

TECNOLOGIAS E DADOS
ABERTOS PARA INOVAÇÃO
EM GOVERNO



Carlos Denner dos Santos Junior
Mario Antonio Ribeiro Dantas
Rogerio Atem de Carvalho
Hélio Lemes Costa Júnior
Letícia Almeida Amaral



Tecnologias e dados abertos para inovação em governo

Crowdsourcing de Dados Abertos Governamentais: Superando o ‘Modelo da Disponibilização’ com Governança Coletiva

Carlos Denner dos Santos Junior

IoT-APP (IoT – Armazenamento e Processamento na área Pública)

Mario Antonio Ribeiro Dantas

Carteira de Cursos Baseada em Tecnologia Blockchain

Rogério Atem de Carvalho

Mobiliza 360: Um experimento com aplicação de realidade virtual como instrumento de empatia para formulação de políticas públicas

Hélio Lemes Costa Júnior e Letícia Almeida Amaral

Expediente



Presidente

Diogo Godinho Ramos Costa

Diretora Executiva

Bruna Santos

Diretor de Desenvolvimento Profissional

Paulo Marques

Diretor de Gestão Interna

Alana Regina Biagi Silva Lisboa

Diretoria de Altos Estudos

Diana Coutinho

Diretor de Educação Executiva

Rodrigo Torres

Diretoria de Inovação

Bruna Santos

Coordenadora-Geral de Inovação

Marizaura Reis de Souza Camões

Equipe GNova – Laboratório de Inovação em Governo

Andrea Marina Lins Lacerda

Carolina Sólía Nasser

Elisabete Ferrarezi

João Augusto Sobreiro Sigora

Joselene Pereira Lemos

Letícia Koepfel Mendonça

Cecília Lariú

Rafael Dick de Brito Dias

Coordenação do Programa Cátedras

Elisabete Ferrarezi

Andrea Marina Lins Lacerda

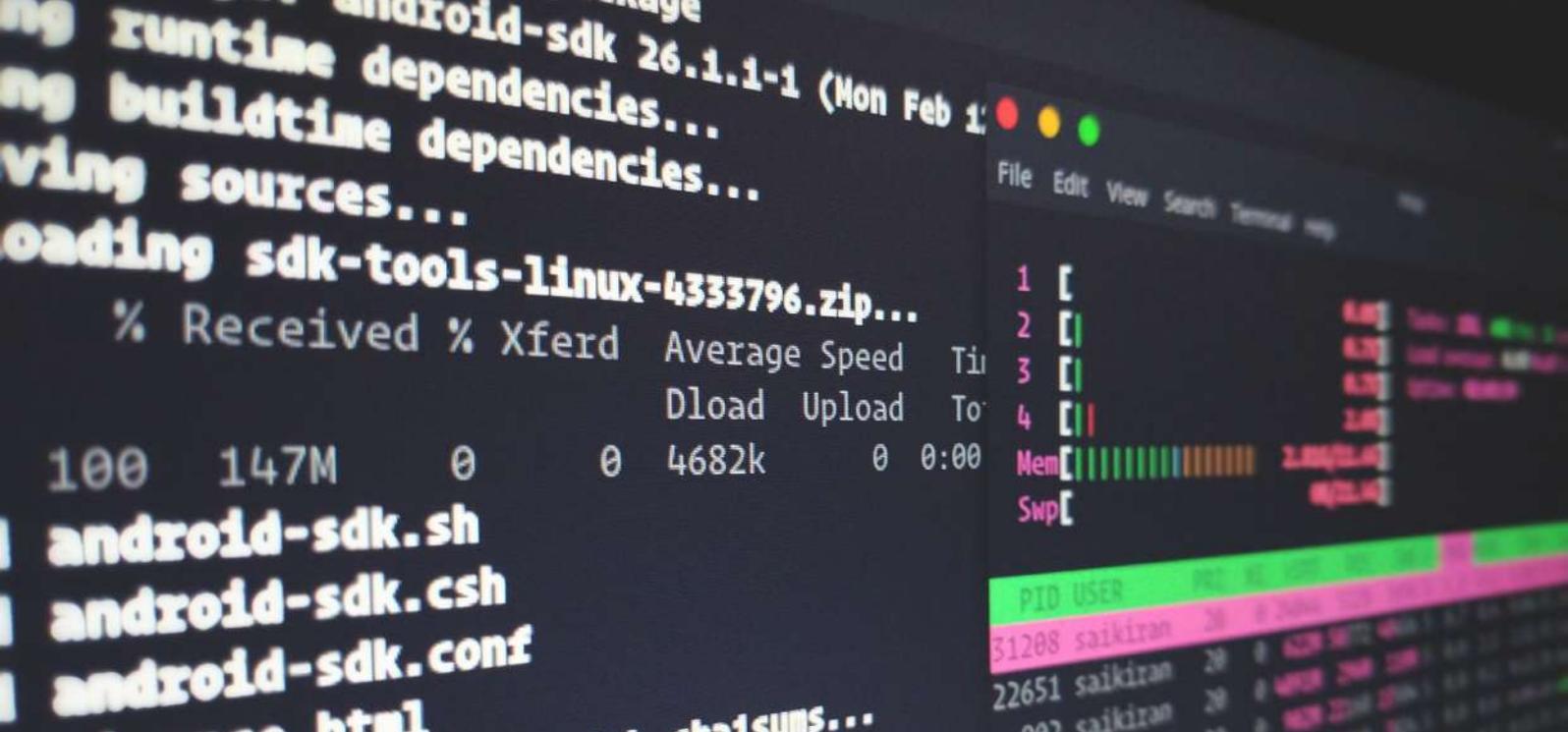
Revisão

Renata Fernandes Mourão

Luiz Augusto Barros de Matos

Projeto Gráfico

Ana Carla Cardoso Gualberto



Tecnologias e dados abertos para inovação em governo

Crowdsourcing de Dados Abertos Governamentais: Superando o 'Modelo da Disponibilização' com Governança Coletiva

Carlos Denner dos Santos Junior

IoT-APP (IoT – Armazenamento e Processamento na área Pública)

Mario Antonio Ribeiro Dantas

Carteira de Cursos Baseada em Tecnologia Blockchain

Rogério Atem de Carvalho

Mobiliza 360: Um experimento com aplicação de realidade virtual como instrumento de empatia para formulação de políticas públicas

Hélio Lemes Costa Júnior e Letícia Almeida Amaral

F374m Ferrarezi, Elisabete
Tecnologia e dados abertos para inovação em governo /
Organizado por Elisabete Ferrarezi, Marina Lins Lacerda. -- Brasília:
Enap, 2021.
102 p. : il. -- (Cadernos Enap, 73)

Inclui bibliografia.
ISSN: 0104-7078

1. Inovação. 2. Dados Abertos. 3. Administração Pública Federal. 4. Primeiro Setor. 5. Ciência e Tecnologia. 6. Políticas Públicas. I. Título. II. Lacerda, Marina Lins. III. Dantas, Mario Antonio Ribeiro. IV. Carvalho, Rogério Atem de. V. Costa Junior, Hélio Lemes. VI. Amaral, Leticia Almeida.

CDU 35:001.895

Ficha catalográfica elaborada por: Tatiane de Oliveira Dias – CRB1/2230

 Enap, 2021

Este trabalho está sob a Licença *Creative Commons* – Atribuição: Não Comercial – Compartilha Igual 4.0 Internacional. As informações e opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista da Escola Nacional de Administração Pública (Enap). É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

Escola Nacional de Administração Pública (Enap)

Diretoria de Pesquisa e Pós-Graduação

Coordenação-Geral de Pesquisa

SAIS – Área 2-A – 70610-900 — Brasília-DF, Brasil



Sumário

Sobre o Gnova.....	8
Apresentação	9
Texto I – Crowdsourcing de Dados Abertos Governamentais: Superando o ‘Modelo da Disponibilização’ com Governança Coletiva.....	10
Texto II – IoT-APP (IoT – Armazenamento & Processamento na área Pública)	34
Texto III – Carteira de cursos baseada em tecnologia blockchain	62
Texto IV – Mobiliza 360: Um experimento com aplicação de realidade virtual como instrumento de empatia para formulação de políticas públicas	83



SOBRE O GNOVA

O Laboratório de Inovação em Governo – GNova, criado em 2016, tem como missão desenvolver soluções inovadoras em projetos com instituições do governo federal para que o serviço público possa melhor responder às demandas da sociedade.

Uma das principais diretrizes do GNova é contribuir para mudar o modo como o Estado se relaciona com os cidadãos na oferta de serviços públicos, colocando o foco nas pessoas. Isso significa reconhecer os problemas e as necessidades dos usuários de serviços e políticas públicas.

Em seus projetos, o GNova utiliza metodologias ágeis e abordagens multidisciplinares inspiradas no design, nas ciências sociais e na economia comportamental e atua em três eixos, prospecção, experimentação e disseminação de inovação em serviços e políticas públicas, com o objetivo de promover a cultura e a prática da inovação na administração pública.

Missão: Promover a inovação no setor público para melhor responder às demandas da sociedade

Visão: inovação como prática transformadora no setor público.

Valores: colaboração, proatividade, abertura ao risco, atuação em rede, empatia e foco no usuário, experimentação e geração de valor público.

Para saber mais, acesse: gnova.enap.gov.br

APRESENTAÇÃO

Esta publicação apresenta os resultados de quatro pesquisas aplicadas que exploram possibilidades de inovação no setor público utilizando novas tecnologias e dados abertos. Os trabalhos foram desenvolvidos no âmbito do Programa Cátedras Brasil Inovação, referente ao edital nº 5/2018, organizado pelo Laboratório de Inovação em Governo - GNova da Escola Nacional de Administração Pública - Enap.

Uma das principais ações de prospecção do GNova é o Programa Cátedras Brasil Inovação, que tem como objetivo fomentar o desenvolvimento de pesquisas aplicadas ao setor público que contribuam para a inovação na gestão e no desenho das políticas públicas. O programa também possibilita à Enap aproximar-se de inovadores com atuação nas universidades e no setor privado, bem como identificar novos campos de atuação.

A proposta é que a pesquisa aplicada contribua para melhor entendimento e enfrentamento de problemas públicos por meio de diagnósticos, protótipos de soluções, novas tecnologias e metodologias que promovam impactos positivos na realidade do setor público e gerem valor para a sociedade.

O programa ofereceu bolsas na modalidade pesquisa, com duração de 12 meses, cujos produtos foram relatórios desenvolvidos por pesquisadores a partir de projetos apresentados para áreas temáticas específicas. Já na modalidade inovação o foco do fomento consistiu em apoiar projetos, com duração de seis meses, para desenvolvimento de protótipo de solução relacionado a serviços públicos.

A seleção de projetos ocorreu por meio de editais públicos, com várias etapas de seleção, contando com comissão de avaliação e realização de entrevistas. No edital 5 de 2018, foram selecionados sete projetos na modalidade pesquisa – nas temáticas Design e inovação no setor público e Perspectivas comportamentais aplicadas ao aprimoramento de programas governamentais e políticas públicas – e dois projetos na modalidade protótipos de inovação. No edital 39 de 2018, feito em parceria com a Escola da Advocacia Geral da União e voltado para bolsas de inovação jurídica, foram selecionados cinco projetos de inovação.

Os pesquisadores contaram com supervisores do GNova acompanhando o andamento dos projetos e seus produtos para que o resultado estivesse mais próximo possível da produção de conhecimento aplicado à administração pública e ajustes fossem feitos tempestivamente. Esse processo possibilitou interlocução fértil entre a equipe e pesquisadores e forneceu apoio logístico para a realização de oficinas de testes de protótipos.

Os temas das pesquisas estão agrupados em seis áreas temáticas e todas serão publicadas:

Laboratórios de Inovação no Setor Público: Mapeamento e diagnóstico de experiências nacionais

Economia Comportamental

Soluções Comportamentais para Adesão a Planos de Previdência Complementar: Aplicações Baseadas na Economia Comportamental

Estratégia-aplicativo: insights comportamentais para poupar

Mapa de evidências em políticas públicas: desafios e reflexões sobre a metodologia

Protótipo de Mapa de evidências para sistemas prisionais: Reflexões sobre a metodologia aplicada

Políticas Informadas por Evidências: barreiras e intervenções

Do lado de dentro: desafios relativos às pesquisas sobre prisões no Brasil

Tecnologias e dados abertos para inovação em governo

IoT-APP (IoT – Armazenamento & Processamento na área Pública)

Carteira de Cursos Baseada em Tecnologia Blockchain

Crowdsourcing de Dados Abertos Governamentais: Superando o ‘Modelo da Disponibilização’ com Governança Coletiva

Mobiliza 360: análise do uso de realidade virtual na construção de empatia

Inovação na área jurídica: ciência de dados e custo oportunidade

Ciência de dados aplicada à análise dos custos associados aos recursos interpostos pela Advocacia Pública Federal

Sistema de apoio à análise de custo de oportunidade em processos jurídicos, baseado em Inteligência Computacional

Judicialização do Benefício de Prestação Continuada (BPC) – uma proposta de fluxograma de gestão processual-administrativa do benefício

Corrupção e composição penal, cível e administrativa: competências e atuação articulada

Sistema Nacional de Ciência, Tecnologia e Inovação

O Novo Sistema Nacional de Ciência, Tecnologia e Inovação: estratégias interdisciplinares para sua implantação e regulamentação.

O primeiro texto desta publicação apresenta “Crowdsourcing de Dados Abertos Governamentais: Superando o Modelo da Disponibilização com Governança Coletiva”, de Carlos Denner dos Santos Junior. O trabalho propõe uma governança de dados de modo mais ágil e colaborativo, em formato de plataforma, co-produzindo e co-gerenciando bases de dados para que possam gerar mais valor público e científico para a sociedade civil e a administração pública.

A pesquisa “IoT-APP: Armazenamento e Processamento na área Pública”, de Mario Antonio Ribeiro Dantas, estudou cenários relativos à automação para armazenamento e análise de grandes volumes de dados abertos em configurações de IoT para a área de atenção primária de saúde pública. Ao final do trabalho, foi feito um teste do protótipo em um cenário orientado à prestação de apoio a uma unidade pública primária de saúde.

A Carteira de Cursos Baseada em Tecnologia Blockchain, de Rogério Atem de Carvalho, apresentou um protótipo de um sistema de emissão de certificados de conclusão de cursos que pode ser integrado ao atual sistema de gestão acadêmica da Enap, bem como a de outras instituições que venham a se integrar a uma rede de emissão e certificação eletrônicas. Para tanto, foi empregada a tecnologia Blockchain, que consiste em livros de registros distribuídos e eletrônicos, capazes de interligar diferentes instituições e sistemas de gestão, possibilitando que qualquer ator, seja egresso, seja uma organização, verifique a autenticidade do certificado de uma maneira segura, rápida e de baixo custo. Essas quatro pesquisas contaram com o acompanhamento de João Augusto Sobreiro Sigora da equipe do GNova.

A última pesquisa é o “Mobiliza 360: Um experimento com aplicação de realidade virtual como instrumento de empatia para formulação de políticas públicas”, de Hélio Lemes Costa Júnior e Letícia Almeida Amaral, que contou com a supervisão de Manuel Bonduki. A realidade virtual, que ainda é pouco utilizada no setor público, possibilita acesso a realidades sociais que dificilmente servidores que elaboram políticas públicas têm acesso. O trabalho comparou como o uso da realidade virtual influencia na construção de empatia a curto e médio prazos e teve como alvo pessoas em situação de rua. Para verificar os efeitos a curto prazo, foram realizados dois estudos, o primeiro utilizando um vídeo 360° e o segundo aplicando uma narrativa escrita.

Com a publicação das pesquisas, a Enap almeja que os resultados gerados sejam disseminados para pesquisadores e servidores para que possam ser adaptados e utilizados tanto na resolução de problemas

na gestão das organizações e nas políticas públicas como no fortalecimento da geração de conhecimento sobre o setor público brasileiro.

CROWDSOURCING DE DADOS ABERTOS GOVERNAMENTAIS: SUPERANDO O ‘MODELO DA DISPONIBILIZAÇÃO’ COM GOVERNANÇA COLETIVA

Resumo

O acesso aos dados governamentais tem sido amplamente demandado pela sociedade. Os dispositivos legais brasileiros e de muitos outros países já existem e definem o acesso a esses dados como um direito fundamental da sociedade civil e um dever do Estado. Esse contexto social é amplificado pelas tecnologias da informação e comunicação e os novos modelos de negócios em startups e organizações não-governamentais. Mas o potencial desses dados está claramente sendo limitado pela prática simples de disponibilização, através da publicação individual e estática dos dados em *sites* ou repositórios.

Destaca-se em plataformas de dados abertos que além de se ‘disponibilizar os dados’, pode-se promover um ecossistema colaborativo de ‘curadoria’ e co-produção em um ambiente que promova mais valor social e econômico ao longo do tempo, à medida que novos colaboradores se juntam a esses ‘projetos de dados’ e geram o efeito de rede. A promoção das políticas públicas de acesso à informação ainda não alcançou os níveis de maturidade presentes em iniciativas da administração pública inspiradas no desenvolvimento de software livre, como o Software Público Brasileiro, por exemplo.

Esta pesquisa propõe aspectos de governança de dados que possa atender a sociedade civil e a administração pública de uma maneira mais moderna e colaborativa, em formato de plataforma, co-produzindo e co-gerenciando bases de dados para que possam gerar ainda mais valor público e científico no coletivo.

Prof. Carlos Denner dos Santos Jr.

Pós-doutor em ciência da computação. Doutor em sistemas de informação. Mestre em administração. Atua nas temáticas de dados abertos e software livre, propriedade intelectual e governança. Esta pesquisa foi realizada com a colaboração de Dyego Alves da Silva (PPGA/UNB, doutorando), e o apoio de Raíssa Paiva Pires (FACE/UNB, bolsista IC DPP/UnB), Edilson Niehues Rodrigues Lima (FACE/UNB, bolsista IC DPI/UnB) e Ateldy Brasil Filho (FGA/UNB, iniciação científica). O SOCIE-DADOS (Laboratório de P&D em Dados Abertos e Software Livre do CNPq/UnB) tem o apoio do ADM, e do PPGA da UnB também.

1. INTRODUÇÃO

O Estado moderno possui uma arquitetura histórica que pode ser analisada na perspectiva da concentração monopolizada do poder e na concentração desse poder dentro das fronteiras de um território definido (WEBER, 1994). A concentração de poder é somente conseguida com o acúmulo de informações (administrativas, geográficas, epidemiológicas etc.) pelo Estado, por esse motivo o Estado moderno é caracterizado por um corpo de funcionários especializados, capazes de manipular essas informações de maneira a produzir os instrumentos de dominação legítima, assim pode-se entender as “sociedades modernas” como “sociedades de informação” (GIDDENS, 1985).

Apesar da gestão da informação ser um processo elementar nos Estados, apenas em períodos recentes, no panorama nacional e internacional, uma importante atividade desse processo está sendo implementada, que é possibilitar à população o acesso às informações. No Brasil, o marco principal desse direito está estabelecido na Constituição Federal (CF) de 1988, que estabelece o acesso à informação como um direito fundamental (BRASIL, 1988).

O marco seguinte, na perspectiva das legislações brasileiras, é a Lei nº 12.527, de 2011 (BRASIL, 2011), conhecida como Lei de Acesso à Informação, que é uma legislação infraconstitucional que regulamenta efetivamente o direito ao acesso à informação. Como desdobramento dessa lei, diversos meios, políticas públicas e plataformas foram desenvolvidos para promover o acesso à informação, como: portais de transparência, manuais, instruções normativas e bases de dados. Destaca-se sobre a necessidade de se desenvolver políticas públicas que contemplem esses dispositivos da Lei de Acesso à Informação, pois sem a adequada política não haverá a efetivação desse direito.

Paralelamente a legislações, oriundas do Poder Legislativo, ocorreram marcos dentro dos órgãos do Executivo federal, destacando-se inicialmente a Portaria Interministerial nº 140 de 2006, do então Ministério de Estado do Controle e da Transparência com o do Ministério de Estado do Planejamento, Orçamento e Gestão. A portaria disciplinava sobre a divulgação de dados e informações pelos órgãos e entidades da administração pública federal, por meio da web, nas páginas de transparência. Essas tinham como objetivo o de possibilitar o acesso aos gastos públicos e incentivar o controle social para que as práticas da administração pública sejam disciplinadas pelos princípios da administração pública (CGU, 2006).

Em 2011, o Brasil em conjunto com oito países fundou a Parceria para o Governo Aberto (em inglês, *Open Government Partnership* – OGP), que é uma iniciativa internacional que pretende difundir e incentivar mundialmente práticas governamentais relacionadas à transparência dos governos, ao acesso à informação pública e à participação social. Como membro da OGP o Brasil já implementou várias iniciativas de governo aberto, que se encontram nos mais variados níveis de maturidade.

Algumas iniciativas promoveram mudanças legais e administrativas que transformaram significativamente a gestão pública do país nos últimos anos, assim o Brasil possui papel de destaque no cenário internacional no que diz respeito ao tema (BRASIL, 2018).

Há diversos benefícios das estratégias de abertura de dados, tais como: a melhoria dos serviços públicos; mais entendimento das atividades governamentais; o gerenciamento mais efetivo dos recursos públicos; o aumento da responsabilização e da prestação de contas; o aumento da integridade pública; a criação de comunidades mais seguras; e, uma maior participação do cidadão na gestão pública (OGP, 2011).

Em cumprimento a essas diretrizes e ordenamentos jurídicos, o governo brasileiro disponibiliza diversas bases de dados e presta vários serviços, a partir de portais de transparência pública, dos sites das organizações e de repositórios de dados, como o Portal Brasileiro de Dados Abertos (PBDA). Disponibilizando, assim, informações sobre os gastos públicos, referentes a servidores, bens, programas, ações governamentais, prestadores de serviços, licitações, contratos, entre muitas outras informações. Essas bases de dados consolidam milhões de informações oriundas das mais variadas organizações governamentais, é nesse contexto que se apresenta o Big Data, que são os repositórios de grandes quantidades de dados.

Essa conjuntura gera e disponibiliza um enorme e crescente volume de dados, sendo de grande complexidade a governança de tal contexto. Destaca-se o PBDA, que é uma ferramenta que disponibiliza o acesso livre de dados e das informações públicas (PBDA, 2018).

O PBDA tem ainda o objetivo de promover a interação entre atores da sociedade e o governo para que juntos possam estudar a melhor forma de utilização dos dados, promovendo impactos positivos sob os pontos de vista social e econômico (PBDA, 2018). Entretanto, a governança dessa política pública não está sendo satisfeita, pois o PBDA, e todos os outros portais do governo brasileiro são fontes estáticas, que apenas disponibilizam os dados, promovem uma baixa interação e colaboração entre as partes interessadas, sendo este um problema de base fundamentalmente de governança.

Nesse contexto, este projeto de pesquisa se propõe a identificar e propor soluções para plataforma com uma governança coletiva para dados abertos governamentais que viabilize a efetiva interação entre as partes interessadas a fim de promover pesquisas, estudos, relatórios e bases mais eficientes e efetivas. Sendo altamente relevante a interação das partes interessadas, buscando promover resultados mais rápidos e expressivos diante da grande diversidade e volume de dados.

As interações entre as partes interessadas serão propostas na perspectiva do *crowdsourcing*, que é o desenvolvimento de redes de colaboração externas à organização para geração de ideias e execução de atividades (WESTERSKI; IGLESIAS; GARCIA, 2012; ZHU; DJURJAGINA; LEKER, 2014).



A utilização do conhecimento externo à organização proporciona uma sinergia coletiva possibilitando maximizar as ideias e a execução de atividades. Diversas organizações têm utilizado o *crowdsourcing* para desenvolver soluções direcionadas a problemas, abrangendo o desenvolvimento de novas tecnologias (KOSOVA; NIEDERBERGER; KOSOVA, 2012).

Há diversas bases de dados nas ciências biológicas e saúde que são geridas com similaridade à perspectiva do *crowdsourcing*; um modelo exemplar desta estrutura é o GenBank, que é um banco de dados localizado no *National Institutes of Health* (NIH), que armazena informação sobre sequências nucleodídicas de aproximadamente 260.000 espécies (BENSON *et al*, 2013). O GenBank faz parte de uma rede de colaboração juntamente com o *European Molecular Biology Laboratory* (EMBL) e o *DNA DataBank of Japan* (DDBJ). Juntos, esses três bancos formam a *International Nucleotide Sequence Database Collaboration* (INSDC), armazenando e trocando informações para reunir as sequências nucleotídicas depositadas nesses bancos e garantir que essas sequências sejam acessadas em todo o mundo. O GenBank pode ser acessado por toda a comunidade científica sem qualquer restrição, e seus dados são implementados e atualizados a partir da submissão dos pesquisadores de todo o mundo (GENEBIO, 2018).

Há ainda o portal Kaggle que é uma plataforma para competições de análise de dados e de modelos preditivos. Neste portal as empresas, governos e pesquisadores publicam problemas que possuem, juntamente com os dados a serem analisados, e, assim, estatísticos de todo

o mundo competem entre si para produzir os melhores modelos, criando um ambiente de colaboração entre cientistas de dados e o vencedor da competição ganha uma premiação financeira. Em junho de 2017, o portal atingiu a marca de 1 milhão de usuários, o que demonstra a magnitude da comunidade envolvida e o potencial das descobertas. A plataforma oferece um ambiente completo para cientistas de dados desde tutoriais básicos, *softwares*, algoritmos, livros, cursos, a possibilidade de formação de equipes de trabalho e diversas competições, promovidas tanto por entes públicos como pela iniciativa privada (KAGGLE, 2017).

Na academia internacional existe iniciativas que buscam o desenvolvimento de ambientes interativos de pesquisas em ciência dos dados, como o Ambiente de Ciência de Dados Moore-Sloan (em inglês Moore-Sloan *Data Science Environments*-MSDSE) formado pela Universidade de Nova York, a Universidade da Califórnia de Berkeley e a Universidade de Washington. O objetivo desta parceria é avançar na descoberta científica intensiva em dados, capacitando pesquisadores para serem muito mais efetivos, utilizando novos métodos, novas ferramentas, novas parcerias e novos caminhos de carreira, através da criação de um “Ambiente de Ciência da Informação” em cada uma das três universidades. Existem outros programas e centros com objetivos similares a estes, como o MSDSE que lista 14 iniciativas em universidades americanas (MSDSE, 2018).

Ainda no âmbito internacional, um grupo de pesquisa, composto por mais de 50 pesquisadores de mais de

30 organizações, liderados por pesquisadores da Universidade de Oxford e da Universidade de Harvard, desenvolveram um padrão comum que permite aos cientistas compartilhar dados de diferentes bancos de dados em áreas que vão desde genética a estudos ambientais. Este padrão possibilita aos pesquisadores de diferentes campos coordenarem as descobertas uns dos outros, combinando diferentes conjuntos de dados (EUREKALERT, 2012).

No Brasil, dentro das universidades, existem apenas grupos pequenos de pesquisas e iniciativas isoladas que atuam na promoção de ambientes interativos de pesquisas de ciência dos dados. Destaca-se, ainda, que potencialmente toda publicação acadêmica, tese ou dissertação, tem uma base de dados associada, mas normalmente ela não é disponibilizada conjuntamente. Os repositórios das universidades públicas também são estáticos em conteúdo científico.

1.1 Problema de pesquisa

O Brasil faz parte do rol de nações que recentemente incluíram em seus ordenamentos jurídicos o direito ao acesso às informações públicas à sociedade civil e estão implementando novas políticas públicas a fim de possibilitar e viabilizar esse direito. O principal meio de promoção de acesso é a web, a partir de portais de transparência e da disponibilização de bases de dados com as mais variadas informações. Destaca-se, nesse contexto, o PBDA que disponibiliza acesso livre à diversas bases de dados sobre inúmeros órgãos públicas, como: agências reguladoras, bancos públicos, instituições do Poder Legislativo, empresas públicas, instituições de ensino entre muitos outros órgãos, que ao todo somam 102 organizações (PBDA, 2018).

O PBDA é de responsabilidade da Infraestrutura Nacional de Dados Abertos (INDA) que é uma política pública para garantir e facilitar o acesso pelos cidadãos, pela sociedade e, em especial, pelas diversas instâncias do setor público, aos dados e informações produzidas ou custodiadas pelo Poder Executivo federal. Assim, a INDA é a responsável por disponibilizar e estabelecer diretrizes para o desenvolvimento, implantação, manutenção e gestão do PBDA (IN4, 2018).

Esta política foi instituída pela Instrução Normativa nº 4 de 2012 (IN4/2012), publicada pela Secretaria de Logística e Tecnologia da Informação do Ministério do Planejamento, Orçamento e Gestão. A IN4/2012 estabelece em seu artigo 1º diversos objetivos, dentre eles o “VII - promover a colaboração entre governos dos diferentes níveis da Federação e entre o Poder Executivo federal e a sociedade, por meio da publicação e do reuso de dados abertos” e o “X - promover a participação social na construção de um ecossistema de reuso e de agregação de valor dos dados públicos” (IN4, 2012). A atual estrutura do PBDA não está satisfazendo de forma apropriada esses dois objetivos, pois o modelo de governança e as ferramentas de colaboração utilizadas para conectar Estado e sociedade são limitadas e estáticas.

1.2 Contextualização e justificativa

Cidadãos bem informados têm melhores condições de conhecer e acessar seus direitos como saúde, educação, benefícios sociais, e de participar de forma ativa no processo de *accountability*. Nessa perspectiva, o acesso às informações, documentos, arquivos bases de dados públicos têm sido cada vez mais reconhecido como um direito em várias partes do mundo; cerca de 90 países possuem leis que o regulamentam (BRASIL, 2011).

O acesso à informação como direito fundamental também é reconhecido por importantes organismos da comunidade internacional, como a Organização das Nações Unidas (ONU) e a Organização dos Estados Americanos (OEA), e diversos tratados, convenções e declarações que o Brasil participa deliberam sobre este direito, como a Declaração Universal dos Direitos Humanos, a Convenção das Nações Unidas contra a Corrupção, a Declaração Interamericana de Princípios de Liberdade de Expressão e o Pacto Internacional dos Direitos Civis e Políticos.

No ordenamento jurídico brasileiro o acesso às informações sob a gestão de órgãos e entidades públicas é um direito fundamental do cidadão e dever do Estado, e está explícito na Constituição Federal de 1988 (BRASIL, 1988) e regulamentado pela Lei de Acesso à Informação (BRASIL, 2011). Em regra, as informações sempre são públicas, devendo o acesso a elas serem restringidos apenas em situações exclusivas e justificáveis.

Assim, as informações e dados geridos pelo Estado em nome da sociedade são um bem público. O acesso a estas e a promoção de meios para o seu uso constituem um dos fundamentos para a consolidação da democracia, ao fomentar meios e a capacitação dos indivíduos de participarem de modo efetivo na gestão pública (BRASIL, 2011).

Esse contexto converge com os ideais do Governo Aberto, que é a forma de comunicação aberta e permanente entre os entes públicos e a sociedade, com base na transparência da administração e no desenvolvimento de trabalhos conjuntos entre governo, sociedade civil e empresarial (FERRER; PESET, 2011).

Como dito, no âmbito da administração pública federal brasileira a principal iniciativa de promoção do governo aberto é o PBDA que atualmente tem dados de 102 organizações públicas (PBDA, 2018). Entretanto, o ambiente implementado por este portal é extremamente estático, não alcançando, assim, em plenitude, os ideais de governo aberto, pois o PBDA exerce basicamente a função de repositório de dados, possibilitando baixa interação entre as partes interessadas. Isso ocorre devido ao uso inadequado das ferramentas, técnicas, estrutura e em alto nível devido à governança empregada.

Essa problemática é observável nas ferramentas colaborativas utilizada pela INDA; as ferramentas são os meios que as partes interessadas podem participar e contribuir para o desenvolvimento dos dados abertos governamentais. As ferramentas são: *streaming* de áudio

e vídeo, listas de discussão por e-mail, wiki, edição de documentos na nuvem, marcador e anotador na nuvem, bate-papo, ferramenta de tíquetes e microblog (PBDA, 2018). Estas ferramentas oferecem baixa interação entre as partes interessadas no sentido de construção de grupos de trabalhos, de comunidade e é inexistente a interação das bases originais com as bases derivadas (bases que foram editadas e transformadas por uma parte interessada), pois não há um ambiente onde as partes interessadas possam disponibilizar os dados de suas pesquisas, para que outras pessoas a reutilizem e assim maximizem o potencial das pesquisas, estudos e descobertas dando assim continuidade.

Destaca-se que a velocidade de aquisição de conhecimento em ciência é impedida pela falta de compartilhamento do processo de pesquisa na sua totalidade, além de todos os seus resultados. Isso torna difícil para os pesquisadores se basearem no trabalho uns dos outros e, sem dúvida, resulta em uma “reinvenção da roda”. Além disso, há um aumento de publicações com “resultados irreproduzíveis”, o que pode levar à desconfiança pública na Ciência (MSDSE, 2018).

À medida que a descoberta científica intensiva em dados se torna mais comum, a reprodutibilidade e a ciência aberta tornam-se mais desafiadoras e importantes. Trabalhar em torno deste tema inclui identificar e promover repositórios para compartilhar dados e fluxos de trabalho, desenvolvendo técnicas para consultar e analisar dados compartilhados e fluxos de trabalho que facilitarão a reutilização, a partir de ferramentas, métodos e *softwares* que suportaram o compartilhamento e reprodutibilidade (Msdse, 2018).

Evans e Campos (2012) ratificam que o principal objetivo do governo aberto é envolver os cidadãos, para tanto as iniciativas atuais dos governos devem ser reavaliadas e novas abordagens devem ser exploradas, pois não basta somente entregar de dados, como ocorre no PBDA. A simples liberação de volumes de dados em um site sem informações sobre por que e como os dados são coletados, organizados e podem ser usados deixa os cidadãos com tarefas hercúleas de determinar sua relevância e confiabilidade.

Neste contexto, as agências executivas, os sites de governo aberto, por exemplo, parecem ser orientados exclusivamente para fornecer dados do governo e acelerar os esforços de liberação da informação. Indiscutivelmente, essa orientação limitou o engajamento dos cidadãos e o escopo geral do governo aberto (EVANS; CAMPOS, 2012).

Dessa forma, para se ter um ambiente adequado de descoberta científica intensiva em dados é essencial tratar a governança e realizar o estudo das partes interessadas, para, a partir de então, propor ferramentas e práticas que suportem compartilhamento, preservação, reusabilidade, rastreamento de procedência, reprodutibilidade dos dados, *software* e fluxos de trabalhos científicos.

Ambientes similares a estes estão em produção em outros países, como o *Open Data Institute* (ODI) que possui sede

no Reino Unido, e defende os dados abertos como um bem público, enfatizando a necessidade de modelos eficazes de governança para protegê-los. O ODI está desenvolvendo definições comuns para descrever como os dados são e como podem ser usados, e devolvem *feedbacks* dos usos dos dados aos governos (ODI, 2018).

2. REFERENCIAL TEÓRICO

2.1 Governo aberto

O governo aberto é um modelo de governança que disponibiliza meios de abertura, transparência e diálogo contínuo entre o governo e seus cidadãos (Veljković; Bogdanović-Dinić; Stoi Enov, 2014). Possibilita o diálogo efetivo entre o Estado e a sociedade civil, bem como promove a participação dos cidadãos na coisa pública (Scholl, 2012).

Conforme a Organização para Cooperação e Desenvolvimento Econômico (OCDE), para a construção de um governo aberto existem três princípios básicos a serem levados em consideração: a *accountability* que é a existência de mecanismos que possibilitem a identificação e responsabilização dos servidores públicos por suas ações; a transparência, que busca a disponibilização de informações confiáveis, relevantes e tempestivas sobre as atividades do governo; e a participação social onde o governo deve promover meios da sociedade civil participar e levar em consideração os seus anseios tanto na proposição quanto na implementação das políticas públicas (UBALDI, 2013).

A OGP ratifica o preceituado pela OCDE, pois para aquela o governo aberto deve se balizar: no aumento da disponibilidade de informações sobre atividades governamentais; em apoiar a participação social; na implementação de padrões nos mais altos níveis de integridade profissional na administração pública; e em aumentar o acesso a novas tecnologias que promovam a transparência e a *accountability* (OGP, 2015). Desta forma, o governo aberto deve disponibilizar os dados e compartilhar informações tornando os cidadãos conhecedores da sua realidade social (HARRISON *et al.*, 2012), e esse contexto se desdobra no conceito de Dados Abertos Governamentais que é apresentado na próxima seção.

2.2 Dados abertos de governo

Os Dados Abertos Governamentais (DAG) são produtos do governo aberto, contribuem para um melhor controle social, fortalecendo a democracia e a cidadania ativa, proporcionando melhorias na administração pública, tais como, inovação, cooperação e a transparência. Para um dado ser considerado aberto ele deve ser acessível, legível por máquina e, a informação produzida, a partir dele, deve ser disseminada a todos (HARRISON *et al.*, 2012).

Segundo a OECD os DAG têm o potencial de aumentar a eficiência do governo, a inovação na prestação de serviços e as operações internas do setor público (UBALDI, 2013),

pois a publicação proporciona a redução de solicitações de informações rotineiramente ao governo, possibilitando que as solicitações sejam acessadas diretamente pela população e trocada entre os entes governamentais.

Os DAG são os dados produzidos ou custodiados pelo governo ou por entidades controladas pelo governo, que podem ser utilizados livremente, reutilizados e redistribuídos por qualquer pessoa. Eles possibilitam a inovação, o crescimento econômico, a produção de novos produtos com valor agregado, a *accountability* e a transparência (YANNOUKAKOU; ARAKA, 2014).

Desse modo, os DAG podem ainda promover a colaboração entre organizações públicas, pois a partir da análise da informação a ser disponibilizada e da identificação do proprietário da informação, pode haver a colaboração e a troca das informações públicas, possibilitando a reengenharia e a simplificação dos procedimentos internos, a fim de posteriormente ocorrer a automatização de processos, a eliminação de gastos redundantes e a redução dos custos transacionais internos (HARRISON *et al.*, 2012).

Nesse sentido, Scholl (2012) acrescenta que a redefinição do papel do governo como fornecedor de informação

permite que a sociedade agregue novos valores aos dados, explorando o potencial de uso e reuso na produção de informação de relevância social, com base nos dados abertos disponíveis.

Observa-se, no contexto nacional, que os dados públicos para serem melhor explorados devem ser disponibilizados sem restrições legais, financeiras, tecnológicas ou políticas para reutilização pela sociedade (DAVIES, 2013). Assim, para que os DAG sejam explorados adequadamente deve-se implementar três pilares: transparência, governança participativa, e proporcionar valor social e comercial (ATTARD *et al.* 2015).

Desse modo, governos locais e nacionais estão voltando-se para dados abertos buscando reduzir seus custos, aumentar a transparência e eficiência, e responder às necessidades da sociedade (HOFFMANN, 2012). Entretanto, não basta apenas disponibilizar os dados; é necessário promover meios para que as partes interessadas deste contexto possam interagir, e, nessa perspectiva, surgem os ecossistemas de dados abertos, que são descritos no próximo tópico.



2.3 Ecossistema de dados

Ecossistema é uma palavra de origem das Ciências Biológicas, mas que tem sido largamente usada para se referir aos sistemas produtivos. A ideia é relativamente recente e é uma influência que os estudos sobre sistemas complexos adaptativos feitos pelas ciências naturais têm feito nas ciências sociais (PELTONIEMI; VUORI, 2004).

Demchenko, Laat e Membrey (2014) definem um ecossistema de dados como um complexo de ferramentas, técnicas e componentes construídos em volta de uma origem de dados específicos e sua aplicação. O ecossistema em questão ainda compreende as seguintes categorias de componentes arquiteturais dos dados: modelos e estruturas, arquitetura, gerenciamento do ciclo de vida e infraestrutura de segurança.

O conceito refere-se à visão do ambiente produtivo dentro de um contexto maior, para além das suas divisas proprietárias. Nesse ambiente há os *stakeholders* que são as partes interessadas e influenciadas pelo sistema produtivo. Um sistema produtivo pode ser visto como uma pessoa exercendo uma atividade, uma área dentro de uma organização, a organização como um todo, um setor econômico e até uma cadeia produtiva (Costa, 2016).

O conceito de ecossistema é composto por muitas especificidades que não se restringe aos *stakeholders*, engloba também as relações de vários níveis entre os diversos *stakeholders*. Ao definir governança como o modelo de gestão que promove a redução ou eliminação dos conflitos de interesses entre as partes, a definição dos *stakeholders* que comporão um ecossistema é uma mensagem clara do gestor do que será considerado nessa tarefa (COSTA, 2016).

Independente de um maior ou menor escopo, o objetivo de se definir um ecossistema *big data* é dar uma visão sistêmica que possibilite decisões mais estratégicas, seja na definição de um modelo de negócio, seja na criação de um produto ou serviço, ou seja na gestão do processo de decisão da própria organização (COSTA, 2016).

O acesso aberto foi motivado, em parte, em resposta ao modelo econômico adotado por editores comerciais que exploravam pesquisas financiadas com recursos públicos e trabalhos acadêmicos voluntários para gerar lucros (LEPTIN, 2012). Observa-se o grande número de *stakeholders* e seus interesses divergentes. Para tanto, o próximo tópico discorre sobre os *stakeholders*.

2.4 Stakeholders

Segundo a perspectiva de *stakeholders*, cabe ao gestor a responsabilidade de identificar todos os atores relevantes para sua organização, que podem afetar ou serem afetados por ela, e, de forma contingencial, desenvolver atitudes, estruturas e mecanismos gerenciais, além de práticas organizacionais, que possam atender os interesses deles de forma simultânea, porém priorizada devido à limitação de recursos (CHAN *et al.*, 2003). Segundo

Freeman (1984) e Donaldson e Preston (1995), em geral os atores relevantes para uma organização são os governos, investidores, grupos políticos, usuários passivos (clientes), comunidades, empregados (voluntários), associações comerciais e fornecedores.

Pode ser custoso levantar todos os *stakeholders* e analisar os interesses que eles possam ter em alguma mudança organizacional no projeto. Mas parece ser imprescindível priorizar ações e esforços, sendo então essencial especificar e entender pelo menos uma vez quem são os envolvidos e afetados pelo projeto, tornando as tarefas de priorização e direcionamento futuras mais realistas e bem-sucedidas. A atenção organizacional não pode ser dada a todos os *stakeholders* de forma isonômica, sendo necessário decidir a quem se dedicar mais e quando, dependendo da assimetria de dependência entre o projeto e o *stakeholder* em foco, algo que é socialmente construído e dinâmico no tempo (CHAN *et al.*, 2003; MITCHELL *et al.*, 1997; PFEFFER; SALANCIK, 2003).

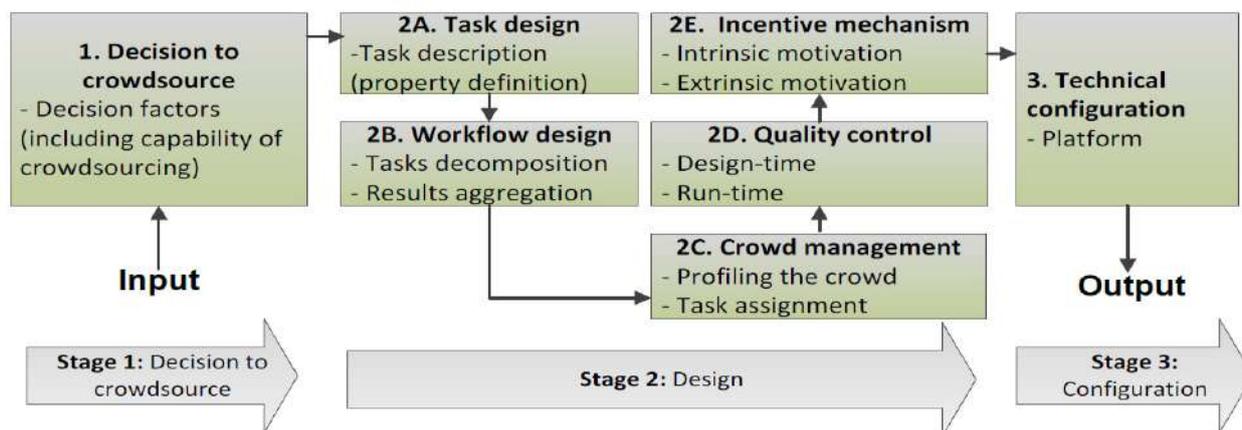
A teoria do *stakeholders* é essencial para o desenvolvimento do *crowdsourcing*, que são ambientes de redes de colaboração para execução de uma atividade. A próxima seção é sobre *crowdsourcing*.

2.5 Crowdsourcing

O *crowdsourcing* representa o ato de uma organização de delegar uma tarefa para uma rede, normalmente grande de pessoas, e feita de forma aberta, podendo ser executada de forma colaborativa ou individual (HOWE, 2006). Nessa perspectiva, *crowdsourcing* é um tema emergente e potencializado pela tecnologia da informação e comunicação. A aplicabilidade do *crowdsourcing* não está somente em seu potencial de terceirizar atividades, mas sim no quanto o processo de terceirização se torna mais rápido, barato e eficiente por meio da conexão de uma grande quantidade de possíveis colaboradores (FERRAZ; GOUVÊA; BARRETO, 2017).

Crowdsourcing é um tipo de atividade participativa realizada pela web em que uma pessoa, uma instituição, uma organização sem fins lucrativos ou uma empresa propõe, por meio de um convite aberto flexível, o compromisso voluntário de execução de uma atividade a um grupo heterogêneo de indivíduos com conhecimentos diversos e em número variado. O empreendimento da tarefa, de complexidade e modulação variáveis, no qual a multidão deve participar contribuindo com seu trabalho, dinheiro, conhecimento e/ou experiência, sempre implica em benefício mútuo. O participante receberá a satisfação de um determinado tipo de necessidade, seja ela econômica, reconhecimento social, autoestima ou o desenvolvimento de competências individuais, enquanto o patrocinador (*crowdsourcer*) irá obter e utilizar a seu favor o que o usuário fornecer, sendo que a maneira de utilização dependerá do tipo de atividade específica (ESTELLÉS-AROLAS; GUEVARA, 2012). Thuan, Antunes, Johnstone (2017) propuseram um modelo de processo de implementação de *crowdsourcing* proposto nas etapas decisão, desenho e configuração. A Figura 1 apresenta esse processo.

Figura 1- Modelo de processo de negócios de *crowdsourcing*



Fonte: Thuan; Antunes; Johnstone (2017)

A primeira fase é a decisão nesse item, e será analisada se a estratégia e o objeto problema são apropriados para a aplicação do *crowdsourcing*. Após tomada essa decisão, a segunda etapa abrange um conjunto de atividades de design necessárias para operacionalizar a decisão. Inclui cinco componentes: projeto de tarefas, projeto de fluxo de trabalho, gerenciamento dos integrantes, controle de qualidade e mecanismo de incentivo. A fase final concentra-se em como configurar um processo de *crowdsourcing* para instanciação em sistemas computacionais envolvendo principalmente uma visão técnica aprofundada (THUAN; ANTUNES; JOHNSTONE, 2017).

2.6 Governança dos dados

Sadiq (2013) declara que o volume de dados gerados no mundo cresce exponencialmente, de forma diversificada e plural, com novos usuários e produtores de dados surgindo a cada dia. Para utilizar esse volume de dados são necessárias adaptações nos processos de negócio, nas fontes de dados, na infraestrutura de tecnologia, nas capacitações e nas mudanças organizacionais. Não se pode apenas levar em consideração a dimensão do volume de dados, mas também a diversidade, o formato, a autenticidade, para que os dados possam ser utilizados e geridos em tempo hábil, com o foco em agregar valor ao negócio.

A governança dos dados pode ser entendida como uma disciplina da área de controle de qualidade que busca a integridade ao processo de gestão, uso, melhoria e proteção de informação organizacional (IBM, 2007). Panian (2010) define que a governança de dados pode ser compreendida como o conjunto de processos, políticas, padrões, estruturas e tecnologias para gerenciar e assegurar disponibilidade, acessibilidade, qualidade, consistência, auditabilidade e segurança nos dados.

Assim é a governança dos dados. Um mecanismo para a implementação de diretrizes para controlar os dados de uma determinada organização, tendo como responsabilidade o controle, o gerenciamento dos dados

e a sua transformação em informação organizacional (BARATA, 2015).

Segundo Mosley *et al.* (2009), a governança de dados visa o planejamento, execução e fiscalização de políticas, práticas e projetos para adquirir, controlar, proteger e alavancar eficazmente o uso dos ativos de dados na organização. Há metodologias de implementação de governança dos dados como: Governança de Dados da Oracle, da IBM (2007); o *Data Governance Institute* (THOMAS, 2009); e, o *Data Management Body of Knowledge* (DAMA, 2015).

A governança de dados é o exercício da autoridade, do controle e da tomada de decisões compartilhada quanto à gestão dos ativos de dados (DMBOK, 2009). Dessa forma, o conceito de governança de dados engloba uma série de atividades, processos, políticas e ações que devem ser realizadas para gerenciar de forma eficaz os dados de uma organização.

3. OBJETIVOS

3.1 Objetivo geral

O objetivo desta pesquisa é identificar e propor soluções de governança de dados abertos para as organizações públicas federais brasileiras a fim de suportar as bases de dados governamentais e proporcionar um ambiente integrado e participativo, desenvolvendo um ecossistema onde as partes interessadas e as bases possam interagir, com características de comunidade e projetos, produzindo insumos e ferramentas para a melhoria da gestão governamental.

