

BRUNO MERCADANTE DELBUONO DOS ANJOS

ANÁLISE DE ARQUITETURA DA INFORMAÇÃO
PARA ESTRATÉGIA DE CRESCIMENTO EM FINTECHS
UTILIZANDO CUSTOMER DATA PLATFORM (CDP)

Monografia apresentada ao Programa de Educação Continuada da Escola Politécnica da Universidade de São Paulo, para obtenção do título de Especialista, pelo Programa de Pós-Graduação em Big Data - Inteligência na Gestão dos Dados.

SÃO PAULO

2022

BRUNO MERCADANTE DELBUONO DOS ANJOS

ANÁLISE DE ARQUITETURA DA INFORMAÇÃO
PARA ESTRATÉGIA DE CRESCIMENTO EM FINTECHS
UTILIZANDO CUSTOMER DATA PLATFORM (CDP)

Monografia apresentada ao Programa de Educação Continuada da Escola Politécnica da Universidade de São Paulo, para obtenção do título de Especialista, pelo Programa de Pós-Graduação em Big Data - Inteligência na Gestão dos Dados.

Área de concentração: Tecnologia da Informação - Big Data

Orientador: Prof. Dr. Marcelo Morandini

SÃO PAULO

2022

Dedico esta monografia à minha querida avó Eunice, por tudo que ela representa para a família Mercadante.

AGRADECIMENTOS

Agradeço a Deus pelo sopro da vida, pela força de vontade que me move e, junto ao Dr Bezerra de Menezes, pelo despertar da espiritualidade em constante evolução dentro de mim.

Agradeço ao meu pai Edson por financiar meus estudos até os 15 anos de idade, possibilitando a base de conhecimento que adquiri para continuar minha trajetória acadêmica.

Agradeço à minha mãe Rita e à minha irmã Isadora - pontos de luz na Terra, que ciência alguma explica - por me provarem diariamente que, ainda que eu fale a língua dos anjos e dos homens, sem amor eu nada seria.

Agradeço à minha irmã Michelle, meu cunhado Rodrigo e meus sobrinhos Júlia e Otávio, por se fazerem presentes mesmo à distância, sendo a família mais amorosa que eu poderia ter.

Agradeço aos meus tios e avós pela referência familiar construída por todos eles em conjunto, porque se hoje sou fruto, tive raízes profundas que sustentaram a base do meu sucesso.

Agradeço à minha namorada Caroline por todo apoio, empenho, amor e compreensão diante dos desafios que permearam a conclusão deste trabalho, desde questões acadêmicas até questões médicas.

Agradeço aos docentes que contribuíram para a conclusão deste trabalho, em especial à coordenadora da especialização Solange Nice e ao orientador Marcelo Morandini, que me deram todo suporte necessário para alcançar meu objetivo.

Agradeço aos amigos e colegas de profissão da DP6 e da Warren Brasil, pelo incentivo diário e pela oportunidade de aprender com cada um à minha volta.

Agradeço aos profissionais da saúde e da ciência que trabalharam arduamente em prol da humanidade em meio à pandemia, da qual fui vítima durante a execução deste trabalho, mas também fui vencedor ao superá-la.

“Sei que o meu trabalho é uma gota no oceano, mas sem ele o oceano seria menor.”

(Madre Teresa de Calcutá)

RESUMO

O conceito de CDP (*Customer Data Platform*) surgiu em 2013 com o avanço de tecnologias voltadas para *Marketing*, em que diferentes perfis sistêmicos passaram a ser observados e desenvolvidos visando um interesse comum: entender o comportamento do cliente. CDP se propõe a mitigar os desafios de tecnologias precedentes, como MDM (*Master Data Management*), CRM (*Customer Relationship Management*), TMS (*Tag Management System*) e DMP (*Data Management Platform*), lidando com escopo específico de rastrear o comportamento do cliente em seus múltiplos pontos de contato com a organização (físicos ou digitais), para flexibilizar a exploração e criação de perfis de cliente, e viabilizar a integração de dados acionáveis com os demais componentes do ecossistema da companhia, colocando o cliente no centro das decisões da empresa. Nesse contexto, o objetivo desta monografia é aprofundar o conhecimento nos desafios que uma CDP demanda e analisar a Arquitetura de Informação compatível ao caso de uso de uma *fintech* brasileira, que visa a adoção de CDP como estratégia de crescimento da sua área de *Growth Marketing*, colocando o cliente no centro da organização. Com o apoio de referências como *DAMA Framework* e *RM-ODP Framework*, a metodologia *Design Science* estabelece a Arquitetura de Informação como artefato para análise e observabilidade diante do contexto da *startup*.

Palavras-Chave: *Fintech*, Marketing de Crescimento, Centralidade no Cliente, Gestão de Dados, Transformação Digital, *Customer Data Platform*, *RM-ODP Framework*, *DAMA Framework*, Arquitetura de Informação, *Design Science*.

ABSTRACT

The concept of CDP (Customer Data Platform) emerged in 2013 with the advancement of technologies aimed at Marketing, in which different systemic profiles began to be observed and developed with a common interest: understanding customer behavior. CDP aims to mitigate the challenges of predecessor technologies such as MDM (Master Data Management), CRM (Customer Relationship Management), TMS (Tag Management System) and DMP (Data Management Platform), dealing with specific scope of tracking customer behavior in its multiple points of contact with the organization (physical or digital), to make the exploration and creation of customer profiles more flexible, and to enable the integration of actionable data with the other components of the company's ecosystem, placing the customer in the center of the company's decisions. In this context, the objective of this monograph is to deepen knowledge on the challenges that a CDP demands and to analyze the Information Architecture compatible with the use case of a Brazilian fintech, which aims to adopt CDP as a growth strategy for its Growth Marketing area, putting the customer at the center of the organization. With the support of references such as DAMA Framework and RM-ODP Framework, the Design Science methodology establishes the Information Architecture as an artifact for analysis and observability in the context of the startup.

Keywords: Fintech, Growth Marketing, Customer Centric, Data Management, Digital Transformation, Customer Data Platform, RM-ODP Framework, DAMA Framework, Information Architecture, Design Science.

LISTA DE FIGURAS

Figura 1 —	Estágios do Cliente no Funil de Growth Marketing, com destaque sobre o que é mensurado.....	21
Figura 2 —	Estágios do Ciclo de Vida dos Dados.....	22
Figura 3 —	Segmento de Serviços Financeiros liderando o ranking de pontuação A&DQ no Brasil, evidenciando a maturidade digital e a demanda por inovação no setor.....	23
Figura 4 —	Gráfico indicando a prática de gestão “Modelos e plataforma tecnológica” em defasagem no mercado brasileiro, com baixo índice se comparada às demais práticas.....	24
Figura 5 —	Exemplo de interação de investidor pessoa física com a CDP.....	25
Figura 6 —	Visões Arquiteturais RM-ODP.....	28
Figura 7 —	Áreas de Conhecimento em Gerenciamento de Dados segundo o DMBOK <i>Framework</i>	30
Figura 8 —	Pirâmide de Peter Aiken utilizando as funções do DMBOK como diretrizes para evolução de gerenciamento de dados nas organizações.....	32
Figura 9 —	Modelo de organização ágil.....	36
Figura 10 —	Organograma da fintech com as principais diretorias do processo decisório da CDP, com ênfase nas gerências responsáveis pelas visões Empresa e Informação.....	38
Figura 11 —	BPMN do fluxo de aquisição de investidores da fintech, segundo relato da Gerência de Aquisição da Diretoria de Growth Marketing.....	45
Figura 12 —	Governança dos Dados da CDP.....	47
Figura 13 —	Modelo Orientado a Agregação para Arquitetura CDP em notação UML.....	57
Figura 14 —	Representação do Modelo de Agregação em formato JSON.....	58

LISTA DE TABELAS

Tabela 1 — Requisitos funcionais (RF) da CDP.....	41
Tabela 2 — Requisitos não-funcionais (RNF) da CDP.....	42
Tabela 3 — Descritivo dos indicadores-chave de desempenho (KPIs) da Gerência de Aquisição.....	45
Tabela 4 — Descritivo das origens de dados da CDP.....	47
Tabela 5 — Lista de Atributos de Usuário para Arquitetura da CDP.....	48
Tabela 6 — Lista de Atributos de Evento para Arquitetura da CDP.....	50
Tabela 7 — Lista de Eventos para Arquitetura da CDP.....	53

LISTA DE ABREVIATURAS E SIGLAS

AI	Arquitetura da Informação
ACID	Atomicidade, Consistência, Isolamento e Durabilidade
BPMN	<i>Business Process Model and Notation</i>
CDP	<i>Customer Data Platform</i>
CMO	<i>Chief Marketing Officer</i>
CMP	<i>Consent Management Platform</i>
CRM	<i>Customer Relationship Management</i>
CRO	<i>Conversion Rate Optimization</i>
CVM	Comissão de Valores Mobiliários
DAMA	<i>Data Management</i>
DMBOK	<i>Data Management Body of Knowledge</i>
DMP	<i>Data Management Platform</i>
GDPR	<i>General Data Protection Regulation</i>
IAAS	<i>Integration as a Service</i>
IHC	Interação Humano Computador
JSON	<i>JavaScript Object Notation</i>
KPI	<i>Key Performance Indicator</i>
LGPD	Lei Geral de Proteção de Dados Pessoais
MDM	<i>Master Data Management</i>
NIST	<i>National Institute of Standards and Technology</i>
NOSQL	<i>Not Only Structured Query Language</i>
OKR	<i>Objectives and Key Results</i>
PII	<i>Personally Identifiable Information</i>
PROV	<i>Provenance</i>
RDBMS	<i>Relational Database Management System</i>

RF	Requisito Funcional
RM-ODP	<i>Reference Model for Open Distributed Platform</i>
RNF	Requisito Não Funcional
SDK	<i>Software Development Kit</i>
SQL	<i>Structured Query Language</i>
TMS	<i>Tag Management System</i>
UML	<i>Unified Modeling Language</i>

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Contexto	13
1.2 Problema	15
1.3 Objetivo	16
1.3.1. Objetivos Específicos	16
1.4 Justificativa	16
1.5 Metodologia	18
2 REFERENCIAL TEÓRICO	19
2.1 <i>Big Data Model</i>	19
2.2 <i>Growth Marketing</i>	20
2.3 Ciclo de Vida dos Dados	21
2.4 Maturidade Digital	23
2.6 <i>Customer Data Platform</i>	24
2.7 <i>RM-ODP Framework</i>	27
2.8 <i>DAMA Framework</i>	28
3 DESENVOLVIMENTO	33
3.1 Contexto de Aplicação	34
3.1.1 Partes Interessadas	34
3.1.1.1 Pessoa Jurídica	34
3.1.1.2 Pessoa Física	39
3.1.2 Requisitos de Sistema	40
3.2 Descrição Arquitetural	44
3.2.1 Visão Empresa	44
3.2.2 Visão Informação	46

3.3 Resultados	59
4 CONCLUSÃO	61
4.1 Conhecimento Adquirido	62
4.2 Oportunidades	63
5 REFERÊNCIAS	64

1 INTRODUÇÃO

A Arquitetura da Informação (AI), oriunda de outras áreas de conhecimento, como Ciência da Computação, Ciência da Informação e Interação Humano Computador (IHC), é uma área multidisciplinar que visa mapear, rotular e organizar dados com o objetivo de promover usabilidade conforme seu contexto de aplicação. Embora os primeiros registros sobre o tema tenham surgido por volta de 1976 por Richard Wurman, o termo ganhou notoriedade e relevância com o surgimento da internet, onde a arquitetura de informação passou a ser vital para a experiência do usuário em ambientes digitais (SILVA, 2013).

Com o avanço da Era da Informação e do *Marketing 4.0*, surge um novo perfil empresarial, as *startups*. Elas se resumem por um modelo de negócio repetível e escalável, sob alta condição de incerteza (SEBRAE, 2021). Segundo Calicchio et al. (2019), o brasileiro tende ao empreendedorismo, mas o Brasil é um dos países mais difíceis do mundo para se estabelecer um negócio, uma vez que $\frac{2}{3}$ das *startups* fecham antes de completar 5 anos de existência (CALICCHIO et al., 2019). Dentre os principais motivadores de fracasso das *startups* estão a falta de clientes e a falta de capital para manter o negócio.

Em contrapartida, os principais fatores diferenciais para o sucesso desses negócios estão as práticas de *Analytics* e *Big Data*, presentes na realidade das *startups* de tecnologia do segmento financeiro, denominadas *fintechs* (CALICCHIO et al., 2019). Estima-se mais de 1000 empresas no setor, acumulando mais de 4,5 bilhões de reais investidos entre 2012 e 2021, evidenciando um mercado movimentado e concorrido (DISTRITO, 2020). Diante do contexto desafiador das *startups*, a estratégia de centralidade no cliente, com *Marketing* de crescimento (*Growth*) orientado a dados, se apresenta como necessidade competitiva, não só diferencial.

Para uma empresa adotar centralidade no cliente em sua estratégia de dados concentrada no modelo de negócio digital, é necessário certo grau de maturidade. *Customer Data Platform* (CDP) aparece como solução tecnológica para sustentar esse modelo. É comum que diversos fornecedores desse tipo de *software* aparentem ser o caminho mais rápido para o sucesso, porém, independente da

aquisição ou construção da solução, o Entendimento de Negócio e a Arquitetura de Informação são requisitos determinantes para tal decisão. Nesse sentido, *RM-ODP Framework* e *DAMA Framework* são base para especificação de Arquitetura da Informação direcionada para o contexto de CDP.

1.1 Contexto

O setor que mais cresce no universo das *startups* é o de tecnologia para o segmento financeiro. Os dados para as empresas são vitais quando o modelo de negócio é sustentado por *Big Data*, o que exige excelência em qualidade no ciclo de vida dos dados, não apenas para explicar o passado por meio de análises descritivas e diagnósticas, mas também para se preparar ao futuro, com análises preditivas e prescritivas avançadas em *Data Science*. Diante desse cenário, as *fintechs* têm revolucionado o segmento financeiro com avanços tecnológicos e novos modelos de negócio (CALICCHIO et al., 2019).

Segundo IBM Institute for Business Value (2018), os *Chief Marketing Officers* (CMOs) preferem o foco no cliente ao invés do foco no produto, com novas tecnologias que permitam coleta e análise de dados simultaneamente, fazendo com que as empresas não sejam apenas orientadas a dados, mas também orientadas a análises, o que na prática são conceitos diferentes associados à centralidade no cliente (IBM INSTITUTE FOR BUSINESS VALUE, 2018). Conforme Gupta et al. (2020), observa-se a necessidade de uma estrutura organizacional que favoreça a geração de valor ao cliente, com tecnologias da nova era que lidem com os desafios de rastreamento da jornada em mídias e dispositivos digitais, enfatizando atenção às conformidades de privacidade de dados (GUPTA et al., 2020).

