

Secretaria de Estado de
Finanças



Governo do Estado de
RONDÔNIA

Jornada de Transformação Data- Driven da SEFIN RO

Apresentação COGEF TI

Agenda

- 1 Contexto data driven

- 2 Jornada SEFIN-RO

- 3 Big Data : Implantação e Arquitetura

- 4 Adoção da cultura data-driven

- 5 Próximos Passos



1

Introdução

Contextualização, conceito *data driven*

O que é Data-Driven

Uma organização orientada por dados é aquela que utiliza as informações como base para tomar as decisões estratégicas e operacionais.

Nessas organizações os dados são seu principal ativo e tem, portanto, uma posição de destaque para guiar as suas ações.



Tomada de Decisão

As decisões são tomadas com análises de dados e não apenas com a intuição e a experiência pessoal



Cultura Data-Driven

Desde a alta gestão até a linha de frente reconhecem a importância dos dados e utilizam para tomar decisões



Infraestrutura Robusta

A organização possui sistemas e capacidade computacional para coletar, armazenar, analisar e visualizar dados de maneira eficiente



Habilidades e Conhecimento

Os colaboradores são treinados em análise de dados e possuem as habilidades necessárias para interpretar e utilizar dados de maneira eficaz

Motivos

01

Os principais ativos das Secretarias de Fazenda são documentos fiscais eletrônicos armazenados em dados

02

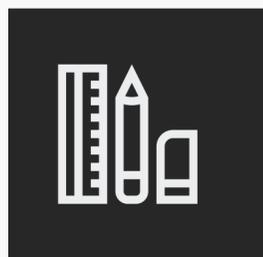
Aumentar a produtividade e reduzir o hiato tributário de maneira direta e indireta.



Segundo Gartner:

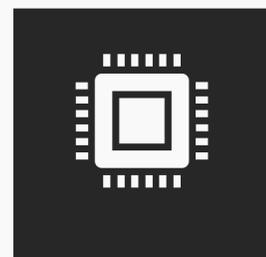
- 1. Os negócios digitais prosperam com base em dados e insights, mas muitas instituições têm as suas ambições inibidas pelos seus comportamentos anteriores. Eles tratam dados e análises (D&A) como apoio e secundários às suas iniciativas de negócios.*
- 2. Muitas vezes as instituições não têm foco em D&A. Muitos ainda não nomearam um executivo responsável pelos dados empresariais e ativos analíticos, pela criação de valor e pelas competências institucionais necessárias para o sucesso.*
- 3. Os líderes das unidades de negócios estão exigindo usar mais dessas ferramentas para permitir que os técnicos de negócios (identificados como “produtores de tecnologia” que não se reportam à TI) produzam aplicativos e automações para suas necessidades de negócios*

Como se tornar uma organização Data-Driven



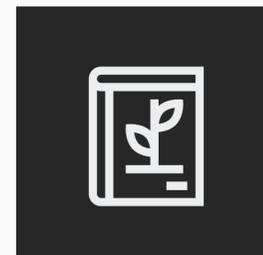
Definir uma visão e estratégia

É importante definir o que a organização quer alcançar com a transformação data-driven e quais são os seus objetivos.



Construir a infraestrutura de dados

A organização precisa investir em sistemas e ferramentas para coletar, armazenar, analisar e visualizar dados de forma eficaz.



Desenvolver as habilidades e conhecimentos

Os colaboradores precisam ser treinados em análise de dados para que possam utilizar os dados de forma eficaz.



