

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
DEPARTAMENTO ACADÊMICO DE INFORMÁTICA  
ESPECIALIZAÇÃO EM CIÊNCIA DE DADOS E SUAS APLICAÇÕES**

**VICTOR DE CERJAT BELTRÃO**

**PROPOSTA DE NOVO PERFIL DE CIENTISTA DE DADOS CITIZEN BASEADO  
NOS CONCEITOS DA GESTÃO DA INFORMAÇÃO**

**CURITIBA**

**2021**

**VICTOR DE CERJAT BELTRÃO**

**PROPOSTA DE NOVO PERFIL DE CIENTISTA DE DADOS CITIZEN BASEADO  
NOS CONCEITOS DA GESTÃO DA INFORMAÇÃO**

**Proposal of Citizen Data Scientist Profile Based on Concepts of Information  
Management**

Trabalho apresentado(a) como requisito parcial à obtenção do título de especialista em Ciência de Dados e suas Aplicações, do Departamento de Informática da Universidade Tecnológica Federal do Paraná.

Orientador(a): Prof(a). Dr(a). Rita Cristina Galarraga Berardi.

**CURITIBA**

**2021**



Ministério da Educação  
**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**  
UTFPR - CAMPUS CURITIBA  
DIRETORIA-GERAL - CAMPUS CURITIBA  
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO - CAMPUS CURITIBA  
DEPARTAMENTO DE APOIO DAS ESPECIALIZAÇÕES LATO-SENSU DOS  
CURSOS DE INFORMÁTICA - CAMPUS CURITIBA  
CURSO DE ESPECIALIZAÇÃO EM CIÊNCIA DE DADOS E SUAS APLICAÇÕES



---

## TERMO DE APROVAÇÃO

### PROPOSTA DE NOVO PERFIL DE CIENTISTA DE DADOS CITIZEN BASEADO NOS CONCEITOS DA GESTÃO DA INFORMAÇÃO

por

**Victor De Cerjat Beltrao**

Este Trabalho de Conclusão de Curso foi apresentado às 19h00 do dia 09 de julho de 2021 por videoconferência como requisito parcial à obtenção do grau de Especialista em Ciência de Dados e suas Aplicações na Universidade Tecnológica Federal do Paraná - UTFPR - Campus Curitiba. O aluno foi arguido pela Banca de Avaliação abaixo assinados. Após deliberação, a Banca de Avaliação considerou o trabalho aprovado.

---

Profa. Dra. Rita Cristina Galarraga Berardi (Presidente/Orientadora – DAINF-CT/ UTFPR-CT)

---

Prof. Dr. Alexandre Reis Graeml (Avaliador 1 – DAINF-CT/ UTFPR-CT)

---

Prof. Dr. Luiz Celso Gomes Junior (Avaliador 2 – DAINF-CT/ UTFPR-CT)

**O Termo de Aprovação assinado encontra-se no sistema SEI- Processo nº 23064.028613/2021-03**

---

Referência: Processo nº 23064.028613/2021-03

SEI nº 2173024

## RESUMO

A alta demanda de digitalização de processos em grandes empresas devido ao cenário atual da indústria 4.0 faz surgir milhares de projetos internos ligados ao uso de grandes massas de dados e ferramentas inovadoras de desenvolvimento em tecnologia de informação. Esses projetos tendem a crescer em quantidade neste novo contexto, porém muitas vezes as empresas não conseguem lidar com os processos internos de gestão de projetos de digitalização da maneira mais apropriada e não possuem uma estrutura designada para atender a estas demandas, criando gargalos e falhas que podem prejudicar o desenvolvimento de muitas iniciativas transformadoras na organização. A Renault propõe atualmente uma estrutura baseada na figura do cientista de dados *citizen*, um profissional que mescla as habilidades da área de negócios e tecnologia da informação para atuar impulsionando os projetos de digitalização da empresa. No entanto, o perfil deste profissional não é bem definido dentro do contexto empresarial, assim como as suas atividades e em que etapas do fluxo de desenvolvimento de projetos ele deve atuar. Com isso, vê-se a oportunidade de aplicar os princípios da gestão da informação para avaliar a execução das etapas de projetos ligados a dados para interpretar possíveis pontos de melhorias, assim como delinear as responsabilidades do novo perfil *citizen* junto as áreas impactadas para maximizar a sua eficiência no desenvolvimento destes projetos de digitalização. As análises realizadas no trabalho permitiram a criação de um novo perfil de cientista de dados *citizen* mais assertivo, agindo em conjunto com um novo fluxo proposto para acompanhamento de projetos de digitalização que permitirá maior eficiência no desenvolvimento das iniciativas digitais da empresa.

**Palavras-chave:** Ciência de dados. Gestão de Projetos. Gestão da Informação. Estrutura empresarial.

## **ABSTRACT**

The high demand for process digitization in large companies due to the current trend of industry 4.0 has led to thousands of internal projects linked to the use of large quantities of data and innovative information technology development tools. These projects tend to grow in quantity in this new context, but often companies are unable to deal with the management process of these internal projects in the most appropriate way and do not have a structure designed to meet these demands, creating bottlenecks and failures that can hinder the development of many transformative initiatives in the organization. Renault currently proposes a structure based on the figure of the citizen data scientist, a professional who merges the skills of the business and information technology area to aid in boosting the company's digitization projects. However, the profile of this professional is not well defined within the business context, as well as its activities and in what stages of the project development flow he should act. Thus, the opportunity is presented to apply the principles of information management to evaluate the execution of data-related project steps to interpret possible points of improvement, as well as to outline the responsibilities of the new citizen profile along with the impacted areas to maximize their efficiency in the development of these digitization projects. The analyses carried out allowed the creation of a new profile of data science citizen, which is more assertive, acting inside a new proposed flow of monitoring for digitization projects that will allow greater efficiency in the development of the company's digital initiatives.

**Keywords:** Data Science. Project Management. Information Management. Enterprise Structure.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Interação entre diferentes áreas de trabalho na ciência de dados.....	13
Figura 2 - Proposta de localização de equipes de ciência de dados.....	15
Figura 3 - Fases do processo de análise de dados.....	17
Figura 4 - Habilidades do cientista de dados citizen .....	20
Figura 5 - Matriz de priorização GUT .....	22
Figura 6 - Evolução da estrutura organizacional interna de pilotos de digitalização .	24
Figura 7 - Evolução da estrutura organizacional interna para criação do data science citizen .....	25
Figura 8 - Fluxo de projetos de digitalização - Caso 1.....	26
Figura 9 - Fluxo de projetos de digitalização - Caso 2.....	28
Figura 10 - Fluxo de projetos de digitalização - Caso 3 .....	30
Figura 11 - Fluxo proposto de projetos de digitalização para grandes empresas .....	36

## LISTA DE TABELAS

Tabela 1 - Comparação entre papel da GI e TI nas empresas.....	19
Tabela 2 – Matriz de priorização GUT.....	32

## LISTA DE QUADROS

Quadro 1 - Dado, informação e conhecimento.....	16
Quadro 2 - Proposta de novo perfil data <i>citizen</i> .....	35

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>9</b>
<b>2 REVISÃO BIBLIOGRÁFICA</b> .....	<b>12</b>
2.1 CIÊNCIA DE DADOS NAS DIFERENTES ÁREAS EMPRESARIAIS .....	12
2.1.1 Profissionais da área de dados .....	12
2.1.2 Inserção da ciência de dados na organização empresarial .....	14
2.2 GESTÃO DA INFORMAÇÃO .....	15
2.2.1 Gestão da informação na análise de dados .....	18
2.3 CIENTISTA DE DADOS CITIZEN.....	19
<b>3 METODOLOGIA</b> .....	<b>21</b>
<b>4 ANÁLISE</b> .....	<b>23</b>
4.1 CENÁRIO ATUAL CIENTISTA DE DADOS CITIZEN - CASO RENAULT .....	23
4.2 ESTUDOS DE CASO.....	25
4.2.1 Estudo de caso – Projeto 1 .....	25
4.2.2 Estudo de caso – Projeto 2 .....	27
4.2.3 Estudo de caso – Projeto 3 .....	29
4.3 MATRIZ DE PRIORIZAÇÃO DE PROBLEMAS ATUAIS .....	31
<b>5 PROPOSTA</b> .....	<b>34</b>
5.1 NOVO PERFIL CITIZEN E FLUXO DE INICIATIVAS DE DIGITALIZAÇÃO....	34
5.2 AVALIAÇÃO EXTERNA SOBRE A PROPOSTA .....	37
<b>6 CONCLUSÃO</b> .....	<b>40</b>
<b>REFERÊNCIAS</b> .....	<b>42</b>



## 1 INTRODUÇÃO

A empresa Renault é uma multinacional francesa do ramo automotivo. Faz parte do grupo Aliança, composto principalmente pelas outras empresas Nissan e Mitsubishi, além de outras marcas da subdivisão europeia como Dacia e Alpine. Cada uma das marcas possui independência para estabelecer a própria metodologia de gestão e estrutura de trabalho, de tal forma que especificidades existem sobre o ponto de vista estratégico para cada uma das marcas. Portanto, o foco deste estudo será somente a Renault, visto que a iniciativa de digitalização da empresa é focada somente para sua própria organização.

A organização empresarial da Renault é baseada na estrutura hierárquica, sendo o CEO o responsável que se encontra no topo da instituição. Abaixo dele, diversas diretorias são estabelecidas para responder às múltiplas atividades e exigências ligadas direta ou indiretamente à atividade produtiva da montadora. Estas ramificações são atreladas às grandes áreas empresariais, como por exemplo direções de RH, comercial e financeiro. Cada uma destas grandes áreas também possui suas subdivisões internas, que são definidas conforme a quantidade e necessidade específica de gestão organizacional de cada uma delas.

Especificamente para a grande área da engenharia, dentro dela encontramos uma camada de subdivisões de engenharia determinadas pela principal atividade desempenhada pelos engenheiros e técnicos designados para a área em questão. As diferentes subáreas da engenharia tratam de aspectos muitas vezes bem distintos, desde a concepção até a fase final de validação do produto. Dentre tantas atividades envolvidas, têm-se diferentes níveis de relação com ferramentas digitais e conseqüentemente exigências de conhecimento relativas à tecnologia da informação e ciência de dados. Entretanto, a tendência atual de transformação digital nas corporações estabelece a necessidade de uma maior disseminação da cultura de dados aos colaboradores dos mais diferentes setores da empresa, bem como a busca por mais aplicações digitais e projetos de melhoria baseados em dados.

