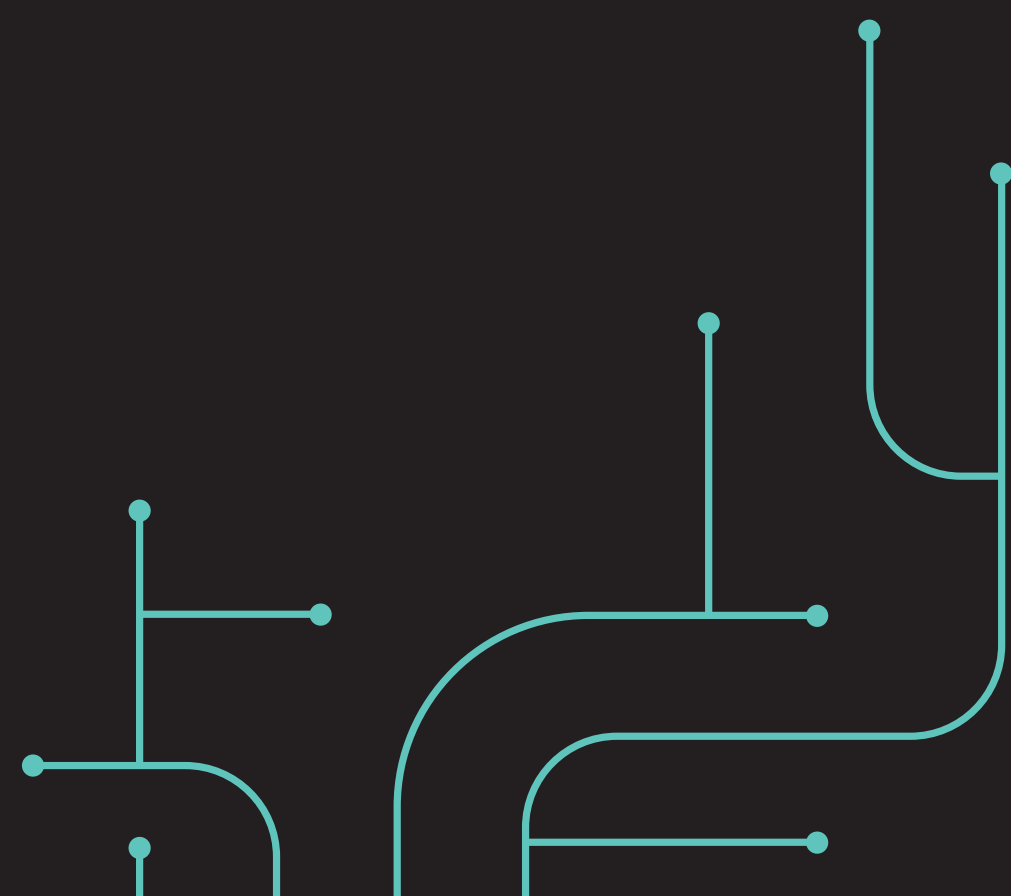


OPERAÇÕES ANALÍTICAS COMO  
DIFERENCIAL ESTRATÉGICO

### data-robots



AUTOMAÇÃO DA  
CIÊNCIA DE DADOS







# criminosos analíticos



O LADO SOMBRIO  
DO BIG DATA

*“Nesse exato momento, criminosos estão roubando seus dados, e analisando a melhor forma de aplicar um golpe. É difícil perceber isso, pois eles são invisíveis nos parâmetros atuais de crimes que conhecemos.”*

CAPPRA, Ricardo.

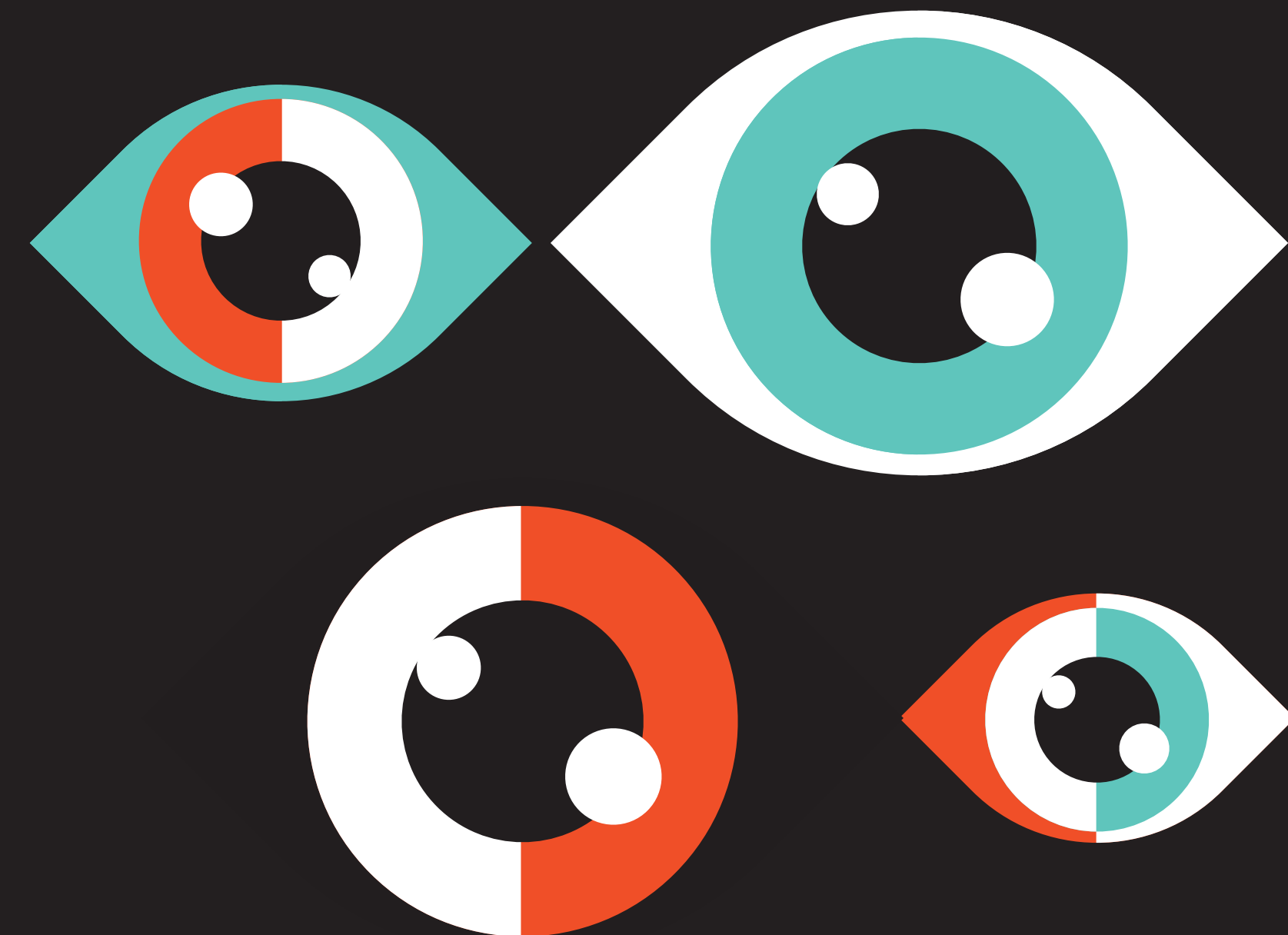


# economia da vigilância

**Imagine um local onde tudo que acontece é monitorado...**

Sensores coletam - ininterruptamente - “pegadas” de comportamento humano durante 24 horas os 7 dias da semana, esses dados capturados incluem suas compras, deslocamentos para o trabalho e até a hora que você deixa seus filhos na escola. Esses monitoramentos continuam enquanto você chega ao trabalho, se conecta no e-mail da empresa e envia uma mensagem para o RH com um planejamento de período para as férias. Enquanto faz isso, você imediatamente visita um site e pesquisa possíveis destinos para sua viagem, envia uma mensagem para seus filhos perguntado qual o local que gostariam de viajar nas férias. Seus filhos imediatamente fazem uma busca na internet por destinos, e, em suas próprias redes sociais, pedem dicas para viagem de férias. Vários amigos deles interagem com aquele post e espalham essa informação no intuito de ajudar nas dicas, levando para várias outras redes privadas.

**Nesse breve roteiro, houve um risco de ocorrência de - no mínimo - 9 crimes virtuais.**





# crimes invisíveis

Atualmente, somente 10% dos crimes nos EUA são associados a categoria de CyberCrimes, sendo que as denúncias mais comuns são sobre:

1o lugar: Não-Pagamento/Não-Recebimento de compras on-line;

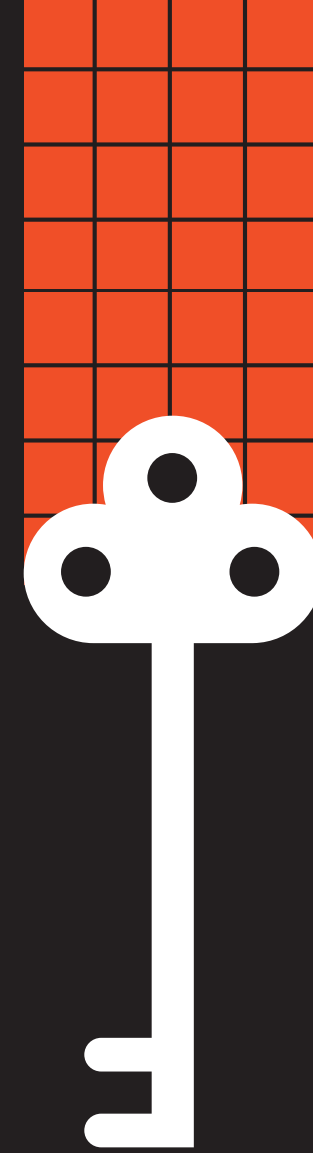
2o lugar: crimes de Extorsão que acontecem na internet ;

3o lugar: roubo de dados pessoais.

No ranking de denúncias de cybercrime por país, nas 3 primeiras posições aparecem EUA, Índia e Canadá, e em 7a posição está o Brasil.

Os cybercrimes não são registrados de forma adequada. Em muitos países os registros policiais não distinguem um crime off-line de um crime on-line, isso obviamente dificulta as análises estatísticas desse tipo de crime. Por isso, quando esses crimes são relatados, geralmente são classificados dentro de categorias tradicionais da justiça, sem diferenciar o meio em que ele foi realizado.

Mas o tipo de investigação necessária nos crimes off-line e on-line são muito diferentes, pois os criminosos virtuais estão adotando algumas técnicas muito avançadas para lidar com grandes volumes de dados dos usuários.



Os criminosos de ambientes digitais estão estudando formas mais eficientes, e menos rastreáveis, para coletar e analisar dados de indivíduos, classificando cada navegação, transação e registro de atividade, aprendendo sobre o comportamento - individual e coletivo - para no momento mais propício cometer algum tipo de crime.

Esses criminosos sabem que se forem mais analíticos poderão cometer um crime com menos riscos de serem flagrados, pois a justiça ainda não possui investigadores especialistas em cybercrime suficientes para rastrear e capturar esse tipo de criminoso. São **Crimes invisíveis**, não categorizados adequadamente nos registros policiais, cometidos por especialistas muito bem treinados em ferramentas de altíssima complexidade técnica.

A maioria desses “novos criminosos” aprendem tudo na internet, gratuitamente, em cursos de ciência de dados avançados, disponíveis em plataformas acadêmicas de reconhecidos institutos ao redor do mundo. Eles estudam as falhas dos sistemas de segurança e se aproveitam da fragilidade dos usuários. Muitos desses crimes podem passar despercebidos, às vezes com valores irrisórios (centavos), mas que pode impactar milhões de pessoas que realizaram transações ao redor do mundo.

**O esforço para roubar uma senha é quase o mesmo para roubar 1 milhão de senhas.**

Através de um spyware instalado em um site é possível coletar informações privadas e abrir um canal para o rastreamento contínuo da vida digital dos usuários da internet.



# vulnerabilidade digital

Para exemplificar, voltamos ao nosso roteiro original para apontar onde ocorreram alguns dos “momentos de vulnerabilidade” dos usuários.

Listamos apenas algumas possibilidades, simulando a fragilidade do usuário em um ato normal no ambiente conectado, nessa situação hipotética, o número de crimes pode ser muito maior.

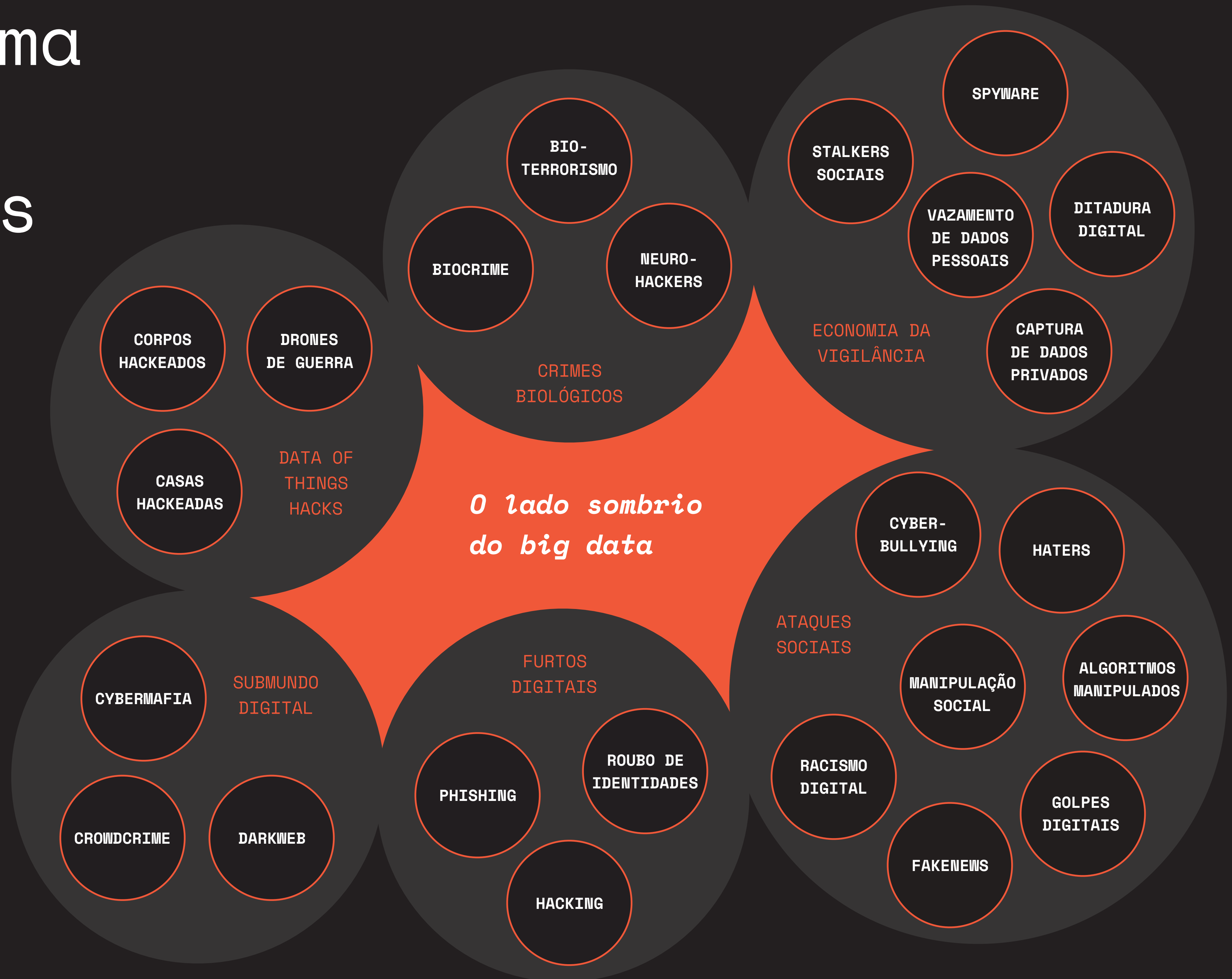
Ação do usuário	Registros de usuários	Fragilidade do usuário	Crimes
Uso de aplicativos de rotas ou serviços de mobilidade urbana	Esses serviços armazenam históricos de navegação e em tempo real registram a localização do usuário	Roubo de senha de acesso para monitoramento em tempo real	Roubo de senha; identificação da localização atual (do usuários e parentes); Uso do serviço indevidamente;...
Serviço de localização do smartphone sempre ativo	Em razão do aplicativo de mobilidade, os serviços de localização normalmente ficam ativados	Coletas de dados de localização a partir de redes wifi e através de navegações realizadas em sites da internet	Roubo de senha; identificação da localização atual (do usuários e parentes); captura de dados de localização em ambientes públicos;...
Identificação de parentes nas redes sociais	Em suas redes sociais, normalmente o usuário identifica seus parentes nos cadastros do seu perfil	Esse tipo de identificação permite um monitoramento dos passos da família em suas redes sociais, além da contínua busca de falhas de segurança para roubos de senhas	Monitoramento indevido da família; roubo de senhas familiares; Captura de registros particulares e invasões de dispositivos móveis;...
Navegação na internet enquanto logado nas Redes Sociais	Com o login em uma rede social por exemplo (se aplica a qualquer serviço que você esteja logado), toda a navegação simultânea é identificada e associada à sua conta	Os dados de navegação dizem muito dos hábitos e até sobre compras dos usuários, permitindo desde captura de senhas até roubos de perfis	Roubo de senhas; roubo de perfis sociais, monitoramento de jornada de navegação; captura de registros particulares;...
Uso de senhas “padrões” em diferentes serviços de internet/apps	Com o monitoramento de registros de navegação e similaridade de senhas é possível coletar todos os tipos de dados de usuários	O roubo de senhas é um canal aberto para monitoramento e constante registro do usuário, permitindo aguardar o momento mais adequado para realização de um crime	Roubo de senhas; vazamento de informações pessoais; análise de perfis para sequestros; interceptações financeiras;...



# ecossistema de crimes analíticos

Mapeamos o ecossistema de crimes digitais, com um breve descritivo de cada um deles.

*Importante ressaltar que existem muitas outras formas de crimes digitais, aqui estão os diretamente relacionados com os criminosos analíticos, ou seja, aqueles que o ponto de partida são os dados.*





# dicionário de crimes virtuais

## [ECONOMIA DA VIGILÂNCIA]

*O controle social baseado na vigilância de indivíduos e grupos*

**Stalkers sociais:** Monitoramento de atividades de indivíduos, que ocorre de forma invasiva e contínua, também pode ser classificado como um crime de perseguição.

**Captura de dados privados:** Invadir contas para captura e armazenamento de dados privados, sejam de empresas ou pessoas físicas, é um tipo de crime que recentemente foi devidamente reconhecido nas leis de proteção de dados e privacidade dos usuários. Essas leis estão surgindo nos últimos anos ao redor do mundo todo, o Brasil terá a sua vigorando a partir de Agosto de 2020.

**Vazamento de dados pessoais:** Divulgar dados pessoais de forma indevida também está referenciado nas leis de privacidade, é ilegal e a responsabilidade é de quem cometeu o armazenamento e o vazamento dessas informações.

**Spyware:** São robôs "invisíveis", que ficam instalados em sites para capturar de forma sistêmica os dados pessoais dos usuários, podem ser capturado senhas, dados privados ou ainda os registros de navegação na internet.

**Ditadura digital:** A ditadura digital não é necessariamente classificada como um tipo de crime pois ela unifica muito das técnicas criminosas que estão nessa lista (spyware, stalkers, manipulação social,...), mas como é algo que está ocorrendo neste momento na China e Rússia, achamos importante colocar nessa lista.

## [FURTOS DIGITAIS]

*Crimes que utilizam técnicas avançadas para roubo na web*

**Phishing:** Sites e emails falsos especialmente criados para roubo de senhas de usuários da internet.

**Roubo de identidade:** Apropriação indevida de documentos e identidades digitais, para realização de crimes, esses ataques costumam ocorrer através de Perfis Sociais Clonados (Fake Profile).

**Hacking:** Apesar desse termo ser muito amplo e abranger quase todos os crimes, aqui estão normalmente presentes as atividades específicas de programação/codificação, ou seja, aqui são construídos os códigos para realização dos crimes. Por isso ouvimos falar tanto dos hackers, que são os construtores de rotinas para viabilização dos crimes.

## [ATAQUES SOCIAIS]

*Ataques de caráter emocional, ou abuso social, à indivíduos ou grupos*

**CyberBullying:** Esse é um tipo de crime de abuso emocional que costuma ser mais perigosos para crianças e jovens.

**Haters:** Um tipo de crime que costuma passar despercebido, a proliferação do ódio costuma se apresentar em forma de agressão escrita nos perfis sociais, mas a justiça em alguns lugares do mundo já vêm considerando como crime esses ataques reais.

**Manipulação social:** A manipulação social geralmente ocorre através de alteração ou influência no conteúdo (por exemplo fake news) que está sendo entregue para cada indivíduo na internet.

**Algoritmos manipulados:** Crimes de manipulação social em escala, que atinge um grande número de pessoas através de softwares desenvolvidos especificamente para esse fim.

**Golpes digitais:** Golpes como extorsão, chantagem e pedidos de recompensa são geralmente aplicados em grupos menos instruídos com relação ao uso da internet, facilitando a aplicação de técnicas e ludibriando os usuários finais.

**Fakenews:** Criação e propagação de notícias falsas é um tipo de crime, pois pode se tornar um mecanismo de manipulação social.

**Racismo digital:** O racismo digital é um dos crimes mais reportados no mundo.

## [DATA OF THINGS HACKS]

*Hackeando a internet das coisas*

**Casas hackeadas:** Invasão de dispositivos conectados no wifi das casas dos usuários é um ato que tem ocorrido continuamente em países mais conectados, o objetivo do criminoso normalmente é capturar dados, controlar dispositivo e monitorar os hábitos pessoais. Por exemplo, entender que uma pessoa chega em casa às 18hrs e hackear o portão da garagem para cometer algum tipo de crime.

**Corpos hackeados:** Quanto mais utilizarmos recursos tecnológicos conectados aos nossos corpos (próteses tecnológicas) maior será o risco de um crime de invasão desses softwares, para os mais diferentes fins, como ameaça de vida e extorsão, por exemplo.

**Drones de Guerra:** Um dos principais instrumentos de guerras atuais são os drones (para vigilância ou mesmo como arma), o uso deles em uma guerra não é necessariamente crime, mas o hacking desse instrumento - algo que acontece em larga escala - transfere o controle dessa arma para um criminoso/hacker.

## [CRIMES BIOLÓGICOS]

*Biologia é tecnologia da informação, isso potencializa crimes de dados*

**BioCrime:** Crimes de pesquisa e manipulação de material genético, por exemplo, ainda são pouco regulados pela justiça.