Adotar a cultura data-driven

A organização precisa criar uma cultura que valorize o uso de dados na tomada de decisões em todos os níveis

“A transformação digital não tem a ver com tecnologia, tem a ver com estratégia e novas maneiras de pensar”

David L. Roger (2017)





2

Jornada SEFIN-RO

Como a forma de adquirir, tratar e
consumir dados evolui desde 2016

Jornada SEFIN Rondônia

Ponto de Partida

2016 Processamento de malhas fiscais

- Iniciativas localizadas com processamento limitado ao equipamento e acesso restrito aos autores

2019 Ampliação massiva do uso de dados

- Estruturação do Fisconforme e automatização do processo de autorregularização dos contribuintes
- Experiências iniciais com BI e visualização de dados
- Disseminação do uso do SQL como ferramenta de trabalho com aumento de consultas sob demanda

- Início de capacitação em Ciência de Dados
- Realização de Workshop sobre “Oportunidades de uso de Analytics”

2017 Implantação do ambiente analítico

- Migração para o banco Oracle
- Criação de ambiente analítico dedicado para a Gerência de Fiscalização
- Estruturação de dados importantes no novo ambiente

2020 Início de direcionamento institucional

- Organização do catálogo de soluções de tecnologia (In. 5/2020)
- “Ampliar a capacidade do uso de ferramentas de BI e Analytics” é eleito como projeto estratégico
- Primeiras conversas sobre metodologias e práticas de desenvolvimento

→ 2021 2022 2023 2024

Jornada SEFIN Rondônia

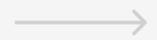
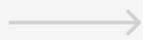
2021 Avanços em práticas de projeto

- Estabelecimento de metodologias para o desenvolvimento das soluções e realização de projetos de referência
- Portifólio para acompanhamento dos projetos
- Mapeamento dos dados de interesse em áreas de assunto
- Migração para nova ferramenta de BI

2023 Estabelecimento de novas bases

- Implementação da metodologia OKR para objetivos estratégicos
- **Implementação do Centro de Excelência em Analytics (CoE)** ▶
- Capacitação em ferramentas analíticas
- **Implantação da infraestrutura e ambiente do Big Data e desenvolvimento de solução piloto** ▶

2020



2024

2022 Avanços em governança

- Realização de seminários internos sobre dados
- Início da capacitação em Análise de Dados e Engenharia de Dados
- MBA em Projetos e Processos
- Evolução das práticas de desenvolvimento de pipelines (observabilidade, desacoplamento)
- Nomeação de Unidades Supervisoras de Dados



3

Big Data

Implantação e Arquitetura

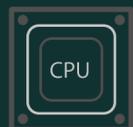
Implementação da arquitetura de **Big Data**

Motivações



Diminuir o armazenamento em ambiente Oracle

Por questões de **volume** e **custo** a Secretaria estabeleceu como objetivo diminuir a necessidade e dependência de investir em storage no ambiente transacional



Capacidade e tecnologia de processamento

Expectativa de que soluções analíticas processem com maior velocidade e que soluções mais flexíveis e robustas possam ser construídas no novo ambiente



Prover infraestrutura exclusiva para dados

A coexistência e concorrência de recursos entre os ambientes transacionais e o analíticos limita iniciativas relacionadas à governança de dados e traz dificuldades na gestão de ambientes com propósitos diferentes

Premissas



On-premisses

A implementação do Big Data foi pensada como expansão da nossa infraestrutura existente onde os serviços de dados assumem uma posição central e demais sistemas convergem para um repositório único de documentos e escriturações fiscais.

Nesse contexto, o equipamento foi adquirido para ser compatível com as soluções de virtualização existentes (hiperconvergência).



Open source

No ambiente on-premisses, a utilização de componentes de código aberto baseados no ecossistema Hadoop permite flexibilidade e adaptabilidade além de eliminar custos com licenças de software.

Os principais fornecedores de plataformas de dados utilizam soluções baseadas no ecossistema Hadoop.