Segundo Martins et al. (2019), Dados e *Analytics* são uma das práticas de gestão mais desafiadoras para a transformação digital no Brasil, em que apenas 12% das empresas pesquisadas são capazes de estruturar e integrar dados de maneira robusta com soluções direcionadas para centralidade no cliente, como DMPs (*Data Management Platform*) e CDPs (*Customer Data Platform*) (MARTINS et al., 2019), ambas plataformas que compõem o ecossistema de dados de uma empresa líder do

segmento financeiro. Isso evidencia o grau de maturidade exigido para implantar tais soluções que alavancam o potencial das organizações.

Na visão de Gartner (2016), a expectativa para atingir maturidade com CDP como solução, direcionada para Marketing Digital, era de 5 anos. Contudo, após 5 anos da primeira divulgação sobre o tema, Gartner (2021) afirma novamente o prazo de mais 5 anos como estimativa, sendo essa a expectativa para atingir maturidade com CDP, reforçando a percepção de dificuldade do mercado em implementar a solução durante o período entre 2016 e 2021 (GARTNER, 2016, 2021).

Aprofundando no conceito de *Customer Data Platform* (CDP), ele se define por “*software* empacotado que cria um banco de dados de cliente persistente e unificado que pode ser acessado por outros sistemas [...]” (EARLEY, 2018). Com a solução empacotada, devidamente estruturada para gerar valor ao *Marketing* ou à área responsável pelo crescimento da empresa, analistas têm mais flexibilidade para criar perfis de cliente e torná-los acionáveis, permitindo a customização da experiência do consumidor de acordo interesses estratégicos (EARLEY, 2018).

Ainda, observa-se CDP com foco primário na integração de dados, mas nota-se que CDP objetiva expandir seu escopo para análises preditivas, campanhas de marketing e entregas de mensagem para o cliente, a depender da estratégia de dados da organização (EARLEY, 2018). Desse modo, entende-se que uma CDP pode ser completa, considerando os princípios das funções de Gerenciamento de Dados estabelecidos pelo DMBOK, aplicados especificamente para centralidade no cliente (DAMA INTERNATIONAL, 2017).

Para Linington (1995), estruturar um sistema distribuído pode ser substancialmente complexo, com diferentes considerações em pauta. Portanto, deve-se elaborar essa estruturação modularmente quando possível, havendo clareza de especificação nas diferentes perspectivas dos módulos (LININGTON, 1995). Por CDP ser representada por um *software*, entende-se que o RM-ODP *Framework* (*Reference Model for Open Distributed Platform*) e seus pontos de vista de arquitetura (*Viewpoints*) podem contribuir para sua implementação.

Para tal, é fundamental que as visões Negócio e Informação sejam restringidas ao seu caso de uso. Nesse sentido, o *DAMA Framework (Data Management)* também compreende Negócio e Informação como facetas da área de conhecimento Arquitetura de Dados, identificando as necessidades de dados da empresa independentemente da estrutura (DAMA INTERNATIONAL, 2017).

Arquitetar dados para CDP, seja em *startups* ou grandes corporações, demanda um modelo de dados de cliente direcionado para tal solução, que mapeia os principais fatores que compõem o perfil e a jornada do cliente no contexto de negócio em que ele está inserido. Esse mapeamento é comumente utilizado pelos analistas de dados da área responsável pelo crescimento da empresa em expansão, geralmente *Growth Marketing* em *startups*, que se orienta por dados para o completo entendimento sobre quem é o cliente e como ele se comporta (TAIL, 2020).

O modelo de CDP além de registrar informações explícitas do cliente, como dados pessoais inseridos sistemicamente, também registra informações implícitas, como dados de inferência gerados pelo uso de um dispositivo digital. A especificação do modelo classifica as informações em atributos de cliente, eventos comportamentais e metadados gerais, independente dos diversos casos de uso em que a solução pode ser aplicada (LEMNISK, 2020).

1.2 Problema

O alto grau de maturidade exigido no mercado para desenvolver ou implantar CDP, especialmente devido à sua complexidade de Arquitetura de Informação, tende a gerar dificuldades às partes interessadas que visam obter visão unificada de seus clientes para extrair valor de dados acionáveis. Nesse sentido, tais desafios podem ocasionar problemas de entendimento de negócio e de dados, levando a projetos mal planejados e executados. Em um cenário de startup, principalmente do segmento financeiro, pode determinar o seu fracasso perante a concorrência.

1.3 Objetivo

O objetivo geral da monografia é analisar uma Arquitetura de Informação direcionada ao caso de uso de CDP aplicada à área de *Growth Marketing* de uma *fintech* brasileira, em que *RM-ODP Framework* e *DAMA Framework* são utilizados como referência para estabelecer as visões Empresa e Informação. Assim, busca-se fornecer conteúdo com potencial de auxiliar as atividades dos arquitetos de dados e demais interessados na implementação de CDP.

1.3.1. Objetivos Específicos

Para cumprir o objetivo geral, foram definidos os objetivos específicos a seguir:

- Apresentar *Customer Data Platform*;
- Apresentar *RM-ODP Framework*;
- Apresentar *DAMA Framework*;
- Apresentar requisitos funcionais e não funcionais de CDP ao contexto de *Growth Marketing* em *fintech* nacional;
- Apresentar visões arquiteturais Empresa e Informação, para entendimento de Negócio e Dados, respectivamente;
- Apresentar análise de artefato para desenvolvimento ou aquisição de CDP, a partir da Arquitetura de Informação estabelecida ao contexto em questão;

1.4 Justificativa

O termo *Customer Data Platform* aparece pela primeira vez figurando o mercado de tecnologia em 2013, citado por David Raab, anos antes das menções realizadas por Gartner (2016) e Martins et al. (2019) (GARTNER, 2016; MARTINS et al., 2019). Ao longo do tempo, com a crescente demanda das empresas por dados para gerar valor de negócio através do Marketing Digital, os *cookies* de navegadores *web* foram peça-chave no rastreamento comportamental de clientes em ambientes digitais até o avanço de discussões éticas e regulações espalhadas pelo globo sobre Privacidade de Dados entrarem em vigor (PADÍN, 2021).

Direcionadas para tratamento de dados pessoais, iniciativas como General Data Protection Regulation (GDPR) no continente europeu e Lei Geral de Proteção de Dados Pessoais (LGPD) no Brasil (BOTTINO, 2020), incentivaram o mercado tecnológico a assimilar que o uso de CDP pode contribuir para Transparência de Dados Pessoais (COLETI, 2020) aos investidores, seja de forma autônoma ou integrada a uma CMP (Consent Management Platform).

Diversos fornecedores de plataformas de dados adotaram reposicionamento de marca, direcionando seu modelo de negócio para atender as necessidades que uma *Customer Data Platform* visa suprir e se auto intitulado CDPs. Com isso, uma variedade de ferramentas, sistemas e plataformas, previamente construídas para um propósito já estabelecido no mercado (CRMs, DMPs, TMSs e afins), passaram a se comprometer com outro escopo, sendo ele:

- Reconhecimento da unicidade do cliente em múltiplos dispositivos;
- Gestão de perfis e audiências a partir de atributos e ações do cliente;
- Integração de dados acionáveis para múltiplas atividades de CRO;

O resultado desse movimento de mercado é a presença de soluções com diferentes vieses funcionais de origem, se adaptando a um novo conceito como proposta de valor e diferentes fornecedores competindo entre si.

O CDP Institute, liderado por David Raab, se propõe a estabelecer critérios para a categorização de *Customer Data Platform*, definindo requisitos funcionais, classificando vieses de origem, mapeando variações de casos de uso, comparando e certificando empresas que se enquadram nesses critérios. Enfim, o instituto promove o tema para a comunidade e nivela o conhecimento dos profissionais interessados nesta pauta, colaborando para com a compreensão tanto das empresas contratantes, que visam adotar a centralidade no cliente como estratégia digital, quanto dos fornecedores, que visam ser reconhecidos como CDP. (CDP INSTITUTE, 2021).

Apesar do vasto conteúdo elaborado e compartilhado pelo órgão, há uma carência de conteúdo quando o enfoque é Arquitetura, tema presente em fóruns de *Big Data* como o NIST (National Institute of Standards and Technology), comumente citado

como instituto regulatório de referência a repositórios como *Data Lakes* e *Data Warehouses* em geral (NIST, 2021).

1.5 Metodologia

A metodologia *Design Science* propõe um *framework* para, a partir da compreensão de um contexto social com potencial problema ou limitação conhecida, estabelecer um protótipo de artefato visando aplicá-lo a esse contexto. Para tal, observa-se eventuais reflexos desse artefato e suas capacidades de resolução. Intermediando o desenvolvimento do artefato e a avaliação do seu impacto, há a pesquisa acadêmica por referenciais teóricos correlatos, artefatos similares desenvolvidos pela academia e impactos previamente observados pelo uso de outras soluções. Portanto, o *Design Science* direciona o projeto para gestão de conhecimento em ambas as etapas do ciclo proposto pelo *framework*: desenvolvimento e análise (WIERINGA, 2014).

2 REFERENCIAL TEÓRICO

Nesta seção, é apresentada a revisão bibliográfica dos temas pertinentes para a monografia em questão. Os temas são abordados a partir dos seguintes tópicos: *Big Data Model*, *Growth Marketing*, Ciclo de Vida dos Dados, Maturidade Digital, *Customer Data Platform*, *RM-ODP Framework*, *DAMA Framework*.

2.1 *Big Data Model*

O termo *Big Data* é conhecido desde os anos 2000, porém ao longo do tempo foi distorcido pela falta de compreensão. Mudanças tecnológicas, como computação em nuvem, *smartphones*, redes sociais e avanços de conexão com a internet, permitiram às pessoas gerarem dados em grande variedade, com mais velocidade e volume. Assim ficou conhecido o termo dos 3 Vs (*Volume*, *Variety*, *Velocity*), que resume a expressão *Big Data*, exigindo processamento distribuído para lidar com a complexidade em questão (DAMA INTERNATIONAL, 2017).

Seguindo o avanço do *Big Data*, encontra-se o termo *Data Science*, que surge com o propósito de analisar o produto gerado pelo avanço de *Big Data*, de modo a não apenas se concentrar em análises descritivas e diagnósticas, que explicam eventos passados (tarefa até então tradicional de Business Intelligence, ou BI), mas também a olhar para trás com o intuito de promover análises preditivas e prescritivas, isto é, projetando o futuro (DAMA INTERNATIONAL, 2017).

Para que práticas de *Big Data* sustentem as práticas de *Data Science* em contexto de negócio, é fundamental um *Big Data Model*, que tradicionalmente lida com a compreensão da interação entre pessoas e dados sistêmicos (devidamente armazenados) mas considera NoSQL (Not Only Structured Query Language) como distinção da modelagem relacional, sendo mais complexo pelo modo como estrutura o armazenamento no banco de dados (FOWLER; SADALAGE, 2012).

A abordagem de Modelagem Orientada a Agregação, referenciada também como metamodelo por Fowler e Sadalage (2012), está presente em 3 das 4 categorias de ecossistema NoSQL:

- Chave-Valor - consulta a apenas um agregado completo por pesquisa, baseada em chave única para busca;
- *Documentos* - consulta à parte do agregado em vez de tudo, baseada na estrutura interna do documento;
- Família de Colunas - consulta à estrutura agregada de dois níveis, baseada em chave de linha do primeiro nível e chave de coluna do segundo nível;
- Grafos: consulta não associada à agregação, baseada em interconexões e relacionamentos complexos, constituídos de nós e arestas.

Importante ressaltar que a Modelagem Orientada a Agregação e as diferentes categorias de modelo previamente mencionadas não se aplicam ao contexto de *Relational Database Management System* (RDBMS), em que dados são manipulados via SQL (*Structured Query Language*) e a ênfase de uso se dá pela garantia ACID (Atomicidade, Consistência, Isolamento e Durabilidade) para um ambiente transacional. CDP não possui características de um RDBMS tradicional, sendo uma solução Big Data com foco analítico.

2.2 Growth Marketing

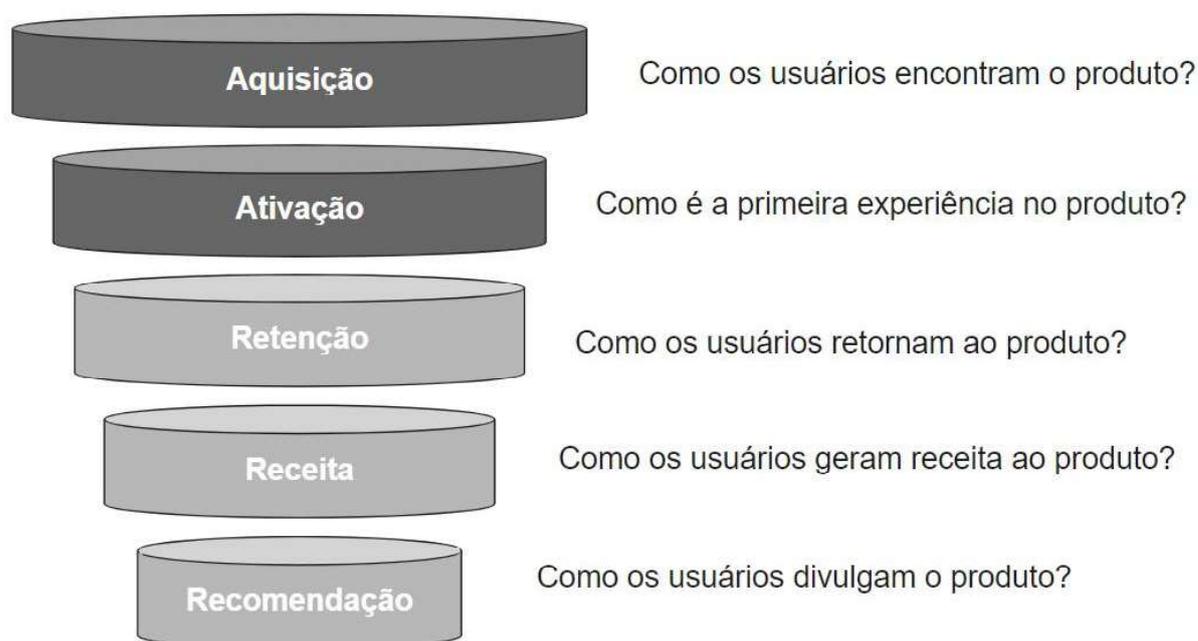
A estratégia de *marketing* inovadora das empresas que visam crescimento acelerado recebe o nome de *Growth Hacking*. Nesse sentido, o termo também é atribuído ao setor das empresas (comumente *startups*) que têm a responsabilidade de expandir o crescimento da base de clientes e, conseqüentemente, da receita gerada por eles (ELLIS; BROWN, 2018). Diante desse conceito, dois fatores são essenciais para o sucesso da estratégia:

- *Cliente* - quem gera receita e justifica a existência do negócio;
- *Dados* - o que orienta as rápidas decisões de *Growth Marketing*;

A seção 2.3 detalha Ciclo de Vida de Dados, mas para representar o Ciclo de Vida do Cliente, é importante destacar nesta seção o Funil de *Growth*, termo denominado para representar os estágios de um cliente na perspectiva da empresa que almeja o crescimento acelerado com centralidade no cliente (ELLIS; BROWN, 2018). A Figura

1, a seguir, relata os 5 estágios do cliente para *Growth Marketing*, frisando o que cada estágio visa medir, com destaque para Aquisição e Ativação, objetivo primário do conteúdo desta monografia.

