

# Jornadas Digitais CCTD

Ciencia de dados

---

## #02





# Programação

- Retrospectiva
- Palestra
- Perguntas
- Encerramento e convite para próximo

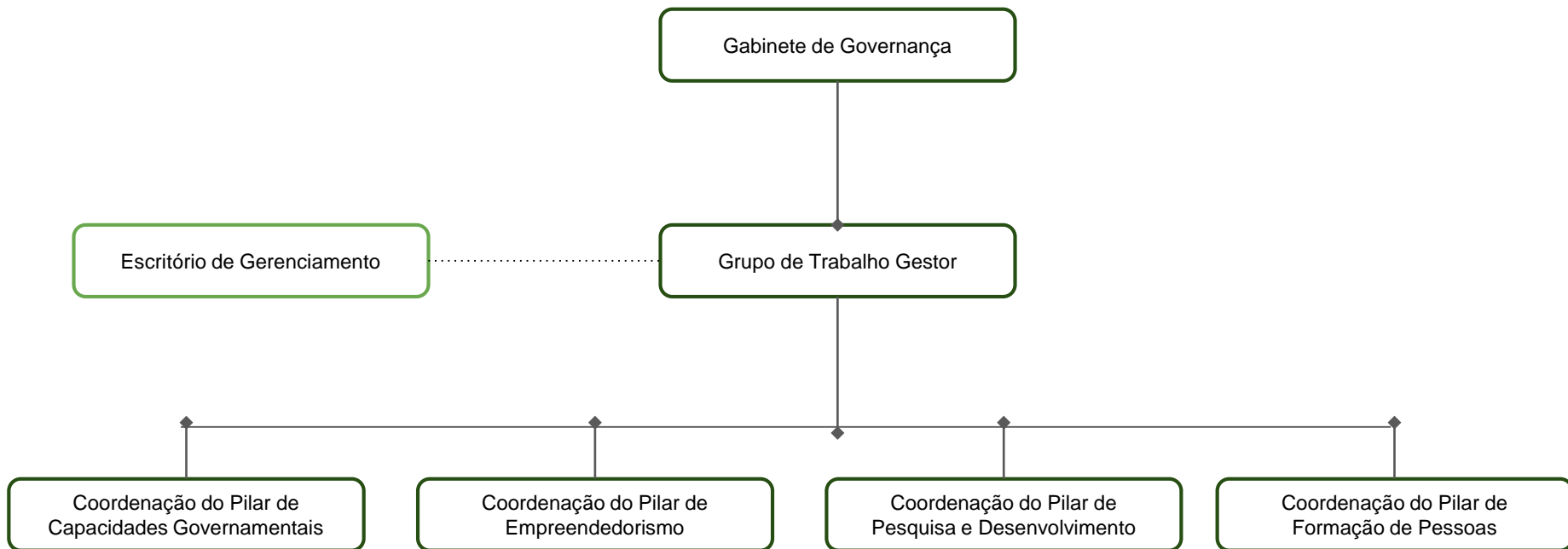


CENTRO DE COMPETÊNCIA PARA  
TRANSFORMAÇÃO DIGITAL DO CEARÁ

## Objetivo

O objetivo do CCTD, enquanto programa de coordenação de políticas públicas, é transformar o Ceará em uma referência na adoção da transformação digital para impacto no cidadão, ensino e empreendedorismo.

## Estrutura de Governança



## Pilares do CCTD

Capacidades  
Governamentais

Secretaria do  
Planejamento e Gestão |  
Casa Civil

Startups e  
Empreendedorismo

Secretaria da Ciência,  
Tecnologia e Educação  
Superior (Secitece)

Pesquisa e  
Desenvolvimento

Fundação Cearense de  
Apoio ao  
Desenvolvimento  
Científico e Tecnológico  
(Funcap)

Formação de Pessoas

Secretaria do  
Desenvolvimento  
Econômico e Trabalho  
(Sedet)

A decorative graphic on the left side of the slide, consisting of a network of white lines and small circles on a dark green background, resembling a circuit board or a neural network.

# Retrospectiva

---



# Agenda Jornadas Digitais

12 de agosto - ~~LGPD (Mariana Zonari e Fernando Gebara)~~

26 de agosto - **Ciencia de dados (Jairo da Silva Junior)**

09 de setembro - DevOps

23 de setembro - Estratégias de Migração para a nuvem

07 de outubro - modernização de backup

21 de outubro – Experiencia do Usuário (UX)

04 de novembro - Nuvem híbrida

18 de novembro - Containers

02 de dezembro – Inteligencia Artificial e aprendizagem de máquina

16 de dezembro - apps serveless





# Quem participa das Jornadas Digitais?

O público-alvo é composto por servidores e empregados públicos e terceirizados da área de TI das Secretarias, vinculadas e agencias reguladoras do estado do Ceará.