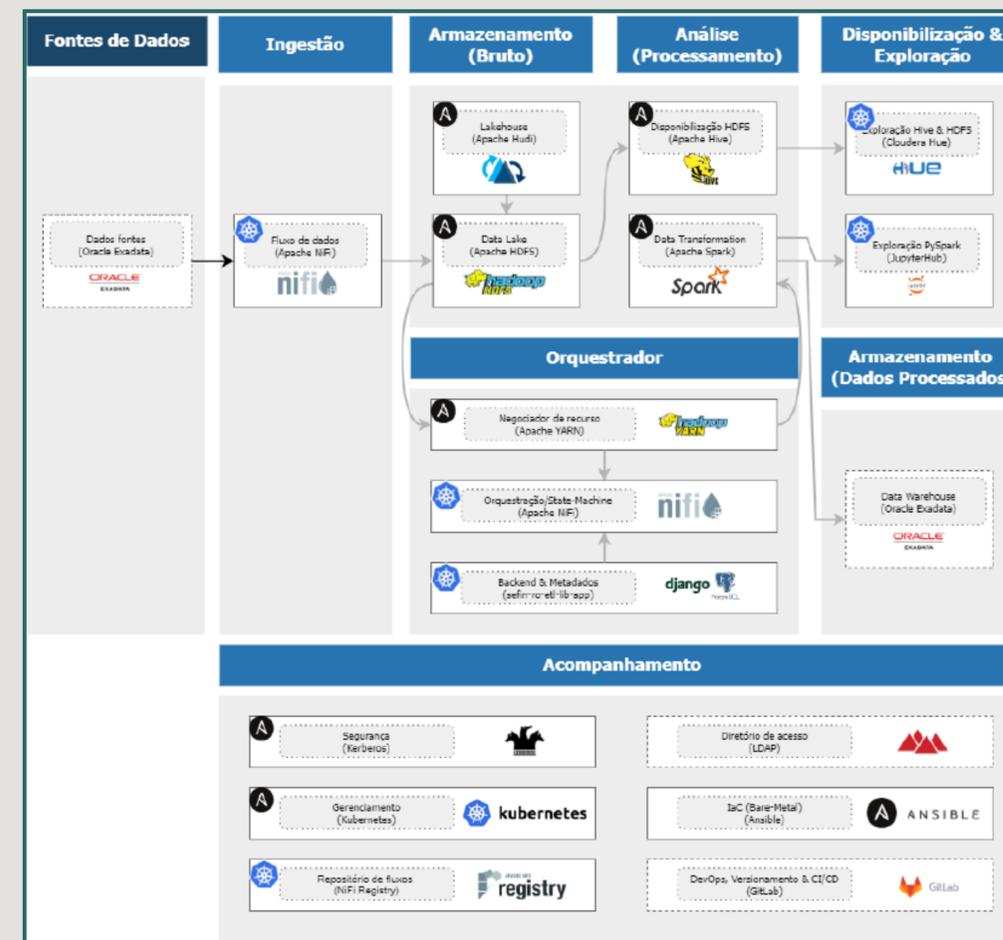
Implementação da arquitetura de Big Data

Equipamento

Fabricante	Nutanix
Modelo	NX-8150N-G8
Chassis	6
Nós/Servidores	6
CONFIGURAÇÕES POR SERVIDOR	
Processadores	02 x Intel Xeon Gold 6346, 3,1 GHz GHz, 16 núcleos (Total 32 Núcleos)
Memória	16x 64 GB 3.200 MHz DDR4 RDIMM (1024 GB)
Discos SSD	12x 7.68 TB

Fabricante	Nutanix
Modelo	NX-3155GN-G8
Chassis	2
Nós/Servidores	2
CONFIGURAÇÕES POR SERVIDOR	
Processadores	02 x Intel Xeon Gold 6326, 2,9 GHz GHz, 16 núcleos (Total 32 Núcleos)
Memória	16x 64 GB 3.200 MHz DDR4 RDIMM (1024 GB)
Discos SSD	6x 7.68 TB
GPU	01 x NVIDIA A100 GPU Card (80GB)

Arquitetura de serviços



Desenvolvimento

Malhas desenvolvidas	
<p>DIV_VAL_C500</p>	Essa tabela representa as notas que estão com divergências de valor entre a EFD e a base de NF3e
<p>DIV_VAL_C590</p>	Essa tabela representa as notas que estão com divergências de valor entre os grupos de itens agrupados na EFD e a base de NF3e.
<p>DIV_EMI</p>	Verifica notas válidas que não possuem escrituração válida no grupo c500 e escriturações válidas que não possuem notas correspondentes na base de nf3e
<p>DIV_VPROD</p>	Verifica dentro da nota se o somatório dos valores dos produtos totaliza de maneira correta o valor de produtos total.