Os profissionais de setores não especializados em TI (Tecnologia da Informação) têm sua carga de trabalho diretamente voltada para a atividade ao

qual estão designados e possuem o conhecimento técnico e prático da sua área de negócios. Por este fato, muitas vezes não têm o tempo ou fundamentos de tecnologia da informação suficientes para desenvolver projetos de digitalização, apesar de conhecerem os maiores pontos fracos e potenciais disruptivos da organização.

Por outro lado, profissionais das áreas de TI possuem todas as ferramentas e carga de trabalho necessárias para o desenvolvimento de projetos de digitalização desejados pela corporação. No entanto, tanto para a Renault como para outras empresas de grande porte, o contato e troca de informações entre as áreas de negócios e as equipes de TI são pouco desenvolvidos e complicados, muito devido à diferença dos aspectos de seus respectivos trabalhos. Isto por sua vez dificulta a parceria da área de informação com outras equipes para identificar os problemas e fornecer as ferramentas requisitadas para promover mudanças realmente impactantes.

Em função desta demanda cada vez mais proeminente de digitalização e automatização de tarefas em áreas não ligadas diretamente à tecnologia da informação que surge a proposta do grupo de trabalho Citizen Data Science. A criação deste profissional foi primeiramente sugerida pela empresa Gartner, sendo um profissional fora do contexto da TI, com conhecimento suficiente para desenvolvimento de análises de dados, que atua diretamente na área de negócios (IDOINE e TAPADINHAS, 2016). A iniciativa na Renault portanto tem por sua função ser um grupo de trabalho composto por profissionais da área de negócios e da TI com objetivo claro de impulsionar a criação de novos projetos envolvendo ferramentas digitais, bem como disseminar o conhecimento e a cultura de utilização de dados para resolução de problemas das mais diferentes áreas do conhecimento. Entretanto, por ser uma estrutura nova e complexa de gestão de projetos entre duas áreas muito distintas, não existe clareza sobre as diferentes funções esperadas para cada perfil de profissional deste grupo, bem como as etapas e a melhor forma de acompanhamento dos projetos criados e seguidos através deste grupo. Tendo isso, se apresenta como boa oportunidade a utilização dos conceitos de GI (Gestão da Informação) para direcionar os objetivos de cada profissional no âmbito Citizen e promover o melhor uso dos fluxos preconizados para maximizar o aproveitamento dos dados disponíveis na empresa.

Este trabalho tem por objetivo realizar um estudo sobre o procedimento de criação e acompanhamento de projetos de digitalização segundo a abordagem organizacional Citizen Data Science, sob a visão de processos relacionados à GI. Além disso, também será proposto um novo perfil de profissional, mesclando as habilidades encontradas no cientista de dados *citizen* e no gestor da informação para adaptar-se de maneira eficiente ao novo procedimento sugerido. O processo de criação e acompanhamento de projetos de digitalização é essencial para impulsionar a eficiência dos processos internos da empresa, além de promover a busca por conhecimento através da análise das grandes quantidades de dados gerados por estes processos. A análise deste processo deve ser realizada para compreender os gargalos e falhas atualmente presentes que podem estar impedindo as empresas de obter maior qualidade nas análises e uso de seus dados.

## 2 REVISÃO BIBLIOGRÁFICA

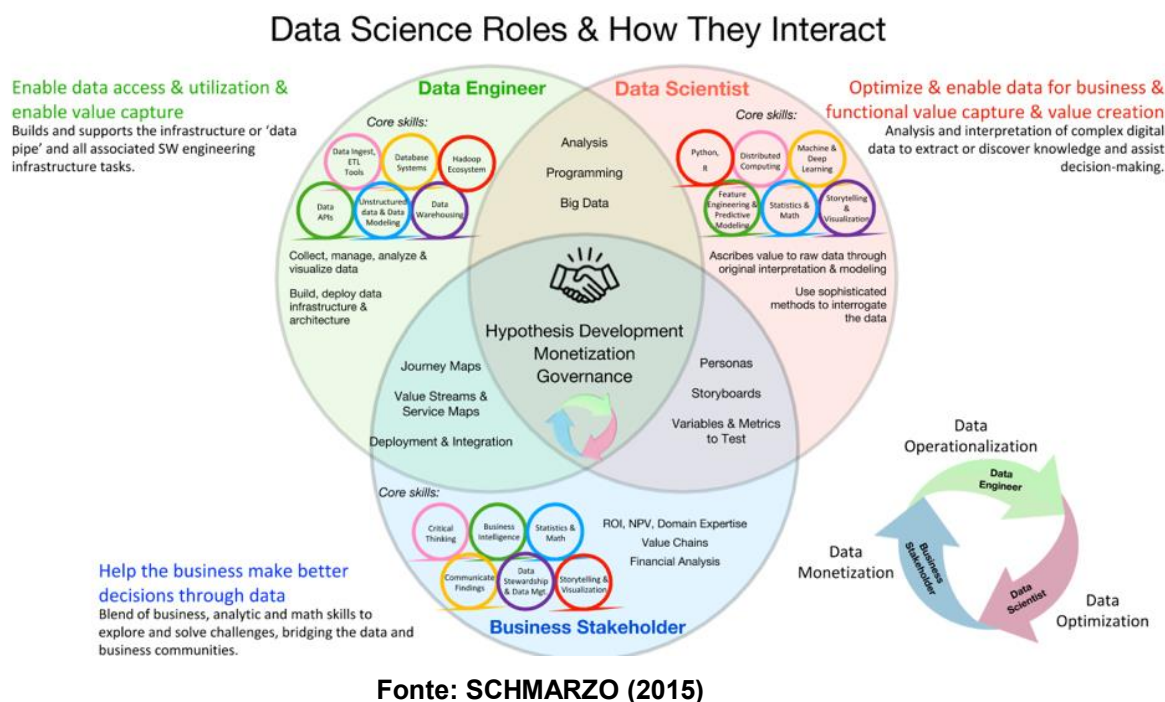
### 2.1 CIÊNCIA DE DADOS NAS DIFERENTES ÁREAS EMPRESARIAIS

#### 2.1.1 Profissionais da área de dados

Por exigência da função e por sua própria definição, os profissionais de ciência de dados possuem uma diversidade de habilidades técnicas. Eles demonstram uma abordagem diferenciada sobre o domínio em que se pretende constituir a sua análise de dados, obtida da prática e vivência no ambiente da atuação, ou de profunda busca nas teorias e modelos de interpretação deste ambiente. Além desta, os profissionais apresentam também um bom conhecimento das ciências computacionais para o desenvolvimento de algoritmos e modelos de análise, aliado ao bom conhecimento matemático para prover a base e a direção correta dentro do seu desenvolvimento (MOHANTY *et al.*, 2013).

Dentro de uma equipe de ciência de dados se encontram múltiplos profissionais com tarefas distintas e multidisciplinares, mas que trabalham em conjunto para obter um produto ou análise baseada em dados. Existem diversas formas de categorizar os papéis envolvidos na literatura, dependendo da profundidade de análise. Para os fins deste estudo, têm-se atenção para uma visão mais ampla, muito ligada às áreas do conhecimento envolvidas na ciência de dados, apresentando três funções gerais que constituem o alicerce desta área do conhecimento, que podem ser vistas na Figura 1.

Figura 1 - Interação entre diferentes áreas de trabalho na ciência de dados



Primeiramente, destacam-se os *Business Stakeholders*, que pertencem a área de negócios (AdN), onde se criam realmente os dados ligados ao objetivo principal da empresa. Dentro desta se encontram principalmente os analistas de negócios, que vivenciam os problemas do dia a dia dos processos empresariais e promovem análises mais reativas sobre estes. Destacam-se os analistas de business intelligence, que muitas vezes fazem uso de dados e ferramentas de visualização avançadas para construção de métricas.

Na sequência, temos os *Data Scientists* (DS), que pertencem a área de ciência de dados, responsável por análises mais profundas sobre a informação produzida na área de negócios. Destacam-se os cientistas de dados e os engenheiros de machine learning, que utilizam de programação e modelagem de dados para produzir introspecções diferenciadas não atingíveis através somente de visualização dos dados.

Por último, temos a área de *Data Engineering*, ou engenharia de dados, cujo objetivo é construir e manter a infraestrutura de dados para promover o acesso aos mesmos. Os principais papéis envolvidos são os engenheiros de dados e analistas de bancos de dados, que devem entregar os dados de forma acessível às outras áreas envolvidas.

### 2.1.2 Inserção da ciência de dados na organização empresarial

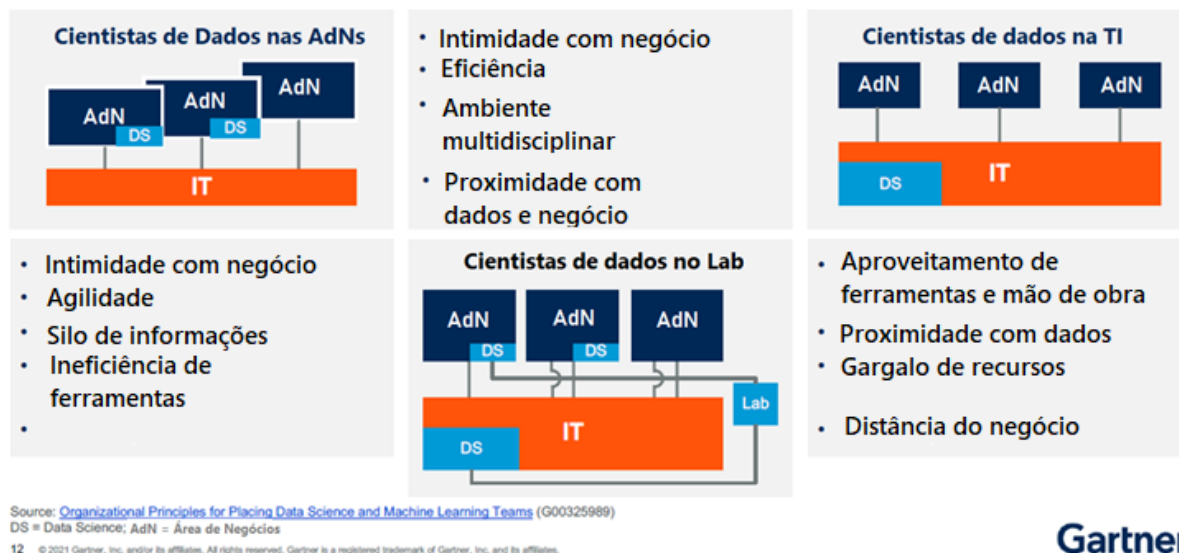
Devido a esta alta flexibilidade da ciência de dados, muitas empresas se questionam sobre o pertencimento destes profissionais dentro de suas equipes. Para obter o máximo de eficiência no seu trabalho e o melhor aproveitamento dos ativos de dados coletados, a área de dados deve ter acesso facilitado aos bancos de dados, conhecimento sobre o catálogo e localização deles, ademais do contato direto com a área de negócios para compreensão dos principais problemas e soluções tecnológicas possíveis para saná-las.