**Convide as outras pessoas da sua equipe!**

# INFOS SOBRE DECLARAÇÃO/CERTIFICADO

- Será gerado um certificado ao final dos 10 webinários que listará todos os que você participou
- Caso precise de uma declaração de um workshop específico, mande um email para [ava@egp.ce.gov.br](mailto:ava@egp.ce.gov.br)
- Essas e outras informações estarão no hotsite que a EGP está finalizando e que segunda feira estará disponível

# Apresentação



**Jairo da Silva Junior**



# Ciência de Dados na AWS

## Transformando dados em insights

Jairo da Silva Junior [jrsilva@amazon.com](mailto:jrsilva@amazon.com)

Solutions Architect



***“Invention requires two things:***

- 1. The ability to try a lot of experiments*
- 2. not having to live with the collateral damage of failed experiments.”*

***- Andy Jassy (Amazon Web Services CEO)***

# Ciência de Dados?

Processos e sistemas para extrair **conhecimento** ou **insights** de dados, estruturados ou não.

# Desafio 1...

Qual o **problema de negócio** a ser resolvido?

## Lean

Um **problema** não vale a pena ser resolvido até que você possa **provar** que ele existe.





Visuals



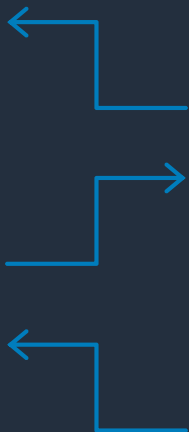
Press  
Release



FAQ



Cliente



## *Working backwards*

Use para aumentar a clareza e não  
para documentar o que já decidiu



## Quem é o cliente?

Seja específico sobre seu contexto e necessidades.





### Working Backwards Questions

Who is the **customer**?

What is the customer problem or **opportunity**?

Is the most important customer **benefit** clear?

How do you know what customers **need** or **want**?

What does the customer **experience** look like?

Wiki > Working Backwards

# Artefatos do processo Working Backwards

## Press Release

## FAQs

## Visuals

Amazon Campus

Amazon.com confidential

Amazon Campus Launches First Multi-Channel Student Store at the University of California Berkeley

[www.amazon.com/Campus&berkeley.amazon.com](http://www.amazon.com/Campus&berkeley.amazon.com)

SEATTLE (BUSINESS WIRE) – August 31st, 2014 – Since returning to school for the Fall semester, Berkeley experienced a fresh take on the college bookstore as the new Berkeley Student Store opened at an on-campus location.

The Amazon-run online Student Store brings Amazon's broad selection, low prices, and convenient shipping to campus and offers a number of additional student-focused features. With the launch of the Student Store, students can now pay for any purchase on Amazon with their CallCard, buy course materials in any Learning Management System or directly on Amazon.com, pick up orders at convenient locations across campus, and have access to an Amazon Student Lounge.

"I've been an Amazon Student member for two years," explained Senior Aubrey **Chen**. "Getting it was huge. I've been ordering everything from Amazon since I got back to school. I just never have to go to a bookstore."

Students can access these campus-specific payment methods and shipping benefits, along with the full range of Amazon's products, by signing up for Amazon Student and selecting "Berkeley" as their campus. Students can also sign up for Amazon Student and select "Berkeley" as their campus within the Berkeley Learning Management System.

In place of a traditional campus bookstore, Amazon has deployed lockers across campus and built the space that features a package pick-up desk, technology showrooms, and lounge area for students to

Amazon Campus

Amazon.com confidential

### Student Customer FAQs

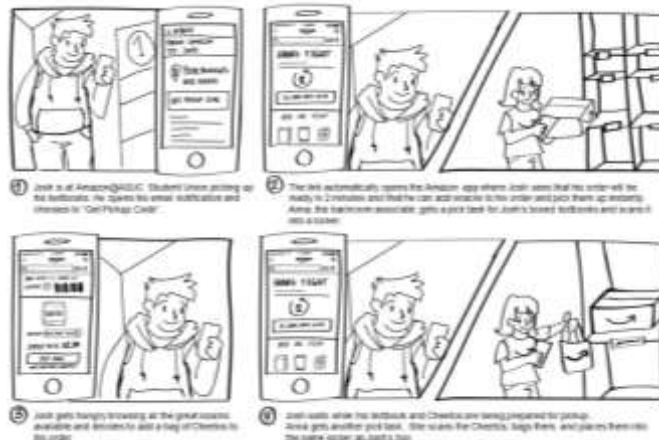
**Q4: What do I get when Amazon is on my campus?** Students get a number of benefits when Amazon is on campus. The co-branded experience on Amazon, students receive Free One Day shipping on textbooks to campus, the ability to pay with their university ID card, and the ability to pick up Amazon orders on campus. On campus, all students will have access to an Amazon Student Lounge.

**Q5: How do I activate the co-branded experience?** Students at client universities will have four methods to activate the co-branded experience: 1) visiting the co-branded subdomain directly by typing the URL [berkeley.amazon.com](http://berkeley.amazon.com), 2) clicking an external link, 3) joining the Amazon Student program and selecting their school, 4) shopping in their school's Learning Management System, 5) clicking on a Student Store merchandising placement on customers have opted in, they will see a persistent co-branded toolbar on the top of the page whenever Customers can turn the experience off at any time by navigating to the "Account Settings" menu in "Your Account".