**BioTerrorismo:** É o terrorismo praticado por meio da liberação ou disseminação intencional de agentes biológicos, vírus ou toxinas; estes podem ser utilizados in natura ou numa forma modificada pelo homem.

**NeuroHackers:** São ataques diretos as funções do cérebro, que usam aparelhos diretamente conectados aos seres humanos.

## [SUBMUNDO DIGITAL]

*Ambientes "invisíveis" da internet, onde os crimes são planejados e estruturados tecnicamente*

**Darkweb:** É uma internet paralela, que é impossível de ser rastreada, normalmente usada apenas por especialistas em técnicas muito avançadas. Nesse espaço acontecem os maiores fluxos de informações ilegais, desde planejamentos de sequestros até ataques terroristas.

**Cybermafia:** Assim como no offline, na internet existem grupos de controle do crime organizado com os mais diferentes fins, atualmente os principais grupos se encontram na Rússia e Indonésia. Normalmente esses grupos controlam ambientes específicos da darkweb.

**Crowdcrime:** Como o próprio nome diz, são espaços para encomendas de crimes com financiamento coletivo, na Darkweb existem aplicações específicas para organizar e planejar esse tipo de crime.

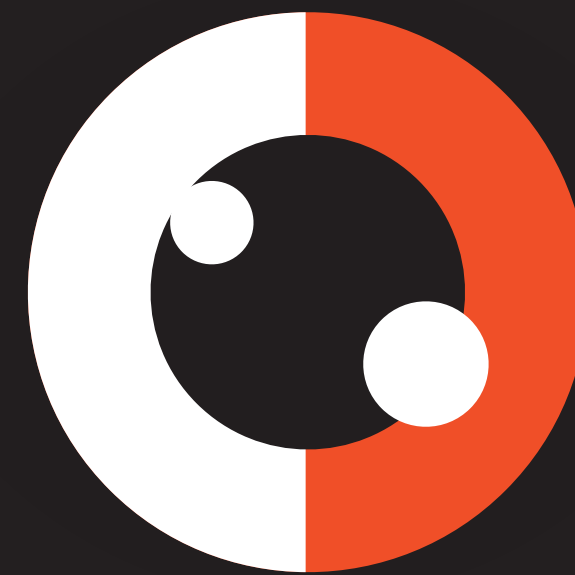


# 5 minutos para hackear seus dados

Fato é que a sua atual localização - identificada no seu smartphone - pode se tornar a cena de um crime, pois um especialista leva apenas 5 minutos para hackear um dispositivo tecnológico qualquer.

O poder do *big data* está à disposição de todos. A expansão tecnológica trouxe muitos benefícios e muitas novas conexões, isso obviamente trouxe junto uma maior vulnerabilidade dos usuários finais.

As análises individuais e sociais que os **criminosos analíticos** estão fazendo transformam a internet em um verdadeiro "big brother" em tempo real.



Existe uma falácia sobre a neutralidade matemática dos algoritmos, dizem que eles vão resolver todos os problemas. A verdade é que a configuração desses sistemas é feita por humanos, atrás de um computador, em algum lugar do mundo.

Existem "escolas de crime digital" preparando hackers para que suas infrações sejam cada vez mais difíceis de serem percebidas pelos usuários comuns da internet, mas a justiça também está se especializando em técnicas avançadas que impeçam a evolução dos crimes digitais.

O mundo precisa de um novo sistema de justiça e segurança, que considere esses novos tipos de crimes. Nossos corpos, que serão complementados por recursos tecnológicos para serem mais eficientes, são os próximos alvos desses criminosos analíticos. *Imagine um sequestrador controlando a frequência do seu coração através do hacking de um marcapasso e lhe extorquindo uma quantia financeira como recompensa.*



# análise de impacto na sociedade e organizações

Os analistas do Cappra Institute investigaram o potencial do impacto desse tema considerando 3 fatores:

**Novidade:** o quanto o assunto é novo

**Atividade:** o quanto o assunto é popular

**Impacto:** o quanto vai impactar em 2020

Com essa análise desenvolvemos um **Score** que representa o impacto desses 3 fatores, considerando uma variação entre 0 e 10.

Esse score é um bom indicador de senso de urgência de ação nesse tema.

**SCORE**

**8**

**NOVIDADE**

**ATIVIDADE**

**IMPACTO**

**<criminosos analíticos>**

**2020**

**CURIOSIDADES**

**CAPPRA INSTITUTE FOR DATA SCIENCE**

Google Trends Comparar

crimes virtuais Termos de pesquisa hackers Termos de pesquisa delegacia de crime... Termos de pesquisa

Brasil Últimos 12 meses Todas as categorias Pesquisa na Web

Interesse ao longo do tempo

Rank	Consulta relacionada	Variação	Rank	Consulta relacionada	Variação
1	delegacia de crimes virtuais rio de janeiro	Mais 190%	1	hackers moro	Aumento repentino
2	delegacia especializada em crimes virtuais	Mais 150%	2	hackers sérgio moro	Aumento repentino
3	delegacia virtual	Mais 100%	3	dilma hackers	Aumento repentino
4	delegacia crimes virtuais sp	Mais 90%	4	glenn greenwald	Aumento repentino
5	delegacia de crimes virtuais sp	Mais 50%	5	hackers araraquara	Aumento repentino

## <criminosos analíticos>

### CASO RECENTE

**G1** ECONOMIA

TECNOLOGIA

### 'Spoofing': como foi a invasão do celular de Sérgio Moro, segundo a decisão judicial que mandou prender 4 suspeitos

Fragilidade técnica do sistema de caixa postal permitiu acesso ao código que permitia entrar no Telegram das vítimas.

Por Altieres Rohr e Thiago Lavado  
24/07/2019 19h16 - Atualizado há 5 meses

Realizadas as diligências investigativas, a autoridade policial logrou inferir como o invasor teve acesso ao código enviado pelos servidores do aplicativo Telegram para a sincronização do serviço Telegram Web relativo às contas invadidas. O Telegram permite que o usuário solicite o código de acesso via ligação telefônica com posterior envio de chamada de voz contendo o código para ativação do serviço Web, cuja mensagem fica gravada na caixa postal das vítimas. O invasor então realiza diversas ligações para o número alvo, a fim de que a linha fique ocupada, e a ligação contendo o código de ativação do serviço Telegram Web é direcionada para a caixa postal da vítima.

Decisão da Justiça aponta como hackers teriam invadido aplicativo de mensagens no celular de autoridades — Foto: Reprodução





## POR QUE VOCÊ DEVE OBSERVAR ESSA TENDÊNCIA EM 2020?

Um mundo com muitas oportunidades de conexões tecnológicas também eleva a quantidade de brechas e falhas dos sistemas, isso cria uma oportunidade para os criminosos analíticos se beneficiarem do uso de técnicas avançadas para a realização de crimes.

Obviamente, que muitos dados e técnicas avançadas também vão permitir a reação mais eficiente da justiça, seja através de sistemas, leis ou até da antecipação de um crime que ainda nem foi realizado - como tem ocorrido na prevenção de ataques terroristas com a destruição antecipada de células que trocam informações no planejamento desses crimes.

A guerra contra o cybercrime já está acontecendo, mas como a tecnologia está disponível para todos, será cada vez mais comum ouvirmos falar sobre esses crimes no cotidiano de nossas vidas.

## RECOMENDAÇÕES:

- Pesquise sobre tipos de crimes virtuais que estão mais relacionados com suas atividades e vida, isso vai ajudar na identificação das suas fragilidades.
- Crie “mecanismos de segurança” para dificultar o acesso aos seus dados e rotinas.
- Proteja suas senhas e altere de forma contínua seus códigos secretos, isso dificulta o roubo dos seus dados.
- Evite disponibilizar seus dados em ambientes abertos e evite ao máximo a conexão em redes públicas, essas costumam ser um ambiente de alto risco para instalação de rotinas criminosas.
- Verifique a Lei Geral de Proteção dos Dados Pessoais (LGPD), ela entra em vigor no dia 20 de agosto de 2020 no Brasil e vai modificar a forma como empresas e pessoas físicas lidam com informações privadas.

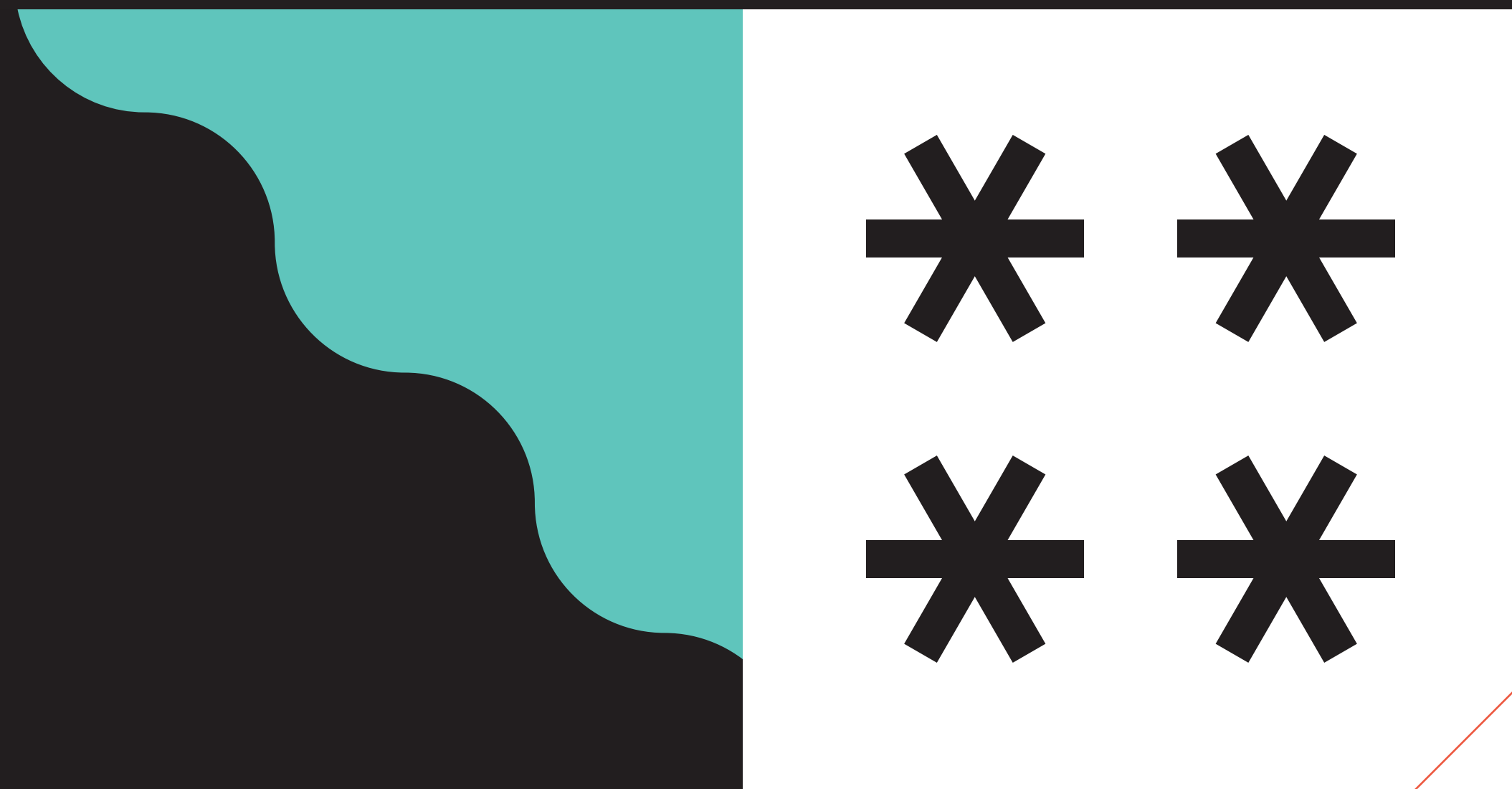
## NA PRÁTICA:

Em 30 dias: Limpe todos seus históricos de navegação, desinstale plugins e aplicativos com risco e mantenha um controle de senhas contínuo.

Em 6 meses: Conheça um pouco mais sobre a Lei Geral de Proteção dos Dados Pessoais (LGPD), visite o site do próprio governo para isso.

Em 12 meses: Crie rotinas de segurança pessoal para uma vida digital mais tranquila.





# detox de informação



ELIMINANDO O EXCESSO  
DE INFORMAÇÃO

*“Precisamos aprender a  
desapegar de informação ruim.  
A sociedades está muito  
impressionada com o volume de  
dados, enquanto deveria estar  
focada na qualidade da  
informação.”*

CAPPRA, Ricardo.



# excesso de informação

Sem dúvida nenhuma o termo da área de tecnologia mais citado na última década foi **Big Data**.

Esse termo foi usado das mais diferentes maneiras, fosse para exemplificar o bom uso dos dados nos negócios ou para referenciar um grande vazamento de dados pessoais. No Data Thinking 2019 explicamos o conceito Big Data e destacamos um dos impactos na sociedade: o Infoxiation - intoxicação por consumo excessivo de informação (você encontra esse estudo completo em [www.cappra.institute/2019](http://www.cappra.institute/2019)).

Mas lidar com tanta informação exige um elevado gasto de energia do cérebro humano, que são incríveis para investigar hipóteses, identificar highlights relevantes e ter insights a partir das descobertas - coisas que cérebros artificiais não fazem tão bem. Por outro lado, falhamos ao analisar grandes volumes de dados, não conseguimos processar informação em alta velocidade e não conseguimos distinguir - *de forma eficiente* - o que é mais relevante quando estamos diante de muitos dados.

Não percebemos, mas o nosso cérebro precisa trabalhar muito para selecionar a informação que vamos consumir. Esse esforço costuma ser elevado por causa do volume de informação que está ao nosso alcance, por isso precisamos buscar métodos para ser mais eficientes no “tratamento dos dados” que vamos usar. Para aprimorar a gestão do conhecimento e melhorar as decisões que tomamos diariamente, precisamos de um processo de **detox de informação**.

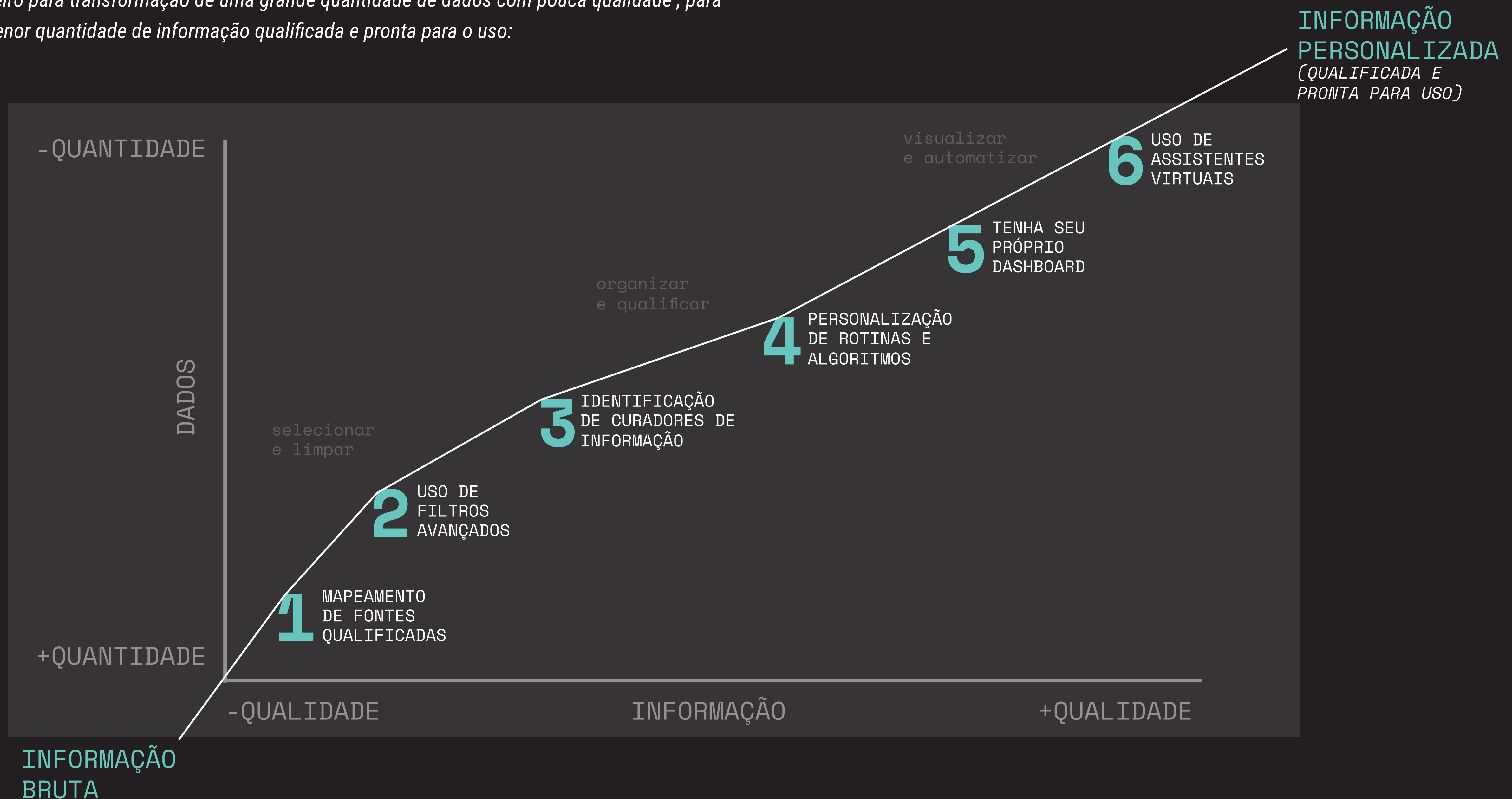
*Investigamos algumas formas para esse consumo de informação ser mais eficiente, e vamos explorar aqui essa jornada.*





# detox de informação

Um roteiro para transformação de uma grande quantidade de dados com pouca qualidade, para uma menor quantidade de informação qualificada e pronta para o uso:



# etapa 1

## *selecionar e limpar*

### [MAPEAMENTO DE FONTES QUALIFICADAS]

Identificar as fontes de informação contínuas mais relevantes, com baixa taxa de viés no tratamento dos dados e com o devido cuidado para eliminação de fake news é uma tarefa que exige tempo e muita atenção. Tanto em ambientes corporativos quanto em nossa vida pessoal, acabamos sendo influenciados por outras pessoas que possuem os mesmos gostos e preferências, fazendo com que a informação normalmente possua um direcionamento prévio para responder as questões que já conhecemos. Em muitos casos, acaba servindo mais com uma confirmação de nossa opinião do que de fato uma informação qualificada para nos ajudar na decisão. Determinar um universo muito fechado de fontes qualificadas costuma nos mostrar um cenário mais confortável para nossos olhos, mas com sério risco de construção de *bolhas de conhecimento*.

**Seja nos negócios ou na vida pessoal, é preciso revisar as fontes de dados que alimentam os seus painéis de informação, principalmente aquelas que você consome no dia-a-dia. Tome cuidado para seus próprios vícios não deturparem os dados que você está analisando.**

### [USO DE FILTROS AVANÇADOS]

As empresas de software/internet estão fazendo um esforço muito grande para transferir as funcionalidades de filtros para os usuários finais. Ferramentas de busca na internet, Redes Sociais, Ferramentas de Business Intelligence, Softwares de gestão empresarial, todos estão trabalhando para *empoderar* o usuário final. Esses recursos estão sendo apresentados em formas de filtros, para que funcione como um serviço *Self-Service de informação*.

**Verifique se está usando esses recursos de forma adequada, navegue um pouco pelos softwares/apps que você acessa diariamente e descubra essas funcionalidades que podem reduzir muito a quantidade de informação - desnecessária - que você está consumindo todos os dias.**

A tendência é que as empresas de software transfiram na íntegra essas funcionalidades para os usuários finais, não deixe que alguém filtre por você limitando assim sua visão do mundo (vale a pena conferir o TEDx do Ricardo Cappra sobre o tema “Algoritmos: você está no controle?” <https://youtu.be/wBzxWPo1Fc> ).



# etapa 2

## *organizar e qualificar*

### [IDENTIFICAÇÃO DE CURADORES DE INFORMAÇÃO]

Selecionar as fontes e informações relevantes é algo muito trabalhoso, exige atenção, tempo para determinar os critérios, e avaliação da qualidade da informação. No trabalho, essa função acaba sendo realizada por assistentes ou outras áreas especializadas que limpam e tratam os dados brutos, mas esse processo acaba não sendo muito eficiente.

**O ideal seria a seleção de uma curadoria qualificada, pessoas/empresas/softwares que trabalham classificando a informação para que chegue mais qualificada nas mãos de quem irá consumir e usar.**

Hoje existem serviços de curadoria de informação: empresas de pesquisas/análises, especialistas em conteúdo, jornalistas, referências técnicas, influenciadores no tema, entre outros que realizam essa atividade. Para reduzir o volume de dados precisamos buscar mais curadores, que nos ajudem na árdua tarefa de consumir informação relevante e de qualidade.

### [PERSONALIZAÇÃO DE ROTINAS E ALGORITMOS]

Para programadores, o sequenciamento de rotinas é objetivo fim de qualquer software, mas costumamos deixar esse recurso restrito ao mundo dos códigos. Existem algumas ferramentas que permitem esse sequenciamento sem necessariamente conhecer linguagens de programação, e que ajudam no sequenciamento de atividades. Essa tarefa pode ser realizada com objetivo de filtro e redução da quantidade de informação, e esse tipo de aplicação costuma ser muito eficiente.

**Uma ferramenta muito fácil de experimentar é o IFTTT ([ifttt.com](https://ifttt.com)). Você pode, por exemplo: coletar todas as menções do twitter que citam a palavra big data, monitorar todos os posts de um determinado especialista no assunto, classificar todos os seus email que possuem a palavra big data, e organizar isso diretamente em uma planilha do excel, classificando como você preferir.**

É apenas um exemplo, mas uma possibilidade real de você fazer sua própria personalização de rotina para consumo de conteúdo e transformação em informação. É um construtor de algoritmos, experimente.

# etapa 3

## *visualizar e automatizar*

### [TENHA SEU PRÓPRIO DASHBOARD]

Ao circularmos em escritórios de grandes empresas e startups digitais, encontramos muitos dashboards espalhados, com informações já selecionadas para o consumo coletivo. Esses mecanismos servem para reduzir a informação e deixar à disposição - e à vista - aquelas que são mais importantes. O único problema é que, normalmente, quem definiu a informação que está ali foi um grupo restrito da empresa, e isso não representa a necessidade de consumo individual de informação.