Figura 1 - Estágios do Cliente no Funil de *Growth Marketing*, com destaque sobre o que é mensurado



Fonte: autoria própria

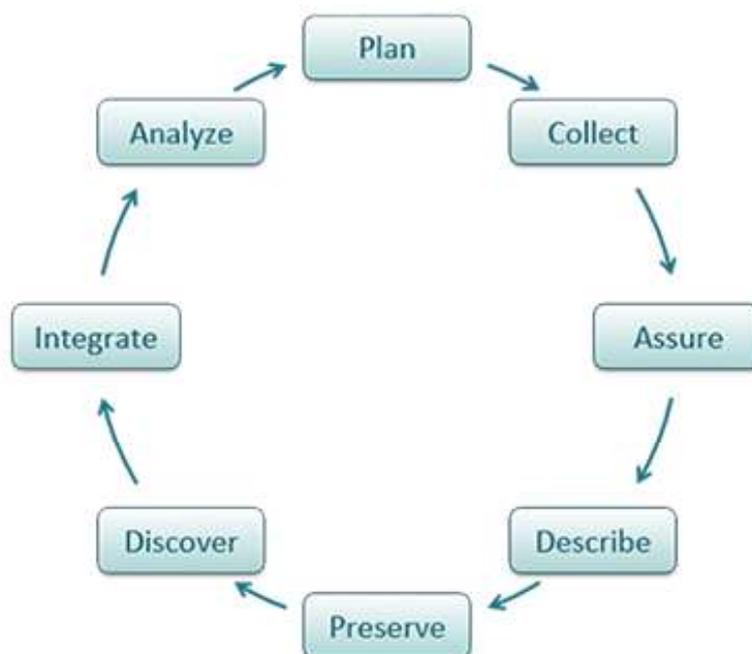
2.3 Ciclo de Vida dos Dados

Todo dado tem um ciclo de vida, com início, meio e fim, sendo ele temporal. Visando a proposição da arquitetura de dados contida nesta monografia, o Ciclo de Vida dos Dados desenvolvido pela DataONE Leadership Team (2021), propõe uma visão de alto nível das diferentes etapas do processo de Gestão de Dados, que pode ser definido por um conjunto de práticas de planejamento, execução e supervisão de políticas para preservar o valor do dado durante seu ciclo, atribuindo papéis e responsabilidades aos profissionais envolvidos.

Para o contexto de *Customer Data Platform* como sistema provedor de serviço (abordado na Seção 2.5 desta monografia), o processo cíclico e contínuo se aplica pelos estágios descritos a seguir e ilustrados na Figura 2 (DATAONE, 2021).

1. *Planejamento* - definição de dados a serem gerenciados e manipulados com o propósito de gerar valor;
2. *Coleta* - registro de dados manualmente ou baseado em sensores e gatilhos;
3. *Garantia* - inspeção e verificação de qualidade de dados;
4. *Descrição* - definição de metadados para representação dos dados;
5. *Preservação* - armazenamento durável de dados em repositório;
6. *Descoberta* - extração de dados relevantes e úteis para o negócio;
7. *Integração* - combinação de dados de diferentes origens;
8. *Análise* - análises descritivas, diagnósticas, preditivas e prescritivas, conforme objeto de uso e propósito.

Figura 2 - Estágios do Ciclo de Vida dos Dados

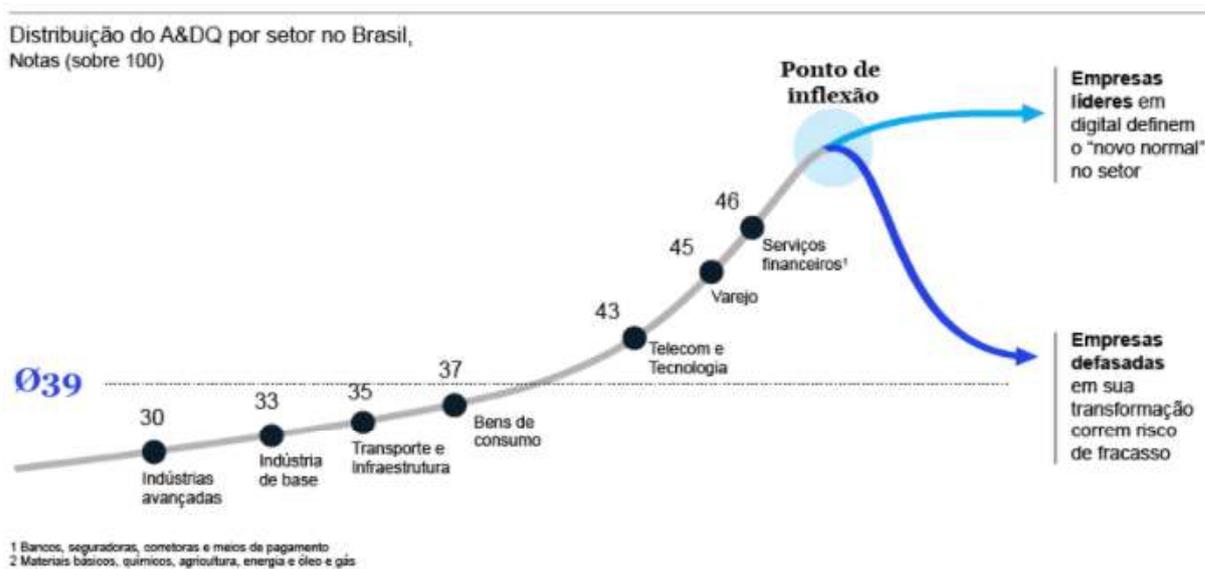


Fonte: DataONE (2021)

2.4 Maturidade Digital

O estudo “*Transformações digitais no Brasil: Insights sobre o nível de maturidade digital das empresas no país*”, realizado pela Martins et al. (2019), avalia a maturidade digital de 124 empresas brasileiras de médio e grande porte, em 7 diferentes setores do Brasil. Destaca-se Serviços Financeiros como líder em maturidade digital no país (MARTINS et al., 2019). A Figura 3 evidencia a maturidade digital e a demanda por inovação no setor.

Figura 3 - Segmento de Serviços Financeiros liderando o ranking de pontuação A&DQ no Brasil, evidenciando a maturidade digital e a demanda por inovação no setor

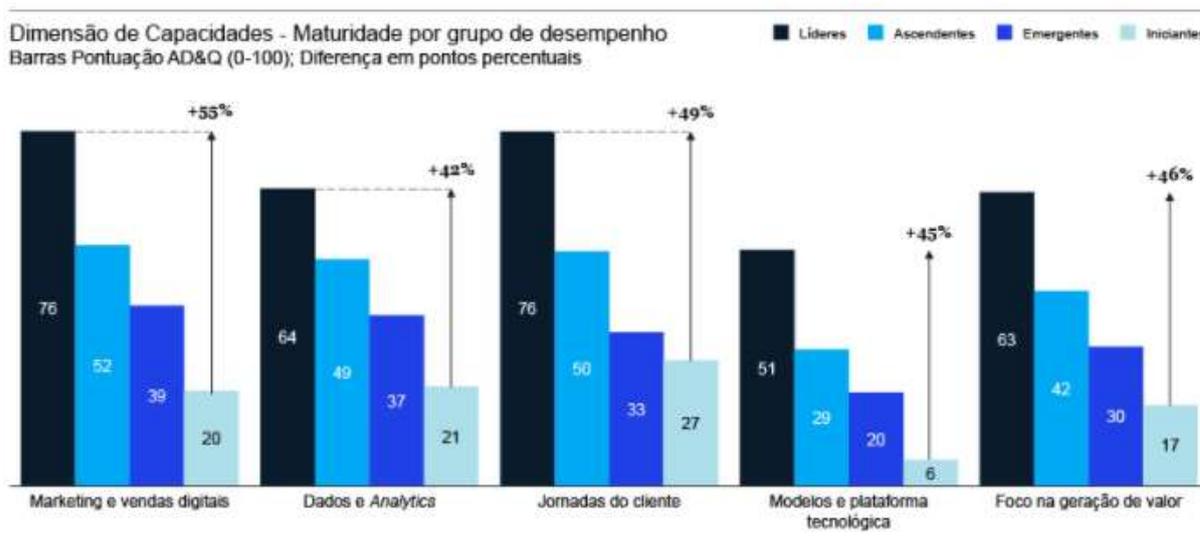


Fonte: Martins et al. (2019)

Na classificação “*Analytics and Digital Quotient*” (A&DQ) adotada pela consultoria, definem-se 4 pilares fundamentais para uma empresa alcançar a excelência no seu processo de digitalização e 22 práticas de gestão associadas a esses pilares, com categorias que determinam o estágio de maturidade das empresas participantes da pesquisa (MARTINS et al., 2019). Dos diferentes tópicos que o estudo ressalta, a dimensão “*Capacidades*” aponta um déficit significativo na prática de gestão “*Modelos e plataforma tecnológica*”. Tal fato evidencia que as empresas enfrentam dificuldades para estabelecer uma arquitetura tecnológica robusta capaz de suportar

estratégias organizacionais que visam adotar a centralidade no cliente para tomar suas decisões baseadas em dados.

Figura 4 - Gráfico indicando a prática de gestão “Modelos e plataforma tecnológica” em defasagem no mercado brasileiro, com baixo índice se comparada às demais práticas



Fonte: Martins et al. (2019)

Para alcançar a maturidade digital, o estudo conclui que as companhias devem se atentar às práticas avaliadas, mencionando *Customer Data Platform* (CDP) como diferencial de uma instituição financeira líder pertencente à pesquisa, que alia *MarTech* e *Big Data* (MARTINS et al., 2019).

2.5 Customer Data Platform

Customer Data Platform pode ser compreendido como um conceito de centralidade de cliente ou *software* estabelecido, em que diferentes componentes tecnológicos ou recursos sistêmicos são integrados para prover armazenamento de dados de todos os pontos de contato do cliente durante sua jornada com a empresa, viabilizando a criação e detecção de perfis consumidores por profissionais de *Marketing*, Produto, Dados e demais interessados em melhorar a experiência dos clientes da marca (CDP INSTITUTE, 2021).

Através da visão unificada proporcionada com a CDP, seu propósito concentra-se em aumentar a rentabilidade da companhia, integrando os dados descobertos a campanhas e comunicações direcionadas, testes de hipóteses multivariados e personalização de conteúdo, além de facilitar a gestão de consentimento para captura e manipulação de dados dos usuários de plataformas digitais (TAIL, 2020).

Figura 5 - Exemplo de interação de investidor pessoa física com a CDP



Fonte: autoria própria

De acordo com o CDP Institute (2021), há 4 tipos de CDP como *software*, cada qual abrangendo um escopo maior que o nível anterior, sendo (CDP INSTITUTE, 2021):

1. *Dados* - executa as funções mínimas de *software* CDP, conectando dados de diferentes origens, reconhecendo os clientes unicamente e armazenando esses dados em um repositório para criação de audiências baseadas em regras aplicadas, acessíveis a outros sistemas. Este tipo de CDP comumente deriva de ferramentas de *Web Analytics* ou *Tag Manager System (TMS)*;
2. *Análise* - além de abranger o escopo das CDPs de Dados, possuem recursos de análise por meio de aprendizagem de máquina, mapeamento de jornada,

modelagem preditiva e atribuição de receita para as audiências descobertas. Geralmente se assemelham à *Data Management Platform* (DMP);

3. *Campanha* - além de abranger o escopo das CDPs de Análise, são capazes de gerenciar ações específicas para as audiências descobertas ou para os clientes individualizados, viabilizando dados para serviços de mensagens personalizadas, campanhas de *marketing* ou recomendações de produtos em tempo real. Geralmente derivam de *Customer Relationship Manager* (CRM);
4. *Entrega* - além de abranger o escopo das CDPs de Campanha, são especializadas na execução de determinadas tarefas que podem compreender mensagens diretas em canais ou práticas de *Conversion Rate Optimization* (CRO) realizadas comumente por outras ferramentas.

A modelo de dados de cliente, abordado por Earley (2018) e com CDP Institute sendo órgão pilar do tema, propõe mecanismos para compreensão de dados gerados no ecossistema empresarial, reforçando as características (EARLEY, 2018):

- *Explícitas* - quando um usuário insere deliberadamente uma informação na base de dados;
- *Implícitas* - quando por meio de monitoramento ou inferência (e de maneira consentida) informações do cliente são registradas.

Neste contexto de aplicação, a Arquitetura da Informação é direcionada para o rastreamento de clientes em plataformas digitais, permitindo que perfis não técnicos da “pessoa jurídica” (gerentes de produtos, analistas de marketing e afins) compreendam os dados, manipulando fácil e corretamente as informações no ecossistema, gerando assim valor aos investidores “pessoa física” da *fintech* que acessam as plataformas digitais da *startup*. O modelo de dados para CDP classifica as informações conforme os tópicos listados abaixo (EARLEY, 2018):

- *Eventos* - dados que representam ações comportamentais do cliente;
- *Atributos* - dados característicos do cliente ou do contexto da ação;
- *Metadados* - dados que catalogam eventos e atributos gerais.

Vale frisar que os metadados previstos para *Customer Data Platform* podem ser associados ao conceito de Proveniência de Dados (ALPER; CURCIN; FAIRWEATHER, 2018), principalmente pela característica de interação dos potenciais investidores da *startup* com a plataforma tecnológica digital (KHAN; HASAN, 2015), como visualizar telas, apertar botões, preencher formulários e afins. Porém, o escopo desta monografia não tem caráter de PROV, ou *Provenance Data Model* (W3C, 2013), mas sim de *Customer Data Model* (EARLEY, 2018).

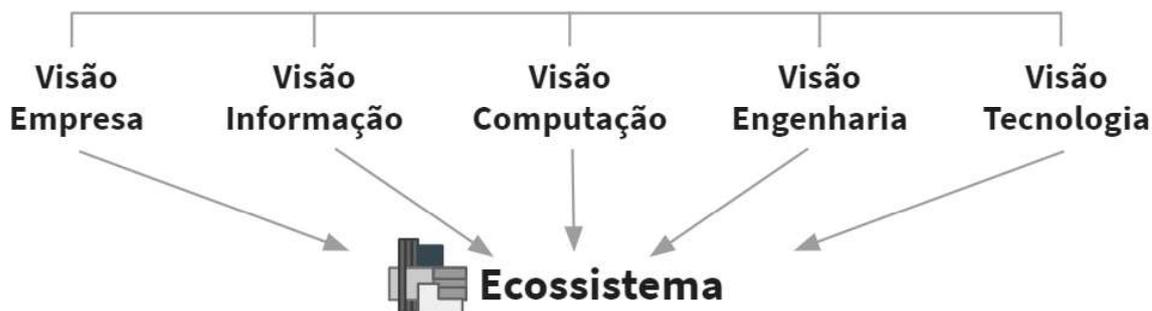
2.6 RM-ODP Framework

RM-ODP *Framework* é um modelo de referência que orienta a padronização de sistemas com processamento distribuído. Entende-se como padrão o vocabulário e a terminologia consensuais que emprega, além de métodos e visões arquiteturais estabelecidos para integrações sistêmicas, permitindo o desenvolvimento de uma solução rápida e incrementalmente (PUTMAN, 2000).

