



# Como trazemos inteligência para o app do Guiabolso

- 1 Arquitetura de machine learning & artificial intelligence
- 2 Integração do AI/ML com o backend
- 3 Integração do backend com aplicativo e experiência do usuário

## CHARACTERISTICS OF A

# DIGITAL BUSINESS



Disrupting existing  
business operations



Blurring the digital  
and physical world



More influence  
of customers



"Supermaneuverable"  
business processes



Agility, speed, scale,  
responsiveness



Combining people, business  
and physical world



0110010101110010■0  
0000011000010111000  
1010011000010111000  
1011011100110011100  
1110010■0111011101  
0100000011100000110  
1000001101001011011  
0110000101110000011

**DATA**  
**IS THE NEW OIL**

# Data Monetization

Capitalizing on Your Data Assets





# Data Driven Decisions



Data  
Processing



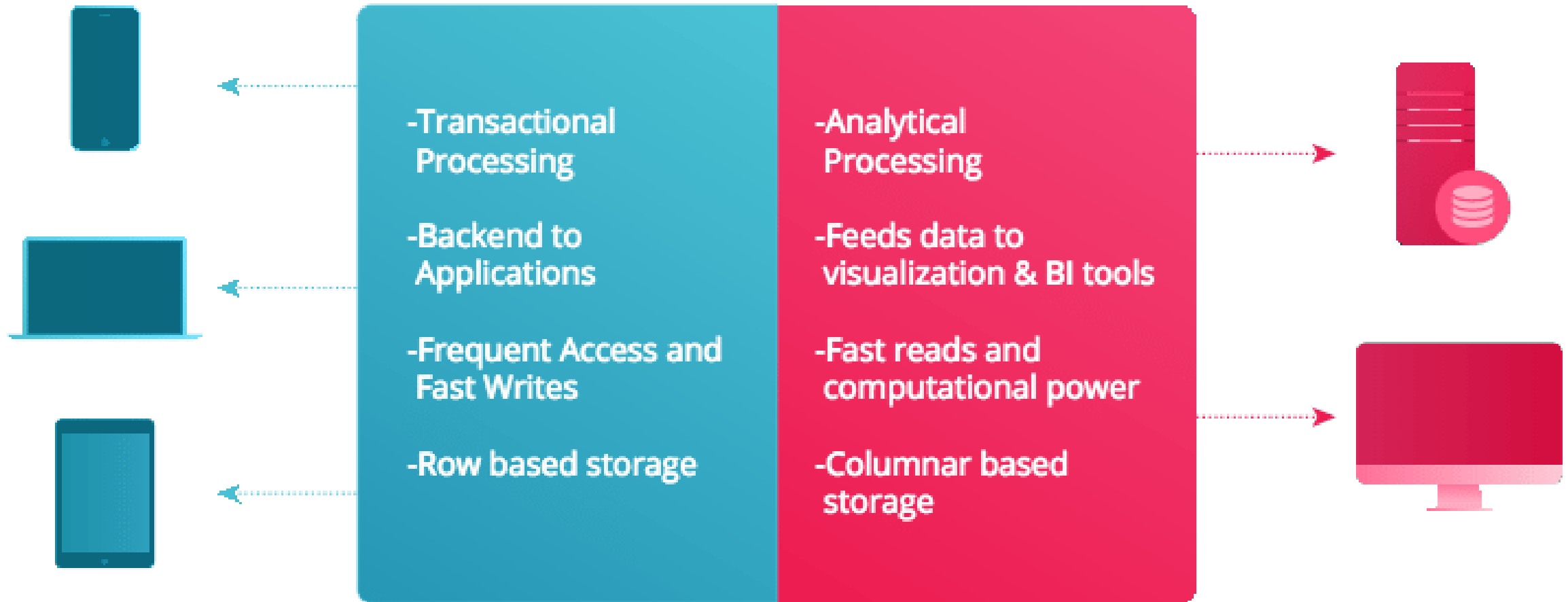
Data  
Analytics



Data Driven  
Decisions



# OLTP versus OLAP







# BIG DATA



**VOLUME**

DATA SIZE



**VELOCITY**

SPEED OF CHANGE



**VARIETY**

DIFFERENT FORMS  
OF DATA SOURCES



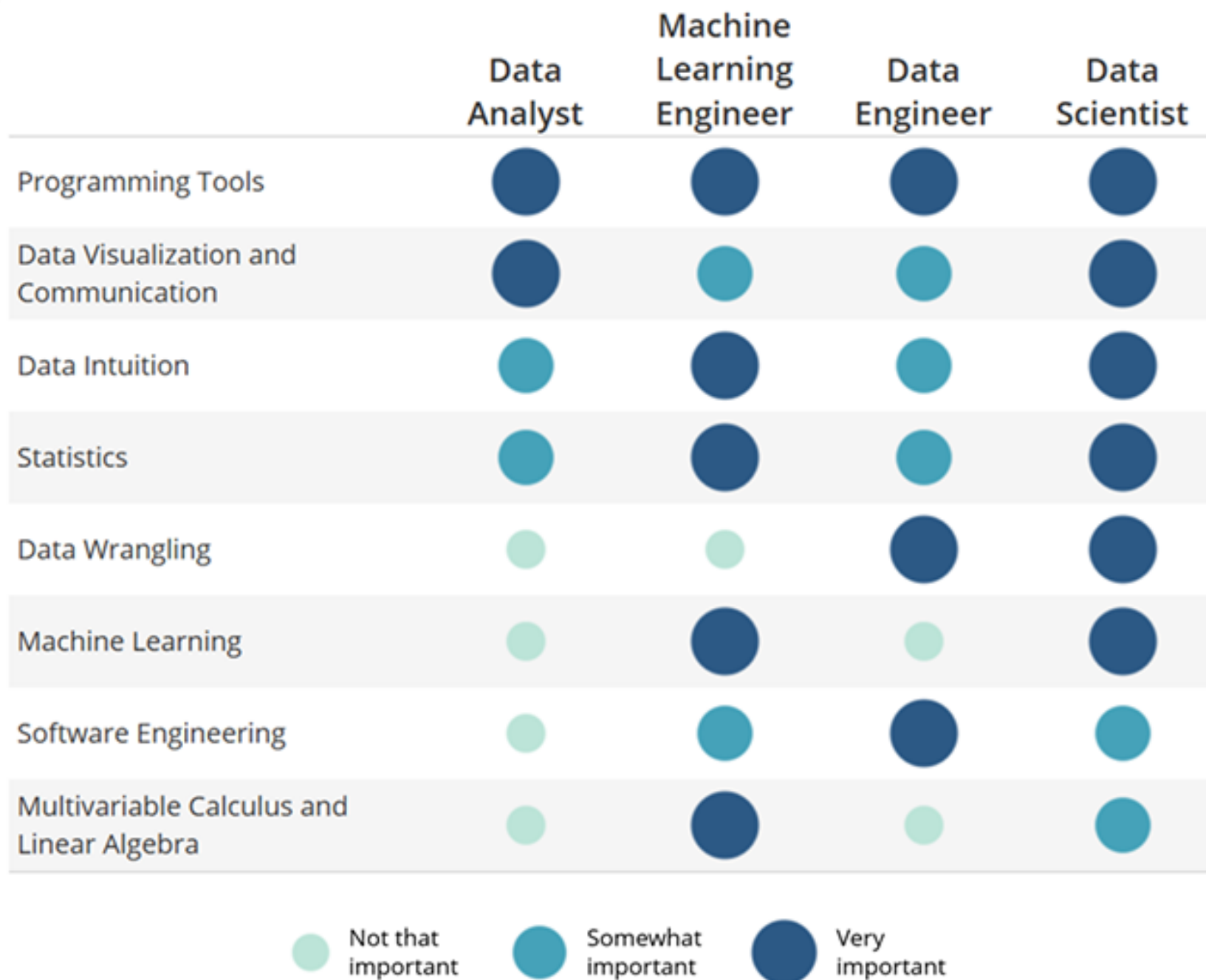
**VERACITY**

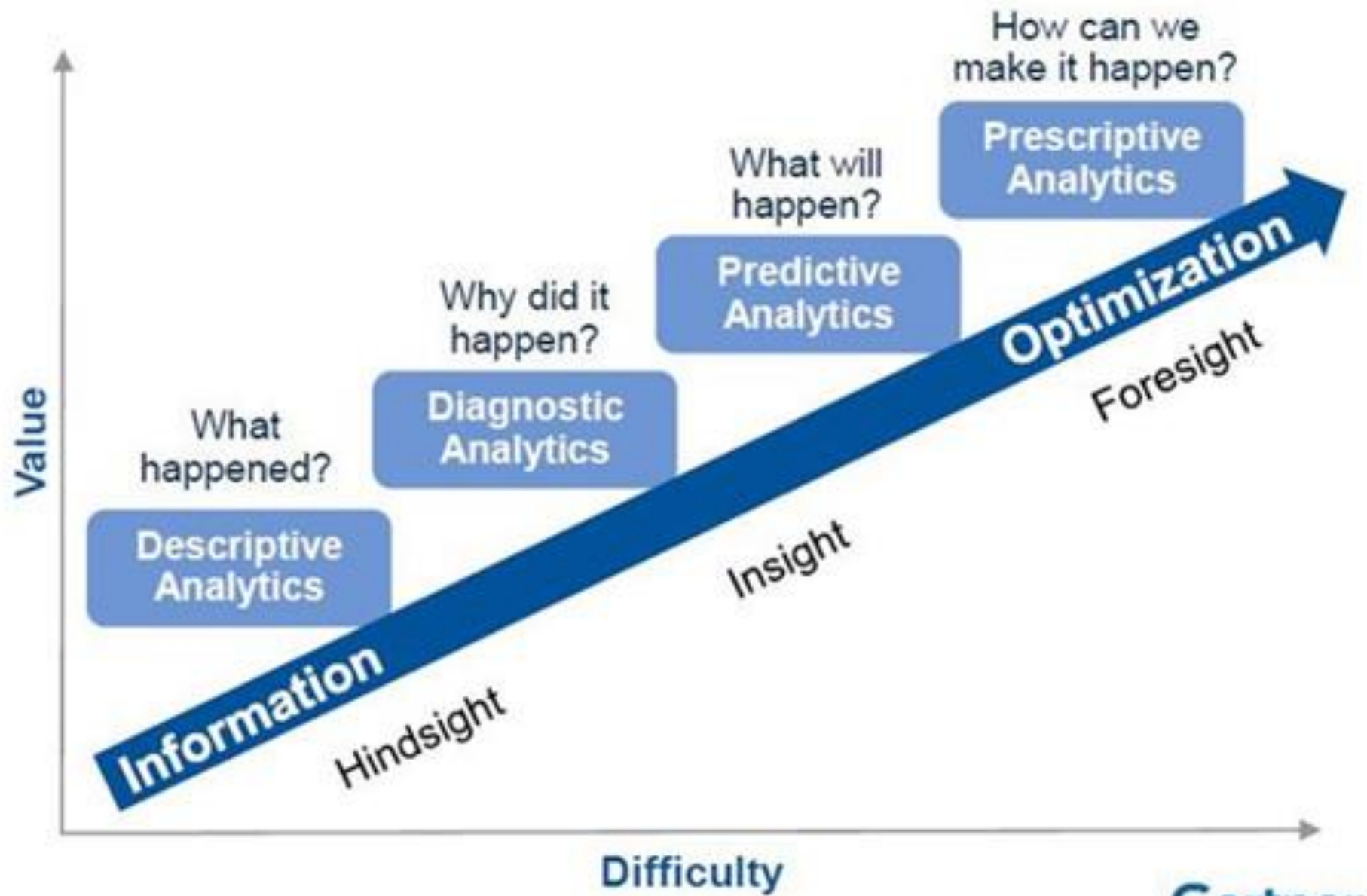
UNCERTAINTY OF  
DATA

# Arquitetura de machine learning & artificial intelligence

The background of the image is a dark blue overlay featuring a hand pointing towards the center. The hand is positioned on the right side, with the index finger extended. The background is filled with various white line-art icons and charts, including bar graphs, line graphs, a pie chart, a network diagram, and a mail icon. The overall theme is data and technology.