**As ferramentas de visualização de informação estão cada vez mais simples de usar, por exemplo, o Google disponibiliza gratuitamente o Data Studio ([datastudio.google.com](https://datastudio.google.com)) e permite que você crie dashboards para visualizar informações de suas planilhas, docs, fontes externas de dados,...**

Não é complexo construir um dashboard e isso permite que você estruture a informação do jeito que você quiser.

### [USO DE ASSISTENTES VIRTUAIS]

Os assistentes virtuais estão ficando mais inteligentes, mas ainda estamos usando muito pouco desses recursos que podem melhorar muito a qualidade do nosso consumo de informação. Esse tipo de serviço pode realizar o trabalho de tratamento dos dados e ainda automatizar o processo de consumo da informação.

**Imagine que a rotina do exemplo anterior, criada no IFTTT para coletar e classificar o conteúdo sobre big data, poderia ser lido um resumo automaticamente enquanto você escova seus dentes pela manhã.**

Trata-se de um filtro automatizado, que entra na sua rotina de consumo de informação sem maiores esforços. Você pode criar um comando chamado: “me fale as coisas mais recentes sobre big data”, isso ativa o uso do filtro de informação personalizado por você, e Alexa/Google/Siri automatizam a leitura desse conteúdo. É algo novo, mas muito mais simples do que parece.



# o hábito libertador de decidir a partir dos dados

O escritor *Hans Rosling* fala em seu livro **Factfulness** sobre **o hábito libertador de só ter opiniões baseadas em fatos**, mas isso exige um esforço prévio de qualificar a informação que estamos consumindo no nosso dia-a-dia, é inevitável um processo de redução da quantidade de informação que chega até nós.

Existem dois caminhos para essa redução de informação: **deixar que outras pessoas - e máquinas - escolham aquilo que chegará até nós** ou **participarmos ativamente desse processo de redução**, fazendo com que seja mais personalizado.

Essa é uma escolha individual, em alguns casos podemos usar filtros genéricos. No caso de informações cruciais, deveríamos estar no controle da qualificação dos dados e transformação em informação, pois vai interferir diretamente em nosso processo de decisão.





# análise de impacto na sociedade e organizações

Os analistas do Cappra Institute investigaram o potencial do impacto desse tema considerando 3 fatores:

**Novidade:** o quanto o assunto é novo

**Atividade:** o quanto o assunto é popular

**Impacto:** o quanto vai impactar em 2020

Com essa análise desenvolvemos um **Score** que representa o impacto desses 3 fatores, considerando uma variação entre 0 e 10.

Esse score é um bom indicador de senso de urgência de ação nesse tema.

**<detox de informação>**

**SCORE**  
**8**

**NOVIDADE**  
[Progress bar]

**ATIVIDADE**  
[Progress bar]

**IMPACTO**  
[Progress bar]

**crise de ansiedade**  
Termo de pesquisa

Brasil | Nos últimos 5 anos | Todas as categorias

Interesse ao longo do tempo

Interesse por sub-região

- Rio de Janeiro
- São Paulo
- Amapá
- Pernambuco
- Paraná

**crescimento de dados no mundo**  
(Exabytes)

Source: IDC's Digital Universe, sponsored by EMC, December 2012

**2020**

**CURIOSIDADES**

**CAPPA INSTITUTE FOR DATA SCIENCE**

## <detox de informação>

### CASO RECENTE

**NEGÓCIOS** COLUNISTAS REVISTA PODCASTS

VIDA

#### Consumo excessivo de mídias sociais está ligado a doenças mentais

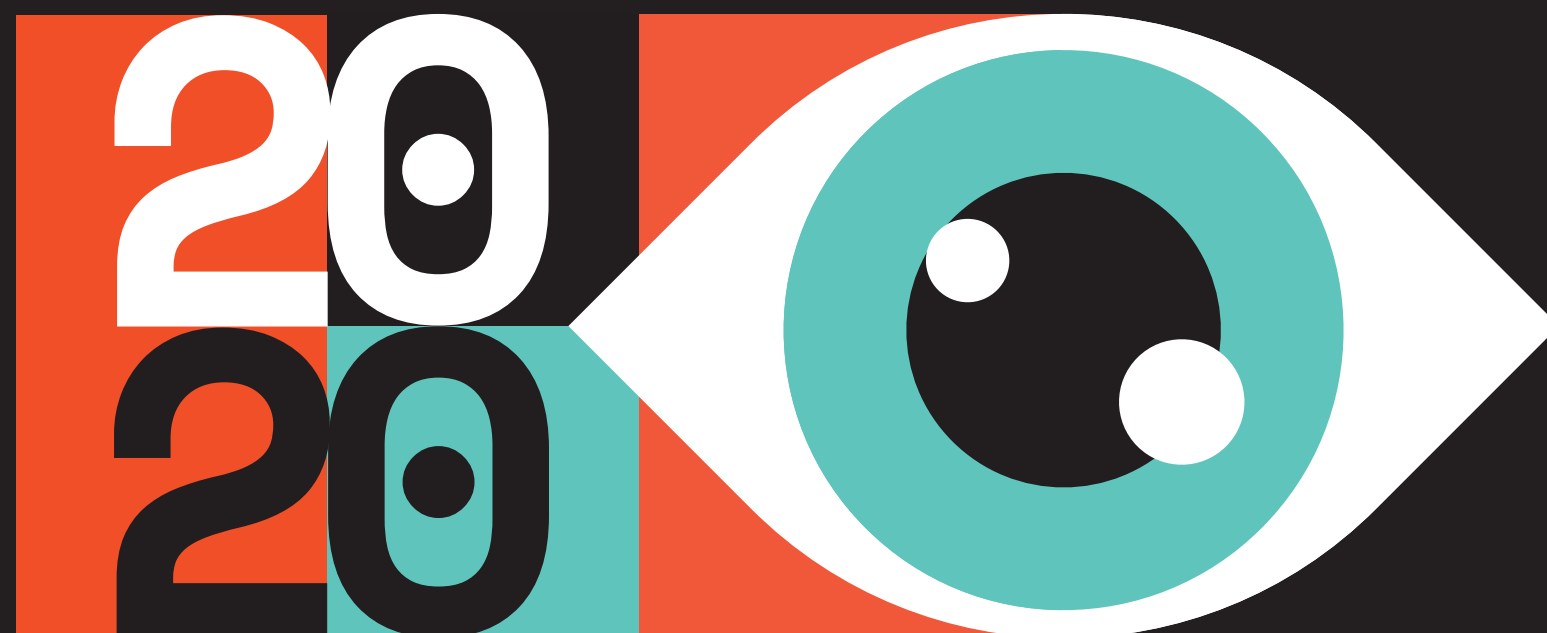
Jovens relatam problemas com ansiedade, depressão, sono e medo

Jovens que usam redes sociais em excesso podem apresentar algum distúrbio mental (Foto: Thinkstock)

Quanto tempo por dia você passa nas redes sociais? Talvez seja hora de repensar esse hábito. Em 20 de maio, inicia-se a "semana de conscientização sobre a saúde mental". No Brasil, 5,8% da população (ou 11,5 milhões de pessoas) sofrem de depressão, segundo a Organização Mundial da Saúde (OMS). Isso é mais do que a média mundial, que é de 4,4% (ou 322 milhões de pessoas). No Reino Unido, cerca de um quarto dos adultos foram diagnosticados em algum momento com distúrbio psiquiátrico, o que custa à economia cerca de 4,5% do PIB por ano.

Essas doenças têm muitas causas, mas segundo pesquisadores, quando apresentadas em jovens, estão ligadas ao consumo em excesso de mídias sociais.





# <detox de informação>

## POR QUE VOCÊ DEVE OBSERVAR ESSA TENDÊNCIA EM 2020?

Daqui para frente a quantidade de fontes de informações não vai reduzir, bem pelo contrário, vai aumentar. Todos os dias a sociedade e as organizações geram volumes absurdos de dados, que estão sendo transformados em informações - sendo essas muitas vezes manipuladas ou simplesmente não filtradas de forma adequada.

Em 2019 passamos por uma extensa crise de disseminação de fake news, as ações de grandes veículos de comunicação e internet não dão conta de eliminar esse enorme ruído. Os indivíduos precisam participar ativamente desse processo de transformação de dados em informação, ou seremos sempre influenciados por terceiros em nosso processo de decisão.

## RECOMENDAÇÕES:

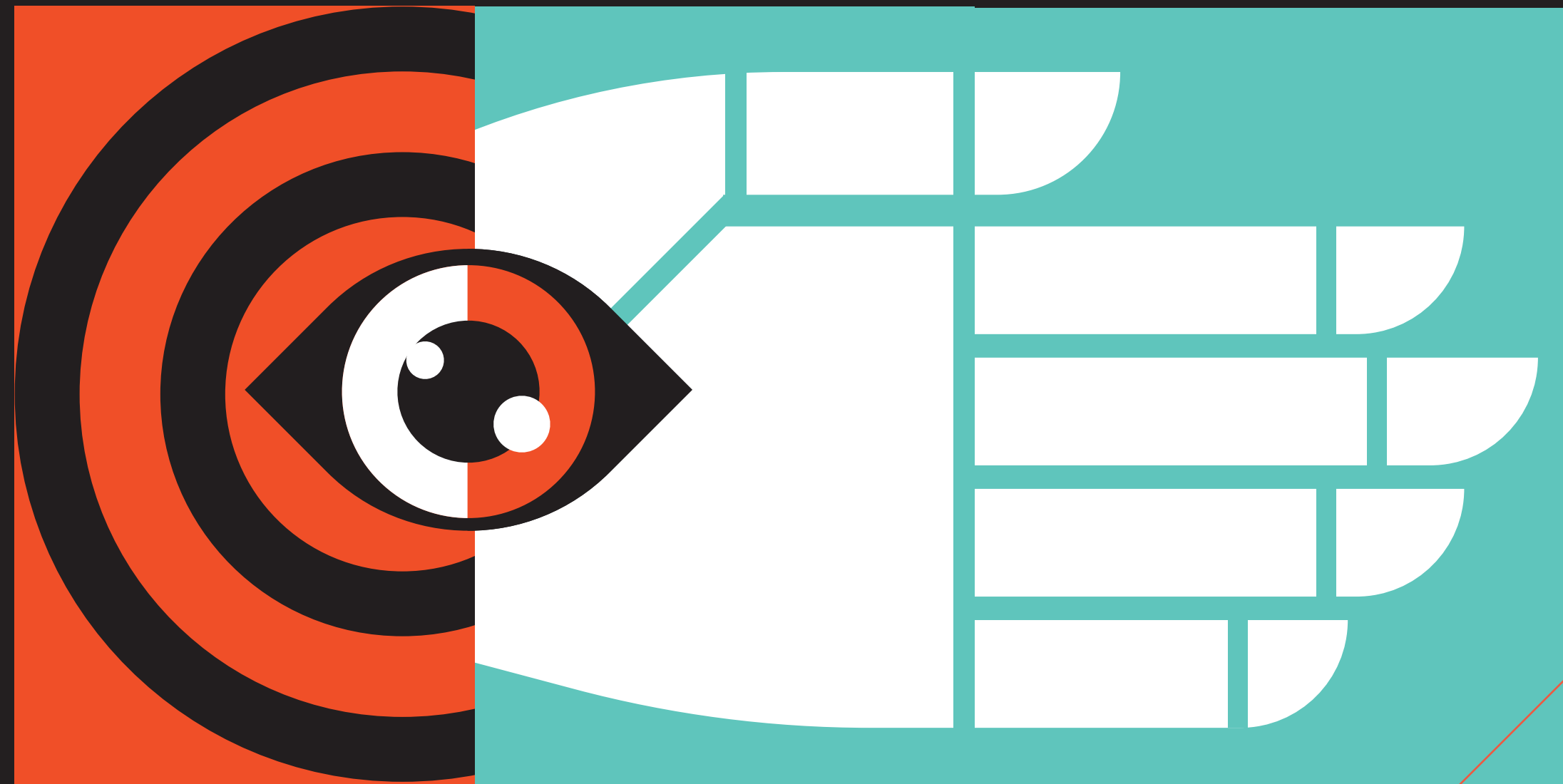
- Verifique as informações que são cruciais para suas decisões e identifique aquelas que você participa ativamente do processo de "limpeza", qualificação e organização.
- Crie rotinas para consumo de informação qualificada, comece pela seleção das fontes de dados.
- Não deixe os filtros de informação nas mãos de terceiros, participe ativamente desse processo de redução (seja através de uso de filtros ou seleção de curadores), isso evitará manipulações inadequadas.
- Aprenda a criar suas próprias rotinas para geração de informação qualificada, a automação desse processo é mais fácil do que você imagina.
- Automatize o máximo que puder para consumo dessas informações, seja através dos assistentes pessoais lendo para você ou de dashboards, isso vai reduzir seu esforço diário nessa tarefa.

## NA PRÁTICA:

Em 30 dias: Liste suas fontes de consumo diário de informação - seja pessoal ou profissional - e marque aquelas que são seguras, relevantes, importantes ou urgentes.

Em 6 meses: Automatize uma rotina de filtro de informação, isso vai melhorar a qualidade das suas decisões.

Em 1 ano: Crie seu próprio dashboard, pode ser com suas informações do trabalho, da academia... Assim vai reduzir o desperdício de energia no consumo desse tipo de informação e qualificar seu processo de tomada de decisão.



# superpoderes cognitivos



CÉREBROS ARTIFICIAIS PARA  
AJUDAR CÉREBROS ORGÂNICOS

*“O smartphone é um complemento artificial do seu cérebro, ele ajuda você a armazenar e gerenciar informações. Agora, esse cérebro artificial também pode conversar com você, executar tarefas e até analisar suas decisões.”*

CAPPRA, Ricardo.



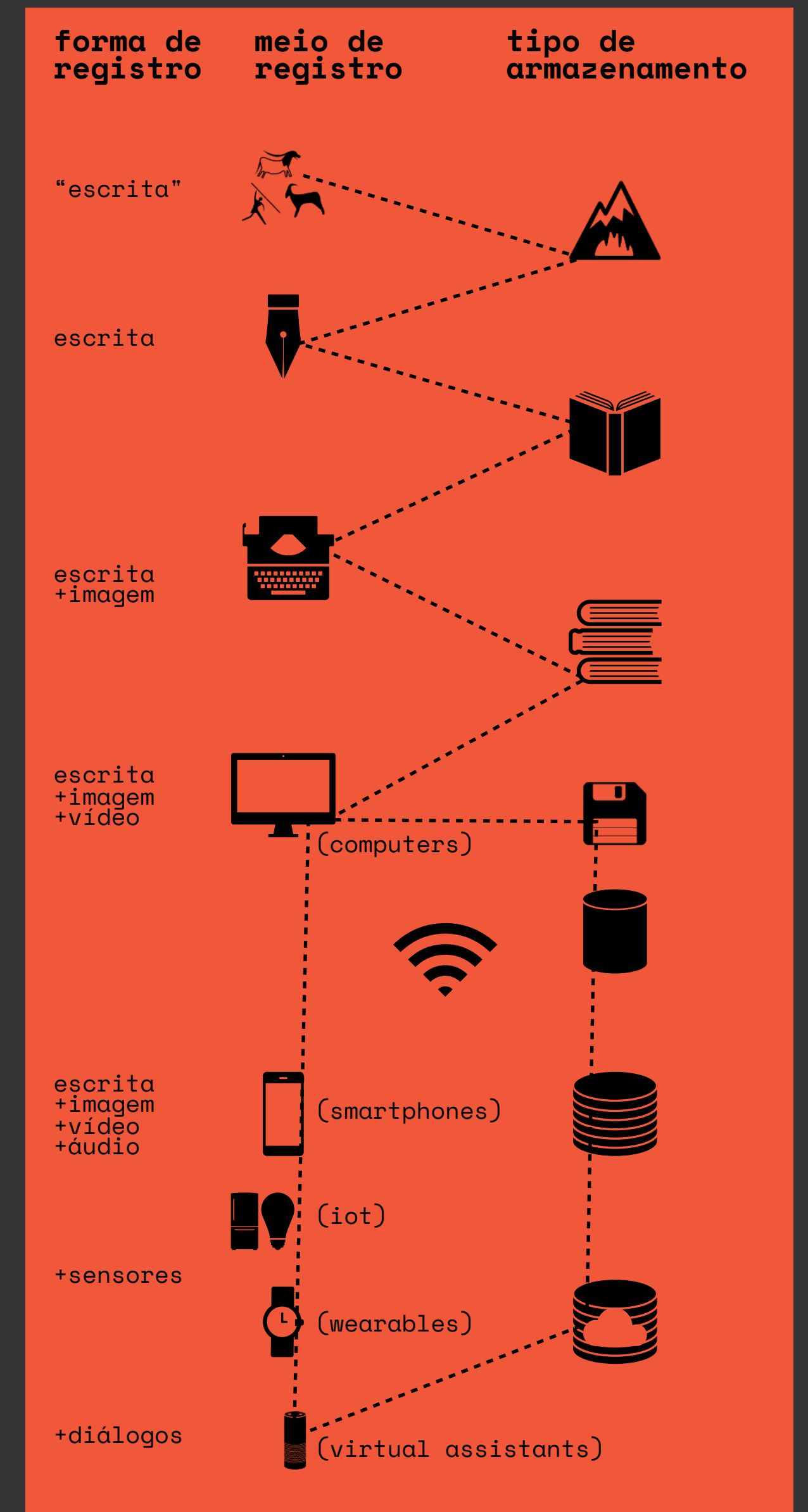
# humanidade limitada = cognição limitada

**A tecnologia sempre foi pensada por humanos, e isso obviamente é um limitador para a evolução tecnológica.**

Sempre que um ser humano pensar na solução de um problema, ele inevitavelmente limita-se ao universo que conhece, aquilo que o afeta diretamente ou impacta a vida das pessoas que estão ao seu redor. Pensar no aprendizado de máquina exige uma análise de como o ser humano aprende e registra conhecimento.

## A FORMA QUE OS HUMANOS ARMAZENAM CONHECIMENTO CONTINUA EVOLUINDO

Nossas primeiras formas de registro de conhecimento são desenhos, armazenados em pedras. Nossa primeira evolução é para a escrita à mão, armazenada geralmente em documentos e livros. Posteriormente desenvolvemos tecnologia para acelerar - e padronizar - o armazenamento de conhecimento, então as máquinas de escrever foram criadas para armazenamento do conhecimento do mundo. Em seguida, por necessidade de formas mais eficientes de acessar a informação, chegamos ao computadores, que além da digitação para armazenar conhecimento de forma mais rápida, ainda trouxe o benefício de sistemas de busca mais rápidos para acessar a informação. Mas aquele grande teclado foi reduzido para equipamentos que podemos carregar em qualquer lugar, onde a digitação exige um esforço cada vez menor. Atualmente os smartphones coletam muito do conhecimento da sociedade, são registros individuais que quando unificados tornam-se grandes bancos de conhecimento do mundo, tudo na internet - armazenado em nuvem - para uma gestão mais eficiente de todo esse conhecimento. Atualmente a sociedade está reduzindo o esforço de escrita, substituindo por emojis e áudios.



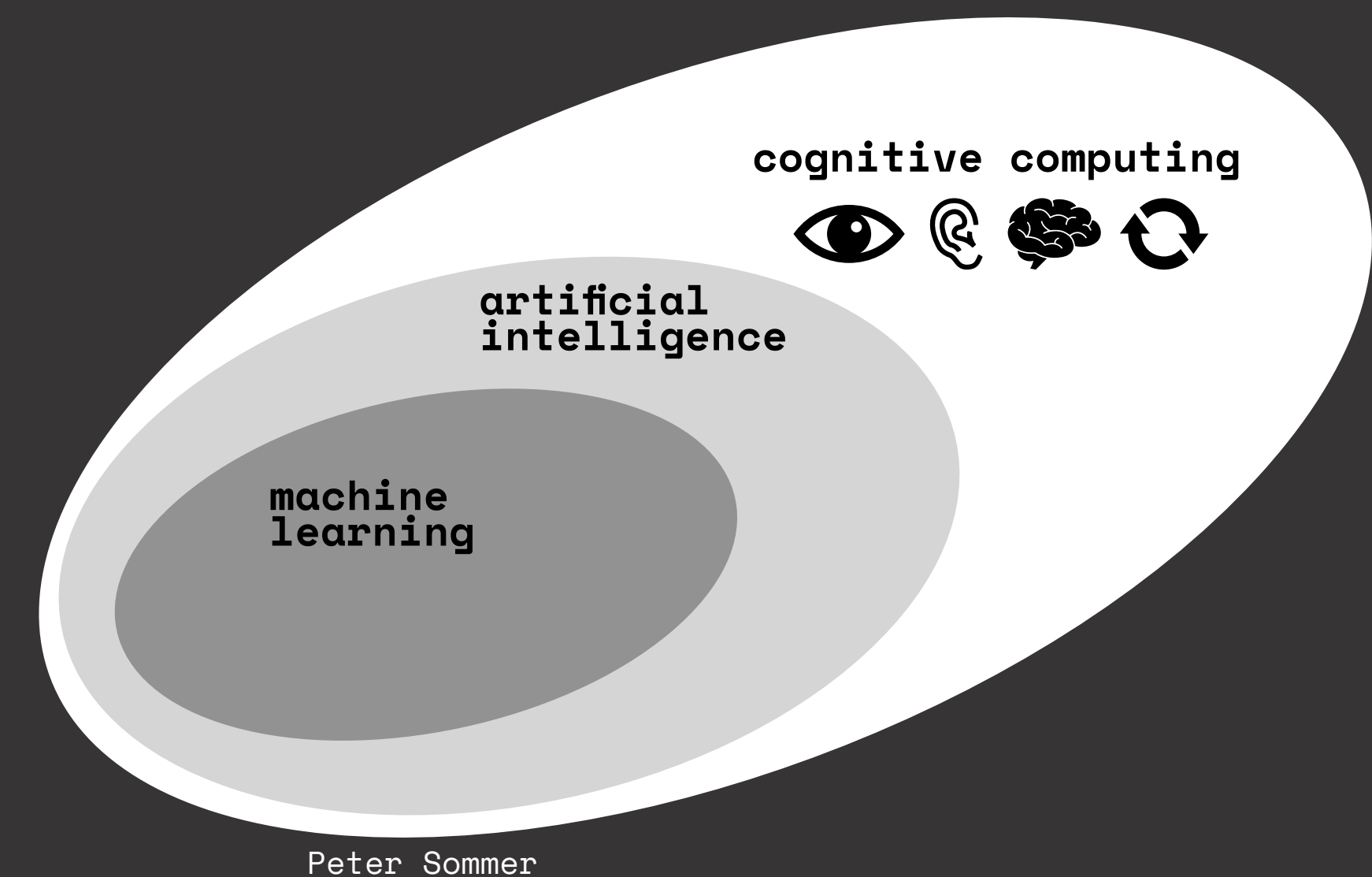
# digitalar é um trabalho burocrático e desnecessário

## Digitalar é uma tarefa muito trabalhosa.

Os humanos não inserem novas informações em seus cérebros através da digitação, não seria prático adquirir conhecimento dessa forma. Nós lemos, ouvimos e até escrevemos - nesse caso como um reforço de registro para nossos cérebros - mas na prática tudo que está ao nosso redor está de alguma forma sendo armazenado. Adquirimos conhecimento através dos nossos 5 sentidos, e quando precisamos recuperar algo nossos cérebros ativam um mecanismo similar ao "buscar na internet". Através da visão, inserimos em nossos cérebros as imagens e textos que capturamos. Através da audição capturamos sons e conversas, e armazenamos tudo isso para uma possível consulta futura daquela informação.