3.2 Objetivo geral

OE1 - Analisar modelos de governança de dados e desenvolver a fim de estabelecer as melhores práticas;

OE2 - Identificar os *stakeholders* relacionados à governança de dados abertos governamentais;

OE3 - Identificar as licenças e os licenciamentos relacionados às bases de dados considerando as motivações e interesses das partes que se envolveriam no projeto comunitário;

OE4 – Analisar plataformas que são *case* de sucesso em governança de dados;

OE5 – Descrever a governança de uma plataforma colaborativa de gestão do conhecimento de dados no contexto da administração pública.

4. MÉTODO

Quanto à modalidade de pesquisa, do ponto de vista da sua natureza, para Gil (2006), a pesquisa aplicada objetiva gerar conhecimentos dirigidos à solução de problemas específicos. Adicionalmente, Collis e Hussey (2005) contribuem afirmando que este tipo de pesquisa visa à aplicação de suas descobertas a um problema. Esta pesquisa possui natureza aplicada, pois pretende-se adquirir conhecimento dos modelos de governanças existentes para aplicá-los no âmbito da APF.

Dos objetivos para Triviños (1987), a pesquisa descritiva exige do investigador uma série de informações e pretende descrever os fatos e fenômenos de determinada realidade (TRIVIÑOS, 1987). Gil (2006) alega que a pesquisa descritiva visa a descrever as características de determinada população ou fenômeno, ou o estabelecimento de relações entre variáveis. Sendo assim, trata-se de uma pesquisa descritiva, pois deseja-se entender como a atual governança dos dados abertos governamentais é adotada.

Sobre a abordagem do problema, Gil (2006) afirma que a pesquisa qualitativa busca descrever eventos visando a compreensão dos fenômenos da perspectiva dos participantes da situação. Já a pesquisa quantitativa considera tudo que possa ser contável, o que significa traduzir em números opiniões e informações para classificá-las e analisá-las, e requer que sejam utilizados recursos e técnicas estatísticas (GIL, 2006). Nesta pesquisa pretende-se utilizar abordagem qualitativa para analisar e avaliar características da governança atual empregada na APF.

Para Fonseca (2002) a pesquisa documental utiliza fontes constituídas por material já elaborado, constituído por fontes mais diversificadas e dispersas, sem tratamento analítico, dentre estas: tabelas estatísticas, jornais, revistas, relatórios, documentos oficiais, cartas, entre outros (FONSECA, 2002). Trata-se de uma pesquisa documental por utilizar materiais oficiais disponibilizados pelo Governo Federal brasileiro por meio dos sites dos órgãos.

Finalmente, quanto à estruturação da pesquisa, deseja-se:

- Buscar, elencar e analisar modelos de governança de dados adequados à APF, alcançando assim o objetivo “OE1: analisar modelos de governança de dados e desenvolver uma taxonomia a fim de estabelecer as melhores práticas”;
- Executar o processo sistemático de: análise documental, identificar, classificar, estabelecer níveis de influências, interesses e as demandas dos *stakeholders* que evidenciam os valores, princípios e processos que regem os mecanismos da governança de dados, resultando no objetivo “OE2: identificar os *stakeholders* relacionados à governança de dados abertos governamentais”;
- Para o objetivo “OE3: identificar as licenças e os licenciamentos relacionados às bases de dados governamentais considerando as motivações e interesses das partes que se envolveriam no projeto comunitário”, elencando e estudando as licenças de bases de dados e os níveis de permissões. Caso não sejam identificadas licenças específicas de bases de dados, serão estudadas licenças de softwares livres e verificadas se são potencialmente adaptáveis a dados abertos governamentais;
- Buscar, analisar e realizar benchmarking de portais de governança dos dados, obtendo a conclusão do objetivo “OE4: analisar plataformas que são *case* de sucesso em governança de dados”;
- Por fim, pretende-se descrever a governança dos dados de um portal colaborativo. Para tanto, irá mapear os processos, definir atividades, definir ferramentas e técnicas, alcançando o objetivo “OE5: descrever a governança de uma plataforma colaborativa de gestão do conhecimento de dados no contexto da administração pública”.

5. DESENVOLVIMENTO

Esta seção apresenta os resultados obtidos.

5.1 Benchmarking de Portais de Dados Abertos

Utilizando a técnica de benchmarking foram analisadas 32 plataformas, conforme Tabela 3, com características de *crowdsourcing* de dados abertos. Dentre elas, 19 mantidas por organizações públicas e 13 por organizações privadas.

Para a seleção das plataformas analisadas utilizou-se a classificação da *Open Data Barometer* que é um instituto que classifica os governos quanto à abertura dos dados (ODB, 2019). Foram analisados os portais dos governos federais dos cinco países (Canadá, Reino Unido, Austrália, França e Coreia) melhores classificados. Na seleção das outras plataformas foram feitas buscas para identificar

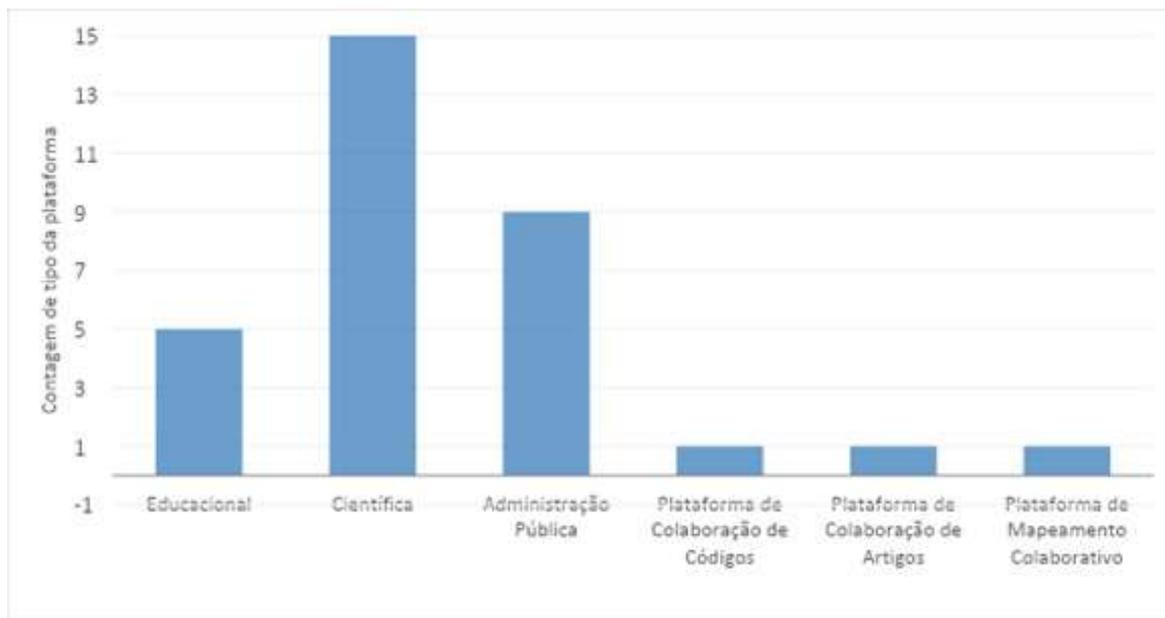
plataformas colaborativas de dados, e os critérios de seleção destas foram: a) plataformas que se destacam no campo dos dados abertos demonstrando alto nível de colaboração; b) plataformas reconhecidas em sua área de atuação; e c) conveniência e acessibilidade.

Para cada caso existente, foram analisados os seguintes aspectos: tipo da plataforma; licença; funcionalidades; stakeholders; como contribuir; validação; tipo de

organização (pública ou privada); permissão de escrita ou edição por terceiros; se a plataforma era software livre e a licença do software.

Categorizamos as plataformas a partir das características dos dados disponibilizados. A maioria das plataformas selecionadas possuía um foco científico (15), seguido por administração pública (9) e educacional (5), conforme está sistematizado na Figura 1.

Figura 1 - Tipos de plataformas analisadas



Fonte: dados primários.

Do grande percentual de plataformas científica, 46,8% se justifica, pois estas apresentam melhores práticas em disponibilização de mecanismos de colaboração. A abordagem em favor do acesso aberto do conhecimento científico está consistente em muitos países e áreas de conhecimento, tendo como ponto de referência a publicação em 2002 da Declaração de Budapeste (Budapest Open Access Initiative - BOAI), onde se estabeleceu o conceito e as metodologias de disponibilização aberta. As estratégias centrais propostas na BOIA são: periódicos de acesso aberto e auto arquivamento da produção científica em repositórios, que são comumente chamadas de Via Dourada e Via Verde, respectivamente (SANTOS *et al.*, 2017).

As plataformas governamentais correspondem a 28%. Este elevado quantitativo se justifica devido às demandas da sociedade e a redefinição do papel do governo com relação à gestão e governança dos dados públicos.

A Tabela 1 apresenta as principais licenças de dados abertos utilizadas nas plataformas analisadas e destacam-se como as licenças mais frequentes: a Creative Commons, a Open Data Commons e as licenças personalizadas.



Tabela 1 – Licenças para dados dos casos existentes

Código	Licenças dos Casos Existentes	Definição	Organização Autora da Licença
LD1	Creative Commons CC-Zero (CC0)	Dedicação ao Domínio Público. O autor deste trabalho renunciou seus direitos autorais, tornando como domínio público, para que o destinatário possa usar o trabalho livremente para qualquer finalidade sem pedir a permissão do autor de tal obra.	Creative Commons
LD2	Creative Commons Attribution 4.0 (CC-BY-4.0)	A licença Creative Commons Attribution permite a redistribuição e reutilização de um trabalho licenciado com a condição de que o criador seja devidamente creditado.	Creative Commons
LD3	Attribution Non Commercial 4.0 International (CC BY-NC 4.0)	Permite que outros remixem, adaptem e criem a partir do seu trabalho para fins não comerciais, e embora os novos trabalhos tenham de lhe atribuir o devido crédito e não possam ser usados para fins comerciais, os usuários não têm de licenciar esses trabalhos derivados sob os mesmos termos.	Creative Commons
LD4	GNU General Public License (GPL)	A GPL permite que os programas sejam distribuídos e reaproveitados, mantendo a recíproca total em todos seus derivados.	Free Software Foundation
LD5	Open Database License (ODbL)	Permite compartilhar, criar, adaptar. Enquanto deve-se: atribuir crédito, manter a licença, manter aberto.	Open Knowledge Foundation
LD6	Open Government Licence (OGL)	O OGL permite que qualquer pessoa copie, publique, distribua, transmita e adapte o trabalho licenciado e o explore comercialmente e não comercialmente. Em troca, o trabalho licenciado tem que reconhecer a fonte do trabalho. Compatível com o Creative Commons Attribution.	The National Archives
LD7	Open Database Connectivity (ODBC)	O ODBC permite acesso a sistemas gerenciadores de bancos de dados (SGBD). É definido um conjunto de interfaces que possibilita o uso de linguagens de programação como Visual Basic, Delphi, Visual C++, Java, entre outras, capazes de utilizar estas interfaces, com o objetivo de ter acesso a uma variedade de bases de dados diferentes sem a necessidade de codificar métodos de acesso especializados.	Microsoft e Simba Technologies
LD8	Creative Commons Attribution 3.0 New Zealand licence (CC BY 3.0 NZ)	A Creative Commons Attribution 3.0 New Zealand possibilita que você visualize, edite, modifique, traduza e distribua trabalhos em todo o mundo, desde que você credite o autor original.	Creative Commons
LD9	Attribution -ShareAlike 3.0 Unported (CC BY-SA 3.0)	A licença Creative Commons Attribution – ShareAlike é uma das licenças copyleft criadas pelo projeto Creative Commons. Um trabalho licenciado desta maneira concede todas as quatro liberdades listadas na definição de obras culturais livres: a liberdade de usar e executar o trabalho, estudar o trabalho e aplicar as informações, redistribuir cópias e distribuir obras derivadas.	Creative Commons
LD10	Creative Commons CC-BY	Esta licença permite que outros distribuam, adaptem, remixem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original. É uma das mais flexíveis licenças e é a mais aberta no quesito das permissões e acessos, também permitindo a utilização dos conteúdos para fins comerciais. Assim, terceiros podem conseguir lucros com o trabalho de outras pessoas a qualquer momento, sem que o criador tenha muito controle.	Creative Commons
LD11	Mozilla Public License (Version 2.0)	A Mozilla Public License (MPL) é uma licença de código aberto e software livre criada pela Mozilla Foundation. É uma licença de copyleft, caracterizada como um meio termo entre licenças de software livre permissivas e a Licença Pública Geral GNU (GPL), que procura equilibrar as preocupações de desenvolvedores proprietários e de código aberto.	Mozilla Foundation
LD12	Open Data Commons Public Domain Dedication and License (PDDL)	A Licença e Dedicção de Domínio Público (ODD) do Open Data Commons (ODC) é usada para colocar conjuntos de dados no domínio público. Não é uma licença de software e não pode ser usada com o objetivo de colocar software no domínio público.	Open Data Commons

LD13	Atribuição-Compartilhada Igual 4.0 Internacional (CC BY-SA 4.0)	É uma licença que exige que adaptações ou cópias do trabalho sejam lançadas sob licença igual ou similar à original.	Creative Commons
LD15	Atribuição-Compartilhada Igual 2.0 Genérica (CC BY-SA 2.0)	É uma licença que exige que adaptações ou cópias do trabalho sejam lançadas sob licença igual ou similar à original.	Creative Commons
LD16	CC0 1.0 Universal (CC0 1.0)	CC0 é o "equivalente moral" mais próximo do domínio público; ele é licenciado para todos os usos, sem restrições ou qualificações, além de que o autor do trabalho ainda possui os direitos autorais naquelas jurisdições onde a alienação voluntária (liberação para o domínio público) é legalmente impossível.	Creative Commons
LD17	CC BY NC 3.0	Essa licença permite que ocorra o compartilhamento (copie e redistribua o material em qualquer mídia ou formato) e a adaptação (remixar, transformar e construir sobre o material) para qualquer finalidade, mesmo comercialmente, desde que se dê o crédito apropriado, promova um link para a licença e indique se mudanças foram feitas. Não se pode usar o material para fins comerciais.	Creative Commons
LD18	Creative Commons Attribution 2.1 Japan License	Nessa licença, o usuário é livre para copiar e redistribuir o material em qualquer lugar e em qualquer forma, transformar, adaptar e desenvolver o material para qualquer finalidade, mesmo comercialmente. O licenciante não pode anular essas liberdades desde que o usuário siga os termos preferidos na licença.	Creative Commons

Fonte: elaboração própria.

Das 32 plataformas analisadas, em 65,6% são utilizadas as licenças da Creative Commons e a licença mais recorrente é a do tipo CC-BY. Essa licença é recomendada para proporcionar uma maior disseminação e uso dos materiais licenciados. Ela é a mais aberta no quesito das permissões e acessos e também permite a utilização dos conteúdos para fins comerciais. Assim, terceiros podem conseguir lucros com o trabalho de outras pessoas a qualquer momento, sem que o criador tenha muito controle.

Quanto ao licenciamento das plataformas, softwares em si, 43,8% são *softwares* livres, ou seja, no qual é concedido liberdade ao usuário para executar, acessar e modificar o código fonte, e redistribuir cópias modificadas ou não. Entre as licenças para os *softwares* livres destaca-se a GPL. A Tabela 2 apresenta as licenças de softwares sob a quais as plataformas foram desenvolvidas.

Tabela 2 – Licenças dos softwares de cada plataforma

Código	Licenças dos Casos Existentes	Definição	Organização Autora da Licença
LS1	GNU Affero General Public License	A GNU Affero General Public License é uma licença de software livre que tem como propósito ser uma licença pouco modificada da GNU GPL e suprir as necessidades de propor liberdade em softwares onde não existe acesso direto ao binário/código-objeto.	Free Software Foundation
LS2	MIT	A licença MIT é uma licença permissiva que pode ser usada para software proprietário ou livre. Seu texto é claro ao abordar os direitos que são transferidos, afirmando que qualquer pessoa que possui arquivos com documentos associados e cópia do software pode tratar com eles sem restrição ou limitação nos quesitos de modificação e publicação, por exemplo. Todavia, deve-se manter uma cópia da licença nas cópias do software e o lembrete de copyright.	Instituto de Tecnologia de Massachusetts
LS3	CC0	Dedicação ao Domínio Público. O autor deste trabalho renunciou seus direitos autorais, tornando-o como domínio público, para que o destinatário possa usar o trabalho livremente para qualquer finalidade sem pedir a permissão do autor de tal obra	Creative Commons
LS4	Apache 2.0	A Licença Apache é uma licença de software livre permissiva. Todo software produzido pela Apache Software Foundation ou algum dos seus projetos ou projetos externos são licenciados em consonância com as regras da licença Apache. Essa licença não é uma copyleft, pois ela permite a distribuição e a utilização do código fonte no software open source e no proprietário. Algumas versões pedem a inclusão do aviso de copyright e um disclaimer.	Apache Software Foundation

LS5	GNU General Public License 2.0	A GPL é uma das licenças com maior uso pelos projetos de software livre, geralmente devido à sua adoção no sistema operacional GNU/Linux e em seu projeto. Essa licença baseia-se em 4 liberdades: a liberdade de executar o programa, para qualquer objetivo; estudar e adaptar o programa para as necessidades; distribuir novamente mais cópias a fim de ajudar outras pessoas e aperfeiçoar o programa para que todos os usuários se beneficiem disso.	Free Software Foundation
LS6	BSD 3-Clause	A licença BSD é uma licença de código aberto com poucas restrições se comparada a outras licenças, como a GNU General Public License ou mesmo as restrições mais usadas determinadas pelo copyright, sendo assim, de certa forma, próxima do domínio público.	Regents of the University of California
LS7	Mozilla Public License Version 2.0	A Mozilla Public License (MPL) é uma licença de código aberto e software livre criada pela Mozilla Foundation. É uma licença de copyleft, caracterizada como um meio termo entre licenças de software livre permissivas e a Licença Pública Geral GNU (GPL), que procura equilibrar as preocupações de desenvolvedores proprietários e de código aberto.	Mozilla Foundation

Fonte: elaboração própria.

A Tabela 3 apresenta uma compilação das informações relativas ao: tipo de plataforma e licenças (dados e software) e verificou-se uma característica comum em várias plataformas do tipo educacional, científica e administração pública: elas não permitem a escrita ou edição por pessoas não cadastradas. Isso pode ser

explicado pela preservação da confiabilidade e segurança das plataformas. Nas plataformas do tipo administração pública, a validação é de responsabilidade do órgão que fornece o conjunto de dados, enquanto nas plataformas educacionais e científicas a validação geralmente é feita pelo próprio usuário.

Tabela 3 - Compilação das informações das plataformas

Caso Existente	Tipo de Plataforma	Licença da Plataforma	Licenças das Bases de Dados
Kaggle	Educacional	LS3, LS5	LD4, LD5, LD10
GenBank	Científica	LS5	LD4
Portal Brasileiro de Dados Abertos (PBDA)*	Administração Pública	LS1, LS3, LS5	LD1, LD3, LD5, LD6, LD12
DNA Data Bank of Japan (DDBJ)	Científica	Não identificado	LD18
European Bioinformatics Institute (EMBL-EBI)	Científica	LS3	LD1, LD10
London Datastore (Reino Unido)	Administração Pública	Não identificado	LD6
Open Government (Canada)	Administração Pública	LS2	LD6
ParisData (França)	Administração Pública	Não identificado	LD7
GitHub**	Plataforma de Colaboração de códigos	LS1, LS2, LS3, LS4, LS5, LS6, LS7	LD4, LD10, LD11
iNaturalist	Científica	LS3	LD1, LD3
USAspending.gov	Administração Pública	LS3	LD1, LD6
Data.Go.KR - Official Korean Government Portal ***	Administração Pública	LS3	LD1, LD10, LD13, LD15
LINZ Data Service	Administração Pública	Não identificado	LD8
Australian Government Catalogue	Administração Pública	LS2, LS4, LS5	LD8
Wikipedia	Plataforma de Colaboração de artigos	LS5	LD9
Catalogue of Life	Científica	LS2	Não identificado

OpenStreetMap	Plataforma de Mapeamento colaborativo	LS2	LD5, LD15
Figshare	Científica	LS3	LD1, LD10
Dryad	Científica	LS3	LD1
Zenodo	Científica	LS5	LD4
Neuroscience Experiments System (NES)	Científica	LS7	LD11
Universidade Aberta do Sistema Único de Saúde (UNA-SUS)	Educacional	LS3	LD1
Oxford University Research Archive (ORA)	Educacional	LS3	LD10
Curriculo+	Educacional	LS3	LD10
CyVerse	Científica	LS3	LD12, LD16
Plataforma MEC de Recursos Educacionais Digitais (MEC RED)	Educacional	LS2, LS5	LD13
DataHub	Científica	LS2	LD12
Social Platform for Open Data (SPOD)	Científica	LS2	LD14
Portal de Dados Abertos da Universidade Federal de Pernambuco	Administração Pública	LS2	LD2, LD8
Mendeley Data	Científica	LS3	LD1, LD2, LD17
DKAN	Científica	LS5	LD4
Data World****	Científica	LS3	LD1, LD10, LD12

Fonte: elaboração própria.

As plataformas Portal Brasileiro de Dados Abertos (PBDA)*, GitHub**, Data.Go.KR – Official Korean Government Portal*** e Data World**** possuem grande variedade de licenças de dados, além das apresentadas na Tabela 3. O PBDA, além das licenças relacionadas na Tabela 3, disponibiliza dados sobre as licenças: Creative Commons Atribuição; Creative Commons Atribuição e Compartilhamento, Licença de Atribuição do Open Data Commons, entre outras. Esta grande variedade de licenças ocorre devido ao governo brasileiro não estabelecer quais licenças as organizações públicas devem utilizar e, destaca-se, ainda, que há no PBDA diversos dados que foram publicados sem licença.

A plataforma Github possui além das licenças apresentadas na Tabela 3 as seguintes licenças: BSD 2- Clause license, Affero GPL, Eclipse Public License v1.0, LGPL v3, GPL v3, LGPL v2.1 e Artistic License 2.0. O elevado número de licenças ocorre devido aos tipos de dados (códigos fontes) que são publicados. Já a plataforma Data World*** possui as licenças CDLA-Permissive-1.0, ODC- BY, CDLA-Sharing-1.0 e ODC-ODbL.

Quanto às funcionalidades das plataformas são notórias que em grande parte há uma função de catálogo científico, como o de DNA (DNA Data Bank of Japan, European Bioinformatics Institute e GenBank) e espécies de seres vivos (iNaturalist, Catalogue of Life). Além disso, é evidente

a variedade de tipos de dados disponibilizados e das funcionalidades como: APIs abertas e visualizações de dados, para permitir que o usuário utilize dados abertos com facilidade. Estes casos foram apresentados por governos mais “fiéis” à política de dados abertos em seus portais, como: London Datastore (Reino Unido), ParisData (França), USAspending.gov (Estados Unidos), Data.Go.KR (Coréia do Sul) e Australian Government Catalogue (Austrália).

Entre os *stakeholders*, além dos governos, destaca-se a presença de seis universidades como mantenedoras das plataformas, a saber: University of Oxford (Dryad e Oxford University Research Archive), Universidade de São Paulo (Neuroscience Experiments System), University of Arizona (CyVerse), Universidade Federal de Santa Catarina (Plataforma MEC de Recursos Educacionais Digitais), Universidade Federal do Paraná (Plataforma MEC de Recursos Educacionais Digitais) e a Universidade Federal de Pernambuco (Portal de Dados Abertos da Universidade Federal de Pernambuco). A participação de universidades como *stakeholders* indica a importância do movimento de abertura de dados no meio acadêmico para permitir que dados sejam utilizados em produção de conhecimento, análises e inovação.

Para a contribuição quase todos os portais exigem uma forma de cadastro, com exceção dos governamentais. Nestes são feitos um modelo de colaboração e validação

interno através de entidades de cada governo. Para a validação de outros portais, é possível que os mantenedores validem os dados disponibilizados, como no GenBank, DNA Data Bank of Japan, European Bioinformatics Institute, LINZ Data Service, Figshare e Currículo+. Em outros casos, afere-se que a validação é de responsabilidade de quem disponibilizou os dados.

Em termos de coprodução e cogerenciamento, a amostra analisada é modesta. Apenas em 15,6% dos 32 portais existe a possibilidade de escrita ou edição de dados por terceiros. O que demonstra uma baixa maturidade quanto à colaboração.

5.1.1 Melhores plataformas

Esta seção apresenta características e funcionalidades das sete plataformas que selecionamos com as que possuem as melhores práticas quanto à colaboração e interação. As plataformas são: a Social Platform for Open Data (SPOD), Kaggle, Neuroscience Experiments System, Wikipedia, Github, Open Street Map e DKAN.

O SPOD permite interações sociais entre usuários de dados abertos e dados do governo. Há uma ferramenta de “salas públicas” para que os usuários adicionem diferentes tipos de dados. Essa plataforma melhora a compreensão e aumenta a acessibilidade de dados abertos, favorecendo assim a colaboração social através de um ambiente favorável aos cidadãos.

O Kaggle é uma plataforma educacional de conjunto de dados públicos, competições de aprendizado de máquina e ensino de aprendizagem de máquinas. Há uma grande interação entre os usuários, tendo em vista que nessa plataforma empresas e instituições publicam desafios de análise de dados onde usuários do mundo todo competem entre si para produzir os melhores resultados. A partir do cadastro, o usuário pode criar um conjunto de dados, enviar um arquivo e convidar pessoas para contribuir.

O Neuroscience Experiments System é uma plataforma científica de código aberto que visa ajudar os laboratórios de pesquisa em neurociência nos procedimentos de rotina para coleta de dados. Essa plataforma possui foco nas necessidades da sociedade no sentido de fornecer um único repositório para os dados experimentais de todo um laboratório de pesquisa, grupo ou projeto, além de proporcionar uma estrutura modular e interface web que fornecem um uso intuitivo de suas funcionalidades de gerenciamento de dados.

A Wikipedia é uma plataforma de colaboração de artigos cujo conteúdo é livre e construído de forma colaborativa. A interação é destaque, pois qualquer usuário (cadastrado ou não) pode criar um artigo nessa plataforma. Basta seguir o guia de criação de artigos e seguir os passos indicados caso o tópico que vai ser criado esteja de acordo com os requisitos. A colaboração vem de todas as partes do mundo.

O Github é uma plataforma de colaboração de repositórios de código-fonte que possui controle de versão, fóruns e muitas outras funcionalidades. Por meio do cadastro

na plataforma, pode-se criar um repositório e enviar o código-fonte via linha de comando. Os repositórios podem ser públicos ou privados. Os públicos são frequentemente usados para o compartilhamento de software de código aberto. A boa prática de interação se dá pelos programadores, utilitários ou qualquer usuário cadastrado no Github poderem contribuir em projetos privados e/ou Open Source de qualquer lugar do mundo.

O Open Street Map é uma plataforma de mapeamento colaborativo onde é disponibilizado um mapa livre e editável. Para contribuir é necessário criar uma conta e recolher dados utilizando um GPS ou outros métodos. O Bing Maps colabora com as imagens aéreas e é possível ainda editar informações. Há grande integração por meio de mapeamento colaborativo a fim de criar um mapa livre e editável do mundo.

O DKAN é uma plataforma científica de código aberto com um conjunto completo de recursos de catalogação, publicação e visualização que permitem que as organizações publiquem facilmente os dados para o público. O usuário se cadastra na plataforma, seleciona o tipo de conteúdo que o interessa, carrega arquivos de dados, adiciona metadados e pode publicar. Possibilita, assim, alto nível de colaboração.

5.1.1.1 Análise das partes interessadas

Este tópico apresenta uma análise das partes interessadas das sete plataformas selecionadas. Foi inicialmente realizada uma busca por comunidades, pessoas ou organizações envolvidas no desenvolvimento e operação dos portais que pudessem de alguma forma serem afetadas pelos projetos, ou que exercessem qualquer tipo de influência sobre o projeto, suas entregas e sua respectiva equipe. As informações levantadas foram: parte interessada, organização, nome da parte interessada, e-mail, posição na organização, função no projeto, principais responsabilidades, principais expectativas/interesses, matriz poder/interesse, nível de engajamento e referências.

Primeiramente, com relação ao PBDA (Portal Brasileiro de Dados Abertos), por se tratar de uma iniciativa pública, tem como órgão “guarda-chuva” a Infraestrutura Nacional de Dados Abertos (INDA), que tem por finalidade gerir a política de dados abertos do Poder Executivo federal, em seu comitê gestor, possui membros de 9 instituições públicas, um membro da sociedade civil e um do setor acadêmico. Dentre as 9 instituições, atualmente estão: Ministério da Saúde, IBGE, Ministério da Cidadania, Controladoria-Geral da União, Ministério da Educação, Ministério da Economia, Casa Civil da Presidência da República, Ministério da Ciência, Tecnologia, Inovações e Comunicações, Secretaria de Governo da Presidência da República e Ministério da Justiça. Inicialmente, a plataforma foi desenvolvida pela Secretaria de Tecnologia da Informação, pasta essa que possui vínculo direto com o Ministério do Planejamento, Desenvolvimento e Gestão. Percebe-se, portanto, a participação direta da administração pública no desenvolvimento, manutenção e fomento da plataforma. No decorrer dos anos, diversos órgãos e instituições públicas aderiram à plataforma da INDA, sendo 2017 o mais profícuo de todos.

Figura 3 - Planos de aderência de órgãos publicados pelo site da INDA



Fonte: <http://wiki.dados.gov.br/Balanco-da-INDA-em-2018.ashx>

Observa-se que existe uma aderência à plataforma, demonstrando que possui credibilidade frente a diversos órgãos que alimentam a plataforma e devido às exigências legais impostas às organizações.

Com relação ao Social Platform for Open Data (SPOD), foi identificado um cenário similar ao PBDA, no que tange a participação estatal na plataforma. Foi inicialmente desenvolvida pela iniciativa ROUTE-TO-PA, que reuniu diversos interessados - do ponto de vista científico e prático - nos tópicos abordados pelo projeto ROUTE-TO-PA durante toda a sua duração. É formado de maneira informal e consiste inicialmente de pessoas que foram ativamente envolvidas em todas as fases do ROUTE-TO-PA, contribuindo para o *know-how* do projeto. O projeto do SPOD recebeu financiamento direto da União Europeia, em um programa chamado “*European Union’s Horizon 2020*”, que visa financiar projetos, que envolvam inovações tecnológicas, que possam beneficiar as relações entre indivíduos/organizações no âmbito da União Europeia. Das 15 pessoas envolvidas no grupo principal de desenvolvimento, 6 estão vinculadas diretamente à universidades europeias enquanto as outras 9 são de empresas privadas e iniciativas comunitárias, demonstrando uma forte presença das universidades.

É importante salientar que neste portal, o apoio da União Europeia, por meio do programa de pesquisa e inovação *European Union’s Horizon 2020*, foi de fundamental importância para um desenvolvimento satisfatório da plataforma como um todo, que, apesar de possuir também auxílio de organizações privadas, estas últimas possuem aspecto de auxílio mais técnico e de pessoal do que financeiro em si.

Sobre a plataforma Kaggle inicialmente ocorreu um massivo financiamento privado de diversas organizações de “investidores anjos” que realizam a ponte entre startups de grande potenciais investidores. Ao contrário

das plataformas anteriores, o Kaggle não possui qualquer tipo de apoio por parte do governo americano, e não possui qualquer universidade entre seus financiadores, ainda mais agora que foi recentemente comprado pela Alphabet, em 2017.

A startup surgiu, inicialmente, como criadora de modelos preditivos, por meio de competições, para auxiliar a resolução de problemas para organizações como: NASA, Wikipedia e Ford. Tendo sido criada por Anthony Goldbloom, que atualmente encontra-se no posto de CEO da Kaggle, o que demonstra sua capacidade de liderar o projeto, apesar de ter sido adquirido por uma grande organização. O que se pode afirmar em relação ao que foi relatado é que existe uma cultura de incentivo ao investimento as startups que, apesar de não ser algo formalizado, é bastante difundido em território americano.

No âmbito de iniciativa brasileira, o Neuroscience Experiments System (NES) possui características similares ao SPOD e ao PBDA, com uma predominância de investimento público que se demonstra pela iniciativa vinda de universidades públicas ou de órgãos públicos, que neste caso encontra afinidade com o mundo acadêmico. A primeira versão da plataforma foi desenvolvida pela Universidade Federal do Rio de Janeiro (UFRJ), atualmente evoluiu bastante e é patrocinada atualmente pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), dentro de seu Centro de Pesquisa, Inovação e Difusão.

Assim como a plataforma SPOD, no que tange pesquisadores, o NES possui um polo de apoio que se encontra em grande parte na região sudeste do país, se considerarmos sua atuação nacional, com apoio em grande parte pela USP, UFRJ, UNICAMP e UFRN. Já em caráter internacional, o NES possui pesquisadores na França, Itália, Uruguai e Países Baixos, onde se encontram as organizações que possuem apoiadores que atuam

como pesquisadores associados seniores: Université Paris Descartes, Université de Cergy Pontoise, Istituto Nazionale di Fisica Nucleare/INFN, Sapienza Università di Roma, CNRS, Universidad de la Republica, Universiteit Utrecht, UFRN.

O Wikipedia, por sua vez, talvez seja a plataforma mais conhecida entre as citadas neste corrente estudo, sendo de iniciativa americana, possui a característica fundamental de se bancar por meio de doações realizadas tanto por entidades e organizações quanto por indivíduos. O mesmo vale para as contribuições na plataforma, que são conhecidas por serem inteiramente realizadas pelos usuários, onde os mesmos editam, criam e desenvolvem todo o ambiente de artigos existentes.

A organização que cuida diretamente dos recursos é a Wikimedia Foundation, que inclusive organiza as outras subsidiárias, agindo como intermediária e sistematizadora da estabilidade da plataforma em diversas línguas e países. É interessante notar que, novamente, o governo não se

envolve com as plataformas impulsionadas dentro de solo americano, garantindo uma certa liberdade de ação para os usuários.

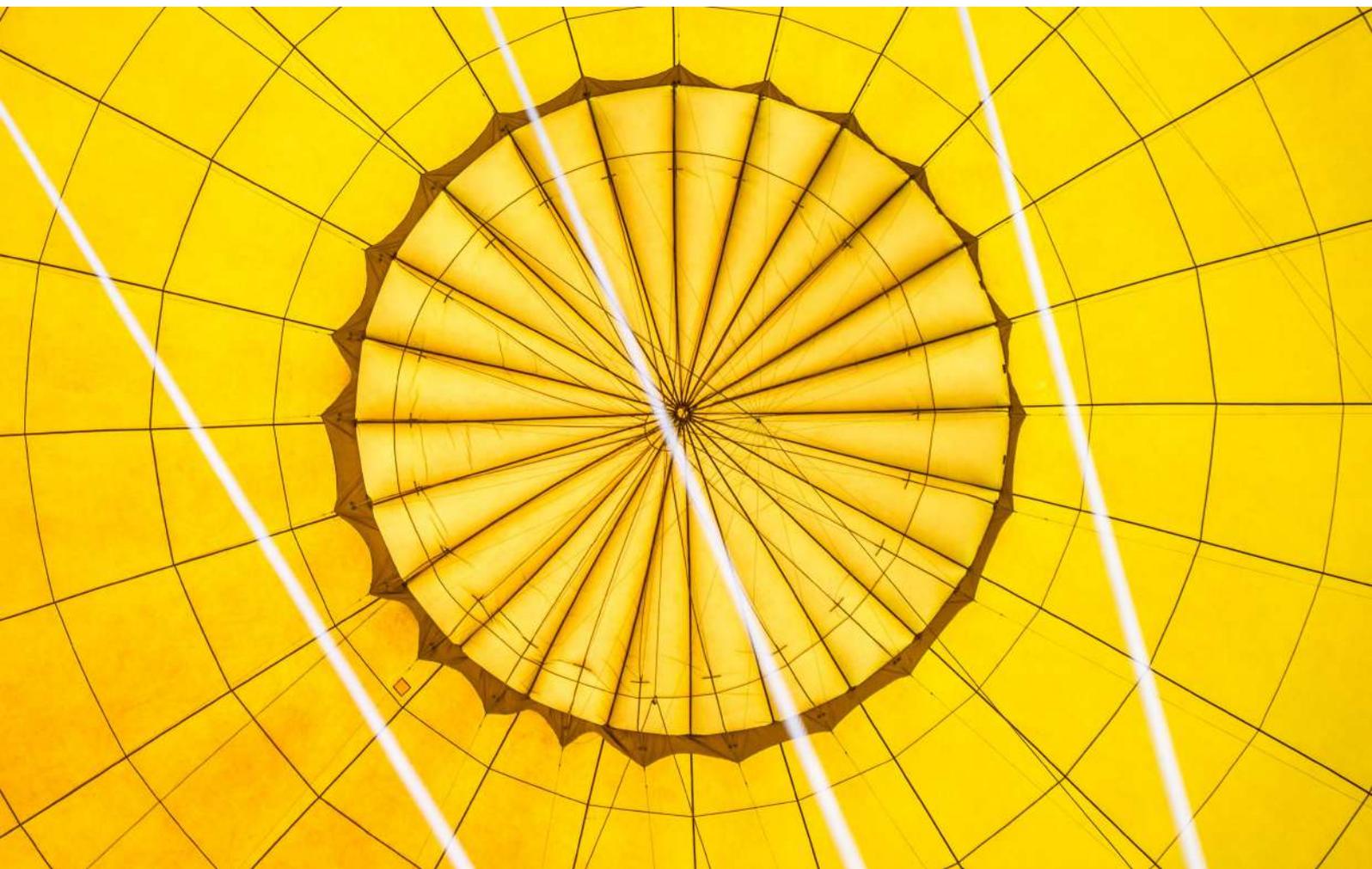
O Github possui um cenário muito similar ao da Kaggle, apesar de lidarem com tipos distintos de dados, um dos seus investidores iniciais são os mesmos, tal como a Khosla Ventures. É interessante analisar que enquanto o Kaggle foi adquirido recentemente pela Google, o Github foi adquirido pela Microsoft e, tal informação, nos mostra que o contexto das *startups* faz com que as grandes corporações e os grandes investidores sempre acabem fomentando o ambiente de inovação, em um contexto de “concorrência” onde as melhores ideias competem pelos melhores investimentos e aportes financeiros. O Github não possui investimento realizado por universidades e nem pelo governo americano.

Por fim, temos o Open Street Map (OSM), que em síntese possui uma mescla de características do Wikipedia e do Github, conforme a Tabela 4.

Tabela 4 – Comparação OMS, Wikipedia e Github

Características	Open Street Map	Wikipedia	Github
É desenvolvido com base em um sistema de colaboração	Sim	Sim	Sim
Aceita doações monetárias	Sim	Sim	Não
Preza pela praticidade	Sim	Sim	Sim

Fonte: dados primários.



De característica relacionada ao Wikipedia, existe a cultura da colaboração, já que a plataforma do OSM é diretamente alimentada pelos usuários. Qualquer um que possua GPS pode auxiliar na inserção e edição de dados na plataforma e, existe também a questão das doações como meio de subsistência da plataforma, e que, recentemente, recebeu ainda por cima um aporte anônimo de milhões de dólares em Bitcoins, importando em uma noção de que existe, sim, um ambiente propício para colaboração não só técnico no ambiente das plataformas de dados abertos, mas também monetário.

Com relação à similaridade com o Github, existe o fato de que a plataforma, como iniciativa, impulsionou um movimento fortíssimo de apropriação da mesma por parte dos usuários na sua utilização. Atualmente, o Open Street Map é um modelo difundido para diversos países e regiões do planeta, sendo um exemplo de transparência tanto de dados quanto de utilização apropriada do dinheiro que é doado para os mesmos.

5.2 Análise das licenças para DAG

Os dados abertos governamentais geralmente são de fácil acesso e estão disponíveis em plataformas governamentais. Entretanto, ter acesso à base de dados não é suficiente para autorizar a pessoa a fazer o que quiser. Nesse contexto, as licenças são documentos que regulamentam as liberdades, direitos e obrigações sobre os dados, assim as licenças habilitam quais atividades e processamentos podem ser realizados.

Existem diversos termos que uma licença de dados pode conter, mas comumente existe um reaproveitamento de licenças que são aceitas pela comunidade. Entretanto, há uma proliferação de licenças, o que traz um impacto negativo, pois dificulta o entendimento dos usuários, podendo limitar a interoperabilidade entre as bases e gerar insegurança jurídica.

Destaca-se que o Governo Federal brasileiro não define um rol de licenças que podem ser utilizadas. Isto reflete em uma total desordem na regulamentação das licenças, conforme observamos que no Portal Brasileiro de Dados Abertos há dados licenciados com 14 tipos de licenças distintas (PBDA, 2018). A Tabela 5 apresenta as licenças utilizadas e a quantidade de bases de dados que as utilizam.

Tabela 5 – Licenças utilizadas no PBDA

LICENÇA	QUANTIDADE DE BASE
Licença Aberta para Bases de Dados (ODbL) do Open Data Commons	3359
Creative Commons Atribuição	1025
Licença não especificada	784
Outra (Domínio Público)	305
Outra (Aberta)	255
Creative Commons Atribuição e Compartilhamento pela mesma Licença	167
Creative Commons Não-Comercial (qualquer)	109
Creative Commons CCZero	19
Licença e Dedicção ao Domínio Público do Open Data Commons (PDDL)	6
Licença de Atribuição do Open Data Commons	2
Licença GNU para Documentação Livre	2
Outra (Não Aberta)	2
Open Government Licence do Reino Unido (OGL)	1
Outra (Atribuição)	1

Fonte: PBDA, 2018

No âmbito internacional, as licenças são utilizadas pelos países melhores classificados na Open Data Barometer (ODB), o instituto que classifica os países quanto à abertura dos dados (ODB, 2019) e apresenta uma grande diversidade. A Tabela 6 apresenta uma compilação das licenças utilizadas

nos portais dos governos federais de 13 países, destes, 10 são os melhores classificados pela ODB. Observa-se que mesmo os portais disponibilizando dados de natureza semelhante, diversas licenças são utilizadas, sendo ainda recorrente dentro de um portal utilizar várias licenças.

Tabela 6 – Licenças utilizadas por governos internacionais

País	Portal	LICENÇAS	OBSERVAÇÕES
França	https://www.data.gouv.fr/fr/licences	-Open License - Personalizado -Open Data Commons - Licença de banco de dados ODC aberto (ODbL) versão 1.0 *Legislação francesa delimita que só pode usar essas duas licenças	Para evitar a proliferação de licenças, a lei para a República digital prevê a criação de uma lista, criada por decreto de licenças que podem ser usadas pelos governos para reutilização gratuita de informação pública, que são dados ou código fonte de um software (artigo D.323-2-1 do código das relações entre o público e a administração (CRPA). As administrações que desejam usar uma licença que não aparece no parágrafo anterior devem primeiro obter aprovação sob as condições estabelecidas no artigo D.323-2-2 doCRPA
Reino Unido	Data.gov.uk http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/ https://www.gov.uk/government/collections/how-to-publish-central-government-transparency-data	-Open Government Licence v3.0 - Personalizada	Estes termos são compatíveis com a Creative Commons Attribution License 4.0 e com a Licença de Atribuição do Open Data Commons, sendo que ambos licenciam direitos de copyright e banco de dados. Isso significa que, quando as informações são adaptadas e licenciadas sob uma dessas licenças, você atende automaticamente às condições do OGL quando cumpre a outra licença. O OGLv3.0 é compatível com Open Definition.
Itália	Dati.gov.it https://www.dati.gov.it/content/italian-open-data-license-v20 https://www.dati.gov.it/content/italian-open-data-license-domande-risposte	-Creative Commons Attribution 4.0 International (CC-BY 4.0) (12732) -Creative Commons CC0 1.0 Universal - Public Domain Dedication (CC0 1.0) (737) -Creative Commons Attribution - ShareAlike 4.0 International (CC- BY-SA 4.0)(365) -Creative Commons Attribuzione 2.5 Italia (CC-BY 2.5 IT) (73) -Creative Commons Attribuzione 3.0 Italia (CC-BY 3.0 IT) (5) -Italian Open Data License 2.0 (3562) Personalizada -Italian Open Data License 1.0 (32) - Personalizada -Open Data Commons Open Database License 1.0 (1)	É possível combinar os dados divulgados no IODL com informações liberadas com outras licenças? Sim. A licença do IODL 2.0, em particular, permite que você crie trabalhos derivados também usando dados lançados com outras licenças « atribuição » (que exigem apenas a indicação da fonte), como as licenças Creative Commons Attribution (CC- BY) e o Atributo de Commons de Dados Abertos (ODC- BY). Todas as outras licenças Creative Commons e Open Data Commons que incluem um requisito de “Atribuição” (por exemplo, CC BY-SA ou ODC ODbL) também podem ser usadas para liberar trabalhos derivados.
Alemanha	Govdata.dev https://www.govdata.de/web/guest/lizenzen	-Licença de dados Alemanha - Atribuição - Versão 2.0 - Personalizada -Licença de dados Alemanha - Zero - Versão 2.0 - Personalizada	Somente essas duas opções de licença.
Noruega	Data.norge.no http://data.norge.no/datahotellet	Licença de Dados Públicos da Noruega - NLOD - Personalizada -Creative Commons Attribution 4.0 Internacional Creative Commons Dedicção de Domínio Público (CC0 1.0)	Todos os dados contidos na propriedade de dados devem ser licenciados com qualquer uma das seguintes licenças abertas: Licença de Dados Públicos da Noruega - NLOD, Creative Commons Attribution 4.0 Internacional ou Creative Commons Dedicção de Domínio Público (CC0 1.0)
Espanha	Datos.gob.es http://datos.gob.es/es/documentacion?f%5B0%5D=field_doc_tx_tipo%3A316	Não existe licença, dados abertos é baseado na legislação nacional espanhola.	

Bélgica	Data.gov.be https://data.gov.be/fr/aspects-legaux	Licença Open Data Flandre Etalab Open License Licença Open Data Bruxelles Licença específica CC Zero (CC 0) Statbel Open License Licença aberta específica Licença Open Data Bruges CC Compartilhamento de equivalência de CC (CC BY SA) GNU Free Documentation License ODC Public Domain DL CC Attribution (CC BY) Atribuição de ODC (ODC-BY)	Não estabelecem quais licenças podem ser usadas, então as diversas organizações utilizam diversas licenças.
Canadá	https://open.canada.ca/en/open-government-licence-canada	-Open Government Licence - Personalizada	
Estados Unidos	https://project-open-data.cio.gov/open-licenses/	-Creative Commons Zero Dedicção de Domínio Público (CC0), -Dedicção e Licença de Domínio Público do Open Data Commons (PDDL) -Commons de dados (ODC-By) -Licença aberta de banco de dados (ODbL) do Open Data Commons Atribuição Creative Commons (CC BY) -Creative Commons Attribution-ShareAlike(CC BY-SA) -GNU Free Documentation License	Não estabelecem quais licenças podem ser usadas, então as diversas organizações utilizam diversas licenças.
Coreia do Sul	https://www.data.go.kr/guide/guide/policy.do	-Creative Commons	
Austrália	https://toolkit.data.gov.au/index.php/Main_Page	-Creative Commons Attribution 3.0 Australia	
Nova Zelândia	https://www.data.govt.nz/terms-of-use/	-Creative Commons Attribution 4.0 International	
México	https://datos.gob.mx/libreusomx	-Creative Commons BY	

Em 6 dos portais analisados, os governos desenvolveram licenças próprias e personalizadas. Esta característica pode ser vista como controversa ao se questionar em alto nível quais as necessidades e especificidades que foram identificadas que levaram a definir estas licenças personalizadas, sendo necessário uma análise detalhada das mesmas.

As legislações governamentais e as licenças que tratam das políticas de uso dos dados são itens críticos no contexto dos DAG, entretanto é possível observar com frequência que os governos não as definem e as organizam adequadamente. Assim, um grande número de licenças demonstra a urgência de se definir um arcabouço comum e limitado de licenças, caso contrário gerará insegurança jurídica e será um limitador para as partes interessadas.

O uso de licenças padrões, Creative Commons e Open Data Commons, geralmente são melhores opções do que as licenças personalizadas, pois podem proporcionar eficiência organizacional e redução de custos, além de gerar uma maior interoperabilidade dos dados, assim como um maior

conhecimento das partes interessadas dos termos das licenças, gerando mais conformidade (KORN *et al.*, 2011).

Observa-se que o PBDA apresenta baixo nível de maturidade com relação à definição de licenças de dados, não há qualquer definição por parte do Governo Federal brasileiro de um rol de licenças que devam ser utilizadas. Isto gera uma anomalia, onde centenas de bases de dados são disponibilizadas no PBDA sem sequer ter uma licença especificada, logo demonstra ser emergente que o governo atue nesta problemática. Desta forma, é urgente que o governo brasileiro realize estudos a fim de definir as licenças a serem utilizadas.

5.3 Propostas de melhorias da tecnologia do PBDA

Com o clamor social por transparência, participação social e fiscalização sob os governos, a descoberta científica intensiva em dados abertos se torna mais importante e desafiadora. Assim, a partir das análises feitas nesta pesquisa, propomos funcionalidades que podem ser

implementadas em uma plataforma de Dados Abertos Governamentais (DAG) para que seja mais colaborativa.

As funcionalidades propostas buscarão implementar técnicas, boas práticas e processos que promoverão um ecossistema colaborativo de “curadoria”, coprodução, compartilhamento, rastreamento de procedência e reprodutibilidade. Isso proporcionará que os dados abertos promovam mais valor social e econômico ao longo do tempo, desenvolvendo assim uma melhoria em um serviço público.

Por se tratar de DAG, os serviços disponíveis devem ser de acesso livre sem ônus, buscando a promoção de uma melhor gestão e governança dos DAG. As propostas buscam atender o Governo Federal, sendo uma solução que poderá aprimorar o PBDA. Assim, trata-se de uma proposta de melhoria de uma tecnologia existente.