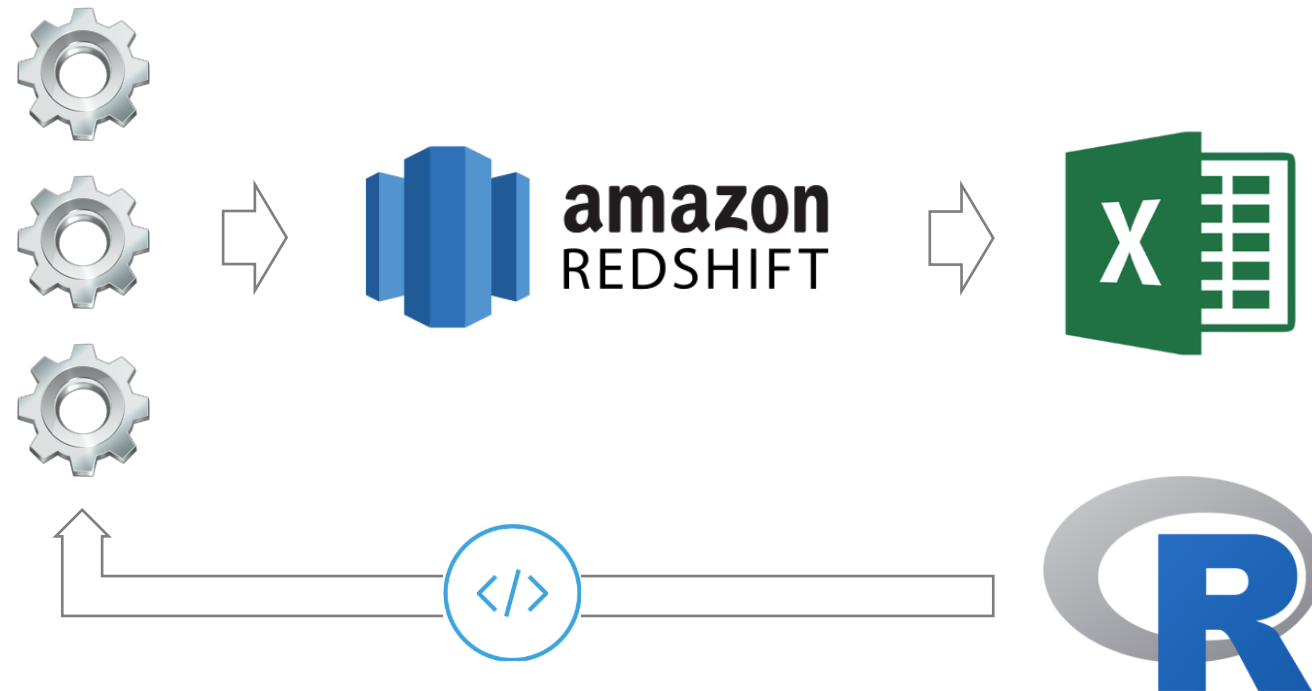
# **True DATA DEMOCRATIZATION Has Arrived**