Esse contexto explica o atual momento que vivemos com as máquinas, agora elas podem aprender através das interações humanas, não só através da escrita. Os assistentes pessoais são os primeiros grandes mecanismos de captura desse conhecimento em forma de diálogos (Alexa, Siri, Google, Cortana,...) se espalham pelo mundo para aprender e nos ajudar nas atividades rotineiras, essas máquinas não precisam que os registros sejam escritos para armazenar conhecimento.

A primeira fase dessa tecnologia, ainda é baseado no padrão de armazenamento de conhecimento antigo, ou seja, segue um longo roteiro: 1) a máquina captura a voz (*cognitive computing*); 2) transforma em texto; 3) processa a informação (*artificial intelligence*); 4) armazena/aprende (*machine learning*); 5) encontra uma resposta no banco de dados - que está em forma de texto; 6) transforma em voz; 7) e responde em forma de diálogo; tudo isso em milésimos de segundos.



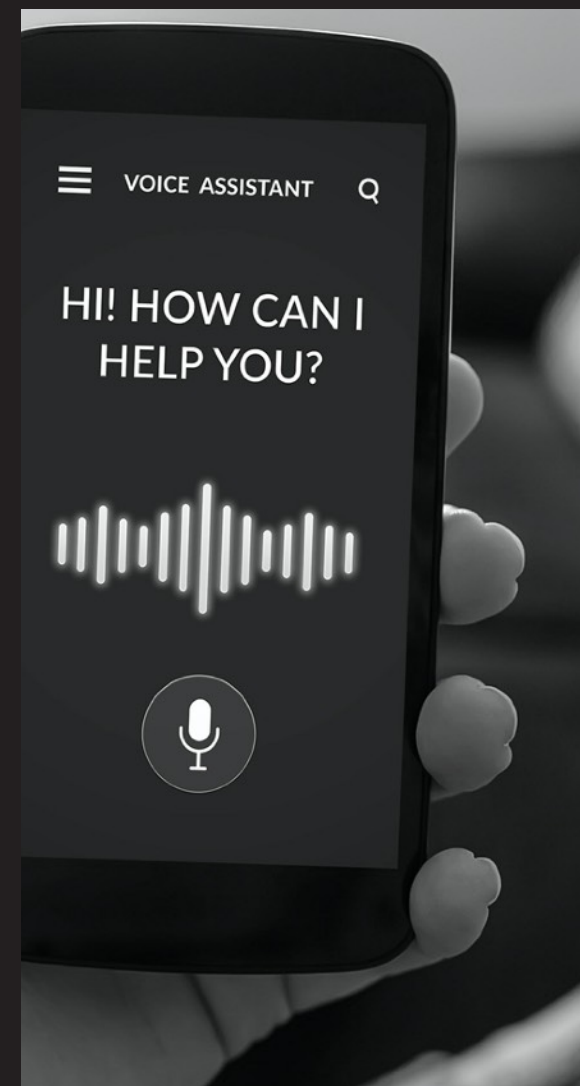


# tudo gera conhecimento

## [APRENDIZADO POR DIÁLOGO]

Em um futuro muito próximo - com processamento quântico e armazenamento mais eficiente da informação - as máquinas não vão mais precisar guardar os registros em formato de texto, será tudo armazenado em forma de áudio, sem necessariamente a transformação disso em texto - *isso atualmente ocorre para reduzir o volume de armazenamento nos bancos de dados.*

Novas técnicas científicas também farão com que entendimento e tratamento semântico da informação sejam mais eficientes, pois hoje essas máquinas ainda falham na compreensão dos registros de áudio. O aprendizado por diálogo está abrindo uma nova janela na forma como as máquinas aprendem - e humanos armazenam conhecimento, será um armazenamento de conhecimento mais informal, e não vamos achar estranho “falar com uma máquina”.



## [O DADOS DE TODAS AS COISAS]

Em paralelo ao armazenamento de conhecimento em formato de diálogo, ocorreu uma grande evolução tecnológica na coleta e armazenamento de informação através de smartphones, sensores (ruas, lojas, eletrodomésticos,...).

Pequenos dispositivos coletam imagens, sinais, áudios de forma ininterrupta, os Dados das Coisas - *chamamos de Data of Things, para associar a Internet of Things.* Todos os novos recursos tecnológicos possuem a captação de dados: sua nova geladeira, sua TV, sua cafeteira, as lâmpadas da sua casa, e qualquer dispositivo com possibilidade de conexão com a internet (seja um Wi-Fi, bluetooth,...).



# análises cognitivas

Esses novos dados estão armazenando um tipo de conhecimento sobre a sociedade - e sobre as organizações - que antes não era possível registrar, pois a velocidade de coleta e recursos de banco de dados eram limitados, mas agora todas essas informações estão ensinando algo sobre nosso comportamento, preferências e hábitos.

Estamos vivendo um momento muito especial, onde a capacidade de cognição humana está sendo amplificada todos os dias, com armazenamento contínuo através de máquinas que poderão no futuro nos ajudar a solucionar problemas considerados impossíveis - até agora.

A evolução analítica ganha um novo capítulo por causa desses novos recursos, agora as análises serão mais avançadas. Haverá uma menor dependência da descrição feita por humanos, a coleta direta dos seus estímulos, diálogos e reações permitirão novas descobertas e a evolução mais rápida dos padrões de cognição humana.





# análise de impacto na sociedade e organizações

Os analistas do Cappra Institute investigaram o potencial do impacto desse tema considerando 3 fatores:

**Novidade:** o quanto o assunto é novo

**Atividade:** o quanto o assunto é popular

**Impacto:** o quanto vai impactar em 2020

Com essa análise desenvolvemos um **Score** que representa o impacto desses 3 fatores, considerando uma variação entre 0 e 10.

Esse score é um bom indicador de senso de urgência de ação nesse tema.



SCORE

# 6

NOVIDADE

ATIVIDADE

IMPACTO

## <superpoderes cognitivos>




CAPPA INSTITUTE FOR DATA SCIENCE

<superpoderes cognitivos>

### CASO RECENTE

DESCUBRA A REDE IT TRENDS

Gestão Carreira Tendências Notícias Transformação Digital Opinião

Home > Tendências

## Assistentes pessoais virtuais ganham o mundo corporativo

Diálogos facilitarão o dia a dia. Tendência chega para ficar, escreve especialista

Marcelo Arakaki\*  
24/04/2019 às 17h12

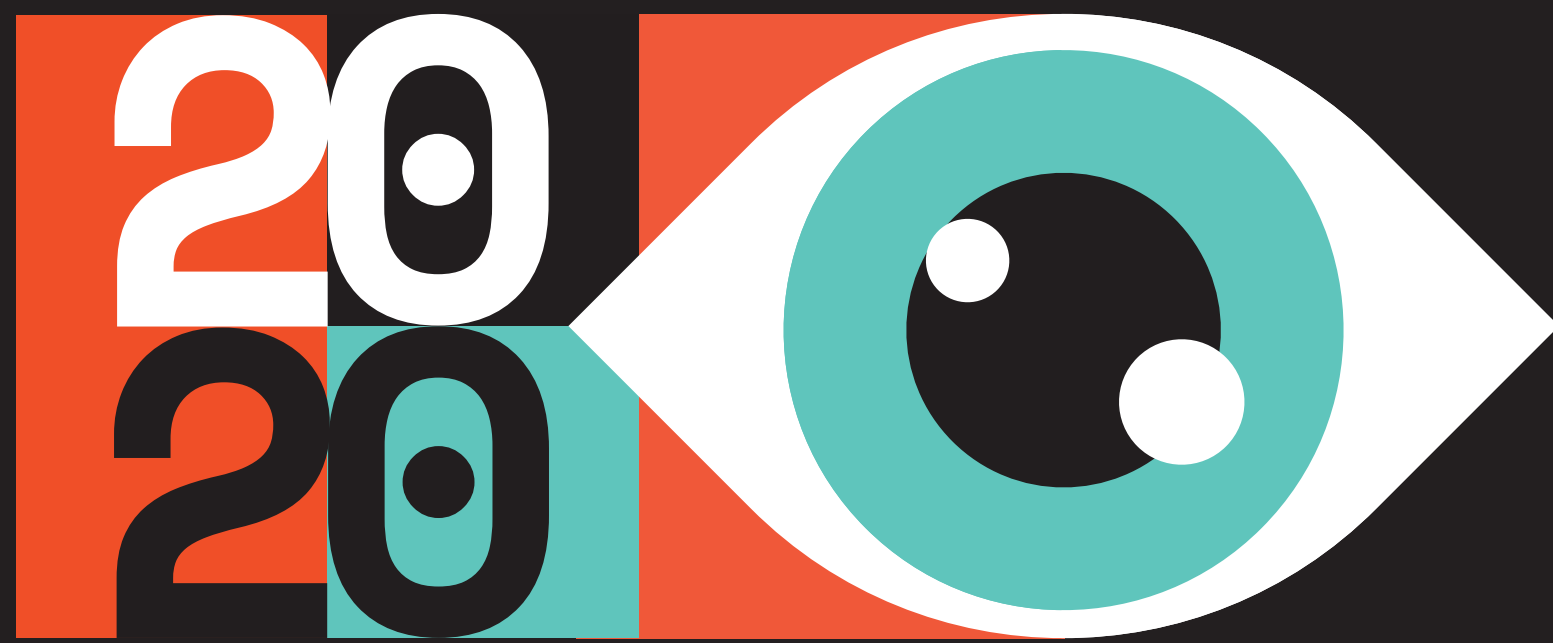


Foto: Shutterstock

<https://cio.com.br/assistentes->

Já faz um bom tempo que as conversas com robôs para resolver questões do dia a dia deixaram de ser apenas cenas de filmes de ficção científica. Aqueles diálogos que nos faziam pensar como essas invenções facilitarão o nosso dia a dia. Afinal, em uma conversa "falada", emitimos cerca de quatro vezes mais palavras do que quando nos comunicamos pela escrita. Essa é uma das razões pelas quais os assistentes virtuais pessoais (APVs) são uma realidade e chegaram para ficar.

CURIOSIDADES



## <superpoderes cognitivos>

### POR QUE VOCÊ DEVE OBSERVAR ESSA TENDÊNCIA EM 2020?

Os Superpoderes Cognitivos já estão à disposição, ainda nas mãos de poucos privilegiados, mas o armazenamento de conhecimento já está ocorrendo dentro de nossas casas. O uso ainda é limitado por parte dos usuários finais dessa tecnologia, poucos utilizam os recursos de *assistente virtual*, então não tiveram a oportunidade de experimentar os benefícios. Esses assistentes podem fazer tarefas banais como cantar uma canção de natal, mas também podem ajudar você a organizar melhor sua vida através de uma rotina automatizada - *imagine uma pessoa com deficiência visual que recebe auxílio desse tipo de tecnologia para realização de tarefas caseiras antes impossíveis para ela.*

Mas precisamos admitir que é estranho usar assistentes virtuais, falar com máquinas - através de voz/diálogo - para ajudar em alguma tarefa ainda não é um hábito. Em breve perceberemos que esses superpoderes cognitivos podem ser usados para potencializar as nossas capacidades, reduzindo as limitações que nossos cérebros orgânicos possuem. Os cérebros eletrônicos podem amplificar a capacidade humana de armazenamento, e potencializar o acesso e uso dessa informação.

### RECOMENDAÇÕES:

- Experimente algumas formas novas de registro de conhecimento, por voz por exemplo, isso fará você compreender o funcionamento desses recursos;
- Você não precisa de recursos altamente tecnológicos, seu smartphone já possui esse recurso de assistente virtual;
- Se quiser ampliar essa experiência, coloque uma Alexa, Google Home,... em sua casa;
- Você precisa vencer a barreira de "conversar com uma máquina";
- Se você tem um negócio e usa dados, prepare-se para lidar com esse novo tipo de informação armazenada. Os processos de coleta, tratamento e análise dos dados precisarão ser revisitados e atualizados.

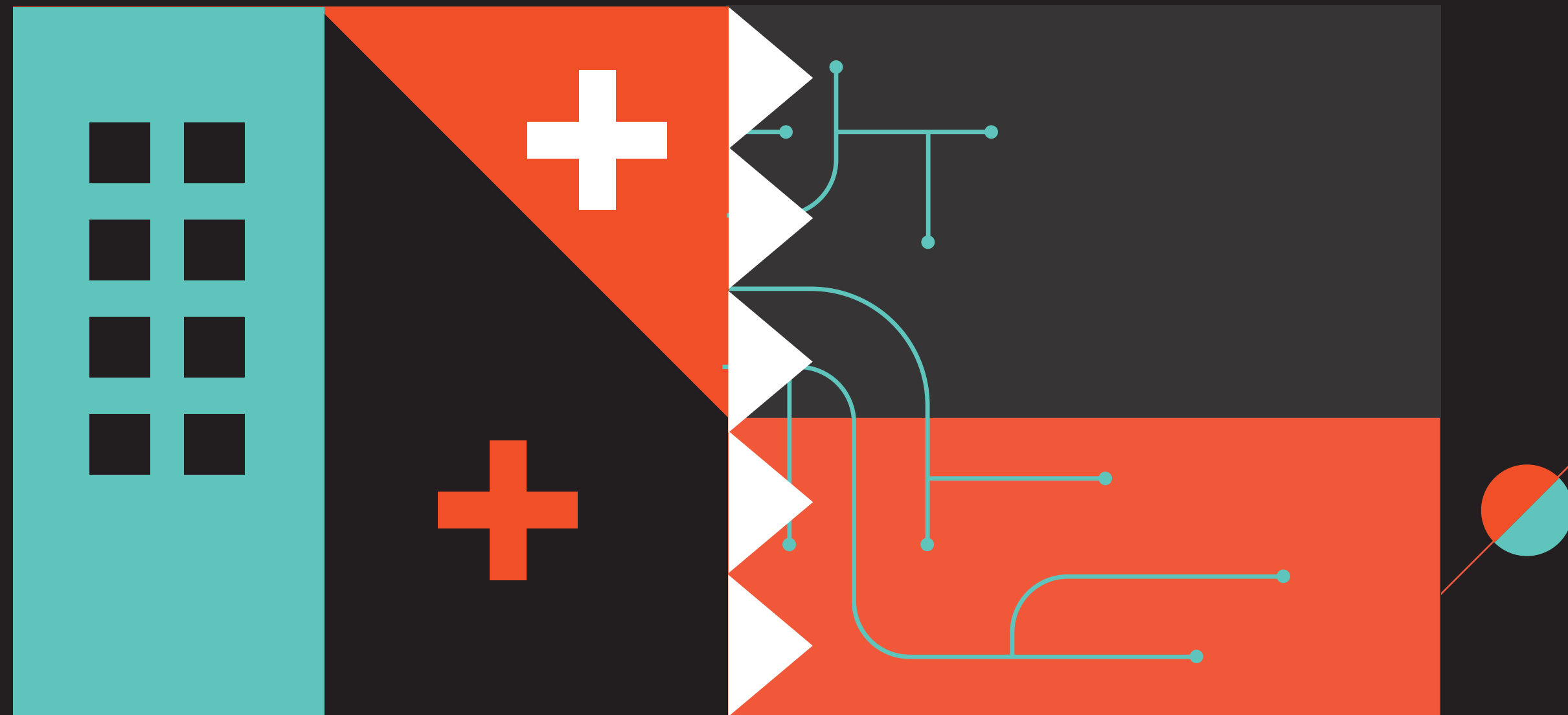
### NA PRÁTICA:

Em 30 dias: Faça alguns testes com a assistente virtual no seu smartphone, você vai se divertir fazendo isso.

Em 6 meses: Identifique aqueles recursos que podem potencializar o uso desse tipo de tecnologia em sua vida, seja em uma automação na sua casa ou no seu negócio. Por exemplo, uma assistente virtual pode participar ativamente de uma reunião, armazenando conhecimento para posterior análise.

Em 1 ano: Verifique aquelas práticas que foram mais interessantes e crie um plano de automação baseado nisso, essa ação vai expandir a atual capacidade de cognição.





# infororganization

*“Existe uma nova geração de empresas, elas funcionam como organismos vivos. Assim que recebem novas informações, elas se adaptam para serem mais eficientes, e sobreviverem.”*

CAPPRA, Ricardo.

ORGANISMOS ORIENTADOS  
POR INFORMAÇÃO

# a falha nos modelos de negócio

Estamos presenciando um momento de transformação gigantesco nos negócios, a gestão baseada em processos, competências e sistemas, está falhando.

Podemos perceber isso em todos os setores, empresas consolidadas e estruturadas estão concorrendo com pequenas startups sem nenhum histórico no setor de atuação. Mercados inteiros estão sendo dominados por modelos mais ágeis, por empresas de poucas pessoas, estabelecidas em qualquer canto do mundo.

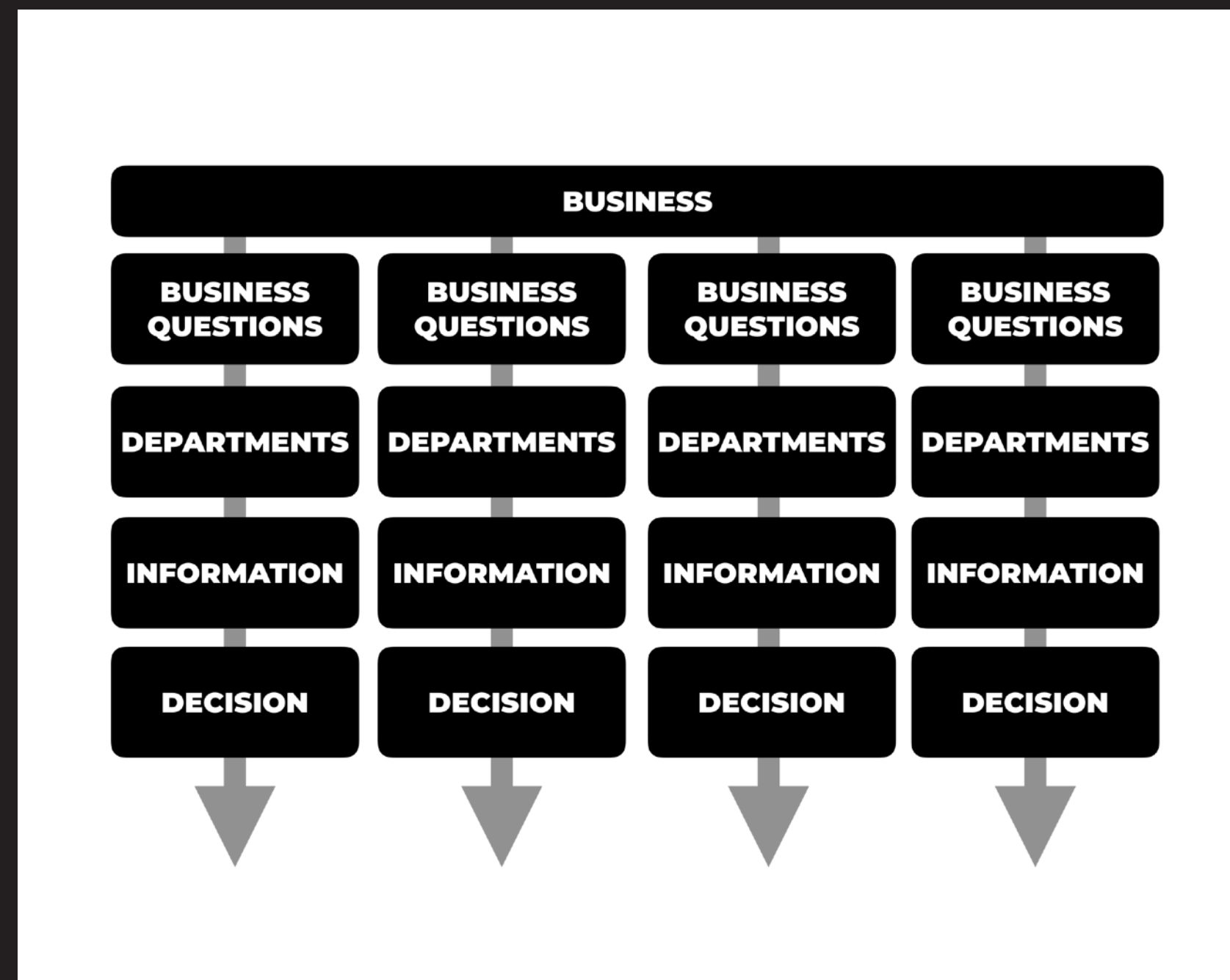
Enquanto isso, as empresas tradicionais tentam competir contratando jovens talentos com habilidades em digital. Essas pessoas são colocadas em escritórios e treinadas para seguir os processos para o melhor funcionamento do atual negócio, isso foi ensinado nos últimos 30 anos nas principais escolas de gestão do mundo... Mas não está mais funcionando.