Segundo Krensky e Linden (2017), existem diferentes métodos de inserção dos profissionais de ciência de dados dentro das empresas. No primeiro método, temos estes profissionais inseridos diretamente nas áreas de negócio, facilitando principalmente a agilidade no desenvolvimento de soluções e a aproximação com os atores dos processos cruciais da empresa. No entanto, esta estratégia promove a criação dos silos de informação, dificultando o compartilhamento de iniciativas e a eficiência no uso das ferramentas disponibilizadas pela TI. Na segunda proposta, temos os cientistas de dados inseridos diretamente na área de TI. Este método privilegia o uso das ferramentas e tecnologias de informática, promovendo soluções mais eficientes e com maior embasamento técnico. Apesar disso, o distanciamento com as equipes de negócio dificulta o enfoque em problemas encontrados pela maior parte dos profissionais da empresa, além de criar um gargalo na área de TI devido à exigência de mais mão de obra técnica para dar vazão às necessidades da área de negócio.

A terceira estratégia propõe uma abordagem híbrida, com os cientistas de dados atuando em uma instância específica de criação de projetos, contendo uma parcela de atores das áreas de negócio e de TI. Esta abordagem, embora seja mais disruptiva e muitas vezes custosa devido à abertura de uma nova área dentro da empresa, é ainda mais eficaz pela proximidade entre as áreas. Assim, viabiliza-se uma maior troca de conhecimento, uma maior assertividade na solução para os reais problemas da empresa e uma facilidade na utilização das técnicas de TI. Um comparativo entre as diferentes propostas citadas pode ser visto na Figura 2.

Figura 2 - Proposta de localização de equipes de ciência de dados

## Where to Put the Data Science Team(s)?



Gartner.

Fonte: Adaptado de Krensky e Linder (2017).

## 2.2 GESTÃO DA INFORMAÇÃO

As empresas possuem o desejo de embarcar no novo contexto da indústria 4.0 e aproveitar ao máximo a oportunidade de extrair a maior quantidade de dados possíveis de seus processos e produtos. Ao mesmo tempo, é insuficiente somente ter à disposição essa grande massa de dados sem ter ao seu dispor meios com os quais usufruir destes e torná-los informações com conteúdo suficiente para a tomada de decisões.

A gestão da informação (GI) é um conceito relativamente novo que tem por função primária assegurar a qualidade e entrega de grandes quantidades de dados às pessoas necessárias para analisá-los (CONEGLIAN, GONÇALVES E SEGUNDO, 2017). É essencial a presença deste tipo de governança de dados dentro da empresa, impactando diretamente na promoção do uso correto da informação e na sua conversão em verdadeiro poder de tomada de decisões dentro da empresa.

A transformação do dado em algo de valor para a companhia é um fator determinante no processo de digitalização. Segundo Davenport e Prusak, o dado por si só não será fonte de informação necessariamente, pois é preciso

algum tipo de processamento e análise sobre ele para que se possa extrair conhecimento (DAVENPORT E PRUSAK, 1998).

São estes conceitos de dado, informação e conhecimento que pontuam Davenport e Prusak para distinguir as diferentes fases do dado e a verdadeira utilidade dele. O Quadro 1 apresenta um resumo destas distinções. O dado é simplesmente uma observação crua sobre o estado do mundo, não trazendo qualquer intuição a mais sobre ele. Por sua vez, a informação é fruto de uma intervenção do ser humano, provendo uma síntese sobre o propósito de existência deste dado. Nesta forma, a informação carrega uma análise inicial e um contexto do dado dentro do ambiente. Por último, tem-se o conhecimento, que exige um processo mais denso de avaliação da informação. O conhecimento é fruto da criação de valor através da informação e muitas vezes é capaz de promover a inteligência competitiva, fator determinante para alavancar vantagem para empresas no mercado (Costa e Silva, 1999).

**Quadro 1 - Dado, informação e conhecimento**

<b>Dados, Informação e Conhecimento</b>		
<b>Dados</b>	<b>Informação</b>	<b>Conhecimento</b>
<p>Simple observações sobre o estado do mundo</p>	<p>Dados dotados de relevância e propósito</p>	<p>Informação valiosa da mente humana Inclui reflexão, síntese, contexto</p>
<ul style="list-style-type: none"> <li>• Facilmente estruturado</li> <li>• Facilmente obtido por máquinas</li> <li>• Frequentemente quantificado</li> <li>• Facilmente transferível</li> </ul>	<ul style="list-style-type: none"> <li>• Requer unidade de análise</li> <li>• Exige consenso em relação ao significado</li> <li>• Exige necessariamente a mediação humana</li> </ul>	<ul style="list-style-type: none"> <li>• De difícil estruturação</li> <li>• De difícil captura em máquinas</li> <li>• Frequentemente tácito</li> <li>• De difícil transferência</li> </ul>

**Fonte: Davenport e Prusak(1998).**

A GI é uma atividade essencial para qualquer tipo de análise de dados a ser realizada dentro da organização. Sua existência se justifica pela necessidade de garantir a qualidade da entrega dos resultados do início até o



fim da análise, promovendo a transformação dos dados em informação e auxiliando as equipes a gerarem conhecimento. Para que isso seja possível, a GI promove o suporte nas diferentes etapas que compõem este fluxo de transformação. Bugembe estabelece seis fases distintas neste fluxo, representadas na Figura 3 (BUGEMBE, 2016).

**Figura 3 - Fases do processo de análise de dados**



**Fonte: Adaptado de Coneglian, Gonçalves e Segundo (2017).**

A fase das fontes é a primeira, pois será nesse momento que serão definidas as bases de dados utilizadas durante a análise. Por tal motivo, a escolha das fontes pode levar ao sucesso ou ao fracasso de uma análise, pois não se consegue obter algum valor de dados incoerentes com o que se deseja (CONEGLIAN, GONÇALVES e SEGUNDO, 2017). Na segunda fase, devem-se definir questões relativas à captura e ao armazenamento dos dados. Em especial para grandes empresas, uma vasta quantidade de bases de dados existe e muitas vezes os dados não são facilmente disponíveis por questões de segurança, portanto nesta etapa é crucial a presença da GI para a estratégia de armazenamento e acesso aos dados posteriormente.

A fase de processamento e fusão trata de como os dados capturados serão transformados em dados estruturados, que permitirá a reutilização e fácil uso durante as análises. No passo seguinte sobre acesso, o objetivo é disponibilizar os dados para os gestores e analistas da área de negócios, visando obter retornos quanto aos dados obtidos pelos atores que são os maiores conhecedores do negócio em questão.

Na fase de análise, são aplicados algoritmos que levarão a resultados, com objetivo de produzir análises ou novas ferramentas para gerar conhecimento. Por fim, têm-se a fase de exposição, que tem por objetivo apresentar os resultados obtidos na fase anterior aos diretores e responsáveis. Nessa fase, é importante demonstrar ganhos, pois é evidência para a validação de todas as etapas anteriores.

### 2.2.1 Gestão da informação na análise de dados

A GI funciona como agente intermediário entre a equipe de negócios e a equipe de TI, juntando as diferentes frentes para garantir a execução das atividades digitais. Os perfis GI têm um conhecimento misto de teoria computacional e as exigências de processamento de dados, além do conhecimento da área de negócios.

Através desta combinação de perfis, a GI torna-se um acelerador do fluxo de projeto. No início, é capaz de mapear e identificar ativos de dados em perímetros externos à TI, ao mesmo tempo em que conhece o catálogo de dados e o caminho para acessá-los. Na sequência, auxilia na definição de como capturar e garantir a extração dinâmica dos dados segundo a necessidade interpretada ao projeto, bem como entende das regras de negócio e consegue contribuir na formatação e uso correto dos dados disponíveis. Ainda, é capaz de julgar com embasamento a análise do ponto de vista de dados, visto que conhece os metadados que determinam a qualidade da informação obtida. Por fim, é também responsável pela apresentação do conteúdo obtido na análise, pois consegue combinar o lado de negócios e de dados envolvido no conteúdo final, além de expor com propriedade através de ferramentas de visualização.

O papel da GI na análise de dados, por sua definição, não deve se estender a realmente analisar os dados, nem a criar os algoritmos e ferramentas de análise, mas sim ser a ligação capaz de certificar a qualidade dos dados e o bom uso destes pelas duas partes envolvidas. Um comparativo entre as atividades de GI e externas com base no fluxo apresentado anteriormente pode ser visto na Tabela 1.