**Q6: What do I get if I sign up for Amazon Student when Amazon is on my campus?** Amazon Student receives the standard benefits of the program plus Free One Day shipping to the campus area on millions of top-selling items. Benefits include Free Two-Day shipping on Prime eligible items, unlimited instant streaming of Prime video content to the Kindle Owners' Lending Library. The cost of a Student membership will remain \$35/year.

**Q7: Why aren't there any "grab and go" items available for purchase at the Student Lounge?** At launch, products to customers "on demand" at the Student Lounge because we do not have a Point of Sale system in place for sale on site. However, we intend to offer this in the future for a small selection of top-selling items. Customers may order items displayed in the Lounge by using the bar code scan feature in the Amazon app.

**Q8: Why is Free One Day shipping limited to campus addresses? Why don't I get Free One Day shipping everywhere?** By working directly with a university, Amazon is able to reduce costs associated with shipping orders to campus to pass on to all students at the school via Free One Day shipping for textbooks, and to our Amazon Student Store on millions of other eligible products. However, in order to provide these benefits at no additional cost to the university, we offer this benefit to those zip codes served from our on-campus delivery station. Members always receive Free Two Day shipping on Prime eligible items, to campus or most other addresses.



# FAQs

Inclui FAQs do **cliente**  
e **interno**

Inclui perguntas difíceis

Compartilhe o seu *Press Release*  
o quanto antes para reunir  
questões

## Sample PRFAQ Questions

What are you launching today?

What is the customer problem or opportunity?

Why should a customer use product X?

What is the primary customer benefit?

How do you know what your customers need or want?

What does the customer experience look like?

How does this product relate to/work with other products in our portfolio?

What would your customers most like in the product?

[Internal] What do you want someone sitting across the table at lunch saying about your product?

[Internal] Who is the target customer for your product?

[Internal] What will customers be most disappointed with in your product?

[Internal] Why do you believe the first release will be successful?

[Internal] What use cases would customers not want to move to your product, and why? What % of the market segment do these use cases comprise?

[Internal] Are we currently making any decisions that are irreversible?

[Internal] What is your plan to make customers a glowing public reference of your service?

[Internal] Will customers interact with your new service/feature through a UI?

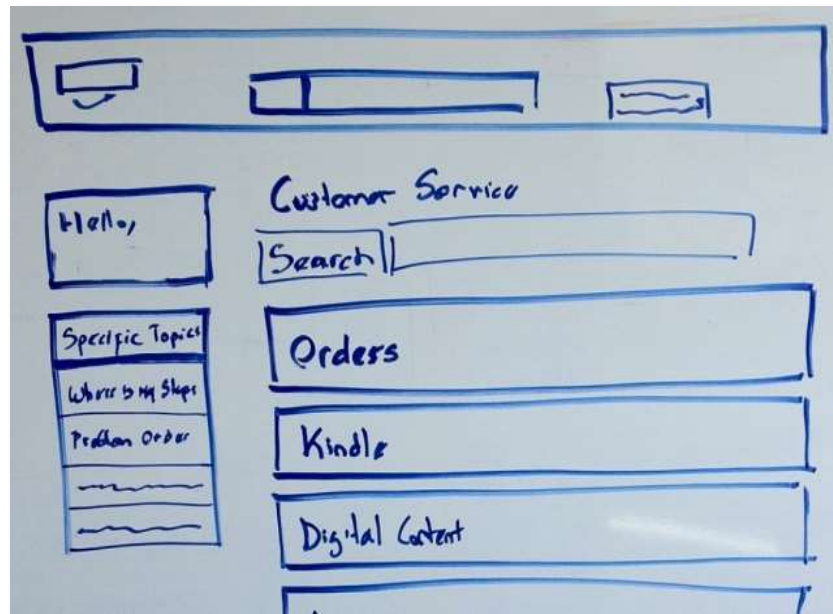
# Visuals

Mostra a experiência do cliente

Ideia aproximada – desenho aproximado

Aumente a fidelidade para amadurecer sua ideia

Não tenha medo de ser disruptivo





Nós lemos,  
discutimos,  
debateamos,  
e fazemos  
perguntas

Knowing that Amazon will send them the right-sized diaper at the right time.

Customer Benefit

Babies need an average of about 3,800 disposable diapers across seven sizes in three years, and remain between a few weeks to several months. A baby might need just one box of size 1 diapers but need as many as eight boxes of size 4 diapers. To complicate matters, weight — and not age — is the best predictor of a baby's diaper size.

With Amazon's Subscribe & Save program, customers previously created diaper subscriptions in a specific size which was replenished until the customer decided it was time to move up a size. This meant that parents needed to anticipate the change in diaper size and either cancel their existing subscription to create a new one, or go to the Manage Your Subscriptions page to update the existing diaper subscription with the new size. This caused endless frustration for parents, added to the stress of having a new baby, and the experience ran counter to the concept of a subscription that should run smoothly with minimal touch points.