NF3e
caso de uso

25 tabelas/
modelos de dados/
Scripts (pyspark)

4 camadas
raw/trustes/stage/
refined

8TB
RAM

490TB
SSD

128
CPU
cores

80GB
GPU

13
componentes
principais

2
clusters
Hadoop e K8s

3
ambientes
dev/hml/prd

ADD
IaC, CI/CD,
Segurança,
Lineage

Implementação da arquitetura de **Big Data**

Investimento

Equipamento

5.303.860,00

Implementação da arquitetura

638.940,00

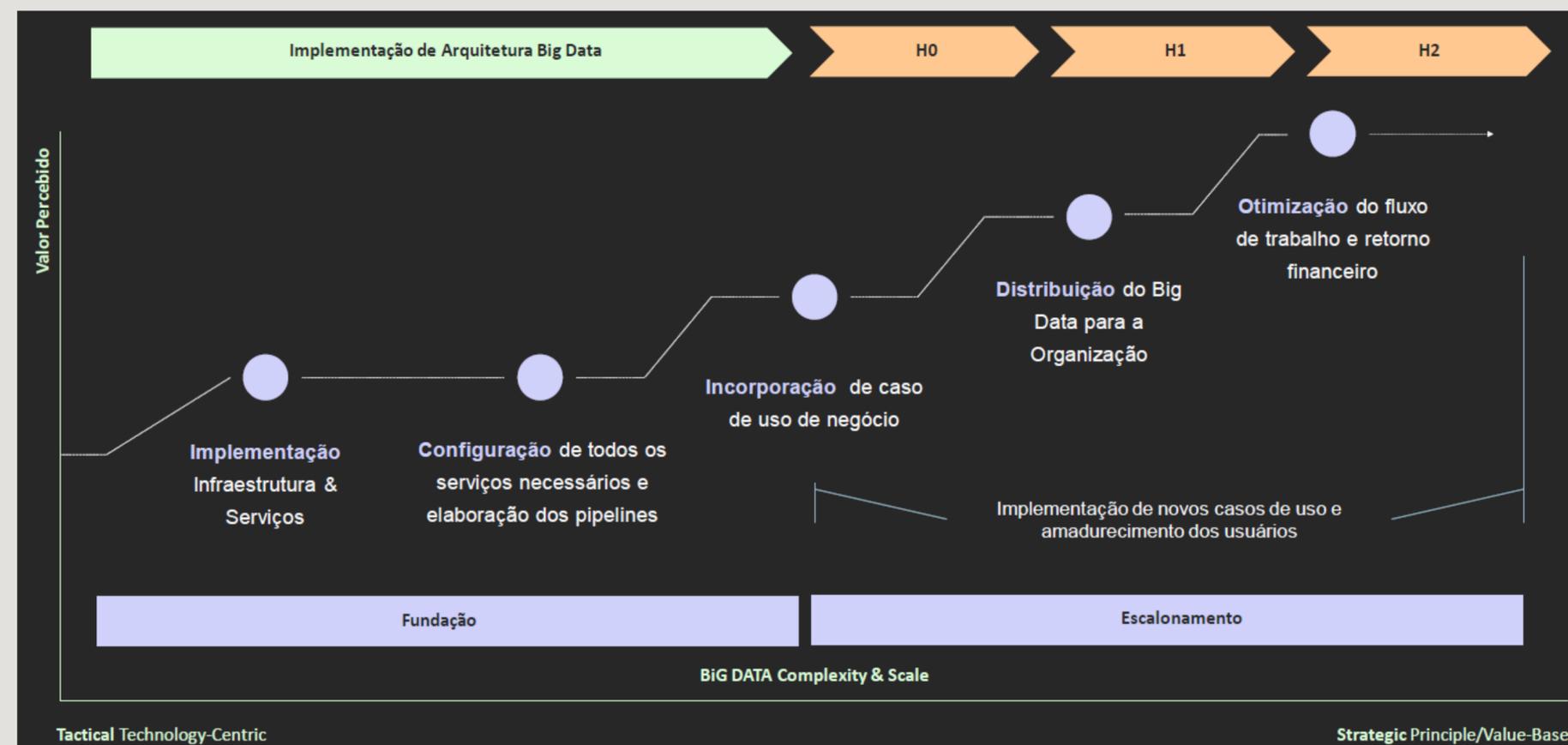
Desenvolvimento do caso de uso

365.700,00

Total do investimento

6.308.500,00

Processo de percepção de valor





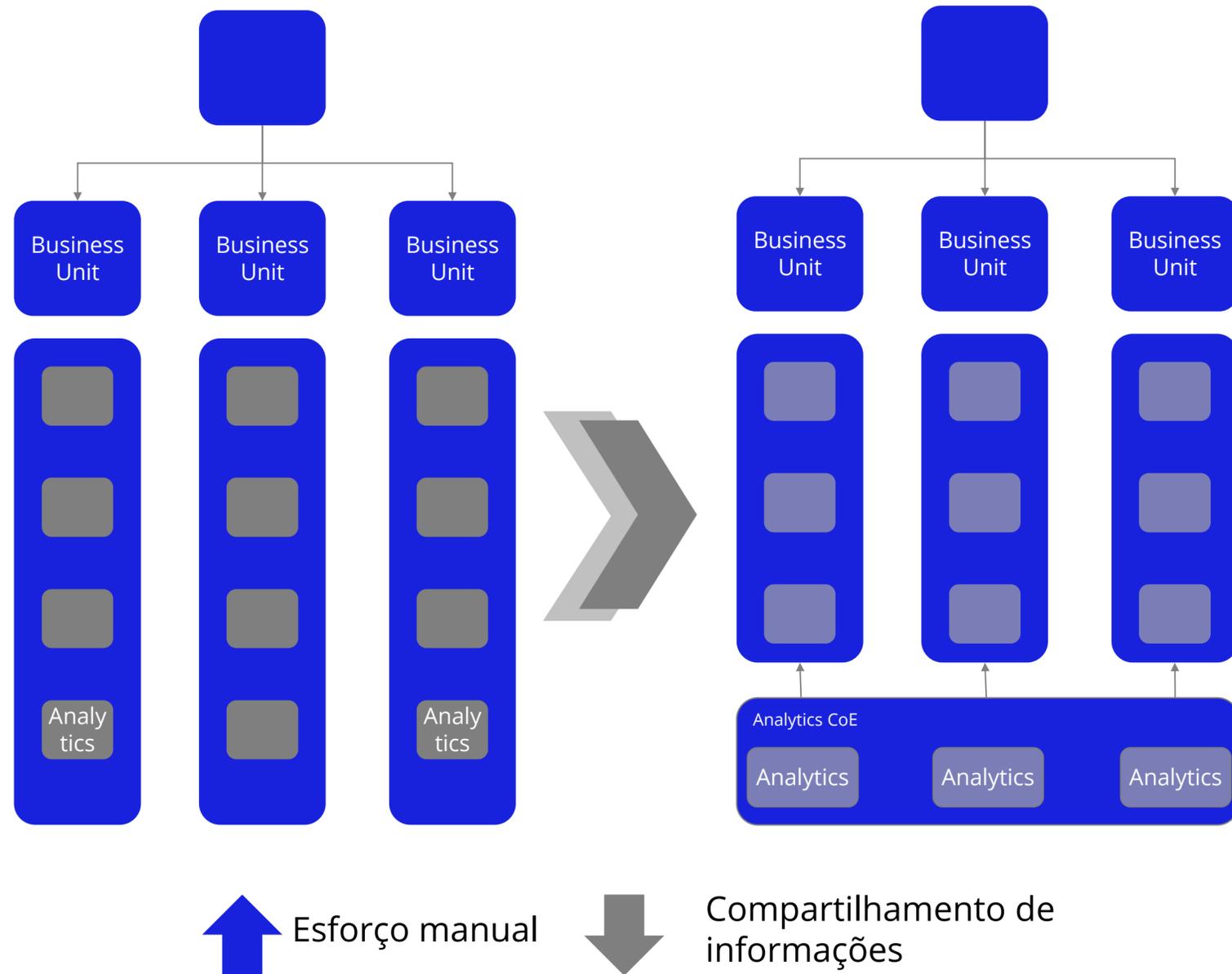
4

Adoção da cultura

Data-Driven

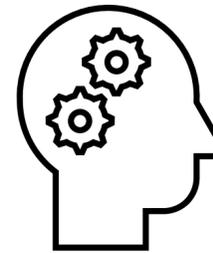
Centros de Excelência (CoE)

O que é um Centro de Excelência (CoE)



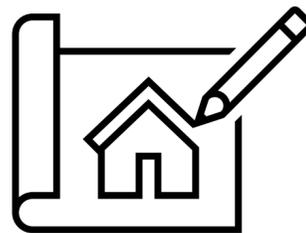
“Um Center of Excellence é um time com colaboradores dedicados, gerenciados por um ponto central em comum – separado das áreas funcionais que o Centro suporta. Também conhecido como Centro de Competências ou Capacidades, o CoE é o time que lidera a exploração e a adoção de novas ferramentas, técnicas ou práticas tecnológicas”

(Five Guiding Principles of a Successful Center of Excellence)