Dos seus elementos básicos de padronização RM-ODP possui a ISO/IEC 10746-3 (2009), que divide a Arquitetura de Solução em 5 partes, denominadas Visões Arquiteturais, para que diferentes *stakeholders*, com diferentes papéis e responsabilidades, possam avaliar a solução no escopo de sua competência, visando satisfazer o propósito base do *framework* que é atender aos requisitos de negócio. Entendem-se por Visões Arquiteturais de RM-ODP os estágios listados a seguir e ilustrados na Figura 6 (ISO/IEC 10746-3, 2009; PUTMAN, 2000).

1. *Empresa* - visão de negócio que define o propósito da arquitetura;
2. *Informação* - visão que descreve os dados que compõem a solução;
3. *Computação* - visão que define ações computacionais da solução;
4. *Engenharia* - visão que aborda componentes e artefatos da solução;
5. *Tecnologia* - visão que especifica as soluções técnicas da arquitetura.

Figura 6 - Visões Arquiteturais RM-ODP



Fonte: autoria própria

No contexto de aplicação em questão, consistido em uma *startup fintech* do mercado de investimentos, a ênfase do conceito de arquitetura do RM-ODP *Framework* está nas Visões Empresa e Informação, isto é, o entendimento de negócio em que a *Customer Data Platform* é vista como solução de valor para alavancar a maturidade digital da companhia, além do entendimento de dados presentes no ciclo de vida da informação para compor a CDP. Embora ambas as visões não sejam completamente independentes, são complementares e suficientes para a especificação da Arquitetura de Informação.

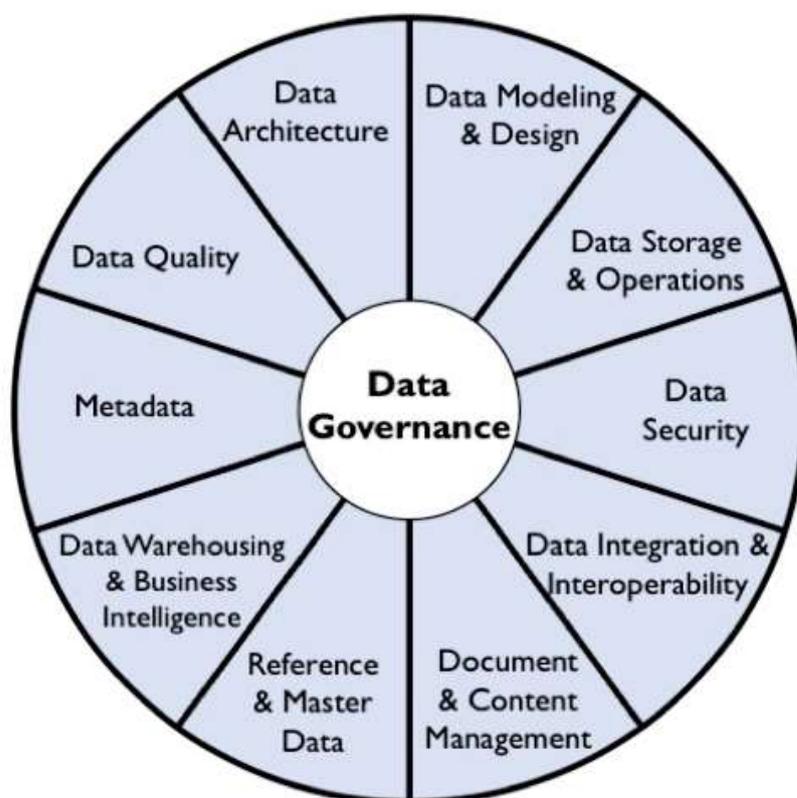
2.7 DAMA Framework

O DAMA International (2017) propõe um *framework* para orientar os profissionais com diferentes funções, de ordem técnica ou não, e estabelece áreas de conhecimento para Gestão de Dados, visando atender o nível de especialização de cada tema. A seguir são listadas cada área de conhecimento pertencente ao *framework*, com destaque à função Arquitetura, elemento norteador do DMBOK para entendimento de Negócio e Dados (DAMA INTERNATIONAL, 2017):

- *Governança* - exerce autoridade sobre os dados, sendo a base para as demais funções, estabelecendo estratégias, papéis e responsabilidades, riscos e contingências, políticas, processos e regulações;

- *Arquitetura* - define os componentes sistêmicos que compõem o ciclo de vida dos dados, estabelecendo escopos de arquitetura em visões de negócio, dados, aplicação e tecnologia;
- *Modelagem* - projeta o entendimento dos dados, de modo a representá-los em diferentes esquemas, conforme os requisitos de negócio;
- *Armazenamento e Operações* - estabelece as práticas de implementação e suporte a dados transacionais, operando o ciclo de vida desde o armazenamento até atividades de backup e exclusão;
- *Segurança* - visa garantir procedimentos associados a autenticação, autorização, acesso e auditoria de ativos de dados e informações seguindo as devidas adequações;
- *Integração e Interoperabilidade* - promove a comunicação de dados entre sistemas, pessoas e tecnologias, com ênfase em consolidações e compartilhamentos de informações;
- *Conteúdos e Documentos* - enfatiza a manipulação de dados não relacionais (não estruturados ou semiestruturados) durante o ciclo de vida, independente de formato ou mídia de armazenamento;
- *Dados Mestres* - gerencia dados vitais da organização, amplamente utilizados por diferentes setores e providos por diferentes fontes;
- *Inteligência de Negócio* - proporciona estrutura de dados consistente para suporte a decisão, sustentando relatórios, explorações e análises;
- *Metadados* - compreende a completa descrição para organização e entendimento dos dados da organização, desde a catalogação das informações até relacionamentos, tecnologias e processos que os envolvem;
- *Qualidade* - estabelece critérios de sucesso e promove práticas para satisfazer as necessidades dos consumidores de dados da organização, monitorando pessoas, processos e tecnologias.

Figura 7 - Áreas de Conhecimento em Gerenciamento de Dados segundo o DMBOK *Framework*



Fonte: DAMA International (2017)

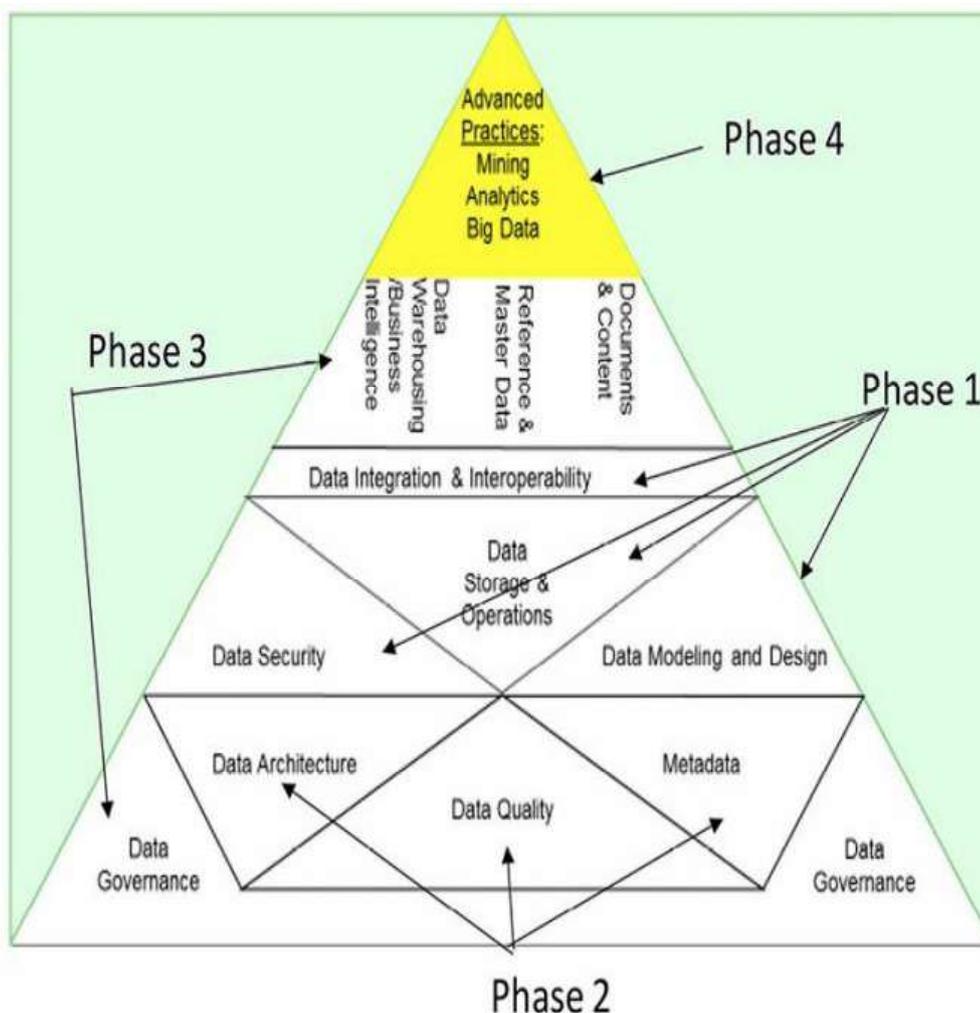
Empregar todas as funções de dados previstas pelo DMBOK para Gestão de Dados não é tarefa trivial. Naturalmente as empresas têm dificuldade para definir, desde o início, uma estratégia robusta de dados e, principalmente, de colocá-las em prática.

É comum que as práticas de gerenciamento de dados evoluam organicamente, à medida em que a empresa evolua também. No contexto de empresas startups esse comportamento é ainda mais evidente, considerando seu crescimento em escala, muitas vezes exponencial, e orçamento reduzido para investimento em pessoas, processos e tecnologias para Gestão de Dados, haja vista que o foco do investimento concentra-se no núcleo que sustenta o modelo de negócio.

A pirâmide de Peter Aiken (Figura 8) retrata o crescimento orgânico das empresas no geral e associa esse crescimento às funções da Gestão de Dados propostas pelo DAMA *Framework*, auxiliando as companhias a atingirem o patamar mais almejado, em que práticas de dados e análises são conjuntas, centralizando as decisões da organização (DAMA INTERNATIONAL, 2017). A pirâmide é composta por 4 fases:

- *Fase 1* - foco em requisitos mínimos de dados para o sistema transacional da empresa funcionar com segurança;
- *Fase 2* - foco em arquitetura para que o ciclo de vida dos dados da empresa seja estabelecido com qualidade;
- *Fase 3* - foco em governança de dados para que iniciativas estratégicas e tomadas de decisão sejam aplicadas;
- *Fase 4* - foco em extrair o máximo de valor dos dados por meio das práticas analíticas previamente pavimentadas pelo modelo piramidal;

Figura 8 - Pirâmide de Peter Aiken utilizando as funções do DMBOK como diretrizes para evolução de gerenciamento de dados nas organizações



Fonte: DAMA International (2017)

Para o contexto de aplicação desta monografia, ressalta-se a semelhança do modelo piramidal com a evolução orgânica da *fintech* em questão. Também vale destacar que a implementação de uma *Customer Data Platform*, detalhada conceitualmente na Seção 2.5, exige grau de maturidade da organização, em que aspectos de Arquitetura de Dados são aplicados antes de práticas avançadas de *Big Data* e *Analytics* da Gestão de Dados. A seguir, o capítulo 3 apresenta as atividades realizadas diante do contexto de *startup fintech*.

3 DESENVOLVIMENTO

Neste capítulo, são apresentadas as Visões Arquiteturais de Empresa e Informação, ambas baseadas em *RM-ODP Framework*, com viés de modelo referencial de desenvolvimento de software, e *DAMA Framework*, com viés de gerenciamento de dados organizacionais. As visões apresentam-se, respectivamente, com a proposta de viabilizar o entendimento de Negócio e Dados da solução que sustenta a centralidade no cliente da *startup*.

A organização do escopo em questão é cuidadosamente apelidada de *Fintech Y*, por razões de sigilo, em que o foco está no rastreamento comportamental digital de clientes no fluxo de aquisição da corretora de investimentos guiado pela sua área de *Growth Marketing*. A Arquitetura de Informação analisada tem como propósito o caso de uso de gestão de audiências para personalização de conteúdo, visando promover experiência digital mais aderente ao público-alvo. Entende-se que há todo o insumo necessário para avaliar:

- Implementação de solução *Integration as a Service (IaaS)*, acoplada a ferramentas de *Digital Analytics* e *Conversion Rate Optimization (CRO)* terceiras que reproduzam em conjunto o conceito de CDP;
- Implementação de solução *Customer Data Platform* terceira, rotulada tanto pelo fornecedor quanto pelo mercado como *software CDP*;
- Implementação de solução *Customer Data Platform* nativa, desenvolvida pela própria *startup* para atender requisitos específicos;

Considera-se que a implementação de *Customer Data Platform* permite crescimento em escala tanto para outros casos de uso quanto para outros setores da *startup*, sendo flexível para eventuais mudanças. Entende-se também que embora haja um contexto de aplicação e um escopo definido para Arquitetura da Informação, a análise contida neste documento permite ser referencial para *startups* de diferentes segmentos que consideram CDP como estratégia de centralidade no cliente para ações de *Growth Marketing*.

3.1 Contexto de Aplicação

Visto ao exposto, a proposta arquitetural tem como contexto a área de *Growth Marketing* de uma *fintech* do mercado de investimentos que busca aumentar sua base de clientes, a satisfação deles e o patrimônio sob gestão. Pelo conceito de centralidade no cliente, considera-se também que existe um interesse indireto desse cliente, objeto central da estratégia, dado que é a sua experiência a principal beneficiada. Se, por um lado, há o interesse da *fintech* em ofertar a melhor experiência digital aos seus potenciais clientes do segmento, representando uma pessoa jurídica, por outro, há também o interesse de clientes pessoa física interessados em utilizar serviços digitais cada vez mais direcionados e customizados para seu perfil investidor. A seguir, são detalhadas as partes interessadas.

3.1.1 Partes Interessadas

As partes interessadas pela *Customer Data Platform*, seja direta ou indiretamente, podem ser resumidas em:

1. *Pessoa Jurídica* - representa o escopo de atuação organizacional, destacando os principais avaliadores decisores da solução em âmbito corporativo, além dos operadores diretos da plataforma;
2. *Pessoas Físicas* - representa os clientes das *fintechs*, destacando os potenciais investidores, interessados na qualificação de sua experiência digital ao realizarem seus investimentos;

3.1.1.1 Pessoa Jurídica

A empresa referência, interessada diretamente em CDP, atua como *fintech* no mercado de investimentos há mais de 5 anos. Por ser nativamente digital, dispensa o processo de digitalização de bancos tradicionais com seu modelo de negócio associado à estratégia online e às capacidades tecnológicas modernizadas, compatíveis com a mentalidade ágil de uma *startup*.