Not a good name for this

With the introduction of the new auto-sizing feature, customers who subscribe to diapers will be asked to provide their child's weight, in addition to the existing child info. They will then be able to set up a single subscription that will take them through the various sizes at the right time. Customers will be notified before the next box of diapers is due to ship that it has automatically sized up — if it's too soon, there will be a simple one-click process to keep the current size.

Could a customer really say this?  
"I love using Subscribe & Save. But for my older child, I have had to return boxes that were one size too small, which actually increased my work load," says Mara Steiner, who has used Subscribe & Save for her baby needs for several years. "As I expect my second child, I am very excited at the thought of having the right size diapers show up at my door each month. New parents have enough to do without having to keep calculating which size of diapers they should be buying for their baby."

"We want parents to always have the right sized diaper for their child at the right time," said Greg, director of Subscribe & Save. "This single subscription experience for diapers will ensure that customers can depend upon Subscribe & Save to do the work for them, and never have to worry about changing size as their children grow."

Amazon's Subscribe & Save ships thousands of items to their customers on a regular basis with additional discounts. With the diaper subscription service, the program has simplified the experience for new parents.

To learn more about Subscribe & Save's diaper subscriptions auto-sizing, go to [www.amazon.com/diaper-subscription](http://www.amazon.com/diaper-subscription)

What about mobile?

Nós lemos,  
discutimos,  
debateamos,  
e fazemos  
perguntas

Knowing that Amazon will send them the right-sized diaper at the right time.

Customer Benefit

Babies need an average of about 3,800 disposable diapers across seven sizes in three years, and remain between a few weeks to several months. A baby might need just one box of size 1 diapers but need as many as eight boxes of size 4 diapers. To complicate matters, weight — and not age — is the best predictor of a baby's diaper size.

With Amazon's Subscribe & Save program, customers previously created diaper subscriptions in a specific size which was replenished until the customer decided it was time to move up a size. This meant that parents needed to anticipate the change in diaper size and either cancel their existing subscription to create a new one, or go to the Manage Your Subscriptions page to update the existing diaper subscription with the new size. This caused endless frustration for parents, added to the stress of having a new baby, and the experience ran counter to the concept of a subscription that should run smoothly with minimal touch points.

Not a good name for this

With the introduction of the new auto-sizing feature, customers who subscribe to diapers will be asked to provide their child's weight, in addition to the existing child info. They will then be able to set up a single subscription that will take them through the various sizes at the right time. Customers will be notified before the next box of diapers is due to ship that it has automatically sized up — if it's too soon, there will be a simple one-click process to keep the current size.

Could a customer really say this?  
"I love using Subscribe & Save. But for my older child, I have had to return boxes that were one size too small, which actually increased my work load," says Mara Steiner, who has used Subscribe & Save for her baby needs for several years. "As I expect my second child, I am very excited at the thought of having the right size diapers show up at my door each month. New parents have enough to do without having to keep calculating which size of diapers they should be buying for their baby."

"We want parents to always have the right sized diaper for their child at the right time," said Greg, director of Subscribe & Save. "This single subscription experience for diapers will ensure that customers can depend upon Subscribe & Save to do the work for them, and never have to worry about changing size as their children grow."

Amazon's Subscribe & Save ships thousands of items to their customers on a regular basis with additional discounts. With the diaper subscription service, the program has simplified the experience for new parents.

To learn more about Subscribe & Save's diaper subscriptions auto-sizing, go to [www.amazon.com/diaper-subscription](http://www.amazon.com/diaper-subscription)

What about mobile?