**Tabela 1 - Comparação entre papel da GI e TI nas empresas**

<b>Etapas de processo</b>	<b>Papel GI</b>	<b>Papel TI</b>
Fonte	Mapear e identificar ativos de informação junto aos gestores	Fornecer acesso às bases de dados, catalogar ativos, gerar dado
Captura e Armazenamento	Definir o que precisa ser adicionado ao sistema e prioridades de construção.	Criar o sistema de armazenamento e atualizações.
Processamento e Fusão	Auxiliar sobre as regras de negócio contidas nas tabelas.	Gerar script de limpeza.
Acesso	Extrair os dados e repassar para as demais áreas	Extrair e disponibilizar os dados para obter a informação
Análise	Analisar o resultado do processamento garantido a exatidão dos dados	Criar um algoritmo que atenda a necessidade do demandante.
Exposição	Auxiliar na forma com que os resultados devem ser apresentados	Gerar o relatório com os resultados de forma com que a área de negócio entenda

**Fonte: Autoria própria**

### 2.3 CIENTISTA DE DADOS CITIZEN

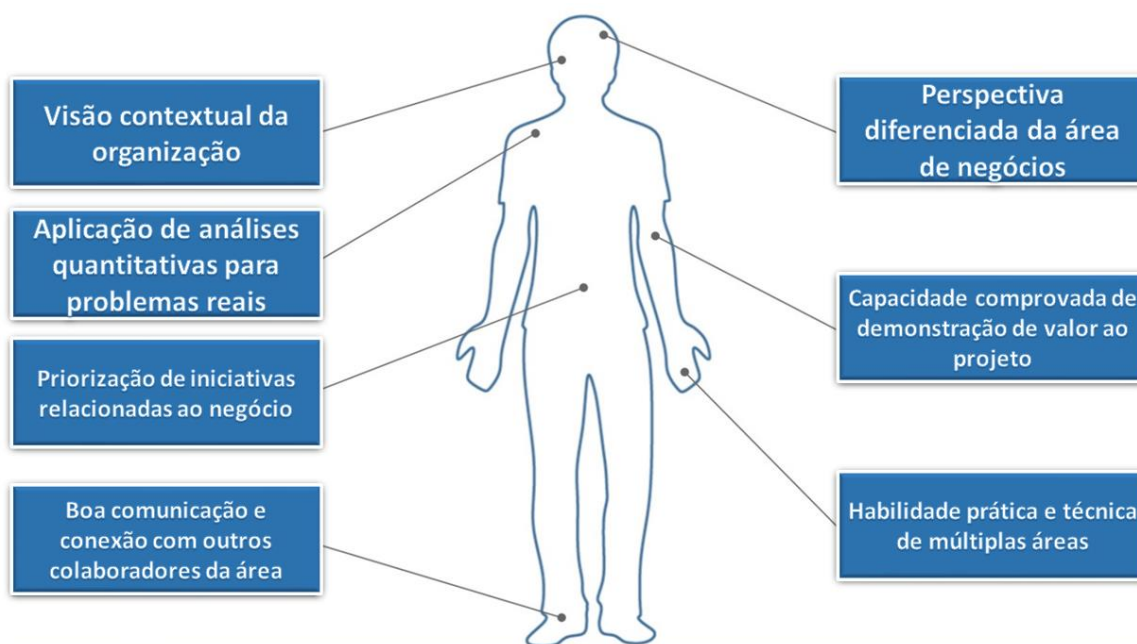
Muitas empresas possuem colaboradores com habilidades analíticas e conhecimento na análise de dados cujas funções não estão diretamente ligadas à TI. Estes funcionários estão inseridos na linha de negócios e, portanto, não possuem acesso direto às ferramentas de desenvolvimento da TI, mas poderiam fazer o uso destas caso fossem fornecidos os acessos e treinamento necessários. Ademais, com a correta disseminação de ferramentas e utilização de dados para embasamento do negócio, se cria uma cultura de digitalização e aprendizado dentro da empresa.

O cientista de dados citizen é um termo cunhado por Gartner para designar um profissional fora do contexto da TI, com conhecimento suficiente

para desenvolvimento de análises de dados, que atua diretamente na área de negócios (IDOINE e TAPADINHAS, 2016). Sua função é aderir às ferramentas preconizadas pela TI para seu benefício, protagonizando projetos internos dentro da sua própria direção e impulsionando outros colaboradores a adotarem uma cultura voltada para dados. A Figura 4 demonstra a identificação das diferentes qualidades de um cientista de dados citizen.

O citizen, apesar de ter habilidades mais superficiais nas técnicas de análise de dados, possui um vasto conhecimento prático das funções de negócio próximas a ele, demonstrando com grande facilidade o valor das iniciativas digitais dentro do contexto da organização. Torna-se mais fácil apresentar os ganhos das análises aos líderes da corporação e incentivar a criação de novas aplicações para outros colaboradores.

**Figura 4 - Habilidades do cientista de dados citizen**



**Fonte: Adaptado de Idoine e Tapadinhas (2016)**

### 3 METODOLOGIA

Inicialmente, será apresentada e analisada a iniciativa citizen nos moldes propostos pela Renault, explicando o contexto da estrutura atual e a adaptação planejada para criação deste profissional na empresa. Na sequência, serão realizadas entrevistas com 3 indivíduos de diferentes áreas de negócios, que por sua vez tenham realizado iniciativas de digitalização dentro do contexto da empresa. As entrevistas visam obter os caminhos utilizados e as etapas realizadas desde a concepção da ideia até a finalização da fase de análise. Através disso, será avaliado o percurso do projeto dentro do escopo preconizado pela GI, entendendo os pontos positivos e negativos do fluxo atual, para delinear as possíveis falhas e melhorias possíveis no processo de acompanhamento e execução de projeto de digitalização.

Em seguida, será efetuada uma análise dos principais problemas encontrados em cada etapa de GI para as 3 iniciativas avaliadas anteriormente, utilizando a ferramenta de matriz de priorização GUT (Gravidade, Urgência e Tendência). Deste modo, será possível elencar as maiores dificuldades atuais no processo, baseando-se em três fatores principais. O primeiro, a gravidade dos problemas, explicando qual o grau em que afetam a execução da etapa de GI. O segundo, a urgência, determinando qual o nível de reatividade necessário para contornar os problemas sem que o processo seja prejudicado. Por último, a tendência, que explica qual a probabilidade de acontecer os problemas novamente ou quando é previsto que os problemas elencados voltem a acontecer. Os diferentes níveis de criticidade de cada fator variam de 1 a 5 conforme a Figura 5. Cada um dos níveis será atribuído de maneira intuitiva conforme a interpretação do problema e do seu resultado em relação ao fator, sendo calculado ao final da análise o valor final GUT através da multiplicação de todos os fatores para cada etapa verificada.

**Figura 5 - Matriz de priorização GUT**



**Fonte: Justo (2019)**

Na sequência, baseando-se na priorização das atividades e dificuldades mais importantes elencadas na etapa anterior, será proposto um novo perfil de profissional, mesclando as habilidades encontradas no cientista de dados citizen e no gestor da informação. Além disso, a proposta contemplará também um fluxo otimizado de entrada e acompanhamento de iniciativas digitais para grandes empresas com base na abrangência das atividades do novo profissional e nos moldes preconizados da GI.

Por último, uma avaliação junto a atores das equipes de TI e de áreas de negócios será realizada, visualizando o levantamento de opiniões e sugestões sobre novo perfil e fluxo propostos neste estudo.

## 4 ANÁLISE

### 4.1 CENÁRIO ATUAL CIENTISTA DE DADOS CITIZEN - CASO RENAULT

A estrutura atual de uma direção do setor de negócios da Renault é voltada para a atuação nas áreas específicas às quais ela tem responsabilidade. Os profissionais que pertencem a estas áreas têm sua principal carga de trabalho voltada para a atividade principal da área, que deve ser diretamente ligada ao objetivo final da empresa de manutenção da comercialização de veículos. Entretanto, não só da atividade dita base do empregado se constrói a carga de trabalho, pois há necessidade de muitas outras atividades que auxiliam no desenvolvimento da empresa e de seus colaboradores.

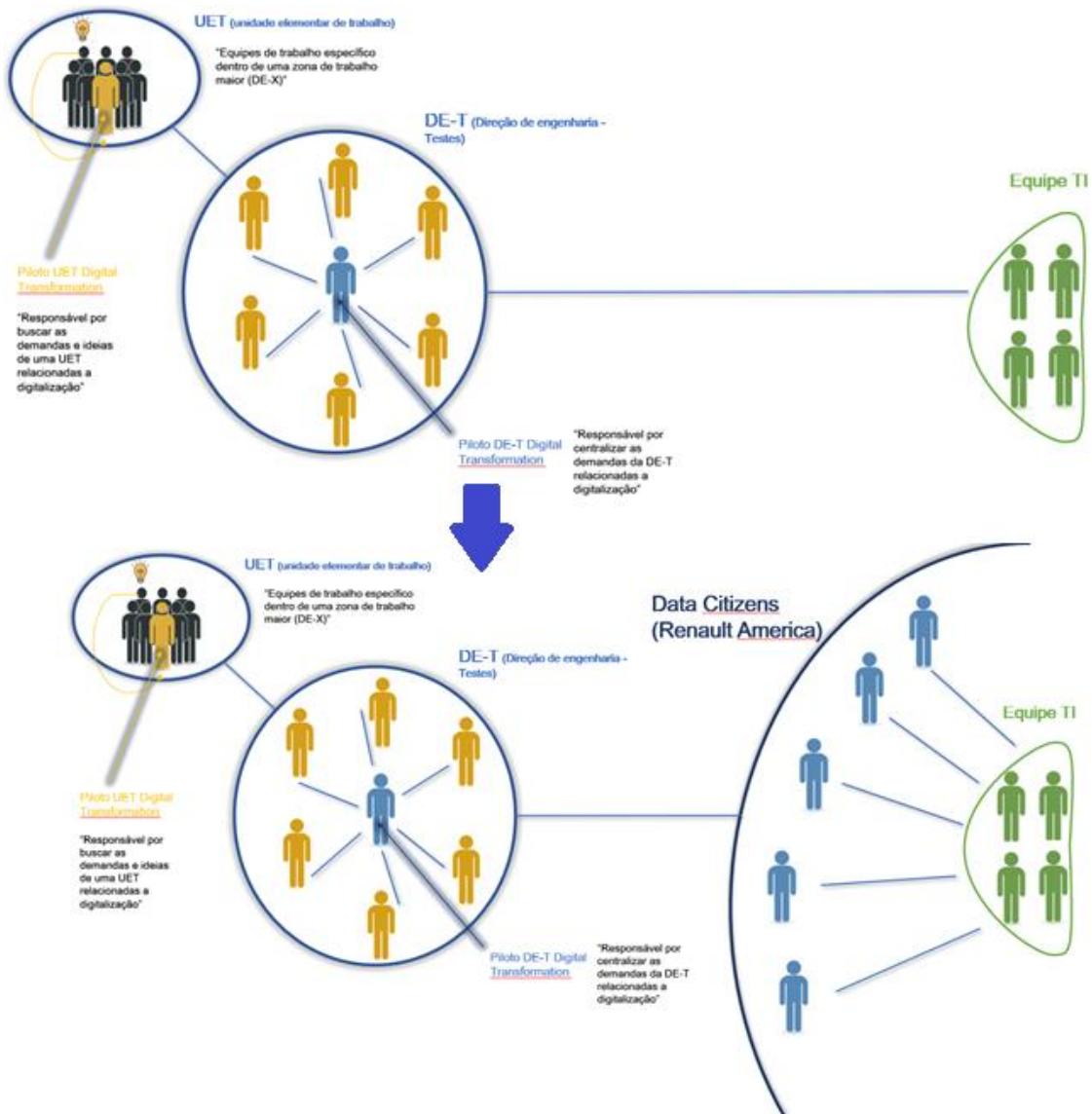
Dentre as atividades consideradas auxiliares, se encontra o incentivo à criação de projetos de inovação. A inovação é essencial para impulsionar um ambiente de melhoria contínua, promovendo parcerias, desenvolvimento de novos produtos e colocando a empresa sempre na vanguarda do mercado onde atua. A inovação está intimamente ligada à digitalização, em especial no atual contexto de transformação da indústria movida pela quarta revolução industrial, bem como a transformação do trabalho, devido à adoção de métodos remotos e híbridos de atuação. Os pilotos de inovação têm a função de promover a criação de projetos de inovação e digitalização dentro de suas direções, os quais muitas vezes fazem parte de objetivos anuais das áreas de negócios determinados por gestores e diretores. O piloto de inovação da direção cria iniciativas para impulsionar a criação destes projetos, além de garantir o suporte por parte de diferentes instâncias necessárias à sua evolução.