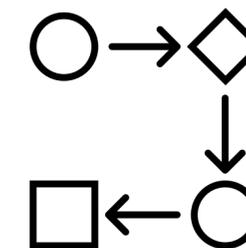
Pessoas

Capacidade de Gestão e Execução dos Projetos e Processos do CoE



Projetos

Iniciativas para ampliar capacidade da organização em Analytics

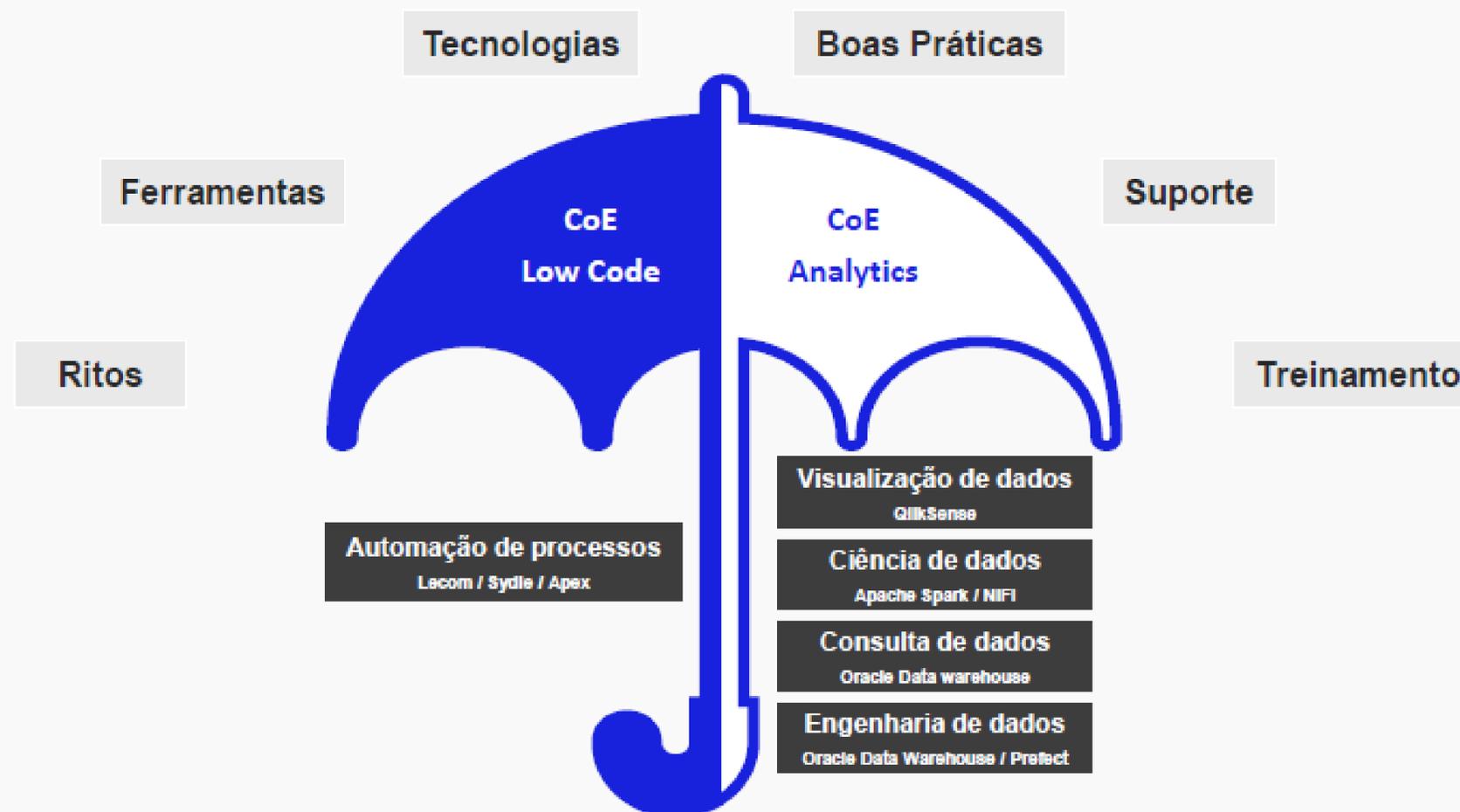


Processos

Rotinas, ritos e artefatos que fazem com que o CoE funcione da forma esperada

Os CoEs são figuras centrais para o gerenciamento de demandas de tecnologia

“Guardiões da Governança”





Visão geral

CoE - Analytics

PAPÉIS E RESPONSABILIDADES

ESTRATÉGICO

- Patrocinador do Projeto
- Coordenador do CoE Analytics
- Líder Técnico

TÁTICO

- Gestor do Negócio (CRE)
- Gestor do Negócio (COTES)
- Gestor de Tecnologia

OPERACIONAL

- Product Owners
- Representante do Processo
- Analista de Dados
- Cientista de Dados
- Engenheiro de Dados

Data Specialists

LEGENDA:

- Atuação fixa no CoE
- Atuação sob demanda

RITOS DE GOVERNANÇA

Clique no rito para visualizar o detalhamento dos ritos

- Acompanhamento do projeto
- Governança do CoE
- Status Mensal
- Reunião de Refinamento

PLATAFORMAS UTILIZADAS



Visualização de dados



Banco de dados



Desenvolvimento de software

ARTEFATOS/DOCUMENTAÇÃO

Clique aqui para visualizar o detalhamento dos artefatos

Coleta e priorização de demandas

Planejamento

Preparação e Execução

Sustentação e Melhorias

PORTFÓLIO DE PRODUTOS



- Consulta de Dados
- Consulta
 - Suporte a consulta existente



- Ciência de dados
- Análise de dados simples
 - Análise de dados complexa



- Visualização de Dados
- Painel simples
 - Painel complexo
 - Manutenção/Incremento de painel



- Engenharia de Dados
- Disponibilização de dados
 - Manutenção ou melhoria de fluxos de dados
 - Migração de Dados

FLUXO DE PRODUÇÃO DO COE

Clique aqui para visualizar o detalhamento do fluxo

Coleta e Priorização da demanda

- Coleta de demanda ativa e/ou formulário
- Priorização considera o Impacto x Esforço Estimado