Sharpie



## Desafio 2...

Preparação de dados representa 80% do trabalho de um cientista de dados

# Por quê?



---

Crescendo  
exponencialmente



---

Novas fontes



---

Maior diversidade



---

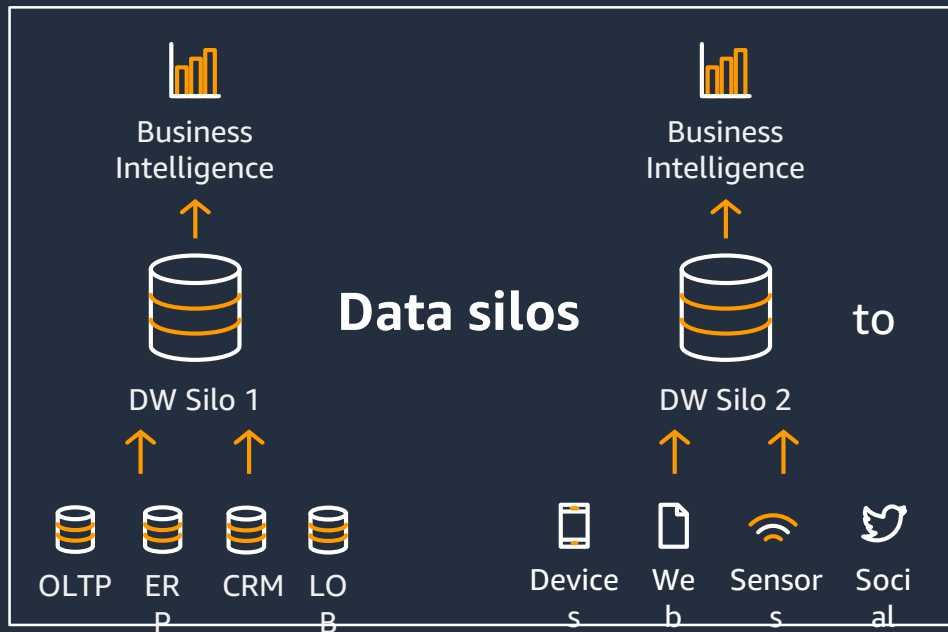
Usado por muitas  
pessoas



---

E aplicações

# Do Data Warehouse tradicional ao Lake House



# Arquitetura de dados moderna



**Data Lakes escaláveis**

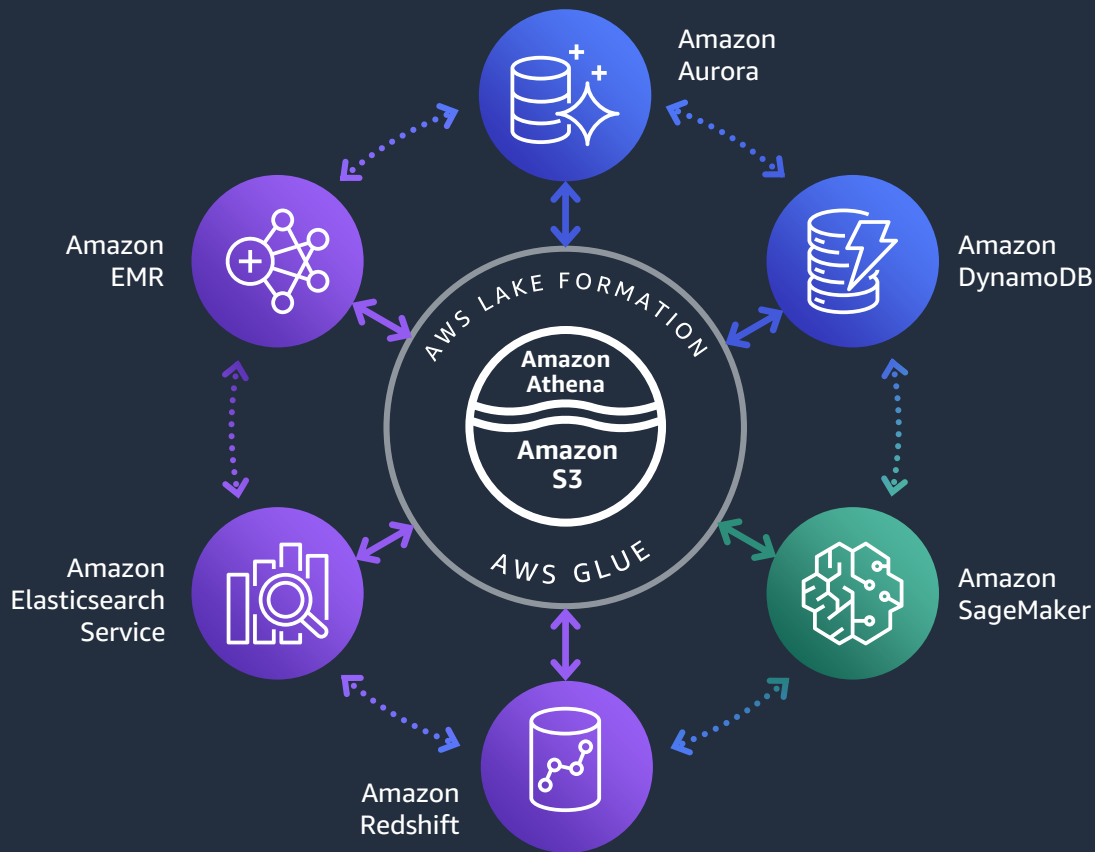
**Serviços de dados com propósito específico**