A organização atual se encontra em evolução para adaptar-se à proposta de criação do perfil cientista de dados *citizen*. Com isso, o piloto de digitalização da direção se tornaria também um *citizen*, aproximando-se da equipe de TI para tornar ainda mais fácil e dinâmica a criação e acompanhamento de iniciativas de digitalização. O *citizen* deve ser capaz de entender as regras do negócio e os dados envolvidos, bem como também

conhecer os atores corretos de cada área para auxiliar o responsável da ideia no desenvolvimento do projeto.

A Figura 6 apresenta a modificação proposta na estrutura de trabalho atual para o surgimento do perfil de cientista de dados *citizen* na Renault.

**Figura 6 - Evolução da estrutura organizacional interna de pilotos de digitalização**



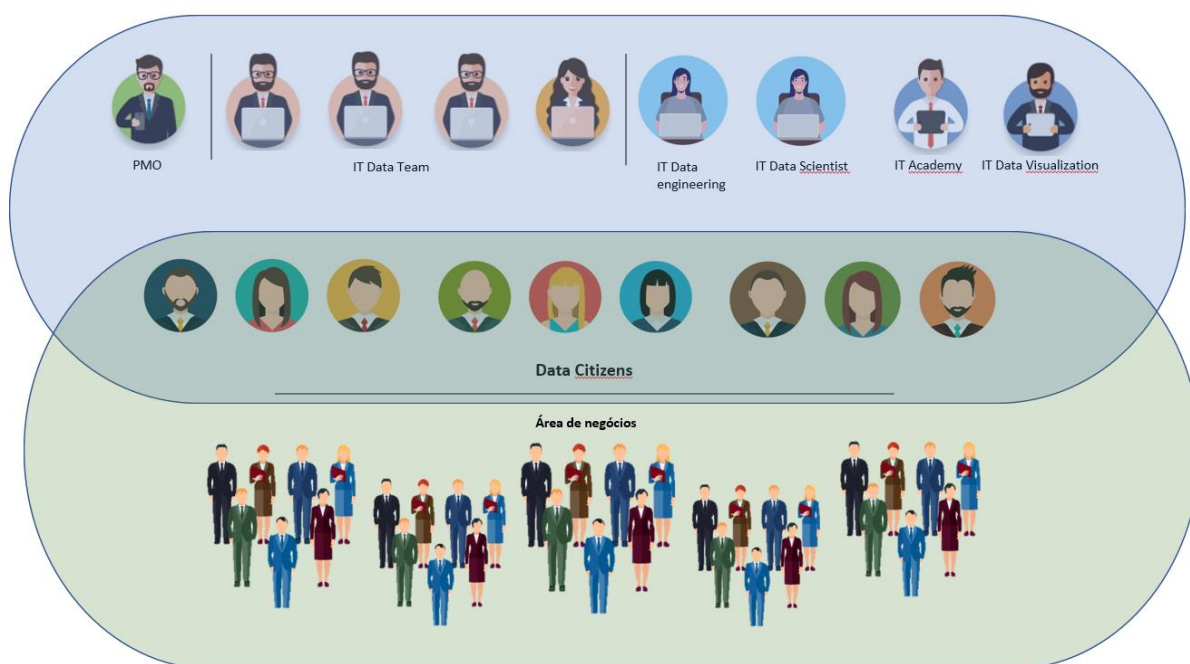
Fonte: Autoria própria

O perfil *citizen* nos moldes atuais desenhados pela companhia não são profissionais dedicados à atividade de ciência de dados, assim como os pilotos de digitalização. Deve-se ressaltar que é uma tarefa adicional para colaboradores que têm interesse em dados, mas não há um incentivo direto em se tornar um *citizen*. De toda forma, ele é vital para o funcionamento da



estratégia de digitalização da empresa, se tornando o elo entre a área de TI e a área de negócios. A Figura 7 apresenta as diferentes funções relacionadas a dados de acordo com a nova estratégia, destacando a atividade *citizen* como mencionada anteriormente.

**Figura 7 - Evolução da estrutura organizacional interna para criação do data science citizen**



**Fonte: Autoria própria**

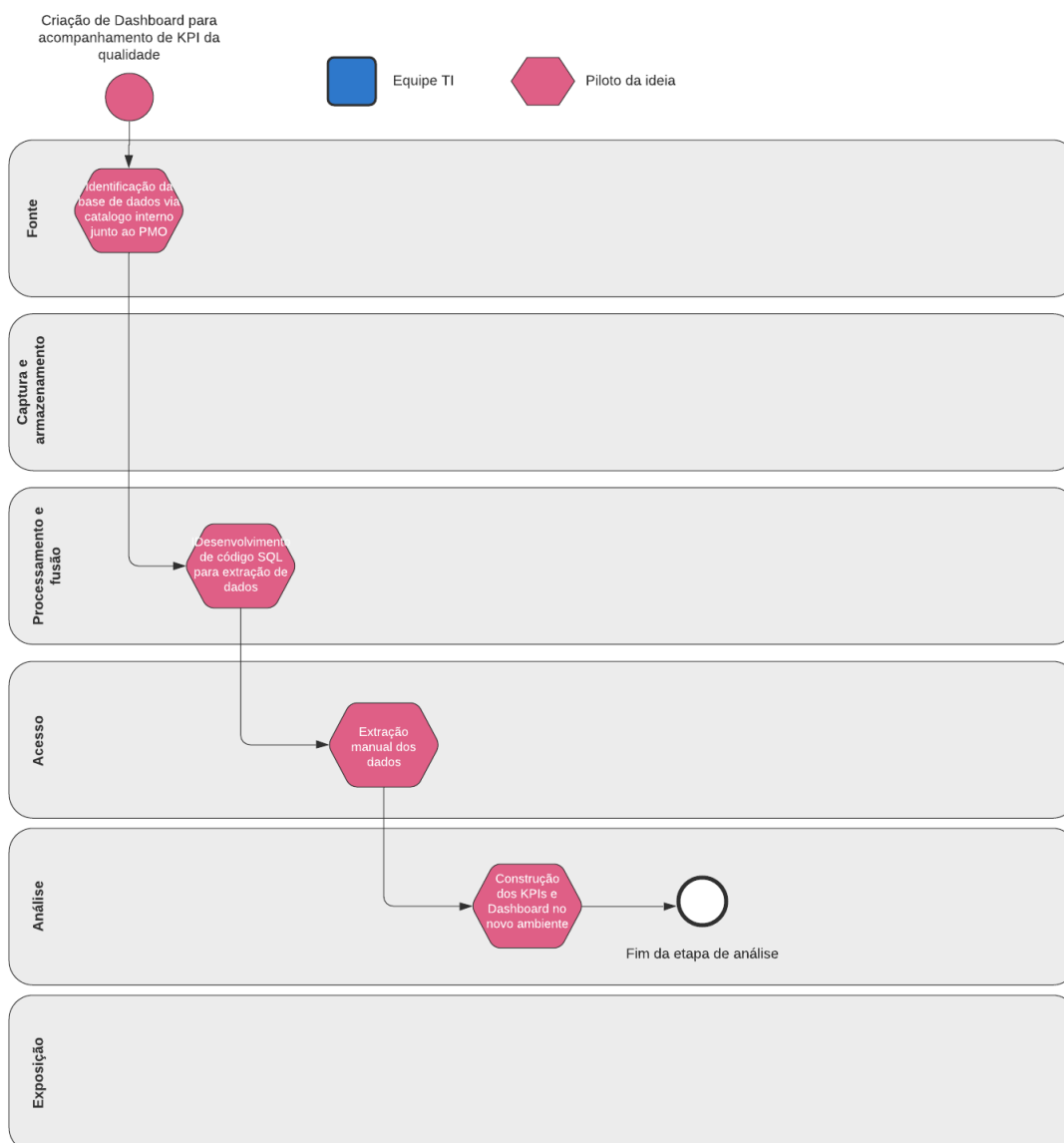
Apesar da criação do cientista de dados *citizen*, suas atividades dentro do fluxo de criação de projetos de digitalização não são claras. Para se aproveitar ao máximo a proposta, deve-se tomar como base as definições estabelecidas na gestão da informação e os conceitos definidos por Gartner para concretizar o verdadeiro propósito deste perfil dentro da empresa.

## 4.2 ESTUDOS DE CASO

### 4.2.1 Estudo de caso – Projeto 1

O primeiro projeto analisado tem suas etapas apresentadas na Figura 8. Trata-se de uma iniciativa de criação de *Dashboard* para o acompanhamento de indicadores de performance na área de qualidade do cliente da empresa.