Dessa forma, a partir dos achados identificados e das análises, essa pesquisa propõe as seguintes funcionalidades e características:

- Interações sociais entre usuários e governo;
- Funcionalidade que permita aos usuários adicionar diferentes tipos de dados;
- Competições de análise de dados acessível a todos;
- Integração *top-down* entre esferas governamentais e órgãos, com o objetivo de ultrapassar o partidário político e a fragmentação da organização e seus setores;
- Permite que vários usuários sejam comantenedores de um conjunto de dados públicos;
- APIs abertas;
- Ferramenta para envolver as partes interessadas e usá-las para apoiar discussões baseadas em dados. Dessa maneira, os usuários poderão criar visualizações diretamente de conjuntos de dados armazenados nos portais, filtrando e agrupando seu conteúdo e publicando o resultado nas discussões em andamento para apoiar as argumentações;
- Funções que permitam a filtragem e o agrupamento de dados e visualizações;
- Salas, locais públicos virtuais onde as partes interessadas possam se encontrar e discutir sobre um tópico específico, formando ingressarem comunidades de interesse existentes, nas quais as discussões são suportadas por dados e suas visualizações por meio da criação e compartilhamento;
- Cocriação de dados: as partes interessadas podem conhecer e criar novos conjuntos de dados em colaboração, podendo adicionar dados de forma colaborativa e interativa ao conjunto de dados, que pode ser criado a partir do zero ou importando um já existente;

- Ambientes de divulgação, onde todos os artefatos, documentos e achados na plataforma possam ser exibidos publicamente fora da plataforma, para acesso livre a qualquer pessoa.

Conforme o que foi descrito na análise das partes interessadas dos portais abordados, percebe-se um diferencial cultural em relação às plataformas brasileiras e as americanas, sendo importante salientar que no caso da SPOD, de origem europeia, existem similaridades, tal diferença possui relação com a forma de financiamento. Uma ferramenta de extremo sucesso (levando em consideração sua massiva utilização e certa credibilidade), o Wikipedia serve como um bom parâmetro para proposição de como uma plataforma deve operar para engajar as partes interessadas a se relacionarem na construção de conhecimento, servindo assim como modelo de participação colaborativa.

Quanto à aspectos de governança propõe-se para o desenvolvimento de uma nova versão do PBDA:

- Modelo de financiamento por doações individuais, similares às do wikipedia, em conjunto com apoio governamental;
- Modelo de financiamento que envolva pacotes de benefícios concedidos para aqueles que se disponham a serem doadores mensais regulares. Com uma estruturação de uma “carreira de patronato”;
- Fomento de convênios a serem realizados entre empresas privadas e a plataforma, para construção de ferramentas de governança que permita uma maior transparência por parte dos dados de ambas as partes, incluindo, assim, organizações com outras características, como ONG’s, OS’s, etc ;
- Fomento à participação de escolas, para que, desde cedo alunos aprendam linguagens básicas de programação e formas de tornar um ambiente de educação mais transparente, participativo, convidativo e prático;
- Envolvimento direto de universidades públicas com a instituição de projetos de participação tanto de alunos quanto de professores;
- Desenvolvimento de uma cultura de competição construtiva, que envolva a sociedade civil e a academia.

6. CONCLUSÃO

Uma governança adequada possibilita uma melhor gestão e a participação social na coisa pública, permitindo que a organização execute ações estratégicas para alcançar os objetivos desejados com menor esforço. Para isso, o planejamento destas ações é primordial e as decisões para a definição, acompanhamento e execução destas mudanças podem ser apoiadas pela governança dos dados.

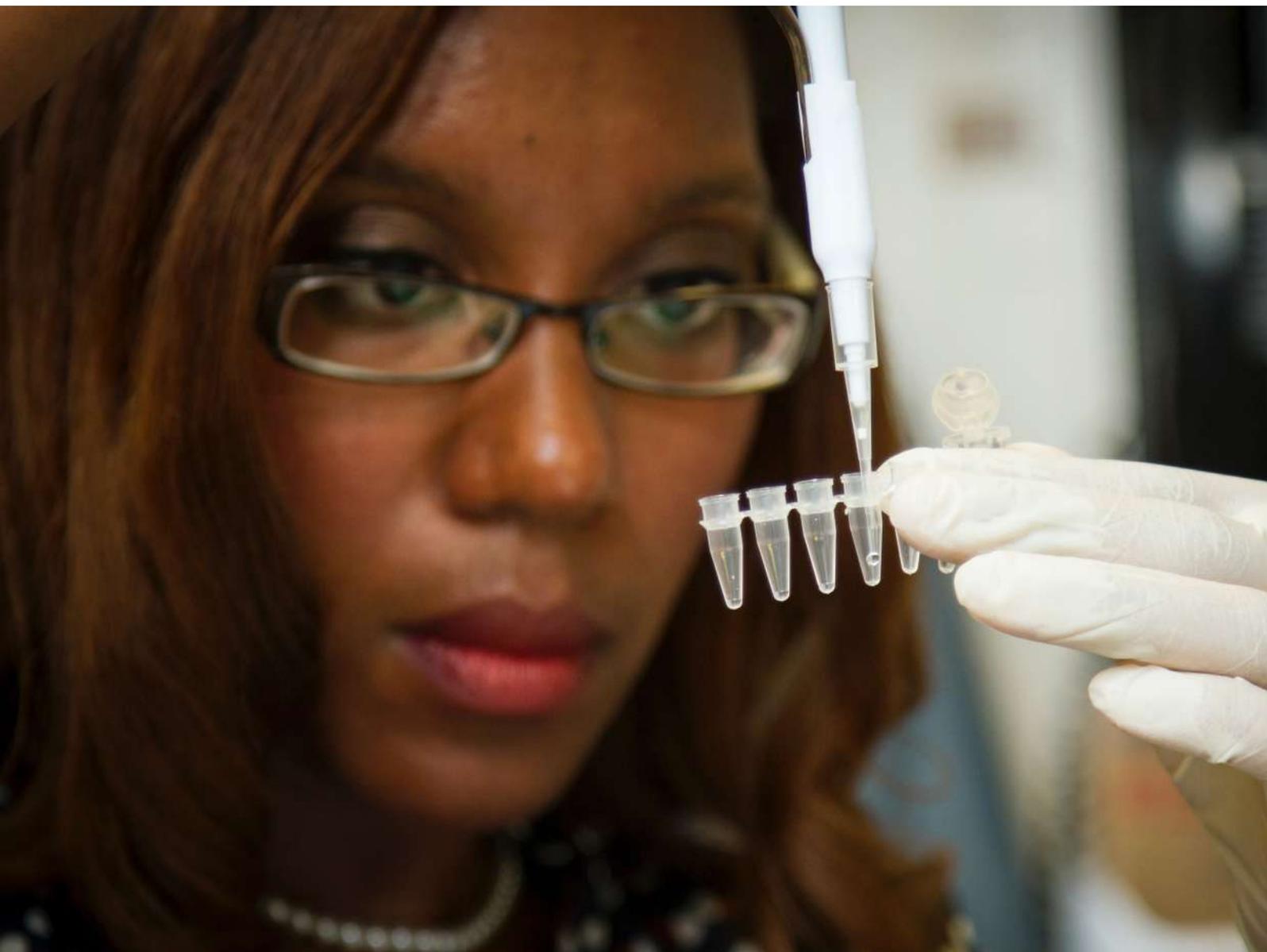
Esta pesquisa propôs uma série de questões e soluções essenciais a serem aprimoradas quanto ao contexto dos dados nas ações estratégicas dos diferentes órgãos da APF para que os projetos, processos e decisões sejam mais assertivas e eficazes, melhorando o serviço e a gestão do negócio destes órgãos para que possam cumprir suas funções finalísticas de forma otimizada e para que estes avanços tragam um retorno positivo à sociedade.

A administração pública federal poderá utilizar os achados desta pesquisa para subsidiar a implementação de novas funcionalidades e versões dos portais públicos e para atualizar as políticas públicas relativas ao acesso à informação.

Além disso, as análises realizadas sobre os *stakeholders* e as licenças possibilitam um entendimento melhor do ambiente e das partes interessadas envolvidas, podendo servir de base para novas políticas, processos, atividades e pesquisas sobre o tema. Pois não há o que se falar em transparência, participação, *accountability* e dados abertos governamentais em um ambiente que não haja uma governança e gestão adequadas, pessoas aptas, recursos disponíveis, políticas públicas atuais, enfim, um ecossistema de dados governamentais adequado e em constante adaptação às transformações contínuas.

Dessa forma, fica claro que alinhar as características de *crowdsourcing* à temática de dados abertos é de extrema importância, pois um sistema colaborativo tem a capacidade de suportar as organizações, e as partes interessadas a inovar e solucionar problemas externos e internos. Como vantagem, fica evidente que o *crowdsourcing* proporciona novos conhecimentos a baixo ou zero custo, além da possibilidade de ideias novas dos usuários e contribuições, podendo gerar inovações. A estratégia de uma plataforma colaborativa permite maior interação e colaboração, tendo em vista que os usuários podem colaborar e, dessa forma, se sentem parte de um projeto, deixando o perfil de apenas consumidor. Isto tem impacto direto no engajamento da sociedade.

Uma possibilidade de trabalho futuro é o estudo aprofundado das licenças de dados abertos buscando uma solução para a anomalia que ocorre no governo brasileiro quanto a não definição de licenças. Outra recomendação é utilizar os achados desta pesquisa como uma das fontes para desenhar, levantar e analisar os requisitos de desenvolvimento de software de uma plataforma de compartilhamento “Dados Abertos Governamentais” para o Governo Federal do Brasil, que crie um ambiente de dados abertos que atenda a sociedade civil e a administração pública de uma maneira mais moderna e colaborativa, coproduzindo e cogerenciando as bases de dados para que tenham mais valor público e científico no coletivo.



REFERÊNCIAS BIBLIOGRÁFICAS

ATTARD, J.; ORLANDI, F.; SCERRI, S.; AUER, S. *A systematic review of open government data initiatives*. *Government Information Quarterly*, v. 32, n. 4, p. 399- 418, 2015.

BARATA, André Montóia. *Governança de dados em organizações brasileiras: uma avaliação corporativa entre os benefícios previstos na literatura e os obtidos pelas organizações*. 2015. 154 f. Dissertação (mestrado) - Curso de PósGraduação em Sistemas de Informação, Universidade de São Paulo, São Paulo, 2015.

BENSON, D. A., CAVANAUGH, M., CLARK, K., KARSCH-MIZRACHI, I., LIPMAN, D. J., OSTELL, J.; SAYERS, E. W. GenBank. *Nucleic acids research*, 41(D1), D36-D42. 2012.

BRASIL, Controladoria-Geral da União. *Acesso à informação pública: uma introdução à Lei 12.527, de 18 de novembro de 2011*. Brasília: Imprensa Nacional, 2011.

BRASIL, Controladoria-Geral. *Acesso à informação pública: uma introdução à Lei nº 12.527 de 18 de novembro de 2011*. In: *Acesso à informação pública: uma introdução à Lei nº 12.527 de 18 de novembro de 2011*. 2012.

BRASIL. CGU. *Parceria para governo aberto*. 2018. Disponível em: <<http://www.cgu.gov.br/assuntos/transparencia-publica/governo-aberto>>. Acesso em: 22 jan.

BRASIL. Constituição da República Federativa do Brasil, 1988. *Centro de Documentação e Informação, Coordenação de Publicações*, 2004.

CONTROLADORIA-GERAL DA UNIÃO – CGU. *Páginas da Transparência Pública*. Brasília, 2006. Disponível em: <<http://www3.transparencia.gov.br/>>. Acesso em: 17 mar. 2018.

CONTROLADORIA-GERAL DA UNIÃO – CGU. *Portal da Transparência*. Brasília, 2012. Disponível em: <<http://www.portaldatransparencia.gov.br/>>. Acesso em: 05 fev. 2018.

COSTA, L. S. *Big Data estratégico: um framework para gestão sistêmica do ecossistema Big Data*. Rio de Janeiro: COPPE UFRJ (Mestrado em Engenharia de Produção), 2016.

DAVIES, T. *Open Data Barometer 2013 Global Report*. *World Wide Web Foundation and Open Data Institute*. 2013. Disponível em: <http://www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf>>. Acesso em: 14 mar. 2018.

DEMCHENKO, Y.; GRUENGARD, E.; KLOUS, S. *Instructional model for building effective Big Data curricula for online and campus education*. In: 1ST IEEE STC CC AND RDA Workshop on Curricula and Teaching Methods In Cloud Computing, Big Data, and Data Science, Singapore, dez. 2014. Disponível em: <https://www.researchgate.net/publication/273945502_Ins-vtructional_Model_for_Building_Effective_BigData_Curricula_for_online_and_Campus_Education>. Acesso em: 10 mar. 2018.

DMBOK.MOSLEY, M.; BRACKETT, M.; EARLEY, S.; HENDERSON, D. *The DAMA Guide to The Data Management Body of Knowledge: DAMA - DMBOK Guide*. Estados Unidos: technics publications, 2009.

EUREKALERT. *Oxford, Harvard scientists lead data-sharing effort*. 2012. Disponível em: <https://www.eurekalert.org/pub_releases/2012-01/hu-ohs012612.php>. Acesso em: 22 mar. 2018.

EVANS, A. M.; CAMPOS, A. Open government initiatives: challenges of citizen participation. *Journal of Policy Analysis and Management*, 32(1), 172-185, 2013.

FERRAZ, I. R., GOUVÊA, M. A.; BARRETO, I. F. Fatores determinantes da participação em iniciativas de crowdsourcing. *Revista Gestão & Tecnologia*, 17(1), 107-129, 2017.

FERRER-SAPENA, Antonia; PESET, Fernanda; LEIXANDBENAVENT, Rafael. Access to and reuse of public data: open data and open government. *Profesional de la Informacion*, v.20, Issue: 3, pp. 260-269, 2011.

GENEBIO. *Search Results for: genbank*. 2018. Disponível em: <<http://www.genebio.ufba.br/?s=genbank&submit=>>>. Acesso em: 22 jan. 2018.

GIDDENS, A. *The Nation-State and Violence: volume 2 of a contemporary critique of historical materialism*. Cambridge: Polity Press, 1985.

HARRISON, Teresa M. *et al. Open government and e-government: democratic challenges from a public value perspective*. *Information Polity*, v. 17, n. 2, p. 83-97, 2012.

HICSS. HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES 2018. 2018. Disponível em:

<<https://scholarspace.manoa.hawaii.edu/handle/10125/49887>>. Acesso em: 22 mar. 2018.

HOFFMANN, L. *Data mining meets city hall*. *Communications of the ACM*, 55(6), 19-21. 2012.

HOWE, J. The rise of crowdsourcing. *Wired Magazine*, v. 14, n. 6, jun. 2006. Disponível em: <http://www.wired.com/wired/archive/14.06/crowds.html>. Acesso em: 07 de março de 2018.

IBM. *The IBM Data Governance Council Maturity Model: building a roadmap for effective data governance*. Nova York: IBM, 16 p.,2007.

ICIST. INTERNATIONAL CONFERENCE ON INFORMATION SOCIETY AND TECHNOLOGY. 2018. Disponível em: <<http://www.eventiotic.com/eventiotic/conference/icist2018>>. Acesso em: 22 mar. 2018.

KAGGLE. *We've passed 1 million members*. 2017. Disponível em: <<http://blog.kaggle.com/2017/06/06/weve-passed-1-million-members/>>. Acesso em: 22 mar. 2018.

KORN, Naomi; OPPENHEIM, Charles. *Licensing open data: a practical guide*. Higher Education Funding Council for England. [Consulta: 14/2/2014]. Disponível em: <http://>

discovery. ac. uk/files/pdf/Licensing_open_data_a_practical_guide. pdf, 2011.

KOSOVA, Gülüm *et al.* Genome-wide association study identifies candidate genes for male fertility traits in humans. *The American Journal of Human Genetics*, v. 90, n. 6, p. 950-961, 2012.

LEPTIN, M. *Open access—pass the buck.* 2012.

MSDSE. *Themes.* 2018. Disponível em: <<http://msdse.org/themes/>>. Acesso em: 22 mar. 2018. ODI OPEN DATA INSTITUTE. *Mapping data ecosystems.* 2018. Disponível em: <<https://theodi.org/article/mapping-data-ecosystems/>>. Acesso em: 22 mar. 2018.

OGP - OPEN GOVERNMENT PARTNERSHIP. *Open by default, policy by the people, accountability for results,* 2011. Disponível em: <http://www.opengovpartnership.org/sites/default/files/091116_OGP_Booklet_digital.pdf> Acesso em: 17 mar. 2018.

OGP. OPEN GOVERNMENT PARTNERSHIP. *About OGP.* Disponível em: <<http://www.opengovpartnership.org/about>>. Acesso em: 22 mar. 2018.

PANIAN, Zeljko. *Some practical experiences in data governance.* World Academy of Science, Engineering and Technology, S.i., v. 6, n. 2, p. 939-946, 2010.

PBDA - PORTAL BRASILEIRO DE DADOS ABERTOS. *Sobre o dados.gov.br.* 2018. Disponível em: <<http://dados.gov.br/pagina/sobre>>. Acesso em: 22 mar. 2018.

PELTONIEMI, Mirva; VUORI, Elisa. *Business ecosystem as the new approach to complex adaptive business environments.* In: Proceedings of eBusiness research forum. 2004.

SADIQ, S. *Handbook of Data Quality Research and Practice.* Brisbane, Austrália: Springer, 2013.

SCHOLL, H. J. *Five trends that matter: challenges to 21st century electronic government.* *Information Polity*, v. 17, p. 317-327, 2012.

SCHOLL, H. J. *Five trends that matter: challenges to 21st century electronic government.* *Information Polity*, v. 17, p. 317-327, 2012.

THUAN, N. H.; ANTUNES, P.; JOHNSTONE, D. A process model for establishing business process crowdsourcing. *Australasian Journal of Information Systems*, 21, 2017

UBALDI, B. *Open government data: towards empirical analysis of open government data initiatives.* OECD Working Papers on Public Governance, n. 22, p. 01, 2013.

UBALDI, B. *Open government data: towards empirical analysis of open government data initiatives.* OECD Working Papers on Public Governance, n. 22, p. 01, 2013.

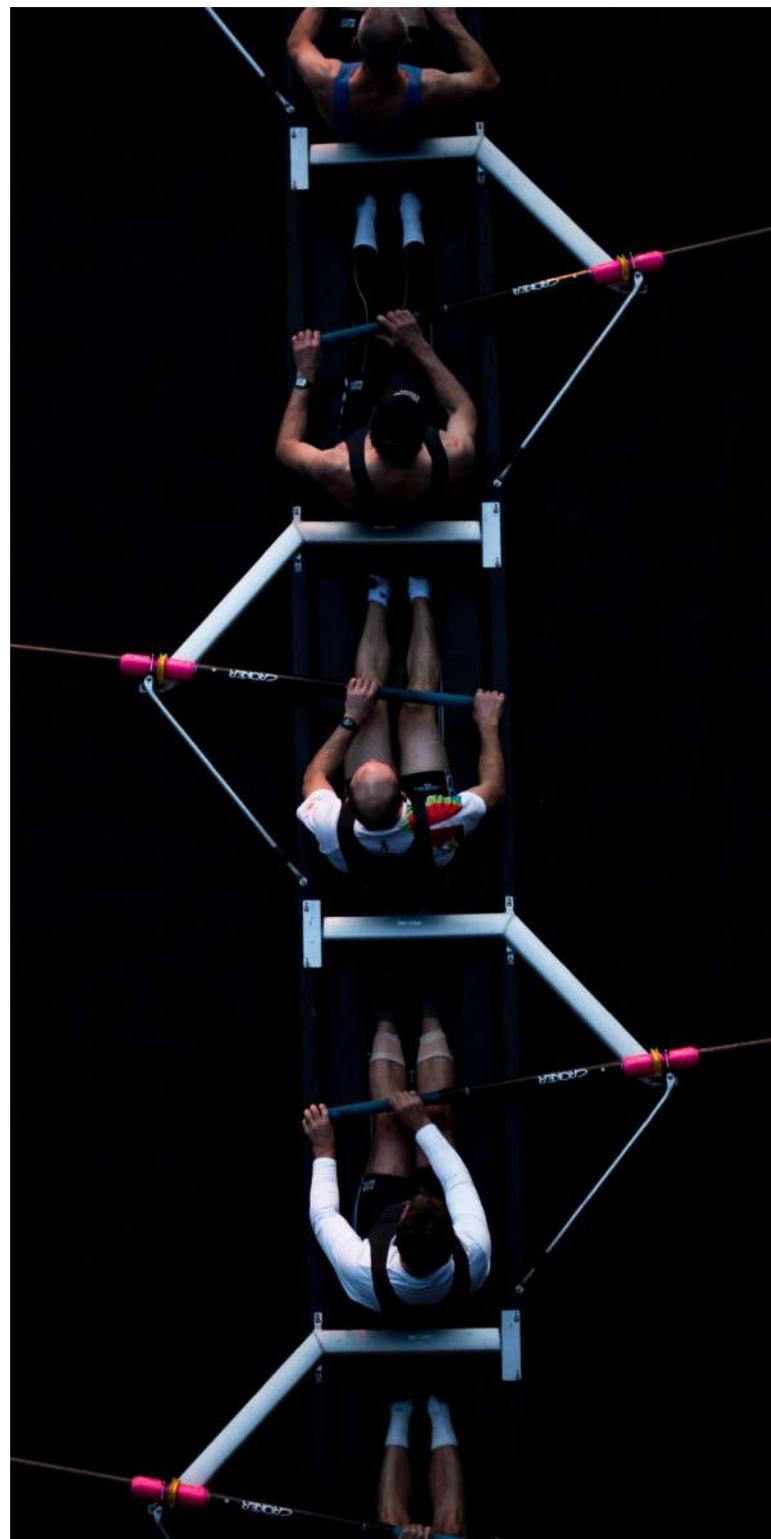
VELJKOVIĆ, N.; BOGDANOVIĆ-DINIĆ, S.; STOJENOV, L. B. *Government: an open data perspective.* *Government Information Quarterly*, v. 31, n. 2, p. 278-290, 2014.

WEBER, M. *Economía y Sociedad.* Madrid: fondo de cultura económica, 1994.

WESTERSKI, A.; IGLESIAS, C. A.; GARCIA, J. E. *Idea relationship analysis in open innovation crowdsourcing systems.* In: Collaborative Computing: networking, applications and worksharing (CollaborateCom), 2012 8th International Conference on. IEEE, p. 289- 296.

YANNOUKAKOU, A.; ARAKA, I. *Access to government information: right to information and open government data synergy.* *Procedia - Social and Behavioral Sciences*, v. 147, p. 332- 340, 2014.

ZHU, H.; DJURJAGINA, K.; LEKER, J. Innovative behaviour types and their influence on individual crowdsourcing performances. *International Journal of Innovation Management*, v. 18, n.6, p. 1440015, 2014.



IOT-APP (IOT – ARMAZENAMENTO & PROCESSAMENTO NA ÁREA PÚBLICA)

Mario Antonio

Resumo

O projeto denominado de **IoT-APP** (*IoT – Armazenamento & Processamento na área Pública*) propôs um estudo que pudesse auxiliar no entendimento de como se proceder na captura, no armazenamento de **dados abertos** em larga escala na área de saúde, além do processamento dos dados, considerando-se a abordagem de IoT para o segmento da saúde pública.

Desta forma, a proposta de pesquisa, que teve como objetivo o desenvolvimento de uma investigação na área de sistemas de computação, com ênfase nos dispositivos *smartbands* e o posterior armazenamento paralelo e/ou distribuído de sinais vitais pessoais abertos para a área de atenção primária de saúde.

Em outras palavras, o objeto de pesquisa foi estudar cenários relativos a **automação** para armazenamento e **análise de grandes volumes de dados abertos** em configurações de **IoT** para a área de atenção primária de saúde pública. Ao final do trabalho de pesquisa ocorreu uma pequena implementação de um protótipo de um cenário orientado a prestação de apoio a uma unidade pública primária de saúde.

INTRODUÇÃO

Mario Antonio

Professor. PhD em Ciência da Computação. Pós-doutorado em Ciência da Computação, Bolsista da Cátedra de Pesquisa ENAP. Professor Titular da Universidade Federal de Juiz de Fora.

Atualmente, verificam-se no mundo inúmeras tendências e avanços tecnológicos que podem ser traduzidos, se bem desenvolvidos e empregados, em uma melhoria significativa na qualidade de vida da sociedade moderna. Exemplos de tais tecnologias são aquelas que encontramos em abordagens tais como na realidade virtual e aumentada, na utilização da inteligência artificial, na adoção do paradigma conhecido como Internet das Coisas (*Internet of Things* - IoT), na análise de grandes volumes de dados, no paradigma conhecido como *blockchain*, nas indústrias aditivas, no uso dos sensores com funções díspares, em uma automação orientada à melhor eficiência, não só de processos, mas, principalmente das pessoas, no uso da geolocalização e na utilização da robótica. Desta forma, imagina-se que esses paradigmas devam começar a despertar nos agentes públicos um maior interesse na utilização dessas abordagens pelo potencial transformador para área pública.

Estudos recentes de grandes corporações de TI, como por exemplo, Cisco [1] e IBM [2], indicam que diariamente são gerados 2.5 *exabytes* de bytes de dados por dia. Estima-se que em 2020 esse número atinja a casa dos 40 *yottabytes*, o que significa aproximadamente 5.200 gigabytes para cada pessoa na terra. Esta grande quantidade de informação deverá ter uma parcela majoritariamente originada da abordagem Internet das Coisas (IoT), através de coisas e pessoas.

Como sugerido em [3], é importante que haja uma discussão do processamento distribuído do IoT, utilizando-se configurações de alto desempenho, tais como, por exemplo, os *clouds* e os *fogs* computacionais, visando a possibilidade de atendimento de diferentes solicitações em tempos compatíveis com suas requisições. Desta diferenciação surge, também, a necessidade de uma forma diferencial na maneira pela qual deve-se **armazenar** essa quantidade de dados, volume esse até então não imaginado, e como efetuar o seu **processamento**.

Um dos desafios mais discutidos atualmente no meio técnico-científico com certeza é o paradigma *big data*. Esse paradigma, muitas vezes mal interpretado, é entendido como se fosse uma única metodologia, que pode por si só responder às necessidades de armazenamento, manipulação e processamento dos dados estruturados (SQL) e não-estruturados (NOSQL). Por outro lado, em nosso dia-a-dia somos levados ao desafio de melhor entender como armazenar uma quantidade massiva de dados gerados por diversas tecnologias e formas diferenciadas de tratá-las e processá-las.

Quanto ao diferencial de armazenamento para o paradigma de *big data* muitos esperam que os chamados 5V's (volume, variedade, veracidade, velocidade e valor) sejam tratados de maneira única com alto desempenho e com facilidades de tolerância às falhas. Os dados paralelos e/ou distribuídos, o número de nodos clientes da configuração, o número de servidores de metadados, o número de servidores de armazenamento e a configuração da rede de interconexão são elementos a serem considerados em uma modelagem de uma proposta que possa atender aos requisitos de armazenamento de uma forma inovadora e eficiente para o paradigma *big data*. Em outras palavras, têm-se atualmente grupos de pesquisas que trabalham com as abordagens de NOSQL e NEWSQL como se fossem as soluções para natureza do problema de *big data*. Todavia, muitas vezes esses dois paradigmas têm seu nicho de mercado e podem não comportar o tratamento de uma forma aberta às informações quando consideramos a casa dos *yottabytes*.

Em nossas pesquisas anteriores, como ilustrado em [4] e [5], por exemplo, apontamos a importância do mapeamento da qualidade de experiência (QoE), através do uso da qualidade de serviço (QoS) para configurações das arquiteturas de bancos de dados distribuídos. Fizemos, também, um levantamento bibliográfico indicando a relevância em pesquisas científicas dos parâmetros de desempenho, eficiência energética, configurações de

nuvem e *big data*. Como se apresenta em [6], os serviços de *IoT* devem ser baseados em infraestruturas orientadas ao contexto, visando um diferencial no provimento dos serviços. Em outra pesquisa ilustrada em [7] existe uma ratificação da necessidade de uma maior especialização da infraestrutura para, por exemplo, o caso dos ambientes domiciliares assistidos (*Ambient Assisted Living- AAL*). Essas pesquisas reconhecidas na área acadêmica internacional apontam para a necessidade de novas abordagens que possam trabalhar o armazenamento e o processamento de uma forma diferenciada, com desempenho, sob a ótica do consumidor das informações (QoE) e com qualidade (QoS).

De outra forma, pode-se indicar que não adianta, por exemplo, no setor público, haver um grande esforço no sentido de coleta de informações, nas mais diferentes áreas (segurança, saúde, educação, agricultura, controle de vias públicas) se não houver pesquisas que indiquem como **armazenar** de maneira diferenciada o que deverá ser **processado** dentro de um **tempo apropriado** pelo agente público. Em adição, sem que os dados não possam ser interoperáveis entre diferentes plataformas, ou seja, que o paradigma de *open data* não seja observado. Conforme já comentado, alguns trabalhos focam muitas vezes em ferramentas do tipo NOSQL e NEWSQL sem considerar a abordagem de grandes volumes de dados que só podem ser tratados por sistemas de arquivos distribuídos e paralelos. De outra forma, é necessário que pesquisas relacionadas ao **armazenamento** sejam desenvolvidas, independentes de ferramentas, para que o **processamento** seja efetuado considerando-se a **restrição temporal do usuário**.

Um exemplo clássico sobre o tema de grandes volumes de dados atuais, conhecido internacionalmente, é o caso do processamento dos cartões de crédito no horário de almoço em país como a China. No período do almoço o número de pessoas utilizando-se do cartão de crédito pode ser estimado em **n vezes toda a população do Brasil**. Assim, surge um problema de pesquisa:

- Como **armazenar** eficientemente os dados, de uma forma aberta, para que o **processamento** possa ser efetuado com uma **restrição temporal** (isto é, **horário do almoço e em minutos**)?

Com foco nos desafios públicos, a presente proposta pode ser considerada como uma proposta potencial de transformação para diversas áreas, como por exemplo:

- a. Saúde pública: no controle remoto de pacientes, através de seus dados, considerando-se a abordagem de **Ambientes Domiciliares Assistidos (AAL)** [25]; controle de uma grande quantidade de pessoas geograficamente distribuídas e poder submeter seus dados de saúde para análises analíticas constantes deste grande volume de dados (*big data*) para profissionais remotamente localizados;
- b. Segurança pública: o paradigma de *big data analytics* onde indivíduos, identidades, veículos, armas e outros itens poderão ser processados

com mais eficiência, por diferentes agências públicas, por causa do maior volume de dados armazenados de forma correta e que possam ser trocados por entidades públicas diferentes (polícia civil, militar e federal, por exemplo);

- c. Serviços públicos: como água, luz, coleta de lixo e gás poderão ser melhores entendidos em termos de padrão se o armazenamento puder representar uma forma de interoperabilidade para o processo de ETL (Extração, Transformação e Carregamento), como ilustrado em [29].

Desta forma, o presente documento visa apresentar uma proposta de pesquisa, que tem como objetivo o desenvolvimento de uma investigação na área de sistemas de computação, com ênfase no armazenamento paralelo e/ou distribuído, de dados de **sensores** pessoais, que possa prover uma modelagem com **automação** para armazenamento e **análise de grandes volumes de dados** em configurações de **IoT**. A proposta, denominada de **IoT-APP** (*IoT – Armazenamento & Processamento na área Pública*) visa propor uma modelagem que possa auxiliar no entendimento de como funciona o armazenamento de **dados abertos** em larga escala e que auxilie o processamento de dados da área pública oriundos da abordagem de *IoT*. Espera-se, ao final do presente estudo, a implementação de protótipo de um cenário com um paradigma orientado à prestação de um serviço público.

1. CONTEXTUALIZAÇÃO E JUSTIFICATIVA

A ideia que o armazenamento é tão somente uma operação simples de leitura/escrita de dados para sistemas especializados de armazenamento não corresponde mais ao grande desafio atual de conexão, transmissão, armazenar, *restore* e o replicar dos dados na casa dos *exabytes*, caminhando para *yottabytes*. O trabalho de pesquisa apresentado no conceituado periódico internacional ACM Computing Surveys [7] indica claramente que dada a quantidade e complexidade dos dados que atualmente são manipulados, abordagens que considerem o conhecimento do usuário (QoE) e a qualidade do serviço exigido (QoS) são importantes fatores para a realização, com sucesso, dos processos de armazenamento, consulta e manipulação das informações.

Por outro lado, a **cooperação** entre os sistemas distribuídos tais como os *clusters*, e *grids* computacionais, têm se tornado um processo complexo à medida que mais recursos podem ser agregados em *clouds* e as aplicações podem fazer uso desses recursos heterogêneos distribuídos. Importante ressaltar que, na maioria das vezes, sob diferentes autoridades e políticas de uso, além do crescimento da solicitação de eficiência energética e formas díspares de armazenamento [7].

Quando a configuração de processamento, e conseqüente armazenamento, devem-se, também, observar que podemos ter as configurações distribuídas, paralelas

e paralelas distribuídas como possíveis alvos de processamento e armazenamento. Desta forma, operações tais como armazenamento e *restore* de arquivos, por exemplo, devem considerar esquemas de tamanhos de faixas (*stripes sizes*) que sejam mais apropriados para os ambientes distribuídos e aqueles em que são orientadas as configurações paralelas.

A abordagem de IoT é caracterizada por possuir cenários que requerem as três configurações de armazenamento (distribuídas, paralelas e paralelas distribuídas) e que devem ser estudadas para cada tipo, ou conjunto de aplicação. Parâmetros tais como números de nodos da configuração, quantidade e tipo de servidores de armazenamento (distribuídos ou paralelos), as redes de interconexão existentes devem ser consideradas quanto da proposta de um projeto visando o armazenamento e, em especial, quando considerarmos a **Internet das Coisas**.

2. OBJETIVO GERAL E ESPECÍFICOS

A presente proposta de pesquisa tem como objetivo geral tentar responder a seguinte pergunta:

- Como **armazenar** um grande volume de dados abertos, na casa dos *yottabytes*, de maneira que os mesmos possam **ser processados** observando uma determinada restrição **temporal requerida** por um serviço público?

O exemplo do almoço de milhões de pessoas na China, passando seus cartões de crédito, é um cenário interessante de ilustração do desafio proposto como objetivo geral da pesquisa, denominada como **IoT-APP** (*IoT - Armazenamento & Processamento na área Pública*), considerando-se a abordagem de **dados abertos**. Para a presente proposta de pesquisa como indicado no recente renomado periódico internacional [36], existe atualmente um grande desafio na coordenação dos pequenos sensores, utilizados para monitoração de ambientes e pessoas, e o grande volume de dados (**abertos**) para os cuidados de saúde de uma forma digital.

Como objetivos específicos que se propõem como suporte para o desenvolvimento da presente proposta, temos:

- Efetuar um estudo de pesquisa (*survey*) detalhado sobre **sensores** e **dispositivos IoT** existentes no Brasil e exterior que possa ser utilizado na proposta ora apresentada;
- Envolver o maior número de orientados de graduação, mestrado e doutorado, que tenham interesse pelo tema, visando criar um grupo com uma competência diferenciada no tema proposto nesta pesquisa com foco no problema do serviço público;
- Propor um ambiente que possa representar uma configuração característica pública com sensores em uma configuração IoT e que tenha

uma automação para prover análise de grandes **dados abertos**;

- Modelar os componentes envolvidos e propor uma abordagem que seja capaz de responder às solicitações de uma aplicação de serviço público de *IoT*;
- Desenvolver um protótipo que trate as abordagens pesquisadas e que por meio de experimentos empíricos indique maneiras de execução de aplicações como melhores índices de desempenho para o armazenamento e processamento, referentes ao cenário público adotado;
- Disponibilizar os resultados em forma software aberto e de artigo para a Enap, se caso for de interesse divulgar os resultados da pesquisa em fóruns de interesse público, visando não somente a divulgação da pesquisa, mas, também, obter comentários e sugestões que possam prover intercâmbios com outros grupos nacionais e internacionais de pesquisa.

3. REFERENCIAL TEÓRICO

Nesta seção vamos apresentar, inicialmente, a abordagem de armazenamento com foco em aspectos que são importantes para um diferencial de desempenho. Desta forma, ilustramos os componentes dos nodos de computação, a rede de interconexão, dos nodos e dispositivos de armazenamento. Nesta subseção é dado um foco especial em distinguir as formas de armazenamento distribuído e paralelo, posto que os requisitos diferentes são cada vez mais recursivos das aplicações e menos conhecidos por aqueles que requerem um armazenamento diferenciado (optou-se pelo tema sistemas de arquivos, posto ser uma pesquisa nova com relação a grandes volumes de dados. As referências sobre NoSQL e NEWSQL são encontradas mais facilmente na literatura). Na segunda subseção, apresentamos a configuração típica conhecida como **Ambientes Domiciliares Assistidos**, também encontrada na literatura internacional como *Ambient Assisted Living* (AAL). O AAL é um ambiente interessante que pode ser considerado um exemplo característico de *IoT*. Esse ambiente é rico em desafios, posto aos seus componentes de sensores de monitoração de ambientes, sensores de sinais vitais, e a necessidade de redes de conexão com qualidade de serviços para o envio das informações e o requisito de armazenamento distribuído e/ou paralelo, dependendo do tipo de controle.

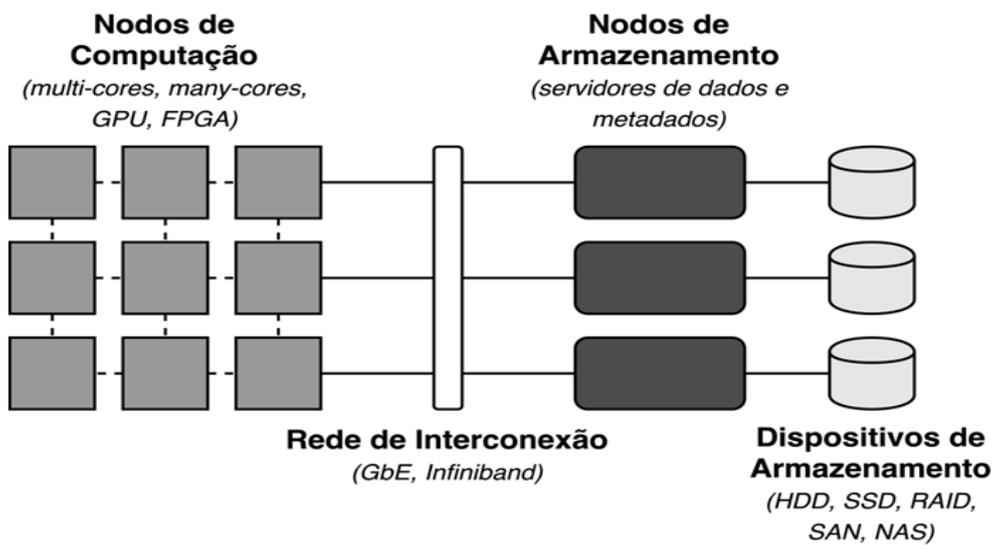
3.1 Armazenamento distribuído e paralelo

A abordagem de armazenamento ao longo dos anos vem mudando rapidamente devido a uma série de fatores, tais como: as tecnologias utilizadas nos dispositivos de armazenamento (exemplos são HDD, SAN e SSD), as redes

de interconexão (exemplos GbE e *Infiniband*) a forma de armazenamento (exemplos são as formas paralelas e distribuídas), nodos de computação (exemplos são os multicomputadores, multiprocessadores e arquiteturas

híbridas). A Figura 1 ilustra os componentes de uma configuração distribuída e/ou paralela de armazenamento com a possibilidade de utilização de díspares elementos.

Figura 1- Elementos em uma configuração distribuída e/ou paralela de armazenamento

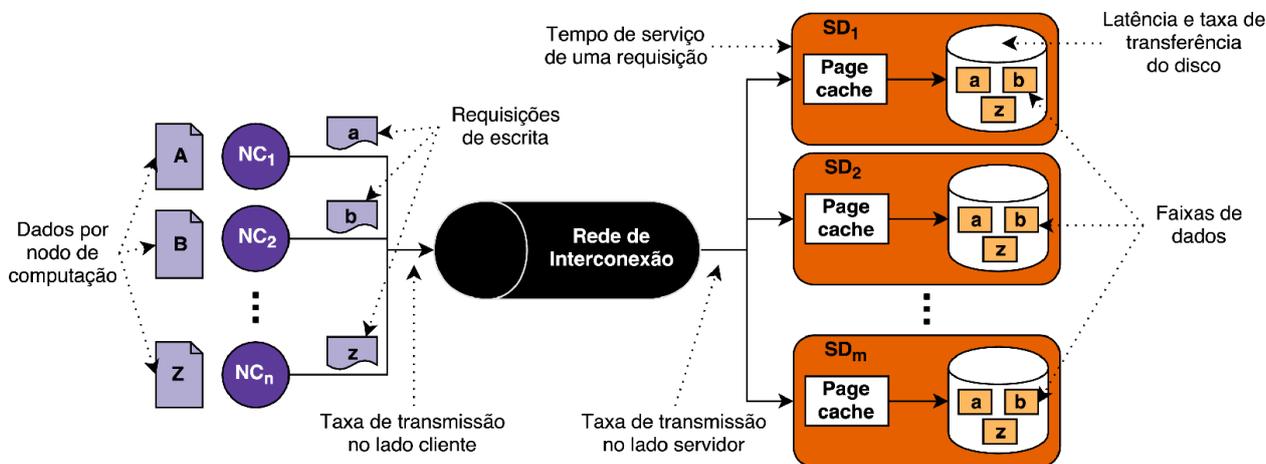


Fonte: Autor

Um modelo lógico, da Figura 1, de uma configuração de armazenamento distribuída e/ou paralela, pode ser visualizado com os elementos cliente, metadados e dados (propriamente ditos) sendo alocados em componentes da configuração. A Figura 1 ilustra de uma forma mais macro os componentes e suas possíveis formas de distribuição de elementos e suas conexões. Por outro lado, a preocupação daqueles envolvidos com a pesquisa em armazenamento é caracterizar as requisições, taxas de transmissão (tanto

no cliente como no servidor), forma de armazenamento nos dispositivos e as latências e faixas dos dados. Em outras palavras, a diferenciação do como armazenar é um aspecto de preocupação da pesquisa. Na Figura 2 ilustra-se de uma forma mais detalhada uma configuração de armazenamento distribuída. Na figura pode-se verificar um maior grau de detalhes dos elementos de um armazenamento, como foco distribuído, conforme sugerido no modelo lógico.

Figura 2- Exemplo de configuração de armazenamento distribuída.



Fonte: Autor

Em nossos trabalhos de pesquisas apresentados em [21], [22], [23] e [24], demonstramos a importância do entendimento dos inúmeros aspectos relativos aos sistemas paralelos e distribuídos em termos de armazenamento. A caracterização do sistema de arquivo paralelo para utilização em ambientes virtualizados apresentado em [21]. Na pesquisa publicada em [22] é tratado o impacto do *caching* do sistema operacional nos sistemas de arquivos paralelos. Já em [23] ilustra-se o importante conhecimento de múltiplos fatores para uma configuração diferenciada de um sistema de arquivo paralelo. Quanto aos aspectos relativos ao armazenamento distribuído, em [24] tratamos o tópico de persistência de dados de imagens médicas DICOM em sistemas hierárquicos.

4.2 Ambientes domiciliares assistidos

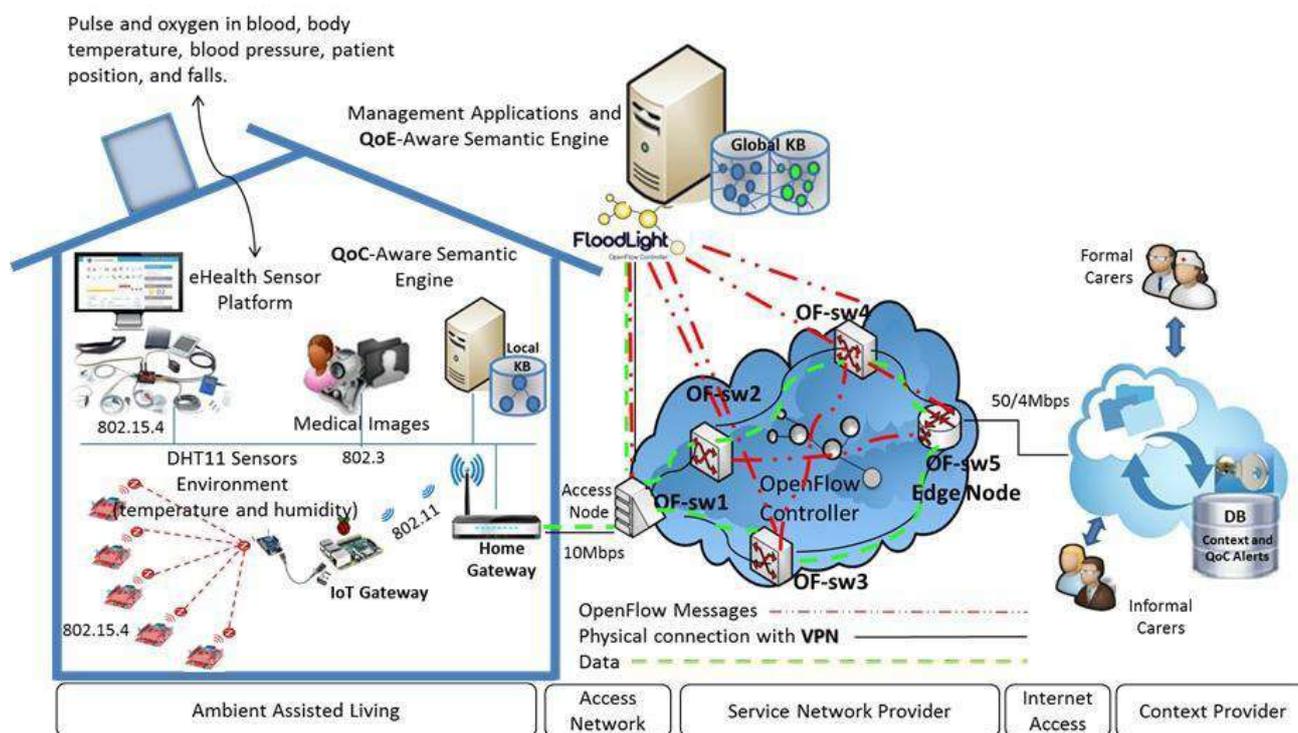
A abordagem de *Internet of Things* engloba um vasto ecossistema de aplicações, tais como as chamadas configurações *smart: cars, cities, energy, farms, grids, health, homes, parking* dentre outros esforços.

Nossos esforços de pesquisa na área se concentram majoritariamente na melhoria da qualidade de vida das pessoas, abordagem essa chamada de *Ambientes Domiciliares Assistidos*, também, conhecida na literatura internacional como *Ambient Assisted Living (AAL)*.

Nas pesquisas apresentadas em [25], [26], [27] e [28] procuramos contribuir com esforços na área de qualidade de parâmetros, gerenciamento de contextos para AAL, e monitoração dos ambientes com foco em *e-health*. Por outro lado, em [29] apresentamos uma pesquisa mais global na qual o foco está na *back-end* do processamento e armazenamento de uma grande quantidade de dados.

A Figura 3, oriunda do trabalho apresentado em [26], ilustra um cenário ampliado de uma configuração AAL, posto que além do ambiente propriamente dito, ilustra-se a rede de acesso, o provedor de serviço de rede, o acesso à internet e, por fim, o provedor de contexto.

Figura 3- Cenário lógico de um Ambiente Domiciliar Assistido [26].



Fonte: autor – Fonte: [26]

Por outro lado, em [26] são representados os componentes arquiteturais de proposta descrita no projeto de pesquisa intitulado *An eHealth Context Management and Distribution Approach in AAL environments*. Nesse projeto, que é um exemplo de IoT na saúde, demonstramos que a complexidade do ambiente e a grande quantidade de dados a serem armazenados devem ser suportados por abordagens diferenciadas utilizando-se semânticas com orientação à qualidade de contexto (QoC) e uma forma apropriada de armazenamento distribuído e/ou paralelo.

Apresentamos a seguir uma síntese de alguns trabalhos de pesquisa que de alguma forma são importantes para o entendimento da proposta deste projeto e que ainda apontam para desafios abertos na área. Na literatura, encontram-se inúmeros e díspares trabalhos que procuram indicar preocupações no sentido da nossa proposta de pesquisa. Exemplos de trabalhos científicos que podem ser citados são aqueles que discutem visões conjunturais, elementos arquiteturais, aplicações e desafios em aberto. Desta forma, a seguir são comentados alguns trabalhos

mais recentes que podem prover uma base de ideias do horizonte do desafio de armazenamento com grande escalabilidade, peculiar nos ambientes de IoT.

Como a pesquisa apresentada em [8] ressalta, a existência da abordagem IoT traz um novo desafio, no qual os sistemas de informação e comunicação estão invisivelmente embutidos no ambiente que nos rodeia e com isto resulta na geração de enormes quantidades de dados que devem ser armazenados, processados e apresentados de uma forma até então não imaginada, eficiente e facilmente interpretável. Este modelo é composto de serviços que são entregáveis de um modo semelhante para produtos tradicionais, todavia com uma complexidade muito maior. A computação agora integra dispositivos de monitorização, dispositivos de armazenamento, ferramentas de análise, plataformas de visualização e só, então, uma entrega ao cliente.