**Movimentação transparente de dados**

**Governança Unificada**

**Desempenho e otimização de custos**

# Arquitetura de dados moderna na AWS



**Data Lakes escaláveis**

**Serviços de dados com propósito específico**

**Movimentação transparente de dados**

**Governança unificada**

**Desempenho e otimização de custos**

# Data Lakes escaláveis



# Amazon S3 é a escolha mais popular para Data Lakes



# Serviços de dados com propósito específico





# Serviços de dados com propósito específico

Otimize os custos, desempenho e escala para o seu caso de uso



Amazon  
Athena

---

Consultas interativas



Amazon  
EMR

---

Big Data



Amazon  
Elasticsearch  
Service

---

Logs e busca textual



Amazon  
Kinesis e  
Amazon MSK

---

Análise em tempo real



Amazon  
Redshift

---

Data warehousing

# Movimentação transparente dos dados



# Movimentação transparente dos dados

Mova seus dados, em escala, para onde você mais precisa deles



Extract,  
transform, load



Preparação visual dos  
dados



Replicação de  
dados



De/para DW e DL



Consultas  
federadas

# AWS Glue

## Integração de dados simples, escalável e sem servidor



### Conecte a mais fontes de dados

Faça ingestão de dados de centenas de fontes



### Simplifique a orquestração de workflows

Execute e gerencie milhares de jobs de integração de dados



### Sem servidor para gerenciar

Pague apenas pelos recursos que seu job consumir

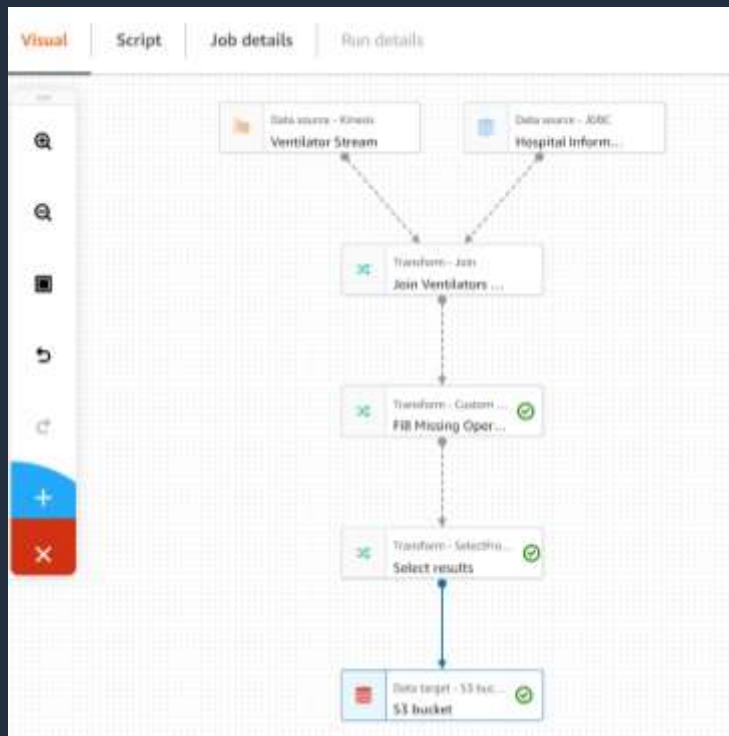


### Simplifique o desenvolvimento

Desenvolva e gerencie jobs de integração de dados visualmente

# AWS Glue Studio

Crie, execute e monitore Jobs ETL do AWS Glue



Crie AWS Glue jobs **visualmente** sem código

Monitore **1,000s de jobs** através de uma **interface única**

**Processamento distribuído** sem curva de aprendizado

Transformações avançadas **usando snippets de código**

# AWS Glue DataBrew

## Preparação visual de dados para analytics e machine learning

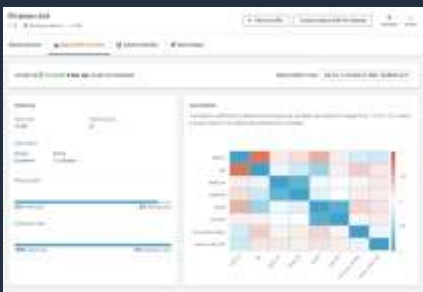


Limpe e normalize dados com uma interface visual

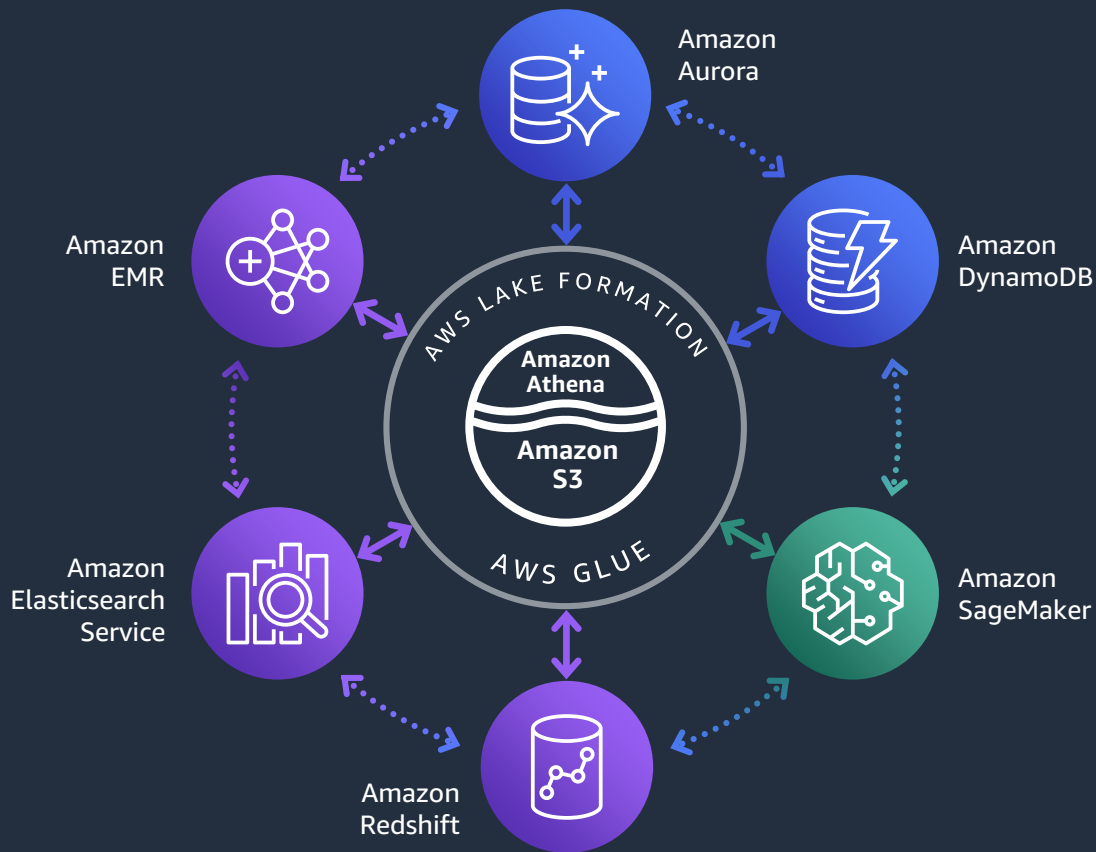
250+ transformações prontas sem necessidade de codificar