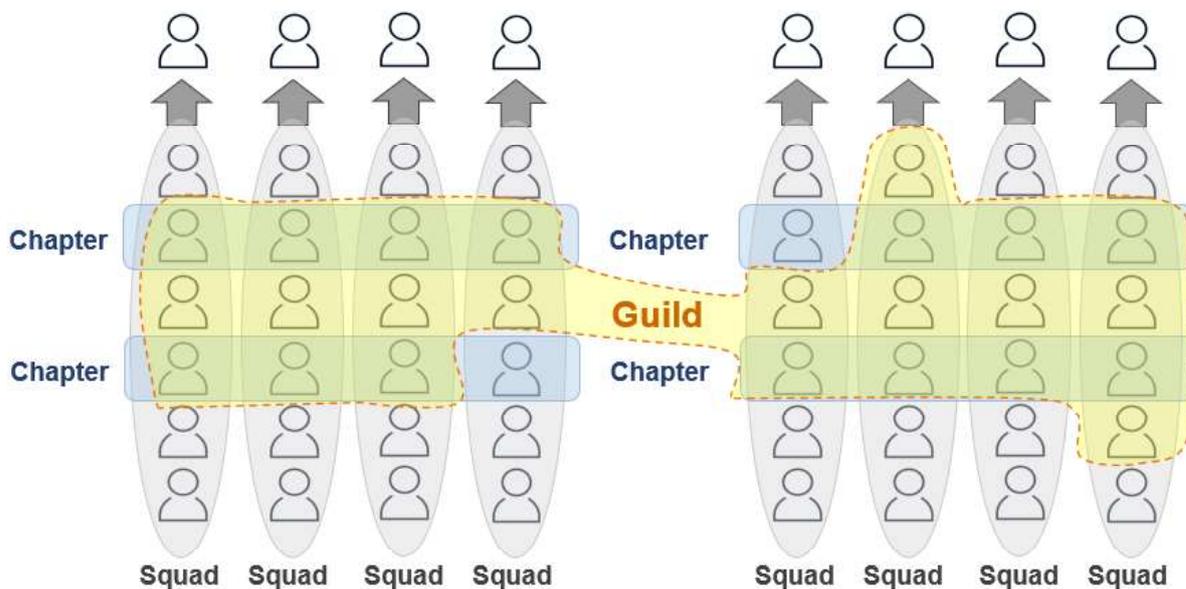
Apesar de sua característica jovem no mercado financeiro, o público-alvo é amplo e também se adapta à quebra de paradigma denominada “desbancarização”, uma vez que a proposta de valor da *fintech* é democratizar o mercado de investimentos no Brasil. Para alcançar o seu objetivo, descomplicando a experiência dos brasileiros no segmento, a *Fintech Y* atua em 3 escopos:

1. *Administração* - criação de fundos de investimentos, tornando-os acessíveis através corretoras de investimento;
2. *Corretagem* - oferta de fundos de investimento de outras administradoras, produtos da Bolsa de Valores e do Tesouro Nacional;
3. *Gestão* - alocação e distribuição de recursos para clientes investidores, conforme o perfil investidor deles;

A empresa está distribuída num modelo organizacional guiado pela metodologia *Objectives and Key Results* (OKRs), permitindo que a estratégia da empresa esteja alinhada às diferentes células que compõem a companhia, em especial o setor de *Growth Marketing*, principal operador da CDP em análise. A *startup* conta com aproximadamente 600 colaboradores, sendo 20 *stakeholders* momentâneos da CDP, podendo escalar para toda a organização no longo prazo. A estrutura organizacional da *Fintech Y* é composta por:

1. *Squads* - estrutura multidisciplinar autônoma, focada em um objetivo comum;
2. *Chapters* - estrutura composta por colaboradores de competência comum;
3. *Tribes* - estrutura matricial composta por diferentes squads e chapters;
4. *Guilds* - estrutura composta por colaboradores com interesses comuns.

Figura 9 - Modelo de organização ágil



Fonte: autoria própria

As diretorias de *Growth Marketing* e Engenharia são as principais partes interessadas, pois exercem papéis avaliadores da arquitetura e tomadores de decisão. Ambas as áreas são definidoras de requisitos funcionais e não funcionais, respectivamente. Estão listados a seguir, de forma mais aprofundada, o papel e a percepção de cada diretoria, com suas respectivas gerências:

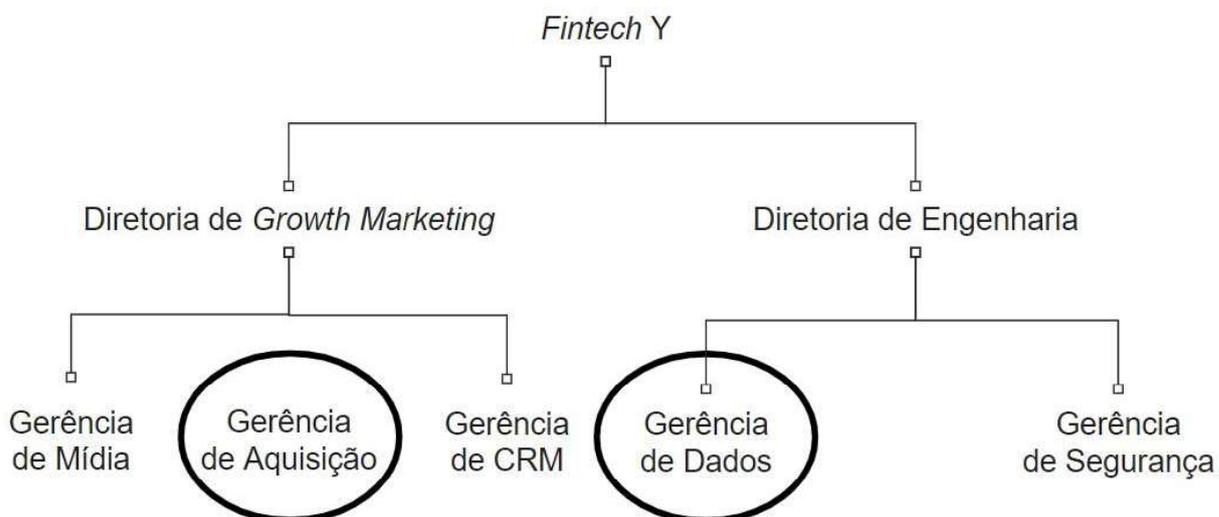
1. *Diretoria de Growth Marketing* - responsável por nortear os esforços da *startup* para que ela continue se estabelecendo no mercado e se sobressaindo em relação aos demais concorrentes do segmento. Monitora o funil de aquisição e ativação da *fintech*, sustentando a comunicação e o engajamento com potenciais clientes. Vê como benefício da CDP a flexibilidade para criação de perfis e gestão de audiências, a partir de atributos e ações comportamentais dos investidores, para otimizar o desempenho das campanhas pagas em mídias digitais, potencializar práticas de CRO e aumentar a sinergia com a plataforma de CRM, com réguas de comunicação mais assertivas;
2. *Gerência de Mídia* - responsável pela estratégia de veiculação de campanhas em mídias pagas, atraindo clientes para a *startup*. Vê a CDP como grande

benefício devido à necessidade recorrente de instalações de pixels de mídia (baseados em *cookies* de navegadores) para sustentar os mecanismos de ofertas de mídia para os clientes na internet. Com a CDP, além de maior precisão no rastreamento digital do cliente com *first party data*, há maior facilidade na segmentação de público-alvo das campanhas e na integração com ferramentas anunciantes no mercado publicitário, visando ativação de dados (MARTINS, E.; FUKUMOTO, M. S., 2021);

3. *Gerência de Aquisição* - responsável pela condução estratégica de todo o fluxo de aquisição digital da *startup*, explorando a melhor experiência digital possível para abertura de conta e primeiro depósito do cliente. Conduz a equipe operacional a atingir os resultados estipulados pela diretoria. A CDP certamente é vista como um produto de dados base para todas as atividades de *Conversion Rate Optimization* (CRO), incluindo testes A/B e multivariados, personalização de conteúdo e sistemas de recomendação para produtos compatíveis para com o cliente. Representa o usuário final da CDP pela perspectiva de pessoa jurídica, sendo assim ator operacional da *startup*;
4. *Gerência de CRM* - responsável pela estratégia de comunicação e relacionamento com o cliente, considerando todos os diferentes pontos de contato com ele dentro da organização. Vê como principal benefício da CDP a sinergia que a plataforma de dados pode oferecer, separando o escopo de cadastro de clientes e criação de perfis (função mínima de uma CDP) do escopo de criação de réguas de comunicação (função de uma CDP do tipo campanha ou solução CRM), permitindo o aprimoramento do processo de segmentação de audiências para comunicação;
5. *Diretoria de Engenharia* - responsável pela estratégia de tecnologia dos diferentes escopos técnicos da *startup*, como infraestrutura, *software*, rede, nuvem e afins, viabilizando o modelo de negócio da companhia através de plataformas tecnológicas. Vê como principal benefício da CDP a otimização de desempenho das aplicações nos canais digitais, concentrando a captura de dados em um único *software* e reduzindo a latência de rede na navegação dos clientes da *fintech*;

6. *Gerência de Dados* - responsável por toda estratégia de dados e análise da *startup*, munindo as áreas de negócio com dados qualitativos e pessoas capacitadas nesta disciplina. Vê como principal benefício da CDP a governança de dados que a plataforma oferece, bem como o ecossistema centralizado no cliente para gestão e análise de dados, incentivando a cultura de orientação a dados para tomadas de decisão. Dentro da companhia, atua diretamente em atividades de entendimento de Negócio e Dados;
7. *Gerência de Segurança* - responsável pelas conformidades regulatórias da *startup* e disseminação de práticas de segurança na companhia, garantindo confidencialidade, integridade, autenticidade e disponibilidade da informação. Vê como principal benefício da CDP uma fonte única da verdade para dados pessoais, possibilitando a gestão de consentimento do cliente via *Content Management Platform* (CMP) (SANTOS et al., 2021), anonimização e criptografia de dados, bem como rastreamento de bases legais.

Figura 10 - Organograma da fintech com as principais diretorias do processo decisório da CDP, com ênfase nas gerências responsáveis pelas visões Empresa e Informação



Fonte: autoria própria

Para o contexto de aplicação da *fintech*, considerando o apoio dos *frameworks* RM-ODP e DAMA, a Gerência de Aquisição é a principal contribuinte para estabelecer a Visão Empresa, enquanto a Gerência de Dados é a principal contribuinte para a Visão Informação.

3.1.1.2 Pessoa Física

O cliente digital pessoa física interessado indiretamente pela CDP, dado o benefício da estratégia de centralidade proporcionado à sua experiência digital como investidor, é também cliente pessoa física da *startup*, a pessoa jurídica interessada na arquitetura para promover a melhor experiência aos seus clientes. Neste cenário, a pessoa física é o usuário final interessado no mercado de investimentos, que busca através da *startup* ter a melhor experiência digital possível.

Para categorizar os investidores em perfis e audiências, seja explorando públicos-alvos ou exercitando a criação de “*personas*” (SIQUEIRA, 2013), a *startup* se utiliza de alguns atributos fornecidos pelos investidores durante a abertura de conta, processo legalizado e respaldado no Brasil por entidades regulatórias do setor, como Banco Central, Anbima e Comissão de Valores Mobiliários (CVM, 2011). Dos principais atributos de clientes, destacam-se:

Potencial Financeiro:

- a. Patrimônio abaixo de R\$50.000,00;
- b. Patrimônio entre R\$50.000,00 e R\$300.000,00;
- c. Patrimônio acima de R\$300.000,00.

2. *Perfil Investidor:*

- a. Propensão a pouco risco para menor rendimento;
- b. Propensão equilibrada entre riscos e ganhos;
- c. Propensão a alto risco para maior recompensa.

Dos atributos citados, o Perfil Investidor é resultado de uma análise baseada em diversas outras informações relevantes do cliente, como histórico investidor no mercado financeiro e dados demográficos (escolaridade, ocupação e afins). Há também informações cadastrais caracterizadas como pessoalmente identificáveis, ou *Personally Identifiable Information* (PII), mas elas não interferem na avaliação do Perfil Investidor e são devidamente consideradas na seção 3.2.2 Visão Informação.

3.1.2 Requisitos de Sistema

Para o escopo *Customer Data Platform*, o fator norteador da plataforma é atender requisitos de negócio para pessoas físicas. Sob esta ótica, a análise da Arquitetura de Informação é direcionada para o momento de transição da jornada do cliente entre Aquisição e Ativação (topo do funil de *Growth Marketing*), quando efetivamente é realizado um depósito financeiro na conta da *corretora* e, com esse ato, passa a ser considerado cliente ativo, propenso a rentabilizar a *startup* através da criação de uma carteira de investimento.

Os requisitos funcionais e não funcionais da solução, que compreendem o caso de uso em questão, foram relatados pelas partes interessadas da pessoa jurídica para o cenário de rastreamento do comportamento da pessoa física em canais digitais, sendo eles dispositivos de navegação *desktop* e *mobile*, ambas tecnologias provedoras de dados para *Customer Data Platform*.

Para rastreamento comportamental em *website* via *desktop*, a *CDP* necessita de comunicação com a aplicação por meio de implementação de código de instalação no próprio *website* da *startup*. Para o rastreamento comportamental em aplicativo *mobile*, a *CDP* necessita estabelecer comunicação com a aplicação por meio de implementação SDK (*Software Development Kit*) nos aplicativos disponibilizados nas lojas de seus respectivos *players*. Ambos os meios de rastreamento obedecem a critérios de segurança e privacidade em benefício das pessoas jurídica e física.

A partir de entrevistas realizadas com representantes da *Fintech Y* no 2º Semestre do ano de 2021, foram oferecidos insumos para análise de requisitos da CDP. Os resultados foram apresentados na Tabela 1 e na Tabela 2 de requisitos funcionais e não funcionais, respectivamente. As tabelas estão organizadas em ordem numérica com o descritivo compilado das respostas obtidas na entrevista. A divisão em duas tabelas tem a finalidade de auxiliar arquitetos e desenvolvedores a atenderem às suas necessidades particulares em seus devidos escopos.

Tabela 1 - Requisitos funcionais (RF) da CDP.

Requisito	Descritivo
RF001	A CDP deve permitir que o usuário digital (cliente potencial) não seja rastreado se assim o desejar
RF002	A CDP deve cadastrar o usuário digital (investidor potencial) como cliente da <i>startup</i> , em caso de consentimento
RF003	A CDP deve reconhecer o usuário digital (investidor potencial) como cliente da <i>startup</i> em qualquer dispositivo após o seu cadastro
RF004	A CDP deve ofertar dados para mensagens personalizadas ao potencial cliente da startup durante sua experiência digital em qualquer dispositivo
RF005	A CDP deve ofertar dados para conteúdo educacional e financeiro personalizado, conforme gostos e interesses do investidor
RF006	A CDP deve ofertar dados para produtos financeiros personalizados, conforme perfil investidor do cliente
RF007	A CDP deve registrar as informações e interações dos potenciais investidores nas plataformas digitais <i>web</i> e <i>app</i> conforme interesse de observabilidade do negócio
RF008	A CDP deve permitir os analistas da <i>startup</i> a terem uma visão unificada da jornada dos investidores
RF009	A CDP deve permitir segmentação de audiência, via interface gráfica, com base em interações e informações de investidores, possibilitando a criação de perfis e audiências
RF010	A CDP deve permitir dados acionáveis em seu ecossistema, isto é, ativação de audiências para testes A/B, personalização de conteúdo, régua de comunicação, veiculação de mídia e sistemas de recomendação de produtos financeiros

RF011	A CDP deve possuir catálogo de dados para os analistas consultarem as informações de seus investidores
RF012	A CDP deve permitir análise exploratória dos dados dos investidores em seu ecossistema

Fonte: autoria própria

Na Tabela 2 abaixo, a coluna “Descritivo” de cada requisito possui um rótulo destacado em negrito (exemplo: Desempenho) para auxiliar o entendimento adequado dos requisitos. O DMBOK possui uma indicação da importância de requisitos funcionais e não funcionais (DAMA INTERNATIONAL, 2017).