**Figura 8 - Fluxo de projetos de digitalização - Caso 1**



**Fonte: Autoria própria**

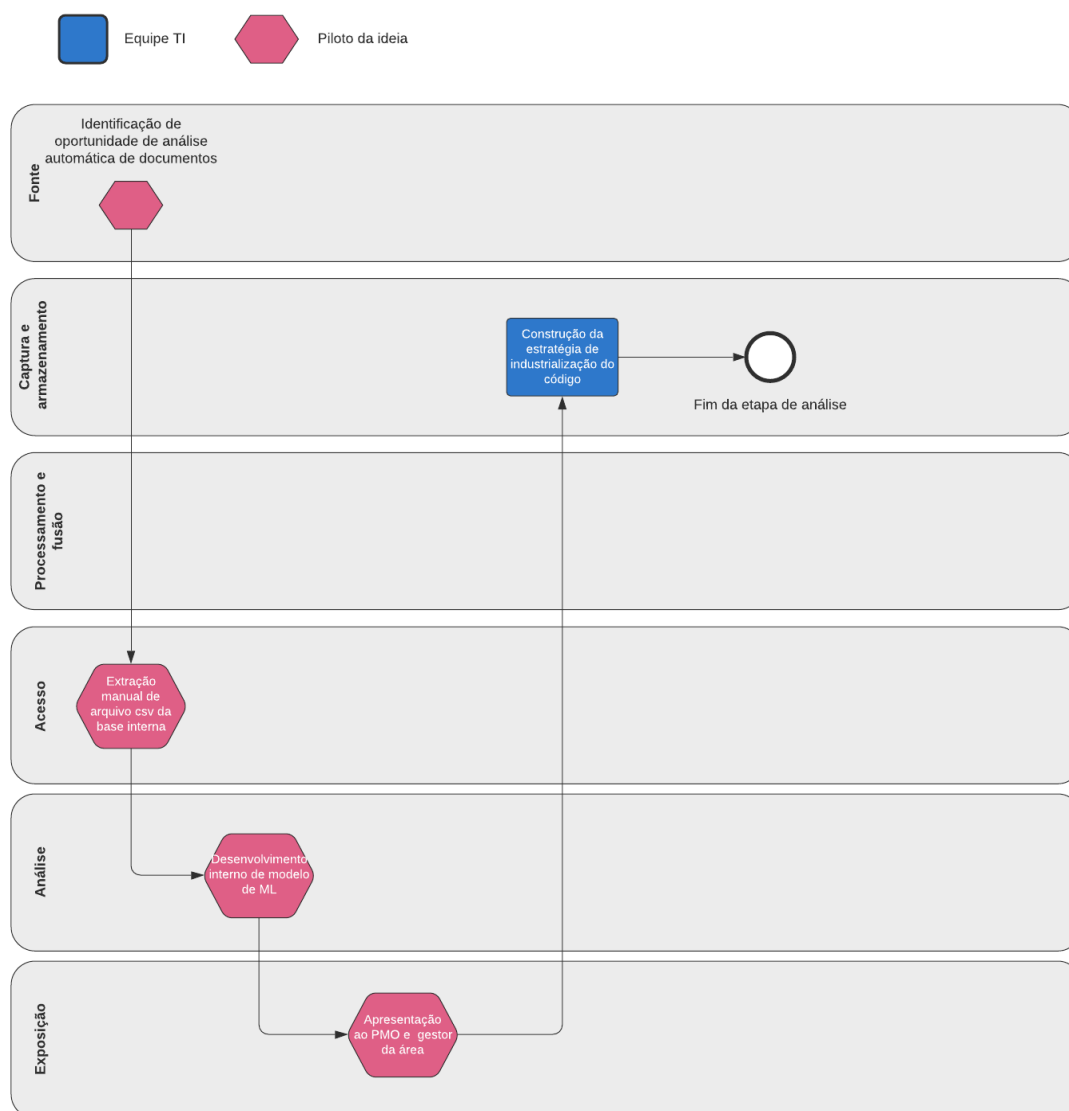
Destaca-se primeiramente a definição do projeto, que buscava a evolução de uma ferramenta de visualização anterior a uma nova devido à mudança da ferramenta oficial da companhia. O piloto da ideia já possuía um formato estabelecido, bastando obter o suporte da TI para transformação para a nova plataforma com maiores possibilidades de conexão de base dados. A ideia foi levada diretamente ao PMO (*Project Management Office*), que na sequência designou a iniciativa ao ponto focal de engenharia de dados da

empresa. Devido à criticidade do projeto, havia necessidade de agilidade no desenvolvimento. Se propôs então a criação de um código de extração SQL, que foi realizado pelo próprio dono da ideia, para na sequência ser utilizado para obtenção dos dados manualmente. Entretanto, na etapa de construção do sistema de armazenamento, depende-se do engenheiro de dados, cujo *know-how* atual do processo na nova plataforma é de sua responsabilidade. Nesta etapa, cria-se um gargalo no fluxo visto que a orientação e treinamento para realizar o trabalho vem somente de um profissional, que possui uma alta carga de trabalho. O piloto, portanto, fez uma extração manual dos dados e realizou os cálculos de KPI (*Key Performance Indicator*) e design do *Dashboard* sobre este rascunho. Ao final do projeto, não se teve o ganho esperado de conexão com a base de dados oficial, optando-se pela estratégia de atualização manual dos dados de visualização e criando um silo de dados na área de negócios.

#### 4.2.2 Estudo de caso – Projeto 2

O primeiro projeto analisado tem suas etapas apresentadas na Figura 9. Trata-se de uma iniciativa de criação de *Dashboard* para o acompanhamento de indicadores de performance na área de qualidade do cliente da empresa.

**Figura 9 - Fluxo de projetos de digitalização - Caso 2**



**Fonte: Autoria própria**

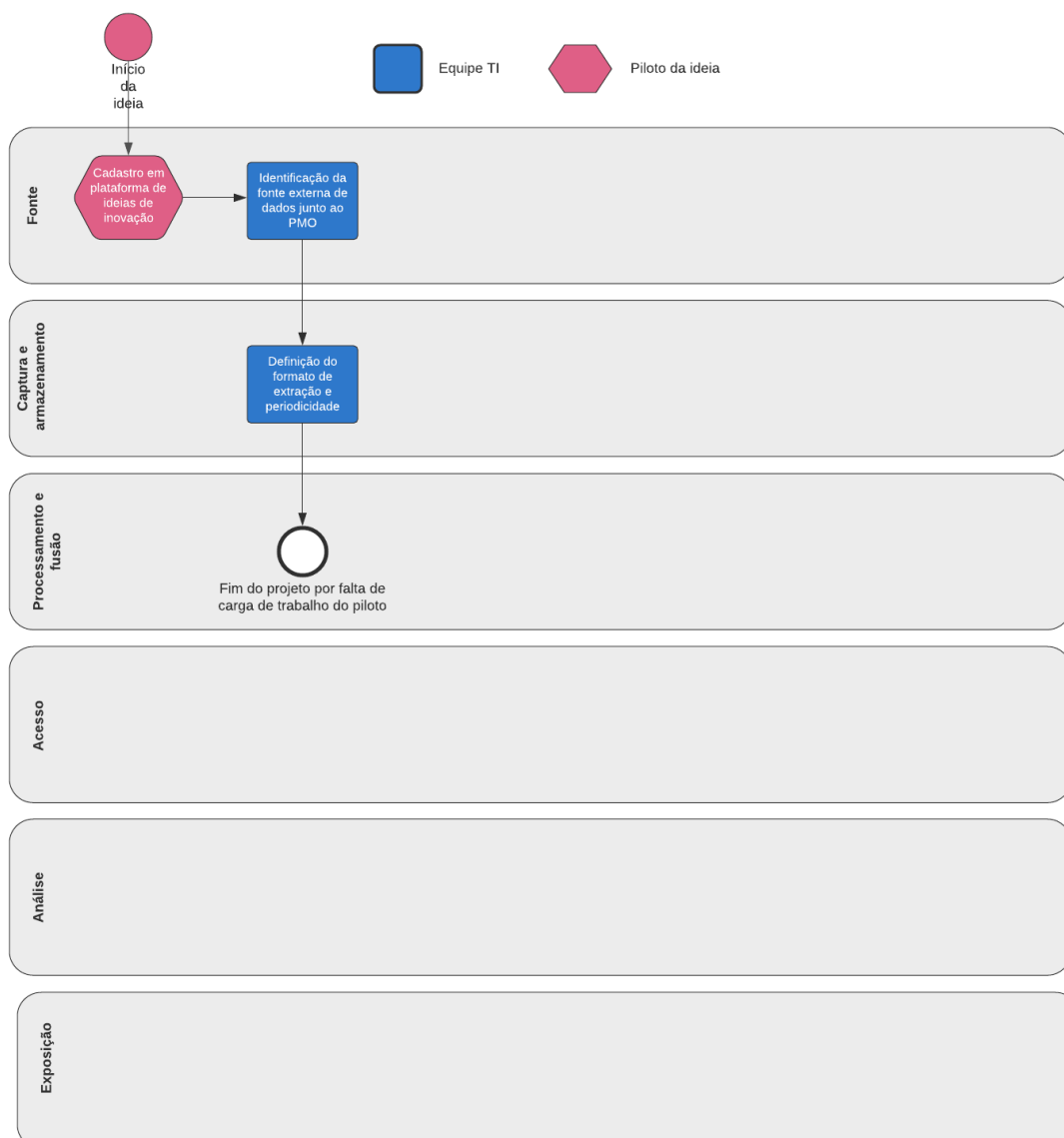
Este projeto foi considerado pioneiro no sentido de desenvolvimento, pois o piloto não estabeleceu contato inicial com a TI para suporte como preconiza o fluxo atual. O próprio piloto fez a extração manual da base de dados a qual tinha acesso na época, excluindo as etapas iniciais de planejamento da extração e verificação dos catálogos de dados. Na sequência, o tratamento de limpeza e produção da análise também foi realizado na própria máquina, contrariando o procedimento correto de instalação de máquina na plataforma oficial da empresa. No momento em que foi apresentado ao PMO, a ideia estava concretizada e a análise, em um estado muito avançado. Deste

modo, foi preciso transferir o código produzido para a plataforma oficial, além de validar as etapas iniciais de acesso, processamento e estratégia de industrialização do modelo após a apresentação dos ganhos com a utilização do modelo. Além disso, a etapa de estratégia de armazenamento não realizada anteriormente teve que ser retomada após o fim das análises para que fosse determinado o método de industrialização do código e desenho da arquitetura, que por sua vez atrasou mais a finalização da análise. Deve-se citar que à época, a metodologia para instâncias de modelagem de aprendizado de máquina ainda era muito embrionária, e este projeto foi um catalisador para o surgimento dos data citizens e dos processos atuais de fluxo de ideias de digitalização.