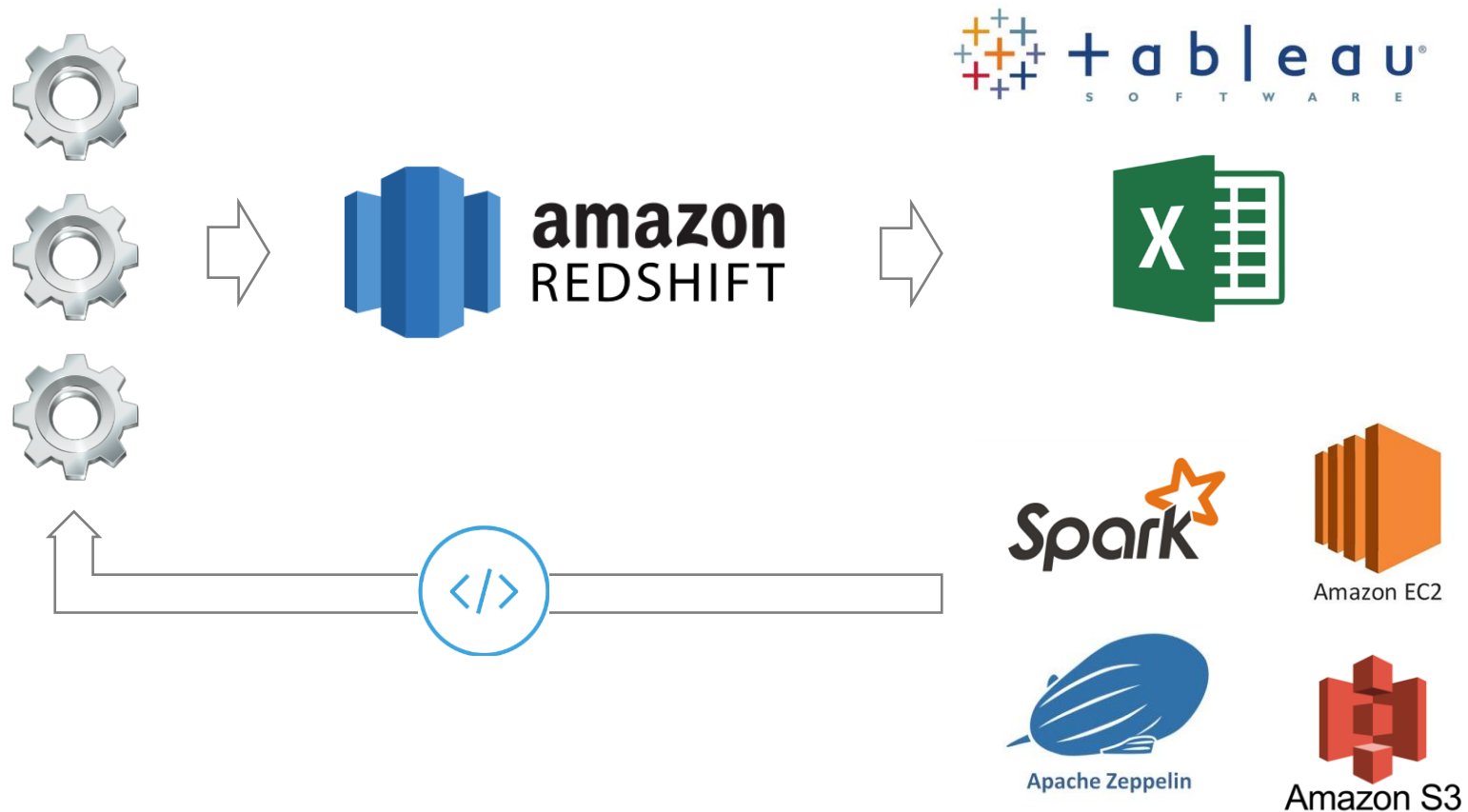




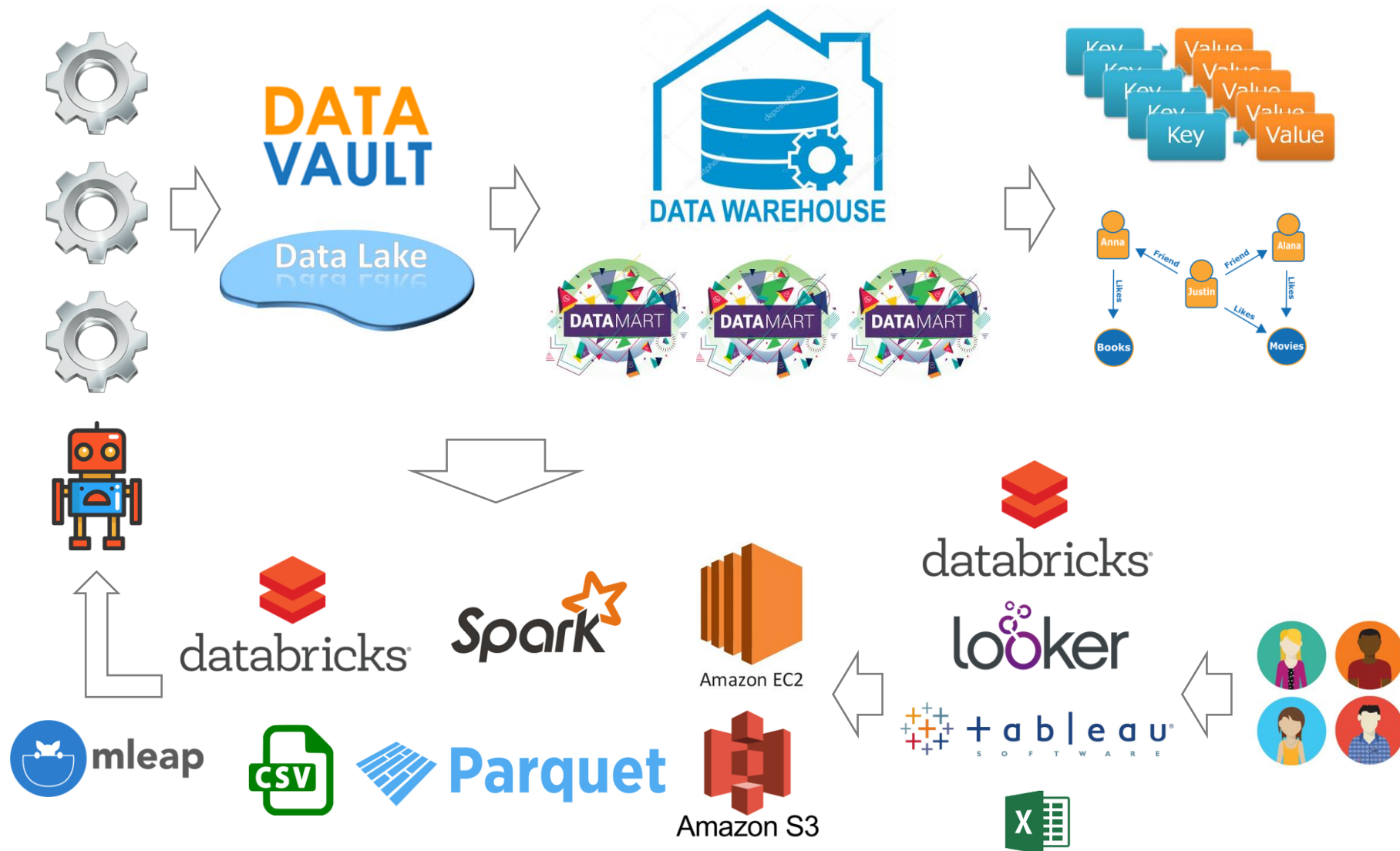
# Arquitetura big data analytics v0



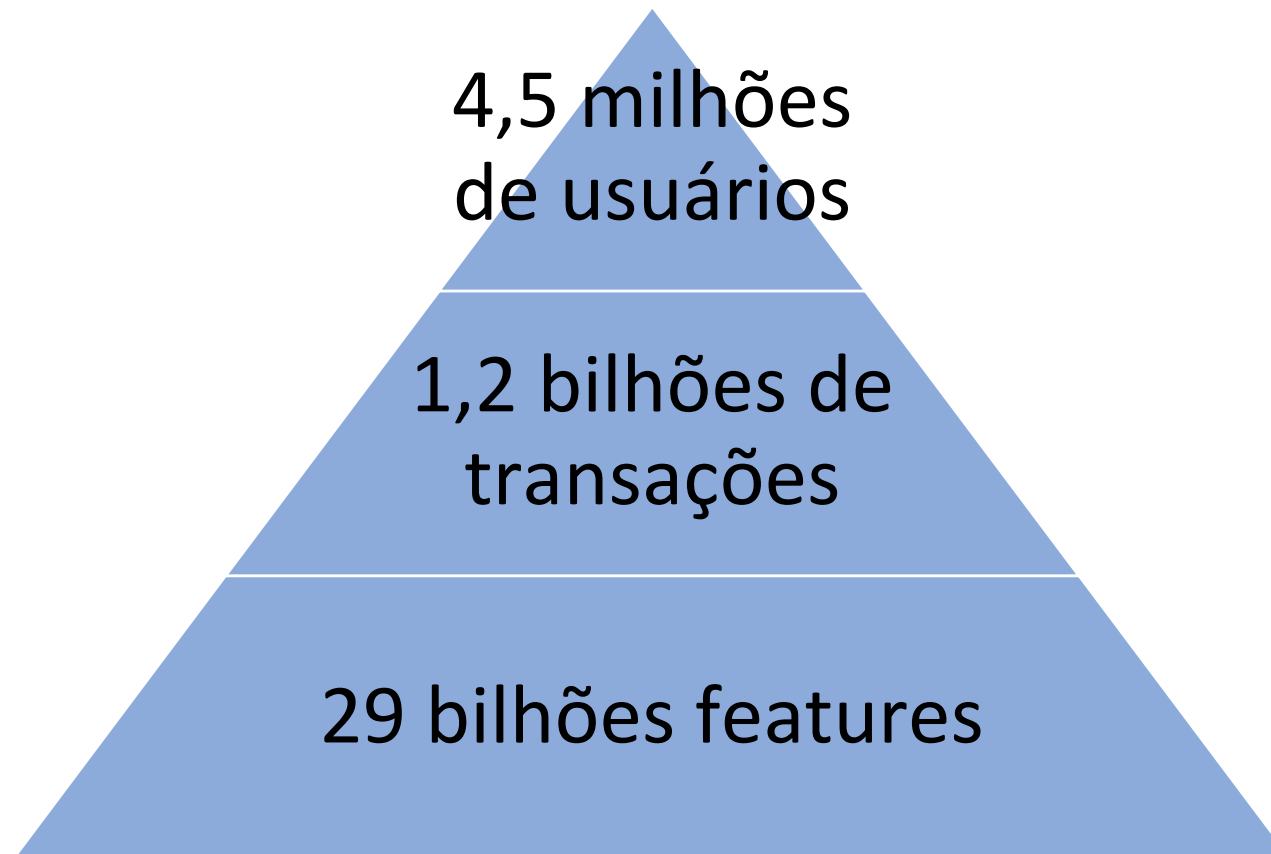
# Arquitetura big data analytics v1



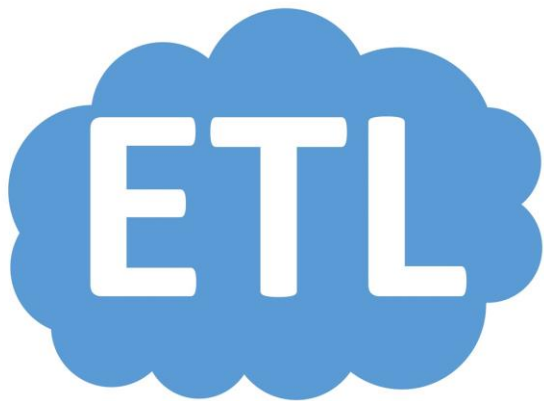
# Arquitetura big data analytics v2



# Volume de dados



# Plataforma de propósito geral

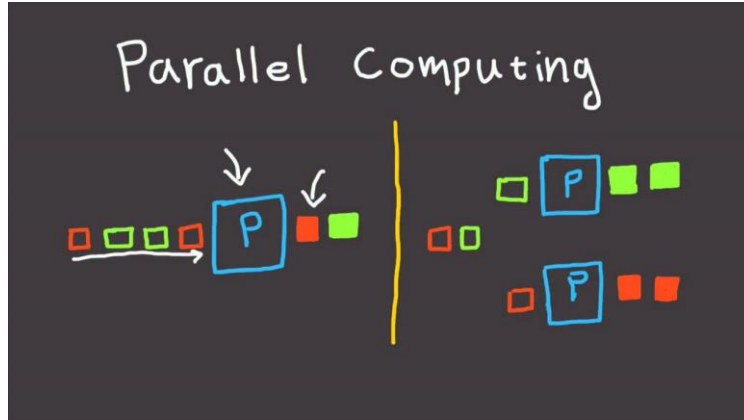




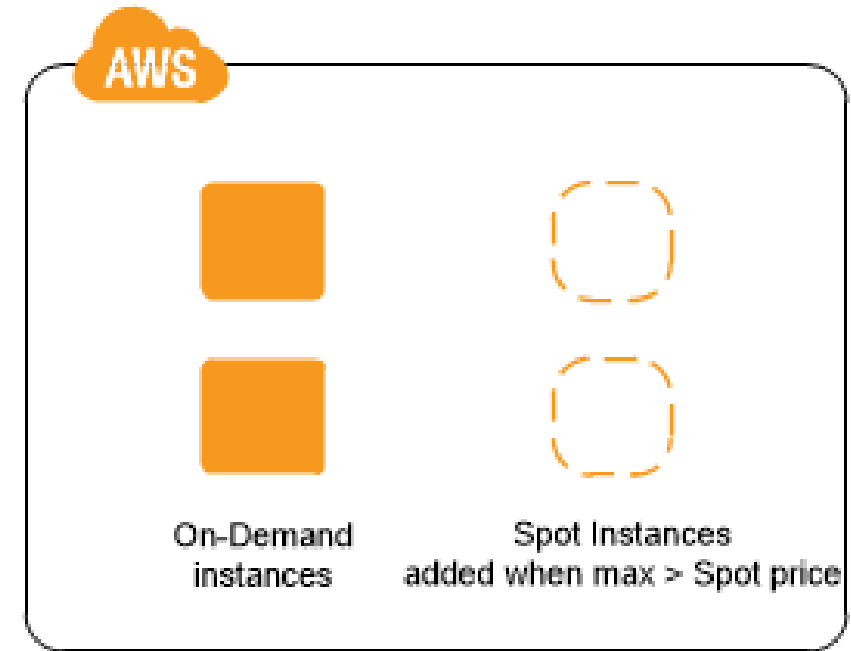
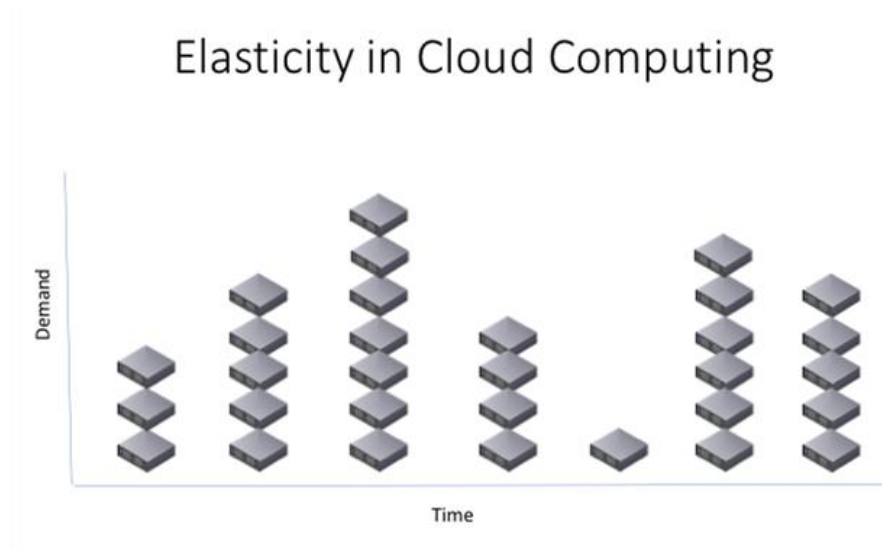
# Múltiplas linguagens atendem vários perfis de usuários



# Computação paralela e distribuída de alta performance

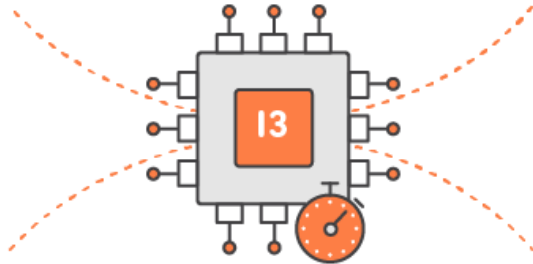


# Grande uso de auto scaling + instâncias spot

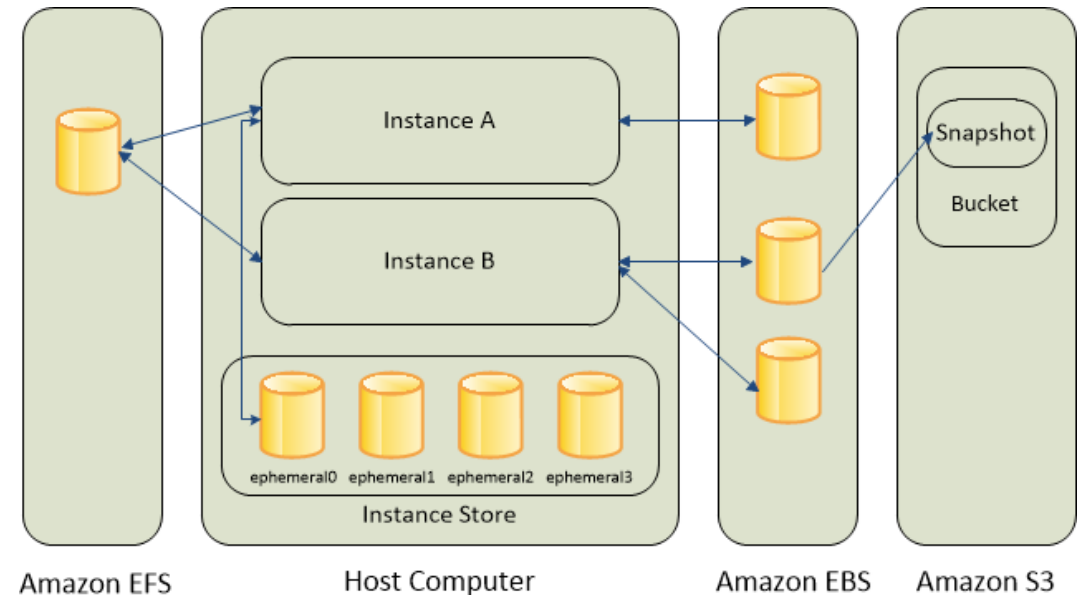


```
spark.databricks.enableAggressiveAutoscaling true
```