Tabela 2 - Requisitos não-funcionais (RNF) da CDP.

Requisito	Descritivo
RNF001	Desempenho - A CDP deve suportar a captura e oferta de dados para até 100k usuários digitais (investidores) simultâneos ao mês
RNF002	Desempenho - A CDP deve suportar captura de dados para até 10 milhões de interações por mês
RNF003	Desempenho - A CDP deve ter baixa latência de processamento, próximo ao tempo real, considerando o ciclo de vida da informação desde a captura até a ativação do dado no digital
RNF004	Disponibilidade - A CDP deve ter alta disponibilidade sistêmica de 99.99%, representando downtime máximo de aproximadamente 52 min/ano e 4 min/mês
RNF005	Manutenibilidade - A CDP deve acompanhar o nível de suporte de TI em N2, para apoio de implantação e acompanhamento de incidentes
RNF006	Manutenibilidade - A CDP deve ser orientada a micro-serviços, com baixo acoplamento entre camadas, permitindo incremento sistêmico para novas partes interessadas
RNF007	Manutenibilidade - A CDP deve ser expansível para integração com bancos de dados transacionais da startup com conectores JDBC
RNF008	Manutenibilidade - A CDP deve acompanhar os níveis de suporte de TI em N1, para apoio às dúvidas dos usuários de dados da solução
RNF009	Interoperabilidade - A CDP deve integrar-se com dispositivos digitais (Web e App) por meio de biblioteca javascript e SDKs Android e iOS para realizar captura de dados diretamente nas origens

RNF010	Interoperabilidade - A CDP deve integrar-se com as ferramentas MarTech da startup via API para elaboração de testes AB e personalização de conteúdo
RNF011	Interoperabilidade - A CDP deve integrar-se com o Data Lake da startup por meio de cargas incrementais, permitindo exportação de dados brutos ao repositório central da <i>startup</i>
RNF012	Usabilidade - A CDP deve possuir mecanismos amigáveis de manipulação de dados, dispensando conhecimento técnico dos usuários que extraem valor do seu uso
RNF013	Usabilidade - A CDP deve acompanhar treinamento para manipulação de dados, com baixa curva de aprendizagem para novos usuários
RNF014	Disponibilidade - A CDP deve oferecer durabilidade aos dados, permitindo consumo sempre que requisitado por um usuário
RNF015	Disponibilidade - A CDP deve proporcionar baixa latência de processamento, próxima à camada de tempo real, considerando todo o ciclo de vida da informação
RNF016	Portabilidade - A CDP deve oferecer mecanismos para portabilizar seus dados de repositório, considerando uma possível quebra de vínculo legal ou desvio estratégico de rota
RNF017	Segurança - A CDP deve integrar-se à plataforma de gestão de consentimento (CMP) da <i>startup</i>
RNF018	Segurança - A CDP deve possuir mecanismo de criptografia de dados sensíveis utilizando chave-própria da <i>startup</i>
RNF019	Segurança - A CDP deve possuir mecanismos de autenticação <i>OAuth</i> para os usuários que manipularem os dados do ecossistema
RNF020	Conformidade - A CDP deve permitir exportação de dados pelo investidor se esta ação for de seu interesse
RNF021	Conformidade - A CDP deve permitir a exclusão de dados pessoais do investidor se assim requisitado por ele

Fonte: autoria própria

3.2 Descrição Arquitetural

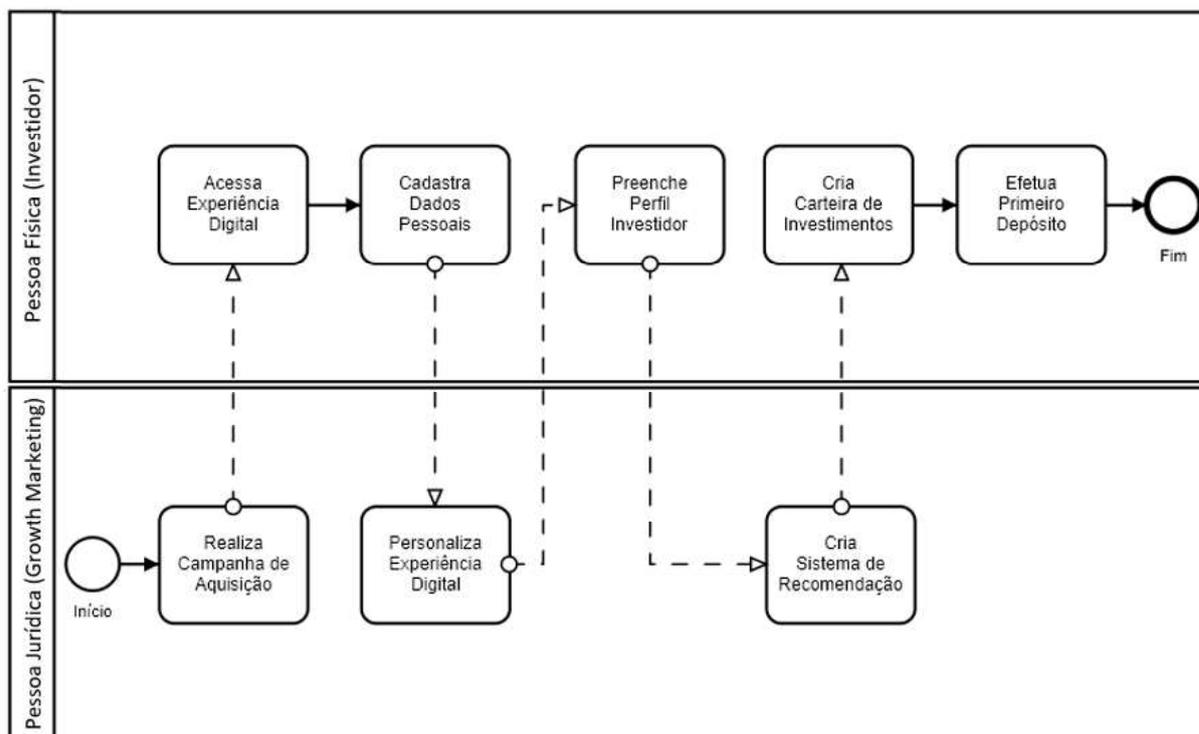
Nesta seção são apresentadas visões arquiteturais orientadas pelos *frameworks* RM-ODP e DAMA, direcionadas para um modelo de dados de cliente compatível com *software* CDP. Ressalta-se que o artefato em questão, proposto pela metodologia *Design Science*, é representado pela especificação de arquitetura das visões Empresa e Informação (PUTMAN, 2000). A análise arquitetural do objeto desta monografia respeita a metodologia *Design Science* quanto à observabilidade em contexto de aplicação. Portanto, considera-se a Arquitetura da Informação como artefato e a análise empírica subsequente como resultados (WIERINGA, 2014).

3.2.1 Visão Empresa

A Visão Empresa caracteriza-se pelo Entendimento de Negócio do contexto de aplicação da solução, no caso, o objetivo primário da Gerência de Aquisição, seu fluxo operacional, bem como a experiência digital oferecida ao investidor. Esta visão configura-se imprescindível para a compreensão do todo e, conseqüentemente, da elaboração da Arquitetura de Informação, haja vista a relação intrínseca existente entre Dados e Negócio.

Para a composição da Visão Empresa, como insumo para Arquitetura da Informação, é retratado o processo de negócio da Gerência de Aquisição através do *Business Process Model Notation*, ou BPMN, notação gráfica utilizada em metodologias e práticas de Gerenciamento de Processos, no qual ambos os atores, pessoas jurídica e física, estão inseridos. A seguir, na Figura 9, o processo pode ser compreendido mais detalhadamente, representando a transição das etapas do funil de *Growth Marketing* (descrito na seção 2.3), de Aquisição para Ativação.

Figura 11 - BPMN do fluxo de aquisição de investidores da *fintech*, segundo relato da Gerência de Aquisição da Diretoria de *Growth Marketing*



Fonte: autoria própria

Dentro do processo previamente citado, alguns indicadores-chave de desempenho, também conhecidos como *Key Performance Indicators* (KPIs), são acompanhados pela Gerência de Aquisição por direcionarem os esforços empregados pela área para atingir suas metas. A seguir, a Tabela 3 detalha os indicadores da gerência:

Tabela 3 - Descritivo dos indicadores-chave de desempenho (KPIs) da Gerência de Aquisição

KPI	Descritivo	Volume Médio Mensal
Usuários Ativos	Monitora quantos potenciais clientes da startup acessam as plataformas digitais	82k
Cadastros Realizados	Monitora quantos potenciais clientes da startup registram ao menos uma informação de contato	25k

Contas Abertas	Monitora quantos potenciais clientes da startup preenchem o Perfil Investidor para abertura de conta	5k
Primeiro Depósito Efetuado	Monitora quantos clientes são efetivados realizando primeiro depósito na conta da startup	1.5k
Valor de Primeiro Depósito	Monitora a quantia em reais (R\$) obtida por meio da ativação dos clientes ao realizar primeiro depósito	15M

Fonte: autoria própria

3.2.2 Visão Informação

A Visão Informação, principal objeto de análise desta monografia, caracteriza-se pelo Entendimento de Dados do contexto de aplicação da CDP como solução. O *framework* RM-ODP, como o próprio autor Linington (1995) diz, é pouco restritivo para as visões arquiteturais Empresa e Informação, exatamente com a finalidade de torná-lo mais flexível para diversos casos de uso (LININGTON, 1995). Por ser oriundo da Arquitetura de *Software*, o modelo de referência contribui para avaliação da construção de uma CDP.

Já do ponto de vista da Gestão de Dados, o DAMA *framework* tem uma área de conhecimento dedicada para Arquitetura de Dados, sustentada por outra área de conhecimento, a Governança de Dados, que direciona o entendimento da relação de pessoas com os dados, bem como o detalhamento das origens provedoras de informação. O aprofundamento desse entendimento permite avaliação da maturidade de dados da *fintech*, além da sua capacidade de sustentar *Big Data* e *Analytics* com soluções de *Customer Data Platform* terceiras ou nativas.

A Figura 12 retrata a governança dos dados da CDP na corretora de investimentos. Ela relaciona a origem dos dados com os papéis propostos pela Governança de Dados abordada no DMBOK, também presente na pirâmide de Peter Aiken do DAMA *Framework*, destacando cada papel.

Figura 12 - Governança dos Dados da CDP



Fonte: autoria própria

A Tabela 4 descreve as origens de dados para CDP, conforme relato da Gerência de Aquisição e da Diretoria de *Growth Marketing*. Para tal, a tabela relaciona origem, frequência de atualização dos dados, granularidade dos dados, registros/mês e volumetria média em GB/mês.

Tabela 4 - Descritivo das origens de dados da CDP

Origem	Frequência dos Dados	Granularidade dos Dados	Volumetria Média (registros/mês)	Volumetria Média (GB/mês)
Site	Tempo Real	Evento comportamental	250k	0,5GB
App Android	Tempo Real Aproximado	Evento comportamental	500k	1GB
App iOS	Tempo Real Aproximado	Evento comportamental	100k	0,2GB

Fonte: autoria própria

De acordo com Earley (2018), *Customer Data Platform* demanda um modelo de dados específico como Arquitetura de Informação para poder representar clientes, audiências e, assim, compreender as necessidades do mundo real para que as descobertas se traduzam em melhorias e mudanças. (EARLEY, 2018).

Visto ao exposto, a Tabela 5 lista os atributos de usuário, conforme o contexto de aplicação, seguindo o padrão de representação de Earley (2018). Além do

detalhamento de todos os atributos presentes no processo de negócio, há um exemplo de valor de preenchimento para cada atributo.

Tabela 5 - Lista de Atributos de Usuário para Arquitetura da CDP

Origem	Atributo	Descrição	Exemplo
Site, App Android, App iOS	<i>customer_id</i>	Define código de identificação do investidor após abertura de conta	"410310049"
Site, App Android, App iOS	<i>user_id</i>	Define código de identificação do investidor na plataforma digital	"556c5c6f-450a-4c15-8760-bd5f49b38413"
Site, App Android, App iOS	<i>full_name</i>	Define o nome completo encriptado do investidor	"33e2220e1367bae3689580e0ec3f152f893b1225f2425ded5d58cf6684dc7ffc"
Site, App Android, App iOS	<i>nickname</i>	Define o apelido escolhido pelo investidor	"Bruno Mercadante"
Site, App Android, App iOS	<i>cpf</i>	Define Cadastro de Pessoa Física (CPF) encriptado do investidor	"wCy4I7sYzAz07GS4ynAiUw"
Site, App Android, App iOS	<i>birth_date</i>	Define data de nascimento do investidor	"1994-10-04"
Site, App Android, App iOS	<i>nationality</i>	Define país de origem do investidor	"Brazil"
Site, App Android, App iOS	<i>naturalness</i>	Define cidade e estado de origem do investidor	"Sao Paulo / SP"
Site, App Android, App iOS	<i>marital_status</i>	Define estado civil do investidor	"Single"

Site, App Android, App iOS	<i>gender</i>	Define sexo do investidor segundo RG	"Male"
Site, App Android, App iOS	<i>equity</i>	Define valor total de patrimônio do investidor	"1200000.00"
Site, App Android, App iOS	<i>monthly_income</i>	Define renda mensal do investidor	"20000.00"
Site, App Android, App iOS	<i>business_segment</i>	Define classificação do segmento de cliente baseado em total de patrimônio	"Yx Investor"
Site, App Android, App iOS	<i>schooling_degree</i>	Define grau de escolaridade do investidor	"Postgraduate"
Site, App Android, App iOS	<i>financial_knowledge_degree</i>	Declaração do investidor sobre sua familiaridade com o mercado financeiro	"High"
Site, App Android, App iOS	<i>main_investment_objective</i>	Declaração do investidor sobre foco principal de investimento	"Increase Equity"
Site, App Android, App iOS	<i>withdraw_time</i>	Declaração do investidor sobre quando pretende resgatar seus investimentos	"Over 10 years"
Site, App Android, App iOS	<i>financial_changes_reaction</i>	Declaração do investidor sobre comportamento diante de oscilações financeiras	"Portfolio Reallocation"
Site, App Android, App iOS	<i>investment_frequency</i>	Declaração do investidor sobre periodicidade de investimento	"Monthly"

Site, App Android, App iOS	<i>investor_profile</i>	Define categoria do investidor baseada nas declarações do investidor	"Moderate"
Site, App Android, App iOS	<i>politically_exposed_person</i>	Define se o investidor tem patrimônio fiscalizado pelo Estado por serviço público	"No"
Site, App Android, App iOS	<i>qualified_investor</i>	Define se o investidor possui a partir de 1 milhão de reais em patrimônio	"Yes"
Site, App Android, App iOS	<i>total_deposited</i>	Define total cumulativo do valor já depositado pelo investidor	"1000000.00"

Fonte: autoria própria

Observa-se na Tabela 6 os atributos de evento representados para a *Customer Data Platform* da empresa referência.