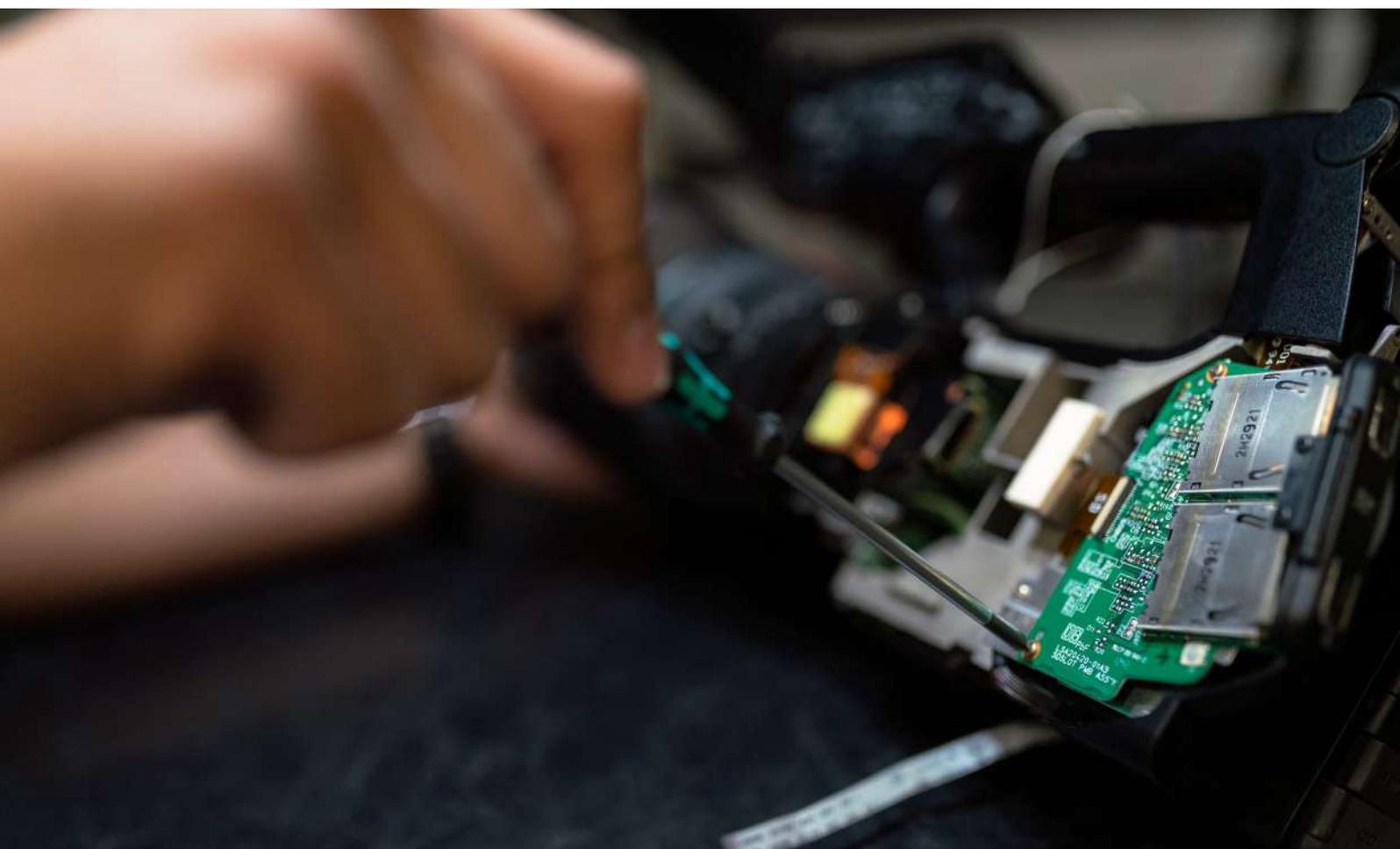
Em [9] é apresentado um estudo sobre *big data* no tocante as aplicações de dados intensivos, desafios, técnicas e tecnologias. Os pesquisadores chamam a atenção que com disposições de dados diversificados, tais como redes de sensores, telescópios, experimentos científicos e instrumentos de alto rendimento, os conjuntos de dados aumentam a uma taxa exponencial, o que leva aos novos desafios que surgem de captura de dados e armazenamento de dados para análise e visualização de dados.

Em [10] é apresentada uma pesquisa orientada a segurança, e a escalabilidade orientada aos ambientes de IoT. Os autores propõem uma configuração que é caracterizada por quatro componentes: clientes, um nodo responsável por disparar solicitações (*dispatcher*), nodos gerenciadores e nodos pares. O cliente é aquele que tem

o arquivo original; o *dispatcher* mantém as referências (endereços IP) dos nodos pares gerenciadores que são responsáveis por grupos; os nodos pares que representam os dispositivos de armazenamento. As operações no sistema de armazenamento sugerido podem ser categorizadas em três níveis: armazenamento e *restore* de arquivos, configuração para transmissão e transferência de dados e replicação compartilhada. Apesar de a proposta ser interessante, existe uma limitação por não considerar a alta disponibilidade do nodo *dispatcher* nem uma rede de interconexão que seja escalável para os nodos pares.

A pesquisa apresentada em [11] ratifica que *big data* e *cloud computing* são desafios de interesse tanto da indústria como da academia. Os autores comentam que as transações comerciais e *loggings* também são bons exemplos de estudos de casos. O conjunto de dados nestes aplicativos são caracterizados por grandes quantidades de dados que são igualmente muito grandes em tamanho e requerem capacidades de processamento em larga escala. Portanto, os requisitos de processamento residem não só na segurança, mas também em termos de eficiência. Assim, a proposta de pesquisa do grupo é a verificação de integridade externa para grandes dados terceirizados em nuvem e **Internet das Coisas**.

Uma proposta de um algoritmo adaptativo para aplicações de monitoramento de grandes massas de dados que se adapta aos intervalos de amostragem e à frequência das alterações de características de dados e às necessidades de um administrador de *cloud* é descrito em [12]. A pesquisa possui uma limitação, pois a abordagem considera a avaliação dos limites somente do método proposto, além do ambiente ser muito específico.



A trabalho apresentado em [13] enumera que as áreas de pesquisa em IoT podem ser sumarizadas como: arquitetura, comunicação abordando a descoberta e processamento de dados, o gerenciamento de dados, segurança e privacidade. A pesquisa ainda indica que uma série de soluções destinadas a resolver esses desafios têm sido propostas, porém eles não são exaustivos e não cobrem todos os vários aspectos.

Distefano, Merlino e Puliafito em seu trabalho [14], propõem um paradigma utilitário *top-down* para a **Internet das Coisas** a partir do IOT-A arquitetura de referência e o Sensing, e a atuação como uma abordagem de serviço (SAaaS). Destina-se a implementação de uma nuvem de detecção e do registro e agregação de recursos de detecção de redes de sensores, e dispositivos móveis pessoais. Em outras palavras, a principal contribuição do trabalho é a concepção e implementação de uma perspectiva de utilidade do paradigma da **Internet das Coisas** através de um *framework* orientado-a-serviços tipo SaaS para a detecção de nuvens.

A relação entre *big data*, *cloud computing*, sistemas de armazenamento de grande volumes de dados e tecnologia Hadoop são discutidos [15]. Esse tipo de discussão caracteriza claramente a preocupação com ambientes distribuídos de larga escala, nos quais configurações de *IoT* são um dos grandes componentes. Além disso, esse trabalho trata dos desafios de pesquisa com foco em escalabilidade, disponibilidade, integridade dos dados, transformação de dados, qualidade de dados, heterogeneidade de dados, privacidade legal e questões regulatórias e de governança.

O trabalho de [16] ressalta que uma dos principais desafios dos sistemas paralelos e distribuídos está em análise quanto a grande quantidade de dados das novas aplicações. Repositórios de dados para tais aplicações excedem atualmente *exabytes* e estão aumentando sua magnitude rapidamente em tamanho. Desafios sugeridos são métodos que considerem essas características e auxiliem o desenvolvimento de *software*.

Os autores da pesquisa encontrada em [17] discutem abordagens e ambientes para a realização de análises em nuvens para aplicações de *big data*. A discussão gira em torno de quatro áreas importantes de análise e *big data*, a saber: (i) gerenciamento de dados e apoiar arquiteturas; (ii) desenvolvimento do modelo e de pontuação; (ii) a visualização e interação do usuário; e (iv) os modelos de negócios. Os autores, entre outras conclusões, afirmam que a abordagem *big data* pode ter vários níveis de preocupações em termos de variedade, velocidade, volume, e veracidade. Portanto, é importante entender os requisitos a fim de escolher ferramentas adequadas de *big data*.

O trabalho apresentado em [18] é interessante posto a extensa avaliação que indica que: (i) coexistindo máquinas virtuais em servidores diminui a vazão dos discos; (ii) o desempenho em *clusters* físicos é significativamente melhor do que em *clusters* virtuais; (iii) a degradação do desempenho, devido à separação dos serviços depende

dos dados para a taxa de computação a ser realizada; (iv) o progresso de conclusão de aplicação correlaciona-se com o consumo de energia e o poder de consumo é fortemente orientado ao tipo de aplicação. Por fim, é apresentado um conjunto de discussões sobre as implicações do uso de ambientes de nuvem para análise de grande quantidade de dados.

Com uma contribuição denominada como *A formal model and analysis of an IoT protocol* [19], Aziz apresenta um modelo formal do MQ Telemetry Transport versão 3.1 do protocolo baseado em uma álgebra de processos de transmissão de mensagens cronometradas. São explicadas as escolhas de modelagem que foram realizadas, inclusive apontando ambiguidades na especificação do protocolo original, e realizou-se uma análise estática do modelo de protocolo formal, que é baseado em uma aproximação de uma semântica nome de substituição para álgebra. A análise revela que o protocolo funciona corretamente como especificado contra os dois primeiros modos de qualidade de serviço de operação proporcionando, pelo menos uma vez, a semântica de entrega para os assinantes. No entanto, foi descoberto que a terceira e mais alta qualidade de semântica de serviços é propensa a erros de maneira ambígua em certos aspectos da sua especificação. Finalmente, sugere-se um aperfeiçoamento deste nível de QoS para o protocolo.

Shaev no trabalho de pesquisa com o título "*From the Sociology of Things to the "Internet of Things"*" [20] traz uma contribuição interessante, posto que os desenvolvimentos tecnológicos devem ser postos em prática com o uso de uma semântica, então precisamos entender como o ser humano entende o IoT. O autor comenta que em uma situação de rápido desenvolvimento das tecnologias da informação sobre a existência de um tal fenômeno cunhado como a "Internet das Coisas". Tal abordagem que captura não só a conjugação de indivíduos humanos na rede de informação, mas também sistemas e dispositivos de *hardware* que muitas vezes pode atuar sem intervenção humana. Tal fenômeno requer um repensar filosófico, e sugere um representar, o de se concentrar em aspectos perceptivos e corporais da vida no mundo moderno.

Finalizando, da revisão bibliográfica realizada fica claro o vasto espectro de caminhos de pesquisas que estão sendo trilhados sob os desafios de *big data* e IoT. Aspectos esses: orientados a propostas de ambientes especiais; estudos ainda iniciais de como devem ser tratados os grandes volumes de dados; propostas de estudos iniciais sobre segurança e escalabilidade em *IoT*; algoritmos específicos para problemas pontuais; propostas de arquiteturas e protocolos de comunicação; sensoriamento de dispositivos e possibilidade de enlace com configurações de *clouds* específicos; estudos sobre *big data*, *cloud computing*, sistemas de armazenamento de grande volumes de dados e tecnologia de Map-Reduce (exemplo: Hadoop); desafios identificados para o desenvolvimento de *software* em uma era *exabyte/yottabyte*; ambientes corporativos e a melhor forma de trabalhar em configurações de *big data*; proposta de modelos formais de protocolos para comunicação de dispositivos IoT; o ser humano e as coisas, e como esse relacionado deve ser pensado.

4. METODOLOGIA

A metodologia prevista para esse projeto será o da prototipação, através de um esforço de desenvolvimento de um conjunto de pacotes de softwares que possam englobar as abordagens propostas no modelo denominado de IoT-APP. Posteriormente, disponibilizando ao final um aplicativo. Para que o esforço de pesquisa atinja o sucesso em um período de um ano, deverá ser considerada a experiência bem-sucedida dos nossos trabalhos prévios de pesquisas realizadas e apresentadas nas referências de [21] a [29], aliados a pesquisa apresentada em [30,31,32]. Cabe ressaltar que as pesquisas mencionadas têm um escopo diferente do ora proposto. Desta forma, não se constituindo assim de um autoplágio. Esses trabalhos serviram de base de conhecimento para a presente proposta.

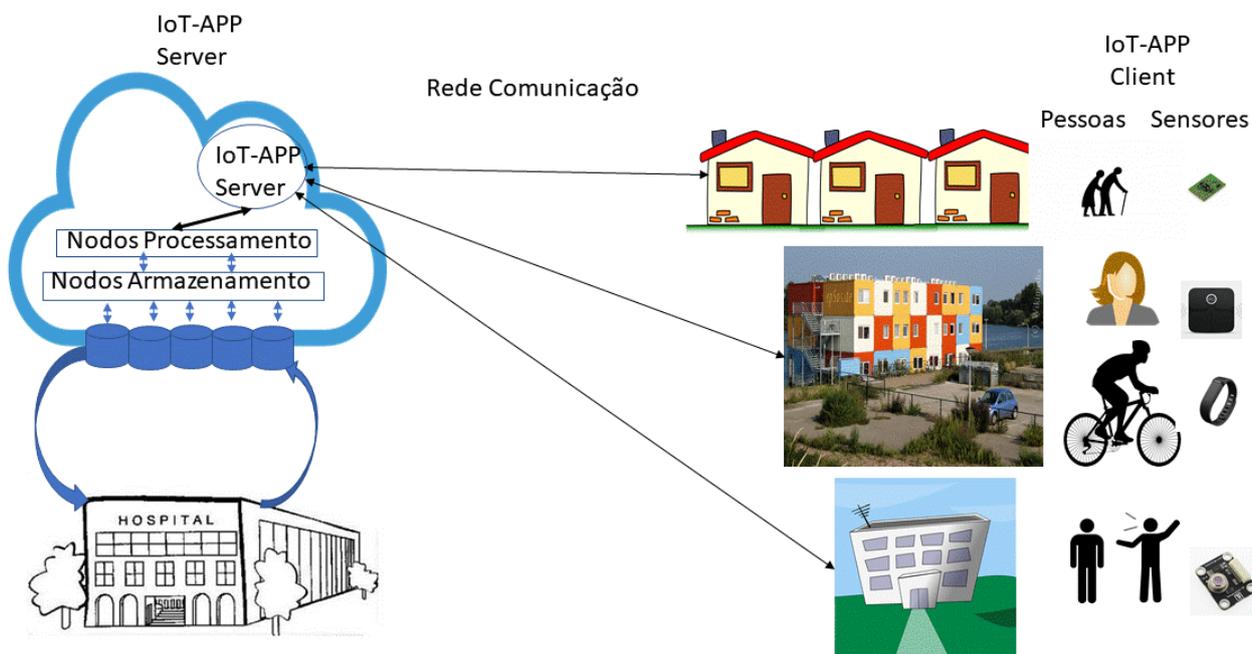
O cenário proposto, conforme ilustrado na Figura 4, é composto dos seguintes componentes:

- a. IoT-APP Client: componente de software que fará a interface entre os **sensores** que deverão coletar dados de sinais vitais de diferentes tipos de dispositivos **IoT** das pessoas, tais como idosos, atletas e pessoas jovens. Esse módulo, também, atuará como roteador para a conexão com o IoT-APP Server;

- a. Rede de comunicação: esta poderá ser caracterizada por uma Wi-Fi ou GSM, visando a transmissão dos dados das pessoas para o módulo IoT-Server;
- a. IoT-APP Server: este componente tem três funções básicas: recebimento das informações do módulo IoT-APP Client, envio dessas informações para o sistema de armazenamento paralelo/distribuído e provimento de uma interface com a unidade de saúde.

Na unidade de saúde (hospital ou UPAs) os grandes volumes de dados (*big data*) existentes no IoT-Server deverão ser **automatizados** em termos de transferência, segundo solicitação de agentes públicos de saúde para **análise** desses **grandes volumes de dados IoT**. Espera-se instalar um módulo de software (ou interface para um prontuário médico eletrônico) para o agente público em unidade pública de saúde, onde o agente público poderá trabalhar com uma grande quantidade de dados de uma grande quantidade de pessoas sem a necessidade de sair de seu ambiente de trabalho. O ganho em termos de qualidade de informação e operacional, no sentido da não necessidade de deslocamento de profissionais para muitas e dispersas áreas geográficas (e muitas vezes perigosas), e melhor diagnóstico, são fator diferenciais que a proposta possa impactar no serviço público de saúde básica.

Figura 4 - Cenário lógico da proposta IoT-APP



Fonte: Autor

Com relação ao ambiente experimental, num primeiro momento será efetuado um esforço de obtenção de sensores e dispositivos IoT de baixo custo, obtidos do resultado da seção 3 dos objetivos gerais. No passo seguinte, a agregação dos pacotes de software e experimentos empíricos realizados quanto ao armazenamento paralelo

e/ou distribuído. Na concepção para o desenvolvimento do ambiente proposto imagina-se que esse possa ser acessado através de uma interface de uma aplicação móvel.

5. AMBIENTES E RESULTADOS EXPERIMENTAIS

Muito importante relatar que o esforço inicial dessa pesquisa foi caracterizado por efetuar diversos contatos com fornecedores de produtos que pudessem apoiar a pesquisa do projeto com dispositivos fabricados ou comercializados no Brasil. A ideia seria que pudéssemos empregar os estudos de casos de possíveis provedores dos dispositivos no IoT-APP. Um exemplo bem interessante de tal dispositivo é o ObaBox [37]. Essa empresa brasileira proveu uma solução chamada de telefone para 3ª idade que se adequa muito bem aos propósitos da prova de conceito desse projeto de pesquisa. Adicionalmente, foram contatadas empresas do segmento de saúde e balanças, visando, de forma semelhante, uma aproximação como proposta de parceria de pesquisa.

Acredita-se que não só pessoas da 3ª idade, mas, também, com baixa escolaridade podem utilizar-se dos dispositivos facilmente, devido a sugestiva interface de contato com os aplicativos. Por outro lado, os fabricantes de dispositivos nas áreas de saúde e balanças, também, poderiam ter interesses na abordagem, posto a capacidade de captação e armazenamento temporários de dados de, por exemplo, aparelhos de pressão e balanças. Estes podem ser elementos muito ricos no processo de coleta e armazenamento *ad hoc* dos dados.

Todavia, apesar de os inúmeros e-mails não terem obtido nenhum retorno para a sugestão de apresentação da proposta de parceria, as empresas foram contatadas (e usualmente) passaram contatos de seus setores de P&D, que nunca nos responderam.

Desta forma, partiu-se para basear a pesquisa em cima dos dois dispositivos móveis adquiridos e *smart band* e telefone celulares de pesquisadores e discentes que colaboram com trabalhos desenvolvidos para a proposta do IoT APP.

A seguir, são apresentados os estudos de casos que são complementares como um todo para a pesquisa IoT APP e que não necessariamente tenham ocorrido em sequência. Em outras palavras, excetuando-se o primeiro estudo de caso os demais ocorreram em paralelo visando atingirmos os tempos limite da pesquisa.

As seções a seguir apresentam as características do desenvolvimento dos conceitos e testes dos componentes IoT-APP Client, IoT-APP Server e derivações que ocorreram.

5.1 Estudo de Caso Fitbit

O primeiro estudo de caso experimental foi a utilização de um dispositivo *smart band* popular, o Fitbit Alta [38]. Ilustrado na Figura 5 com o objetivo de captar alguns sinais vitais e entender os procedimentos energéticos relativos ao seu carregamento, os de visualização e a possível captura aberta dos dados monitorados. Justifica-se o uso desse *smart band* observando-se, por exemplo, a Figura 4, posto que o mesmo é um dispositivo IoT bastante utilizado para captura de dados nos exercícios físicos e para os esportes. O dispositivo utilizado de um pesquisador do NEnC, do Departamento de Ciência da Computação (DCC), da UFJF, foi considerado como o *wearable* padrão para as comparações subseqüentes nesse documento. Relevante observar que a escolha do mesmo teve como justificativa ser um exemplo disponível de mercado, similar a outros tais como os da Apple [39] e da Samsung [40]. Adicionalmente, é interessante observar a semelhança de abordagem dos dispositivos quanto disponibilização dos dados capturados.

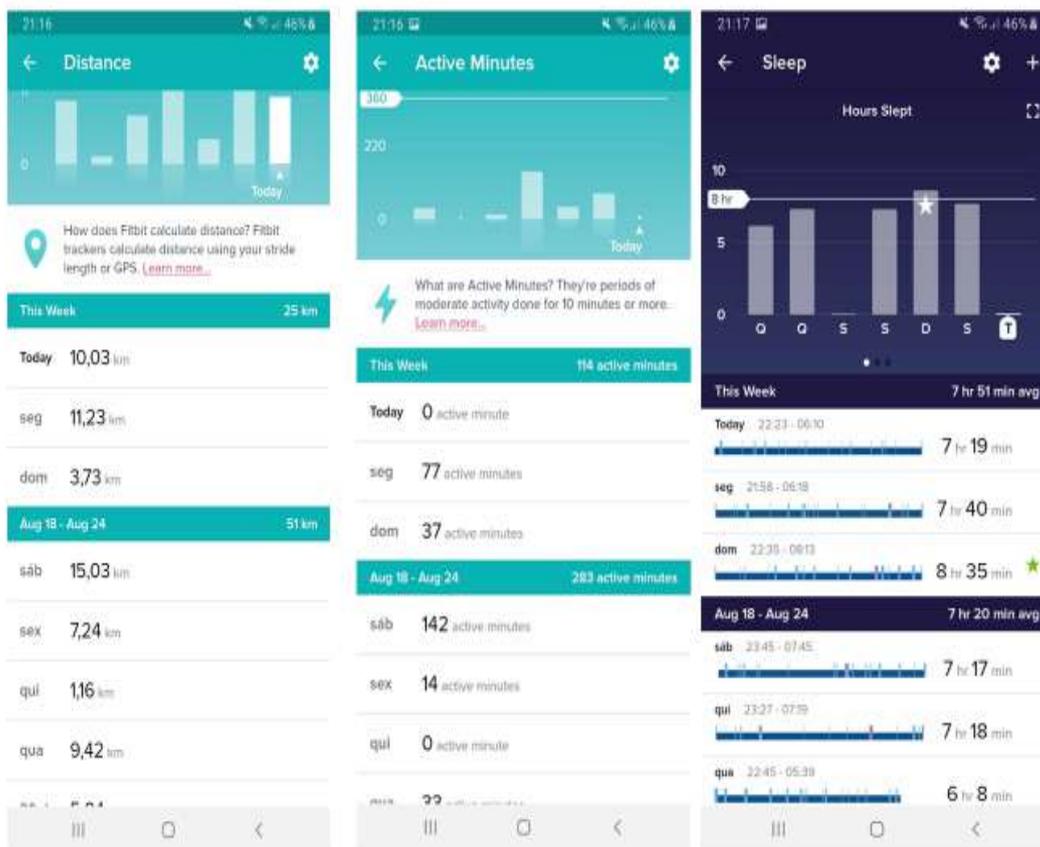
Figura 5 - Fitbit Alta Smart Band



A Figura 6 apresenta três gráficos que indicam, respectivamente, a distância, minutos ativos e desempenho no sono em uma monitoração do smart

band Fitbit. O objetivo dessas ilustrações é indicar como são apresentadas as informações.

Figura 6 - Monitoração Smart Band Fitbit



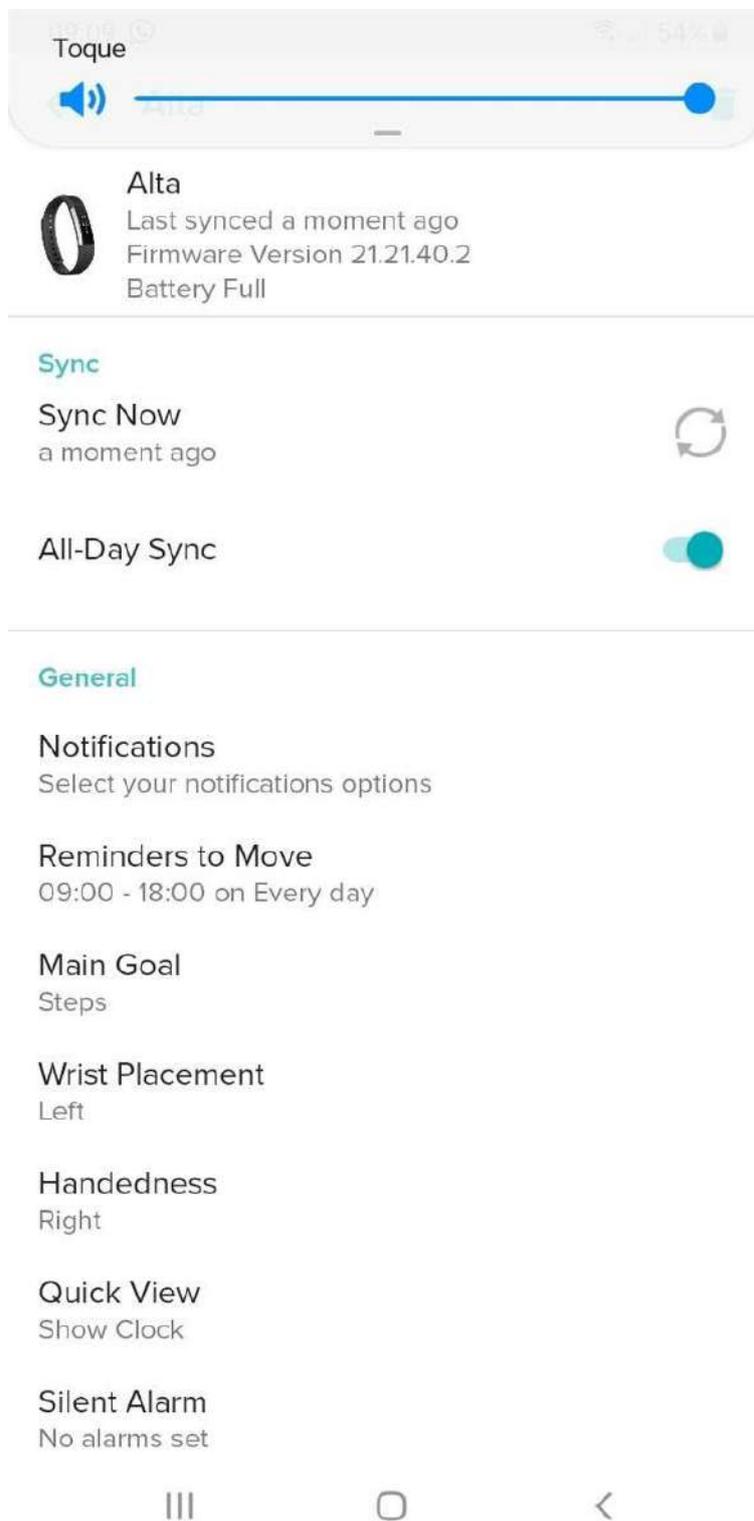
Fonte: Autor



A forma de sincronização para a captação dos sinais vitais nesse dispositivo é apresentada na Figura 7. Esse aspecto, considerando-se um projeto para atendimento de serviços públicos, apresenta uma grande restrição. Ocorre que o sincronismo é efetuado depois do acesso a um servidor central da Fitbit, que em outras palavras indica que os

dados capturados e ilustrados na Figura 6 estão sob a guarda da empresa Fitbit, ou seja, os dados, apesar de serem dos usuários, estão **fechados**. De outra forma, os dados não são interoperáveis (abertos para o usuário) visando que esses possam ser utilizados para fins de armazenamento e data *analytics* de uma política pública.

Figura 7 - Processo de sincronização do Smart Band Fitbit



Fonte: Autor

5.2 Estudo de Caso Kaihai

Em um segundo momento, visto a impossibilidade de captura dos dados monitorados de uma forma aberta, efetuou-se uma pesquisa de mercado visando a aquisição de dois *smart bands* que dispusessem da facilidade de provimento dos dados de maneira aberta e de baixo custo. Esses dois parâmetros são pilares para o sucesso da proposta lógica do projeto IoT-APP, posto que o uso será um serviço público com percepções de cunho de dados abertos e modicidade.

A aquisição dos dois *smart bands*, conforme mencionado, foi baseada no cenário lógico da proposta IoT-APP, como ilustrado na Figura 4. Um dos dispositivos móveis iria ser utilizado por um pesquisador no lado do usuário, no

interfaceando do módulo do IoT-APP Client e, também, junto a um projeto de prontuário médico eletrônico (IoT-APP Server).

Após uma pesquisa documental, via web, com os principais requisitos de provimento à captura de dados abertos e modicidade, foi encontrado o fabricante Kaihai. Após aquisição efetuada, recebemos os dispositivos depois de 60 dias. A Figura 8 serve como ilustração para prover uma ideia dos *smart bands*, cujo custo total dos dois equipamentos foi de R\$ 251,48.

O estudo de caso foi batizado de *Kihai*, uma vez que esse foi o fabricante encontrado para provimento dos dispositivos IoT móveis.

(a) Figura 8 - Dois dispositivos Kaihai IP68



Figura 8 - (b) Kaihai Smart Band



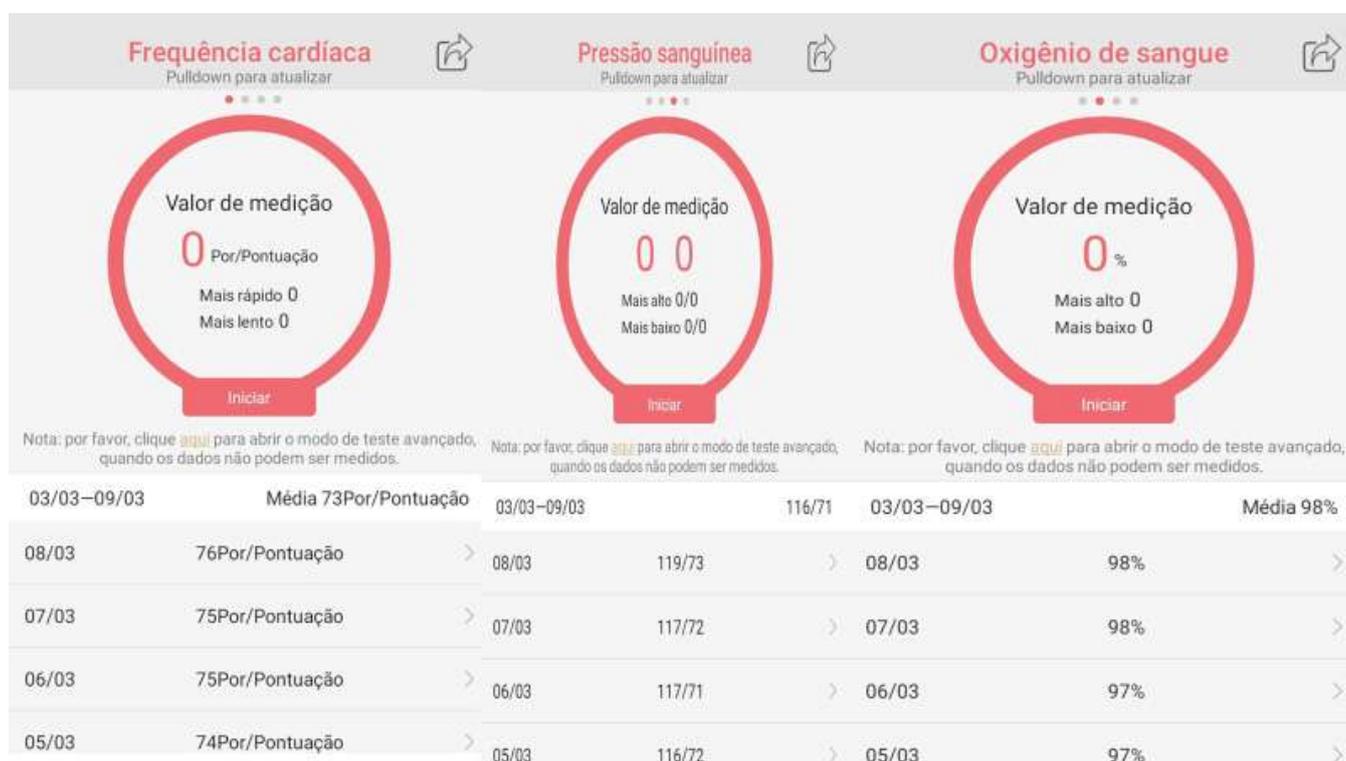
No estudo de caso *Kihai*, o modo de sincronização é semelhante ao *wearable* da FitBit, ou seja, o aplicativo Fitcloud, encontrado na Apple Store e Google Play, solicita o cadastramento do usuário. Efetuado esse cadastramento, o usuário poderá ter acesso a um modo de configuração que fará o sincronismo entre o aparelho móvel e o dispositivo IoT *smart band*. Desta forma, neste primeiro requisito não existe uma diferença entre os dois ambientes.

Prosseguindo a utilização dos equipamentos, conforme mencionado anteriormente, um dos *smart band* foi empregado para os testes, como um usuário final fosse o proprietário do dispositivo. Então, a ideia foi efetuar uma análise experimental como se uma pessoa sem muita instrução fosse o usuário do dispositivo.

Quanto à forma de carregamento do dispositivo, mais uma vez foi verificada uma similaridade com o dispositivo FitBit do experimento 1. A vida média da bateria se mostrou adequada para um uso intenso de operações para uns 4 dias, incluindo-se indicações de recebimentos de e-mails e chamadas ao telefone celular. Por outro lado, quando essas últimas duas funções foram desconsideradas a bateria apresentou uma melhoria.

Com relação às facilidades de monitoração do *smart band Kihai*, foram consideradas a frequência cardíaca, pressão sanguínea, oxigenação, sono e exercícios. Na Figura 9 é possível visualizar três saídas relativas aos três parâmetros. Essas telas indicam uma forma fácil e amigável para leitura e entendimento dos dados captados.

Figura 9 - Capturas frequência cardíaca, pressão e oxigenação



Fonte: Autor

Com relação à Figura 10, esta apresenta saídas relativas aos parâmetros de sono e exercícios. Até o momento, as facilidades do *smart band* empregado nos estudos

experimentais tinham características similares ao estudo de caso da seção 6.1.

Figura 10 - Capturas do sono e exercícios

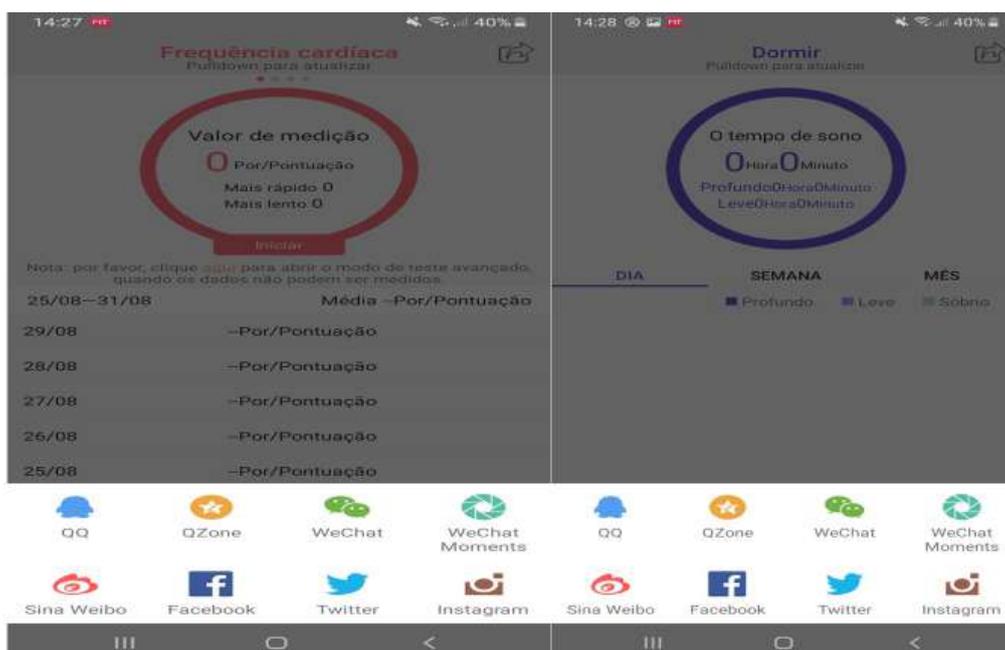


Fonte: Autor

Um diferencial interessante e desejável para o desenvolvimento de um projeto público do dispositivo é ilustrado na Figura 11. Nesta figura é possível constatar que os parâmetros monitorados, no caso da figura como exemplo temos a frequência cardíaca e sono que são

passíveis de serem redirecionados para uma série de aplicativos. Uma outra dimensão de visão dessa facilidade é o entendimento que os dados capturados não precisam ser armazenados em uma plataforma fechada da empresa fornecedora do dispositivo.

Figura 11 - Capturas com portabilidade



Fonte: Autor

Esse diferencial que deve ser compulsório em qualquer disposto provedor de informação dos dados para uma aplicação de terceiros abertos é um aspecto muito importante, principalmente para uma aplicação de serviço público com foco em dados abertos. Acrescendo-se a essa facilidade o baixo custo do dispositivo, neste caso R\$ 125,00.

Após terminada a fase de conhecimento e testes do dispositivo, iniciou-se uma fase de investigação e experimentos com relação às características de facilidades de “abertura” do *hardware*. De outra forma, foi iniciado um desenvolvimento de sincronização do equipamento,

visando que os dados pudessem ser passados para um aplicativo “aberto”, como a proposta IoT-APP Client. A finalidade de tal esforço é um *re-design*, ou utilização de uma aplicação nossa, no caso IoT-APP Client, onde deseje-se que nossas interfaces amigáveis e mais diretas sejam capazes de consumir os dados monitorados. Com esse objetivo temos a Figura 12 apresentando uma visão dos componentes do IoT-APP Client.

Na Figura 12 ilustra-se uma proposta, desenvolvida para esse projeto, onde sinais vitais básicos são considerados. É possível incrementar com novas facilidade existentes nas opções de capturas dos *smart bands*.

Figura 12 - Tela inicial do IoT-APP Client



Fonte: Autor

A Figura 13 apresenta uma visualização da tela de sincronização do IoT-APP Client. Acredita-se que essa

forma de conexão é mais fácil para usuários finais com pouca instrução ou mais idosos.



Figura 13 - Tela sincronização do IoT-APP Client



Fonte: Autor

Finalizando, nesta seção, apresentamos as características do *smart band Kihai*, que foram testadas e depois utilizadas como elementos primários da arquitetura IoT-APP Client. Sugeriu-se, também, conforme apresentado nas Figuras 12 e 13, que uma customização para uma aplicação pública é possível.

5.3 Estudo de Caso Prontuário Eletrônico

Nesta seção apresenta-se o esforço levando-se em conta o lado onde o prontuário eletrônico é a interface com a entidade final de armazenamento dos dados públicos. Relembrando, através da Figura 4, essa fase tem como característica um interfaceamento com os dados que deverão ser destinados ao sistema de armazenamento final, ou seja, parte do IoT-Server.

Também é importante mencionar que o segundo dispositivo *Kaihai* adquirido foi utilizado nesse esforço, no sentido de entender o sincronismo do lado cliente com o lado servidor.

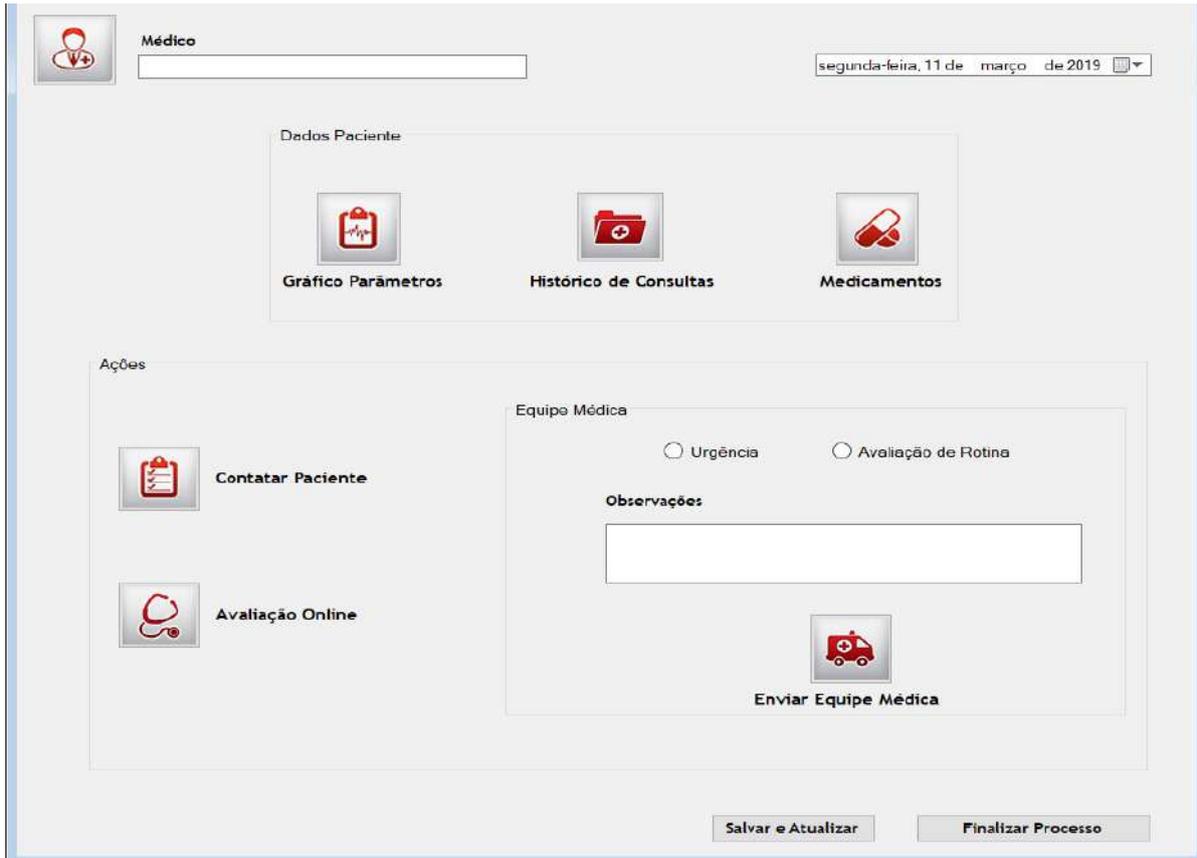
Esta fase do projeto de pesquisa foi desenvolvida em parceria com o professor Fabricio Mendonça, do Departamento de Ciência da Computação, da UFJF, e do médico Marcelo Quesado (hospital Monte Sinai e hospital regional João Penido, Juiz de Fora, MG). Foi enriquecedora essa abordagem, pois foi possível desenvolver o interfaceamento dos dados coletados com prontuário eletrônico que é utilizado em uma unidade pública de saúde. A ideia da inserção desses dados, conforme temos advogado, é que os mesmos possam servir como uma

boa base de monitoração de indivíduos, inclusive por agentes de saúde básica, que devem fazer centenas de medidas em pessoas. Em outras palavras, com a inserção automatizada de sinais vitais abertos através do IoT APP os profissionais de saúde das unidades básicas podem ter uma maior abrangência nos cuidados, uma vez que o serviço de medição com dados abertos permite-lhes dar uma atenção aos casos onde se verificam cuidados diferenciados. Outro aspecto importante é que os dados serão armazenados em diretórios individualizados, onde ferramentas de *data analytics* poderão prover predições diferenciadas de patologias.

Como forma de ilustrar melhor tais funcionalidades no sistema em desenvolvimento são apresentadas a seguir telas do sistema que ilustram essa integração da plataforma de IoT com o sistema de prontuário eletrônico.

A Figura 14 apresenta a interface do médico que ocupa o cargo de coordenador da central de monitoramento de dados do prontuário eletrônico. Tal central é composta por esse médico coordenador e uma equipe de enfermeiros (agentes de saúde) que monitoram os dados dos pacientes coletados dos *smarts bands*. Por meio dessa interface é possível consultar os dados dos pacientes, tais como: os parâmetros coletados, histórico de consultas e medicamentos administrados. Através desta interface ocorrem também as ações de intervenção geradas após a análise dos dados coletados, que podem ser desde uma avaliação on-line com o paciente até o envio de uma equipe médica à residência do paciente.

Figura 14 - Interface do médico coordenador da central de monitoramento

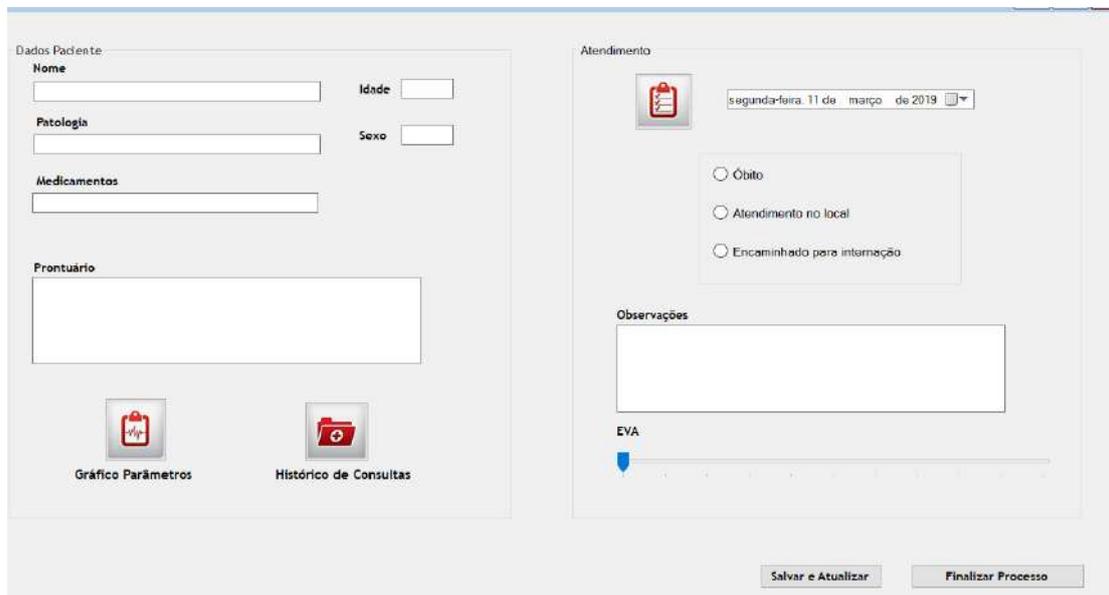


Fonte: Autor

A Figura 15 ilustra a interface do agente de saúde responsável pelo atendimento do paciente em seu domicílio, depois que uma ação de intervenção médica foi disparada. Tal interface é executada em um aplicativo de um dispositivo móvel para que o agente possa registrar as

informações do atendimento (medicamentos, prontuário, observações e EVA – escala de dor do paciente) e, também, finalizar o atendimento registrando um dos 3 (três) estados possíveis: óbito, atendimento no local ou encaminhamento para internação para internação.

Figura 15 - Interface do agente de saúde responsável pelo atendimento



Fonte: Autor

A Figura 16 mostra um exemplo de uma análise histórica e estatística dos dados coletados e monitorados do paciente atendido. Entre os parâmetros analisados têm-se: temperatura corporal, frequência cardíaca (apenas esses

2 são mostrados na Figura 16), ritmo cardíaco, dosagem de glicose percutânea, pressão sanguínea, oxigenação, frequência respiratória, controle do sono e dos exercícios praticados pelo paciente.

Figura 16: - Interface da análise histórica e estatística do paciente



Fonte: Autor

Concluindo, nesta seção apresentamos os esforços de pesquisa no módulo IoT-APP Server, o qual apresenta a prototipagem das telas necessárias do prontuário eletrônico e que pode ser alimentado em grande parte por dados captados de um *smart band*, como aquele apresentado na seção anterior. É relevante sinalizar que essa interface entre nosso grupo de pesquisa, com a participação do professor Fabricio Mendonça e do médico Marcelo Quesado (hospital Monte Sinai e hospital regional João Penido, Juiz de Fora, MG), se intensificou a partir dessa pesquisa. Existe atualmente um esforço de se conseguir recursos financeiros para que possamos desenvolver todas as interfaces prototipadas.

5.4 Estudos de Casos de Transformação Digital

Nesta etapa da pesquisa apresentamos dois trabalhos que são derivados do estudo original, que ao tomarem conhecimento do projeto solicitaram ajuda e apoio em uma dimensão de **transformação digital**.