# Grande uso de caching e in-memory



**spark.databricks.io.cache.enabled true**



# Shared feature store para ML

- Relational Data

UserID	UserName	Email	Password	FirstName	LastName
1	Chuck.Norris@icb.bg	Chuck.Norris@icb.bg	[enc]	Chuck	Norris
2	Agent.007@icb.bg	Agent.007@icb.bg	[enc]	Agent	007

- EAV Tables

- Entity

EntityId	EntityName
1	Chuck.Norris@icb.bg
2	Agent.007@icb.bg

- Entity Attribute

AttributeId	Name
1	Email
2	Password

- Value

EntityId	AttributeId	Value
1	1	Chuck.Norris@icb.bg
1	2	[enc]



A	B	C
A1	B1	C1
A2	B2	C2
A3	B3	C3

row-oriented storage

A1	B1	C1	A2	B2	C2	A3	B3	C3
----	----	----	----	----	----	----	----	----

column-oriented storage

A1	A2	A3	B1	B2	B3	C1	C2	C3
----	----	----	----	----	----	----	----	----

Partition Pruning

a	b	c
a5	b5	c5
a6	b6	c6
a7	b7	c7
a8	b8	c8
a9	b9	c9

a	b	c
a5	b5	c5
a6	b6	c6
a7	b7	c7
a8	b8	c8
a9	b9	c9

Column selection

a	b	c
a0	b0	c0
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4

Predicate Pushdown

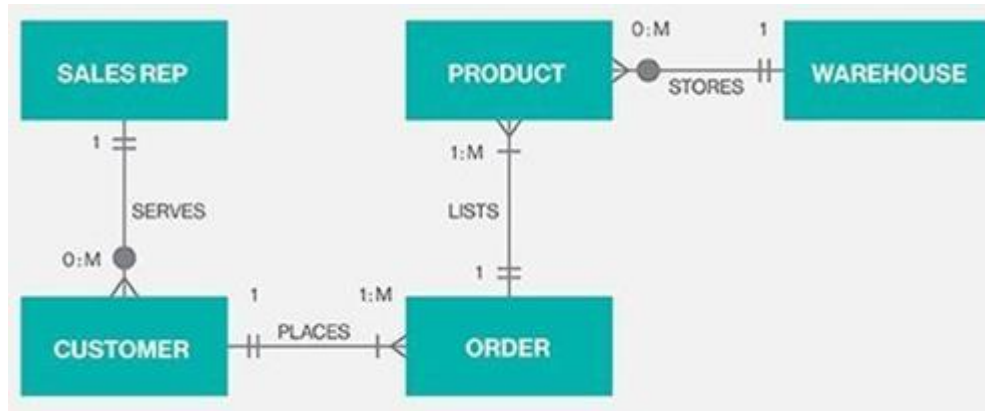
a	b	c
a0	b0	c0
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4

Read only the data you need

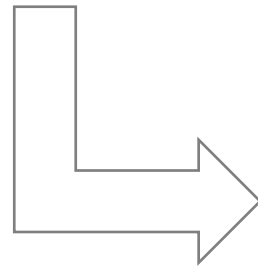
a	b	c
a0	b0	c0
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4



# Automação da engenharia de features

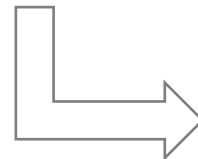


- Orientado a evento
- Conceito de “shared feature store”
- Dezenas de milhares de features



- Frequência
- Recência
- Valores
- Classes
- Estatísticas

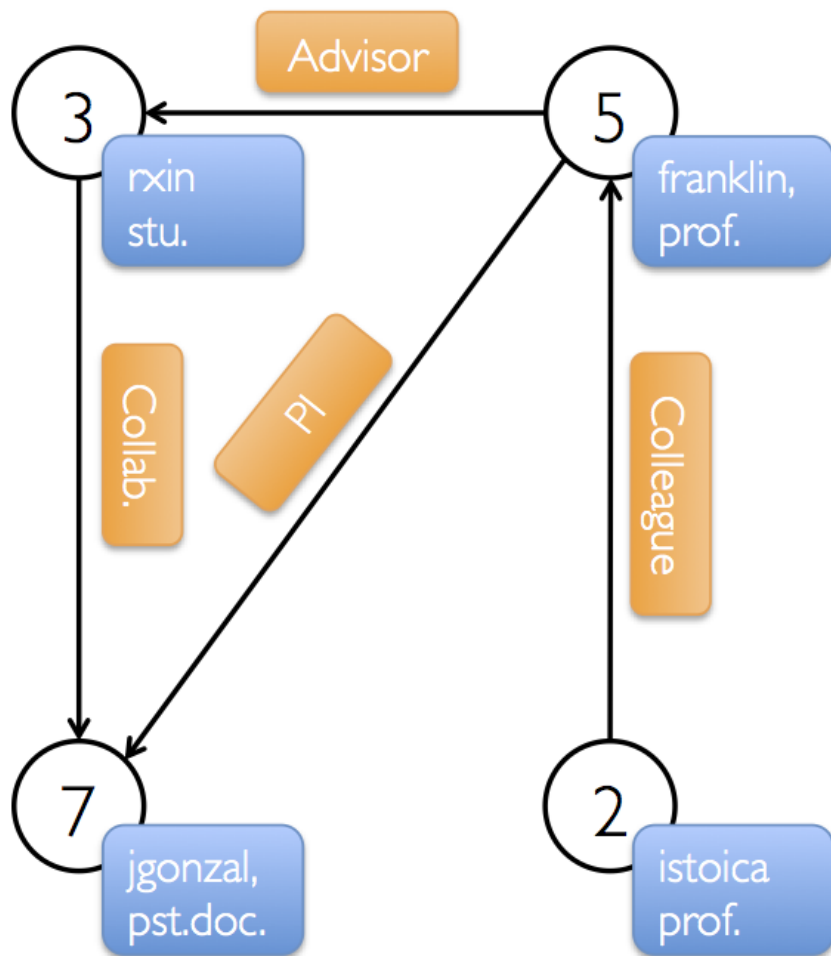
Ex.: valor médio de transações nos últimos 90 dias por categoria e moeda



$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_m x_m$$

# Graph feature store para ML

Property Graph



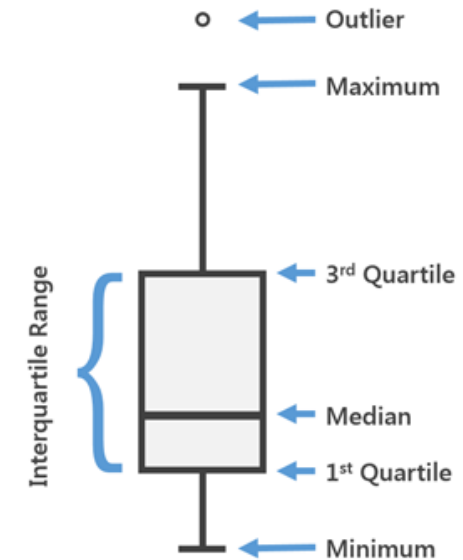
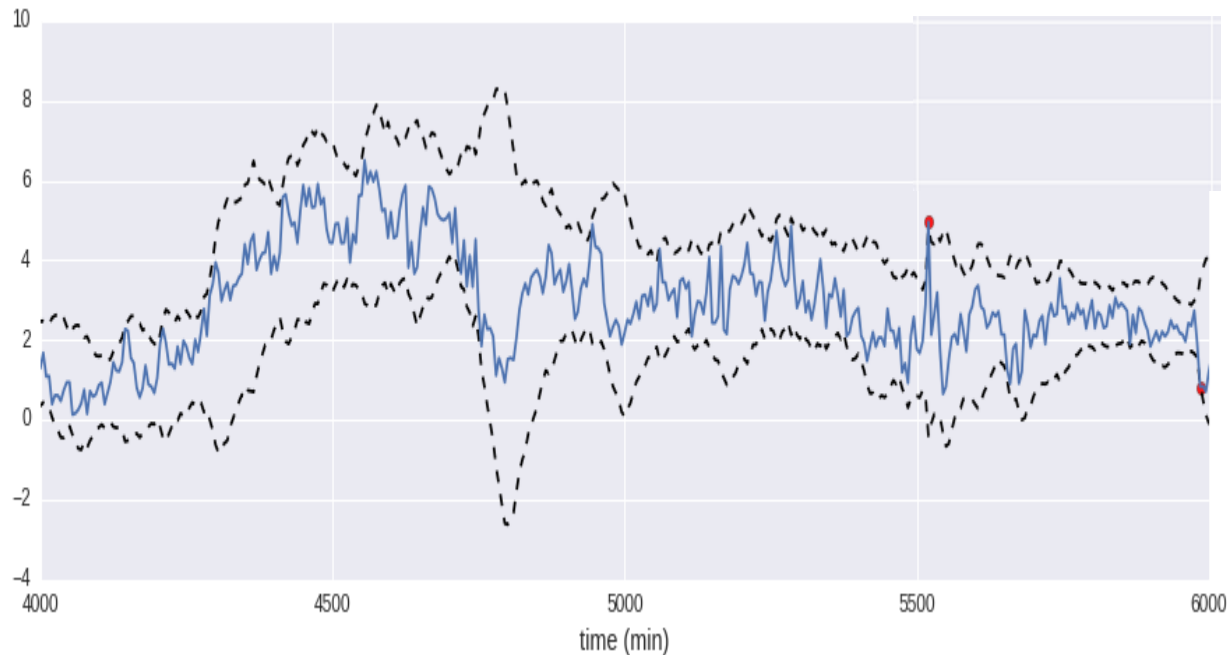
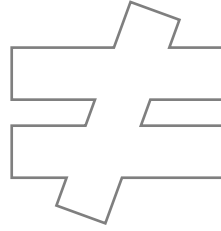
Vertex Table

Id	Property (V)
3	(rxin, student)
7	(jgonzal, postdoc)
5	(franklin, professor)
2	(istoica, professor)