Tabela 6 - Lista de Atributos de Evento para Arquitetura da CDP

Origem	Atributo	Descrição	Exemplo
Site, App Android, App iOS	<i>name</i>	Define o nome do evento realizado pelo investidor	"Viewed Signup"
Site, App Android, App iOS	<i>trigger</i>	Define tecnicamente o gatilho sistêmico feito pelo investidor para computar o evento	"Load"
Site, App Android, App iOS	<i>timestamp</i>	Define o tempo em milissegundos de registro do evento	"1635470942511"

Site, App Android, App iOS	<i>device</i>	Objeto para associação lógica dos atributos de dispositivo	-
Site, App Android, App iOS	<i>device_id</i>	Define código identificador do dispositivo de navegação	"07GS4ynAiUw"
Site, App Android, App iOS	<i>type</i>	Define se o dispositivo é desktop, tablet ou mobile	"Desktop"
Site, App Android, App iOS	<i>brand</i>	Define a marca do dispositivo	"Motorola"
Site, App Android, App iOS	<i>model</i>	Define o modelo de fábrica do dispositivo	"G9 Plus"
Site, App Android, App iOS	<i>operating_system</i>	Define sistema operacional do dispositivo de navegação	"Windows"
Site, App Android, App iOS	<i>operating_system_v ersion</i>	Define a versão do sistema operacional de navegação	"10"
Site	<i>browser</i>	Define o navegador de uso do investidor	"Chrome"
Site	<i>browser_version</i>	Define a versão do navegador de uso do investidor	"91"
Site, App Android, App iOS	<i>screen_resolution</i>	Define o dimensionamento da tela em pixels	"2400x1080"
Site, App Android, App iOS	<i>display_window</i>	Define o dimensionamento da área visível em pixel durante a navegação	"432x960"
Site, App Android, App iOS	<i>text_language</i>	Define a linguagem configurada pelo investidor no dispositivo	"pt-br"

Site, App Android, App iOS	<i>traffic</i>	Objeto para associação lógica dos atributos de tráfego	-
Site, App Android, App iOS	<i>source</i>	Define a origem do tráfego do usuário digital ao acessar a plataforma digital da <i>startup</i>	"Facebook"
Site, App Android, App iOS	<i>medium</i>	Define por qual meio de campanha o usuário digital acessou a plataforma digital da <i>startup</i>	"cpc"
Site, App Android, App iOS	<i>campaign</i>	Define o nome da campanha que atraiu o usuário digital no momento em que ele acessa a plataforma da <i>startup</i>	"Black Friday 2021"
Site, App Android, App iOS	<i>geolocation</i>	Objeto para associação lógica dos atributos de geolocalização	-
Site, App Android, App iOS	<i>latitude</i>	Define latitude para representação do ponto de geolocalização do investidor	"-23.5318858"
Site, App Android, App iOS	<i>longitude</i>	Define longitude para representação do ponto de geolocalização do investidor	"-46.445952"
Site, App Android, App iOS	<i>context_info</i>	Objeto para associação lógica dos atributos de contexto da experiência do investidor	-
Site, App Android, App iOS	<i>site_app_name</i>	Define o nome do site ou app utilizado pelo investidor	"Fintech Y"
Site, App Android, App iOS	<i>site_app_version</i>	Define a versão do site ou app utilizado pelo investidor	"3.2.8"

Site, App Android, App iOS	<i>domain_package</i>	Define o domínio do site ou o nome do pacote do app	<i>"fintechy.com.br"</i>
Site, App Android, App iOS	<i>page_screen</i>	Define o nome da tela de navegação do investidor	<i>"Signup"</i>
Site, App Android, App iOS	<i>page_screen_detail</i>	Define o path (em Website) e a classe (em App) da tela de navegação do investidor	<i>"/signup"</i>
Site, App Android, App iOS	<i>digital_experience</i>	Define se a experiência do investidor é B2C (autônoma) ou B2B (via consultor)	<i>"b2c"</i>
Site, App Android, App iOS	<i>digital_product</i>	Define em qual produto da fintech o investidor está navegando	<i>"Onboarding"</i>
Site, App Android, App iOS	<i>product_section</i>	Define a área do produto em que o evento foi realizado pelo investidor	<i>"Form"</i>
Site, App Android, App iOS	<i>product_target</i>	Define o conteúdo do elemento de interação, quando aplicável	<i>"Register"</i>

Fonte: autoria própria

Notam-se na Tabela 7 os eventos aplicáveis à *fintech*, com associação de seus respectivos atributos para rastreamento do comportamento do cliente via CDP.

Tabela 7 - Lista de Eventos para Arquitetura da CDP

Origem	Evento	Descrição	Atributos
Site, App Android, App iOS	<i>Viewed Campaign</i>	Ocorre quando o usuário acessa a tela da campanha	<ul style="list-style-type: none"> - <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i>

Site, App Android, App iOS	<i>Viewed Signup</i>	Ocorre quando o usuário acessa a tela de cadastro	<ul style="list-style-type: none"> - <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i>
Site, App Android, App iOS	<i>Filled Personal Data</i>	Ocorre sempre quando o usuário preenche alguma informação de cadastro, registrando seu respectivo parâmetro como atributo de usuário, tecnicamente especificado como <i>user_info</i>	<ul style="list-style-type: none"> - <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i> - <i>user_id</i> - <i>full_name</i> - <i>nickname</i> - <i>cpf</i> - <i>birth_date</i> - <i>nationality</i> - <i>naturalness</i> - <i>marital_status</i> - <i>gender</i>
Site, App Android, App iOS	<i>Finished Signup</i>	Ocorre quando o usuário completa o cadastro	<ul style="list-style-type: none"> - <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i>
Site, App Android, App iOS	<i>Viewed Suitability</i>	Ocorre quando o usuário acessa a tela do perfil investidor	<ul style="list-style-type: none"> - <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i>

Site, App Android, App iOS	<i>Answered Profile Question</i>	Ocorre sempre quando o usuário responde a uma questão do perfil investidor, registrando seu respectivo parâmetro como atributo de usuário, tecnicamente especificado como user_info	<ul style="list-style-type: none"> - name - trigger - timestamp - context_info - geolocation - traffic - device - equity - monthly_income - schooling_degree - financial_knowledge_degree - main_investment_objective - withdraw_time - financial_changes_reaction - investment_frequency - politically_exposed_person - qualified_investor
Site, App Android, App iOS	<i>Finished Suitability</i>	Ocorre quando o usuário conclui o perfil investidor	<ul style="list-style-type: none"> - name - trigger - timestamp - context_info - geolocation - traffic - device - business_segment - investor_profile
Site, App Android, App iOS	<i>Viewed Investor Profile</i>		<ul style="list-style-type: none"> - name - trigger - timestamp - context_info - geolocation - traffic - device
Site, App Android, App iOS	<i>Viewed Portfolio Builder</i>	Ocorre quando o usuário acessa a tela de criação de carteiras	<ul style="list-style-type: none"> - name - trigger - timestamp - context_info - geolocation - traffic - device

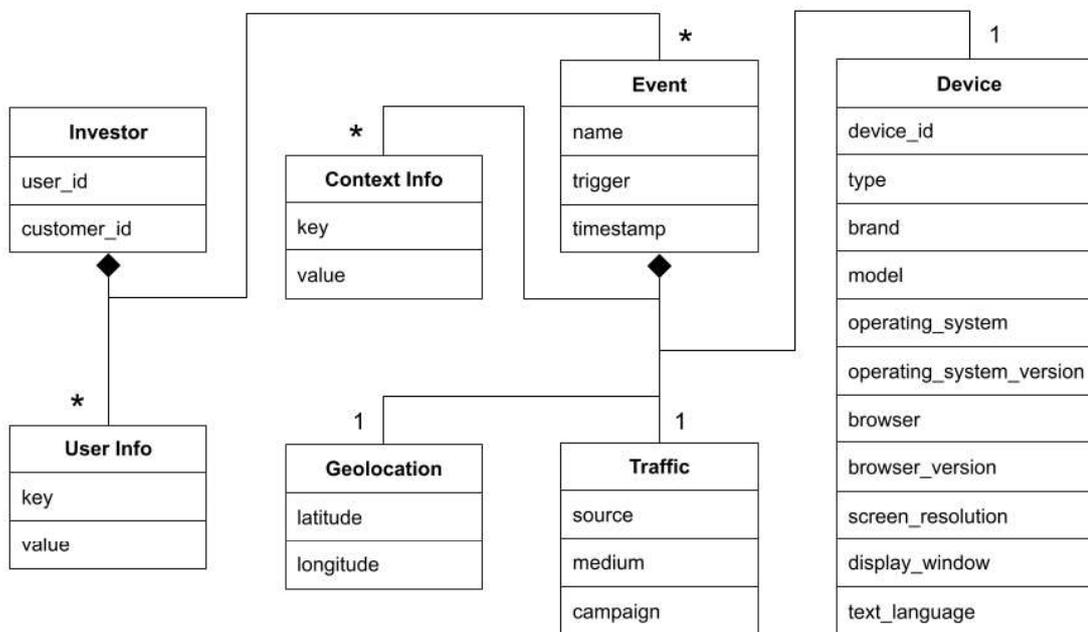
Site, App Android, App iOS	<i>Advanced Builder Step</i>	Ocorre sempre quando o usuário avança em uma etapa da criação da carteira, detalhando o respectivo complemento	- <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i>
Site, App Android, App iOS	<i>Finished Portfolio Builder</i>	Ocorre quando o usuário conclui a criação da carteira	- <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i>
Site, App Android, App iOS	<i>Viewed Banking Data</i>	Ocorre quando o usuário acessa a tela de dados bancários	- <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i>
Site, App Android, App iOS	<i>Accomplished Banking Deposit</i>	Ocorre quando o sistema detecta o depósito do usuário, registrando seu respectivo parâmetro como atributo de usuário, tecnicamente especificado como <i>user_info</i>	- <i>name</i> - <i>trigger</i> - <i>timestamp</i> - <i>context_info</i> - <i>geolocation</i> - <i>traffic</i> - <i>device</i> - <i>customer_id</i> - <i>total_deposited</i>

Fonte: autoria própria

Para transcrever as informações nos moldes de diagramação representando o modelo de dados de cliente para *Customer Data Platform*, a Figura 13 contempla os dados listados em um Modelo Orientado a Agregação, em notação UML (*Unified Modeling Language*) relatado por Fowler e Sadalage (2012) para lidar com dados NoSQL (FOWLER; SADALAGE, 2012). O modelo é característico de *Big Data*,

aplicável para a CDP, e a granularidade dos registros é a nível de Investidor, com cada investidor podendo ter múltiplos atributos e eventos, vide imagem a seguir:

Figura 13 - Modelo Orientado a Agregação para Arquitetura CDP em notação UML



Fonte: autoria própria

Outra representação aplicável ao Modelo Orientado à Agregação é a escrita em notação JSON (*JavaScript Object Notation*), presente na Figura 14, especificação capaz de apresentar informações organizadas a nível de Chave-Valor e Objetos, ambos modelos contemplados na estrutura de Agregação e, conseqüentemente, representando a modelagem CDP (FOWLER; SADALAGE, 2012).

É relevante salientar que a característica NoSQL, sem esquema predefinido (*schemaless*), permite adaptar a solução a diferentes cenários, não apenas ao contexto da *fintech*, como também aos demais produtos da *fintech*, não se restringindo à Gerência de Aquisição sendo escalável na organização, podendo se moldar às necessidades de outras gerências, como Conta Corrente, Bolsa de Valores e Carteiras de Investimento, produtos pertencentes à *Fintech Y*.

A seguir, a Figura 14 apresenta o Modelo de Dados do Cliente para a CDP, Orientado a Agregação, em formato JSON:

Figura 14 - Representação do Modelo de Agregação em formato JSON

```

{
  "investor": {
    "user_id": "556c5c6f-450a-4c15-8760-bd5f49b38413",
    "customer_id": "410310049",
    "user_info": [
      { "key": "full_name", "value": "wCy417sYzAz07GS4ynAiUw" },
      { "key": "nickname", "value": "1994-10-04" },
      { "key": "cpf", "value": "wCy417sYzAz07GS4ynAiUw" },
      { "key": "birth_date", "value": "1994-10-04" },
      { "key": "nationality", "value": "Brazil" },
      { "key": "naturalness", "value": "Sao Paulo / SP" },
      { "key": "marital_status", "value": "Single" },
      { "key": "gender", "value": "Male" },
      { "key": "equity", "value": "1200000.00" },
      { "key": "monthly_income", "value": "20000.00" },
      { "key": "business_segment", "value": "Yx Investor" },
      { "key": "schooling_degree", "value": "Postgraduate" },
      { "key": "financial_knowledge_degree", "value": "High" },
      { "key": "main_investment_objective", "value": "Increase Equity" },
      { "key": "withdrawal_time", "value": "Over 10 years" },
      { "key": "financial_changes_reaction", "value": "Portfolio Reallocation" },
      { "key": "investment_frequency", "value": "Monthly" },
      { "key": "investor_profile", "value": "Moderate" },
      { "key": "politically_exposed_person", "value": "No" },
      { "key": "qualified_investor", "value": "No" },
      { "key": "total_deposited", "value": "100000.00" }
    ],
    "event": [ {
      "name": "Viewed Signup",
      "trigger": "Load",
      "timestamp": "1646455733838",
      "device": {
        "device_id": "07GS4ynAiUw",
        "type": "Desktop",
        "brand": "Motorola",
        "model": "G9 Plus",
        "operating_system": "Windows",
        "operating_system_version": "10",
        "browser": "Chrome",
        "browser_version": "91",
        "screen_resolution": "2400x1080",
        "display_window": "432x960",
        "text_language": "en"
      },
      "traffic": {
        "source": "Facebook",
        "medium": "cpc",
        "campaign": "Black Friday 2021"
      },
      "geolocation": {
        "latitude": "-23.53164476935386",
        "longitude": "-46.443666743453306"
      },
      "context_info": [
        { "key": "site_app_name", "value": "Fintech Y" },
        { "key": "site_app_version", "value": "3.2.8" },
        { "key": "domain_package", "value": "fintechy.com.br" },
        { "key": "page_screen", "value": "Signup" },
        { "key": "page_screen_detail", "value": "/signup" },
        { "key": "digital_experience", "value": "b2c" },
        { "key": "digital_product", "value": "Onboarding" },
        { "key": "product_section", "value": "Form" },
        { "key": "product_target", "value": "Register" }
      ]
    } ]
  }
}

```

Fonte: autoria própria

3.3 Resultados

A partir da imersão no contexto de aplicação da *fintech*, mais precisamente ao cenário de *Growth Marketing*, para entendimento do fluxo de aquisição de clientes investidores, as entrevistas com as partes interessadas foram determinantes para a compreensão do modelo de negócio da companhia e de como CDP é observada como solução para a estratégia de centralidade no cliente, a começar pela análise de requisitos funcionais e não funcionais exercitada em conjunto com a *startup*.