Trabalho 1

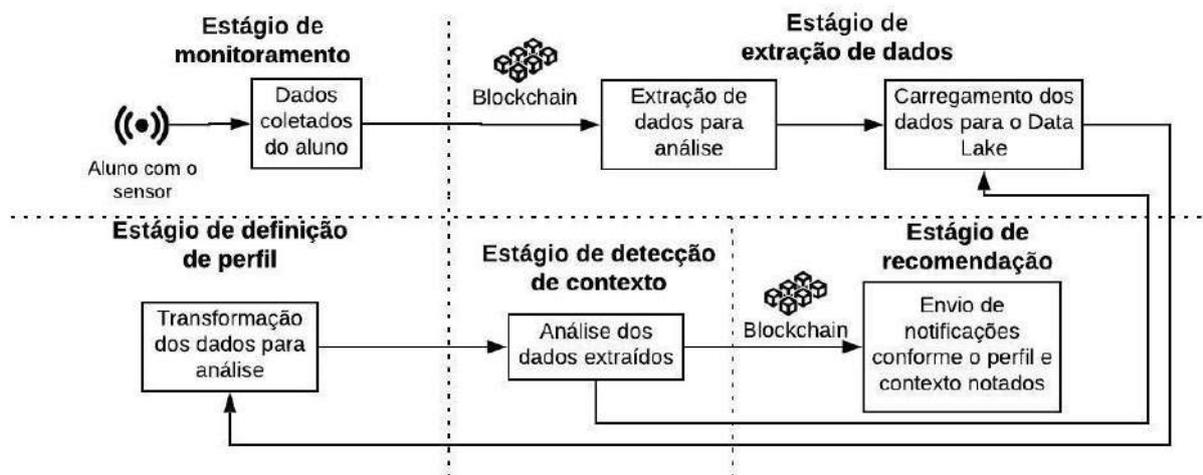
No primeiro estudo de caso foi considerado um contexto do trabalho desenvolvido em um novo ambiente experimental, seguindo as diretrizes da **transformação digital**. Desta forma, foi utilizada a abordagem de IoT para a detecção e controle de ansiedade e nervosismo em

cenários acadêmicos. Durante o processo de formação é reconhecido que os discentes passam por períodos algumas vezes difíceis, onde a angústia e apreensão se elevam, posto os desafios no processo de grande mudança nas suas vidas, usualmente entendendo que é uma fase importante de preparação para as suas carreiras profissionais.

Estresse e ansiedade são fatores que afetam o desempenho do aluno. Desta forma, um modelo denominado como **Hold Up** [41], baseado no conceito de monitoração de pessoas semelhante à proposta de pesquisa foi projetado para a busca de auxílio ao aluno na detecção e controle das emoções, visando ajudar seu desempenho acadêmico. O modelo detecta oscilações na frequência cardíaca por meio de sensores e define o perfil e contexto emocional do aluno. Variações nas frequências cardíacas que podem indicar situações de estresse que, quando não controladas, prejudicam o desempenho do aluno durante a realização das atividades acadêmicas. A proposta foi avaliada com alunos de uma universidade brasileira. Os resultados apontam a viabilidade da proposta e os comentários opinativos forneceram indícios positivos de que esta abordagem pode ser usada em ambientes educacionais.

O modelo é apresentado na Figura 17 e detalhes sobre o desenvolvimento na referência [41], onde agradecemos o apoio da Enap, posto a Cátedra da chamada no. 5/2018. Tal referência foi nosso primeiro artigo aceito no Congresso SBIE 2019.

Figura 17 - Visão geral do modelo *Hold Up* [41]

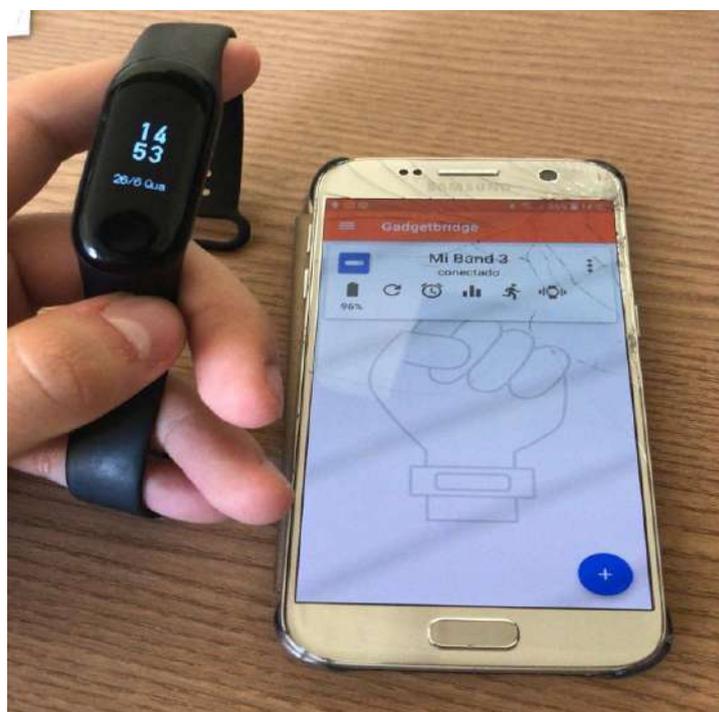


Fonte: [41]

Diante do exposto, conforme apresentado na Figura 18, utilizou-se um dispositivo Mi Band – 3 e um telefone móvel, de propriedade do voluntário do estudo de caso. Importante observar que nossa experiência de projeto foi

um dos pilares base para o desenvolvimento da proposta, além do incentivo aos esforços de sincronização com o dispositivo.

Figura 18 - Interface Sincronização do Dispositivo Mi Band 3

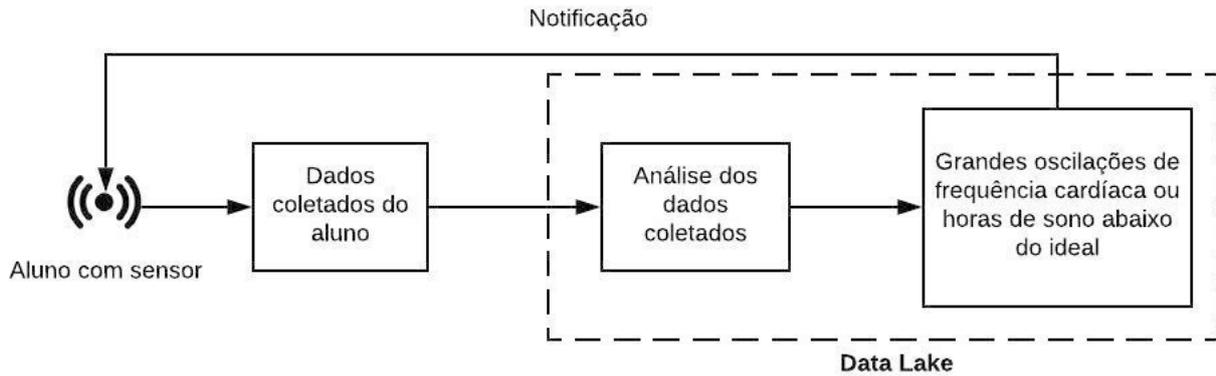


Fonte: Autor

Para a definição do perfil dos alunos foi necessário que cada um deles aferisse sua frequência cardíaca, inicialmente, em uma situação de repouso, para determinarmos qual era sua frequência cardíaca normal quando eles não se encontravam em situações de estresse. Após as medições, calculamos a frequência cardíaca média dos alunos envolvidos neste estudo para que, assim, seus perfis fossem definidos em situações de repouso. Logo, com todos os

alunos examinados nesse estudo, identificou-se que a frequência cardíaca média em situações de repouso ficava na faixa de 60bpm a 70bpm entre todos os participantes. Após a definição do perfil, os alunos foram monitorados durante a execução de atividades acadêmicas e a etapa de execução deste estudo de viabilidade seguiu as etapas ilustradas na Figura 19.

Figura 19 - Etapas de execução



Fonte: [41]

Os dados coletados dos alunos através das *smartband* (Figura 18) foram enviados para o *MongoDB*, por ser um banco de dados *NoSQL* amplamente utilizado em soluções *Data Lake*. Subsequentemente, esses dados foram analisados e o modelo *Hold Up* traçou o contexto dos alunos para identificar seus estados emocionais e verificar a necessidade de envio de alguma recomendação. Assim, os alunos que, em algum momento, apresentaram oscilações de frequência cardíaca ou que obtiveram uma quantidade de horas de sono abaixo do indicado, receberam uma notificação.

Os alunos foram monitorados durante as atividades acadêmicas. A Tabela 1 apresenta as frequências cardíacas de alguns dos alunos voluntários em três momentos da avaliação. Percebe-se que, em geral, os alunos apresentam um aumento na frequência cardíaca durante a atividade acadêmica, mas nem todos apresentaram oscilações, como o caso do aluno 2. Isto é um indicativo de que alguns alunos permanecem calmos, mesmo diante de situações de possível ansiedade como provas, reuniões ou apresentações. Nestes contextos, o modelo não enviou notificações para o aluno.

Tabela I - Ambiente experimental com seis alunos

Alunos	Frequências			Sono		
	1º	2º	3º	Leve	Profundo	Total
Aluno 1	83 bpm	103 bpm	78 bpm	1 h e 59 min	4 h e 17 min	6 h e 16 min
Aluno 2	64 bpm	70 bpm	68 bpm	3 h e 13 min	4 h e 43 min	7 h e 56 min
Aluno 3	115 bpm	93 bpm	53 bpm	6 h e 49 min	2 h e 30 min	8 h e 19 min
Aluno 4	61 bpm	88 bpm	69 bpm	5 h e 24 min	1 h e 17 min	6 h e 41 min
Aluno 5	68 bpm	98 bpm	72 bpm	1 h e 36 min	4 h e 51 min	6 h e 27 min
Aluno 6	70 bpm	101 bpm	65 bpm	4 h e 19 min	4 h e 20 min	8 h e 39 min

Fonte: [41]

Neste primeiro esforço de implementação, definiu-se que era pertinente enviar dois tipos de notificação aos usuários. Uma relacionada a atividades imediatas e outra a atividades preventivas, visto que ambas fazem com que seu nível de estresse abaixo, cabendo assim ao usuário escolher qual delas melhor se adéqua ao ambiente em que o mesmo se encontra. Também foi interessante informar o montante de horas de sono que o mesmo obteve na determinada noite e informá-lo caso o mesmo não tenha obtido quantidade de horas necessárias.

Neste estudo, cerca de 20% dos alunos monitorados não receberam notificações, visto que o modelo não detectou que os mesmos se encontravam em situação de estresse. Por outro lado, aproximadamente 80% dos alunos receberam algum tipo de recomendação, seja para o controle emocional ou para a adequação das horas de sono.

Para verificar a aceitação da proposta por parte dos alunos e a adequação das recomendações foi enviado

um formulário para resposta de todos os indivíduos que participaram na coleta de dados. No formulário constavam os dois grupos de recomendação a serem enviados: atividades imediatas e atividades preventivas. E, também, um breve comentário sobre em que consiste cada uma das atividades para redução de estresse. O formulário foi elaborado com questões como: “Você considera que a recomendação de uma atividade de cada grupo pode ajudar a reduzir seu nível de ansiedade e estresse?”; “As recomendações que lhe foram feitas foram adequadas e úteis para você?”. As respostas das questões seguiram a escala Likert variando de 1 (péssimo) a 5 (ótimo). Os resultados são descritos em sequência.

Em síntese, identificamos que 50% dos alunos observados qualificam a proposta do modelo *Hold Up* como sendo “ótima”.

Os 50% restantes avaliaram a proposta como “boa”. Logo, identifica-se que alguns dos alunos examinados avaliam a proposta como sendo “boa” devido às recomendações que não são úteis para os mesmos perante a situação por eles vivenciada, como a realização de atividades físicas durante uma prova. Os alunos que não receberam recomendações não preencheram o formulário de avaliação. Por outro

lado, através de entrevista informal, foram questionados se eles consideravam importante receber alguma notificação. Alguns deles relataram o interesse de receber dicas mais genéricas para o controle de emoções, como ouvir músicas.

Trabalho 2

Um estudo paralelo, efetuado por um usuário leigo com perfil de engenharia, ao tomar conhecimento de nossos esforços concluiu que poderia ser um estudo de caso experimental interessante, posto a necessidade de controle de variações abruptas de frequências cardíacas. Em outras palavras, esse estudo de caso é interessante posto o conhecimento na área de monitoramento e, também, por conhecer o problema de variações abruptas de frequências.

A Figura 20 apresenta o *smart band* Kaihai Eletro. Este dispositivo, diferente do adquirido para nossos experimentos na proposta IoT-APP, além das funções de pressão, oxigenação e batimentos cardíacos possui uma função de eletrocardiograma. Esta última função, com disparos em vários momentos do dia, visa uma indicação de variabilidades da frequência.

Figura 20 - Kaihi Eletro



Após a utilização do dispositivo por um período determinado foi relatado que o dispositivo não se apresentou conforme indicado no seu manual. De outra forma, em muitas situações as medidas ficaram muito diferentes em relação a outros dispositivos que em paralelo faziam captação dos dados.

Concluindo, o usuário não certificou a utilização desse *smart band* como um dispositivo que pudesse ser uma ferramenta de uma **transformação digital**, para aqueles usuários que requerem inúmeros equipamentos diariamente para um controle de sinais vitais.

5.5 Estudo de Caso de Armazenamento

O grande esforço inicial da pesquisa do projeto IoT-APP era relativo ao tópico de armazenamento dos dados captados dos dispositivos. Todavia, ao longo do processo de investigação do cenário foi verificado que aspectos operacionais, já mencionados neste relatório, como os parceiros comerciais e entidades de saúde, mostraram-se muito complexos de adequação em doze meses.

Nossos esforços foram em vão quanto a aproximação com departamentos de P&D, como no exemplo da Obabex [37], de organizações comerciais.

Por outro lado, de uma forma mais positiva, o projeto foi a alavanca para a conexão junto ao Hospital João Penido Filho. Essa conexão começou a dar bons frutos em termos de parcerias de captação de colaboradores e possíveis recursos.

Todavia, nossos esforços não cessaram por conta dessa falta de apoio de interesse dos possíveis *players*. Decidi fazer efetuar pesquisa através de dados sintéticos. Efetuei a execução em uma infraestrutura internacional de grande porte, o GRID5000 [42] na França, por ser pesquisador com acesso a esses recursos de computação de alto de desempenho.

Esta subseção apresenta os resultados obtidos da experimentação, que gerou um artigo [43] submetido ao evento WSCAD 2019 (Workshop Computação de Alto Desempenho), onde os agradecimentos à Enap foram citados.

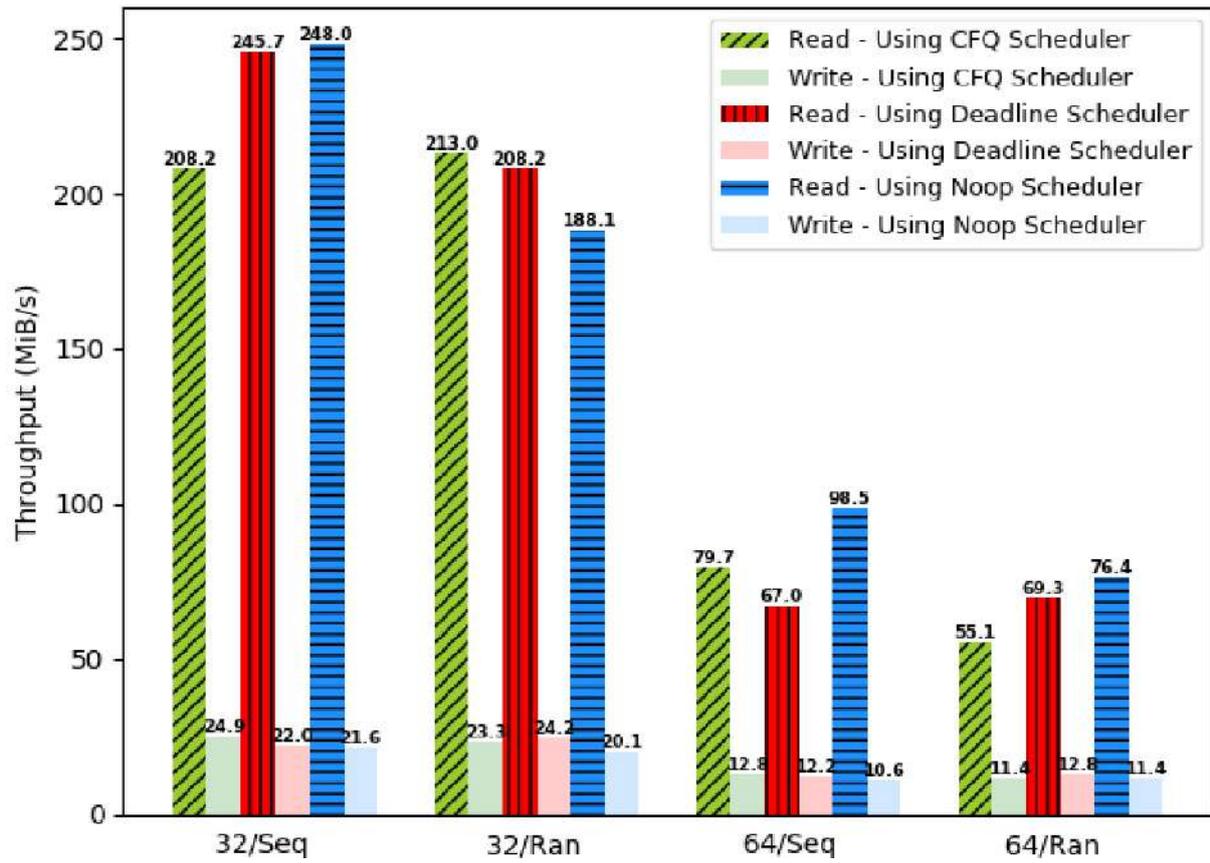
Com os experimentos realizados estamos procurando entender como o rendimento se comporta quando um *benchmark* executa a leitura e operações de gravação de grande quantidade de dados. Desta forma, aproxima-se muito ao cenário desafio do IoT-APP.

Inicialmente, vamos discutir como os dados e informações são apresentados nestes gráficos. Em cada figura apresentada, 24 resultados para as operações de leitura e gravação são ilustrados. Todas as barras com hachuras apresentam o valor da taxa de transferência para a operação de leitura que são relacionados a um escalador específico e a todas as barras sem hachuras, localizadas após a barra da hachura, apresenta o valor da taxa de transferência da operação de gravação para o mesmo escalonador. O valor da taxa de transferência para a operação de gravação sempre vem após a taxa de transferência relacionada à leitura.

Na Figura 21, podemos analisar o valor da taxa de transferência das operações ao armazenar dados e metadados no dispositivo HDD, alternando os escalonadores de E/S. Fazendo uma análise superficial, é possível verificar, em todos os casos que a taxa de transferência que o valor da operação de leitura é maior que a operação de gravação, independentemente do ser usada pelo planejador. Apresentamos no artigo detalhes da média porque acreditamos que os aplicativos DISC e HPC podem tratar e usar dados heterogêneos com diferentes padrões e números de acesso de tarefas no mesmo aplicativo. Entre outras análises, é possível verificar que, quando armazenamos dados e metadados no dispositivo HDD, o escalonador que apresenta o valor mais alto de rendimento médio ao executar a operação de leitura é o *Noop*, e ao executar a operação de gravação é o planejador CFQ.



Figura 21 - Dado e metadado HDD



Fonte: autor

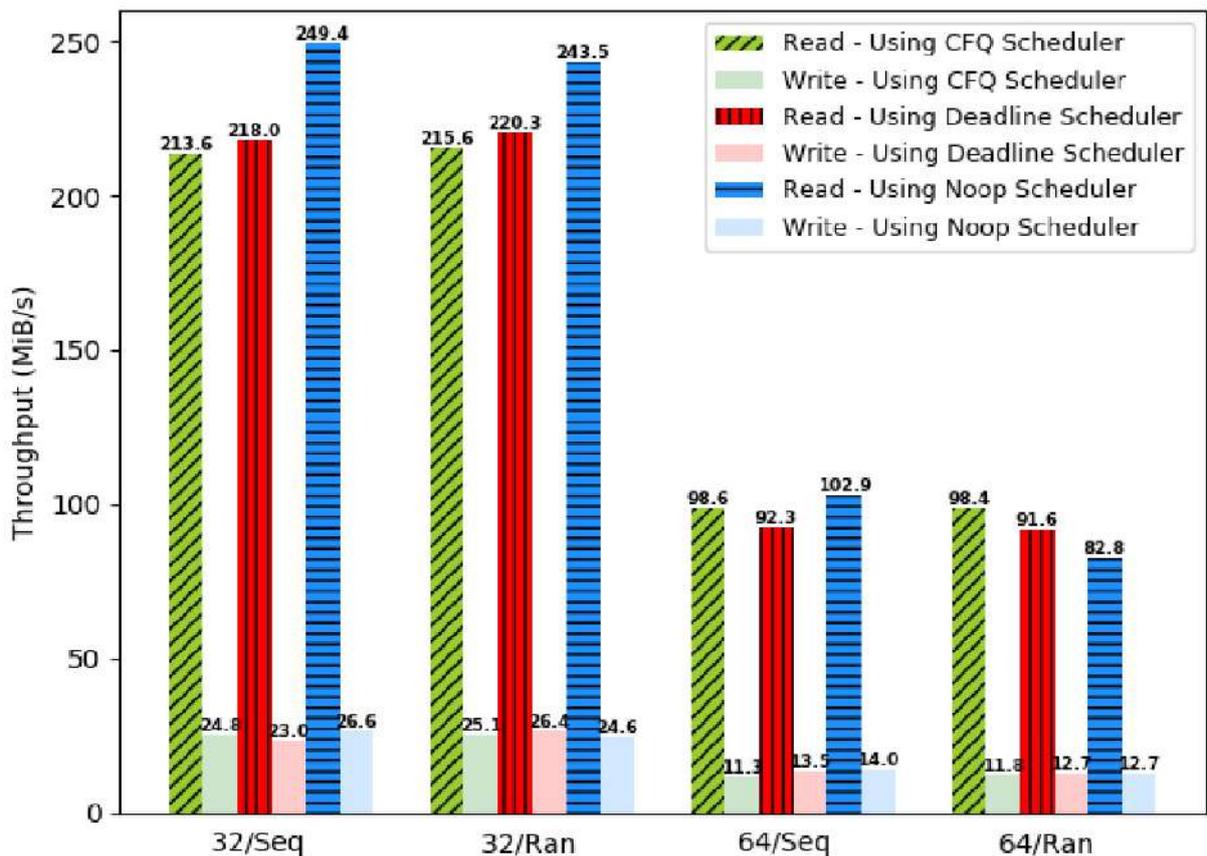


Na Figura 22 podemos analisar o valor da taxa de transferência das operações ao armazenar

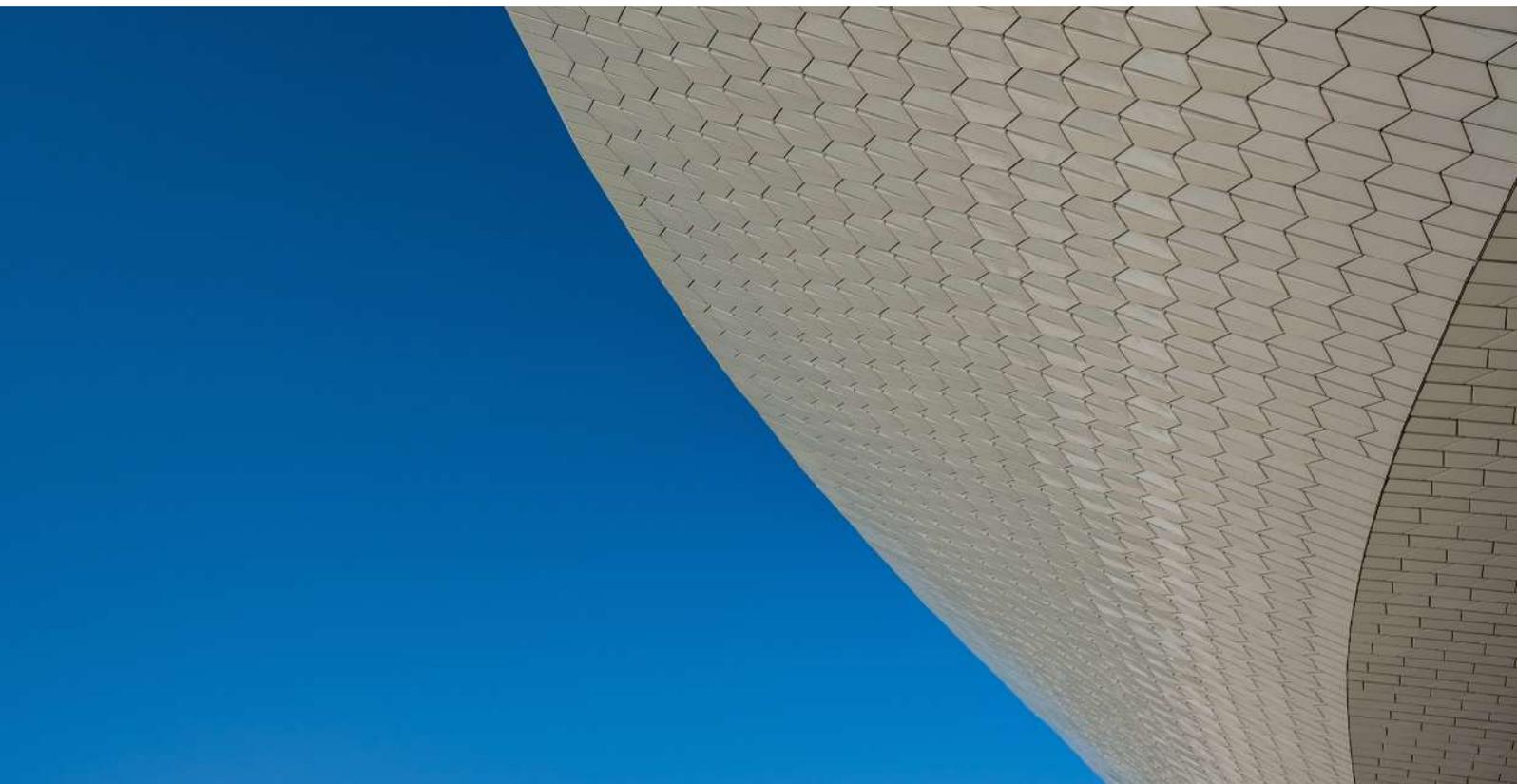
dados no HDD e metadados na comutação do dispositivo SSD e, em seguida, nos escalonadores de E/S. Assim, como aconteceu na Figura 21, na Figura 22 todos os valores ao executar a operação de leitura são maiores que os valores ao executar a operação de gravação. Apresentamos no

artigo, também, uma tabela que tem o valor da taxa de transferência ao executar a operação de leitura para todos os escalonadores maiores do que ao armazenar dados no disco rígido. Nota-se, da mesma forma, que ao armazenar dados no HDD e metadados no SSD, o escalonador que apresenta o maior valor médio de taxa de transferência quando executar operação de leitura e gravação é o escalonador *Noop*.

Figura 22 - Dado no HDD e metadado no SSD



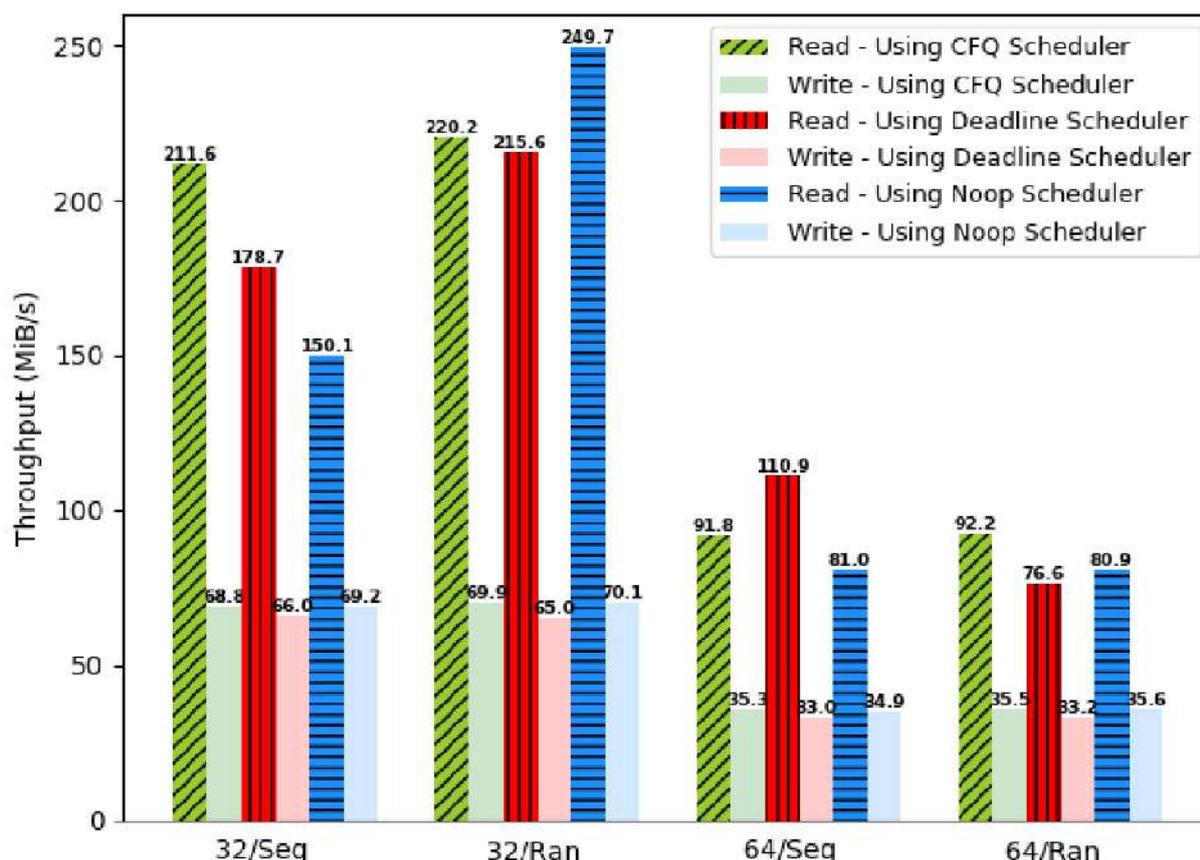
Fonte: autor



Na Figura 23 armazenamos até os dados e metadados no SSD, que é um dispositivo que cria um cenário onde não haja componentes mecânicos alternando, os mesmos escalonadores de E/S. Neste caso, o valor da taxa de transferência ao executar a operação de gravação aumenta significativamente em comparação com as outras duas abordagens apresentadas anteriormente. No entanto, é possível verificar-se, quando executada a operação de leitura, a taxa de transferência não sofreu variações significativas.

Outro resultado interessante pode ser verificado na Figura

Figura 23 - Dado e metadado no SSD



Fonte: autor

6. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

O esforço de pesquisa da proposta do projeto IoT-APP, após o período de doze meses, mostrou-se bem-sucedido de uma forma diferente daquele que imaginado no princípio. Todavia, pode-se considerar que os objetivos foram alcançados e estendidos. Interessante mencionar que desse esforço foram escritos dois artigos científicos, um já aceito e o outro em fase de avaliação nos melhores fóruns de suas áreas no Brasil.

Como resultado dessa proposta de pesquisa esperava-se:

- Descobrir qual a melhor forma de armazenar o grande volume de dados públicos. Em outras palavras, utilizando-se de ferramentas do tipo NoSQL, NEWSQL,

23 se compararmos a taxa de transferência de operação de leitura com as Figura 22 e 21. O resultado apresentado na Figura 21 ao usar um número de tarefas igual a 32 e o padrão de acesso sequencial era menor para prazo e escalonador *Noop* ao armazenar dados e metadados no SSD.

Estes resultados podem nos levar a pensar que se o dispositivo que você está armazenando os dados é um dispositivo SSD, é muito provável que o valor da taxa de transferência da operação de gravação aumentou.

sistemas de arquivos distribuídos ou paralelos, ou um mix dessas tecnologias. Desta forma, provendo uma indicação de facilidades, capacidades e limitações:

- » Este objetivo foi atingido com estudo de caso apresentado nas seções 6.4 e 6.5. Na seção 6.4, onde a captação de dados de voluntários através de *smart bands* foi armazenada em uma base NoSQL que foi utilizada para o processo de recomendação. Na seção 6.5, dados sintéticos foram utilizados em infraestrutura de alto desempenho no Grid5000, na França, e representando um estudo focal na proposta inicial do projeto IoT-APP. Desta forma, o objetivo básico foi alcançado.
- » A captação de um grande volume de dados públicos não foi realizada e descobrimos as limitações político-administrativas para essa execução. Em outras palavras, somente com

esforço maior e com maior patrocínio voluntários se dispõem a participar. Uma possível abordagem seria o provimento de pulseiras sem custo, uma vez que captações de dados fossem permitidas, semelhante ao ocorrido no estudo da seção 6.4.

- Um aplicativo para um serviço público, através do protótipo, que fosse capaz de obter uma grande captação no volume de dados de um grande número de pessoas relacionados aos sinais vitais:
 - » Este desenvolvimento da pesquisa foi atingido através das prototipações apresentadas nas seções 6.2 e 6.3;
 - » Uma dificuldade encontrada e entendida ao longo do projeto foi a existência de diferentes interfaces entre uma aplicação de prontuário eletrônico e aquela que pode ser disponibilizada por um aplicativo cliente. Em outras palavras, esse esforço deve ser abraçado por ambas as partes, o que se iniciou com a proposta desse projeto e foi apresentado para um médico coordenador de um hospital público de médio porte, e mencionado na seção 6.3.
- Que os dados oriundos da abordagem IoT sejam armazenamento de uma forma diferenciada visando uma melhor análise desses grandes volumes de dados por parte dos profissionais de saúde:
 - » Pode-se inferir depois dos estudos de casos apresentados que é possível atingir esse nível, todavia no espaço temporal maior que doze meses. A aproximação já mencionada anteriormente de uma forma mais efetiva poderá dar alcance a essa meta.
 - » Que seja possível convencer os fabricantes de sensores brasileiros que vale a pena uma melhoria e diversificação em seus produtos, visando um melhor e maior uso da captação e disponibilização de dados: um dos pontos de maior preocupação foi o contato com os fabricantes e representantes, foram despendidos inúmeros esforços para essa aproximação, todavia sem nenhuma sinalização de nenhum player.
- Que outros tipos de aplicativos para o serviço público, tais como para a educação e segurança, possam ser desenvolvidos efetuando abordagens semelhantes à proposta desenvolvida.
 - » Os estudos de casos denominados como de **transformação digital** são representativos do sucesso deste tópico. Quanto aos estudantes foi possível verificar o amadurecimento de muitos, quanto ao ponto que o tácito mais tecnologia podem sim representar uma transformação digital do tipo ganha-ganha. Importante mencionar que esse esforço foi certificado com a aprovação de um artigo no SBIE (Simpósio Brasileiro de Informática na Educação) 2019, no qual os agradecimentos a Enap foram incluídos.

- » Por outro lado, o estudo de caso do *smart band* onde as especificações não se confirmaram no dia-a-dia leva-nos a pensar em um tipo de certificação, como era de se esperar para o uso dos dispositivos. Ao optar por um determinado *smart band* foram verificados por nossa parte cerca de vinte equipamentos. Ao contrário, a escolha do leigo foi guiada pela nossa escolha de marca, que não se mostrou eficiente para seu problema.

Como trabalhos futuros, após as experiências dos estudos de casos, acredita-se que um fator a ser observado com cuidado é que o patrocínio institucional é fundamental para a continuação dessa pesquisa. Desta forma, a aproximação com a direção de um hospital público de médio porte irá tornar mais eficaz esse trabalho, devido a todos os desafios apresentados e a prototipagem ilustrada. Por outro lado, é necessário que exista pessoal para que mais esforços possam ser traduzidos em códigos e experimentos. Assim, estamos participando de chamadas que possam prover esse subsídio. Finalizando, uma aproximação com órgãos federais de fomento que possam conhecer os esforços dessa pesquisa ajudaria na dimensão financeira e na aplicação dos conhecimentos adquiridos no projeto que fossem largamente utilizados em desafios de órgãos de saúde.

Finalizando, esse relatório final de pesquisa indica o sucesso a prova de conceito da proposta IoT-APP (*IoT – Armazenamento & Processamento na área Pública*), apesar das dificuldades iniciais. Acredita-se, após os doze meses de pesquisa, que o potencial é muito grande no retorno em diversos níveis para a área de saúde e de educação pública. Os elementos das partes do projeto se mostraram muito interessantes e os dados podem prover um rico *data analytics*, como ilustrado no estudo com os estudantes. Os demais desafios tecnológicos são possíveis de serem atingidos. Todavia, um especial cuidado deve ser observado com empresas e suas interfaces de P&D no entendimento do esforço proposto.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] *Bringing Big data to the Enterprise*, IBM, <http://www-01.ibm.com/software/in/data/bigdata>.
- [2] PEPPER, R.; GARRITY, J. “*The Internet of Everything: how the Network Unleashes the Benefits of Big Data*,” *Global Information Technology Report*, Cisco, pp. 35-42, 2014, <http://blogs.cisco.com/wp-content/uploads/GITR-2014-Cisco-Chapter.pdf>.
- [3] WANG, Lizhe; Ranjan, Rajiv. *Processing Distributed Internet of Things Data in Clouds* IEEE Cloud Computing, January-February..75-80. 2015.
- [4] SOUZA, Ramon Hugo de. DANTAS, Mario Antonio Ribeiro. *Mapping QoE through QoS in an Approach to DDB Architectures: research analysis and conceptualization*. ACM Comput. Surv. 48(2):31, 2015.

- [5] INACIO, Eduardo Camilo;; DANTAS, Mario A. R. *survey into performance and energy efficiency in HPC, Cloud and big data environments*. IJNVO14 (4):299-318,2014.
- [6] CHIN, Won Sang; KIM, Hyun-soo; HEO, Young Ju; JANG, Ju Wook. *A Context-based Future Network Infrastructure for IoT Services*. The 10th International Conference on Future Networks and Communications (FNC 2015).
- [7] DANTAS, Mario. *Computação distribuída de alto-desempenho: redes, clusters e grids computacionais*, Axcel Books, 2005.
- [8] GUBBI, Jayavardhana; BUYA, Rajkumar; MARUSIC, Slaven; PALANISWAMI, Marimuthu. *Internet of Things (IoT): a vision, architectural elements, and future directions, future generation computer systems* 29,1645–1660, 2013.
- [9] CHEN, C.L. Philip; ZHANG, Chun-Yang. *Data-intensive applications, challenges, techniques and technologies: a survey on Big data*, *Information Sciences* 275, 314–347, 2014.
- [10] JIANG, Hai; SHEN, Feng; CHEN, Su; LI, Kuan-Ching; JEONG, Young-Sik. *A secure and scalable storage system for aggregate data in IoT*, *Future Generation Computer Systems* 49, 133–141, 2015.
- [11] LIU, Chang; YANG, Chi; ZANG, Xuyun; CHEN, Jinjun. *External integrity verification for outsourced big data in Cloud and IoT: a big picture, future generation computer systems* 49, 58–67, 2015.
- [12] ANDREOLINI, Mauro; COLAJANNI, Michele; PIETRI, Marcello; TOSI, Stefania. *Adaptive, scalable and reliable monitoring of big data on Clouds*, *J. Parallel Distrib. Comput.* 79–80,67–79, 2015.
- [13] BORGIA, Eleonora. *The Internet of Things vision: key features, applications and open issues*, *Computer Communications* 54,1–31, 2014.
- [14] DISTEFANO, Salvatore a,b; MERLINO, Giovanni; PULIAFITO, Antonio. *A utility paradigm for IoT: the sensing Cloud, Pervasive and Mobile Computing* 20, 127–144, 2015.
- [15] HASHEM, Ibrahim Abaker Targio; YAQOUB, Ibrar; ANUAR, Nor Badrul; MOKHTAR, Salimah; GANI, Abdullah; KHAN, Samee Ullah . *The rise of “big data” on Cloud computing: review and open research issues*, *Information Systems* 47,98–115, 2015.
- [16] KAMBATLA, Karthik; KOLLIAS, Giorgos; KUMAR, Vipin; GRAMA, Ananth. *Trends in big data analytics*, *J. Parallel Distrib. Comput.* 74, 2561–2573, 2014.
- [17] ASSUNÇÃO, Marcos D.; CALHEIROS, Rodrigo N.; BIANCHI, Silvia; NETTO, Marco A.S.; BUYA, Rajkumar. *Big data computing and Clouds: trends and future directions*, *J. Parallel Distrib. Comput.* 79–80, 3–15, 2015.
- [18] FELLER, Eugen; RAMAKRISHNAN, Lavanya; MORIN, Christine. *Performance and energy efficiency of big data applications in Cloud environments: a Hadoop case study*, *J. Parallel Distrib. Comput.* 79–80, 80–89, 2015.
- [19] AZIZ, Benjamin. *A formal model and analysis of an IoT protocol*, *Ad Hoc Networks* 36, 49–57, 2016.
- [20] SHAEV, Yury. *From the Sociology of Things to the “Internet of Things*, *Procedia - Social and Behavioral Sciences* 149, 874 – 878, 2014.
- [21] INACIO, Eduardo Camilo; DANTAS, Mario A. R.; MACEDO, Douglas D. J. *Towards a performance characterization of a parallel file system over virtualized environments*. ISCC2015:595-600.
- [22] INACIO, Eduardo Camilo; DANTAS, Mario A. R. ; BOITO, Francieli Zanon; NAVAU, Philippe Olivier Alexandre; MACEDO, Douglas Dyllon Jeronimo de. *Performance impact of operating systems’ caching parameters on parallel file systems*. SAC2015:2066-2068.
- [23] INACIO, Eduardo Camilo;PILLA, Laércio Lima ;DANTAS, Mario Antonio Ribeiro. *Understanding the Effect of Multiple Factors on a Parallel File System’s Performance*. WETICE 2015: 90-92.
- [24] PRADO, Thiago Coelho; MACEDO, Douglas D. J. de. ; DANTAS, M. A. R.; WANGENHEIM, Aldo von. *Optimization of PACS Data Persistency Using Indexed Hierarchical Data*. *J. Digital Imaging* 27(3):297-308,2014.
- [25] NAZÁRIO, Débora Cabral; TROMEL, Igor Vilas Boas; DANTAS, Mario Antonio Ribeiro; TODESCO, José Leomar. *An Approach to Evaluating Quality of Context Parameters in an Ambient Assisted Living Environment*. CBMS 2014: 515-516.
- [26] SILVA, Madalena Pereira da ; NAZÁRIO, Débora Cabral; DANTAS, M.A.R.; GONÇALVES, Alexandre Leopoldo; PINTO, A.R.; MANERICHI, Guilherme; VANELLI, Bruno. *An eHealth Context Management and Distribution Approach in AAL environments*, CBMS2016.
- [27] NAZÁRIO, Débora Cabral; ANDRADE, Adroaldo de.; BORGES, Lucas; RAMOS, Willian Romeu; TODESCO, José Leomar; DANTAS, Mario Antonio Ribeiro. *An Enhanced Quality of Context Evaluating Approach in the e-Health Sensor Platform*. Q2SWinet@MSWIM 2015: 1-7.
- [28] SILVA, Madalena P. da ; GONÇALVES, Alexandre L. ; DANTAS, M. A. R.; VANELLI, Bruno; MANERICHI, Guilherme; SANTOS, Stephan A. R. D. dos. ; FERRANDIM, Mauri; PINTO, Alex R. *Implementation of IoT for Monitoring Ambient Air in Ubiquitous AAL Environments*. SBESC 2015: 158-161.
- [29] GOMES, Eliza; DANTAS, M.A.R.; MACEDO, Douglas Dyllon J de. ; ROLT, Carlos Roberto De; BROCARD, M. L. ; FOSCHINI, L. . *Towards an infrastructure to support Big Data for a smart city project*. In: 25th IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (Wetice), 2016, Paris, França. Proceedings of the 25th IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises. New Jersey: IEEE PRESS, 2016. p. 107-111.

[30] SIDDEQ, Mohammed Mand; RODRIGUES, Marcos. *Novel 3D compression methods for geometry, connectivity and texture*. 3D Research, 7(13), 2016.

[31] SILVA, Madalena Pereira; DANTAS, M. A. R.; GONÇALVES, Alexandre L. *Autonomic Quality of Experience Management of Software-Defined Networks*, *Revue des Nouvelles Technologies de information*, v. RNTI-L-8, p. 79-98, 2016.

[32] ASSUITI, M.; VOLPATO, F; SILVA, Madalena Pereira; DANTAS, M. A. R. *A Software-Defined Network Configuration Providing Differentiated QoS to an eHealth Environment*. In: *International Conference Parallel and Distributed Processing Techniques and Applications*, 2016, Las Vegas. *Proceedings of the The 22nd International Conference on Parallel and Distributed Processing Techniques and Applications*. Georgia, US: CSREA Press, 2016. p. 237-243.

[33] GUIA DE ESTRUTURAÇÃO E ADMINISTRAÇÃO DO AMBIENTE DE CLUSTER E GRID. Brasília, 2006. <https://www.governoeletronico.gov.br/documentos-e-arquivos/guiacluster.pdf>

[34] ABDI, AGENDA TECNOLÓGICA DE TELEMEDICINA http://ats.abdi.com.br/Publicaes%20ATS%20Sade%20%20Telemedicina/telemedicina%20_PT.pdf, 2016.

[35] MINISTÉRIO DA INDÚSTRIA, COMÉRCIO EXTERIOR E SERVIÇOS E MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES.

<http://www.abinee.org.br/informac/arquivos/maavmdic.pdf> 2016.

[36] IEEE, *Journal of biomedical and health informatics: special issue on small things and Big Data: controversies and challenges in digital healthcare*, to be published in 2018.

[37] OBABOX, <https://wcw.obabox.com.br/>, acesso em agosto 2019.

[38] FITBIT, <https://www.fitbit.com/home>, acesso em agosto 2019.

[39] APPLE, <https://www.apple.com/lae/watch/>, acesso em agosto 2019.

[40] SAMSUNG, <https://www.samsung.com/global/galaxy/gear-fit2/>, acesso em agosto 2019.

[41] IORI, Gabriel; STRÖELE, Victor; DANTAS, Mario “*Hold up: modelo de detecção e controle de emoções em ambientes acadêmicos*”. Trabalho aceito SBIE 2019.

[42] GRID5000, <https://www.grid5000.fr/w/Grid5000:Home>, acesso em agosto 2019.

[43] PIOLI, Laércio ; STRÖELE ,V.; MENEZES, A. ; DANTAS, Mario A. R.; INACIO, Eduardo C.; MÉHAUT , Jean-François. *A study research I/O improvement on storage in a large scale computing facility*. Artigo submetido ao WSCAD 2019.



CARTEIRA DE CURSOS BASEADA EM TECNOLOGIA BLOCKCHAIN

Rogério Atem

Resumo

Este projeto teve como objetivo prototipar um sistema de emissão de certificados de conclusão de cursos que possa ser integrado ao atual sistema de gestão acadêmica da ENAP, bem como a de outras instituições que venham a se integrar a uma rede de emissão e certificação eletrônicas. Para tanto, foi empregada a tecnologia Blockchain, que consiste nos chamados livros de registros distribuídos e eletrônicos, capazes de interligar assim diferentes instituições e diferentes sistemas de gestão, bem como permitir que qualquer ator, seja egresso, seja uma organização, verifique a autenticidade do certificado de uma maneira segura, rápida e de baixíssimo custo.

1. INTRODUÇÃO

Rogério Atem

Professor EBTT. Doutor e Mestre em Engenharia de Produção. Bacharel em Informática. MBA em Innovation Management. Diretor do Polo de Inovação IFF/EMBRAPII

As sociedades modernas são competitivas e ágeis, gerando profissionais ávidos por conhecimento, aprimoramento pessoal e incremento curricular. Uma forma de especialização dos profissionais é através de cursos oferecidos por instituições de ensino, algumas vezes patrocinados pela própria empresa empregadora, mas muitas vezes procurados pelos próprios profissionais que desejam algum tipo de especialização (DUARTE, 2003).

Segundo Børresen *et al.* (2018), a dificuldade de validação de documentos causa uma grande brecha para o mercado de falsificações, o que permite maus profissionais tirarem proveito desse cenário, forjando estudos e formações inexistentes, criando três grandes categorias de documentos falsificados:

- A primeira categoria é o documento que parece ser emitido por uma instituição conhecida, mas ele é produzido ilegalmente, a pessoa que possui o diploma nunca concluiu o curso na instituição.
- Na segunda categoria se enquadram os documentos que realmente são emitidos pelas instituições, mas com informações alteradas.
- A terceira categoria são os documentos emitidos por instituições de fachada, onde nenhum estudo é realmente realizado.

Uma tecnologia se destaca como uma grande candidata à solução dos problemas apresentados, ela é conhecida

como Blockchain. Segundo Tapscott *et al.* (2016), Blockchain é um livro digital incorruptível de transações econômicas que pode ser programado para registrar não apenas transações financeiras, mas virtualmente tudo de valor. Apesar da sua primeira e mais conhecida aplicação ser no cenário de movimentações financeiras, o Blockchain pode ser usado em diferentes cenários para aumentar a segurança da informação, e uma dessas aplicações é na emissão de documentos digitais. (VILNER, 2018).

Tecnicamente, nas palavras de seu criador Satoshi Nakamoto (2008), que é um pseudônimo para uma pessoa ou um grupo de pessoas ou empresas (THE ECONOMIST, 2015), o Blockchain consiste em blocos de informação ligados uns aos outros através de *hashcodes* calculados através da *Proof of Work* (Prova de Trabalho), esses *hashcodes* são baseados nas informações já existente no próprio Blockchain, formando um registro que não pode ser alterado sem refazer a *Proof of Work*. Satoshi Nakamoto apresentou o conceito de Blockchain primeiramente aplicado à transações financeiras, que é a tecnologia que é conhecida hoje como Bitcoin (ACHESON, 2018).