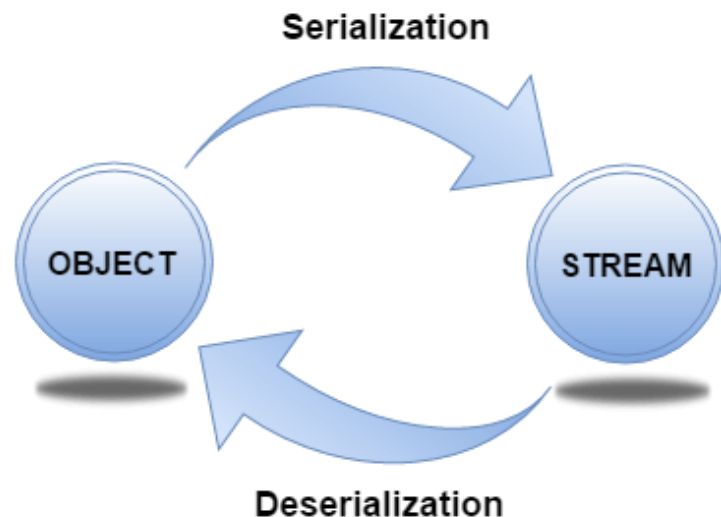
Edge Table

SrcId	DstId	Property (E)
3	7	Collaborator
5	3	Advisor
2	5	Colleague
5	7	PI

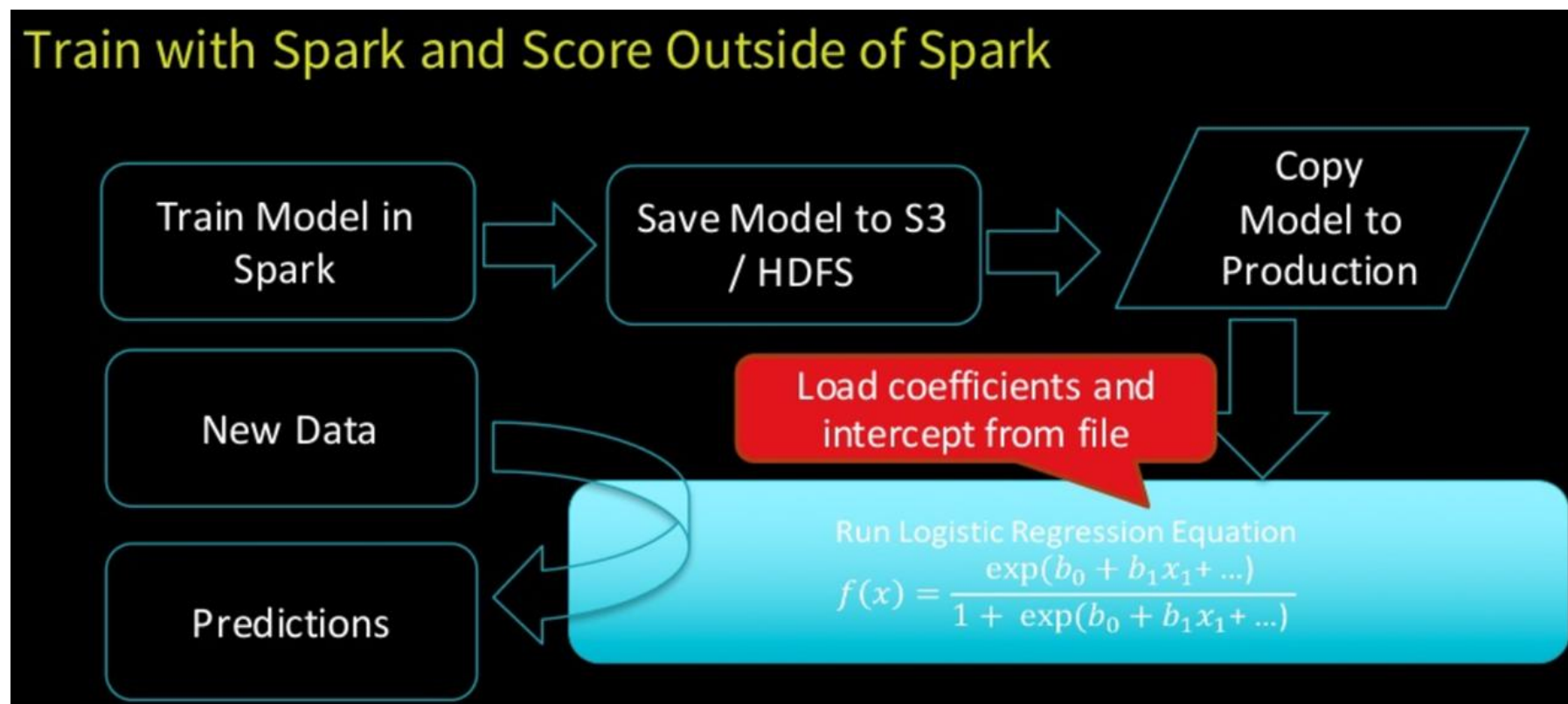
# Data anomaly detection



# Deploy de modelos em produção



# Deploy de modelos em produção (cont.)



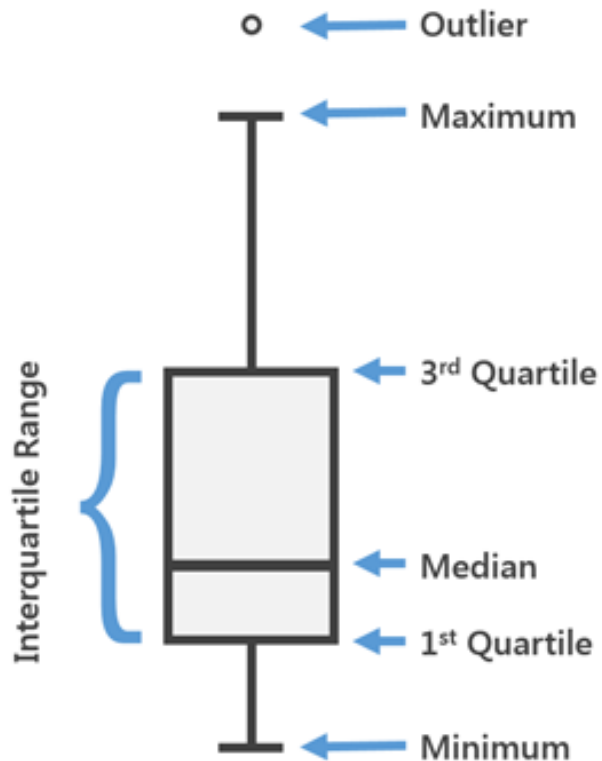
- Fácil de incorporar nos sistemas existentes
- Baixa latência e complexidade
- Baixo overhead

- Decision Tree Classifier
- Decision Tree Regression
- Logistic Regression
- Random Forest Classifier
- Random Forest Regression