Entenda padrões de dados e anomalias

Trabalhe em grandes datasets em escala



# Governança Unificada



# AWS Lake Formation

Construa um Data Lake seguro em dias



## Construa data lakes facilmente

Mova, armazene e catalogue seus dados rapidamente; simplifique o gerenciamento de dados



## Simplifique o controle de acesso

Defina e garanta as políticas de segurança, governança e auditoria de forma centralizada

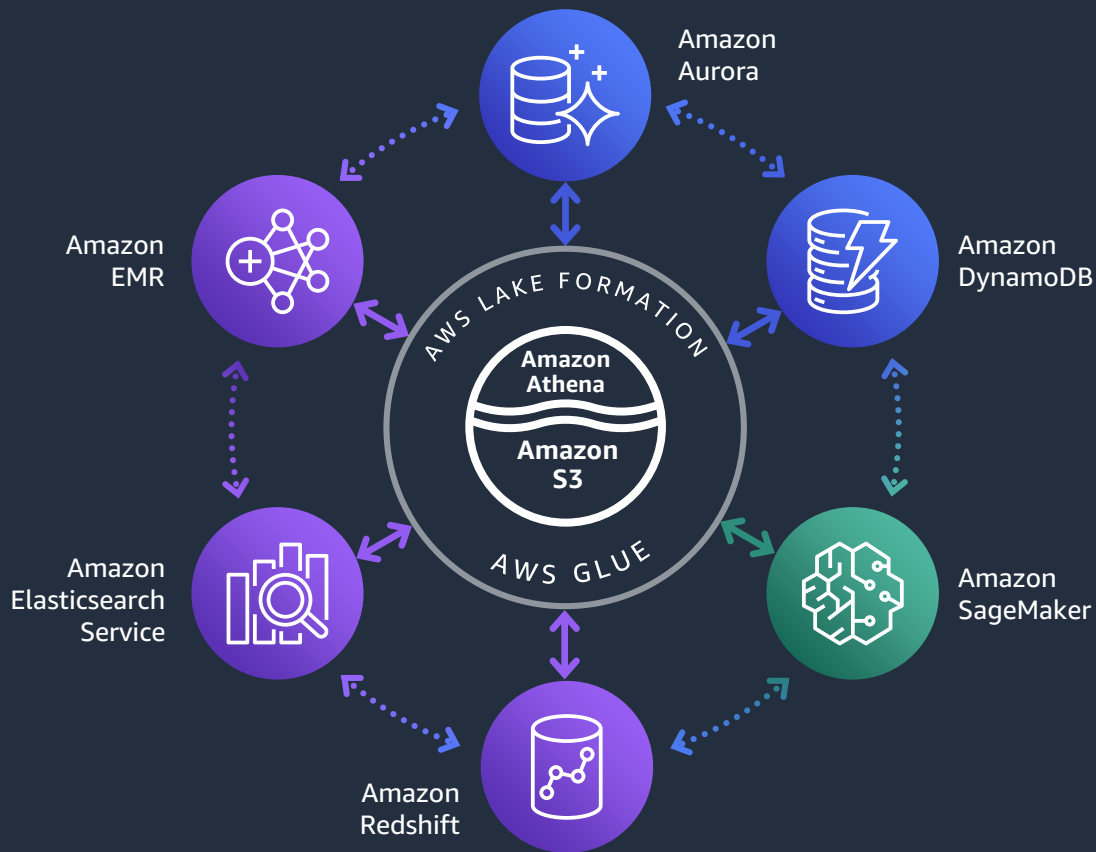


## Acesso self-service aos dados

Compartilhe datasets facilmente e de forma segura com sua organização e parceiros



# Performance e Otimização de custos



# Performance e otimização de custos



---

+ de 200 tipos de instâncias



---

Interfaces de rede de até 100Gbps



---

Sob demanda, reservada  
e/ou spot para reduzir custos



---

Cinco classes de storage  
altamente disponíveis e tierização  
inteligente

# Reinventando business intelligence



# Amazon QuickSight



Serviço de BI escalável,  
embutível e construído para  
a Nuvem



## BI em escala

Sem servidores para gerenciar. Pague por usuários.