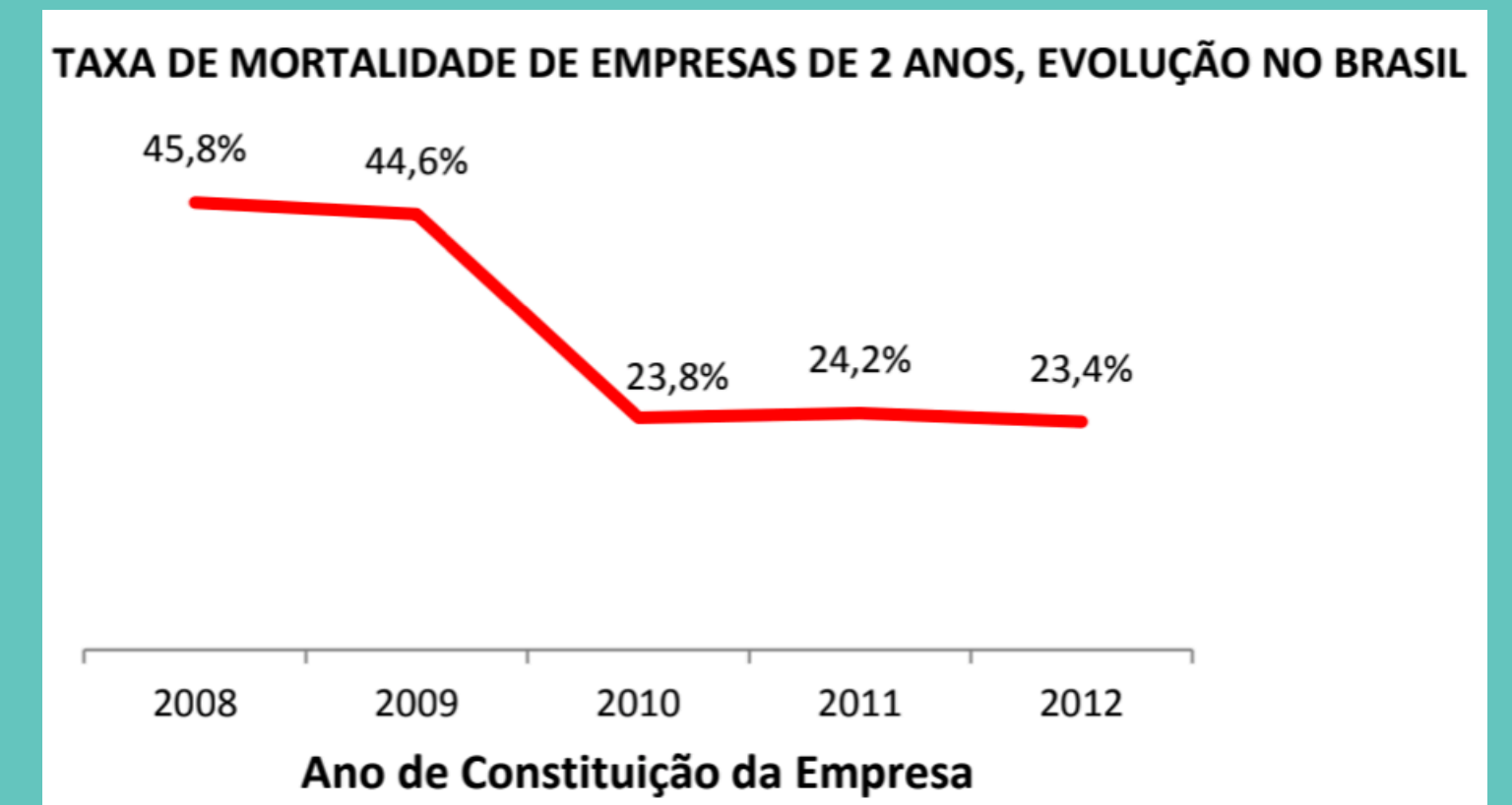


# empresas transformaram-se em organizações

Quando pensamos em uma empresa tradicional, na forma de funcionamento dela, isso logo nos remete para hierarquias funcionais, com fluxos pré-estabelecidos para que todos saibam o que precisam fazer e a quem precisam responder. São estruturas nascidas de modelos militares direcionados pela estrutura de comando/poder, que a indústria adotou como prática para administrar de forma mais eficientes os recursos e, conseqüentemente, os processos.



Agora pense em uma startup, ela nasce de uma ideia. O idealizador conta seu objetivo para algumas pessoas, então reúne um grupo e começam a realizar juntos as tarefas em torno da ideia. Normalmente isso ocorre de forma desestruturada, apenas se organizam para evitar que façam trabalhos repetidos entre eles. Essas organizações, sem processos previamente estabelecidos, colocam seus produtos ainda estágio experimental no mercado. Caso o produto não funcione, mudam e testam diretamente no mercado, até que a ideia se torne uma experiência positiva para seus consumidores. Essas organizações são mais ágeis e dinâmicas, elas não precisam estabelecer processos para serem competitivas, elas funcionam conforme a demanda, adaptando seu modelo de funcionamento às necessidades do mercado.



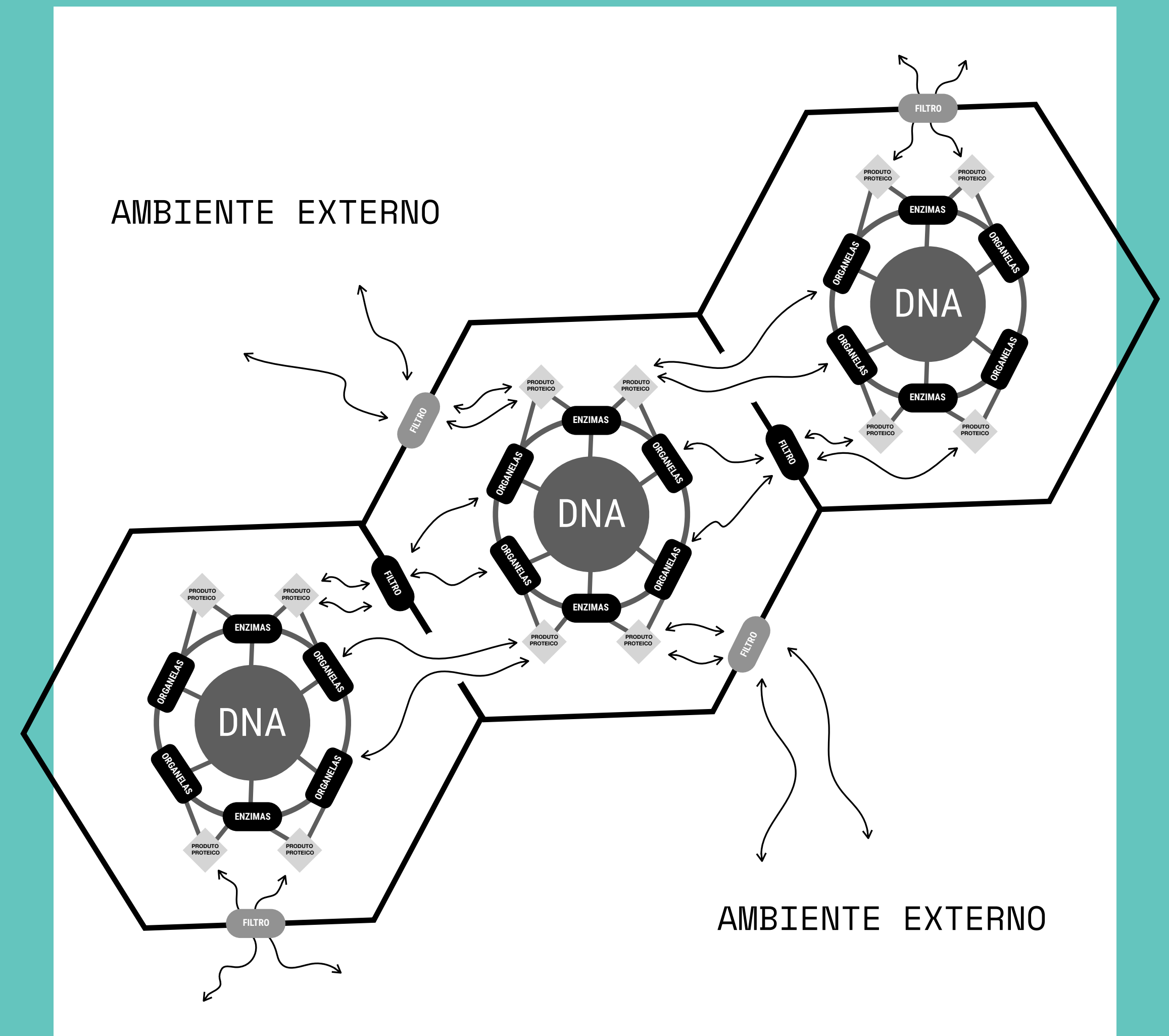
Até 2008, a taxa de mortalidade de empresas de 2 anos era de 45,8%, a partir de 2010 a taxa caiu e se mantém em torno de 23,5%, ou seja, as empresas novas estão sobrevivendo mais nesse mercado mais dinâmico.

**A verdade é que esse tipo de organização baseada em métodos ágeis, independente do segmento, porte e/ou lugar que ela esteja, se adaptam ao ecossistema e sobrevivem.**

# organizações viraram organismos vivos

Isso não significa que essas novas organizações que estão no mercado não tenham processos, elas apenas não usam isso como um mecanismo prévio para funcionamento do negócio, o que elas fazem é ajustar seus fluxos conforme a necessidade para a própria sobrevivência, exatamente como todos os organismos vivos funcionam.

Fomos buscar na biologia uma explicação para o que está acontecendo nos negócios, e aquelas empresas que possuem modelos adaptativos, com sistemas mais orgânicos, são as que estão sobrevivendo nos ambientes complexos e de alto risco, como é o caso do mercado atual. Nosso sistema biológico humano funciona exatamente da mesma forma, as células recebem informação do ambiente externo, filtram e ajustam para que nosso organismo funcione mesmo em uma adversidade, é um sistema comunicativo em loops de feedback, que garante ao nosso organismo uma constante adaptabilidade.





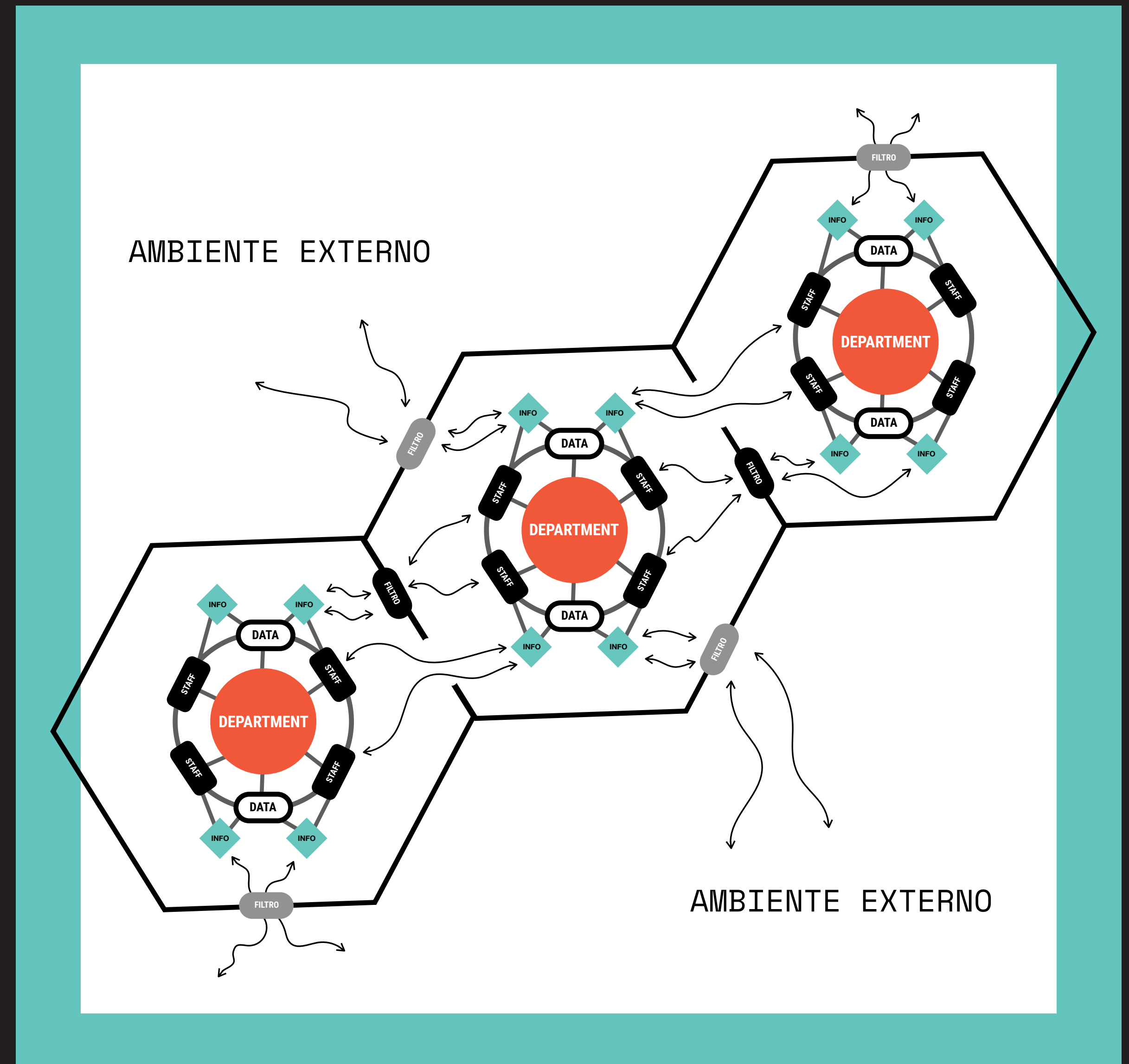
# informação é o combustível desse organismo

Se visualizarmos uma organização como um organismo, rapidamente perceberemos a semelhança entre os dois sistemas, ambos buscam se manter atualizados em tempo real para garantir a sobrevivência, de acordo com as interações que recebem.

Assim como em qualquer organismo vivo, as organizações também se movimentam de acordo com o fluxo de informação que recebem e usam. As empresas tradicionais, com funcionamento baseado em hierarquia e processo, se movimentam de acordo com a orientação que recebem das lideranças. Já as organizações novas usam a informação como combustível, e não o processo, e isso muda tudo.

Essas novas organizações ativam seus sistemas adaptativos assim que recebem novas informações - que chegam de todos os lugares, nos mais diferentes formatos, em grande volume, e em tempo real - e reagem com ações para garantir sua sobrevivência no mercado. Como reagem imediatamente aos estímulos, geram novas informações que ensinam algo novo ao organismo, gerando assim uma cadeia adaptativa completa baseado em informação. Essa forma de funcionamento das organizações não se diferencia em nada de um organismo vivo, quanto mais informação, mais o sistema irá aprender e se adaptar, buscando uma melhor eficiência funcional.

**É o sistema evolutivo dos negócios, baseado em dados e informação.**

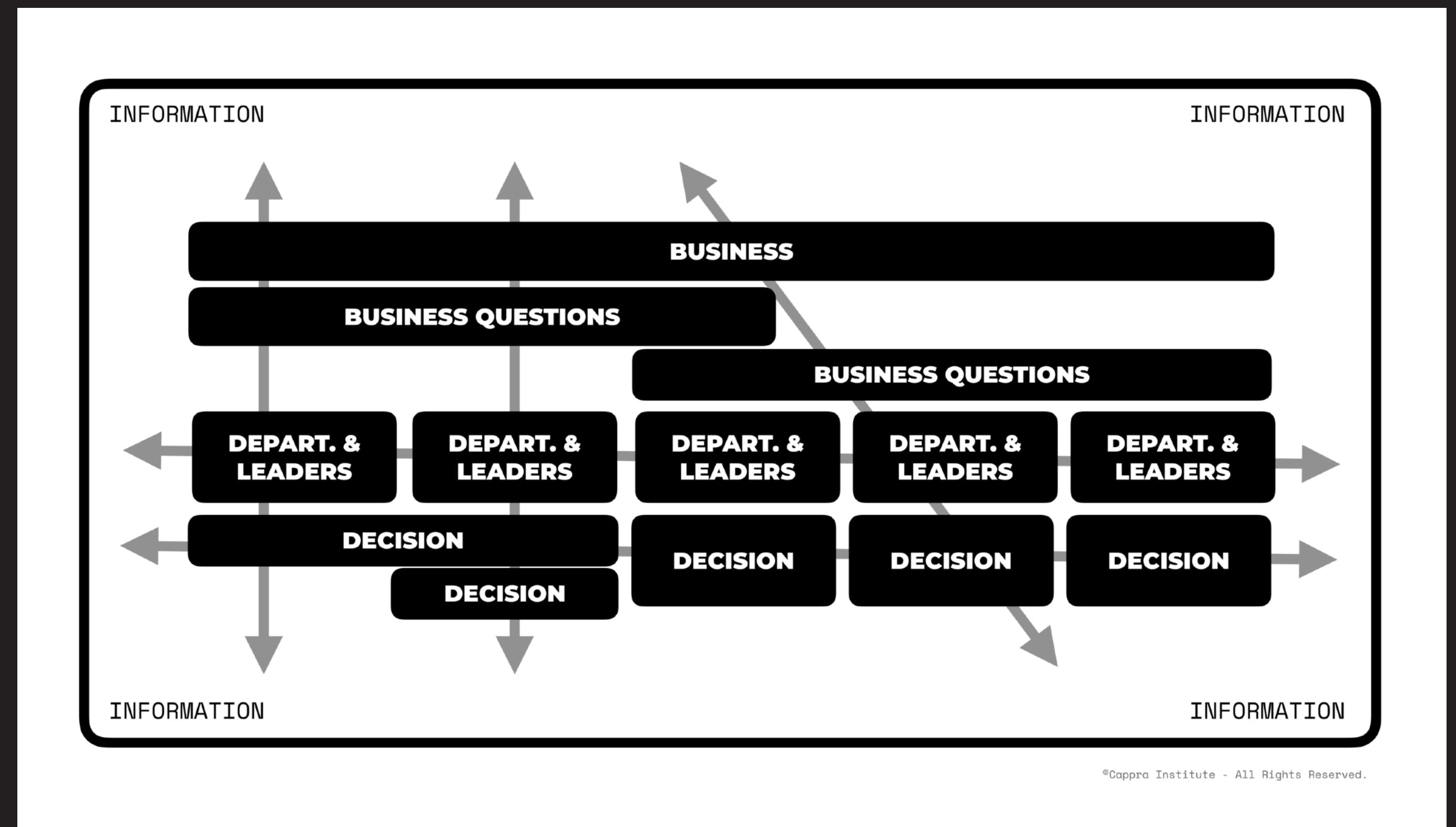


# redesenhando o fluxo de informação

Em um organismos novo, tudo se adapta para a sobrevivência.

Empresas tradicionais, que pré-estabelecem os processos inibem a circulação da informação, fazendo com que o sistema não funcione organicamente e precise de interação humana para liberar o fluxo de ação. É similar a uma *intervenção cirúrgica* cada vez que o organismo recebe um novo lote de informação, quando na verdade deveríamos dar espaço para o sistema se adaptar e reagir rapidamente. A informação deveria permear toda a organização, permitindo novas perguntas de negócio, gerando novas células/áreas para atender as novas demandas, libertando as pessoas para um processo decisório mais fluído e dinâmico, ou seja, orgânico.

Sistemas orgânicos também podem ser medidos, mas vamos precisar de novas métricas, pois os indicadores que usamos nos negócios atuais são baseados baseados em processos e não em fluxo de informação.





# tecnologia da informação é o acelerador dos organismos

A transformação digital criou esses organismos de negócio reagentes à dados, então precisaremos adotar novas formas para lidar com esses novos fluxos.

Não é possível “controlar” a circulação da informação, mas podemos ajudar o organismo a lidar com esses dados de forma mais natural, criando mecanismos receptores para tradução da informação que permita ação instantânea. Os famosos planos de ação, poderiam ser chamados de *planos de reação*, já que esses organismos vivos estão reagindo a toda informação que estão recebendo. O bom uso da tecnologia da informação pode ser um acelerador para administrar o volume de dados que essas células/negócios estão recebendo, tornando os negócios mas dinâmicos e eficientes em seu funcionamento.

Empresas que não funcionarem como organismos sempre terão a dependência da informação já filtrada pelo dono/diretor/gerente, fazendo com que seja um sistema de replicação e não um organismo adaptativo. Sistemas orgânicos são mais eficientes pois criam mecanismos naturais para a sobrevivência, mas precisam de fluxos de informação sem interrupções, com espaço para reações instantâneas e modelos que permitam a adaptação constante para aumentar a eficiência.

Assim como a Teoria da evolução humana, a Teoria da evolução das organizações é totalmente baseada na capacidade de adaptação para sobrevivência, sendo que o agente transformador não é o organismo em si, e sim a informação que ele recebe e usa.





# análise de impacto na sociedade e organizações

Os analistas do Cappra Institute investigaram o potencial do impacto desse tema considerando 3 fatores:

**Novidade:** o quanto o assunto é novo

**Atividade:** o quanto o assunto é popular

**Impacto:** o quanto vai impactar em 2020

Com essa análise desenvolvemos um **Score** que representa o impacto desses 3 fatores, considerando uma variação entre 0 e 10.

Esse score é um bom indicador de senso de urgência de ação nesse tema.

SCORE

# 6

NOVIDADE

ATIVIDADE

IMPACTO

<infororganization>

Google Trends | Pesquisar

Termo de pesquisa	Assuntos relacionados
1. abrir empresa	1. Microempreendedor individual - Assunto
2. empresa mei	2. Cadastro Nacional da Pessoa Jurídica - As...
3. abrir empresa mei	3. Dinheiro - Assunto
4. mei	4. Empréstimo - Assunto
5. abrir mei	5. Empreendedorismo - Assunto

Interesse ao longo do tempo

Design thinking vs. Transformação Digital

Interesse ao longo do tempo

Detalhamento comparado por região

Região	Design thinking	Transformação Digital
1. Polónia	High	Low
2. Taiwan	High	Low
3. Brasil	High	Low
4. México	High	Low
5. Nova Zelândia	High	Low

2020

CURIOSIDADES

## <infororganization>

### CASO RECENTE

EXAME

Ranking do movimento 100 Open Startups mede negócios escaláveis e inovadores mais atrativos para corporações -- e, agora, para investidores anjo

Por Mariana Fonseca  
22 jul 2019, 06h00

Funcionários em startup: 212 startups ranqueadas receberam mais de 142 milhões de reais em investimentos (KatarzynaBialasiewicz/Thinkstock)

A Allya, startup com um aplicativo de descontos personalizados a funcionários, é novamente a mais desejada pelas grandes corporações. É o que mostra a nova edição do ranking 100 Top Open Startups, que elenca os negócios escaláveis e inovadores mais atrativos, entre mais de 8,6 mil startups cadastradas, na opinião de 15 mil executivos de mais de 2,2 mil empresas com faturamento de 100 milhões de reais ou mais.

A pesquisa, elaborada pela rede 100 Open Startups, leva em consideração critérios como posicionamento da startup; plataforma e parcerias; desenvolvimento de fornecedores; e investimento. São considerados contratos ativos entre junho de 2018 e maio de 2019. O limite de faturamento em 2018 ou investimento recebido deve ser de 10 milhões de reais por cada startup.