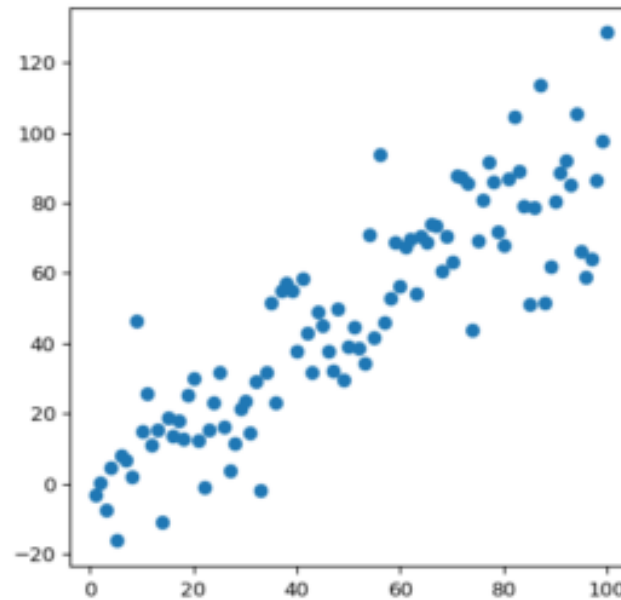


# Automação de análises estatísticas

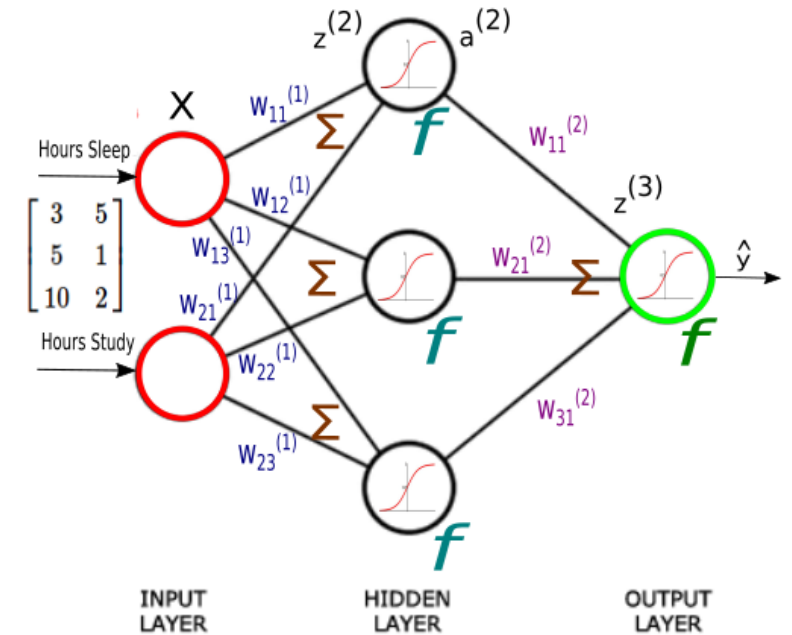
Univariate



Bivariate

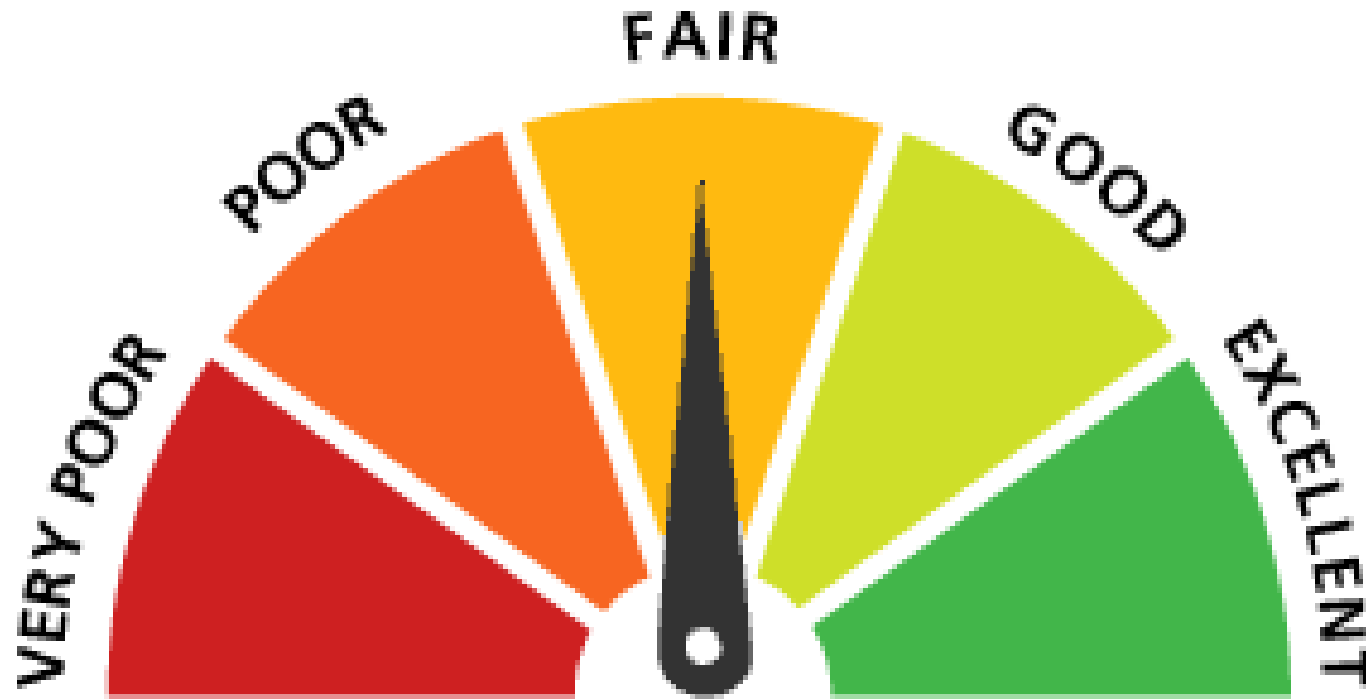


Multivariate



**Use case:  
credit scoring**

# Atribuição de notas de acordo com a probabilidade inadimplência

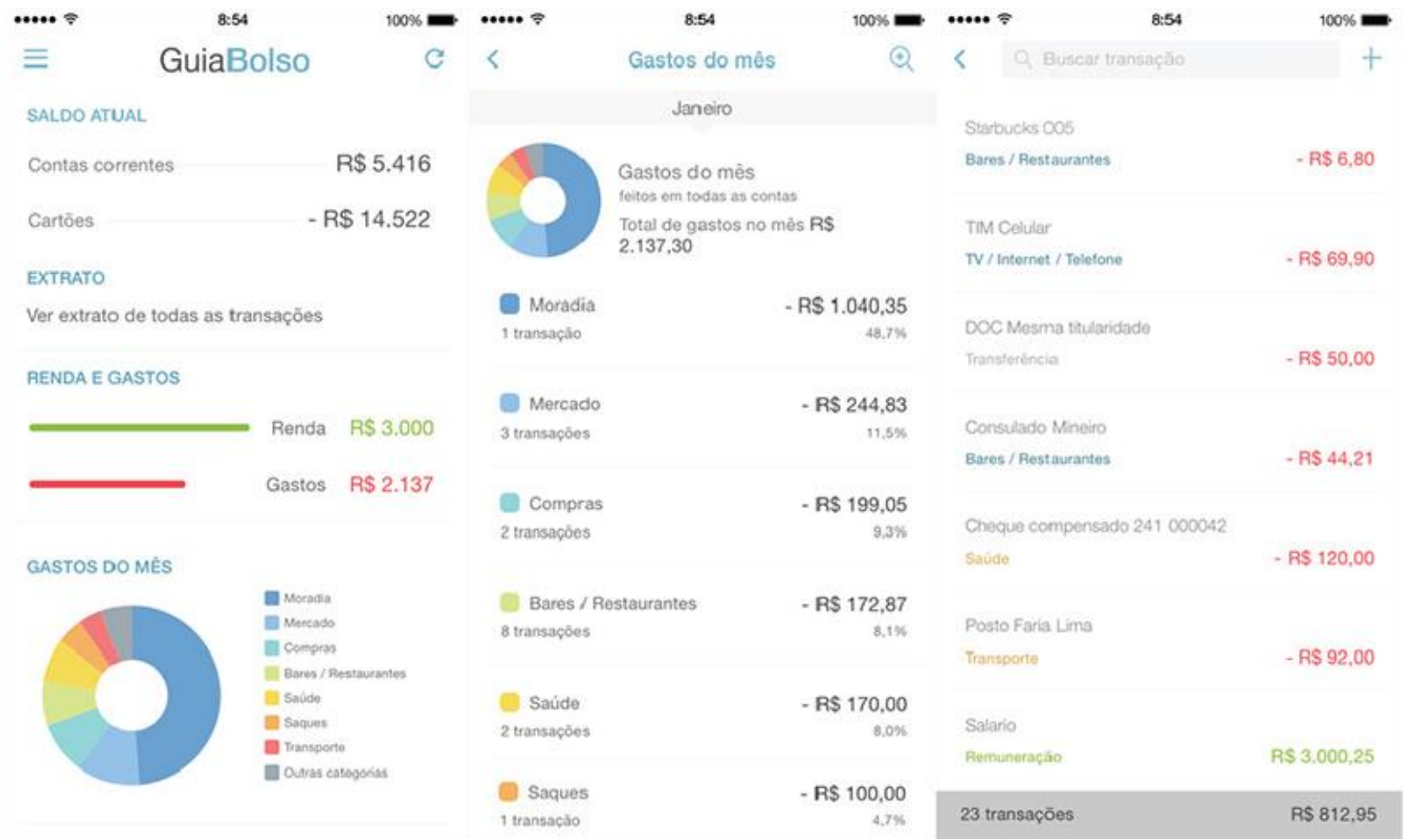


# Fluxo na plataforma



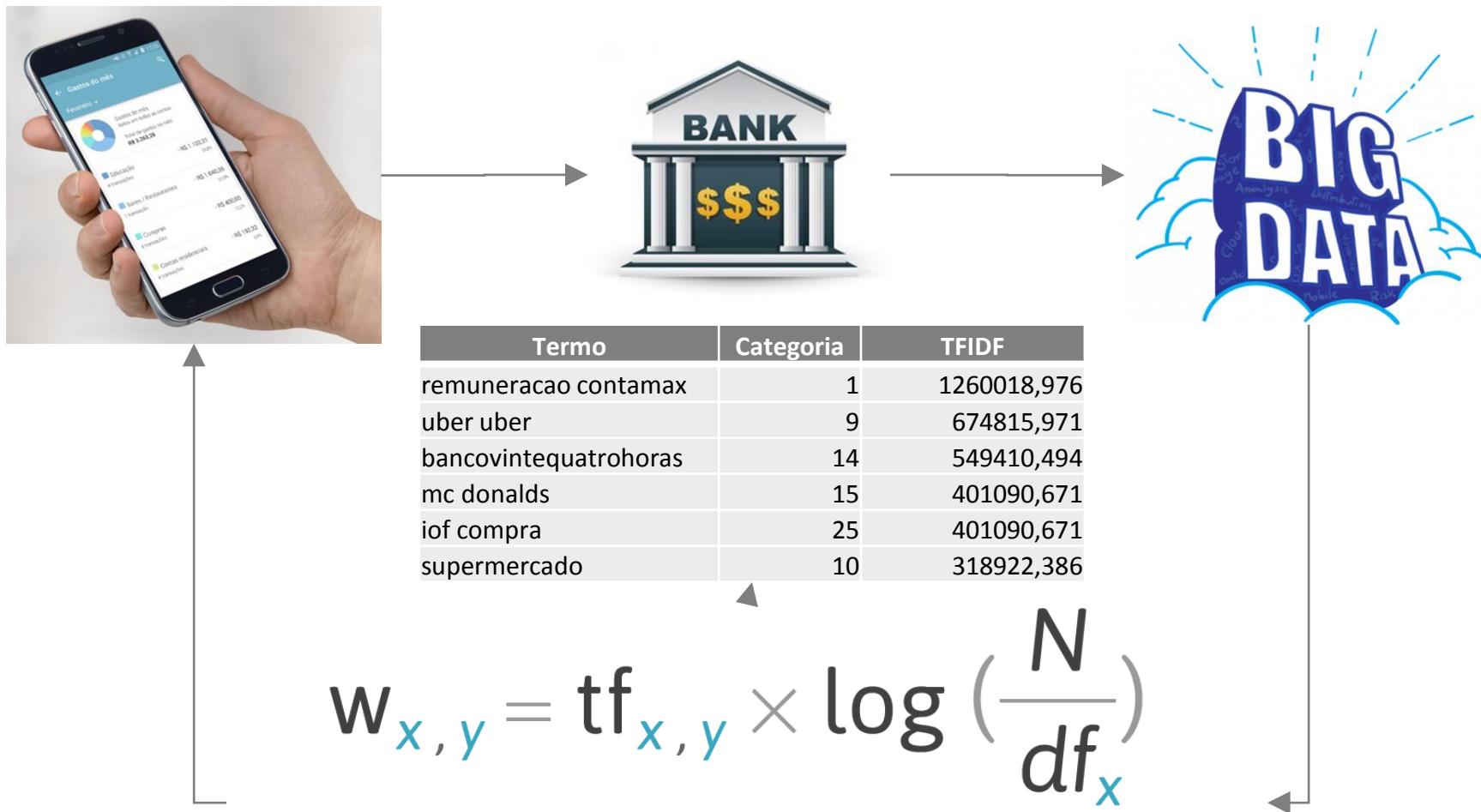
**Use case:  
categorização  
de transações**

# Categorização no aplicativo





# Fluxo na plataforma



**TF-IDF**

Term  $x$  within document  $y$

$tf_{x,y}$  = frequency of  $x$  in  $y$

$df_x$  = number of documents containing  $x$

$N$  = total number of documents

# Google Launchpad

# GuiaBolso é selecionado para programa global de aceleração do Google

Launchpad Studio, do Google, tem duração de quatro meses nos Estados Unidos



5.set.2018 às 18h01

🔊 A- A+

**Paula Soprana**

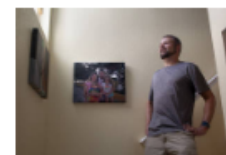
**SÃO PAULO** A startup brasileira GuiaBolso foi uma das selecionadas para o programa de aceleração global do Google, o Launchpad Studio, informou a companhia nesta quarta-feira (5).

Essa nova rodada do programa tem duração de quatro meses nos Estados Unidos e foco em produtos da área de finanças. O objetivo é otimizar serviços que utilizam aprendizado de máquina e blockchain.