Hooper (2018) cita algumas vantagens no uso do Blockchain para persistência de um registro de dados:

- **Transparência:** o registro de dados se torna mais transparente com o uso do Blockchain. Por ser um registro distribuído, todos os participantes da rede compartilham a mesma documentação. A versão compartilhada só pode ser atualizada através de um consenso, o que significa que todos os participantes precisam estar de acordo com a mudança. Para alterar uma única transação já persistida, é necessário atualizar todos os registros subsequentes.
- **Segurança:** o Blockchain, de várias formas, mostra-se mais seguro que outros sistemas de armazenamento de dados. Quando uma transação é aprovada, ela é criptografada e anexada ao bloco, esse bloco é replicado e armazenado em todos os participantes da rede, tornando praticamente inviável que acessos maliciosos comprometam a integridade da informação.
- **Rastreabilidade:** dada a natureza dos dados no Blockchain, a informação é altamente rastreável, sendo que nenhum dado é removido, o que facilita para cenários de auditoria.
- **Eficiência e velocidade:** qualquer informação persistida no bloco é considerada verídica, o que torna o processo de validação tão rápido quanto o tempo de processamento de uma nova informação e persistência no Blockchain.
- **Redução de custos:** o Blockchain torna obsoleto o uso de intermediários para garantir a autenticidade da informação. Outro fator que influencia na redução de custos é a replicação natural da informação, tornando o uso de uma estrutura de *backup* robusta desnecessária.

1.2 Objetivos

Como objetivo geral, este projeto teve desenvolver uma estrutura para emissão de certificados referentes a cursos oferecidos por instituições de ensino. O objetivo principal é desenvolver esta solução utilizando a tecnologia Blockchain, interligando instituições através de uma rede de comunicação.

Além do objetivo geral, o trabalho possui diferentes objetivos específicos, que são:

- Realizar levantamento sobre as implementações de Blockchain usadas atualmente no mercado.
- Definir a arquitetura do projeto.
- Implementar a solução em ambiente de teste no IFF e na Enap.
- Possibilitar futuras extensões ao projeto, permitindo criação de novos módulos e de aplicações cliente, além do uso de informações coletadas para análise de dados e tomada de decisão.

1.3 Estrutura

Este relato se divide em cinco tópicos, enumerados a seguir:

- Após esta Introdução, o segundo tópico apresenta brevemente a tecnologia Blockchain, através de uma breve revisão bibliográfica.
- No terceiro tópico é feito um estudo sobre diferentes tecnologias para implementação de redes *Blockchain*, visando identificar a que mais se adequava à proposta deste trabalho.
- O quarto tópico mostra detalhes do desenvolvimento da solução, levantando quesitos técnicos e limitações da tecnologia usada.
- O quinto e último tópico apresenta as conclusões e possibilidades de trabalhos futuros.

2. BLOCKCHAIN

Blockchain é uma tecnologia descentralizada de gerenciamento de transações e dados desenvolvida primeiro para a criptomoeda Bitcoin, cujo interesse vem aumentando desde quando a ideia foi cunhada, em 2008. A razão para o interesse na Blockchain é sua capacidade de fornecer segurança, anonimato e integridade de dados sem qualquer terceiro no controle das transações. As transações monetárias entre pessoas ou empresas são frequentemente centralizadas e controladas por terceiros. Fazer um pagamento digital ou transferência de moeda requer um banco ou provedor de cartão de crédito como intermediário para concluir a transação. Além disso, uma transação causa uma taxa de um banco ou de uma

empresa de cartão de crédito. O mesmo processo também se aplica em vários outros domínios, como jogos, música, *software* etc. O sistema de transações e todos os dados e informações são controlados e gerenciados por uma organização terceirizada, em vez das duas principais entidades envolvidas na transação (YLI-HUUMO *et al.*, 2016).

O objetivo da tecnologia Blockchain é criar um ambiente descentralizado onde nenhum terceiro esteja no controle das transações e dos dados. Cabe observar que, no caso dos governos, embora haja distribuição de processamento e armazenamento entre os nós da rede, é possível que isso seja feito em uma rede privada controlada por um órgão governamental.

O Blockchain é uma solução de banco de dados distribuída que mantém uma lista crescente de registros de dados que são confirmados pelos nós participantes. Os dados são registrados em um registro público, incluindo informações de todas as transações já concluídas (YLI-HUUMO *et al.*, *op. cit.*). Alta integridade de transações e segurança, bem como a privacidade dos nós, são necessárias para evitar ataques e tentativas de perturbar transações no Blockchain e, adicionalmente, as transações distribuídas exigem poder computacional (SWAN, 2015). Prototipar soluções e testar sua segurança e escalabilidade colaboram no entendimento dessas questões.

Atualmente, a Bitcoin, criptomoeda mais conhecida, é também a aplicação mais comum de Blockchain (COINMARKETCAP, 2016), e o número de transferências nesta moeda cresce constantemente (KONDOR *et al.*, 2014). O Bitcoin usa o mecanismo de infra-estrutura de chave pública (*Public Key Infrastructure*, PKI). Neste mecanismo, o usuário tem um par de chaves públicas e privadas. A chave pública é usada no endereço da carteira do usuário Bitcoin, e a chave privada é para a autenticação do usuário. A transação do Bitcoin consiste na chave pública do remetente, múltiplas chaves públicas do receptor e o valor transferido. Em cerca de dez minutos, a transação será gravada em um bloco. Este novo bloco é então vinculado a um bloco previamente escrito. Todos os blocos, incluindo informações sobre cada transação feita, são armazenados em disco dos usuários, chamados nós. Todos os nós armazenam informações sobre todas as transações gravadas da rede *Bitcoin* e verificam a exatidão de cada nova transação feita usando blocos anteriores. Os nós são recompensados, verificando a exatidão das transações. Este método é chamado de mineração, e é confirmado com *Proof of Work*, que é um dos os principais conceitos da tecnologia Blockchain. Quando todas as transações são confirmadas com sucesso, existe um consenso entre todos os nós. Os novos blocos estão ligados aos blocos anteriores e todos os blocos estão alinhados em uma cadeia contínua (HOUSLEY, 2004)

Para o entendimento da aplicação do Blockchain, o seu uso será analisado no cenário de implantação proposto por este estudo, uma explicação mais abstrata será usada, evidenciando somente pontos cruciais para essa implementação e omitindo detalhes não cruciais para o entendimento da tecnologia.

Dada uma transação, que neste cenário é a emissão de um certificado, um novo bloco é criado com os detalhes do certificado e um cabeçalho, que possui o *hashcode* do bloco atual, do anterior e do próximo, formando uma

lista encadeada e também possui um elemento chamado *Nonce*, que é utilizado na *Proof of Work*, e será detalhado na continuação deste texto.

Figura 1 – Blocos do Blockchain.



Fonte: Melo, 2019.

O *hashcode* do novo bloco a ser inserido no Blockchain é calculado baseado nos detalhes do certificado, e na informação do bloco anterior, porém, a proposta do Blockchain é que o cálculo desse novo *hashcode* exija um poder computacional grande para ser calculado, sendo que para alterar um bloco no meio da corrente seria preciso calcular o *hashcode* de todos os blocos subsequentes novamente, algo que se torna inviável dada a tecnologia de processamento atual dos computadores mais modernos.

Para melhor entendimento, será usado o seguinte cenário: o *hashcode* de cada bloco deve começar com uma sequência pré estabelecida de caracteres para ele ser incluído no Blockchain, e esse *hashcode* é alterado modificando o conteúdo do bloco, neste caso, o *Nonce*. Um cálculo em uma abordagem de tentativa e erro precisa ser feito até que o *Nonce* tenha um valor que faça o *hashcode* do bloco atender ao pré-requisito. O trabalho computacional necessário para calcular o *Nonce* é que é chamado de *Proof of Work*.

A forma de validação de um novo bloco através da *Proof of Work* possui uma grande desvantagem, que é o desperdício de energia, pois um grande poder de processamento é exigido para o cálculo do *Nonce* e todo o trabalho efetuado pelos nós não vencedores é totalmente descartado. Uma forma alternativa à *Proof of Work* para validação de um novo bloco, é a *Proof of Stake*, que em um Blockchain que armazena registros de movimentação financeira, como o Bitcoin, ao invés de cálculo de hashing, as próprias moedas são utilizadas para a validação (KHATWANI, 2018).

Em um Blockchain que faz uso do *Proof of Stake*, Khatwani (*op. cit.*) define o seguinte funcionamento: o validador do próximo bloco é determinado aleatoriamente, e para se tornar um validador, o participante precisa depositar uma quantidade de moedas que é chamada de *stake*,

que funciona como um depósito caução. O tamanho do *stake* determina a probabilidade do participante ser escolhido como o validador do próximo bloco. Quando um participante é escolhido para validar o novo bloco, ele confere se todas as informações no bloco são válidas, em seguida adiciona o bloco ao Blockchain e recebe a recompensa da transação. Caso um validador valide uma transação fraudulenta, ele perderá parte do seu *stake* maior do que a quantia ganha como recompensa.

A vantagem do Blockchain é que o livro público não pode ser modificado ou excluído depois que os dados foram aprovados por todos os nós. É por isso que o Blockchain é bem conhecido por suas características de integridade e segurança de dados. A tecnologia Blockchain também pode ser aplicada a outros tipos de uso. Pode, por exemplo, criar um ambiente para contratos digitais e compartilhamento de dados *peer-to-peer* em um serviço de nuvem. O ponto forte da técnica Blockchain, integridade de dados, é a razão pela qual seu uso se estende também a outros serviços e aplicações, embora apresente alguns desafios (SWAN, *op. cit.*):

- Taxa de transferência: o rendimento potencial de problemas na rede Bitcoin é atualmente maximizado para 7 transações por segundo. Outras redes de processamento de transações como a VISA e o Twitter apresentam 2.000 transações por segundo e o 5.000 transações por segundo respectivamente. Quando a frequência das transações no Blockchain aumenta para níveis semelhantes, a taxa de transferência da rede Blockchain precisa ser melhorada.
- Latência: para criar segurança suficiente para um bloco de transação Bitcoin, é necessário cerca de 10 minutos para concluir uma transação.

Para obter eficiência na segurança, é preciso gastar mais tempo em um bloco, porque ele tem que compensar o custo de ataques com gastos duplos. O gasto duplo é o resultado de gastos bem-sucedidos de dinheiro mais de uma vez. O Bitcoin protege contra o gasto duplo, verificando cada transação adicionada à cadeia de blocos, para garantir que as entradas para a transação não tenham sido gastas anteriormente (BITCOIN, 2016a). Isso torna a latência um grande problema no Blockchain atualmente. Fazer um bloqueio e confirmar a transação deve acontecer em segundos, mantendo a segurança. Para concluir uma transação, por exemplo, na VISA, leva apenas alguns segundos, o que é uma enorme vantagem em relação ao Blockchain.

- Tamanho e largura de banda: no momento, o tamanho de um Blockchain na rede Bitcoin é de mais de 50.000 MB (MegaBytes) (fevereiro de 2016). Quando a taxa de transferência aumenta para os níveis da VISA, Blockchain pode crescer 214 PB (PetaBytes) em cada ano. A comunidade Bitcoin assume que o tamanho de um bloco é de 1 MB e um bloco é criado a cada dez minutos (BITCOIN, 2015). Portanto, há uma limitação no número de transações que podem ser manipuladas (em média 500 transações em um bloco) (ANTONOPOULOS, 2014). Se o Blockchain precisar controlar mais transações, os problemas de tamanho e largura de banda precisam ser resolvidos.
- Segurança: o Blockchain atual tem a possibilidade de um ataque de 51%. Em um ataque de 51%, uma única entidade teria controle total da maior parte da taxa de *hash* de mineração da rede e seria capaz de manipular o Blockchain. Para superar esse problema, mais pesquisas sobre segurança são necessárias.
- Recursos desperdiçados: a mineração de Bitcoins emprega enormes quantidades de energia. O gasto energético no Bitcoin é causado pelo esforço de *Proof of Work*. Existem algumas alternativas nos campos da indústria, como prova de participação. Com a *Proof of Work*, a probabilidade de mineração de um bloco depende do trabalho realizado pelo minerador. No entanto, na Prova de Aposta, o recurso que é comparado é a quantidade de Bitcoin que um minerador detém. Por exemplo, alguém que detenha 1% do Bitcoin pode extrair 1% dos “blocos de prova de aposta” (BITCOIN, 2016b). A questão com recursos despendidos precisa ser resolvida para haver uma mineração mais eficiente no Blockchain.
- Versionamento, *hard forks* e cadeias múltiplas: uma pequena cadeia que consiste em um pequeno número de nós tem uma maior possibilidade de um ataque de 51%. Outro problema surge quando as cadeias são divididas para fins administrativos ou de versão.

Ponto importante a considerar em aplicações Blockchain é se a rede formada deve ser pública ou privada. Segundo a ComputerWorld (2018), no Blockchain público todos podem ler e enviar transações ou participar do processo de consenso no Blockchain, pois não é requerida permissão. Todas as transações são públicas e os usuários podem permanecer anônimos. Já os privados são controlados por uma única organização que determina quem pode ler e enviar transações e participar do processo de consenso. Ainda segundo a ComputerWorld (*op. cit.*) há ainda outros dois tipos de redes de Blockchain. O consórcio de Blockchain, que é controlado por um grupo predefinido e onde o direito de ler e enviar transações para o Blockchain pode ser público ou restrito aos participantes. Os consórcios de Blockchain são considerados “com permissão” e são os mais indicados para a maioria das empresas. Já os Blockchains semi privados são administrados por uma única organização que concede acesso a qualquer usuário que atenda aos critérios preestabelecidos. Embora não seja realmente descentralizado, esse tipo de Blockchain com permissão é mais interessante para casos de uso de B2B (Business to Business) e aplicações governamentais. A principal diferença entre público e privado é o mecanismo de consenso. No público, os usuários não se conhecem, portanto, o nível de confiança é baixo, necessitando uma sobrecarga computacional maior. Assim, a validação de cada transação é bastante demorada. Já na conexão privada, a confiança é maior, pois é baseada na permissão de acesso. Assim, é possível fazer uso de algoritmos compartilhados mais simples e rápidos como alternativa aos previamente demonstrados *Proof of Work* e *Proof of Stake*, resultando, em vez de algumas, milhares de transações por segundo. Além disso, em *Blockchains* privados, os registros das transações podem ser criptografados e estão disponíveis apenas para as partes autorizadas, o que, por sua vez, ajuda a satisfazer os requisitos de privacidade dos participantes.

O modelo de consórcio parece ser o mais adequado para o estudo de caso, em que diferentes instituições participariam do Blockchain, compartilhando certificados, o que diminuiria o volume de processamento na autenticação de transações, sem porém impedir o acesso de leitura aos registros, posto que existem soluções para tanto, como, por exemplo, a proposta por Júnior *et al.* (2018).

Em relação à temática específica deste projeto, a ideia de empregar Blockchain para emitir diplomas já vem sendo implantada há algum tempo por instituições no Brasil e no exterior. Uma das primeiras iniciativas foi a do MIT (Massachusetts Institute of Technology), como relata Sá (2017), que já em outubro de 2017 emitiu seus primeiros diplomas empregando a técnica. Várias instituições a seguiram no mundo, como a Escola de Finanças e Administração de Frankfurt, na Alemanha, a Universidade de Cagliari, na Itália, e a Impacta, no Brasil. Neste ponto, Cardoso e Goya (2018) identificam os problemas dos certificados de cursos impressos:

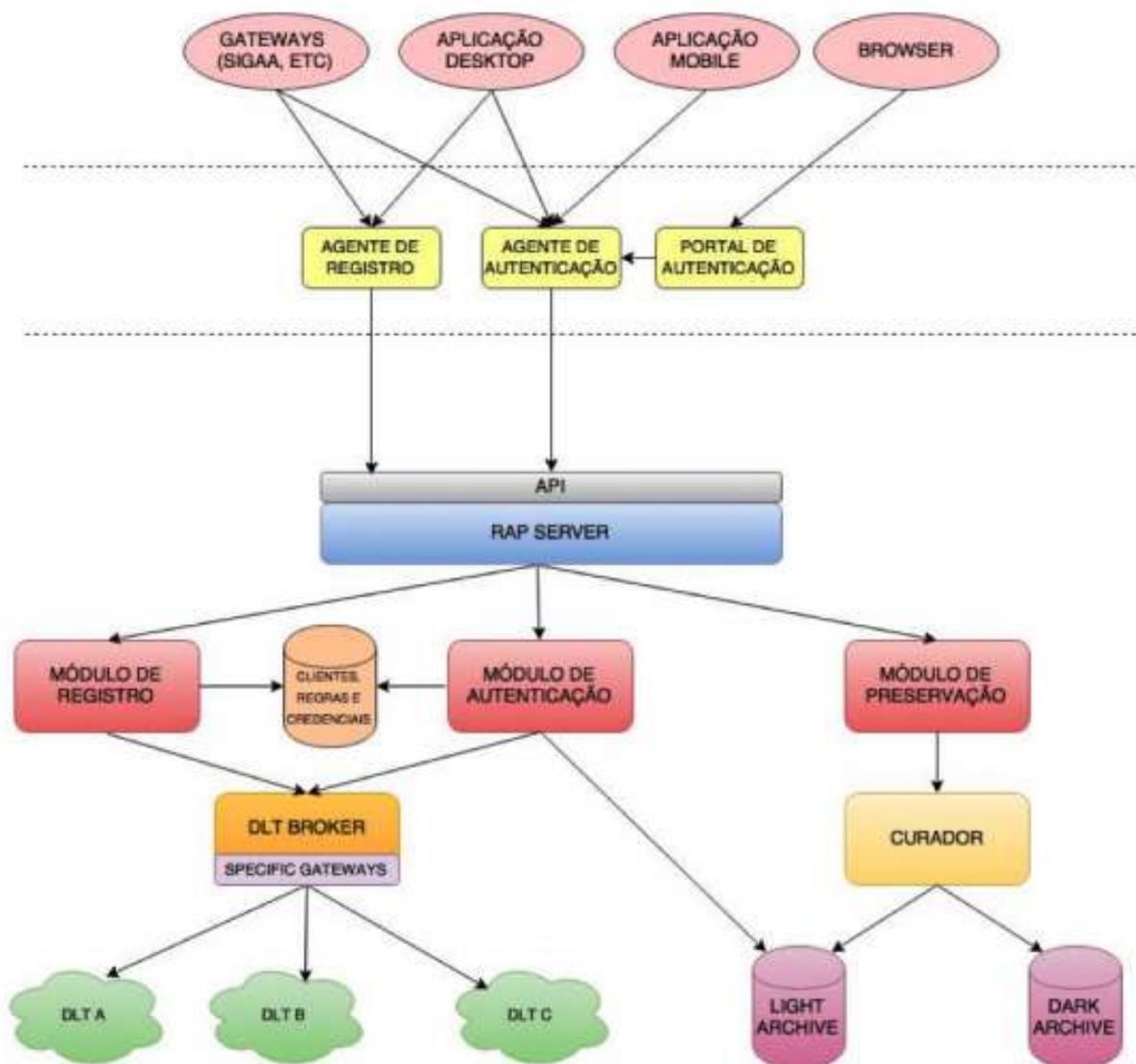
- falsificação, que envolve um mercado bilionário mundial;
- confirmação manual das informações contidas no certificado;

- custos de manutenção dos registros por tempo às vezes indeterminado, que pode ainda gerar adulterações nas informações durante este armazenamento, bem como acessos indevidos à informação;
- limitação das informações contidas em um certificado a um conjunto restrito de metadados.

Cardoso e Goya (*op. cit.*) propõem um modelo conceitual de framework onde as instituições de ensino mantêm os registros (escrita) e os disponibilizam para consulta (leitura), disponibilizando um interface comum de acesso

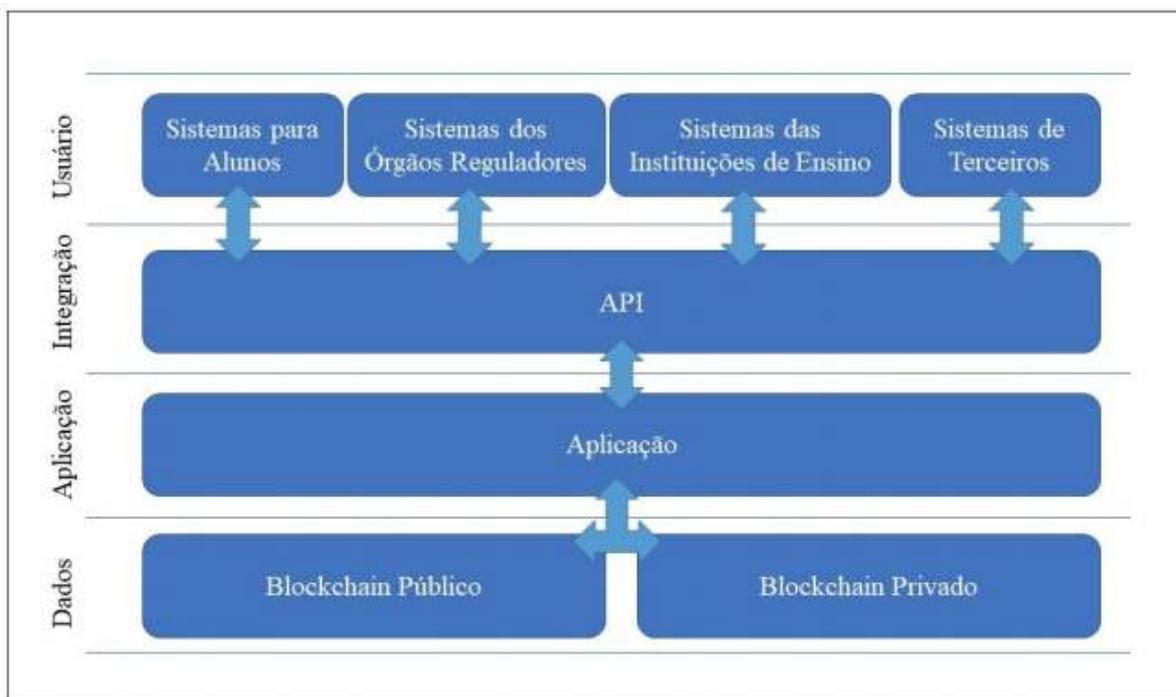
para os participantes da rede. Costa *et al.* (2018) chegam à conclusão similar, citando que o Blockchain pode fornecer prova de propriedade/autoria, prova de integridade e prova de existência, as três dimensões principais da certificação de documentos, e apresentam estudo de caso na implementação de plataforma agnóstica em termos tecnológicos para a proteção de documentos acadêmicos, como na solução de Cardoso e Goya. Esta proposta oferece uma forma de acesso em comum para interação com o Blockchain através de aplicações *Desktop*, *Mobile* e via *Browser*. As Figuras 2 e 3 a seguir, ilustram as soluções propostas por Cardoso e Goya e Costa *et al.*

Figura 2 - Arquitetura do Protótipo



Fonte: Costa, R. *et al.*, 2018).

Figura 3 – Diagrama de Camadas da Arquitetura



Fonte: Cardoso e Goya, 2018.

Em relação à discussão sobre Blockchain versus certificação digital, uma breve e objetiva abordagem é dada pela empresa Link Certificação Digital (2018), que lista os prós e contras de cada tecnologia. Blockchain é segura e barata, mas relativamente de difícil acesso e pouco prática, além de presumir o anonimato. Já o certificado digital é mais prático, acessível e permite identificar as pessoas quando utilizado, mas possui uma estrutura mais onerosa e centralizada. Assim, a empresa indica que o caminho mais viável parece ser fazer com que as duas tecnologias se encontrem em um mesmo ponto, fazendo com que as falhas sejam preenchidas pelas qualidades que uma tecnologia oferece a outra. Mesma conclusão chegaram Costa et al. (op. cit.), que integram Blockchain, certificação digital e preservação de documentos para implementar um serviço de registro capaz, segundo os autores, de interferir ao mínimo nos processos correntes de emissão de documentos acadêmicos. Este serviço de registro possui módulos cliente e módulos servidor, permitindo as funcionalidades tanto de leitura quanto de escrita de dados, dependendo do papel que o usuário tem no serviço. A plataforma oferece uma API (*Application Programming Interface*) REST (*Representational State Transfer*) e um *broker* para as tecnologias Bitcoin e Ethereum, embora outras possam ser incorporadas no futuro.

Cabe ainda ressaltar a existência de um conceito mais genérico, porém de menor granularidade, os *badges* (medalhas), que são reconhecimentos de mérito eletrônicos. A ideia de usar Blockchain para autenticar mérito deu origem ao Open Badge, uma sistemática de emissão de certificados digitais com o padrão de autenticação Blockchain, que visa dar reconhecimento profissional às pessoas de forma transparente e auditável (BRASIL OPEN BADGE, 2018). Um badge, ou “medalha”, possui uma imagem, um nome que identifique

rapidamente a capacitação relativa, a descrição dessa capacitação, os critérios que indicam o por que o ganhador recebeu o *badge*, validade e o emissor. Esses metadados de um badge podem ser empregados como base os metadados para certificados, faltando referências legais da certificação.

3. DEFINIÇÃO DA TECNOLOGIA

Com a publicação do artigo “*Bitcoin: A Peer-to-Peer Electronic Cash System*”, por Satoshi Nakamoto, no ano de dois mil e oito, e, conseqüentemente, o sucesso do Bitcoin, várias outras tecnologias começaram a aparecer no mercado. De acordo com Marr (2018), o uso do Blockchain depois do Bitcoin foi somente para implementações de soluções similares, como o Ethereum, mas, com o tempo, surgiram outras soluções que não tinham como seu principal *asset* uma moeda.

Karthik (2018) define seis soluções de Blockchain como as maiores e mais usadas no mercado, são elas:

- Eris Industries
- Ethereum
- HyperLedger
- Multichain
- Openblockchain
- R3 Corda

Essas seis tecnologias citadas por Karthik como as principais no atual mercado, foram analisadas para a implementação deste projeto, tendo como requisito algumas características como: distribuição gratuita; não voltada à monetização como o Bitcoin; rede privada, já que as primeiras tecnologias tinham como premissa o uso de um rede pública; e suporte à autorização, ou seja, os nós da rede e usuários do Blockchain precisam ser identificáveis.

A primeira analisada e descartada foi a Ethereum, pois sua proposta é a mesma do Bitcoin, oferecer um Blockchain para suportar movimentações financeiras. A R3 Corda, que é projetada especificamente para aplicação de Blockchain em um determinado domínio, especializada para a indústria BFSI (*Bank, Financial Service and Insurance*),

também foi descartada. Outra solução disponível no mercado é a Openblockchain, desenvolvida pela IBM, porém esta veio a se tornar parte da Hyperledger (IBM, 2018), outra tecnologia que será analisada neste estudo, portanto, a análise da Openblockchain se tornou desnecessária. Também excluída deste trabalho, mas por um motivo diferente, foi a solução Eris Industries. Seu *website* e documentação se mostraram instáveis e inacessíveis durante o levantamento das tecnologias realizado neste tópico.

As duas tecnologias que atenderam a todos os requisitos foram Multichain e HyperLedger. O Quadro 1 mostra o resumo do levantamento das seis tecnologias analisadas e o motivo pelo qual não foram selecionadas.

Quadro 1 – Tecnologias selecionadas e eliminadas

TECNOLOGIA	SELECIONADA	MOTIVO (CASO NÃO SELECIONADA)
Eris Industries	Não	<i>Website</i> fora do ar durante o levantamento.
Ethereum	Não	Especializada para movimentações financeiras.
HyperLedger	Sim	
Multichain	Sim	
Openblockchain	Não	Foi incorporada a outra tecnologia da lista.
R3 Corda	Não	Especializada para um ramo diferente.

Fonte: Melo, 2019.



3.1 Tecnologias Candidatas

O HyperLedger surgiu de um esforço colaborativo para evoluir o uso de Blockchains na indústria. É uma colaboração global, na qual a *The Linux Foundation* é a anfitriã e grandes empresas também fazem parte, como líderes na área de finanças, bancos, internet das coisas, manufatura e tecnologia, de acordo com a documentação oficial do produto em seu *website* (2016).

O MultiChain, segundo Greenspan (2015) no *blog* oficial da tecnologia, é uma plataforma para criação de Blockchains privados com o objetivo de oferecer segurança e controle através dessa privacidade. Tanto o HyperLedger quanto o MultiChain, ainda se encontram em constantes atualizações, sendo assim, as informações fornecidas neste estudo são baseadas nas funcionalidades oferecidas pelas tecnologias nas versões oferecidas nas versões HyperLedger 1.3.0 e MultiChain 1.0. O Multichain já disponibilizou uma versão 2.0, mas ainda se encontra na sua fase beta.

3.2 Comparação das Tecnologias

Alguns quesitos foram levados em consideração para a escolha da tecnologia que seria usada para a aplicação projeto. O primeiro desses quesitos foi como cada tecnologia suporta a estruturação dos dados da aplicação, ou seja, se é oferecida uma forma de mapear o modelo da do Blockchain. O HyperLedger oferece um mapeamento orientado a objetos em JavaScript, enquanto o MultiChain trata o asset como uma descrição em JSON (*JavaScript Object Notation*). Para os propósitos deste projeto, ambas as soluções, apesar de diferentes, são viáveis e oferecem vantagens e desvantagens. A vantagem do HyperLedger é uma estrutura mais robusta, podendo ter validações dos dados inseridos no Blockchain, porém, uma atualização no modelo é mais trabalhosa, necessitando uma criação de uma nova versão para qualquer mudança realizada. As vantagens do HyperLedger são justamente as desvantagens do MultiChain: por falta de um modelo pré definido, uma validação de dados teria que ser feita por um intermediário entre o nó e o Blockchain. E pelo fato de não haver

validação do modelo, diferentemente do HyperLedger, uma atualização não implicaria na necessidade de uma nova versão da rede.

O segundo quesito analisado foi o suporte à autorização. O MultiChain trabalha com um conceito de *Streams*, que nada mais é do que rotular alguns blocos do Blockchain como pertencentes a uma *Stream*, podendo aplicar diferentes permissões de leitura e escrita para cada um desses *Streams*. Esse conceito poderia ser usado, por exemplo, criando um *Stream* para cada uma das instituições emissoras de certificados, garantido direito de escrita somente para a instituição detentora do *Stream*. O HyperLedger oferece uma validação baseada em certificados digitais, podendo ser aplicada uma hierarquia de permissões, como exemplo neste projeto; essa hierarquia poderia ser usada para cada instituição ter suas permissões, porém, podendo ainda dar permissões diferentes para cada funcionário. Nesse quesito o HyperLedger se mostrou mais completo e com uma maior escalabilidade.

Como uma das propostas deste projeto é a componentização do Blockchain, oferecendo-o como um serviço para diferentes implementações como aplicações cliente, o terceiro quesito analisado foi se a tecnologia oferece uma API. Ambas disponibilizam acesso ao Blockchain através de uma API REST, porém, o HyperLedger oferece um maior grau de customização desta API, o que não teria muito impacto para cumprir os requisitos deste projeto, no entanto, poderia ser útil em uma futura atualização.

O quarto e último quesito é o uso de *Smart Contracts* pelas tecnologias, que, segundo Chandler (2017) são códigos que são executados no Blockchain para aplicação de regras, lógicas de negócio, e outros variados usos. Nesse quesito, o HyperLedger se destaca, pois ele oferece suporte ao desenvolvimento de *Smart Contracts*, que nessa tecnologia são chamados de *Chain Code*, enquanto o Multichain não oferece esse suporte até a sua versão 1.0.

O Quadro 2 mostra um resumo dos quesitos mais importantes na escolha do HyperLedger sobre o Multichain para a implementação da solução proposta.

Quadro 2 – Comparação entre tecnologias

	HYPERLEDGER	MULTICHAIN
Estrutura de Dados	Estrutura Orientada a Objetos	Representação JSON
Suporte a Autorização	<i>Roles</i> (Multi-hierarquia)	<i>Stream</i> (por Instituição)
API	REST	REST
<i>Smart Contracts</i>	<i>Chain Code</i>	-

Fonte: Melo, 2019

3.3 Tecnologia Selecionada

Para os propósitos deste projeto, o HyperLedger se mostrou mais robusto para atender aos requisitos iniciais, destacando-se no suporte à autenticação, oferecendo vários níveis de permissões. Outra característica do HyperLedger que se destaca é a capacidade de validação dos dados inseridos, não sendo necessário o desenvolvimento de um projeto intermediário; isso é facilitado pelo mapeamento orientado a objetos, que, como já citado, também traz desvantagens, como a necessidade de atualização da versão do modelo em qualquer mudança no mesmo.

Outro requisito que teve grande influência na escolha do HyperLedger foi a possibilidade de uso de *Smart Contracts*, o que o Multichain não oferece. Tendo essa análise sido feita, HyperLedger foi a tecnologia escolhida para a implementação do projeto, pois atende às características dos requisitos. O próximo tópico mostrará como essa implementação foi realizada, descrevendo detalhes técnicos da tecnologia, assim como detalhes do desenvolvimento e também dificuldades encontradas.

4. DESENVOLVIMENTO E IMPLANTAÇÃO

A rede que compõe a comunicação dos participantes do Blockchain é formada por nós que representam as instituições emissoras de certificado. Cada nó, ou cada instituição, possui sua própria e exclusiva réplica do Blockchain.

A ideia principal para que o Ministério da Educação faça parte da rede, é para que ele possa oferecer um mecanismo de leitura do Blockchain, funcionando como um centralizador de informação. A funcionalidade de leitura continuaria

disponível em cada um dos nós, podendo cada instituição escolher exibir os dados somente dos certificados emitidos por ela, ou ainda exibir todos os certificados persistidos no Blockchain. A proposta deste projeto é tornar a leitura do Blockchain altamente flexível.

4.1 Arquitetura funcional

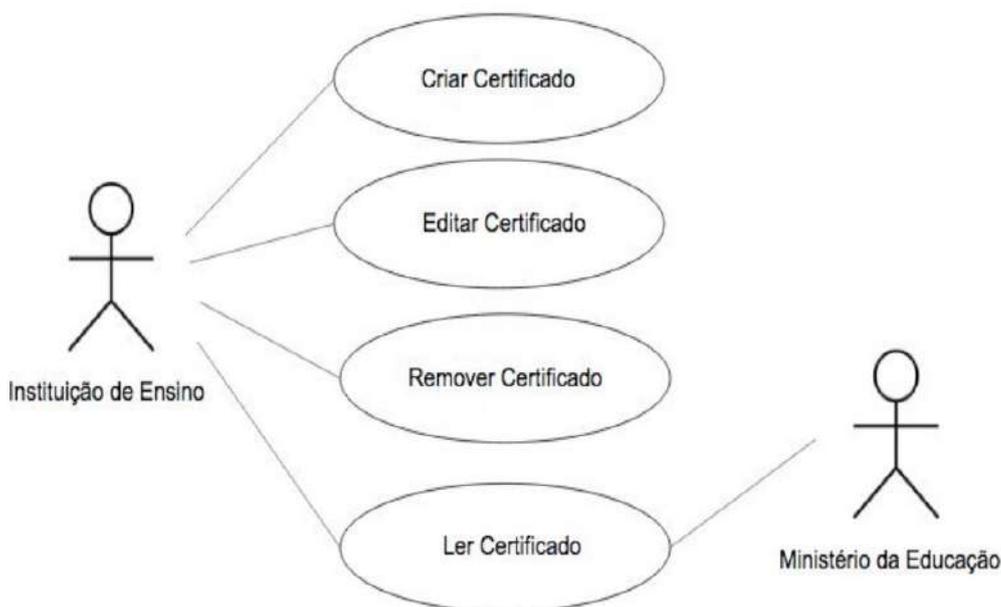
A arquitetura funcional do projeto se divide em três partes, que serão detalhadas após a seguinte breve descrição sobre elas:

- Criação, leitura, edição e deleção dos dados do Blockchain realizadas pelas instituições de ensino.
- Pesquisa sobre os dados do Blockchain disponibilizada pelas instituições de ensino e pelo Ministério da Educação.
- Disponibilização de uma carteira de certificados pelo Ministério da Educação e pelas instituições de ensino para acesso individual do estudante.

4.1.1 Caso de uso: leitura e escrita

A primeira parte da arquitetura funcional é relativa ao acesso de leitura e escrita (criação, edição e deleção) no Blockchain, realizadas pelas instituições de ensino e pelo Ministério da Educação. Toda instituição de ensino participante da rede tem o direito total de leitura e direito parcial de escrita, incluindo somente os certificados emitidos por ela. Este controle de escrita é providenciado através de mecanismos de autenticação e autorização que serão detalhados ainda neste tópico. A Figura 4 mostra o caso de uso de leitura e escrita no Blockchain pelas instituições de ensino e pelo Ministério da Educação

Figura 4 – Caso de uso de leitura e escrita no Blockchain



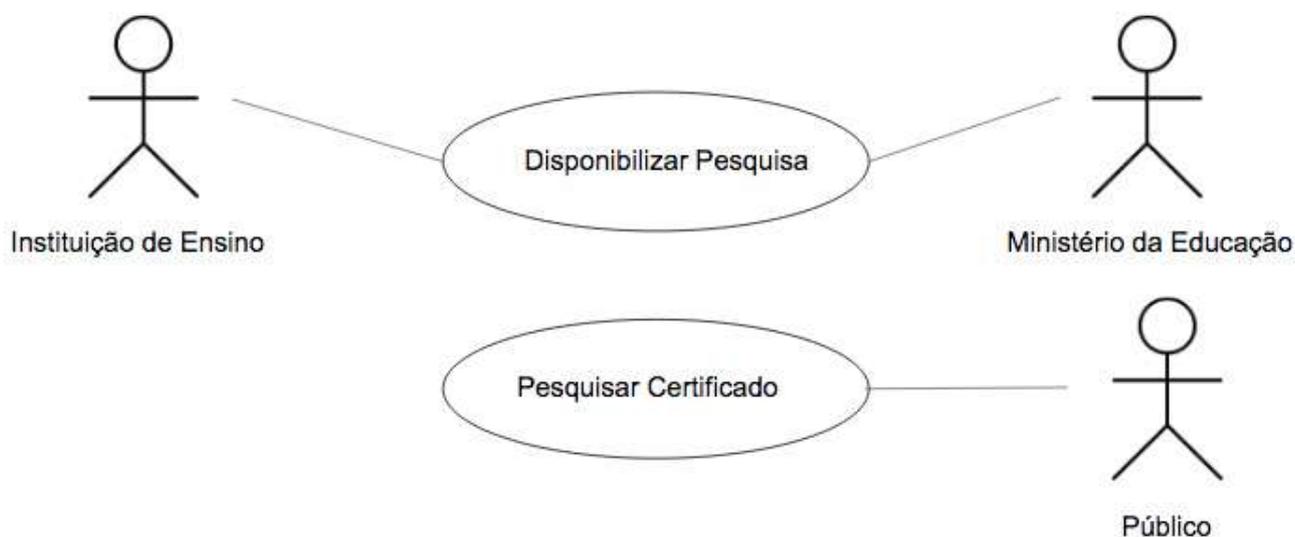
O diagrama mostra a instituição de ensino e o Ministério da Educação como atores do caso de uso, sendo que enquanto as instituições possuem acesso de leitura e escrita, o Ministério da Educação possui somente acesso à leitura, pois não emite certificados.

4.1.2 Caso de uso: pesquisa

A segunda parte da arquitetura funcional é a disponibilização de pesquisa sobre os dados do Blockchain; esta é uma funcionalidade para tornar os dados do Blockchain disponíveis para consulta pública. A proposta principal deste projeto é que o Ministério

da Educação disponibilize essa funcionalidade para pesquisa pública incluindo todo o conteúdo persistido no Blockchain, enquanto as instituições de ensino também possam oferecer uma pesquisa limitada aos certificados emitidos por elas. A Figura 5 mostra o diagrama de caso de uso da disponibilização da pesquisa pública da informação contida no Blockchain sobre os certificados emitidos, as instituições de ensino e o Ministério da Educação são exibidos como atores no diagrama responsáveis por disponibilizar a pesquisa, e o público em geral é representado por um ator responsável por realizar a pesquisa.

Figura 5 – Caso de uso de pesquisa no Blockchain



Fonte: Melo, 2019.

As instituições possuem acesso irrestrito de leitura, cabendo a cada uma implementar a pesquisa de forma mais adequada. Este projeto oferece uma aplicação cliente como exemplo, nessa implementação, a leitura do Blockchain é implementada na visão de um funcionário autenticado no sistema e oferece a visão de todos os certificados no nome do estudante filtrado através do CPF, porém, fornecendo as opções de remoção e edição somente para os certificados emitidos para instituição responsável pela emissão dos mesmos.

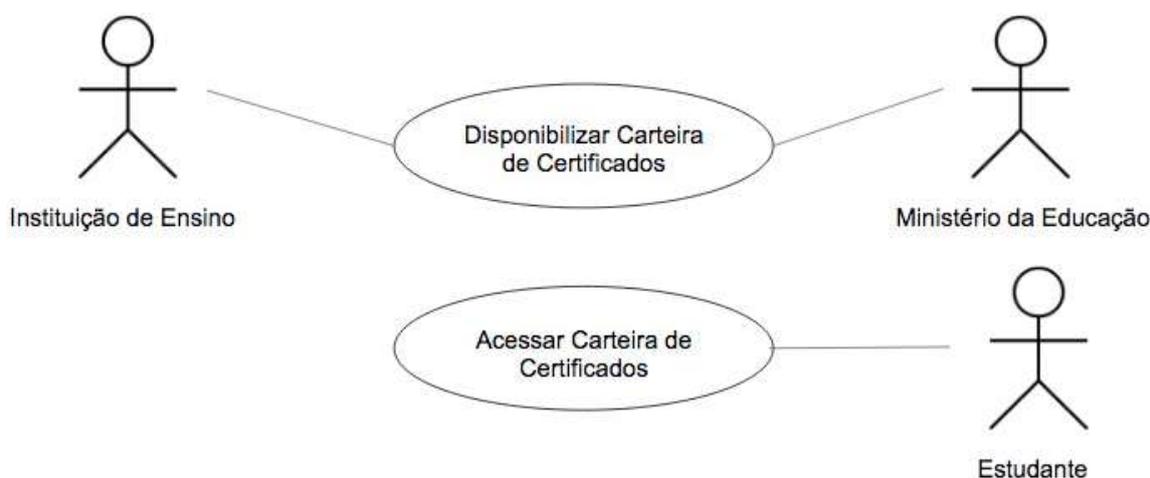
4.1.3 Caso de uso: carteira de certificados

A terceira e última parte da arquitetura funcional é a parte dedicada à carteira de certificados do estudante. Assim como na pesquisa pública, a ideia principal é que esta carteira de certificados abranja todos os certificados conquistados pelo estudante, não dependendo da instituição emissora. A maneira mais intuitiva de alcançar

esse objetivo é o Ministério da Educação disponibilizar tal funcionalidade, agindo como um integrador e centralizador da informação, possibilitando às instituições de disponibilizarem uma carteira de certificados específica para os cursos oferecidos pela própria instituição. Como o acesso à leitura é livre, nada impede as instituições que também tenham acesso de leitura aos certificados emitidos por outras, mantendo a proposta de flexibilidade deste projeto. Na aplicação demonstrativa, que será apresentada neste tópico, a instituição exibe ao estudante todos seus certificados emitidos, concretizando a flexibilidade da arquitetura deste projeto.

A Figura 6 detalha a interação do estudante com a sua carteira de certificados. No diagrama as instituições de ensino e o Ministério da Educação são exibidos como atores responsáveis pela disponibilização da carteira de certificados, e o estudante é representado como um ator que pode acessar a sua própria carteira.

Figura 6 – Caso de uso carteira de certificados



Fonte: Melo, 2019.

4.2. Componentes

Cada nó da rede consiste em três componentes. O primeiro é o próprio Blockchain desenvolvido no HyperLedger Fabric; o acesso ao Blockchain é feito através do segundo componente, um servidor que disponibiliza uma API REST. O terceiro componente é a aplicação cliente que as

instituições emissoras de certificado usarão para acessar o Blockchain através do Servidor REST. Esta aplicação cliente será altamente customizável, permitindo que qualquer instituição desenvolva sua tecnologia, sendo necessário somente usar a API disponibilizada para acesso. A Figura 7 ilustra o relacionamento dos três componentes do nó.

Figura 7 – Componentes do nó Blockchain.



Fonte: Melo, 2019.

4.3 Arquitetura cliente servidor

Os três elementos de cada nó do Blockchain são separados em uma arquitetura Cliente Servidor; a aplicação cliente corresponde à parte Cliente e O Servidor REST e a réplica do Blockchain de cada nó se entendem pela parte do Servidor.

A proposta do projeto é disponibilizar toda a arquitetura entendida como Servidor através de plataformas IaaS (*Infrastructure as a Service*), o que se pode chamar de nuvem para esse cenário, termo utilizado na área de Tecnologia da Informação para definir a distribuição de

serviços de computação – servidores, armazenamento, bancos de dados, redes, software, análises, inteligência e muito mais pela Internet, proporcionando inovações mais rápidas, recursos flexíveis e economia na escala (MICROSOFT, 2019). O Blockchain é replicado para cada instituição devido à característica padrão do protocolo, o que leva também à necessidade de replicar o Servidor REST, em um relacionamento de um-para-um com a réplica do Blockchain. O uso de instâncias exclusivas do Servidor REST para cada nó também tem ligação com o suporte à autorização, que será discutido em detalhes nos tópicos seguintes.

4.4 Modelo

A seguir, será discutido o modelo da aplicação e as decisões tomadas para se chegar ao resultado final.

4.4.1 Terminologia

A tecnologia HyperLedger oferece uma estrutura robusta para criação de um modelo de dados, e é preciso estar habituado com sua terminologia para compreender a capacidade e limitações da plataforma. A seguir são descritos os principais conceitos usados pelo projeto para a criação de um modelo. Os nomes foram mantidos em inglês, já que são usados como palavras chaves no código:

- *Asset* - É o elemento principal do Blockchain, no caso do projeto sendo apresentado é o certificado.
- *Participant* - São elementos que apresentam uma relação de cliente no Blockchain.

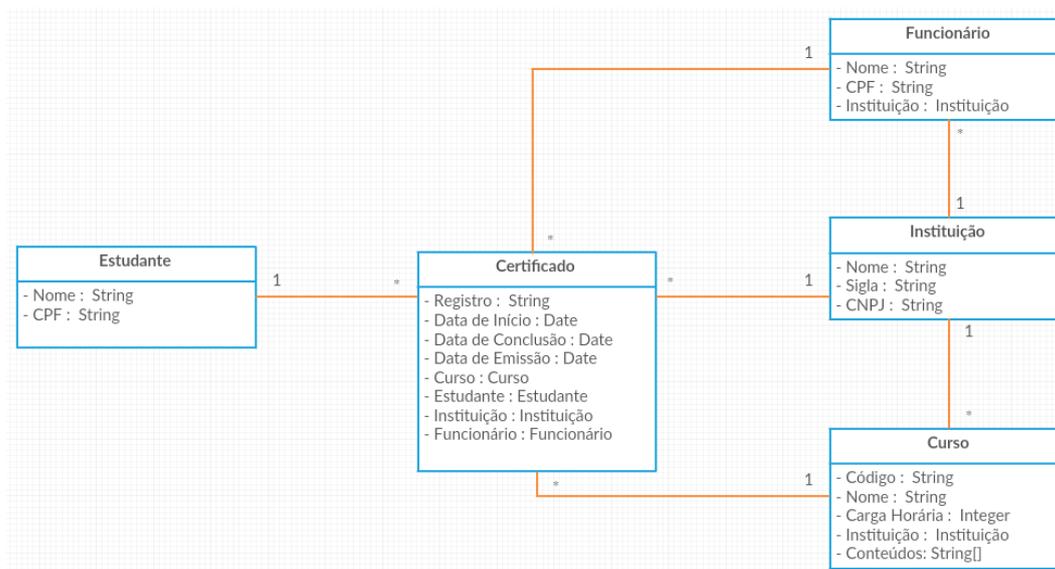
- *Transaction* - Operações realizadas no Blockchain que alteram sua estrutura, geralmente inserindo um novo bloco.
- *Concept* - São usadas para definir um tipo complexo de dados, formado pela composição de informações simples.

O HyperLedger ainda oferece outros elementos para a criação do modelo, mas pela simplicidade da explicação e pelo fato desses elementos não serem usados no projeto, eles não foram mencionados.

4.4.2 Modelo relacional

No Blockchain, é preciso somente armazenar as informações pertinentes ao certificado e outras para controle de segurança e auditoria; por exemplo, para o estudante que conquistou o certificado, é preciso somente seu nome e um identificador único, que nesse caso foi usado o CPF (Cadastro de Pessoa Física), os outros dados cadastrais do estudante que a instituição possui não são necessários para o certificado. Caso haja uma futura necessidade de persistir outros dados, pode-se enriquecer o modelo. A Figura 8 mostra o modelo relacional.

Figura 8 – Modelo relacional adaptado do Blockchain.



Fonte: Melo, 2019.

Esse modelo relacional foi criado devido ao fato do HyperLedger trabalhar com um mapeamento Orientado a Objetos, mas pode-se notar claramente que o *asset* da transação, que é o certificado, é a peça principal que será persistida no bloco do Blockchain. Os outros componentes, além de serem referenciados pelo certificado, possuem relacionamentos que facilitam a segurança e a auditoria, como o relacionamento entre Funcionário e Instituição, assim como Instituição e Curso.

De acordo com os componentes do HyperLedger mostrado no tópico anterior (*Asset*, *Participant* e *Concept*), cada classe

do diagrama é representada por um tipo de elemento. Primeiramente, há o certificado como o *Asset*, que é o elemento principal no Blockchain. As classes Instituição, Estudante e Funcionário, três das quatro classes restantes, são definidas como *Participant*, esse elemento representa entidades que possuem um relacionamento de cliente com o Blockchain, o que significa que quando um aluno ou funcionário interagir com a API REST, o HyperLedger os reconhecerá e aplicará as devidas abordagens de segurança. A Instituição também é definida como *Participant*, pois regras de segurança podem ser aplicadas a nível de instituição ou a nível de funcionário.