Planejamento da Execução

- Detalhamento do planejamento
- Registro de demandas em ferramenta de gestão

Preparação e Execução

- Definição das bases de dados necessárias
- Desenvolvimento da solução

Homologação

- Homologação da solução final
- Produção dos documentos técnicos

Operação Assistida

- Monitoramento da solução em ambiente definido
- Consolidação de resultados
- Em caso de necessidade, seguir para implantação

Implantação (Deployment)

- Definição da arquitetura da solução
- Parametrização e teste em novo ambiente
- Implantação da solução

BACKLOG

Área	Gerência	Projeto	Plataforma
COTES	GEOP	Adicional Cronograma de Desembolso	
COTES	GEOP	Painel de gestão do Orçamento	QlikSense
CRE	NEEC	Mensuração de Contrapartidas de Conformidade	
CRE	GEFIS	Painel de Gestão de Performance Fiscalização	QlikSense
CRE	GEFIS	Implementação da Arquitetura de Armazenamento do Processamento NF3 em Big Data (MVP1 – Batch)	
CRE	GEFIS	Modelo de Avaliação da Assimetria na Arrecadação de ICMS	
CRE	GEAR	Painel de Gestão da Inadimplência	QlikSense

SOLUÇÕES EM PRODUÇÃO

Painel de Acompanhamento de Processos Virtualizados



Painel de Gestão do Nota Legal



Diagnóstico Capacidade Atual (Visão Consolidada da Autoavaliação)

Analisando de forma sintetizada as informações obtidas por meio de *assessments*, foi observada a seguinte conjuntura:

		ÁREA DE CONHECIMENTO				
		1  Engenharia de Dados	2  Ciência de Dados	3  Visualização de Dados	4  Consulta de Dados	
NEGÓCIOS	COTES	Conhecimento	Baixos	Baixos	Baixos	Baixos
		Aptidão e Interesse	Baixos	Baixos	Altos	Médios
	CRE	Conhecimento	Baixos	Médios	Médios	Médios
		Aptidão e Interesse	Baixos	Médios	Altos	Altos
TECNOLOGIA	GETIC	Conhecimento	Médios	Médios	Baixos	Médios
		Aptidão e Interesse	Altos	Médios	Baixos	Médios

CONCLUSÕES	
1  Engenharia de Dados	<ul style="list-style-type: none"> Existe maior interesse, aptidão e capacidade da área de tecnologia na temática do que as áreas de negócios Necessário desenvolver capacidade e aperfeiçoar estrutura técnica
2  Ciência de Dados	<ul style="list-style-type: none"> Nenhuma área demonstrou capacidade suficiente de desenvolver projetos de Ciência de Dados Habilitar capacidade para desenvolver projetos de maior complexidade
3  Visualização de Dados	<ul style="list-style-type: none"> Existe maior interesse, aptidão e capacidade das áreas de negócios na temática do que a área de tecnologia Necessidade de maior capacitação das áreas de negócio no tema
4  Consulta de Dados	<ul style="list-style-type: none"> Conhecimento nessa área de conhecimento habilita o uso de Analytics, sendo essencial disseminá-lo ao máximo Existe diferença entre o conhecimento e o interesse das áreas de negócio no tema, havendo espaço para capacitações



5

Próximos Passos

Resultados estruturantes orientados à D&A

Compromisso Estratégico da SEFIN

É possível perceber a sinergia da temática Analytics com os objetivos estratégicos da SEFIN, conforme os exemplos abaixo:

MAPA ESTRATÉGICO



Ampliar Uso de Dados para tomada de decisão

- Ampliar capacidade de uso das ferramentas de BI e Analytics



Automatizar a Cobrança Administrativa e Judicial

- Implantação da Cobrança Administrativa Ponta-a-Ponta



Implantação de novas soluções para Fiscalização

- Implantação de Novas Soluções para Fiscalização



Incentivar a Conformidade Fiscal dos Contribuintes

- Implantação da Sistemática de Categorização dos Contribuintes



Obrigado!