PATROCÍNIO

**Embratel**

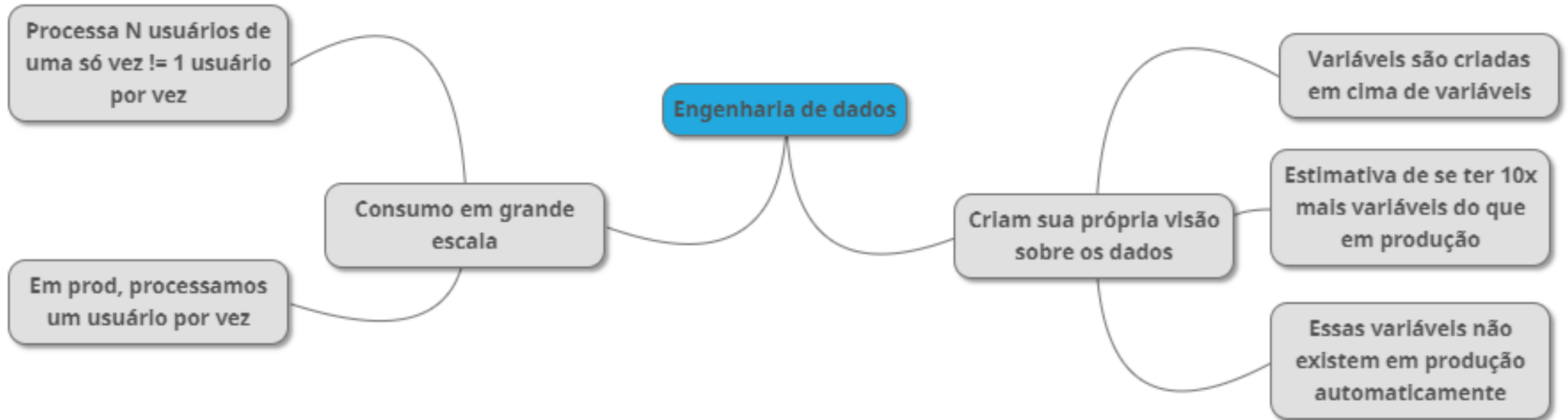
## relacionadas



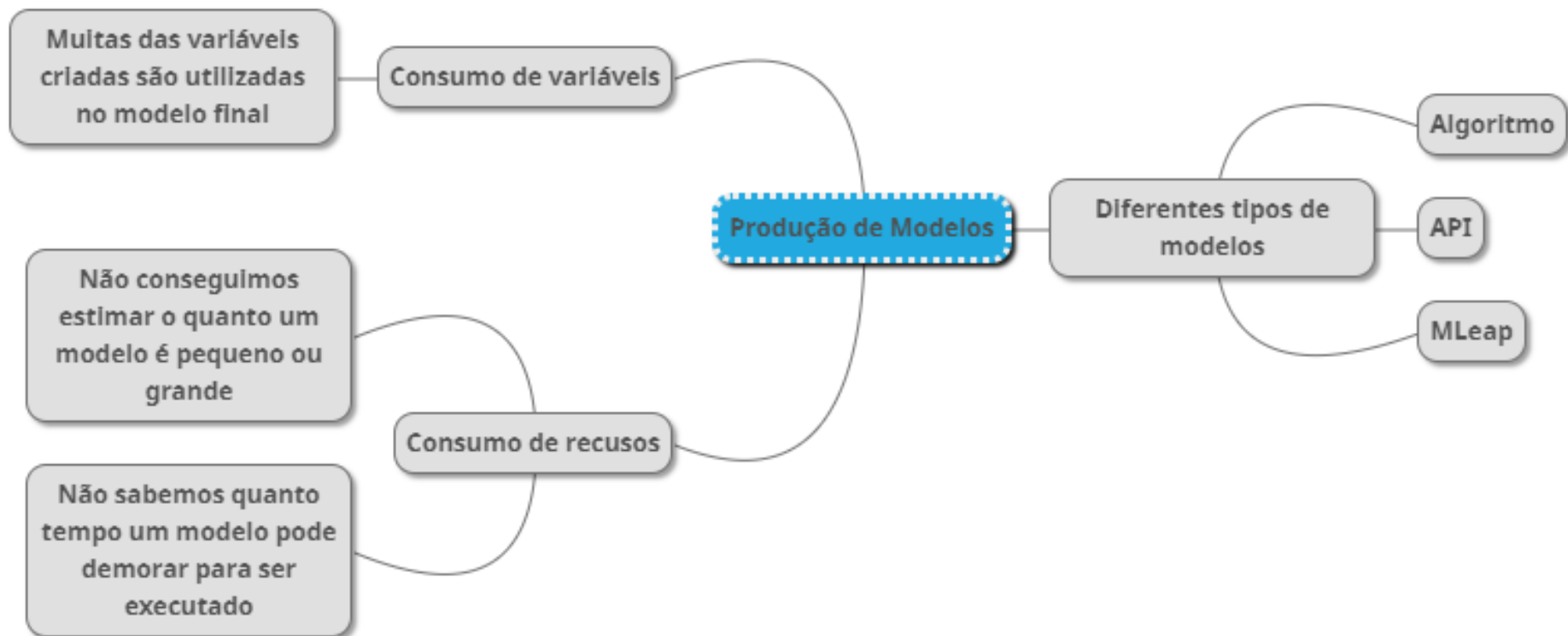
Como se aposentar antes dos 40, com US\$ 1 milhão no banco

# Integração do *AI/ML* com o backend

# Como um backend dev enche a AI/ML no GB



# Como um backend dev enche a AI/ML no GB





**Como  
resolvemos até  
agora?**

# Consumo de variáveis

- 1** Criamos cada variável manualmente no backend.
  - Cada variável é criada em seu serviço responsável pelo domínio
  - Traduzimos de pysql para sql/kotlin
- 2** Único ponto de consumo de variáveis
  - Um sistema é responsável por centralizar as requisições de procura (GraphQL)
  - Isolamento da parte de consumo de modelos com a plataforma
- 3** Criamos variáveis como uma foto dos sistemas naquela hora

# Como gerenciamos recursos

## **1** Rodamos os modelos em paralelo

- Deste modo, o tempo total de geração de modelos seria igual ao modelo mais lento
- Utilizamos um sistema de orquestração de W

## **2** Auto Scaling

- ECS na AWS



# Diferentes tipos de Modelos

**1** Criamos handlers para cada tipo

- API
- Algoritmo manual
- MLeap
- Otimizado no Workflow

**2** Utilizamos o mesmo esquema de input/output para cada modelo

- Variáveis chave/valor



# Integração do backend com aplicativo e experiência do usuário

**Meu app era  
simples... e  
agora?**



# Vá por partes

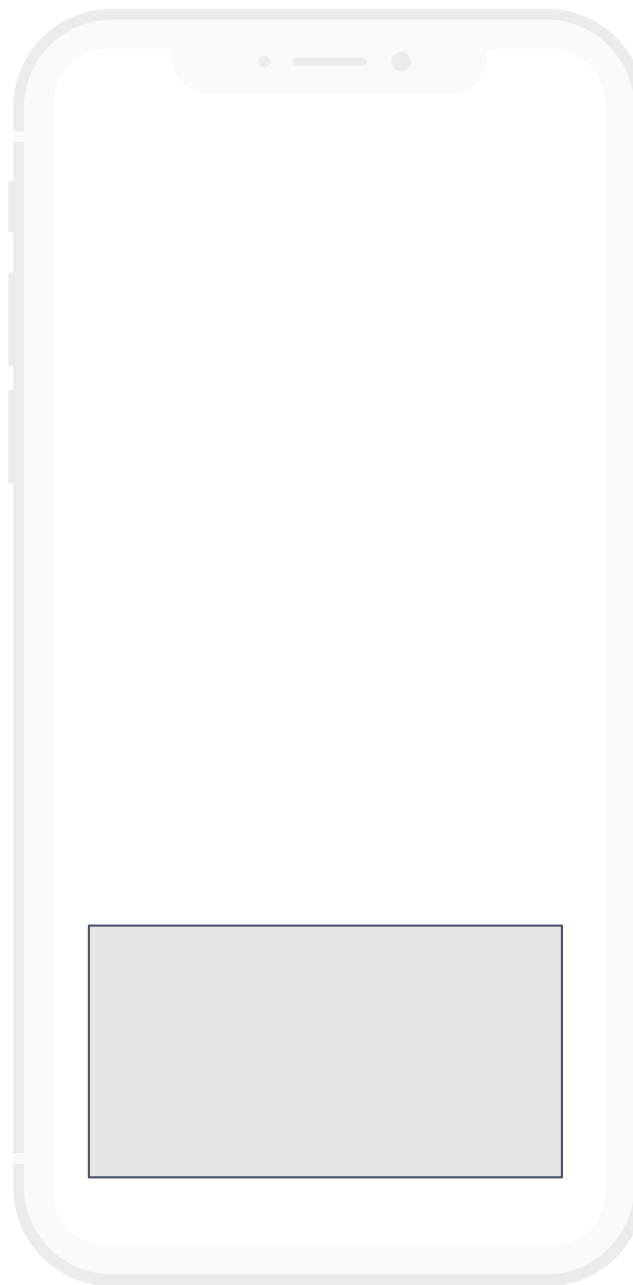
**Se seu app era simples, não coloque previsão do futuro do nada**

## **Siga um caminho lógico**

- 1** Teste em algo menor, um bloco pequeno, um caso específico
- 2** Escale para ver a adesão, faça onboardings
- 3** Vá misturando ou, de alguma maneira, acostumando o usuário

# Primeiro testamos somente na home

Tivemos que construir uma estrutura de  
builders para interpretar dinamicamente o  
conteúdo e entregar uma experiência  
diferente para cada card sem ter que apelar  
para webview