O último elemento é o Curso, que é definido como um *Concept*. Ele é somente um tipo complexo do certificado, possuindo suas propriedades. Durante o desenvolvimento foi encontrado um comportamento inesperado no acesso à API REST devido ao uso do Curso como um *Concept*, e por isso no modelo ele foi definido como *Participant*. Este mesmo problema também impactou na propriedade Registro do certificado, que a princípio seria um *Concept* contendo três valores: Número, Registro e Folha. Até a conclusão do desenvolvimento do modelo, esse problema do HyperLedger não tinha sido resolvido.

4.5 Interface

O servidor REST é o único ponto de interação do cliente com o Blockchain, e ele cumpre esse papel oferecendo uma API que pode ser acessada pelas aplicações clientes. Essa API providencia seis formas de interação com o Blockchain através de cinco verbos HTTP (*Hypertext Transfer Protocol*). O Quadro 3 informa o modo de interação com a API REST a partir dos métodos HTTP e URLs (*Universal Resource Locator*).

Quadro 3 – API REST.

HTTP	URL	DESCRIÇÃO
GET	/Certificado	Retorna todos os certificados, possibilitando filtragem.
GET	/Certificado/{id}	Retorna o certificado representado pelo identificador {id}.
POST	/Certificado	Cria um novo certificado.
PUT	/Certificado/{id}	Atualiza o certificado representado pelo identificador {id}.
DELETE	/Certificado/{id}	Remove o certificado representado pelo identificador {id}.
HEAD	/Certificado/{id}	Consulta se o certificado representado pelo identificador {id} existe.

Fonte: autor

Todas as informações enviadas pelas aplicações clientes através dos métodos HTTP Get, Delete e Head são passadas unicamente pela URL, somente o Post e o Put necessitam passar informações no corpo do protocolo. A API REST aceita tanto informações em formato JSON ou em formato

XML (*eXtensible Markup Language*). A seguir, no Quadro 4 pode-se ver lado a lado uma representação da mesma requisição de criação de um novo elemento feita através do método POST em formato JSON e a mesma requisição em XML.

Quadro 4 – Requisição JSON e XML.

JSON	XML
<pre>[{ "id": "string", "registro": "string", "inicio": "2019-01-06T09:41:20.617Z", "fim": "2019-01-06T09:41:20.617Z", "emissao": "2019-01-06T09:41:20.617Z", "curso": {}, "estudante": {}, "instituicao": {}, "funcionario": {} }]</pre>	<pre><?xml version="1.0"?> <Inline Model> <id>string</id> <registro>string</registro> <inicio>1970-01-01T00:00:00.001Z</inicio> <fim>1970-01-01T00:00:00.001Z</fim> <emissao>1970-01-01T00:00:00.001Z</emissao> <curso></curso> <estudante></estudante> <instituicao></instituicao> <funcionario></funcionario> </Inline Model></pre>

Fonte: autor

Os elementos que são definidos como *Participant* são referenciados na requisição de criação do certificado através de seus identificadores. Isso significa que esses elementos precisam antes já estarem criados no Blockchain. A API REST fornece as mesmas formas de interação relativas ao certificado para os elementos

Participant, e isso significa que também pode-se pesquisar, criar deletar e atualizar dados desses elementos. Então, na criação do certificado é necessário pesquisar se suas dependências já estão criadas no Blockchain; caso alguma não exista, estas precisam ser criadas antes do envio da requisição de criação do certificado.

Para a consistência do modelo, o HyperLedger oferece, além de tipagem das propriedades, o uso de validações através restrições como tamanho de campo, limite mínimo e máximo para propriedades numéricas e também oferece o uso de expressões regulares. Essas expressões regulares podem ser usadas por exemplo para validar um formato de CPF ou do número do registro do certificado.

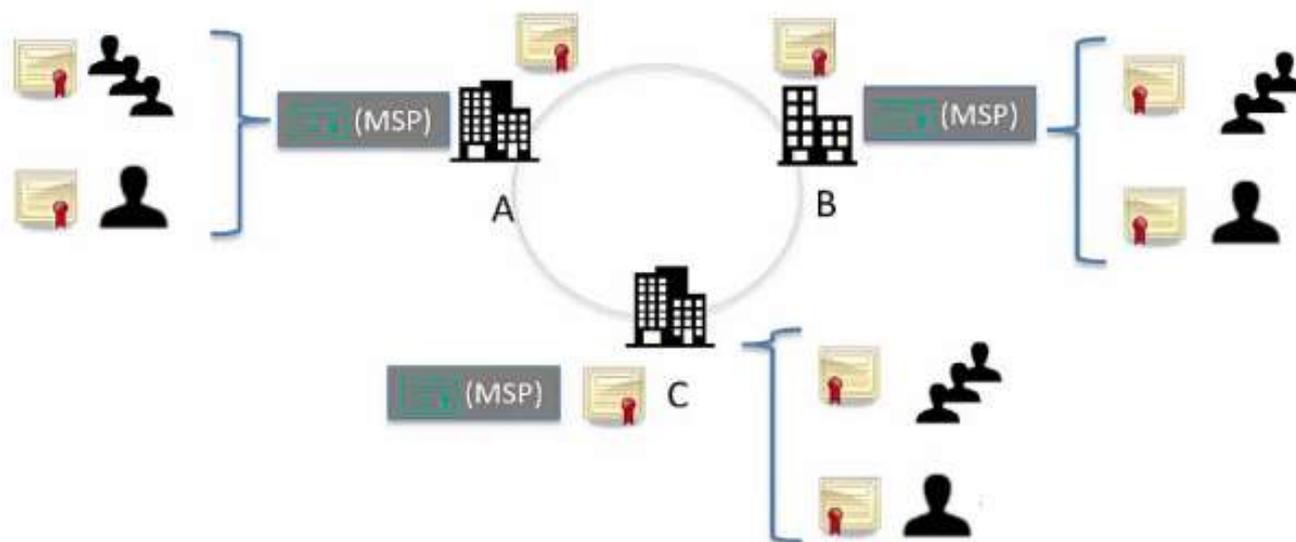
4.6 Segurança (rede privada e suporte a autorização)

Um dos quesitos na escolha da tecnologia para a implementação deste projeto foi o oferecimento de uma rede permissionada, já que a proposta original da tecnologia Blockchain era propor uma rede onde seus nós sejam anônimos (SHARMA, 2018). O HyperLedger oferece uma rede onde não há anonimato entre seus nós, e isto é alcançado através do uso de certificados digitais de chave pública baseados no protocolo X.509.

Para melhor entendimento de como o HyperLedger suporta a autorização, primeiramente é preciso entender o papel desempenhado pelo MSP (*Membership Service Provider*), que, segundo a documentação oficial da tecnologia, é responsável por abstrair todos os mecanismos criptográficos e protocolos por trás da criação e validação de certificados. Em um nível de abstração maior, para a proposta deste projeto, o MSP pode ser entendido como o conjunto de três funções. A primeira é a de CA (*Certificate Authority*), responsável por emitir certificados digitais no Blockchain; a segunda função é a de VA (*Validation Authority*), responsável pela validação dos certificados; e a terceira e última função é a de RA (*Registration Authority*), que é responsável por verificar a requisição de um novo certificado e enviar para o CA (Rouse, 2006).

A proposta deste projeto é que cada nó possua seu MSP exclusivo, responsável pelo gerenciamento de protocolos locais, podendo ser aplicadas vários níveis de autenticação, como a nível de departamento da instituição ou até a nível de funcionário.

Figura 9 – Membership Service Provider



Fonte: Sakhuja (2017).

Na Figura 9 é exibida a integração das instituições com o MSP, sugerindo uma hierarquia de permissões baseadas em certificados digitais de chave pública. A Figura mostra três instituições representadas pelas letras A, B e C, cada uma com um MSP exclusivo com um certificado digital de chave pública que garante o acesso da instituição ao Blockchain. No contexto dentro da instituição, a hierarquia de permissões fica a cargo na mesma. A Figura sugere níveis de permissões individuais, representados pela silhueta de uma pessoa e departamentais, representados pelo conjunto silhuetas.

4.7 Auditoria

A principal característica de um Blockchain é ser um modelo *append-only*. Isso significa que cada nova

informação é registrada no banco como um novo registro, nenhuma atualização sobrescreve outra já existente com o novo valor (CROOK, 2014). Isso permite rever todo histórico de edições. O HyperLedger otimiza a ideia inicial do conceito Blockchain dividindo a informação em dois registros distintos. O primeiro registra todas as transações realizadas em ordem cronológica e o segundo registra o estado atual do *asset*.

Neste projeto, qualquer informação de emissão de certificado ficará persistida nesse registro, facilitando a inspeção do histórico, e, como é característico do Blockchain, esse histórico de registro ficará replicado em cada nó por toda a rede.

4.8 Aplicação cliente demonstrativa

A solução foi desenvolvida para ser altamente customizável no quesito de aceitação de aplicações clientes, e por isso foi facilmente acoplada ao SUAP, plataforma administrativa utilizada pelo IFF; esta integração será abordada no próximo tópico. Durante o desenvolvimento foi criada uma

aplicação cliente demonstrativa, simulando um conjunto de operações denominado CRUD (*Create Read Update Delete*), que são as operações comuns de leitura e escrita, incluindo: criação, leitura, atualização e remoção de dados. Essa aplicação se divide em dois cenários, o primeiro é a visão do estudante sobre sua carteira de certificados.

Figura 10 – Carteira de Certificado na aplicação cliente



Nome	Curso	Carga Horária	Data de Emissão	Instituição	Ação
Jean Melo	Gestão em Ouvidoria	20 horas	19/12/2018	ENAP	Ver
Jean Melo	Gerenciamento de Redes	20 horas	13/12/2018	IFF	Ver
Jean Melo	Lógica de Programação	20 horas	13/12/2018	IFF	Ver

Fonte: Melo, 2019.

Quando o estudante se encontra autenticado no sistema, ele tem acesso a todos os seus certificados emitidos por diferentes instituições. Também cabe à instituição oferecer acesso de pesquisa mesmo para usuários que não fazem parte da instituição, ou seja, pessoas e instituições que não possuem acesso ao sistema através de autenticação.

O próximo cenário é a visão de um funcionário da instituição com permissão de criação e atualização de dados no Blockchain. No seu acesso ao sistema, depois de autenticado, ele tem acesso completo à funcionalidade de leitura de todos os certificados do Blockchain, porém, as permissões de criação, atualização e remoção de dados são limitadas somente aos certificados emitidos pela própria instituição, permissões que o HyperLedger

também permite limitar a nível departamental. É válido lembrar que mesmo as informações de atualização e remoção dos dados são realizadas em um modelo *Append-Only*, sendo que nenhuma informação já persistida no Blockchain é perdida.

As próximas três Figuras mostram detalhes da interface da aplicação cliente na visão de um funcionário autenticado no sistema. A Figura 12 ilustra a visão de um funcionário do IFF autenticado no sistema; este possui permissão de visualização de todos os certificados criados no Blockchain, porém só possui a permissão de escrita para os certificados emitidos pelo IFF. Isso fica claro na coluna “Ação”, onde os certificados emitidos pela Enap não exibem a opção “Remover”.

Figura 11 – Funcionário autenticado na aplicação cliente



Nome	Curso	Carga Horária	Data de Emissão	Instituição	Ação
Rogério Atem	Ética e Serviço Público	20 horas	23/12/2018	ENAP	Ver
Jean Melo	Gestão em Ouvidoria	20 horas	19/12/2018	ENAP	Ver
Jean Melo	Gerenciamento de Redes	20 horas	13/12/2018	IFF	Ver Remover
Jean Melo	Lógica de Programação	20 horas	13/12/2018	IFF	Ver Remover
Rogério Atem	Serviços REST	20 horas	17/12/2018	IFF	Ver Remover

Fonte: Melo, 2019.

A Figura 12 exibe um formulário para emissão de um novo certificado, e este formulário permite o preenchimento de dados importantes do certificado como: nome do estudante, nome do curso, data de início e conclusão do curso, data de emissão do certificado e dados relativos ao

número de registro do certificado. A Figura 13 exibe dados de um certificado já criado; na aplicação demonstrativa, esses dados são exibidos ao clicar na ação “Ver” na lista apresentada na Figura 11.

Figura 12 – Formulário de emissão na aplicação cliente

Ivan Ferreira (IFF)

Listar Certificados
Emitir Certificado
X Sair

Emitir Certificado

Estudante:

Curso:

Data de Início:

Data de Conclusão:

Data de Emissão:

Registro - Número:

Registro - Livro:

Registro - Folha:

Fonte: Melo, 2019.

Figura 13 – Visualização de certificado aplicação cliente

Ivan Ferreira (IFF)

Listar Certificados
Emitir Certificado
X Sair

Certificado

Nome: Jean Melo
Curso: Gestão em Ouvidoria
Instituição: ENAP
Carga Horária: 20 horas
Data de Emissão: 19/12/2018
Data de Início: 03/12/2018
Data de Conclusão: 07/12/2018
Registro: Número: 972 | Livro: 972 (FIC) | Folha: 4
Cadastrado por: 111.111.111-11 (Elis Nogueira)

Fonte: Melo, 2019.

4.9 Integração com o Suap

A primeira integração da solução proposta com uma instituição de ensino foi feita no próprio IFF através do Suap. Foi criado um novo módulo na interface do sistema administrativo e um servidor com a API REST e o Blockchain. Todo o processo, desde preparação do servidor para a API REST e o desenvolvimento do módulo no Suap durou aproximadamente duas semanas, sendo dois profissionais responsáveis pela integração, totalizando cento e sessenta homem / hora. Testes foram utilizados durante o desenvolvimento, seguindo uma abordagem BDD (*Behaviour Driven Development*), e comprovaram a natural integração entre o novo módulo e o Blockchain.

Com somente o IFF compondo a rede, o Blockchain pode ser visto como apenas um banco de dados *Append-Only*, já que as principais propostas do Blockchain incluem a integração entre os nós da rede. O segundo nó da rede é o Enap, sua integração com o Blockchain será detalhada no próximo tópico.

As Figuras 14, 15 e 16 mostram telas do módulo do Suap desenvolvidas para integração com o Blockchain. A Figura 14 mostra a tela de criação de um certificado na visão de um funcionário autenticado no sistema com as devidas permissões. A Figura 14 exibe, também na visão do funcionário, uma lista com os certificados, no caso do exemplo, exibe o certificado criado. A Figura 16 mostra a visualização dos dados do certificado emitido.

Figura 14 – Visualização de certificado aplicação cliente

Início » Emitir certificado

Estudante: *	Helber Ferreira Cisilio dos Santos - 2161080	X
Curso: *	428 - Mestrado Profissional em Sistemas Aplicados à Engenharia e Gestão (CAMPUS CAMPOS CENTRO)	X
Data de Início: *	01/01/2016	
Data de Conclusão: *	31/12/2017	
Data de Emissão: *	01/01/2018	
Registro - Número: *	1216	
Registro - Livro: *	1216	
Registro - Folha: *	1216	

Enviar

Fonte: Melo, 2019.

Figura 15 – Visualização de certificado aplicação cliente

Nome	Curso	Carga Horária	Data de Emissão	Instituição	Ação
Helber Ferreira Cisilio dos Santos	Mestrado Profissional em Sistemas Aplicados à Engenharia e Gestão	32 horas	01/01/2018	REIT	Ver Histórico

Fonte: Melo, 2019.

Figura 16 – Visualização de certificado aplicação cliente.

Certificado

Nome: Helber Ferreira Cisilio dos Santos

Curso: Mestrado Profissional em Sistemas Aplicados à Engenharia e Gestão

Instituição: REIT

Carga Horária: 32 horas

Data de Emissão: 01/01/2018

Data de Início: 01/01/2016

Data de Conclusão: 31/12/2017

Registro: Número: 1216 | Livro: 1216 | Folha: 1216

Cadastrado por: 123.387.547-78 (Helber Ferreira Cisilio dos Santos)

Fonte: Melo, 2019.

4.10 Base para Solução Enap

Para que a solução desenvolvida entrar em produção é necessária uma série de adaptações à instância do Suap em produção na Enap:

- Adaptação do Módulo de Geração de Certificados para ler das Tabelas de Cursos, Diários e Alunos e gerar os dados, que serão enviados utilizando a interface REST definida pelo plugin Python. Este plugin é independente de plataforma e envia um request para o Blockchain, criando o bloco que representa o certificado;
- Alterar o código desenvolvido no Hyperledger para retornar uma chave única, que identifique o bloco e será devolvida para o plugin e deste

para o Módulo de Geração de Certificados, que finalmente armazenará esta chave numa tabela do Suap.

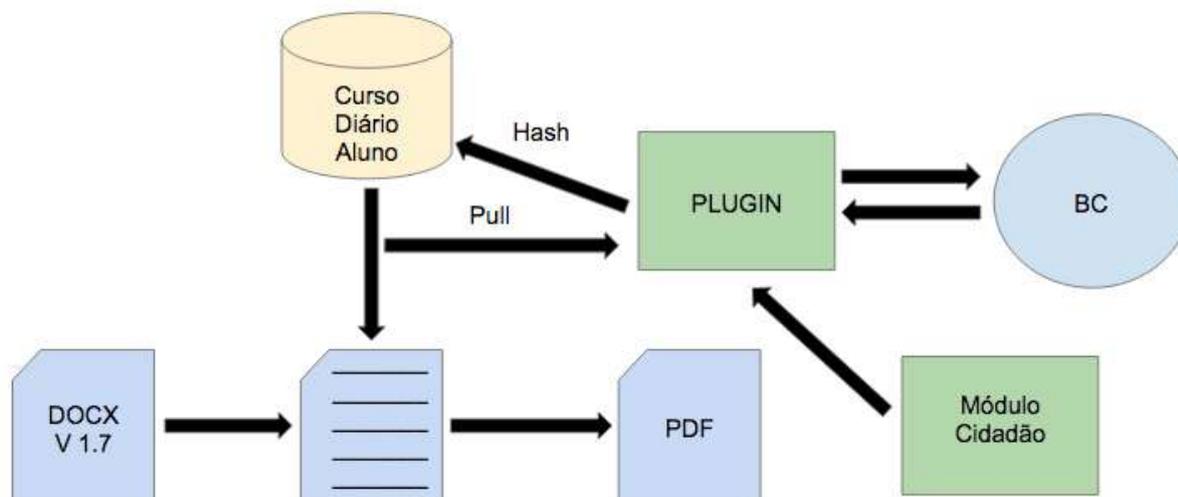
- O plugin será empregado também para fornecer, via REST, dados para qualquer aplicação que deseje acessar a base de certificados da Enap, devendo ainda serem analisadas as questões referentes à segurança, em especial quanto a ataques do tipo *Denial of Service* (DoS).
- Criar e documentar um sistema de versionamento das templates de certificados, o número da versão do certificado deve ser também armazenado como metadado deste certificado. Isso se mostra necessário para emitir segundas-vias de certificados, onde é necessário resgatar o mesmo

padrão do certificado, bem como a assinatura da autoridade competente que o assinou à época da conclusão do curso pelo egresso.

A Figura 17 mostra esquematicamente como funciona a solução aqui descrita. Deve ser salientado que as imagens

dos certificados em PDF não serão armazenadas na Blockchain a princípio, posto que, com os metadados e as templates dos certificados, é sempre possível reconstruir essas imagens, evitando assim o uso excessivo de armazenamento nos nós da rede.

Figura 17 – Arquitetura da Solução Enap



Fonte: Melo, 2019.

5. CONCLUSÕES

Este projeto propôs uma aplicação baseada na tecnologia Blockchain na emissão de certificados acadêmicos. A proposta contempla aspectos técnicos e do domínio específico, mostrando como característica a escalabilidade e componentização do sistema, permitindo qualquer instituição desenvolver seu próprio módulo cliente em sistemas já usados por esta, e, como exemplo, foi desenvolvido um módulo do Suap, sistema administrativo utilizado pela Enap.

Diferentes tecnologias foram investigadas como preparação ao desenvolvimento do protótipo, com uma preocupação maior a respeito do quesito segurança, principalmente abordando o suporte à autorização e auditoria. Duas das tecnologias foram escolhidas para a implementação, quais sejam, HyperLedger e MultiChain. O desenvolvimento do protótipo mostrou que a tecnologia HyperLedger seria a mais adequada para atingir os propósitos deste projeto e uma solução foi desenvolvida baseada nessa tecnologia. Este trabalho detalhou as características e limitações do HyperLedger aplicado ao domínio de interesse, tratando de tópicos como arquitetura do projeto, criação de um modelo, segurança, auditoria, API e integração com uma aplicação cliente demonstrativa.

5.1 Limitações

O estudo destacou algumas limitações tecnológicas e problemas técnicos de menor monta, porém, como a tecnologia está em constante atualização, os problemas

técnicos apresentados neste trabalho se tornam irrelevantes. Outras limitações, referentes ao modelo da tecnologia, foram apresentadas, analisadas, discutidas e finalmente foram encontradas alternativas e soluções para a adequação do projeto.

Cabe dizer que mesmo este estudo tendo sido focado em duas diferentes implementações de Blockchain, suas características, qualidades, pontos fortes e pontos fracos, o conceito de aplicação da emissão de certificados via Blockchain é mais abrangente do que a tecnologia escolhida. O conceito de Blockchain é relativamente novo, ainda existem muitas implementações surgindo no mercado e as existentes continuam em constante evolução, assim como o próprio estudo sobre as capacidades e possibilidades de aplicação do Blockchain em inúmeras áreas.

5.2 Trabalhos Futuros

Um possível trabalho futuro é a integração com Certificação Digital e Preservação de Documentos Digitais, cuja aplicação teve discussão iniciada entre Enap, IFF e UFPB, tendo sido este trabalho e a proposta apresentada por Costa *et al.* (2018) como bases para implementação.

Uma recomendação seria experimentar uma abordagem diferente da usada neste projeto, isto é, ao invés de oferecer somente uma interface REST para integração com aplicações clientes, estudar se é viável e mais adaptável o desenvolvimento de uma aplicação cliente em comum. Isso pode trazer algumas vantagens, como a não necessidade de desenvolvimento de uma aplicação cliente pela instituição, agilizando a implantação do

sistema; mas também traz algumas desvantagens, como a falta de customização da solução, sendo necessário os funcionários responsáveis por emitir o certificado usarem uma segunda aplicação administrativa.

Outra recomendação de trabalho futuro, já baseada na existência do Blockchain com uma carga de informação considerável, seria a aplicação de técnicas de mineração. Através da informação sobre emissão de certificados de diferentes instituições concentrada em um mesmo repositório de dados, seria possível a aplicação desta técnica, utilizando a informação coletada, possibilitando diferentes estudos e análises que podem ajudar na concepção de diferentes relatórios e na tomada de decisão.

REFERÊNCIAS BIBLIOGRÁFICAS

- ACHESON, N. *What is bitcoin?* CoinDesk, 26 jan. 2018. Disponível em: <<https://www.coindesk.com/information/what-is-bitcoin>>. Acesso em: 2 mar. 2019.
- ANTONOPOULOS, A. M. *Mastering Bitcoin: unlocking digital cryptocurrencies*. O'Reilly Media, Inc., 2014.
- BITCOIN. *Bitcoinwiki*. Disponível em: <<https://en.bitcoin.it/>>, 2015. Acesso em: 21 mar. 2018.
- BITCOIN. *Double-spending*. Disponível em: <<https://en.bitcoin.it/wiki/Double-spending>>, 2016a. Acesso em: 21 mar. 2018.
- BITCOIN. *Proof-of-Stake*. Disponível em: <https://en.bitcoin.it/wiki/Proof_of_Stake>, 2016b. Acesso em: 21 mar. 2018.
- BØRRESEN, L.; SKJERVEN, S. *Detecting fake university degrees in a digital world*, 2018. Disponível em: <<https://www.universityworldnews.com/post.php?story=20180911120249317>>. Acesso em: 2 abr. 2019.
- BRASIL OPEN BADGE. *Brasil Open Badge*. Disponível em: <<https://www.brasilopenbadge.com.br/site>> Acesso em: 28 dez. 2018.
- CARDOSO, R.P.; GOYA, D. *Um framework para interoperabilidade de instituições heterogêneas de ensino utilizando Blockchain*. Anais do II Workshop @NUVEM, Universidade Federal do ABC, 2018.
- CHANDLER, R. *Smart Contracts: How to understand smart contracts and be ahead of competition - learn about the future of Blockchain technology*. Wrocław: CreateSpace Independent Publishing Platform, 2017.
- CHENG, S.; DAUB, M.; DOMEYER, A.; LUNDQVIST, M. *Using Blockchain to Improve Data Management in the Public Sector*. Disponível em <<https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/using-blockchain-to-improve-data-management-in-the-public-sector>>, 2017. Acesso em 15 mar. 2018.
- COINMARKETCAP. *Cryptocurrency Market Capitalizations*. Disponível em: <<https://coinmarketcap.com>>, 2016. Acesso em 22 mar. 2018.
- COMPUTERWORLD. *Blockchain privado e público: entenda as principais diferenças*. Disponível em: <<https://computerworld.com.br/2018/03/01/blockchain-privado-e-publico-entenda-principais-diferencas/>>. Acesso em: 30 out. 2018.
- COSTA, R. et al. *Uso não financeiro de Blockchain: um estudo de caso sobre o registro, autenticação e preservação de documentos digitais acadêmicos*. Workshop em Blockchain: Teoria, Tecnologias e Aplicações (WBlockchain_SBRC), [S.l.], v. 1, n. 1/2018.
- CRANE, F. B. *Proof of work, proof of stake and the consensus debate*. Disponível em: <<https://cointelegraph.com/news/proof-of-work-proof-of-stake-and-the-consensus-debate>>. Acesso em: 2 abr. 2019.
- CROOK, P. *Append-only Data Store - FAIMS Mobile Platform User Guide - FAIMS Wiki*. Disponível em: <<https://faimsproject.atlassian.net/wiki/spaces/MobileUser/pages/5865632/Append-only+Data+Store>>. Acesso em: 13 jan. 2019.
- DUARTE, J. *Artigo: A empresa que investe no aprimoramento pessoal de seu empregado.*, 2003. Disponível em: <<https://www.migalhas.com.br/dePeso/16,MI3115,91041-A+empresa+que+investe+no+aprimoramento+pessoal+de+seu+empregado>>. Acesso em: 2 abr. 2019.
- GREENSPAN, G. *Private blockchains are more than “just” shared databases | MultiChain*. Disponível em: <<https://www.multichain.com/blog/2015/10/private-blockchains-shared-databases/>>. Acesso em: 2 mar. 2019.
- HOOPER, M. *Top five blockchain benefits transforming your industry*, 2018 Disponível em: <<https://www.ibm.com/blogs/blockchain/2018/02/top-five-blockchain-benefits-transforming-your-industry/>>. Acesso em: 2 abr. 2019.
- HOUSLEY, R. *Public Key Infrastructure (PKI)*. Disponível em: <<http://dx.doi.org/10.1002/047148296X.tie149>> John Wiley & Sons, Inc., 2004. Acesso em 22 mar. 2018.
- HYPERLEDGER DOCUMENTATION. *About Hyperledger*, 16 set. 2016a. Disponível em: <<https://www.hyperledger.org/about>>. Acesso em: 20 jan. 2019
- HYPERLEDGER DOCUMENTATION. *Introduction — hyperledger-fabric docs master documentation*. Disponível em: <<https://hyperledger-fabric.readthedocs.io/en/release-1.3/whatis.html#permissioned-vs-permissionless-blockchains>>. Acesso em: 13 jan. 2019b.
- HYPERLEDGER DOCUMENTATION. *Membership Service Providers (MSP) — hyperledger-fabric docs master documentation*. Disponível em: <<https://hyperledger-fabric.readthedocs.io/en/release-1.3/msp.html>>. Acesso em: 24 fev. 2019c.

IBM. *Open Blockchain IBM Developer*, [s.d.], 2018 Disponível em: <<https://developer.ibm.com/open/projects/open-blockchain/>>. Acesso em: 2 abr. 2019

JUNIOR, G. M. A. et al. *BNDES Token: uma proposta para rastrear o caminho de recursos do BNDES*. Workshop em Blockchain: Teoria, Tecnologias e Aplicações, (WBlockchain_SBRC), [S.l.], v.1, n.1/2018. Disponível em: <<https://portaldeconteudo.sbc.org.br/index.php/wblockchain/article/view/2355>>. Acesso em: 30 out. 2018.

KARTHIK, K. *6 Blockchain frameworks to build Enterprise Blockchain & how to choose them?* Medium, 21 jan. 2018. Disponível em: <<https://medium.com/hyperlegendary/6-blockchain-frameworks-to-build-enterprise-blockchain-how-to-choose-them-2b7d50ba275c>>. Acesso em: 3 abr. 2019

KHATWANI, S. *What is Proof-Of-Work & Proof-Of-Stake?*, 2018. Disponível em: <<https://coinsutra.com/proof-of-work-vs-proof-of-stake-pow-vs-pos>>. Acesso em: 2 abr. 2019.

KISHIGAMI, J.; FUJIMURA, S.; WATANABE, H.; NAKADAIRA, A.; AKUTSU, A. *The Blockchain-Based Digital Content Distribution System*. IEEE Fifth International Conference on Big Data and Cloud Computing (BDCloud); p. 187–190, 2015.

KONDOR, D.; PÓSFAL, M.; CSABAI, I.; VATTAY, G. *Do the rich get richer? An empirical analysis of the Bitcoin transaction network*. PLoS one, 9(2):e86197, 2014.

LINK CERTIFICAÇÃO, *Digital. Blockchain x Certificado Digital: quem vence essa briga?* Disponível em: <<https://www.linkcertificacao.com.br/blockchain-e-certificado-digital>>. Acesso em 29 dez. 2018.

NAKAMOTO, S. *Bitcoin: a peer-to-peer electronic cash system*, 2008. Disponível em: <<http://bitcoin.com/bitcoin.pdf>>. Acesso em: 22 mar. 2018.

MARR, B. *A Very Brief History Of Blockchain Technology Everyone Should Read*. Disponível em: <<https://www.forbes.com/sites/bernardmarr/2018/02/16/a-very-brief-history-of-blockchain-technology-everyone-should-read/>>. Acesso em: 20 jan. 2019.

MELO, J. F. D. *Desenvolvimento de um Sistema para Emissão de Certificados em Blockchain*. 2019. Dissertação (Mestrado em Mestrado Profissional em Sistemas Aplicados à Engenharia e Gestão) - Instituto

Federal de Educação, Ciência e Tecnologia Fluminense.

MICROSOFT AZURE DOCUMENTATION. *O que é computação em nuvem? Um guia para iniciantes* | Microsoft Azure. Disponível em: <<https://azure.microsoft.com/pt-br/overview/what-is-cloud-computing/>>. Acesso em: 6 jan. 2019.

ROUSE, M. *What is registration authority (RA)? - Definition from WhatIs.com*. Disponível em: <<https://searchsecurity.techtarget.com/definition/registration-authority>>. Acesso em: 24 fev. 2019.

SÁ, V. *MIT Emite Diplomas na Blockchain*. Disponível em: <<https://portaldobitcoin.com/mit-emite-diplomas-na-blockchain-do-bitcoin>>. Acesso em: 20 dez. 2018.

SAKHUJA, R. *Hyperledger Fabric & Composer for Blockchain Development*. Disponível em: <<https://www.udemy.com/hyperledger/>>. Acesso em: 2 mar. 2019.

SHARMA, T. K. *How is Blockchain verifiable by public and yet anonymous?*, 10 jul. 2018. Disponível em: <<https://www.blockchain-council.org/blockchain/how-is-blockchain-verifiable-by-public-and-yet-anonymous/>>. Acesso em: 24 fev. 2019.

SWAN, M. *Blockchain: Blueprint for a New Economy*. O'Reilly Media, Inc., 2015.

TAPSCOTT, D.; TAPSCOTT, A. *Blockchain Revolution: how the technology behind Bitcoin and other Cryptocurrencies is changing the world*. [s.l.] Penguin UK, 2016.

THE ECONOMIST. *Who is Satoshi Nakamoto?* *The Economist*, 2 nov. 2015.

VILNER, Y. *5 Blockchain Product Use Cases To Follow This Year.*, 2018. Disponível em: <<https://www.forbes.com/sites/yoavvilner/2018/06/27/5-blockchain-product-use-cases-to-follow-this-year/#74506df21b60>>. Acesso em: 2 abr. 2019.

YLI-HUUMO, J; KO, D; CHOI, S; PARK, S; SMOLANDER, K. *Where is current research on Blockchain Technology? A Systematic Review*. PLoS ONE, v. 11, n. 10, 2016.



MOBILIZA 360: UM EXPERIMENTO COM APLICAÇÃO DE REALIDADE VIRTUAL COMO INSTRUMENTO DE EMPATIA PARA FORMULAÇÃO DE POLÍTICAS PÚBLICAS

Hélio Lemes Costa Júnior e Letícia Almeida Amaral

Professor Adjunto no Instituto de Ciências Sociais Aplicadas da Universidade Federal de Alfenas, no campus de Varginha – MG. Doutor e mestre em Engenharia de Produção, graduado em Administração e Computação. Chefe do Laboratório de Políticas Públicas para Cidades mais inteligentes e conselheiro da Agência de Inovação na Unifal-MG.

Letícia Almeida Amaral

Pesquisadora no Laboratório de Políticas Públicas para Cidades mais Inteligentes (PP4SCities) no Instituto de Ciências Sociais Aplicadas. É graduada no Bacharelado Interdisciplinar em Ciência e Economia e em Administração Pública, pela Universidade Federal de Alfenas. Atuou no Laboratório de Gestão Simulada em Administração Pública, participando do projeto de pesquisa O2 - Organizações e Outras, com tema sobre coletivos e sustentabilidade. Tem artigos sobre o projeto Mobiliza 360 publicados em evento e revista internacionais.

Resumo

Diversas tecnologias nos permitem ter acesso a conhecimento e situações em qualquer lugar e momento. A realidade virtual é uma delas, que permite acesso com um grau maior de aproximação. A “máquina de empatia final”, como também é chamada, é capaz de fazer com que o usuário vivencie outra realidade a partir de qualquer ponto de vista. Este trabalho buscou comparar como o uso da realidade virtual influencia na construção de empatia a curto e médio prazos e teve como alvo social pessoas em situação de rua. Para verificar os efeitos a curto prazo, foram realizados dois estudos, o primeiro utilizando um vídeo 360° e o segundo aplicando uma narrativa escrita. Os resultados dos estudos mostram que os participantes que tiveram contato com o vídeo 360° demonstraram uma maior empatia com as pessoas em situação de rua do que os participantes que tiveram contato apenas com a narrativa. Para avaliar os efeitos a médio prazo, foi solicitado que os participantes escrevessem uma mensagem, dez dias após a intervenção, aos representantes do Poder Legislativo, emitindo sua opinião e o que era necessário para o Poder Público solucionar e amenizar o problema das pessoas que se encontram em situação de rua. Houve uma baixa adesão a esse experimento, apenas 5% dos participantes enviaram a mensagem, o que dificultou ter um resultado

relevante quanto aos efeitos da realidade virtual no médio prazo. Como ponto positivo desse experimento, ressalta-se que os 5% dos participantes que retornaram a mensagem tiveram contato com a realidade virtual.

Palavras-chave: vídeo 360°, realidade virtual, empatia, pessoas em situação de rua

INTRODUÇÃO

Inovações constantes e o rápido desenvolvimento científico e tecnológico permitiu que fossem criadas inúmeras tecnologias instantâneas e acessíveis que podem ser aproveitadas para obter conhecimento, registrar e transmitir informações e se comunicar em qualquer lugar, a qualquer momento e de diversas formas. No entanto, somente a tecnologia em si não é relevante, é necessário empregá-la de forma que se obtenham soluções para as adversidades que enfrentam diversas áreas importantes, tais como: educação, comunicação, gestão, saúde, entretenimento, entre outras (MORAN, 2007).

A realidade virtual é uma dessas tecnologias, capaz de gerar resultados através de uma inovação de vanguarda de interface, que leva o usuário a um ambiente de imersão e interação, fazendo com que ele vivencie outra realidade (BRAGA, 2001). Os vídeos desenvolvidos para atender a essa tecnologia são gerados por meio de câmeras 360°, o que possibilita que a imersão e a interação sejam feitas com maior proximidade, onde o espectador escolhe qual direção ele deseja observar.

Percebem-se, atualmente, inúmeros benefícios dessa tecnologia, tais como: o auxílio aos profissionais da saúde em procedimentos, tratamentos, diagnósticos, uma participação mais ativa do estudante e a oportunidade de novas experiências na área da educação, contribuição nos processos de treinamento e de tomada de decisões no campo da gestão, entre outros exemplos presentes dentro dessas áreas e em outras diversas (BRAGA, 2001).

A falta de conhecimento acerca da realidade no Brasil é algo preocupante; de acordo com o Instituto IPSOS MORI (2017), o país está na segunda colocação nesse quesito, ficando atrás somente da África do Sul. Esse índice é avaliado através de perguntas sobre temas cotidianos, como: homicídios, gravidez na adolescência, vacinas etc. Ressalta-se que essa percepção pode gerar análises equivocadas sobre os reais problemas do país e, como resultado, soluções inadequadas.

O projeto Mobiliza 360 foi criado em 2018, pelo professor Hélio Lemes Costa, da Universidade Federal de Alfenas, e recebeu apoio do Programa de Cátedras da Escola Nacional de Administração Pública. O objetivo do Mobiliza 360 é avaliar o uso de tecnologias de realidade virtual na gestão pública e difundir o uso dessa tecnologia entre a população e os gestores públicos. Para tanto, foi criado um curso online de produção de conteúdos em 360°, que é oferecido gratuitamente no site do projeto, e foi feita a pesquisa de campo, testando o uso da tecnologia, o que resultou no relato apresentado no presente trabalho.

Este trabalho tem por objetivo averiguar como o uso da realidade virtual 360° impacta as pessoas, e se essa exposição é capaz de promover mais empatia diante de situações adversas. E verificar os impactos causados momentaneamente ou a médio prazo. Este estudo se justifica por sua relevância acadêmica, social e por ser um tema pertinente e atual, uma vez que essas tecnologias podem mudar a forma de ensinar, aprender, agir e tomar decisões.

O presente trabalho é constituído por cinco seções, sendo a primeira referente às considerações gerais; a segunda é o referencial teórico, em que é apresentada uma breve contextualização do tema e das tecnologias relacionadas; a terceira apresenta a metodologia utilizada no desenvolvimento da pesquisa; a quarta refere-se aos resultados obtidos no experimento, e, por fim, são apresentadas as considerações finais, que expõem as conclusões, sugestões e limitações do estudo.

1. REFERENCIAL TEÓRICO

Em 1964, Marshall McLuhan afirmou que “o meio é a mensagem”, ou seja, que a tecnologia usada para transmitir a mensagem é tão importante como a própria mensagem (MCLUHAN, 1979). Por isso, é interessante à ciência observar o impacto que o uso da tecnologia de imagens em 360° e da realidade virtual vem causando na comunicação e quais são suas consequências para as Ciências Sociais.

De acordo com um relatório da empresa americana Digi-Capital (2019), até 2021, tecnologias relacionadas à realidade virtual e realidade aumentada devem movimentar cerca de US\$ 108 bilhões no mercado global. Isso porque são ferramentas únicas e poderosas para criar uma experiência de usuário impactante.

Uma de suas vantagens mais marcantes é a possibilidade de estabelecer uma ligação emocional entre as pessoas e marcas (incluindo aí as ideias). Isso graças às experiências que podem explorar os diversos sentidos e trazer surpresa e encantamento (DIGI-CAPITAL, 2019).

Uma lista, cada vez maior, de meios de comunicação ao redor do mundo inclui conteúdo em 360° no seu portfólio, dedicando recursos tecnológicos e humanos para produções nessa nova mídia (BATHIA, 2015; DOYLE; GELMAN; GILL, 2016).

O que era considerado como inovação radical, de alta tecnologia, subitamente se torna acessível comercialmente e uma fatia maior dos consumidores têm acesso a tais inovações. A rápida evolução da tecnologia da informação faz com que seus custos e complexidade se reduzam significativamente, tornando-se mais acessíveis à população.

É nesse cenário que se encontram as tecnologias de produção de imagem em 360° e o uso de óculos de realidade virtual para visualização dessas imagens. Também a alta penetração dos smartphones, que se tornaram portas

de entrada para a produção e consumo desse tipo de conteúdo, tem auxiliado na sua popularização.

Diversos estudos acompanham a adoção de tais tecnologias, compreendendo seu potencial, assim como reflexos de seu uso nos mais diferentes ambientes e com os mais variados propósitos, como pode ser visto a seguir.

1.1 Imersão

Notadamente, a capacidade de imersão em uma realidade virtual é um dos benefícios das tecnologias citadas, permitindo que alguém que não tem acesso real a um ambiente ou a uma localização possa se sentir imerso naquela realidade, dita, virtual.

Assim, considera-se que a imersão seja capaz de provocar reações emocionais, psicológicas e sensoriais, mais profundas do que apenas a narração de um fato, ou a exposição a imagens bidimensionais, estáticas ou em movimento. Mas será que isso é verdadeiro?

O jornalismo imersivo é um gerador de narrativa em que o receptor da notícia vive em primeira pessoa a história, ele vivencia de fato. Porém, essa ideia de teletransportar o espectador para o local da história, ou do acontecimento, não é inédita, pois no cinema, teatro e em outras artes já existia esse desejo de imergir o usuário na narrativa, sendo apropriado pelo jornalismo mais recentemente, como diz Longhi (2016).

2.2 Empatia

Através da imersão em uma realidade distinta da sua, acredita-se que é possível promover a empatia entre pessoas. A empatia é a capacidade de compartilhar e entender as emoções de outra pessoa, colocando-se no lugar dela, e é um componente obrigatório nas interações sociais bem-sucedidas (HOFFMAN, 2001).

Diversos estudos, citados neste texto, mostram que a empatia aumenta a compreensão e motiva comportamentos pró-sociais. Entusiastas deste tipo de mídia defendem que histórias mais imersivas atraem mais atenção e envolvimento dos espectadores, se comparado aos textos e vídeos tradicionais. A resposta tende a ser mais visceral, e as reações mais emocionais. Os usuários que passam pela experiência de realidade virtual afirmam que a tecnologia os aproxima dos eventos experimentados e quebra as barreiras, inerentemente promovidas por narradores ou intermediários (DOYLE; GELMAN; GILL, 2016).

Em 2015, a realidade virtual foi descrita por Milk (2015) como a “máquina de empatia final”, pois permite que as pessoas vivenciem visceralmente qualquer coisa do ponto de vista de outra pessoa. Estudos anteriores mostraram que a realidade virtual pode aumentar a empatia do usuário em relação aos personagens apresentados em um ambiente virtual (GILLATH *et al.*, 2008; HASLER *et al.*, 2014), incluindo vídeos em realidade virtual de 360° (SHIN, 2018; SHIN; BIOCCHA, 2017) transmitindo experiências ou sentimentos de outra pessoa para o público (SHIN, 2017).

Nos ambientes de realidade virtual, os espectadores podem ter forte empatia com as emoções ou situações de outra pessoa, pois sentem que estão ocupando o mesmo espaço físico e, portanto, estão próximos a essa pessoa (SHIN, 2018). A pesquisa mostrou que os espectadores que experimentam histórias por meio de vídeos de realidade virtual em 360° superam aqueles que leem as mesmas histórias por meio de texto e imagens, não apenas em termos de presença, mas também empatia (SUNDAR; KANG; OPREAN, 2017).

De la Peña *et al.* (2010) argumentam que a capacidade da realidade virtual de obter presença, a sensação subjetiva do usuário de estar dentro de um ambiente virtual imersivo, permitem que os usuários compreendam mais profundamente outras perspectivas além da sua.

Essas experiências colocam os usuários em novos ambientes, mostrando a eles como seria experimentar uma situação específica da perspectiva de outra pessoa. Uma extensa pesquisa mostra que, tendo a perspectiva de outra pessoa (ou seja, imaginando como seria ser outra pessoa), pode ser um método eficaz de promover a empatia e motivar comportamentos pró-sociais (HERRERA *et al.*, 2018).

Esta pesquisa considera que há necessidade de mais evidências empíricas que validem a tomada de perspectiva em realidade virtual como uma forma eficaz de promover a empatia, pois pesquisas anteriores demonstram que a perspectiva tradicional, baseada em narrativas textuais e exposições a imagens convencionais, nem sempre leva para resultados positivos (PIERCE *et al.*, 2013).

Assim, a presente investigação concentra-se em comparar os efeitos de curto e médio prazos de uma tarefa de tomada de perspectiva tradicional com uma tarefa de tomada de perspectiva de realidade virtual e explorar os mecanismos tecnológicos e psicológicos que tornam a tomada de perspectiva um método eficaz de promover empatia e comportamentos pró-sociais.

A realidade virtual é frequentemente usada para simular a realidade física de forma semelhante à do cinema e da televisão, portanto, a tecnologia fornece uma representação real de um mundo, como um ambiente projetado por computador ou filmagem capturada da vida real. “O ‘real’ refere-se apenas à RV em si como uma mídia provendo uma experiência autêntica” (BOLTER; GRUSIN, 2000).

O conceito de imersão em tecnologia digital foi originalmente derivado da concepção de videogames. A imersão tem sido vista como um efeito positivo no sentido em que estabelece uma conexão profunda em um jogo digital entre o jogador e o mundo virtual. De acordo com McRoberts (2017), imersão pode ser definida como um estado de profundo envolvimento ou absorção em uma atividade ou experiência.

Exemplos de organizações de mídia destacam os sentimentos que os espectadores sentiram ao entrar na realidade virtual pela primeira vez. Ao ver, pela primeira vez, uma peça imersiva, Jake Silverstein, editor chefe da revista New York Times, comentou: “Eu editei centenas de histórias sobre refugiados e nunca tive uma experiência como essa” (BOHRER, 2016, p. 36). Quando onze estagiários foram introduzidos à realidade virtual por



meio de Intercâmbio Juvenil de Informações sobre Justiça (JJIE), do projeto Virtual World Journalism, notaram maior sensação de presença e narrativa mais eficaz, como se percebe em um de seus relatos:

É novo, empolgante e mergulha seu público em um evento de notícias, oferecendo a versão mais próxima possível da realidade, é uma maneira mais avançada de apresentar histórias, pessoas reagem a histórias pessoais e emocionais, te mostra como se você estivesse presente. Isso faz você se sentir como se estivesse no mundo real. (BOHRER, 2016, p. 36)

Se os espectadores sentirem que estão lá, eles poderão ter uma experiência bem mais gratificante, devido à qualidade imersiva desse tipo de tecnologia. Muitas peças em realidade virtual podem elicitar essa conexão, e é essa qualidade que um público pode procurar quando se aproxima desse tipo de narrativa.

Uma dessas histórias criadas em realidade virtual sobre o sistema prisional tinha esse efeito de conexão sobre os espectadores. Francesca Panetta, diretora executiva de realidade virtual no jornal britânico *The Guardian* – cujo projeto 6x9 conta com uma experiência em realidade virtual sobre confinamento solitário, diz que a realidade virtual e o conteúdo interativo oferecem uma imersão extra ao *storytelling* no jornalismo (MAYHEW, 2016). Tal imersão extra serve ao jornalismo como uma forma empática de conexão entre espectador e a notícia, e pode ser usada como uma forma de mobilizar pessoas, provocando engajamento real em causas dos mais diversos tipos.

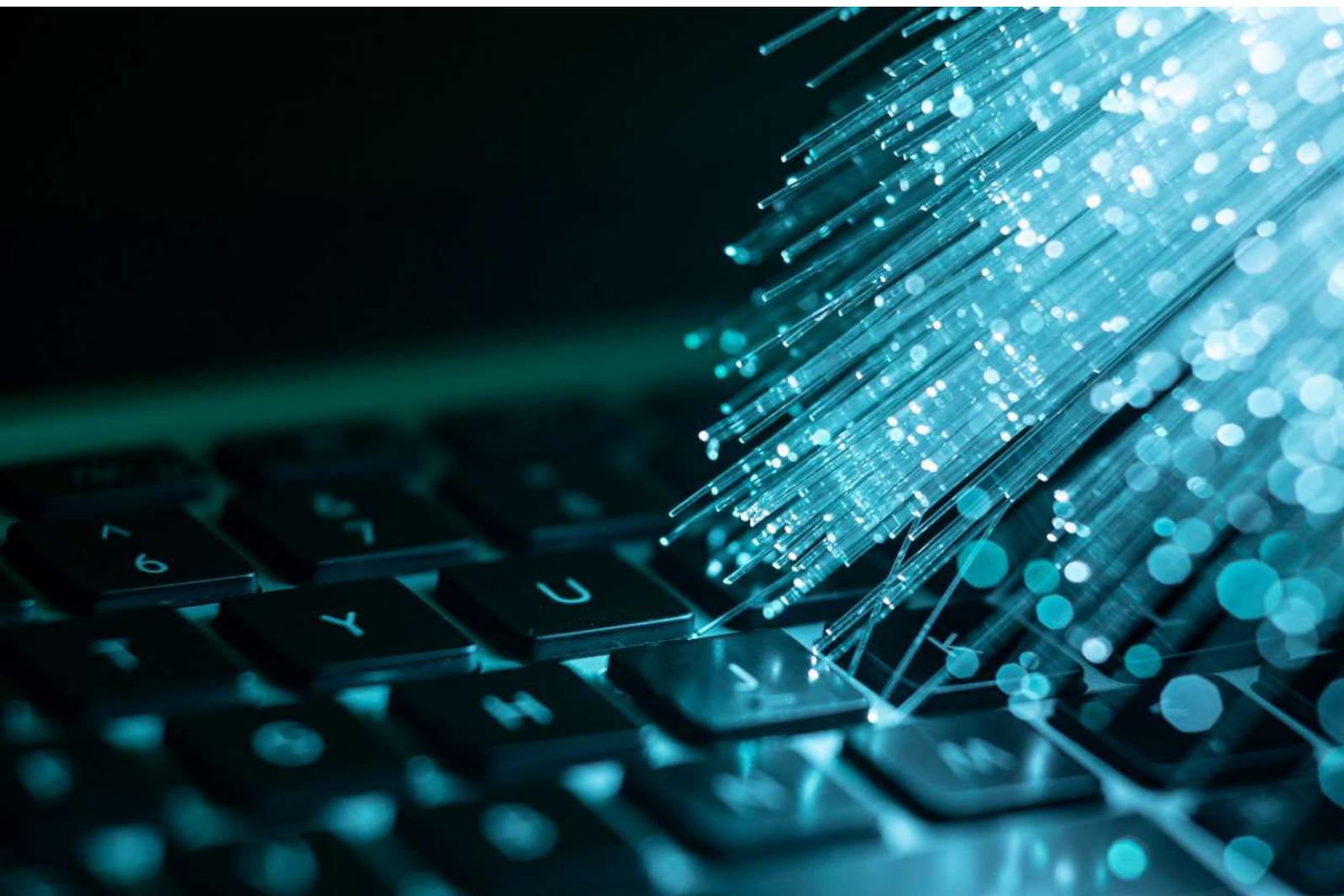
A Jaunt é uma empresa de realidade virtual fundada na Califórnia. Ela permite aos espectadores experimentarem a realidade virtual cinematográfica. Jens Christensen, cofundador da Jaunt, afirma que:

Pensamos que notícias são simplesmente fantásticas para a realidade virtual. Você tem uma visão realmente objetiva, não é filtrada e sente que está realmente lá. Você tem uma noção de perspectiva e escala que não consegue quando está olhando para um retângulo de 16 por 9 (a proporção da tela de cinema). (EVANGELISTA, 2015 apud MAYHEW, 2016 p. 27).