## Análise de dados embutível

Embuta dashboards na sua aplicação



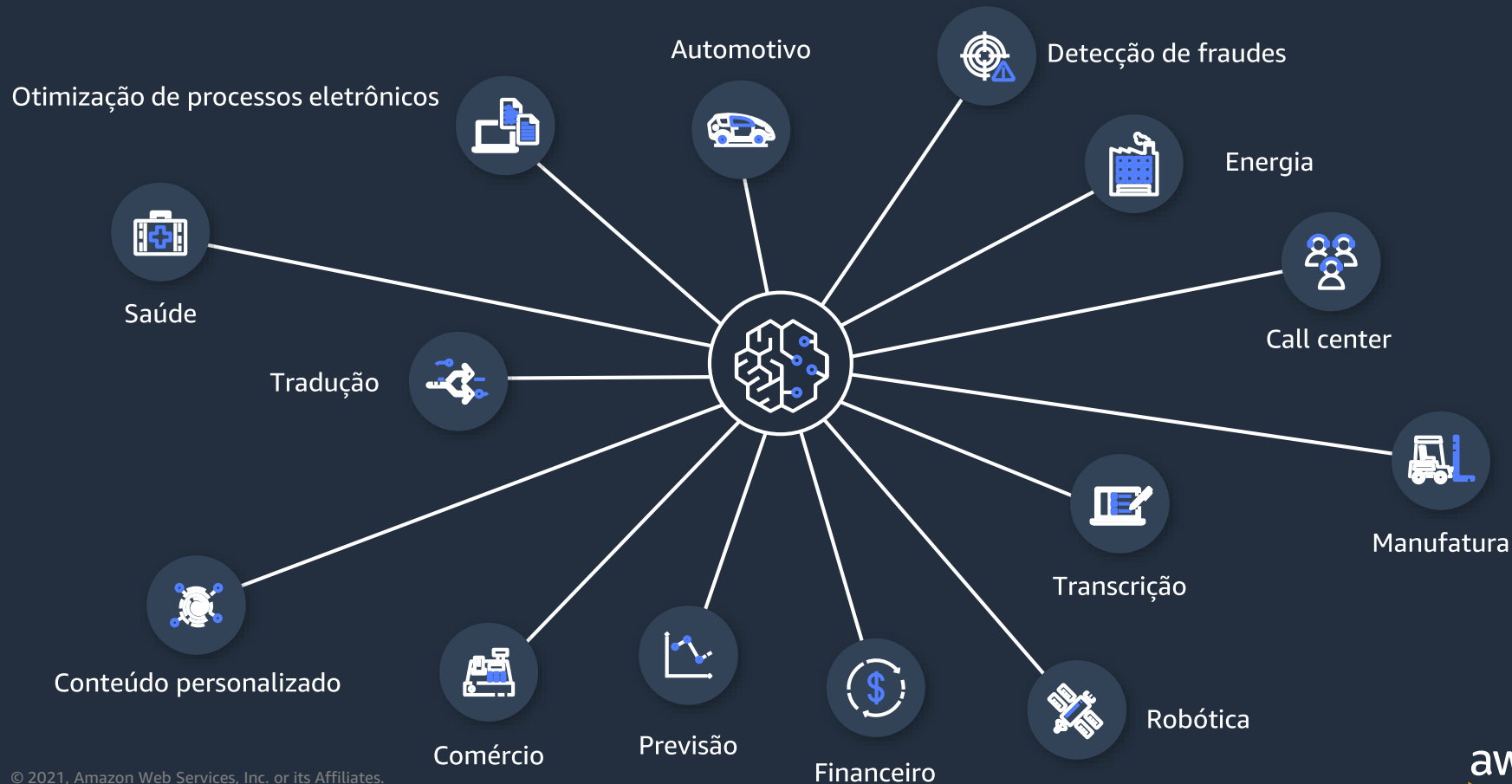
## ML-powered insights

Detecção de anomalias, forecasting e narrativas para interpretar seus dados

## Desafio 3...

Conhecimento em ML é raro; é difícil de  
construir, escalar e manter modelos

# O alcance de ML está aumentando...



# Acelere sua jornada de análise preditiva e de ML



# Amazon SageMaker: Construído para tornar ML **mais acessível**





# Obrigado



# Agenda

- 12 de agosto - LGPD (Mariana Zonari e Fernando Gebara) - ok
- 26 de agosto - Ciência de dados (Jairo da Silva Junior) - ok
- 09 de setembro - **DevOps - Não Perca!**
- 23 de setembro - Estratégias de Migração para a nuvem
- 07 de outubro - modernização de backup
- 21 de outubro – Experiencia do Usuário (UX)
- 04 de novembro - Nuvem híbrida
- 18 de novembro - Containers
- 02 de dezembro – Inteligencia Artificial e aprendizagem de máquina
- 16 de dezembro - apps serveless

Inscrição no hotsite da EGP. O Endereço será enviado na segunda