## POR QUE VOCÊ DEVE OBSERVAR ESSA TENDÊNCIA EM 2020?

Já fazem muitos anos que as startups são os grandes agentes de transformação do mercado, mas é importante compreender que o segredo desses negócios vai além do seu modelo de funcionamento e *espaços modernos*, seu diferencial está na forma como lidam com a informação, isso transforma os processos em algo orgânico.

Empresas com melhores dados, estão criando melhores produtos, atraindo mais consumidores e assim gerando mais dados. As empresas irão se adaptar aos modelos organizacionais baseando-se em sistemas de informação, assim como o genoma, as organizações evoluirão baseando-se na sequência de informação que capturam, analisam e utilizam para sua evolução e sobrevivência.

## RECOMENDAÇÕES:

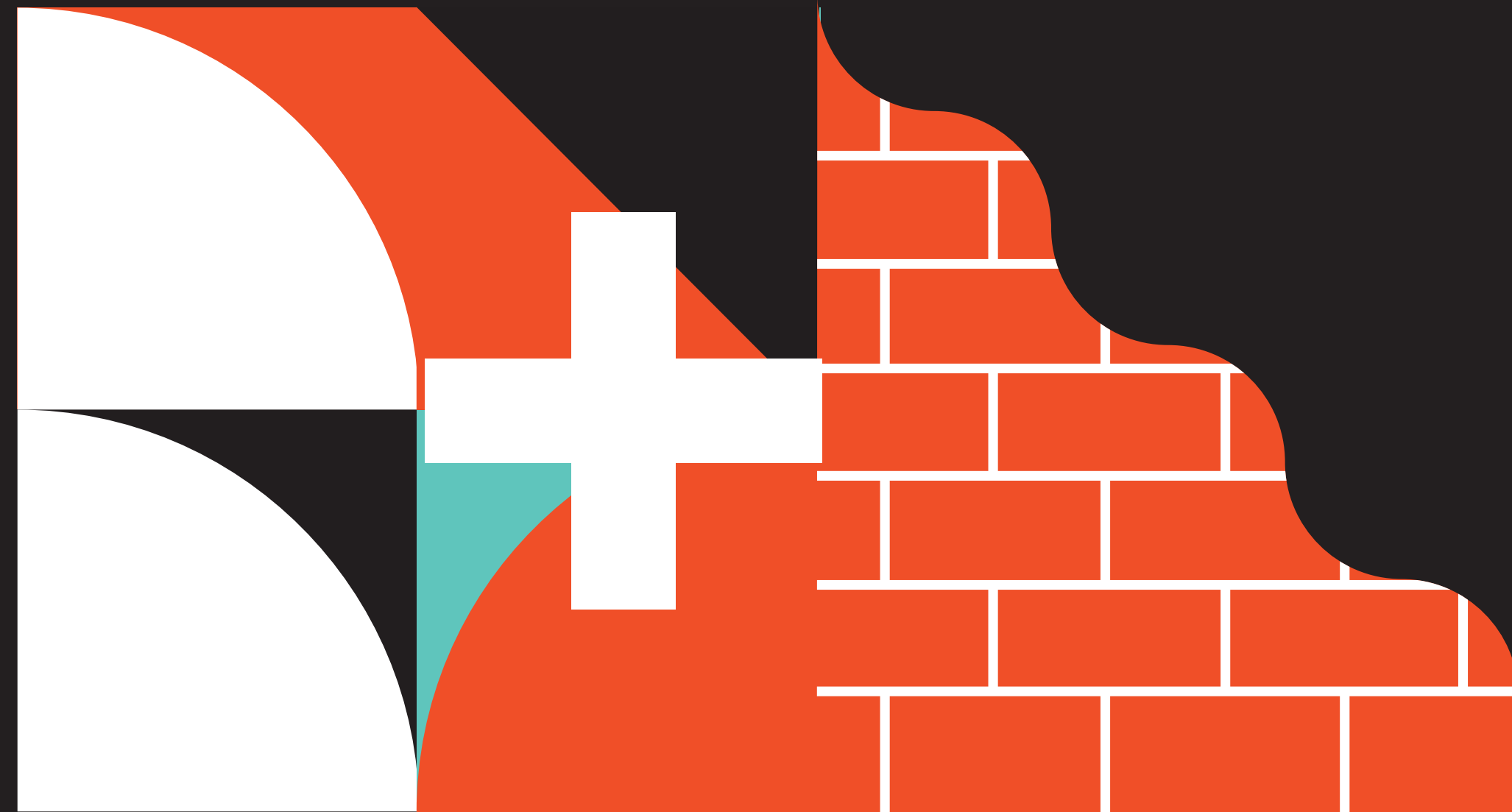
- Elimine as barreiras de entrada e circulação da informação na sua empresa, muitas delas são "invisíveis". Algumas se apresentam em forma de processo, outras como travas psicológicas de sustentação do poder.
- Não trave o negócio em torno dos Processos, deixe que os dados e informação sejam o verdadeiro combustível do negócio.
- Crie planos de reação baseado em informação, ao invés de planos de ação baseado em ideais.
- Use recursos da tecnologia para acelerar o consumo de informação.
- O DNA do seu negócio não precisa mudar. Sistemas orgânicos se adaptam mas não modificam sua essência, apenas potencializam seu funcionamento conforme a qualidade da informação que circula.

## NA PRÁTICA:

Em 30 dias: Identifique - escreva - os principais gaps da circulação da informação no negócio (podem ser pessoas, processos, tecnologias ou políticas).

Em 6 meses: Crie um plano de incentivo para as modificações nos processos que melhorem diretamente o uso/consumo de dados e informação.

Em 1 ano: Modifique algo significativo no organograma do seu negócio, destaque aquilo que comprove que seu modelo de gestão está mais orgânico.



# times data-driven

*“Times de negócio que estão mais preparados para lidar com dados, são mais competitivos e eficientes na era da informação.”*

CAPPRA, Ricardo.



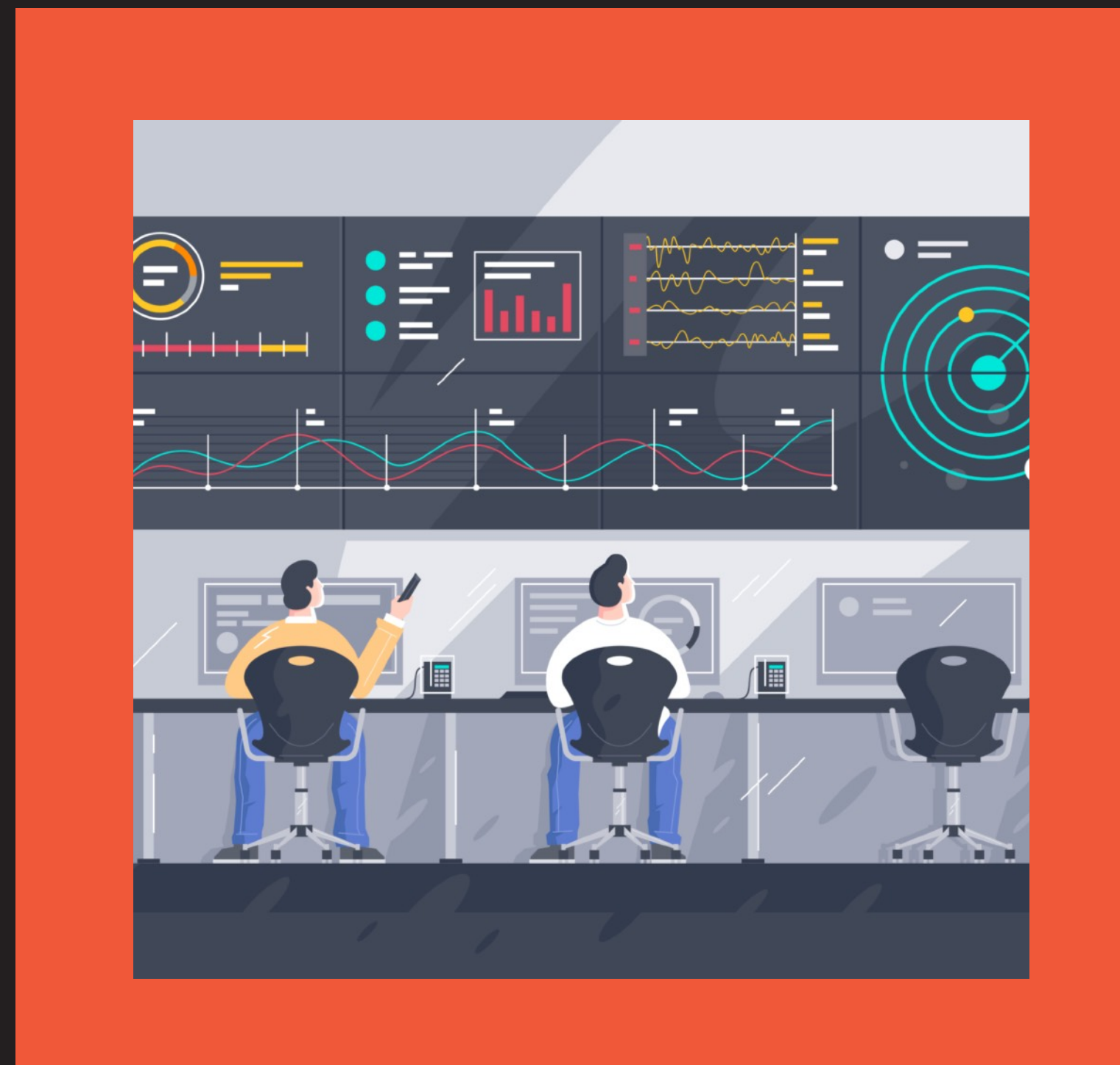
OPERAÇÕES ANALÍTICAS COMO  
DIFERENCIAL ESTRATÉGICO



# preparando uma organização data-driven

Nos laboratórios do Cappra Institute, passamos os últimos 5 anos decodificando a forma como as empresas trabalham com dados.

Nossa intenção é compreender o papel, as relações institucionais/hierárquicas, as expectativas com relação às atividades/habilidades, formações funcionais,..., por isso estudamos aqueles modelos de funcionamento que são mais ou menos eficientes. Esses estudos nos levaram a construir ferramentas específicas para organizações que estão no processo de transformação data-driven, e ficou muito evidente que os principais questionamentos são sobre a formação de operações mais analíticas e desenvolvimento de times data-driven.



Nesse cenário atual de mercado - pós-transformação digital - as perguntas giram em torno do legado gerado pela tecnologia, aquele “amontoado de dados”:

*Mas e agora?!*

*Como vamos nos estruturar para deixar o time de negócios mais eficiente usando esses dados?*

*Quem vai fazer a transformação desses dados em informação?*

*Estamos prontos para ter algoritmos e inteligência artificial?*

*Que área deve cuidar dessa disciplina na empresa?*

*Essas são apenas algumas das perguntas que encontramos ao conversar com donos e líderes de negócios ao redor do mundo.*

Resolvemos trazer um pouco de nossas descobertas para ajudar a esclarecer como as empresas se organizam para suportar a evolução data-driven, algo indispensável para os negócios na Era da Informação.

# times de dados ou times data-driven?!

Em nossos estudos, descobrimos que a maior preocupação das empresas é com a formação de um Time de Dados, ou seja, técnicos/especialistas que trabalham na transformação dos dados em análises. **Cerca de 68% das empresas coloca entre suas prioridades a formação dos Times de Dados**, acreditando que esse é o caminho mais curto para que suas empresas sejam data-driven.

## [TIMES DE DADOS]

Um **Time de Dados eficiente é aquele que atua de forma plena na transformação dos dados em informação qualificada, organizada e visual.**

Esse trabalho é normalmente realizado por especialistas da área de tecnologia, com algum complemento especializado em estatística/matемática dependendo da complexidade das análises envolvidas no negócio.

*Os times de dados precisam de interlocutores para dialogarem de forma mais eficiente com os times de negócio, pois isso aproxima as demandas do trabalho técnico, evitando retrabalho e otimizando a atuação de ambas as partes.*

Em resumo, os Times de Dados precisam possuir as habilidades técnicas para a evolução analítica do negócio ocorrer da forma mais rápida e eficiente possível.

## [TIMES DATA-DRIVEN]

Os esforços atuais na preparação dos times de negócio para lidarem melhor com os dados são muito menores, geralmente aparecem como iniciativas isoladas de treinamento da área de RH, ou de alguns estímulos esporádicos dos presidentes das empresas para adoção da cultura data-driven.

**As empresas são mais eficientes e ágeis quando os dados fazem parte de toda a jornada de tomada de decisão**, desde o momento das perguntas de negócio até o momento da mensuração das decisões tomadas, ou seja, a mentalidade e práticas analíticas precisam fazer parte de todo o ciclo do negócio.

Ter uma operação mais analítica é um diferencial competitivo necessário para a área de negócio, esse tipo de conhecimento não pode ficar restrita somente aos especialistas dos Times de dados, **toda a organização precisa ser data-driven.**



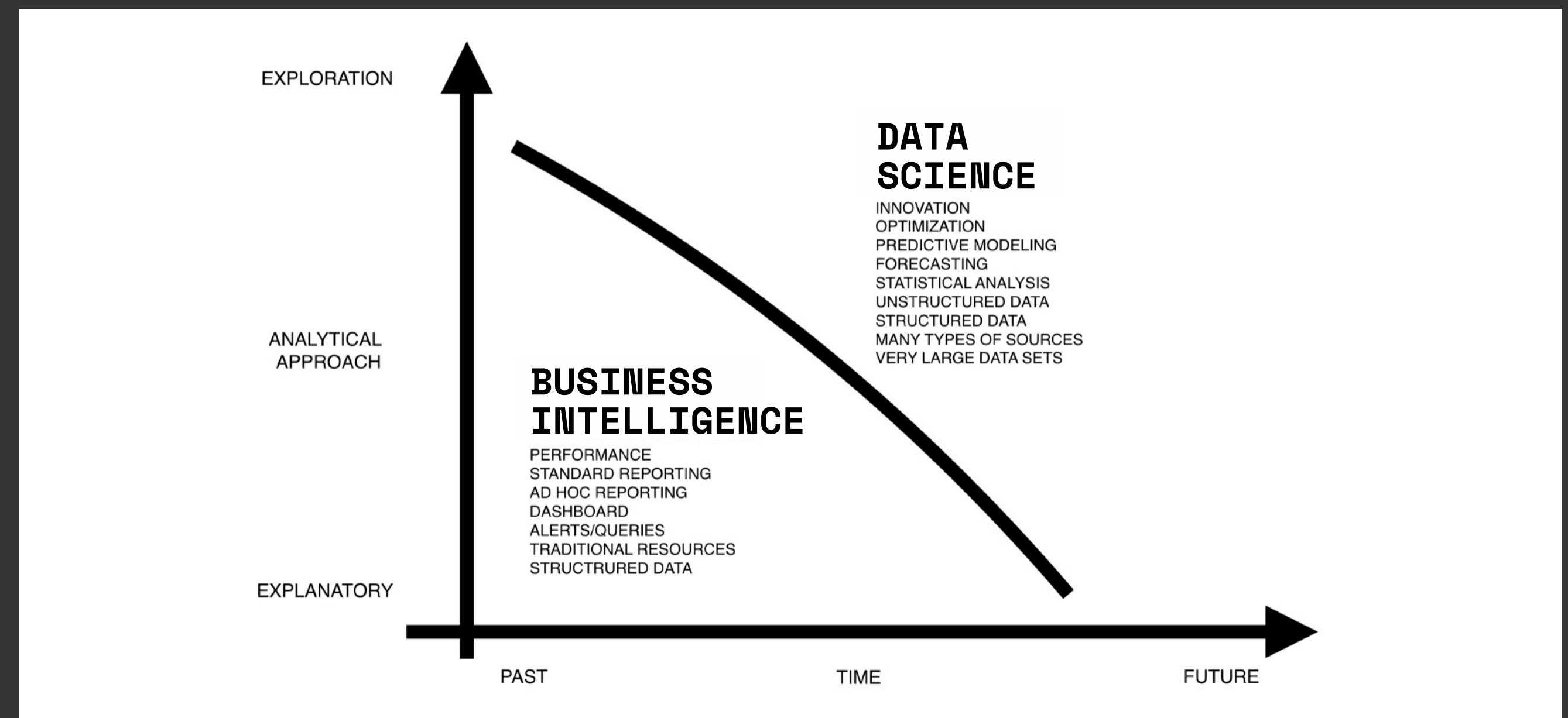
# business intelligence vs data science

Dois termos constantemente geram confusão no linguajar de negócios, aqui vamos esclarecer como os executivos estão percebendo essas duas disciplinas como suporte ao negócio.

O termo **Business Intelligence** (ou simplesmente BI) é algo muito antigo, sempre esteve muito relacionado ao processo de análises de negócio, com cruzamentos de informações e criação de relatórios para dar suporte às decisões.

Já o termo de **Data Science** (ou ciência de dados) é mais recente, ele surge como uma camada mais profunda do tratamento de dados e análises avançadas, usando novas técnicas e tecnologias para melhorar a eficiência analítica.

Ao mapear a evolução da maturidade analítica dentro de empresas de diferentes locais do mundo, descobrimos que existe um certo consenso com relação a classificação das atividades envolvidas e habilidades necessárias, no gráfico podemos ver a classificação mais comum dessas disciplinas. Importante ressaltar que não estamos nos referindo a área de negócio, e sim de atividades relacionadas as áreas de conhecimento.



## [BUSINESS INTELLIGENCE]

A disciplina de Business Intelligence está diretamente relacionada com as atividades de análise de performance - passado e presente. O papel de apoio ao dia-a-dia do negócio, com recursos como dashboards, relatórios, alertas, ..., aparecem aqui como destaques. Essa relação direta com a tomada de decisão torna indispensável manter esse conhecimento em constante evolução, pois a informação precisa ser ágil e sem erros.

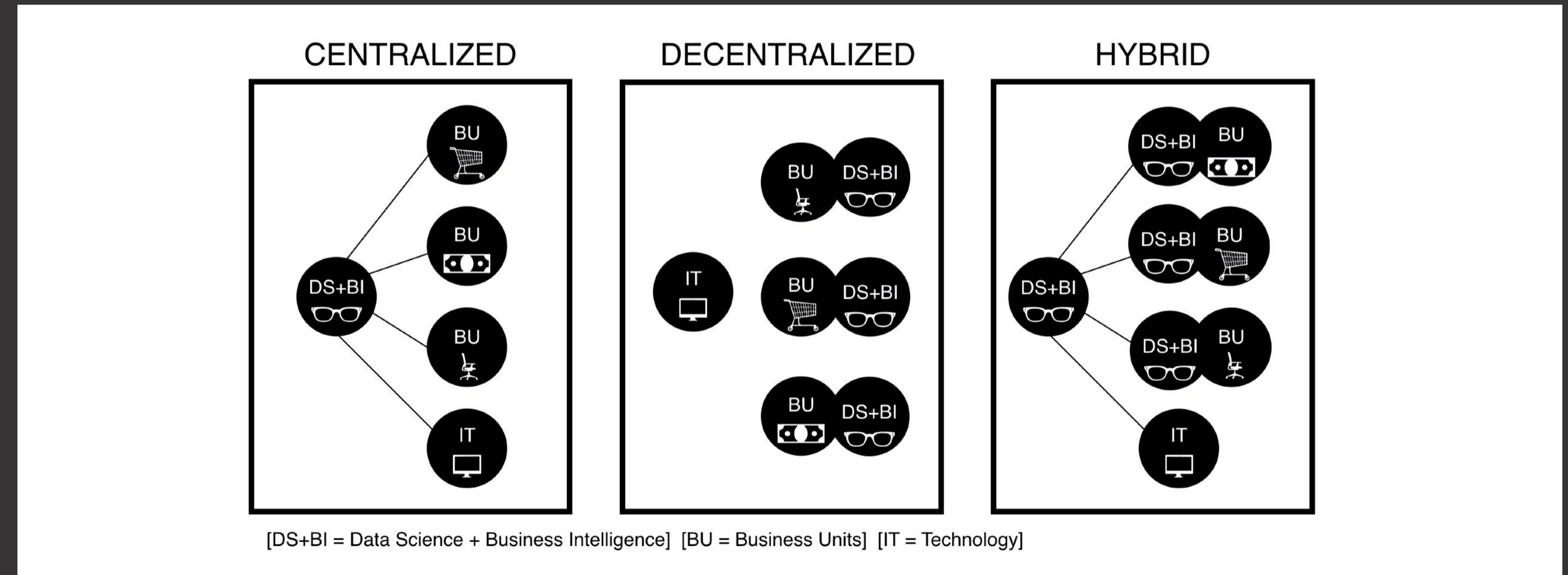
## [DATA SCIENCE]

Data Science está mais relacionada com explorações de cenários futuros, hipóteses ainda não investigadas, geralmente está associada ao uso de técnicas mais avançadas. Modelos matemáticos complexos e tratamentos de dados não estruturados aparecem como uma competência de Data Science, por isso a associação direta com inovação.

# modelos de operações analíticas

Com a evolução da maturidade analítica nas organizações, um dos grandes desafios tem sido ajustar a forma de operação da área de dados, para que ela funcione mais integrada e fluída junto das áreas de negócio.

Em nossos estudos sobre o funcionamento das operações analíticas nas organizações, descobrimos 3 formas mais utilizadas:



## [CENTRALIZADO]

Aqui a área de DS + BI opera como um “prestador de serviços de informação” para as áreas de negócio, é muito parecido com o modelo atual de BI nas empresas, mas com o incremento das competências relacionadas a ciência de dados.

**FORÇAS:** as empresas já estão acostumadas a trabalhar departamentalizado e isso acelera a adoção; as áreas de negócio mantêm sua independência e podem tratar a área de DS+BI como um prestador de serviço; as entregas são feitas em forma de produtos e isso facilita a compreensão dos usuários finais.

**FRAQUEZAS:** desconhecimento da área de negócio das competências analíticas; distância entre o time de dados dos usuários finais; interfaces de comunicação costumam falhar gerando grandes volumes de retrabalho no time de dados.

## [DESCENTRALIZADO]

Aqui a área de DS + BI é “desmontada”, levando as habilidades analíticas para dentro do negócio. Cada área possui seus próprios recursos para trabalhar com dados.

**FORÇAS:** a área de negócio leva as habilidades analíticas para dentro das operações que realiza; o recurso de DS+BI torna-se especialista naquela operação de negócios e isso melhora as entregas de informação qualificada; a independência gera um ambiente de inovação constante; liberdade para criar modelos mais avançados independente do nível das outras áreas do negócio.