#### 4.2.3 Estudo de caso – Projeto 3

O terceiro projeto analisado tem suas etapas apresentadas na Figura 10. Trata-se de uma iniciativa de criação de modelo de *Machine Learning* aplicado à classificação de queixas do cliente para análise interna da equipe de incidentologia.

**Figura 10 - Fluxo de projetos de digitalização - Caso 3**



Este projeto seguiu as etapas iniciais como previsto pelo fluxo, partindo de um cadastro na plataforma de ideias de inovação da companhia, que na sequência deve ser direcionada à instância de tratamento de iniciativas de digitalização. O piloto recebeu suporte do PMO no primeiro contato através da identificação do problema e da base de dados a ser utilizada para o desenvolvimento de um modelo de aprendizado de máquina, incluindo os ativos de dados presentes e a forma como tratá-los. Na sequência, houve também a definição do método de extração dos dados, que vem de uma base externa de um fornecedor, estabelecendo-se o formato e periodicidade do

recebimento dos dados. Entretanto, na etapa de criação do código de limpeza e extração, o projeto foi abandonado, visto que a responsabilidade desta etapa também recai sobre o piloto da ideia, que alegou não haver tempo e carga de trabalho suficiente para a continuação do projeto. Isto demonstra que mesmo seguindo as preconizações corretas conforme a GI espera, há mais fatores envolvidos que podem impedir a finalização de um projeto de digitalização.

#### 4.3 MATRIZ DE PRIORIZAÇÃO DE PROBLEMAS ATUAIS

Os problemas levantados na análise do histórico de projetos anteriores foram utilizados como base para a construção da matriz de priorização de problemas apresentada na Tabela 2. O foco foi estabelecido sobre as etapas definidas pela GI para a eleição das principais dificuldades atuais, cuja importância recai sobre a definição dos pontos a serem atacados com a criação do perfil *citizen*.

Tabela 2 – Matriz de priorização GUT

Etapa	Gravidade	Urgência	Tendência	GUT
Fonte	Grave	Pouco urgente	Piorar no longo prazo	12
	3 Silos de informação	2 Prática atual não prejudica o funcionamento	2 Maior inutilização de grande massa de dados gerada pela empresa	
Captura e armazenamento	Grave	Urgente	Piorar no médio prazo	27
	3 Ausência de estratégia inicial de disponibilização dos dados, falta de mão de obra na engenharia de dados para industrialização	3 Projetos que chegam à industrialização demoram para evoluir	3 Maior carga de projetos de ciência de dados no futuro impactará na capacidade de industrialização	
Processamento e Fusão	Grave	Pouco urgente	Piorar no médio prazo	18
	3 Falta de tempo do cientista de dados para criação de códigos de extração e limpeza de dados, falta de conhecimento dos pilotos para realização da atividade	2 Mais demandas gerarão mais necessidade de limpeza de dados para as análises, gerando gargalo no fluxo, necessidade de conhecimento de limpeza por parte dos pilotos	3 Maior dificuldade no avanço dos projetos	
Acesso	Grave	Urgente	Piorar no curto prazo	36
	3 Falta de conhecimento do catálogo de dados, dificuldade na extração dos dados em bases na nuvem por parte dos pilotos	3 Dificuldade na conexão com dados produz análises focadas internamente, gerando silos de dados	4 Muitas análises dependendo da disposição dos dados são perdidas	
Análise	Muito grave	Muito urgente	Piorar no médio prazo	48
	4 Falta de carga de trabalho voltada para iniciativas digitais nas áreas de negócios	4 Projetos direcionados à fornecedores externos ou abandonados por dificuldade no acompanhamento	3 Projetos direcionados à fornecedores externos ou abandonados por dificuldade no acompanhamento	
Exposição	Grave	Pouco urgente	Sem tendência de piorar	6
	3 Falta de instância de divulgação dos trabalhos desenvolvidos em mais direções, falta de aproveitamento de trabalhos anteriores	2 Desconhecimento de outras equipes dos projetos internos e soluções criadas	1	

Fonte: Autoria própria

O principal problema enfrentado atualmente é a ausência de carga de trabalho voltada para a criação e acompanhamento das iniciativas de digitalização. A definição da responsabilidade sobre a condução do projeto pelo piloto da ideia acaba gerando uma sobrecarga nele, visto que sua principal função não é atrelada à realização dos projetos digitais. Por mais que a capacidade analítica do piloto seja alta, muitas vezes não há tempo suficiente para o desenvolvimento interno do projeto, recaindo sobre propostas de fornecedores externos caso haja orçamento suficiente, ou o completo abandono do projeto.



Outro problema levantado é a dificuldade no acesso aos dados por parte dos pilotos, que não possuem conhecimento suficiente sobre o catálogo de dados produzido pela TI e não conhecem os procedimentos existentes para conexão e extração dos dados presentes em bancos de dados na nuvem. Logo, estes optam por manter as análises e os códigos criados em ambiente interno, criando os silos de dados e a dificuldade no aproveitamento de soluções por outras equipes.

Por fim, o terceiro grande problema está ligado à etapa de captura e armazenamento, desta vez por parte da TI para a industrialização das iniciativas que chegam ao fim do fluxo. Como preconizado pela GI, é essencial planejar o sistema de extração e tratamento de dados antes de iniciar as análises em si, algo que não é previsto ainda no formato atual de acompanhamento dos projetos. Por isso, o processo se torna um gargalo à medida que as iniciativas chegam à industrialização, pois o sistema ainda não é preparado para inserção do código e a conexão direta com o banco de dados.

## 5 PROPOSTA

### 5.1 NOVO PERFIL CITIZEN E FLUXO DE INICIATIVAS DE DIGITALIZAÇÃO

Para atender à grande demanda de iniciativas de digitalização dentro de grandes empresas, promovendo a conexão fundamental entre a área de negócios e TI, se propõe a criação de um perfil *citizen* dentro da empresa, composto de um misto entre as funções de TI e GI. As habilidades e funções deste perfil dentro das etapas de GI podem ser vistas no Quadro 2.

**Quadro 2 - Proposta de novo perfil data *citizen***

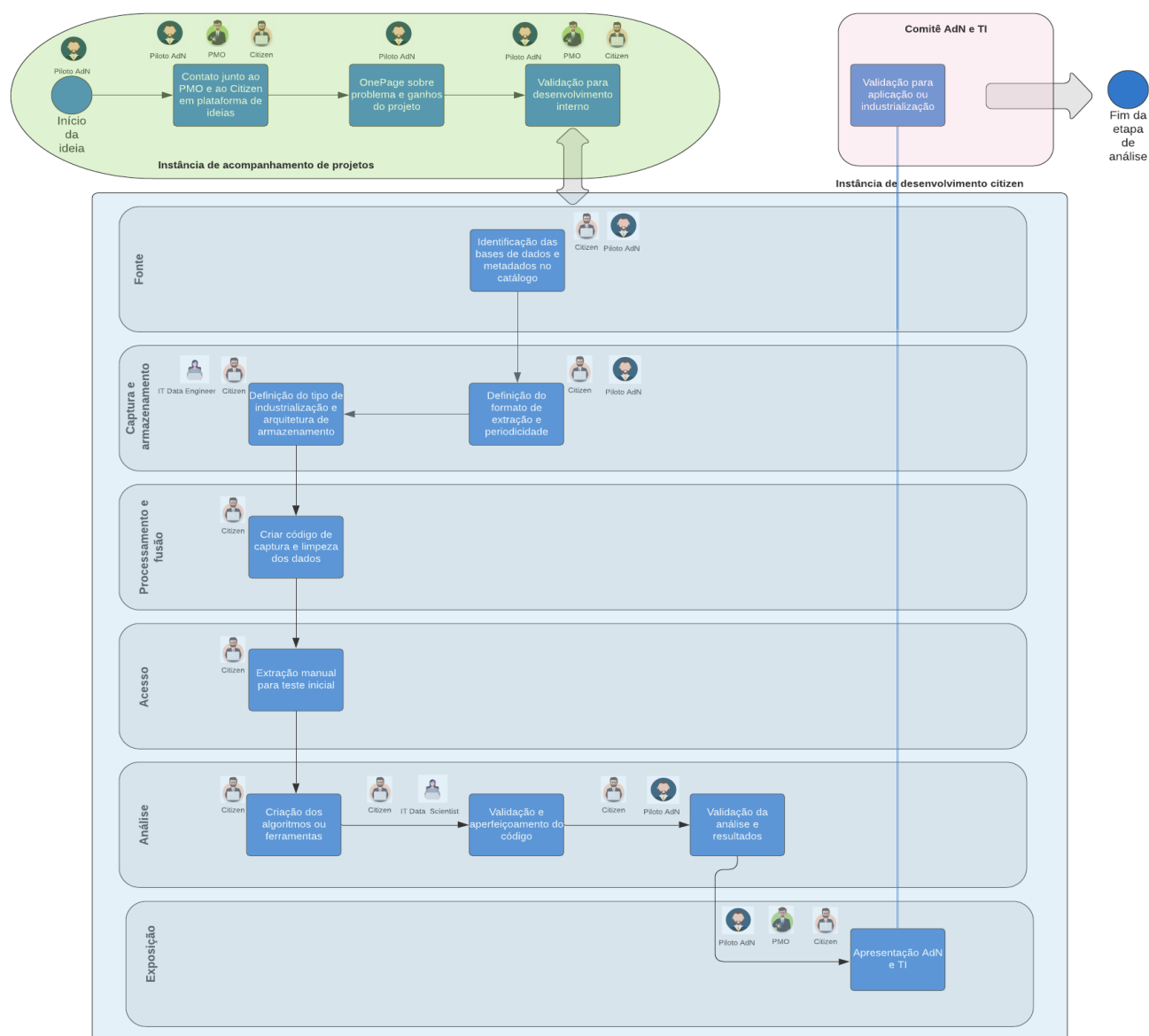
Etapa	Ações	Responsabilidade		
		Área de negócios	Citizen	Área de TI
Fonte	Mapear e identificar ativos de informação junto aos gestores	O	X	
	Fornecer acesso às bases de dados, catalogar ativos, gerar dados			X
Captura e Armazenamento	Definir o que precisa ser adicionado ao sistema e prioridades de construção.	O	X	
	Criar o sistema de armazenamento e atualizações.			X
Processamento e Fusão	Auxiliar sobre as regras de negócio contidas nas tabelas.	X	O	
	Gerar script de limpeza.		X	
Acesso	Extrair os dados e repassar para as demais áreas	O	X	
	Extrair e disponibilizar os dados para obter a informação		X	O
Análise	Analisar o resultado do processamento garantindo a exatidão dos dados	O	X	
	Criar um algoritmo que atenda a necessidade do demandante.		X	O
Exposição	Auxiliar na forma com que os resultados devem ser apresentados	X	O	
	Gerar o relatório com os resultados de forma com que a área de negócio entenda		X	

**Fonte: Autoria própria**

O novo perfil *citizen* tem sua carga de trabalho totalmente voltada para a atividade de acompanhamento e desenvolvimento de projetos de digitalização junto a área de negócios, atuando, portanto, como suporte tanto para a área de negócios quanto para a TI. Para que este consiga impulsionar as iniciativas internas de cada área da empresa, deve pertencer e ser inserido diretamente nos setores externos à TI, promovendo as soluções ao lado dos

pilotos das ideias, ao mesmo tempo em que é capaz de desenvolver os algoritmos e modelos necessários sem necessidade de intervenção da TI. Um novo fluxo para iniciativas de digitalização também é proposto e pode ser visto na Figura 11.