Foi possível compreender, com a evidente análise de requisitos, que a empresa pode evoluir gradativamente seu conceito de *Customer Data Platform*, iniciando pela estruturação das ferramentas *MarTech*, especialmente para práticas de CRM e CRO. A partir do refinamento das ferramentas já em uso, integrando-as às tecnologias de CMP e TMS da *startup*, inicia-se a “Fase 1” do conceito de CDP estabelecido para a *fintech*, baseando a plataforma numa junção de recursos de múltiplas ferramentas de *MarTech*, com potencial para crescimento em escala, seja aderindo ferramenta terceira especializada ou customizando solução nativa. Todos os requisitos funcionais e não funcionais podem ser preenchidos com o devido uso em conjunto das ferramentas já existentes na *startup*.

Com o desenvolvimento das visões arquiteturais Empresa e Informação, propostas por RM-ODP *Framework* (viés *software*) e DAMA *Framework* (viés dados), direcionadas ao modelo de dados CDP, notou-se a relevância de especificar uma arquitetura em diferentes pontos de vista, principalmente de uma solução que exige diferentes partes interessadas para desenvolvimento e uso efetivo, além do grau de maturidade que a solução implica, salientado tanto pelo modelo de maturidade A&DQ da McKinsey (MARTINS et al., 2019) quanto evidenciado no *framework* piramidal de Peter Aiken do DMBOK (DAMA INTERNATIONAL, 2017), haja vista a dificuldade de se implantar CDP dentro de uma estrutura organizacional, observada no Hype Cycle do Gartner (Gartner 2016, 2021) desde 2016.

No ponto de vista Empresa, a arquitetura representada pelo mapeamento do processo em notação BPMN, da principal gerência interessada em operar a CDP, permitiu a interface dos integrantes da gerência com as demais partes interessadas

da *fintech*, facilitando a comunicação e o entendimento entre os envolvidos. Além disso, a própria gerência enxergou oportunidades de otimização de seu processo, independente da aderência ao uso de CDP ou não, uma vez que o conhecimento tácito dos profissionais foi nivelado e inteligível, dado o incentivo dos debates e da especificação arquitetural, traduzida em documentação.

Os KPIs também foram alinhados, com potencial para contribuírem ao critério de sucesso de uma Prova de Conceito (PoC) de implementação efetiva de CDP terceira, representando uma “Fase 2” de CDP dentro da organização. O Ciclo de Vida dos Dados (DATAONE, 2021) também foi revisto à medida em que os processos foram otimizados para a incorporação gradual da CDP na companhia, através do uso adequado das ferramentas de *MarTech*.

No ponto de vista Informação, a arquitetura representada pelo modelo de dados de cliente direcionado para CDP possibilitou observar um esforço até então oculto para a Gerência de Aquisição no aspecto de Arquitetura de Informação, estritamente necessário e independente da decisão de adoção de uma CDP, dado que o rastreamento de clientes em dispositivos digitais exige atividades de taxonomia, desenvolvimento e catalogação, etapas presentes no Ciclo de Vida dos Dados (DATAONE, 2021) mencionado mais detalhadamente na seção 2.3 deste documento.

Também foi possível notar maior grau de entendimento dos profissionais de dados tanto sobre o cliente investidor pessoa física e seus pontos de contato com a *fintech*, conforme o nível de detalhamento incentivado pelo modelo, quanto na distinção do escopo de atributos de evento. A modelagem de dados orientada a agregação também foi destaque, tornando mais compreensível a aplicabilidade de práticas *Big Data* no tema, uma vez que o próprio CDP Institute, apesar da vasta contribuição na difusão do tema CDP para academia e para o mercado, não aborda, do ponto de vista técnico, os desafios de arquitetura a serem superados por uma organização.

4 CONCLUSÃO

Implementar *Customer Data Platform* não é algo trivial. O termo é ainda relativamente recente para referenciar um conceito que propõe um *software* como o coração de um ecossistema de dados, já que o bem mais precioso de uma empresa é o seu cliente. Sendo assim, a centralidade no cliente com CDP tem um alto custo de investimento e, se tratando do universo de *startups*, mais precisamente *fintechs*, pode determinar o seu sucesso ou o seu fracasso.

É de suma importância que pessoas jurídicas e físicas compreendam o que é uma CDP, com clareza sobre o que ela é capaz e, principalmente, sobre o que ela não é capaz de proporcionar. Pessoas jurídicas devem estar atentas à evolução de CDP porque ela almeja e instiga o maior grau de maturidade das organizações, não só no aspecto dados, mas no aspecto digital. Apenas o exercício conjunto entre as partes interessadas para especificar a arquitetura CDP, sem considerar sua implementação efetiva, proporciona benefícios substanciais à organização. Pessoas físicas devem estar atentas também pelo simples fato da transparência e ética das empresas, já que cada vez mais dados de clientes tendem a esse destino.

As visões arquiteturais (Empresa e Informação) abstraídas de *RM-ODP Framework* e *DAMA Framework*, empregadas ao caso de gestão de audiências para personalização de conteúdo com CDP para *Growth Marketing* em *fintech*, superaram as expectativas das partes interessadas, não no sentido de implementação efetiva, mas de conscientização de que o tema exige alto nível de especialização profissional, isto é, demanda pessoas capacitadas para arquitetar uma solução robusta, além de investimento processual e cultural, desde entendimento de negócio até gerenciamento de dados, visando assim atingir a maturidade esperada. CDP como pilar de negócio e informação requer pavimentação, independente de implementação terceirizada ou customizada pela própria companhia.

4.1 Conhecimento Adquirido

Para ser concebido, o documento em questão demandou aprofundamento teórico em uma série de conteúdos que permeiam *Big Data - Inteligência na Gestão dos Dados*, a começar pela percepção de *Customer Data Platform* como um conteúdo relevante para academia e para o mercado, haja vista a expectativa criada com o termo em sua primeira menção e a dificuldade subsequente para que o conceito se estabelecesse com maturidade nas empresas. O estudo teórico permitiu domínio sobre o conceito de CDP, suas funcionalidades, suas classificações e seus casos de uso. Também proporcionou conhecimento sobre entidades referências no tema.

No aspecto arquitetural, o estudo referencial e a análise da especificação permitiram capacidade de síntese do autor, estimulando a resolução de um problema complexo, com diferentes visões arquiteturais, abrangendo seus vieses e objetivos de maneira modular para alinhamento com as devidas partes interessadas, cada qual com sua especialidade, técnica ou não. Além disso, semelhanças entre as visões arquiteturais do *RM-ODP Framework* e a área de conhecimento *Arquitetura de Dados* do *DAMA Framework* foram compreendidas.

O estudo também proporcionou maior entendimento sobre o universo de *startups*, em especial a *fintech* do contexto de aplicação. Foi possível dialogar com diferentes representantes da organização, entender sua estrutura e aprender com os mais diversos profissionais, todos visando a maturidade de dados da empresa e agregando em seus respectivos pontos. Diante do cotidiano de uma *startup* de tecnologia do segmento financeiro, o estudo diário possibilitou comparativo com pesquisas do setor, avaliando grau de maturidade digital, como o A&DQ (MARTINS et al., 2019), além de associar o crescimento orgânico da *startup* ao modelo piramidal de Peter Aiken (DAMA INTERNATIONAL, 2017).

4.2 Oportunidades

A elaboração deste documento, compilando uma problemática a ser resolvida (especificação de arquitetura CDP), com referenciais teóricos pertinentes e reconhecidos pela academia e pelo mercado, ao contexto de uma *fintech* que visa realmente adotar uso de CDP como estratégia de centralidade no cliente, além das pessoas conhecidas ao longo do desenvolvimento da monografia (professores, profissionais do mercado, pesquisadores em geral e afins), diversas oportunidades acadêmicas e profissionais se apresentam.

Em âmbito acadêmico, a expectativa é contribuir para a ciência e para a pesquisa, reconhecendo o autor como especialista pela Universidade de São Paulo. Certamente novas portas se abrirão, com parcerias, projetos, fóruns e discussões abordando o tema apresentado neste material. Em âmbito de mercado, certamente há um domínio significativo e direcionado para *Big Data* - Inteligência na Gestão dos Dados, podendo promover o autor como referência no tema para *startups*, até mesmo prestando consultoria sobre os desafios de dados que envolvem CDP.

5 REFERÊNCIAS

ALPER, P.; CURCIN, V.; FAIRWEATHER, E. **Provenance and Annotation of Data and Processes**. London: Springer, 2018. p. 71–83.

BOTTINO, C. et al. **Lei Geral de Proteção de Dados Pessoais e Resolução de Conflitos: Experiências internacionais e perspectivas para o Brasil**. Rio de Janeiro: Instituto de Tecnologia e Sociedade do Rio, 2020.

CALICCHIO, N. et al. **Brazil Digital Report**. 1.ed. [S.I.]: Mckinsey & Company/Brazil at Silicon Valley, 2019.

CDP INSTITUTE. **Learning-center. What is a CDP?** [S.I.]: Customer Data Platform Institute, 2021. Disponível em: <<https://www.cdpinstitute.org/learning-center/what-is-a-cdp/>>. Acesso em 12 set. 2021.

COLETI, T. A. **TR-Model: um perfil de aplicação de metadados para transparência de dados pessoais em aplicações de software**. 2020. 147 p. Tese (Doutorado) - Escola Politécnica, Universidade de São Paulo, São Paulo, 2020.

CVM. **Instrução CVM 506**. [S.I.]: COMISSÃO DE VALORES MOBILIÁRIOS, 2011. Disponível em: <<https://conteudo.cvm.gov.br/legislacao/instrucoes/inst506.html>>. Acesso em 11 out. 2021.

DAMA INTERNATIONAL. **DAMA - DMBOK - Data Management Body of Knowledge**. New Jersey: Technics Publications, 2017. v. 2.

DATAONE. **Data Life Cycle**. [S.I.]: Data Observation Network for Earth/University of New Mexico, 2021. Disponível em: <<https://old.dataone.org/data-life-cycle>>. Acesso em 04 out. 2021.

DISTRITO. **Distrito Fintech Report 2020**. [S.I.]: Distrito, 2020.

EARLEY, S. **The role of a Customer Data Platform**. Boston: IEEE Computer Society, 2018.

ELLIS, S.; BROWN, M. **Hacking Growth: A estratégia de marketing inovadora das empresas de crescimento mais rápido**. 1.ed. Rio de Janeiro: Alta Books, 2018.

FOWLER, M.; SADALAGE, P. J. **NoSQL Distilled: a brief guide to the emerging world of polyglot persistence**. Crawfordsville: Addison-Wesley Professional, 2012.

GARTNER. Digital Marketing. **Hype Cycle for Digital Marketing**. [S.I.]: Gartner, 2016.

GARTNER. Digital Marketing. **Hype Cycle for Digital Marketing**. [S.I.]: Gartner, 2021.

GUPTA, S. et al. Digital Analytics: Modeling for Insights and New Methods. **Journal of Interactive Marketing**, 28 jun. 2020. J. 51, 2020, p.26-43. <https://doi.org/10.1016/j.intmar.2020.04.003>

IBM INSTITUTE FOR BUSINESS VALUE. Chief Marketing Officer. **The modern marketing mandate**. [S.I.]: IBM Institute for Business Value, 2018.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC 10746-3: Information technology - Open distributed processing - Reference model: Architecture - Part 3**. ISO, 2009.

KHAN, R.; HASAN, R. Fuzzy Authentication Using Interaction Provenance in Service Oriented Computing. **Proceedings - 2015 IEEE International Conference on Services Computing**, SCC 2015, n. October 2020, p. 170–177, 2015. DOI: 10.1109/SCC.2015.32.

LEMNISK. **116 Customer data platform use cases**. [S.I.]: Lemnisk, 2020.

LININGTON, P. F. **RM-ODP: The Architecture**. Canterbury: University of Kent, 1995.

MARTINS, E.; FUKUMOTO, M. S. **Uso de dados na era pós third-party cookies**. [S.I.]: IAB Brasil, 2021. Disponível em: [<https://iabbrasil.com.br/artigo-uso-de-dados-na-era-pos-third-party-cookies/>](https://iabbrasil.com.br/artigo-uso-de-dados-na-era-pos-third-party-cookies/). Acesso em 04 out. 2021.

MARTINS, H. et al. **Transformações digitais no Brasil: Insights sobre o nível de maturidade digital das empresas no país.** [S.I.]: Mckinsey & Company, 2019.

NIST. **Standard Reference Data.** [S.I.]: National Institute of Standards and Technology, 2021. Disponível em: <<https://www.nist.gov/srd>>. Acesso em 05 out. 2021.

PADÍN, S. G. **Cookies first-party e third-party: quais as diferenças?.** [S.I.]: IAB Brasil, 2021. Disponível em: <<https://iabbrasil.com.br/artigo-cookies-first-party-e-third-party-quais-as-diferencas/>>. Acesso em 04 out. 2021.

PUTMAN, J. R. **Architecting with Rm-Odp.** 1.ed. [S.I.]: Prentice Hall, 2000.

SANTOS C. et al. **Consent Management Platforms Under the GDPR: Processors and/or Controllers?.** London: Springer, 2021.

SEBRAE. **O que é Startup.** [S.I.]: Sebrae, 2021. Disponível em: <<https://www.sebrae.com.br/sites/PortalSebrae/artigos/o-que-e-uma-startup,6979b2a178c83410VgnVCM1000003b74010aRCRD>>. Acesso em 11 out. 2021.

SILVA, M. A. T. et al. Arquitetura da Informação para quem e para quem?: uma reflexão a partir da prática em ambientes informacionais digitais. **Encontros Bibli: Revista eletrônica De Biblioteconomia E Ciência Da informação**, Santa Catarina, v. 18, n. 37, p. 283-302, mai./ago., 2013. DOI: 10.5007/1518-2924.2013v18n37p283

SIQUEIRA, A. G. C. **Persona: o que é, como definir e por que criar uma para sua empresa.** [S.I.]: Resultados Digitais, 2013. Disponível em: <<https://resultadosdigitais.com.br/marketing/persona-o-que-e/>>. Acesso em 15 fev. 2022.

TAIL. **O que é CDP: Customer Data Platform.** [S.I.]: Tail, 2020.

W3C. **PROV-DM: The PROV Data Model.** [S.I.]: W3C, 2013. Disponível em: <<https://www.w3.org/TR/prov-dm/>>. Acesso em 19 nov. 2021.

WIERINGA, R. J. **Design science methodology for information systems and software engineering**. London: Springer, 2014.