**FRAQUEZAS:** esse formato possui um alto índice de retrabalho para os recursos de DS+BI, pois criam modelos praticamente iguais em diferentes áreas; esse formato facilita o *desnivelamento* da maturidade analítica da empresa como um todo; as áreas de negócio que contratam melhores recursos levarão vantagem.

## [HÍBRIDO]

Nesse formato a área de DS+BI possui recursos dentro da área de negócio, mas mantém uma célula independente para gestão do conhecimento especialista.

**FORÇAS:** agilidade no atendimento de demandas de negócio especialistas; existência de uma célula dedicada a pesquisa e desenvolvimento para gestão do conhecimento especialista; ciclos de desenvolvimento analítico mais uniformes no negócio.

**FRAQUEZAS:** conflitos na gestão matricial dos recursos; custos dessa operação, pois inevitavelmente existirão sobreposições de atividades; empresas não estão habituadas em investirem em uma área de pesquisa e questionam os investimentos relacionados; esse modelo pode “retardar” a evolução de uma determinada área de negócio por políticas que determinam o uso de padrões de ferramentas e métodos.



# análise de impacto na sociedade e organizações

Os analistas do Cappra Institute investigaram o potencial do impacto desse tema considerando 3 fatores:

**Novidade:** o quanto o assunto é novo

**Atividade:** o quanto o assunto é popular

**Impacto:** o quanto vai impactar em 2020

Com essa análise desenvolvemos um **Score** que representa o impacto desses 3 fatores, considerando uma variação entre 0 e 10.

Esse score é um bom indicador de senso de urgência de ação nesse tema.

**<operações analíticas>**

**SCORE**  
**7**

**NOVIDADE**  
[Progress bar]

**ATIVIDADE**  
[Progress bar]

**IMPACTO**  
[Progress bar]

Google Trends Pesquisa

Interesse ao longo do tempo

• Data-driven  
Assunto

Todo o mundo No últimos 5 anos Todas as categorias

Interesse por região

Assuntos relacionados	Em ascensão	Consultas relacionadas
1 Inovação - Assunto	1 data driven attribution model	
2 Experiência - Filosofia	2 testing annotations	
3 Cucumber - Software	3 toolica	
4 Aprendizado de máquina - Campo de estu.	4 b7b data driven marketing	
5 Object model - Assunto	5 data driven analytics	

Consultas relacionadas	Em ascensão	Consultas relacionadas	Em ascensão	Consultas relacionadas	Em ascensão
1 data analytics	1 coursera	1 big data hadoop			
2 analytics	2 coursera data science	2 hadoop			
3 hadoop big data	3 data science and analytics	3 big data and analytics			
4 big data and analytics	4 data science in python	4 big data course			
5 big data course	5 data science certificate	5 machine learning			

• business intellige...  
Termo de pesquisa

• data science  
Termo de pesquisa

• big data  
Termo de pesquisa

Todo o mundo 05/12/2009 - 05/01/2020 Todas as categorias Pesquisa na Web

Interesse ao longo do tempo

Média

Observação

CURIOSIDADES

CAPPRA INSTITUTE FOR DATA SCIENCE

<times data-driven>

## CASO RECENTE



### Sua empresa está pronta para o futuro data-driven?

Fabio Coelho, Rafael Russo / Março de 2018 / Mobile, Novas Tecnologias, Métricas, Programática, UX e Design



Todo mundo sabe da importância dos dados. Basta notar que as empresas que mais cresceram nos últimos anos tem no seu core esse compromisso de ser *data-driven*, e existe uma enorme oportunidade para muitas outras chegarem aos estágios mais avançados dessa transformação.

Dados são uma riqueza ainda tão pouco explorada, que ela não está sendo nem mesmo avaliada formalmente nas análises de valor das empresas. É como se a gente avaliasse o solo, mas ignorasse o potencial escondido no subsolo. No entanto, é apenas uma questão de tempo até o valor da massa de dados de um negócio começar a fazer parte do balanço das empresas, assim como o *goodwill* de marca já faz hoje. Segundo a



## POR QUE VOCÊ DEVE OBSERVAR ESSA TENDÊNCIA EM 2020?

É praticamente impossível pensar na evolução da maturidade analíticas das organizações sem pensar nos modelos de funcionamento dentro das atuais hierarquias dos negócios, então esse é um tema que vai permanecer muito ativo durante 2020, seja para pequenas, médias ou grandes empresas.

É natural que negócios mais orgânicos/ágeis - como as startups por exemplo - utilizem essas habilidades de forma mais fluídas, onde as próprias pessoas de negócio buscam especializações para trabalhar com dados, isso vai acelerar a cultura analítica nesses negócios. Por outro lado, empresas que precisam operar grande volumes de processos e informações precisarão de um formato que gere menos fricção no modelo atual, para amenizar a transformação para os colaboradores.

## RECOMENDAÇÕES:

- Não limite a construção de uma operação mais analítica aos seus atuais recursos disponíveis no time de dados.
- Os Dados precisam ser recursos do negócio todo e não uma exclusividade da área de dados, eles precisam ser facilmente acionáveis e estarem prontos para uso.
- As habilidades analíticas precisam estar distribuídas no negócio para acelerar a adoção de uma cultura data-driven
- Você precisa compreender como funciona o atual processo de uso de dados e consumo de informação da empresa, esse poderá ser um limitador para a evolução da maturidade analítica do negócio
- O modelo de operação analítica pode se tornar um grande diferencial competitivo, se as áreas de negócio forem mais ágeis ao usar informação podem tomar decisões melhores.

## NA PRÁTICA:

Em 30 dias: Observe sua organização e identifique a atual forma de funcionamento dos processos e recursos que transformam dados em informação (com foco no suporte à tomada de decisão).

Em 6 meses: Faça conversas com as lideranças que podem conduzir o processo de transformação, experimente alguns modelos de trabalho e instigue testes reais.

Em 1 ano: Monte um plano de negócio considerando DS+BI com um dos pilares fundamentais, seja ele como uma área ou como uma competência distribuída.



20  
20



# data-robots

*“Uma grande parte do trabalho do cientista de dados já está sendo realizada por robôs, de forma muito mais rápida e com menos erros.”*

AUTOMAÇÃO DA  
CIÊNCIA DE DADOS

CAPPRA, Ricardo.





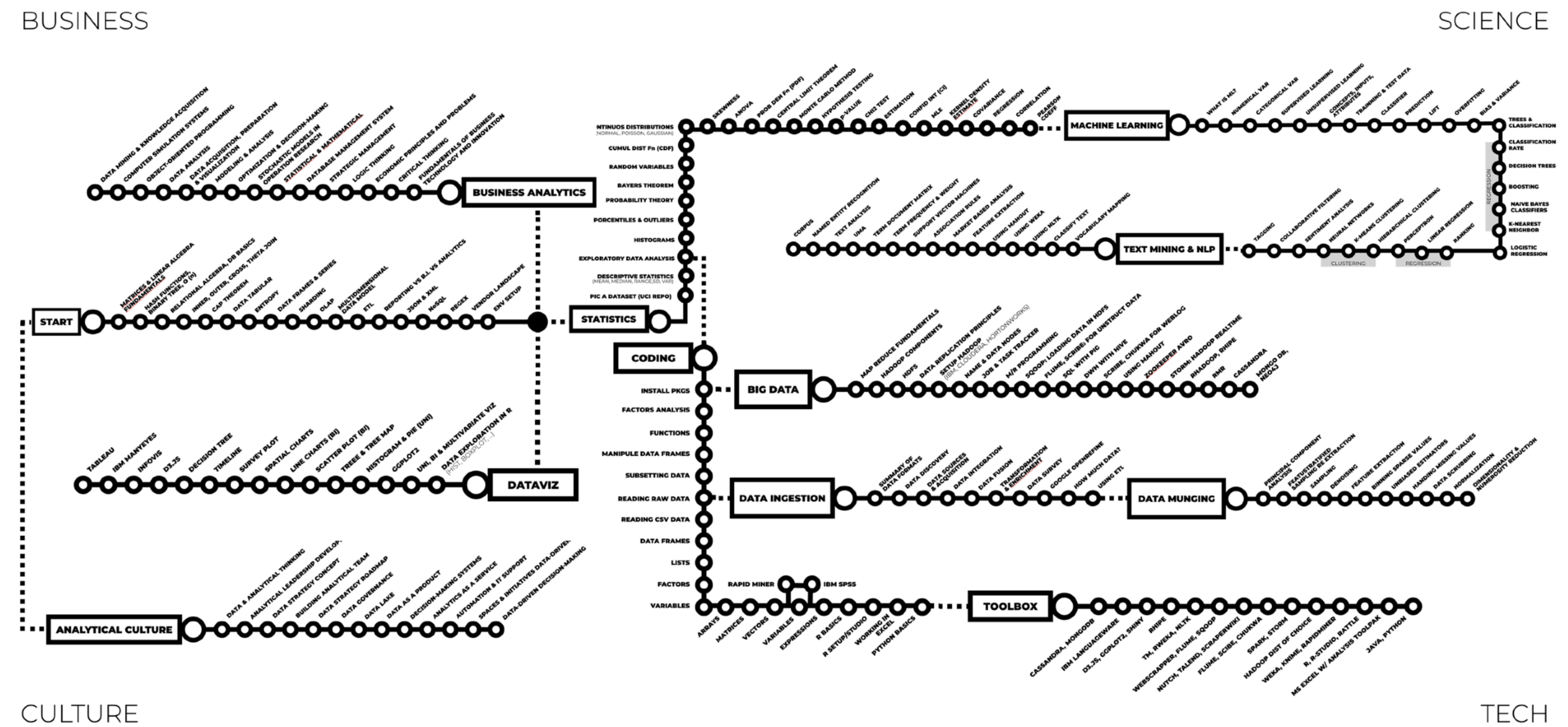
# decodificando a ciência de dados

A partir de uma inspiração de metromap criado por Swami Chandrasekaran para mapear o trabalho na área de ciência de dados, amplifcamos e aprofundamos o estudo investigando as habilidades necessárias.

Estruturamos em forma de um mapa de evolução para a carreira de cientista de dados, roteirizando os conhecimentos necessários e classificando em 4 grupos: **Negócios, Tecnologia, Ciência e Cultura.**

## TRILHAS DE CONHECIMENTO DA CIÊNCIA DE DADOS

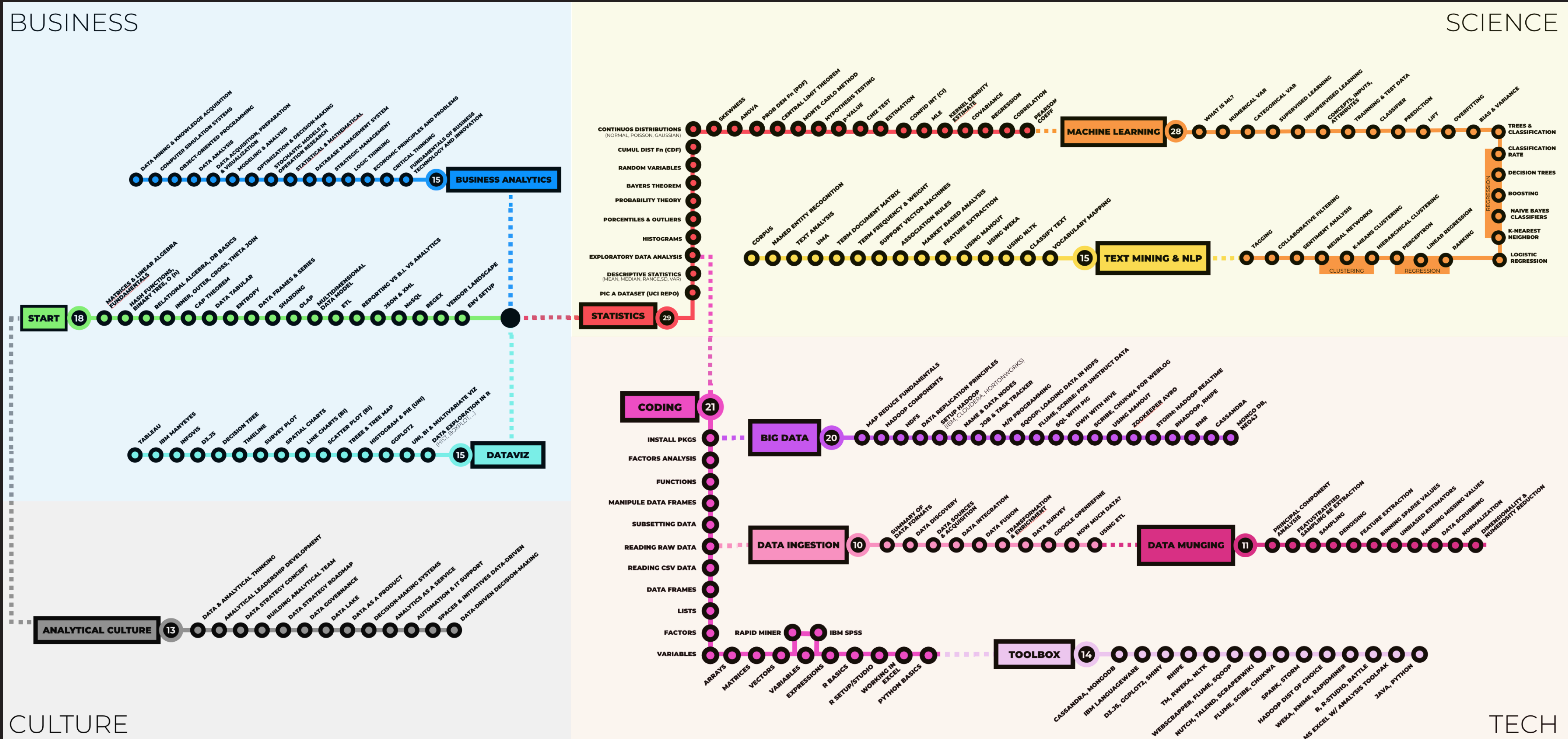
Dentro desses 4 grupos encontram-se diversas ferramentas e habilidades técnicas específicas, que criam um verdadeiro circuito do conhecimento sobre ciência de dados. Esse mapeamento nos permitiu investigar o assunto de forma mais estruturada, para entender a relação entre os temas, interferência das ferramentas nas atividades, entre outras características relacionadas ao tema.





# 206 skills da ciência de dados

O mapa completo da ciência de dados, organizada por área de conhecimento e agrupada por disciplina





# decodificando o cientista de dados

Criamos uma ferramenta para que os profissionais possam analisar seus próprios perfis e entender a amplitude da área de ciência de dados, cada pessoa que participa através do link [skills.cappra.institute](https://skills.cappra.institute), e faz análise do seu perfil individual, contribui com uma pesquisa colaborativa para compreensão do momento atual - e do futuro - dos profissionais que trabalham com dados. Durante todo o ano de 2019 coletamos informações, e essa base de conhecimento permitiu uma profunda investigação de quem atua na área.

A coleta de informações ocorre por nível de conhecimento, para cada skill existente o participante identifica seu nível de habilidade, sendo:

0 - Não conheço

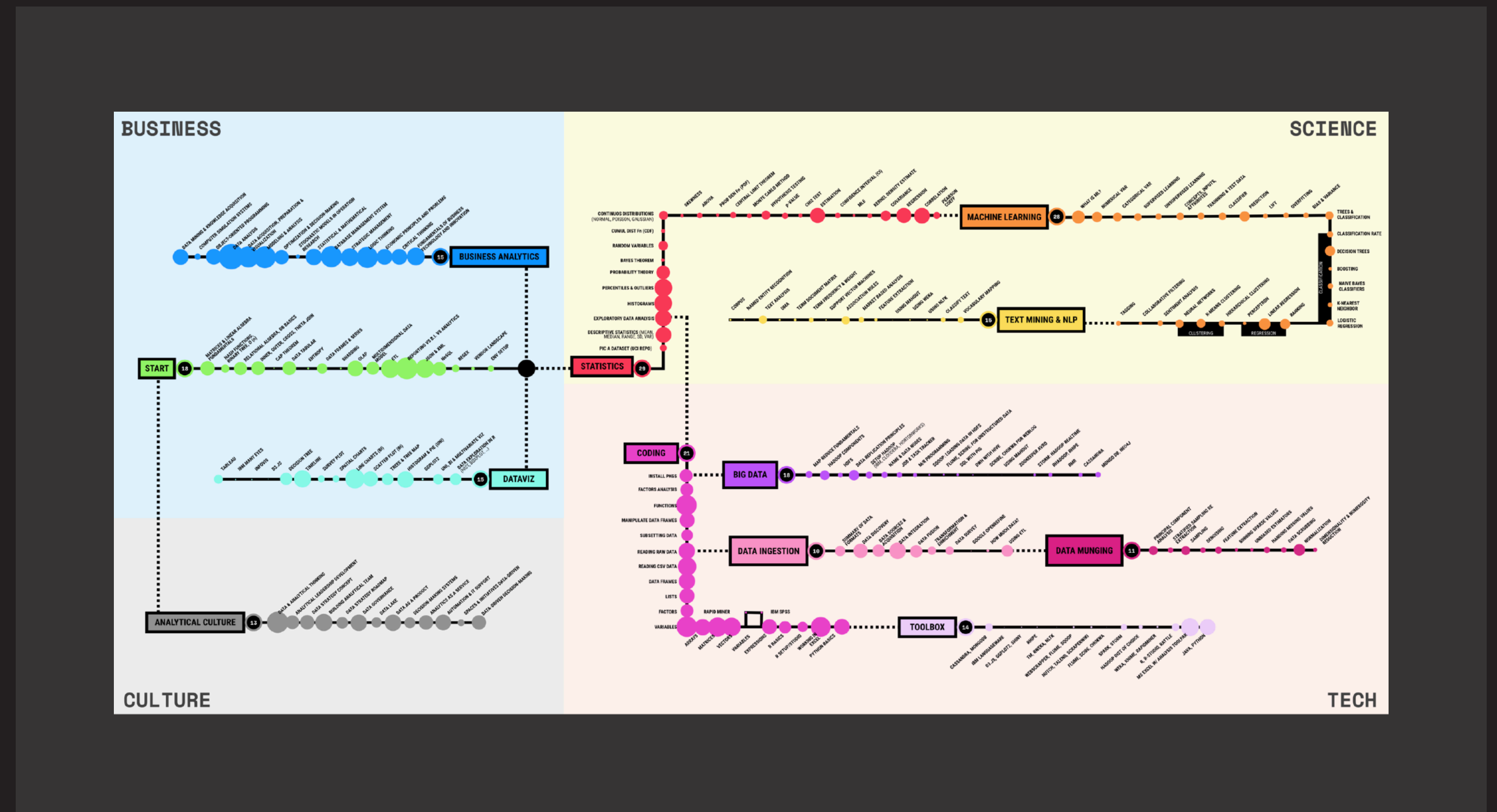
1 - Já ouvi falar

2 - Sei conversar sobre o assunto

3 - Já fiz algo com isso

4 - Uso no meu dia a dia

5 - Consigo ensinar alguém

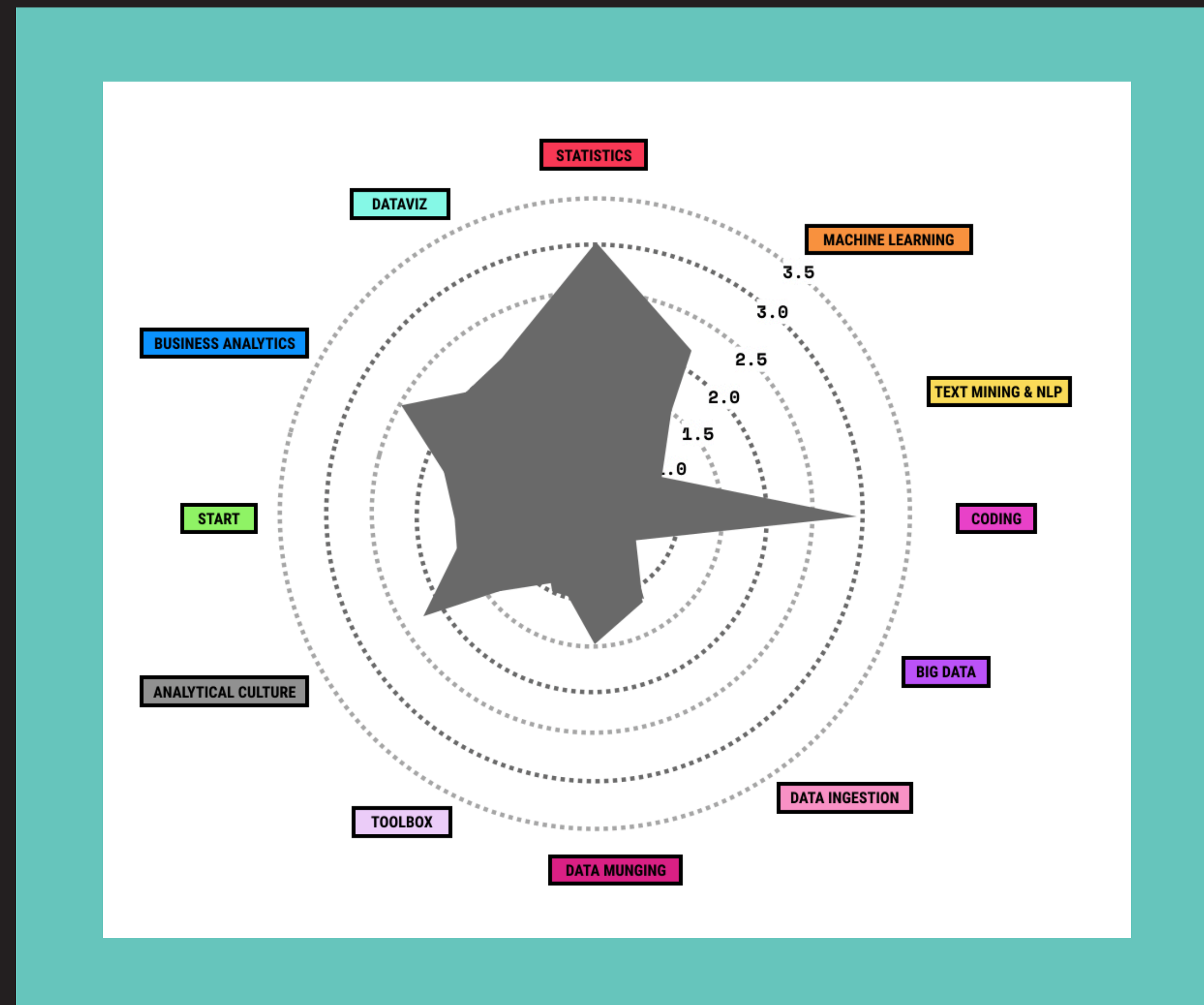


Na representação gráfica do público participante geral, podemos visualizar as habilidades e conhecimentos práticos que esses profissionais possuem, e onde estão mais concentrados os conhecimentos, ou seja, em que atividades esses profissionais estão mais preparados para atuar. É possível perceber isso pela tamanho do círculo representado junto da skill, quanto maior o círculo, mais pessoas reconhecem isso como habilidade já desenvolvida.