**Figura 11 - Fluxo proposto de projetos de digitalização para grandes empresas**



**Fonte: Autoria própria**

Dentre as melhorias propostas através deste novo fluxo, podemos citar primeiramente a definição concreta da instância de acompanhamento de projetos de digitalização. Esta é uma esfera da companhia já existente, mas

que pode ser aperfeiçoada através da delimitação da sua abordagem, limitando-a a selecionar os projetos mais interessantes e com maior proposta de valor e realizando o planejamento e acompanhamento do avanço dos projetos. Dentro disso, com o novo fluxo, é possível visualizar claramente quais etapas são esperadas dentro da instância de desenvolvimento, o que por sua vez facilita o seguimento do avanço, além de promover uma responsabilidade nítida dentro de cada fase do projeto. Esta responsabilidade estará atrelada a cada perfil delineado na ciência de dados, mas restando nitidamente uma função de desenvolvimento técnico dos projetos para o novo perfil *citizen*.

A etapa de captura e armazenamento, que no passado foi esquecida ou deixada para o final do projeto, agora tem sua execução na fase preconizada pela GI, facilitando a estruturação da industrialização final antes do fim do projeto, com auxílio da engenharia de dados somente caso necessário. A etapa de acesso também será facilitada com o uso deste fluxo, pois o *citizen* possui conhecimento vasto do catálogo de dados e as conexões com os colaboradores de TI e *data owners* capazes de entregar uma extração dos dados para início dos trabalhos. Ademais, a fase de análise, antes dispensada por falta de carga ou esquecida por falta de seguimento, tem sua responsabilidade direta ao *citizen*, que responde para a instância de acompanhamento de projetos sobre a avanço do projeto dentro desta etapa, compartilhando os resultados com a área de negócios e o cientista de dados especialista caso necessário.

A exposição dos resultados deve chegar aos diretores da área de negócios e da TI para validação dos ganhos finais e a arquitetura da ferramenta ou modelo desenvolvido. Após esta etapa, se tem o fim da análise e o início da aplicação ou adaptação para funcionamento em ambiente produtivo.

## 5.2 AVALIAÇÃO EXTERNA SOBRE A PROPOSTA

A proposta de novo perfil de *citizen* e novo fluxo de acompanhamento de iniciativas de digitalização foi apresentada para a equipe de dados da TI da Renault, com intuito de obter sugestões e principalmente críticas em relação às mudanças propostas.

Primeiramente, o perfil *citizen* desenvolvido foi bem aceito pela equipe, que indicou que o delineamento das atividades e a abordagem criada se relacionam bem com o esperado por eles. Ademais, é interessante também que o *citizen* tenha no seu âmbito de trabalho a carga de trabalho voltada para criação de algoritmos e ferramentas dentro das iniciativas de digitalização, demonstrando o intuito da equipe de realmente promover a criação de um novo profissional na empresa. Entretanto, foi considerado que a alocação de toda a carga de trabalho para atuação nas iniciativas não é plausível e não deve ser a melhor forma de direcionar a função do *citizen*, realizando algum tipo de divisão entre colaboração dentro da área de negócios e na instância *citizen*, assim garantindo o conhecimento técnico da área e o contato direto com os problemas atuais.

Em relação ao fluxo, a proposta foi bem recebida pela equipe, que avaliou como uma grande oportunidade para definir os procedimentos e etapas esperadas nesta atividade prática de acompanhamento de iniciativas digitais. Em especial, a etapa de captura e armazenamento foi muito bem recebida, pois verificou-se que mais casos podem ter sofrido do mesmo problema de falta de planejamento da estrutura de armazenamento e aplicação industrial antes da etapa de análise em si. A promoção do *citizen* como grande responsável no ciclo de GI e desenvolvimento técnico do projeto também foi bem aceita, com ressalva de que o mesmo deverá ser bem treinado e terá que ter um aprofundamento das técnicas de TI antes de iniciar sua trajetória dentro da área de dados.

Foi pontuado um obstáculo dentro do fluxo disposto que poderia ocasionar dificuldades no avanço das iniciativas, principalmente voltado para a instância de acompanhamento dos projetos. Um grande entrave visto pela equipe seria logo ao início do fluxo, em específico sobre a forma como as ideias devem chegar dos pilotos até o conhecimento do PMO e do *citizen*. Para esta proposta, não há uma definição clara de como será realizado isso, restando a forma de disseminação da existência da instância através de um comunicado aos colaboradores, muitas vezes desconhecido por muitos em algumas áreas. É preciso uma boa estratégia de sensibilização das áreas de negócio e uma definição formal da forma de entrada das ideias dentro do fluxo

para garantir que nenhuma iniciativa seja ignorada ou não chegue às pessoas importantes para que se inicie o projeto.

A equipe também realizou uma remarca sobre a outra faceta do perfil *citizen* relacionada à disseminação da cultura de dados dentro da empresa. Segundo eles, o *citizen* também atua como incentivador da busca do conhecimento das ferramentas preconizadas pela TI e do uso maior dos dados disponibilizados para todos os colaboradores das áreas de negócios. Os *citizens* participam de fóruns de troca de informações e soluções junto a TI, trazendo as novidades e treinamentos para os colaboradores das áreas de negócios. Entretanto, esta atividade do *citizen* é muito complexa de se definir através de procedimentos e menciona-se que o escopo deste projeto é somente para o tratamento do fluxo de iniciativas de digitalização.

## 6 CONCLUSÃO

O estudo deste trabalho consistiu em realizar uma análise sobre o fluxo de desenvolvimento de projetos de digitalização em áreas de negócios especificamente para a Renault e a inclusão do novo profissional cientista de dados *Citizen* para impulsionar tais projetos em grandes empresas, baseando-se nas estruturas da GI.

Através dos resultados dos estudos de caso de projetos anteriores, foi possível avaliar gargalos e falhas no processo utilizado atualmente pela empresa, determinando pontos críticos onde a preconização de etapas de GI poderia atuar para desbloquear ou melhorar o desempenho dos projetos. Além disso, a análise de priorização dos problemas encontrados também proporcionou uma melhor concepção sobre onde a inserção dos *citizens* pode ter o maior benefício, além das reais habilidades e atividades que deveriam ser desempenhadas por eles para obter a máxima eficiência no acompanhamento e desenvolvimento dos projetos de digitalização de áreas de negócio.

A proposta realizada neste trabalho para o novo perfil de cientista de dados *citizen* e o novo fluxo de desenvolvimento de projetos de digitalização com base nos conceitos de GI e nos pontos fracos do fluxo anterior foi bem aceita pela equipe de dados da Renault. As determinações sobre o alcance das atividades de cada ator do processo ficaram bem definidas de acordo com o conhecimento técnico e prático de cada um, além das novas características do perfil *citizen* que mesclam profundamente as noções das áreas de negócio e TI.

É preciso dizer que a implementação de tal profissional dentro do contexto corporativo passa por um profundo desafio de modificação das estruturas de trabalho da empresa e das estruturas de TI para adaptação ao cenário de disponibilização das grandes massas de dados criadas no negócio. Ademais, ainda é necessário avaliar com mais robustez o papel do *citizen* como disseminador da cultura de uso de dados e ferramentas de TI dentro da organização, um dos princípios que poderiam ser avaliados através das proposições da gestão do conhecimento, mas que ultrapassa os objetivos deste trabalho.





## REFERÊNCIAS

BUGEMBE, M. Finding Value in Data: Determining Where Data Science has The Greatest Impact. O'Reilly: Sebastopol, 2016.

CONEGLIAN, C. S.; GONÇALVEZ, P. R. V. A.; SEGUNDO, J. E. S. O profissional da informação na era do big data. Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, v. 22, n. 50, p. 128-143, 2017.

DAVENPORT, T.; PRUSAK, L. Learn how valuable knowledge is acquired, created, bought and bartered. The Australian Library Journal, v. 47, n. 3, p. 268-272, 1998.

JUSTO, A. S. Matriz GUT: entenda o que é e como aplicá-la na priorização dos seus projetos. Euax Consulting, 2 de abril de 2019. Disponível em: <https://www.euax.com.br/2019/04/matriz-gut/>. Acesso em: 10 de junho de 2021.

LINDEN, A.; KRENSKY, P. Organizational Principles for Placing Data Science and Machine Learning Teams. Gartner Research, 18 de setembro de 2017.

MOHANTY, S; JAGADEESH, M.; SRIVATSA, H; 2013. Big Data Imperative: Enterprise Big Data Warehouse, BI Implementations and Analytics, Apress.

PALMER, S. Data Science for the C-Suite. Nova York, Digital Living Press, 2015.

SCHMARZO, B. A Winning Game Plan For Building Your Data Science Team. Data Science Central, 31 de agosto de 2018. Disponível em: <https://www.kdnuggets.com/2018/09/winning-game-plan-building-data-science-team.html>. Acesso em 10 de junho de 2021.

TAPADINHAS, J.; IDOINE, C. Pursue Citizen Data Science to Expand Analytics Use Cases. Gartner Research, 09 de dezembro de 2016.