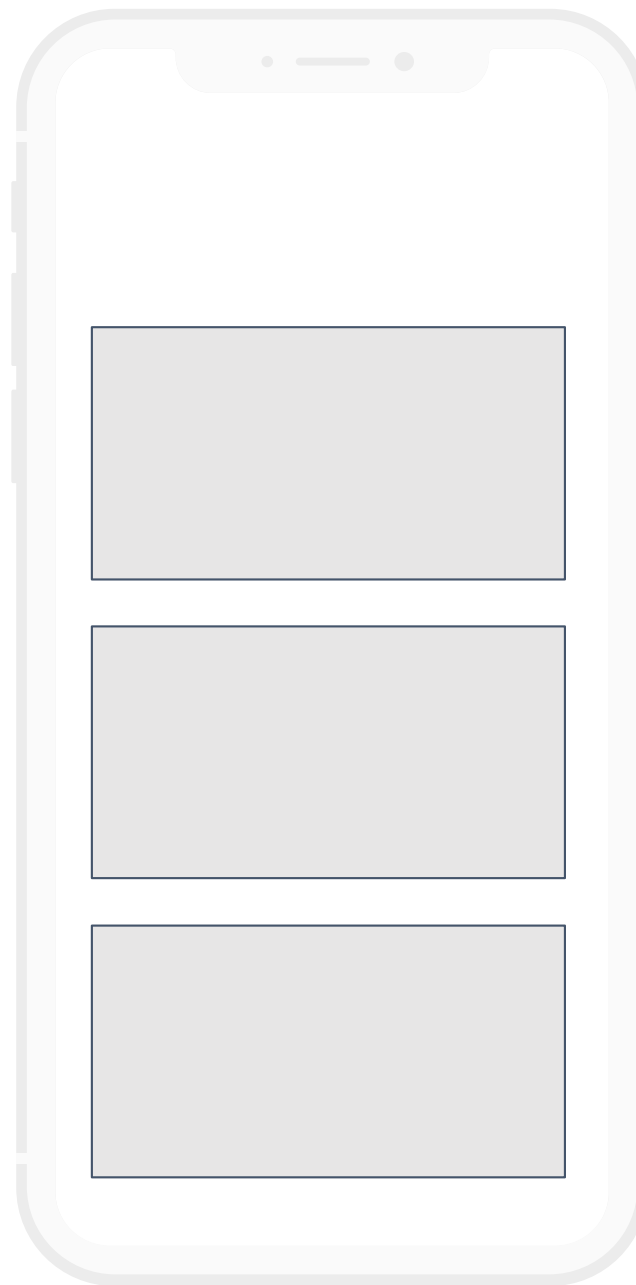


## Começamos no iOS

Não precisávamos  
parar vários  
times, pois temos  
apps nativos

# Depois, fizemos 3 cards

A ideia de criar um conceito de card e foco no reaproveitamento ajudou a escalar

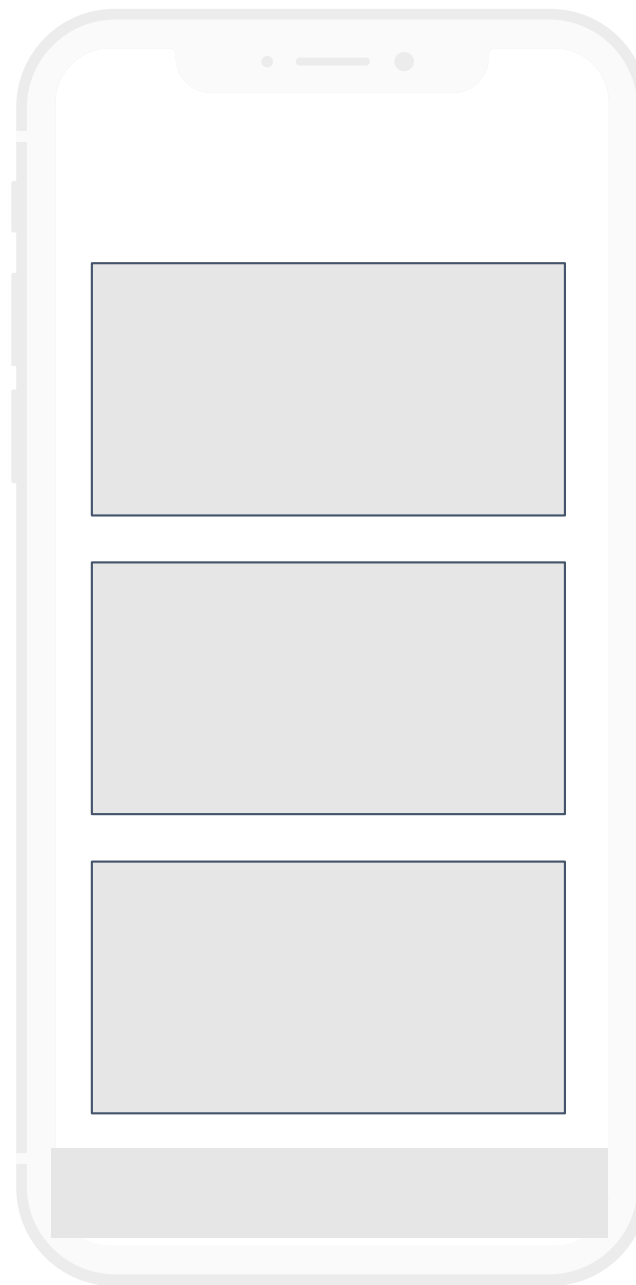


## Conteúdo

Mais difícil do que tecnologia é gerar conteúdo

# Enfim, uma aba GUIA

Só depois que nos sentimos confortáveis  
com o modelo, saímos para experimentação  
com uma tela exclusiva para o guia



## Android

Foi nesse  
momento que  
acionamos um  
time de Android

**Só é smart se  
parece smart**

# IA/ML erra

**Legal, categorizei, mas tá errado... coloquei um card que não é sua cara... te dei uma oferta caída**

## **Dê autonomia pro usuário**

- 1** Tenha sempre uma forma de dar input
- 2** Se puder, torne o erro uma forma de mostrar que liga pra ele
- 3** Quando recomendar algo que talvez não seja a melhor, deixe opções

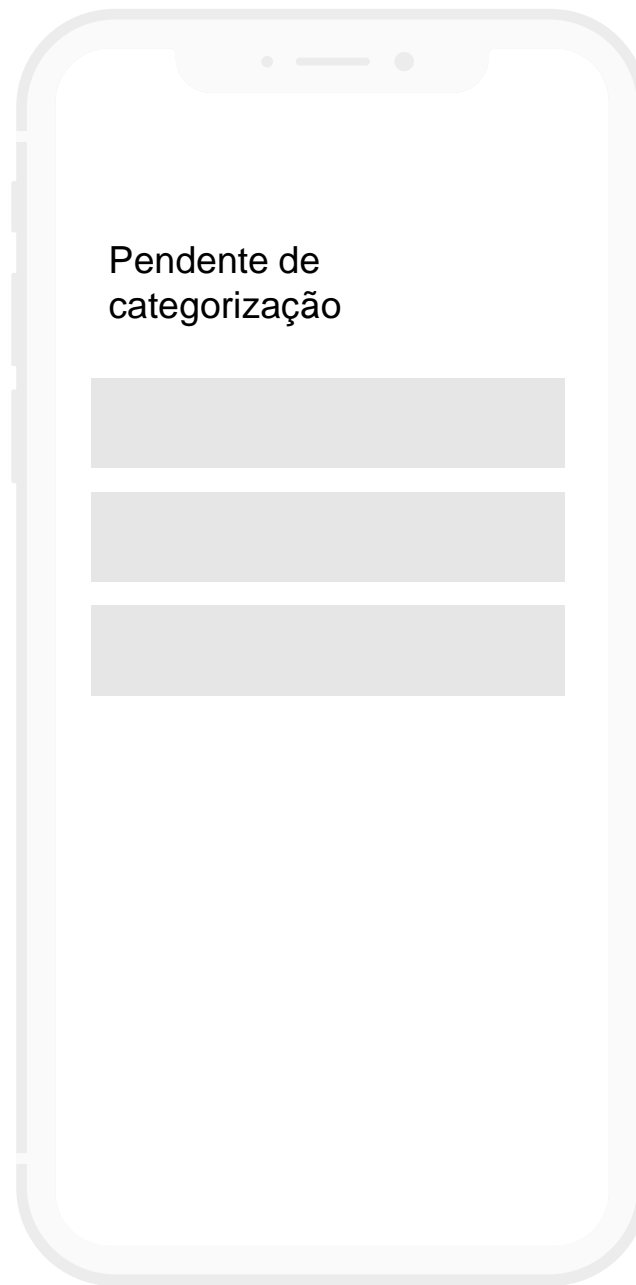
**UAAUU!!**

**MEH!?**



# Deixamos claro que não deu

Quando não conseguimos selecionar uma categoria para o usuário, a gente entrega a experiência de pendência

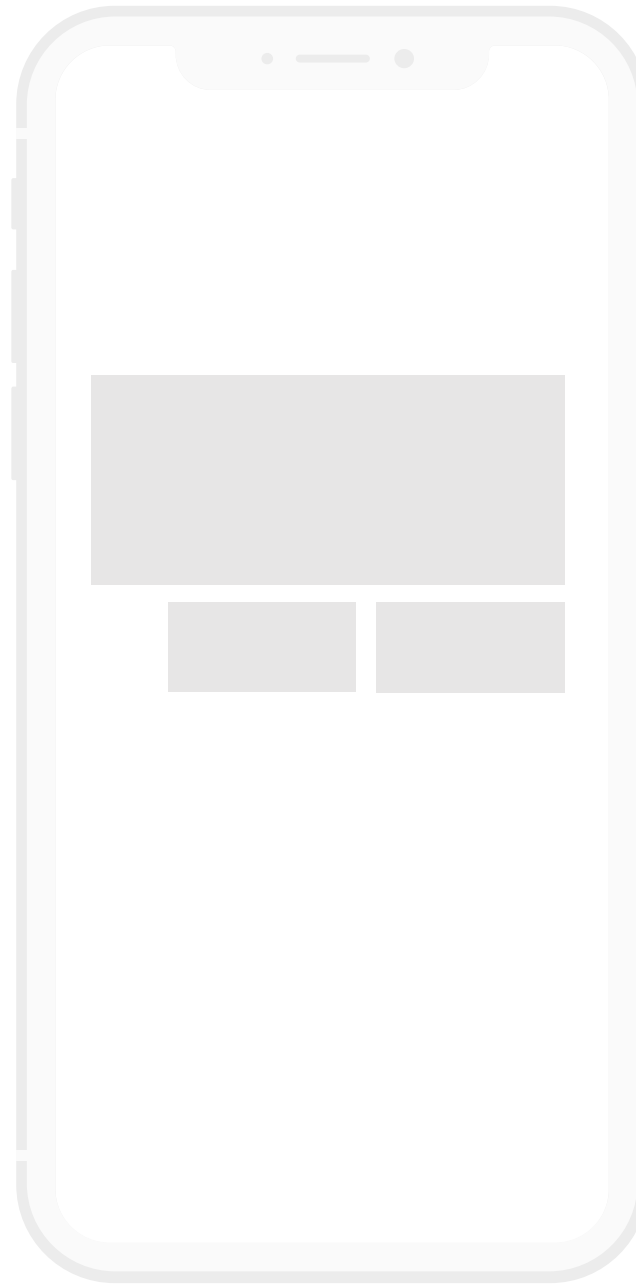


## Animação

Demos um caminho fácil para recategorizar

# Feedback do guia

No guia, você pode dizer se achou aquela informação relevante, assim, conseguimos te entregar algo personalizado não só pela inferência mas pelos seus gostos



## Verbal

Não adianta ter um bruta sistema de inteligência... sua linguagem deve encaixar com o propósito

**Evento, evento,  
evento, evento,  
even..to**

# Invista para coletar eventos

**Lembre-se que para validar algo você precisa de dados e para ter dados... vai ter que capturar**

## **Capture todos os eventos possíveis**

- 1** Invista tempo, esforço, pessoas para garantir consistência e qualidade
- 2** Sai caro... sai muito caro... ler, receber, guardar
- 3** Lembra daquela ferramenta de CRM? Bom, talvez não resolva

# Coleta de eventos in house

## Fizemos nosso próprio engine de eventos

Nós usávamos uma ferramenta e ela não nos atendia, optamos por algo manual e até simples



Não fazemos fila, a gente trabalha on-line com eventos



Usamos nosso próprio protocolo



Graças ao Clean Architecture foi mais fácil substituir o legado



Foi trabalho braçal o esforço para migrar e alinhar todos os eventos

**Real-time ou  
batch**

# Ou gasta em status ou gasta em loader

**Se você decidir usar batch, seu usuário vai ter que ter visibilidade... se usar real-time, vai ter que ter loader**

**As vezes pra impressionar é importante pesar as prioridades**

- 1** Se colocar o processamento no loading, seu app vai parecer lento
- 2** Se colocar em batch, o usuário vai ter que ter um status das info
- 3** Se colocar em real-time... lembre, internet no Brasil não é das melhores

Guiabolso

GUIA FINANCEIRO

QUEM SOMOS ▾

EMPRÉSTIMO

RADAR DE CPF

BLOG

ENTRAR

CADASTRE-SE

A gente faz a conta

que **ninguém** te conta.



ECONOMIZE MAIS

# Perguntas? Sugestões? Críticas?

Thiago Reis

thiago.reis@guiabolso.com.br

[www.linkedin.com/in/thgreis](http://www.linkedin.com/in/thgreis)