# perfil atual dos profissionais de ciência de dados

Se visualizarmos os resultados de nossos estudos em uma análise radial, vamos perceber que o conhecimento geral é concentrado nas habilidades e ferramentas mais populares de mercado, sendo que o conhecimento mais especialista ainda tem muito espaço para desenvolvimento profissional.

Analisando esse gráfico podemos afirmar que existe uma uniformidade com relação ao conhecimento adquirido para atuar na área de ciência de dados. Os profissionais obviamente não conseguem ser especialistas em todas as competências exigidas pelo mercado.





# robôs farão muitas tarefas da ciência de dados

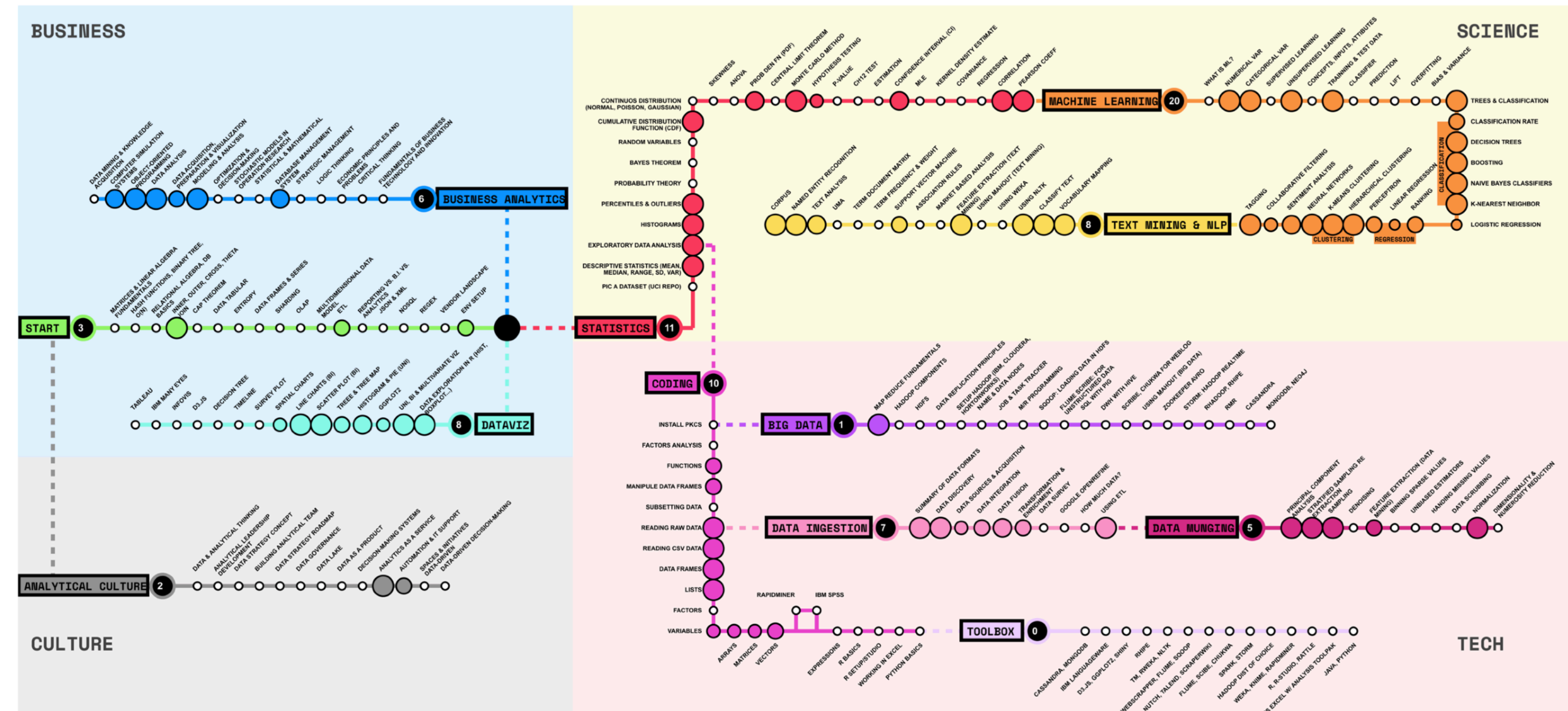
Mas nossas descobertas através desse estudo vão além da trilha de conhecimento.

Investigamos as atividades que já podem ser automatizadas por softwares de mercado, e correlacionamos com as habilidades/skills necessárias ao trabalho do cientista de dados.

Das 206 skills relacionadas ao trabalho do cientista de dados, quase 40% delas (81 skills) já estão em nível avançado de formação, ou seja, essas habilidades já podem ser substituídas para tarefas executadas por máquinas. Importante ressaltar que esse levantamento foi realizado considerando as ferramentas atuais de mercado, ou seja, esse cenário pode se modificar rapidamente com um nível de automação ainda maior das skills da profissão.

## [DATA-ROBOTS]

São Cientistas de Dados Virtuais - robôs e softwares - que automatizam tarefas que são baseadas em repetição e regras pré-definidas. Os data-robots realizam essas tarefas em menor tempo, com maior precisão e menor esforço.



No gráfico podemos visualizar nas trilhas onde data-robots estão mais preparados para atuar. Algumas dessas tarefas já podem ser realizadas com pouquíssima - ou nenhuma - participação humana.

# o novo profissional da ciência de dados

O trabalho do cientista de dados vai ser diferente daqui para frente.

Uma profissão tão nova já enfrenta os mesmos desafios de outras áreas com relação a velocidade de automação de tarefas, podendo ter uma parte do seu trabalho realizado por máquinas muito rápidas e que não cometem erros.

O trabalho de “processamento de dados” é sem dúvida o mais próximo de ser automatizado, já que o humano não é tão eficiente quanto uma máquina na execução dessa tarefa - em razão dos grandes volumes de dados e por se tratar de uma tarefa repetitiva baseada em uma série de regras pré-definidas.

Assim como em qualquer outra profissão, apenas conhecer profundamente algumas skills e operar bem as ferramentas disponíveis não será o suficiente, o profissional precisará encontrar formas de se diferenciar de uma execução mecânica. Ainda existe muito espaço para trabalhar com ciência de dados, trata-se de um mercado em crescimento e que todos os dias apresenta novos desafios, mas é uma ilusão encarar este como um mercado estável.





# análise de impacto na sociedade e organizações

Os analistas do Capra Institute investigaram o potencial do impacto desse tema considerando 3 fatores:

**Novidade:** o quanto o assunto é novo

**Atividade:** o quanto o assunto é popular

**Impacto:** o quanto vai impactar em 2020

Com essa análise desenvolvemos um **Score** que representa o impacto desses 3 fatores, considerando uma variação entre 0 e 10.

Esse score é um bom indicador de senso de urgência de ação nesse tema.

## <data-robots>

SCORE

# 7

NOVIDADE

ATIVIDADE

IMPACTO

CAPRA INSTITUTE FOR DATA SCIENCE

<data-robots>

### CASO RECENTE

Forbes | Billionaires | Innovation | Leadership | Money | Business | Small Business | Lifestyle

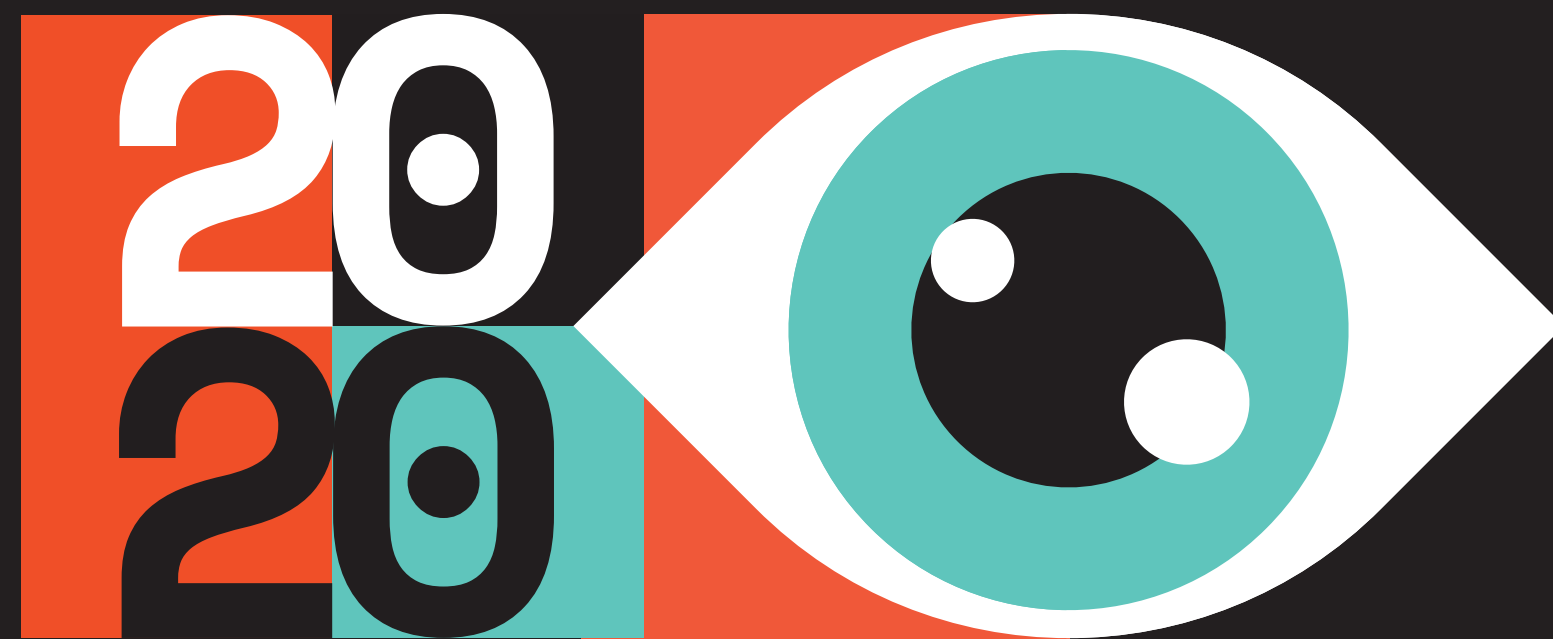
10,838 views | Jul 18, 2019, 12:56pm

## How Automation Is Changing Data Science

Ryohei Fujimaki Contributor @ AI & Big Data  
Ryohei Fujimaki Ph.D. is founder and CEO of dotData

**Data Science: High Value, Even Higher Failures**

Data science can provide a high return on investment across multiple industries and use cases. Whether predicting new target customers, measuring product demand or detecting high product failures - the use cases are nearly as infinite as the problems that face modern businesses. Although data science undoubtedly has significant potential to impact business decision-making, leaders across multiple industries have struggled with getting value from data science projects. In fact, according to research by the Gartner Group, nearly 85 percent of big data projects fail. Even more telling, a 2019 survey by Dimensional Research found that 96% of companies struggle with AI and Machine Learning. While



<data-robots>

### POR QUE VOCÊ DEVE OBSERVAR ESSA TENDÊNCIA EM 2020?

O processo de automação das atividades de ciência de dados está em franca expansão, as gigantes da tecnologia como Microsoft, Google, Oracle, Salesforce, SAP,... já estão disponibilizando muitos desses recursos em seus softwares, sendo que muitos deles podem ser acessados gratuitamente ou dentro dos contratos já estabelecidos.

Um dos maiores concorrentes pela vaga de ciência de dados não é um outro profissional com skills similares, e sim um robô com eficiência de execução inquestionável, e velocidade impossível de ser reproduzida por um humano.

Em 2020 surgirão muitos outros softwares e robôs para executar tarefas da área de ciência de dados, concentre-se nas atividades que exigem pensamento crítico e raciocínio lógico, as máquinas estão mais longe de atingir plenitude nessas habilidades.

### RECOMENDAÇÕES:

- Se você trabalha na área de dados, visite o site [skills.cappra.institute](https://skills.cappra.institute) e crie sua própria trilha de conhecimento.
- Identifique aquelas skills que você é especialista, e observe os riscos de automação dessas tarefas.
- Mantenha-se constantemente atualizado sobre as tendências do mercado, geralmente as novas atividades costumam levar mais tempo para serem automatizadas.
- Estude as ferramentas prontas, muitas vezes você está colocando sua energia em algo que poderia ser realizado por uma máquina.
- Não considere a automação um concorrente de mercado, essa tarefa repetitiva abrirá espaço para você atuar em algo que uma máquina não consegue executar.

### NA PRÁTICA:

Em 30 dias: Faça seu próprio mapa de conhecimento em [skills.cappra.institute](https://skills.cappra.institute)

Em 6 meses: Visite seu mapa criado, e identifique aquelas tarefas que estão sendo automatizadas no dia-a-dia do seu trabalho.

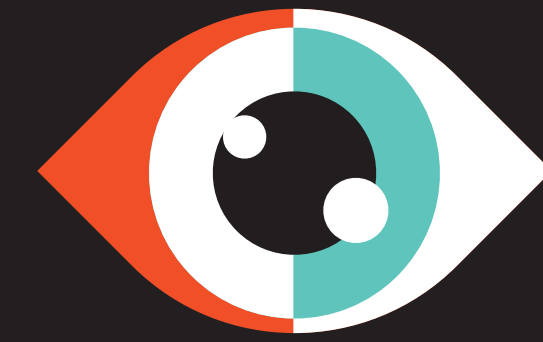
Em 1 ano: Identifique as oportunidades de atuação, os gaps que os softwares não são eficientes para resolver. Atividades que envolvem pensamento crítico e raciocínio lógico são mais difíceis de serem automatizadas. Atualize seu plano de desenvolvimento individual a partir desse mapeamento.



#datathinking2020 @cappra (cappra•institute)

faça agora

20  
20



*Obrigado por ler nosso estudo, esperamos que esse seja o ponto de partida para uma transformação na forma como você lida com dados.*

*A tecnologia da informação está mudando tudo ao nosso redor, por isso as competências como pensamento crítico, raciocínio lógico e habilidades analíticas serão indispensáveis no futuro próximo.*

*Não espere esse ano passar para agir em cima desses temas, escolha aqueles que mais afetam sua vida e/ou profissão, e comece agora mesmo através das listas de recomendações que criamos especialmente para esse Data Thinking.*

*O Cappra Institute desenvolve métodos específicos para aceleração da Cultura Analítica, para acessar outros estudos ou conhecer nossos programas de atualização para você e sua empresa, acesse [www.cappralab.com](http://www.cappralab.com)*

Ricardo Cappra @cappra

Chief Scientist, Cappra Institute for Data Science @cappraLAB

# links & referências

- ▶ <https://tab.uol.com.br/neurohacking/#cerebro-hackeado>
- ▶ <https://g1.globo.com/economia/tecnologia/noticia/2019/07/24/spoofing-como-foi-a-invasao-do-celular-de-sergio-moro-segundo-a-decisao-judicial-que-mandou-prender-4-suspeitos.ghtml>
- ▶ <https://epocanegocios.globo.com/Vida/noticia/2018/05/consumo-excessivo-de-midias-sociais-esta-ligado-doencas-mentais.html>
- ▶ <https://trends.google.com.br>
- ▶ <https://cio.com.br/assistentes-possuam-virtuais-ganham-o-mundo-corporativo/>
- ▶ <https://exame.abril.com.br/pme/as-100-startups-brasileiras-mais-desejadas-pelas-empresas/>
- ▶ <https://www.thinkwithgoogle.com/intl/pt-br/marketing-resources/metricas/sua-empresa-esta-pronta-para-o-futuro-data-driven/>
- ▶ <https://economicgraph.linkedin.com/research/linkedin-2018-emerging-jobs-report>
- ▶ <https://www.bls.gov/emp/graphics/2019/unemployment-rates-and-earnings.htm>
- ▶ <https://www.bls.gov/emp/tables/fastest-growing-occupations.htm>
- ▶ <https://www.bls.gov/opub/btn/volume-7/big-data-adds-up.htm>
- ▶ <https://www.forbes.com/sites/ryoheifujimaki/2019/07/18/how-automation-is-changing-data-science/#542e7a5222fc>
- ▶ <https://www.uipath.com/lp/company/rpa-analyst-reports/forrester-wave-rpa>
- ▶ Algoritmos: você está no controle? | Ricardo Cappra | TEDxBrasilia: <https://www.youtube.com/watch?v=wBz-xWPo1Fc>
- ▶ <https://study.com/academy/lesson/organic-organizational-structure-definition-examples-quiz.html>
- ▶ [https://en.m.wikipedia.org/wiki/Organic\\_organisation](https://en.m.wikipedia.org/wiki/Organic_organisation)
- ▶ Burns, T. & Stalker, G. M. (1961), The Management of Innovation, Tavistock, London. <http://www.leeds.ac.uk/educol/documents/00003223.doc>
- ▶ <https://www.cappra.institute/2019>
- ▶ <http://nirvacana.com/thoughts/2013/07/08/becoming-a-data-scientist/>
- ▶ [www.skills.cappra.institute](http://www.skills.cappra.institute)
- ▶ <https://cappra.com.br/2019/06/05/decodificando-a-profissao-cientista-de-dados/>
- ▶ <https://towardsdatascience.com/the-4-hottest-trends-in-data-science-for-2020-3956cd9fc182>
- ▶ <https://mitsloan.mit.edu/ideas-made-to-matter/are-we-encouraging-wrong-ai>
- ▶ <https://www.equities.com/why-data-science-will-be-the-most-promising-career-in-2020>
- ▶ <https://economicgraph.linkedin.com/blog/how-artificial-intelligence-is-already-impacting-todays-jobs>
- ▶ <https://www.aeseducation.com/careercenter21/what-are-21st-century-skills>
- ▶ <https://ingomierswa.com/2017/05/09/what-artificial-intelligence-and-machine-learning-can-do-and-what-not>
- ▶ <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-ceos-new-technology-agenda>
- ▶ <https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact?cid=soc-app>
- ▶ <https://101.datascience.community/2019/09/26/5-great-data-strategy-resources/>
- ▶ [https://www.sas.com/content/dam/SAS/en\\_us/doc/whitepaper1/5-essential-components-of-data-strategy-108109.pdf](https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/5-essential-components-of-data-strategy-108109.pdf)
- ▶ <https://cisr.mit.edu/reports/create-a-data-strategy/intro.php>
- ▶ <https://www.dataversity.net/what-is-data-strategy/#>
- ▶ <https://hbr.org/2019/07/building-the-ai-powered-organization>
- ▶ <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/breaking-away-the-secrets-to-scaling-analytics>
- ▶ <https://bernardmarr.com/default.asp?contentID=1330>
- ▶ <https://www.gartner.com/en/newsroom/press-releases/2019-02-18-gartner-identifies-top-10-data-and-analytics-technolo>
- ▶ <https://www.kdnuggets.com/2018/09/how-many-data-scientists-are-there.html>
- ▶ <https://www.kdnuggets.com/2015/05/data-scientists-automated-2025.html>
- ▶ <https://101.datascience.community/2016/11/28/data-scientists-data-engineers-software-engineers-the-difference-according-to-linkedin/>
- ▶ [www.cappra.com.br](http://www.cappra.com.br)



# agradecimentos

Essa é uma pesquisa liderada por **Ricardo Cappra**, que mapeia os principais temas de estudos do Cappra Institute, e compila em ensaios que analisam o impacto na sociedade e nos negócios. Os pesquisadores e cientistas associados ao Cappra Institute colaboram com a produção dessa publicação, os mais envolvidos nessa edição foram:

Albertine Weber Carneiro

Anelise Schutz

Anna Cristina Rezende Braga

Caetano Slaviero Pires

Eduardo Santos

Gabriel Rossi Figlarz

Guilherme Rodrigues Machado

Henrique Antonio Pozza

Jessica Nunes Langanz

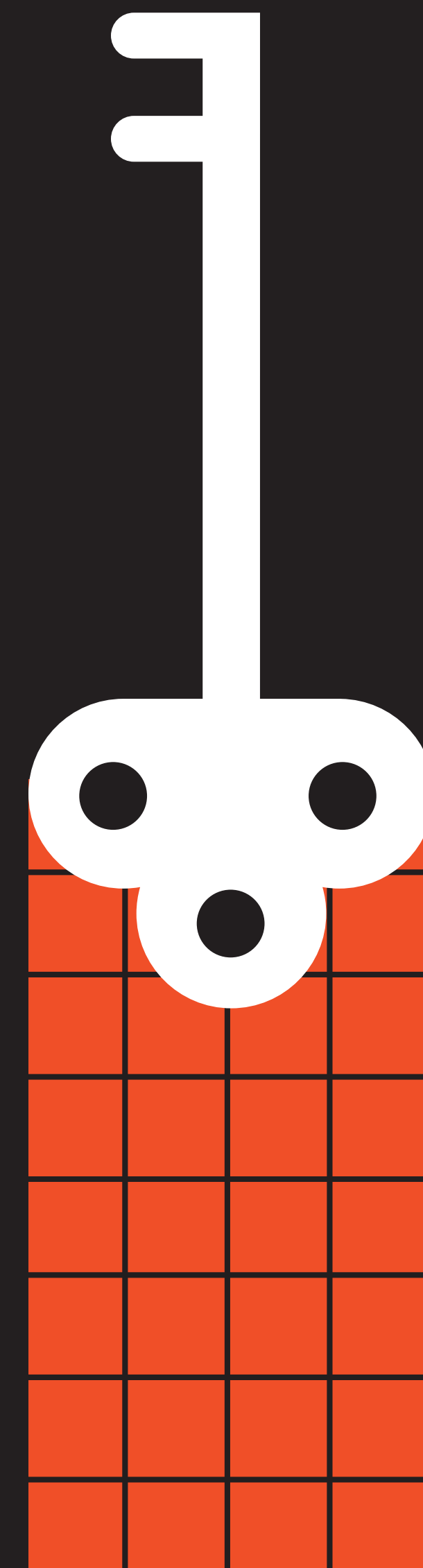
Luciano Firpo

Pedro Venturini

Maurício Diello Guggiana

Tiago Venturieri Barra

Agradecemos à todos que se envolveram - diretamente ou indiretamente - no desenvolvimento e produção desse material.



# data thinking

RICARDO  
CAPPRA  
+

CAPPRA  
INSTITUTE  
FOR DATA  
SCIENCE

20

20