As pessoas possuem níveis variáveis de apatia em relação a eventos globais, com base na distância entre si e o assunto. Contar histórias em 360° pode ser uma maneira de reverter a tendência de distanciamento, os espectadores podem aprender mais sobre o assunto por estarem imersos em um meio diferente.

2.3 Efeitos desejados e indesejados

Atividades tradicionais de tomada de perspectiva imersiva, em que o espectador é convidado a se passar pela outra pessoa, podem efetivamente aumentar a empatia que um observador sente por um alvo específico. A experiência pode ser particularmente poderosa, já que muitas vezes levam a um sentimento de inclusão, reduzem preconceitos e atenuam estereótipos negativos, podem ajudar a criar e manter laços sociais, facilitar interações e motivar comportamento de ajuda, de acordo com pesquisa realizada por um grupo de pesquisadores



dos departamentos de Comunicação e Psicologia, da Universidade de Stanford, na Califórnia em 2018. (HERRERA *et al.*, 2018)

Além disso, a tomada de perspectiva não apenas aumenta a empatia por pessoas específicas, mas pode aumentar a empatia em relação a grupos estigmatizados inteiros. Por exemplo, Batson *et al.* (1997) descobriram que os participantes que foram solicitados a assumir a perspectiva de um membro de um grupo estigmatizado (por exemplo, pessoas diagnosticadas com AIDS ou criminosos condenados) relataram atitudes mais positivas não apenas em relação a um membro específico desse grupo, mas em direção ao grupo estigmatizado como um todo.

Apesar da ampla maioria das pesquisas apontar que a tomada de perspectiva leva a resultados positivos, sob certas circunstâncias ela pode promover o efeito contrário, pois pode aumentar os estereótipos associados a uma pessoa ou a um grupo (TARRANT; CALITRI; WESTON, 2012; SKORINKO; SINCLAIR, 2013; HODGES; WEGNER; LCKES, 1997), podendo até levar pessoas a culparem um grupo estigmatizado por sua própria situação e gerar efeitos diversos dos desejados com a experiência (BATSON, 1997).

2.4 Mídia e empatia

A tentativa de promover empatia pode ter vários objetivos e usar de diversas mídias: impressa, narrativas interativas, videogames, vídeos bidimensionais e, mais recentemente os vídeos em 360°. Esta última mostrou-se ser uma das imersões mais ricas, de acordo com os trabalhos citados e com maior potencial de promover a desejada empatia.

O conteúdo experimentado através da mídia em 360° fornece informações adicionais ao participante sobre uma situação específica, em vez de confiar apenas na imaginação do usuário. Isso pode ser particularmente útil quando os participantes tiveram contato limitado ou têm muito pouca informação sobre uma meta social (GELBACH *et al.* 2015).

Diferentes tipos de mídia também variam em seu nível de imersão e interatividade. Imersão é uma medida objetiva e descritiva da medida em que um meio particular é capaz de criar e sustentar uma “ilusão vívida da realidade” (HERRERA *et al.*, 2018). A interatividade refere-se à medida em que os usuários podem exercer influência sobre o conteúdo do ambiente mediado em tempo real. O fato de que diferentes tipos de mídia, como livros, TVs, computadores e realidade virtual se encaixam em diferentes níveis de imersão e interatividade, pode ajudar a explicar por que pesquisas existentes misturaram resultados relativos a tarefas mediadas de tomada de perspectiva e seus efeitos sobre empatia e comportamentos pró-sociais.

2.5 Empatia para mobilização

A realidade virtual, mediada por óculos apropriados, é uma experiência digital tridimensional com o objetivo de emular a realidade, e este estudo supõe que as histórias contadas

em realidade virtual visam envolver pessoalmente o destinatário, amplificando o envolvimento emocional. O conteúdo de mídia e a forma de mídia contribuem para o engajamento, por exemplo, na criação de proximidade com o outro e evocando o interesse pessoal do destinatário.

A presença simulada estimula a liberdade do destinatário de gerar seu próprio entendimento dos eventos. O que é filmado nesse formato é importante por causa de como os espectadores podem reagir ao material com base em seu campo de experiência. Experimentar um vídeo em realidade virtual é uma experiência intensamente individual e, como resultado, as experiências pessoais dos espectadores mudam a forma como o conteúdo é entendido.

O que se pretende, com a pesquisa de campo, é medir a capacidade de mobilização que a empatia, provocada pela realidade virtual, pode estimular. Medir, através de comparações sistemáticas os impactos empáticos provocados pela mesma informação, experimentadas por mídias supostamente mais e menos imersivas.

3. METODOLOGIA

Trata-se de um estudo de natureza qualitativa e quantitativa, pois “interpreta as informações quantitativas por meio de símbolos numéricos e os dados qualitativos mediante a observação, a interação participativa e a interpretação do discurso dos sujeitos (semântica)” (KNECHTEL, 2014).

Em relação ao objetivo da pesquisa ela é exploratória, uma vez que este tipo de pesquisa busca proporcionar maior familiaridade com o problema, visando torná-lo mais explícito ou a construir hipóteses (GIL, 2007). Quanto aos procedimentos utilizados, pode-se dizer que são experimentais, visto que esse tipo de pesquisa consiste em determinar um objeto de estudo, selecionar as variáveis que são capazes de influenciá-lo, definir as formas de controle e de observação dos efeitos que a variável produz no objeto (GIL, 2007).

Foram realizados dois experimentos com a finalidade de comparar os efeitos das diferentes intervenções a partir da perspectiva da empatia e dos comportamentos pró-sociais a curto e médio prazos, em que o primeiro estudo consiste em uma narrativa e o segundo é um vídeo em 360°. A coleta de dados para averiguar os efeitos das intervenções foi feita através de questionários.

Neste trabalho, escolheu-se como alvo social pessoas que se encontram em situação de rua, pois é um grupo à margem da sociedade com o qual muitas pessoas lutam para simpatizar. Apesar da presença de alguns programas sociais, poucas políticas públicas são realizadas para sanar o problema. Entende-se também que apresentar uma narrativa ou um vídeo em 360° de uma pessoa em situação de rua atinge uma maior parte da população, uma vez que, a partir de determinadas circunstâncias, qualquer um pode tornar-se uma pessoa em situação de rua.

Os experimentos foram realizados na Universidade Federal de Alfenas-MG, no campus de Varginha-MG, e os sujeitos de pesquisa foram os discentes da instituição. A amostra do experimento é de 80 (oitenta) alunos, em que 50% participou do estudo 1 (narrativa) e os outros 50% do estudo 2 (vídeo 360°). O experimento foi realizado nos dias 25 de outubro e 14 de novembro do ano de 2019. No primeiro dia realizou-se o experimento com 57 (cinquenta e sete) discentes, sendo que 33 (trinta e três) alunos participaram do estudo 1 e 24 (vinte e quatro) do estudo 2, e no segundo dia, obteve-se 7 (sete) alunos realizando o estudo 1 e 16 (dezesesseis) o estudo 2, totalizando 23 (vinte e três) discentes. A amostra obtida é de 44 mulheres e 36 homens, e as idades variaram de 18 a 38 anos.

A divulgação do experimento foi realizada através das mídias sociais e também foi enviado um e-mail para todos os discentes do campus de Varginha-MG informando local, dia e horário e comunicando que os discentes que participassem dos estudos receberiam a quantia de R\$ 10,00 (dez reais).

Os alunos foram escolhidos aleatoriamente para participar dos estudos. No primeiro momento os alunos de ambos os experimentos responderam ao questionário pré-intervenção (Apêndice 1), que incluía questões demográficas, como gênero e idade, o e-mail para contato e trazia a indicação do estudo do qual ele iria participar; afirmações para mensurar o Índice de Reatividade Interpessoal (IRI), que foram respondidas através de uma escala de Likert de 5 pontos, variando de “não me descrevem bem” para “me descreve bem”; e afirmações das crenças sobre a escala de empatia, que foram respondidas através de uma escala de Likert de 7 pontos (1 = discordo fortemente, 7 = concordo fortemente).

É importante ressaltar que o Índice de Reatividade (IRI) é uma escala de 28 (vinte e oito) itens e que possui 4 (quatro) subescalas, que buscou medir as diferenças individuais em relação à empatia, avaliando a tendência do participante para adotar o ponto de vista do outro e a angústia pessoal em situações delicadas que envolvem outras pessoas. O questionário possui itens como “Eu às vezes tenho dificuldade de ver as coisas a partir do ponto de vista de outra pessoa” e “Muitas vezes eu fico sensibilizado e preocupado com pessoas menos afortunadas que eu”. No questionário utilizado neste estudo foi excluída a subescala fantasia, pois essa mede especificamente a capacidade de incorporar (através da imaginação) um personagem fictício e sentir o que eles estão se sentindo, algo que era menos relevante para a nossa intervenção.

As crenças sobre a escala de empatia possuem 12 (doze) itens que tinham como finalidade medir o grau em que as pessoas acreditam que a empatia é algo que pode ser controlado. Os itens da amostra incluem “as pessoas sempre podem alterar a quantidade de empatia que sente por outra pessoa” e “quando uma pessoa sente empatia por alguém, ele não consegue parar de sentir empatia, mesmo que ela queira parar.”

Após responderem aos questionários, todos os oitenta alunos passaram pelo experimento da narrativa ou do

vídeo 360°. Destaca-se que, em ambos os estudos, os participantes realizaram a tarefa de imaginar que estavam conversando com uma pessoa em situação de rua, sendo assim, pode-se dizer que as intervenções foram qualitativamente semelhantes, diferenciando apenas a forma que foram usadas para aplicar a intervenção.

No estudo que foi utilizado o vídeo 360°, os discentes puderam experimentar a sensação de estarem conversando com uma pessoa em situação de rua. Os participantes inicialmente estavam sentados durante a experiência, mas podiam se levantar, mover a cabeça e inclinar seu torso para olhar para onde eles queriam. Os discentes nessa condição utilizaram óculos de realidade virtual, um smartphone e um fone de ouvido para a imersão.

O vídeo em 360° que foi apresentado neste estudo foi elaborado pelos autores, usando uma câmera capaz de realizar filmagens com essa tecnologia. Foi realizada uma entrevista com um roteiro semiestruturado com o Luiz Fernando Souza, que se encontra em situação de rua no município de Varginha - MG, na praça José Resende de Paiva, no dia 6 de setembro de 2019, que autorizou a utilização das filmagens para fins de pesquisa (Anexo 1). A entrevista contemplava questões tais como: as maiores dificuldades de estar em situação de rua, se havia programas por parte do setor público para as pessoas que se encontram nessa situação, se há ajuda por parte dos moradores e comerciantes do entorno, entre outras.

Quanto ao estudo que se refere à narrativa, os participantes tiveram acesso a um texto descritivo em terceira pessoa, dos eventos e situações a partir do que ocorreu na entrevista descrita acima. Os alunos nessa condição foram induzidos a imaginar o quão difícil era se encontrar em situação de rua, e para isso procurou-se trazer uma maior riqueza de detalhes para essa experiência.

Com a finalidade de buscar mensurar a empatia a curto prazo, foi solicitado que todos os alunos respondessem a um questionário após as intervenções (Apêndice 2). No segundo questionário, os participantes foram convidados no primeiro momento a avaliar a extensão com a qual eles sentiram algumas emoções ao longo da intervenção, tais como: emoção, desconforto, angústia, confusão, incômodo, entre outras. Foi utilizada uma escala de Likert de 7 pontos (1 = não sentiu, 7 = sentiu extremamente). Logo após, foi solicitado que alunos respondessem sete itens referentes à escala de atitudes em relação às pessoas que se encontram em situação de rua; para mensurar, foi utilizada a escala de 9 pontos de Likert, em que 1 é “discordo fortemente” e 9 representa “concordo fortemente”. Os itens da amostra incluem “A maioria das pessoas em situação de rua poderiam ter evitado chegar a essa situação” e “Nossos representantes não fazem o suficiente para ajudar as pessoas em situação de rua”. Posteriormente foi indagado aos participantes se eles tinham interesse em doar integralmente ou parcialmente a quantia de R\$ 10,00 (dez reais) recebida para participar da pesquisa, para o Centro de Referência Especializada para população em situação de rua em Varginha/MG (Centro POP). No questionário as opções eram doar os R\$ 10,00 (dez reais), doar R\$ 5,00 (cinco reais) ou não doar nenhuma

quantia. Destaca-se que o tempo estimado de ambos os experimentos foi de aproximadamente 15 (quinze) minutos.

Ressalta-se que o recurso inicial para a realização do experimento foi de R\$ 1.100,00 (um mil e cem reais) e os recursos que foram devolvidos pelos participantes, após as intervenções, foram doados integralmente para o Centro de Referência Especializada para população em situação de rua de Varginha-MG.

A análise de dados foi realizada por meio de um software de tabulação e tratamento de dados chamado Statistical Package for the Social Sciences (SPSS). Os dados de todos os 80 questionários foram inseridos no software e foram aplicados e executados procedimentos baseados em técnicas estatísticas visando relacionar a percepção dos alunos quanto à perspectiva da empatia a curto e médio prazo em relação às pessoas que se encontram em situação de rua, relacionando com o estudo que o aluno participou. Os resultados dessas análises serão apresentados na próxima sessão.

Para verificar o efeito dos estudos a médio prazo, foi enviado um e-mail (Apêndice 3) a todos os participantes dez dias após a intervenção, solicitando que os mesmos escrevessem uma mensagem à Câmara de Vereadores de Varginha expressando o que eles pensavam sobre as pessoas em situação de rua, a sua opinião sobre o assunto e quais iniciativas públicas deveriam ocorrer para solucionar ou amenizar o problema. A finalidade dessa tarefa é analisar a extensão do engajamento cívico de cada

participante sobre a questão das pessoas em situação de rua, e realizar a análise de conteúdo para melhor compreender os estados emocionais dos participantes através de sua expressão escrita.

Por fim, é importante salientar que os alunos que participaram do experimento autorizaram a utilização dos seus dados para fins de pesquisa.

4. RESULTADOS

4.1 Análise do questionário pré-intervenção

A partir da análise do primeiro questionário (pré-intervenção), tem-se inicialmente como resultado as questões relacionadas aos dados demográficos dos participantes dos dois estudos. Tem-se como primeiro resultado que mais de 70% da amostra pertencem à faixa etária de 18 e 25 anos. Quanto ao gênero, 55% se identificou como pertencente ao gênero feminino e 45% ao masculino.

Em relação às 21 (vinte e uma) afirmativas que buscavam para mensurar o Índice de Reatividade Interpessoal (IRI) a partir das diferenças individuais em relação à empatia, notou-se um certo equilíbrio nas respostas dos discentes dos estudos 1 e 2, como se pode observar nos quadros abaixo:

Quadro 1 – Questionário pré-intervenção: Escala IRI - Afirmativa 4

Em situações de emergência, sinto-me apreensivo e desinquieto. * Estudo que o aluno participou Tabulação cruzada

% do Total

		Estudo que o aluno participou		Total
		Estudo 1 - Narrativa	Estudo 2 - Vídeo 360°	
Em situações de emergência, sinto-me apreensivo e desinquieto.	A	3,8%		3,8%
	B	8,8%	10,0%	18,8%
	C	6,3%	5,0%	11,3%
	D	13,8%	6,3%	20,0%
	E	17,5%	28,8%	46,3%
Total		50,0%	50,0%	100,0%

Fonte: elaboração própria.

Quadro 2 – Questionário pré-intervenção: Escala IRI - Afirmativa 21

Antes de criticar alguém, eu tento imaginar como me sentiria se estivesse no seu lugar. * Estudo que o aluno participou **Tabulação cruzada**

% do Total

		Estudo que o aluno participou		Total
		Estudo 1 - Narrativa	Estudo 2 - Vídeo 360°	
Antes de criticar alguém, eu tento imaginar como me sentiria se estivesse no seu lugar.	A	2,5%		2,5%
	B	5,0%	2,5%	7,5%
	C	7,5%	1,3%	8,8%
	D	23,8%	31,3%	55,0%
	E	11,3%	15,0%	26,3%
Total		50,0%	50,0%	100,0%

Fonte: elaboração própria.

4.2 Análise do resultado no curto prazo

A partir do questionário pós-intervenção pode-se avaliar os efeitos das intervenções a curto prazo. Com esse questionário pode-se observar os diferentes níveis de emoção e atitudes dos participantes dos estudos 1 e 2 em relação às pessoas em situação de rua. Inicialmente foi indagado aos partícipes com qual intensidade foram experimentadas 8 (oito) emoções, sendo elas: sensibilizado, emocionado, desconfortável, angustiado,

confuso, solidário, incomodado e humanitário. Notou-se que a maior parte dos estudantes que participaram do estudo 2 sentiram com maior intensidade as emoções “positivas”, como: sensibilizado, solidário, humanitário e emocionado. Para mensurar as emoções foi utilizada uma escala de Likert de 7 (sete) pontos, onde 1 (um) representa que não sentiu a emoção e 7 (sete) representa que sentiu extremamente essa emoção. Pode-se observar esse resultado a partir dos Quadros abaixo:

Quadro 3 – Questionário pós-intervenção - Emoção: Sensibilizado

% do Total

		Estudo que o aluno participou		Total
		Estudo 1 - Narrativa	Estudo 2 - Vídeo 360°	
Sensibilizado	1	1,3%		1,3%
	2	1,3%		1,3%
	3	3,8%		3,8%
	4	7,5%		7,5%
	5	5,0%	3,8%	8,8%
	6	10,0%	11,3%	21,3%
	7	21,3%	35,0%	56,3%
Total		50,0%	50,0%	100,0%

Fonte: elaboração própria.

Quadro 4 – Questionário pós-intervenção - Emoção: Humanitário

% do Total

	Estudo que o aluno participou		Total
	Estudo 1 - Narrativa	Estudo 2 - Vídeo 360°	
Humanitário 1	1,3%		1,3%
2	2,5%		2,5%
3	6,3%	2,5%	8,8%
4	10,0%	1,3%	11,3%
5	5,0%	5,0%	10,0%
6	15,0%	10,0%	25,0%
7	10,0%	31,3%	41,3%
Total	50,0%	50,0%	100,0%

Fonte: elaboração própria.

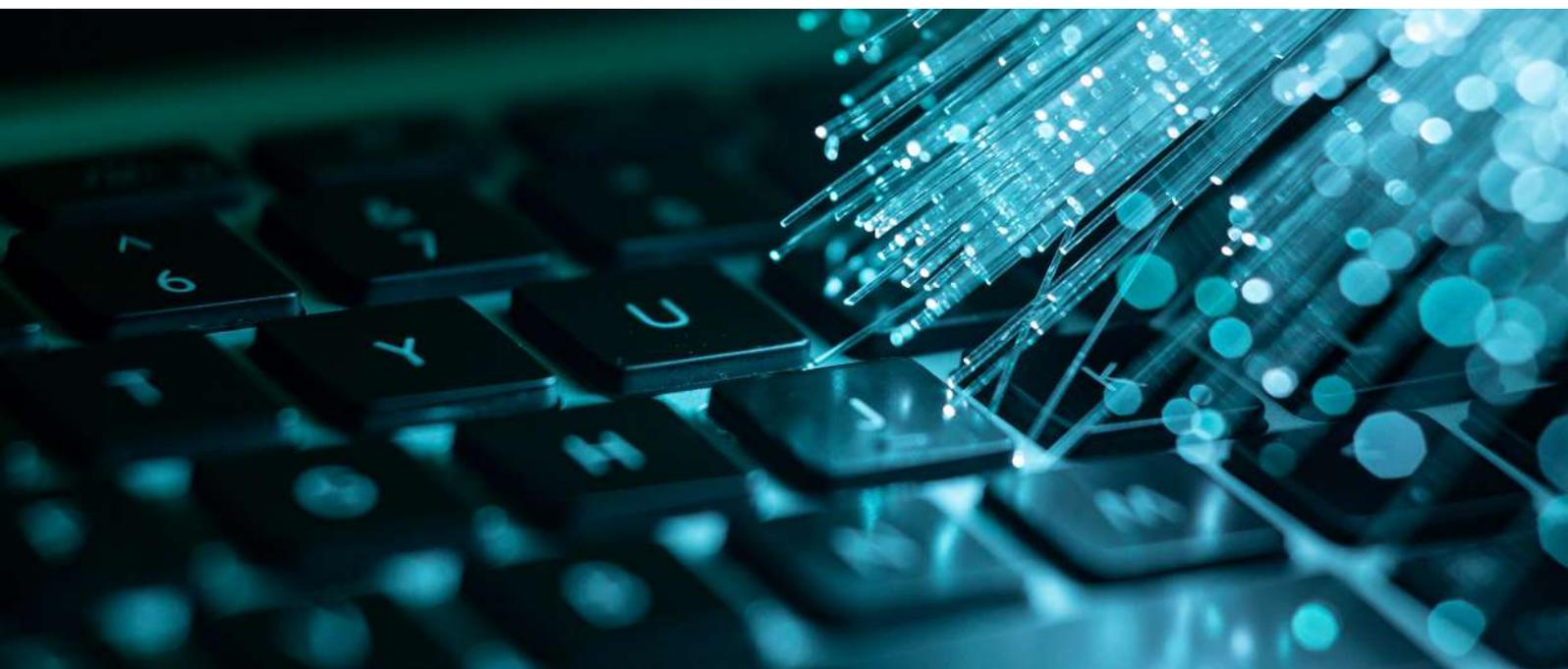
Observa-se através do Quadro 3, que demonstra o quanto os alunos experimentaram a emoção “sensibilizados”, que 21,3% dos alunos que experimentaram essa emoção fortemente haviam participado do estudo 1; no entanto, quando se analisa o percentual dos alunos pertencentes ao estudo 2 que sentiram essa emoção da mesma forma, esse percentual sobe para 35%. Vale destacar que somente os alunos que participaram do estudo 1 manifestaram que não se sentiram sensibilizados ou pouco sensibilizados (3,8%).

No Quadro 4, percebe-se que a maior parte dos indivíduos da amostra tiveram o sentimento humanitário mais intenso, totalizando o percentual de 41,3%. Nota-se que 31,3% do percentual total é representado por alunos que participaram da experiência com o vídeo 360°, e que apenas 10% eram do experimento da narrativa.

A partir desse resultado, percebe-se que o estudo 2 teve um maior impacto nos alunos do que o estudo 1. Inicialmente

pode-se dizer que a imersão proporcionada a partir da realidade virtual, através do vídeo em 360° gerou uma maior aproximação com o nosso alvo social (pessoas em situação de rua) e que os alunos que tiveram a experiência da realidade virtual apresentaram um grau maior de empatia, a partir das escalas utilizadas para mensuração.

Posteriormente, foi solicitado aos alunos que respondessem o quanto concordavam com determinadas afirmações sobre as pessoas que vivem em situação de rua. Para mensurar foi utilizada novamente a escala de 9 pontos de Likert, onde 1 é “discordo fortemente” e 9 representa “concordo fortemente”. Como resposta às 6 (seis) afirmativas, obteve-se um resultado expressivo em relação aos participantes que tiveram contato com a realidade virtual, através do vídeo em 360°. Estes demonstraram um maior nível de empatia a curto prazo do que os alunos que tiveram acesso apenas à narrativa escrita. Pode-se observar tais resultados nos Quadros seguintes:



Quadro 5 – Questionário pós-intervenção - Atitude em relação às pessoas em situação de rua

% do Total

		Estudo que o aluno participou		Total
		Estudo 1 - Narrativa	Estudo 2 - Vídeo 360°	
Quanto você pessoalmente se preocupa com as pessoas em situação de rua?	1	1,3%		1,3%
	2		1,3%	1,3%
	3	6,3%	1,3%	7,5%
	4	5,0%	2,5%	7,5%
	5	3,8%	5,0%	8,8%
	6	10,0%	7,5%	17,5%
	7	11,3%	5,0%	16,3%
	8	7,5%	22,5%	30,0%
	9	5,0%	5,0%	10,0%
Total		50,0%	50,0%	100,0%

Fonte: elaboração própria.

Quadro 6 – Questionário pós-intervenção - Atitude em relação às pessoas em situação de rua

% do Total

		Estudo que o aluno participou		Total
		Estudo 1 - Narrativa	Estudo 2 - Vídeo 360°	
Nossa sociedade não faz o suficiente para ajudar as pessoas em situação de rua.	1	2,5%		2,5%
	2	2,5%		2,5%
	4	2,5%		2,5%
	5	1,3%	2,5%	3,8%
	6	3,8%	1,3%	5,0%
	7	12,5%	5,0%	17,5%
	8	6,3%	6,3%	12,5%
	9	18,8%	35,0%	53,8%
Total		50,0%	50,0%	100,0%

Fonte: elaboração própria.

Quando questionado aos alunos que participaram do experimento o quanto eles se preocupavam com as pessoas em situação de rua (Quadro 5), verificou-se que 40% dos indivíduos da amostra assinalaram as opções 8 e 9 que demonstram uma maior preocupação. Quando se analisa qual o percentual de alunos destes 40% eram pertencentes ao estudo 2, verifica-se que 27,5% corresponde aos alunos que vivenciaram o vídeo 360° e apenas 12,5% eram pertencentes ao experimento da narrativa.

Em relação à afirmação “nossa sociedade não faz o suficiente para ajudar as pessoas em situação de rua” (Quadro 6), nota-se que 53,8% da amostra concorda fortemente com essa afirmativa, no entanto, percebe-se uma disparidade quando

se compara o quanto desse percentual é representado pelo estudo 1 e pelo estudo 2, em que apenas 18,8% dos 53,8% concordam fortemente com a assertiva.

Ressalta-se que em todas as emoções sentidas após a intervenção e em relação às seis assertivas que representavam atitudes em relação aos moradores de rua, observa-se que os participantes que vivenciaram o vídeo 360° mostraram, em sua maioria, de acordo com as escalas pré-definidas, um maior nível de empatia em relação aos participantes que tiveram contato apenas com a narrativa.

Supõe-se, a partir desses resultados, que os indivíduos após participarem do estudo 2 tiveram uma maior

predisposição em colaborar em ações em prol das pessoas que estão em situação de rua, assim como em buscar novas soluções para amenizar e solucionar as adversidades que essas pessoas enfrentam cotidianamente.

Para concluir o questionário pós intervenção, foi perguntado aos participantes de ambos os experimentos

se eles tinham o interesse em realizar a doação do incentivo que eles receberam no início da intervenção, dando as opções a eles de doarem de forma integral (R\$10,00), parcial (R\$5,00) ou não doarem nenhuma quantia. Como resposta a esse questionamento, obteve-se o seguinte resultado:

Quadro 7 - Questionário pós-intervenção - Doação para Centro POP

Doação para o Centro POP * Estudo que o aluno participou Tabulação cruzada

			Estudo que o aluno participou		Total
			Estudo 1 - Narrativa	Estudo 2 - Vídeo 360°	
Doação para o Centro POP	Doou R\$ 10,00	Contagem	23	32	55
		% do Total	28,8%	40,0%	68,8%
	Doou R\$ 5,00	Contagem	3	3	6
		% do Total	3,8%	3,8%	7,5%
	Não doou	Contagem	14	5	19
		% do Total	17,5%	6,3%	23,8%
Total	Contagem	40	40	80	
	% do Total	50,0%	50,0%	100,0%	

Fonte: elaboração própria.

Pode-se perceber a partir desse Quadro que, de forma geral, tivemos um resultado positivo em ambos os estudos, pois 68,8% da nossa amostra doou integralmente o valor anteriormente recebido para participar do estudo. Quando se analisa o quanto esse percentual representa os participantes do estudo 1 e 2, verifica-se que, assim como os demais resultados obtidos através das assertivas, os alunos que vivenciaram o vídeo 360° tiveram um percentual superior (40%) em relação aos alunos que experimentaram apenas a narrativa.

Quando examinados os percentuais de alunos que doaram parcialmente o valor, nota-se que esse resultado é inconclusivo, pois ambos os estudos apresentaram o mesmo valor percentual.

Em relação aos que não doaram, percebe-se uma discrepância, visto que, dos 23,8% dos alunos que não doaram, 17,5% é representado por alunos do estudo 1.

Destaca-se, apenas como curiosidade, que ocorreram duas situações atípicas durante os experimentos: a primeira refere-se a um participante que selecionou que iria doar integralmente o valor, mas não doou. Neste caso, identificou-se o seu questionário e o dado não foi computado como doação real. Na segunda situação, um participante tinha uma dívida de R\$10,00 com outro participante. Assim que o primeiro recebeu o valor entregue pelos pesquisadores, quitou sua dívida com o segundo e este fez então uma doação dupla, de R\$20,00 para o POP.

Como dito anteriormente, o recurso disponibilizado para este estudo foi de R\$1.100,00 (mil e cem reais) e, após este

experimento, restou um saldo de R\$ 580,00 (quinhentos e oitenta reais), que foram doados pelos participantes, na forma de itens de higiene pessoal para o Centro POP, decisão tomada em comum acordo entre os pesquisadores e os gestores da organização.

A partir da premissa inicial de avaliar os efeitos da realidade virtual a curto prazo, pode-se dizer que se obteve um resultado satisfatório, visto que foi possível verificar uma diferença relevante entre os níveis de empatia após a realização das intervenções. Os dados apontaram que os participantes do estudo 2 passaram a sentir mais empatia pelas pessoas em situação de rua comparados aos participantes que tiveram acesso somente à narrativa.

4.3 Análise do resultado no médio prazo

Para realizar a análise do efeitos da realidade virtual em relação à empatia a médio prazo, foi enviado um e-mail a todos os participantes do experimento, solicitando que os mesmos escrevessem uma mensagem para a Câmara de Vereadores de Varginha expressando o que eles pensam sobre a situação das pessoas em situação de rua, sua opinião sobre o assunto e perguntando quais as iniciativas públicas devem ocorrer para solucionar ou amenizar o problema. Como citado anteriormente, os e-mails foram enviados 10 (dez) dias após ocorrer a intervenção e eles tiveram cinco dias úteis como prazo para enviar a mensagem.

Nessa tarefa houve uma baixa adesão, visto que foram obtidas apenas 4 (quatro) mensagens, e essas poucas foram somente de alunos que haviam participado do experimento com o vídeo 360°.

Considera-se, no âmbito dessa análise, que o fato de apenas participantes do estudo 2 enviarem as mensagens é algo relevante para analisar os efeitos da realidade virtual a médio prazo. Inicialmente, pode-se dizer que os alunos que enviaram as mensagens consideram que o setor público deve fazer mais para as pessoas que se encontram em situação de rua.

“Sugiro que elaborem uma Política Pública voltada para amenizar essa situação. Essas pessoas precisam sim de comida, de casa, mas também precisam que sua autoestima e confiança em si mesmo sejam recuperadas. Uma PP voltada para conseguir emprego para essas pessoas seria o ideal. Agradeço a atenção” (Aluno/a participante do estudo 2 – trecho da resposta enviada).

“Percebo que não há aparentemente, serviço qualificado de Albergues e abrigos noturnos para pessoas que se encontram vivendo na rua. Acredito que é extremamente importante tal serviço tendo em vista as condições climáticas de Minas Gerais, sendo até mesmo no verão, como nos tempos de agora, faz frio a noite” (Aluno/a participante do estudo 2 – trecho da resposta enviada).

5. CONSIDERAÇÕES FINAIS

A partir do experimento realizado, percebe-se que a realidade virtual pode gerar empatia maior no curto prazo, se compararmos os seus efeitos no longo prazo. Com base nos dados obtidos na intervenção, nota-se que os participantes que tiveram contato com o vídeo 360° apresentaram maior empatia no curto prazo com as pessoas em situação de rua e demonstraram uma maior proximidade do que em relação aos participantes que tiveram contato com a narrativa escrita. Entretanto, quando se analisa o resultado dos efeitos no longo prazo, percebe-se que, apesar do retorno dos quatro participantes do estudo 2, não se obteve um resultado satisfatório para se presumir que o experimento do vídeo 360° teve um resultado relevante no longo prazo, já que o resultado corresponde a apenas 5% do total.

Como limitação do estudo, pode-se dizer que as atitudes dos participantes em relação às pessoas em situação de rua não foram medidas antes da intervenção. Mesmo os participantes sendo selecionados de forma aleatória, é possível que os participantes já tivessem opiniões bem formadas sobre as pessoas em situação de rua, que os pesquisadores não conseguem identificar. Estudos futuros devem medir preconceitos e atitudes pré-existentes em relação às pessoas em situação de rua, a fim de avaliar com mais precisão o efeito dos diferentes tipos de intervenções e a empatia.

A partir deste trabalho, conclui-se que a realidade virtual pode ser utilizada como uma ferramenta capaz de auxiliar na mobilização, através da empatia, e que esta é capaz de promover maior aproximação de uma realidade que,

muitas vezes, é distante dos cidadãos e dos tomadores de decisão sobre políticas sociais e assistenciais.

Sugere-se que esse experimento seja aplicado aos representantes dos Poderes Executivo e Legislativo em todas as esferas, que tomam suas decisões sobre projetos, programas e políticas públicas pautadas, muitas vezes, apenas nas narrativas escritas. Tais representantes podem estar distantes da realidade percebida pela população afetada pelos variados problemas.

As sugestão não é que eles deixem de conhecer os problemas e os projetos *in loco*, mas sim que a realidade virtual seja capaz de auxiliá-los na tomada de decisões, fazendo dessa uma nova ferramenta para a gestão pública.

Como estudo posterior a este, pretende-se replicar o experimento com os vereadores da Câmara Municipal de Varginha, fazendo os devidos ajustes na metodologia para atender a esta amostra específica. Acredita-se que a aplicação desse estudo para o público em questão será de grande contribuição, uma vez que se constatou que a realidade virtual auxilia no engajamento das pessoas em causas coletivas, na avaliação de cenários e pode apoiá-los no processo de tomada de decisão, demonstrando o potencial do uso dessa tecnologia para esta e outras causas relevantes para a sociedade.

6. REFERÊNCIAS BIBLIOGRÁFICAS

BATSON, C; POLYCARPOU, M; HARMON-JONES, E; IMHOFF, H; MITCHENER, E; BEDNAR, L. *et al.* Empathy and attitudes: can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of personality and social psychology*. 1997 Jan; v. 72, n. 1, p. 105. PMID:9008376. 1997.

BHATIA, G. Virtual reality news is becoming a reality in many newsrooms. Available in: <https://www.poynter.org/news/virtual-reality-news-becoming-reality-many-newsrooms>. Poynter, September 30, 2015.

BOHRER, C. The JJIE *Virtual World Journalism Project*: experimenting with virtual worlds as an emerging journalism platform. Retrieved from <http://digitalcommons.kennesaw.edu/undergradsymposiumksu/2016/Posters/2/> Acesso em 2016

BOLTER, J; GRUSIN, R. *Remediation: understanding new media*. Cambridge: The MIT Press, 2000, 293 p. 2000.

BRAGA, M. Realidade virtual e educação. *Revista de Biologia e Ciências da Terra*, Paraíba, v. 1, n. 1, jan./jun. 2001.

DE la PEÑA, N; WEIL, P; LLOBERA, J; GIANNOPOULOS, E; POMÉS, A; SPALANG, B. *et al.* Immersive journalism: immersive virtual reality for the first-person experience of news. *Presence: Teleoperators and virtual environments*. Aug 1; v.19, n. 4, p. 291–301, 2010.

- DIGI-CAPITAL. *Augmented/Virtual Reality Report Q3 2019*, San Francisco: Digi-Capital, 2019.
- DOYLE, P.; GELMAN, M.; GILL, S. *Viewing the future? Virtual reality in journalism*. Knight Foundation, 2016. Disponível em: http://www.knightfoundation.org/media/uploads/publication_pdfs/VR_report_web.pdf Acesso em: 20 mar. 2016
- DOYLE, P.; GELMAN, M.; GILL, S. *Viewing the Future? Virtual Reality in Journalism*. Available in: <https://knightfoundation.org/reports/vrjournalism>. Knight Foundation, March 13, 2016.
- GEHLBACH, H; MARIETTA, G; KING, A; KARUTZ, C; BAILENSEN, J; DEDE, C. Many ways to walk a mile in another's moccasins: type of social perspective taking and its effect on negotiation outcomes. *Computers in Human Behavior*. 2015 Nov, v. 30, n. 52, p. 523–32. 2015.
- GIL, A. C. *Como elaborar projetos de pesquisa*. 4. ed. São Paulo: Atlas, 2007.
- GILLATH, O; MCCALL, C; SHAVER, P; BLASCOVICH, J. What can virtual reality teach us about prosocial tendencies in real and virtual environments? *Media Psychol.* v. 11, n. 2, p. 259–282, 2008.
- HASLER, B; HIRSHBERGER, G; SHANI-SHERMAN, T; FRIEDMAN, D. Virtual peacemakers: mimicry increases empathy in simulated contact with virtual outgroup members. *Cyberpsychol. Behav. Soc. Netw.* v. 17, n. 12, p. 766–771, 2014.
- HERRERA, F; BAILENSEN, J; WEISZ, E; OGLE, E; ZAKI, J. Building long-term empathy: a large-scale comparison of traditional and virtual reality perspective-taking. *PLoS ONE* 13(10): e0204494. 2018.
- HODGES, S; WEGNER, D; ICKES, W. Automatic and controlled empathy. *Empathic accuracy*. New York: Guilford p. 311-339. 1997.
- HOFFMAN, M. L. *Empathy and moral development: implications for caring and justice*. Cambridge University Press; 2001
- IPSOS MORI. *Os perigos da percepção*. 2017. Disponível em: <https://opinioao.estadao.com.br/noticias/geral,os-perigos-da-percepcao>. Acesso em: 12 nov. 2019
- KNECHTEL, Maria do Rosário. *Metodologia da pesquisa em educação: uma abordagem teórico-prática dialogada*. Curitiba: Intersaberes, 2014
- LONGHI, Raquel Ritter. Narrativas imersivas no webjornalismo: entre interfaces e realidade Virtual. In: SBPJor, 2016, Palhoça. *14º Encontro Nacional de Pesquisadores em Jornalismo*. Palhoça: Unisul, 2016. v. 1. p. 108. Disponível em: <http://migre.me/vX7yH>. Acesso em: 23 nov. 2016
- MAYHEW, F. *Guardian creates in-house virtual reality team*. Press Gazette. London, England. Retrieved from: http://www.pressgazette.co.uk/guardian-creates-inhousevirtualrealityteam/?utm_source=Daily+Lab+email+list&utm_campaign=c42cf57242-dailylabemail3&utm_medium=email&utm_term=0_d68264fd5e-c42cf57242-396038393 Acesso em: out. 2016
- MCLUHAN, M. *Os meios de comunicação como extensões do homem*. 5 ed. São Paulo: Cultrix, 1979.
- MCROBERTS, J. Are we there yet? Media content and sense of presence in nonfiction virtual reality. In: *Studies in Documentary Film*. p. 1-18. 2017.
- MILK, C. *How virtual reality can create the ultimate empathy machine* [Internet]. TED: Ideas worth spreading. https://www.ted.com/talks/chris_milk_how_virtual_reality_can_create_the_ultimate_empathy_machine Acesso em: dez 2015
- MORAN, J. Ensino e aprendizagem inovadores com tecnologias audiovisuais e telemáticas. In: MORAN, J. M.; MASETTO, M. T.; BEHRENS, M. A. *Novas tecnologias e mediação pedagógica*. Campinas: Papirus, 2007.
- PIERCE, J; KILDUFF, G; GALINSKY, A; SIVANATHAN, N. From glue to gasoline: how competition turns perspective takers unethical. *Psychological science*. Oct; v. 24, n. 10, p. 1986–94. pmid:23955353. 2013.
- SHIN, D; BIOCICA, F. Exploring immersive experience in journalism. *New Media Soc.* v. 19, n. 11, p. 1–24, 2017.
- SHIN, D. Empathy and embodied experience in virtual environment: to what extent can virtual reality stimulate empathy and embodied experience? *Comput. Hum. Behav.* v. 78, p. 64–73, 2018
- SHIN, D. The role of affordance in the experience of virtual reality learning: technological and affective affordances in virtual reality. *Telematics Inform.* v. 34, n. 8, p. 1826–1836, 2017.
- SKORINKO, J; SINCLAIR, S. Perspective taking can increase stereotyping: the role of apparent stereotype confirmation. *Journal of Experimental Social Psychology*. 2013 Jan 1;49(1):10–8. 2013
- SUNDAR, S; KANG, J; OPREAN, D. Being there in the midst of the story: how immersive journalism affects our perceptions and cognitions. *Cyberpsychol. Behav. Soc. Netw.* v. 20, n. 11, p. 672–682, 2017
- TARRANT, M; CALITRI, R; WESTON, D. Social identification structures the effects of perspective taking. **Psychological science**. 2012 Sep.;v. 23, n. 9, p. 973–8. pmid:22851441. 2012.

APÊNDICE 1

Experimento Mobiliza (1º Questionário - estudo 2)

*Obrigatório

1. Endereço de e-mail *

2. Qual estudo está participando? *

Marcar apenas uma oval.

Estudo 1 - Narrativa

Estudo 2 - Vídeo 360°

3. Idade *

4. Gênero *

Marcar apenas uma oval.

Masculino

Feminino

Prefiro não dizer

Escala IRI

Utilizando a escala abaixo, indique o seu nível de concordância com cada uma das seguintes afirmações. Não há respostas certas ou erradas.

As declarações a seguir buscam analisar seus pensamentos e sentimentos em uma variedade de situações. Para cada item, você deve responder o quão bem ele o descreve, escolhendo a letra apropriada na escala para cada pergunta: A, B, C, D ou E. Sendo que: A = não descreve me bem e E = me descreve muito bem.

Favor ler cada item com atenção antes de responder. Responder da forma mais honesta possível. Obrigado.

5. Marcar apenas uma oval por linha.

	A	B	C	D	E
Muitas vezes eu fico sensibilizado e preocupado com pessoas menos afortunadas que eu.	<input type="radio"/>				
Às vezes tenho dificuldade de ver determinada situação do ponto de vista de outra pessoa.	<input type="radio"/>				
Às vezes eu não me sensibilizo com outras pessoas quando eles estão tendo problemas.	<input type="radio"/>				
Em situações de emergência, sinto-me apreensivo e desinquieto.	<input type="radio"/>				
Eu tento observar todos os pontos de vista antes de tomar uma decisão.	<input type="radio"/>				
Quando vejo alguém sendo aproveitado, eu tento protegê-lo.	<input type="radio"/>				
Às vezes me sinto impotente quando eu estou no meio de uma situação emocional.	<input type="radio"/>				
Às vezes eu busco entender meus amigos, imaginando determinadas situações a partir de sua perspectiva.	<input type="radio"/>				
Quando vejo alguém se machucar, tenho a tendência em me manter calmo.	<input type="radio"/>				
As perdas, problemas de outras pessoas não costumam me perturbar muito.	<input type="radio"/>				
Se eu tenho certeza que estou certo sobre alguma coisa, não perco tempo ouvindo o argumento de outras pessoas.	<input type="radio"/>				
Estar em uma situação emocional me assusta.	<input type="radio"/>				
Quando vejo alguém sendo tratado injustamente, às vezes eu não sinto comovido.	<input type="radio"/>				
Eu sou geralmente muito eficaz para lidar com situações de emergência.	<input type="radio"/>				
Frequentemente me sinto tocado por coisas que eu vejo acontecer.	<input type="radio"/>				
Eu acredito que existem dois lados para cada questão e tento olhar para os dois.	<input type="radio"/>				
Eu me descreveria como uma pessoa muito sentimental.	<input type="radio"/>				
Eu perco o controle durante as emergências.	<input type="radio"/>				
Quando estou chateado com alguém, geralmente tento "me colocar no lugar dele" por um tempo.	<input type="radio"/>				
Quando vejo alguém que precisa desesperadamente de ajuda em uma emergência, eu fico comovido.	<input type="radio"/>				
Antes de criticar alguém, eu tento imaginar como me sentiria se estivesse no seu lugar.	<input type="radio"/>				

Crenças sobre a escala de empatia

6. Utilizando a escala abaixo, indique o seu acordo com cada uma das seguintes afirmações. Não há respostas certas ou erradas. *

Marcar apenas uma oval por linha.

	1	2	3	4	5	6	7
As pessoas podem ajustar a quantidade de empatia que estão sentindo em qualquer situação.	<input type="radio"/>						
As pessoas podem controlar quantidade de empatia sentem por outras pessoas no momento.	<input type="radio"/>						
Em qualquer situação, as pessoas têm a capacidade de "aumentar" a quantidade de empatia que sentem por alguém.	<input type="radio"/>						
Em qualquer situação, as pessoas têm a capacidade de não sentir empatia por alguém.	<input type="radio"/>						
Quando uma pessoa sente empatia por alguém, não consegue parar de sentir empatia, mesmo que queira parar.	<input type="radio"/>						
Quando uma pessoa não sente empatia por alguém, ela não consegue parar de sentir, mesmo que queira sentir empatia por essa pessoa.	<input type="radio"/>						
As pessoas podem sempre alterar a quantidade de empatia que sentem por outra pessoa.	<input type="radio"/>						
As pessoas não podem mudar quanta empatia elas tendem a sentir pelos outros. Algumas pessoas sentem mais empatias do que as outras, e não há muito o que fazer para mudar.	<input type="radio"/>						
Não importa que alguém seja, eles podem sempre mudar a empatia de uma pessoa.	<input type="radio"/>						
Se uma pessoa é empática ou não, está ligado a sua personalidade. Não pode se pode mudar se uma pessoa é empática ou não.	<input type="radio"/>						
Qualquer um pode mudar e fazer outra pessoa sentir empatia.	<input type="radio"/>						

7. Autorizo a utilização das minhas respostas para fins de pesquisa, podendo essas serem divulgadas e utilizadas em trabalhos acadêmicos, congressos e outros eventos. *

Marcar apenas uma oval.

- Sim
 Não

Obrigado por participar do estudo!

APÊNDICE 2

Experimento Mobiliza (2º Questionário)

*Obrigatório

1. Endereço de e-mail *

2. Qual estudo participou? *

Marcar apenas uma oval.

- Estudo 1 - Narrativa
 Estudo 2 - Vídeo 360º

Empatia e angústia pessoal

Na escala: 1 significa menor relevância e 7 significa maior relevância.

3. Por favor, indicar a extensão em que você experimentou as seguintes emoções. *

Marcar apenas uma oval por linha.

	1	2	3	4	5	6	7
Sensibilizado	<input type="radio"/>						
Emocionado	<input type="radio"/>						
Desconfortável	<input type="radio"/>						
Confuso	<input type="radio"/>						
Angustiado	<input type="radio"/>						
Solidário	<input type="radio"/>						
Incomodado	<input type="radio"/>						
Humanitário	<input type="radio"/>						

4. Atitudes em relação as pessoas que vivem em situação de rua *

Marcar apenas uma oval por linha.

	1	2	3	4	5	6	7	8	9
Para a maioria das pessoas em situação de rua, é sua própria culpa que elas estejam desabrigadas.	<input type="radio"/>								
A maioria das pessoas em situação de rua poderiam ter evitado essa condição.	<input type="radio"/>								
Nossa sociedade não faz o suficiente para ajudar as pessoas em situação de rua.	<input type="radio"/>								
A nossa sociedade deve fazer mais para proteger o bem-estar das pessoas em situação de rua.	<input type="radio"/>								
Quanto você pessoalmente se preocupa com as pessoas em situação de rua?	<input type="radio"/>								
Comparado com outros problemas sociais que enfrentamos hoje (por exemplo, crime, educação, drogas, AIDS, proteção ambiental, entre outros), como você classificaria a importância de ajudar as pessoas em situação de rua?	<input type="radio"/>								

5. Você tem interesse em doar integralmente ou parcialmente o dinheiro recebido para participar deste experimento para o Centro de Referência Especializada para população em situação de rua em Varginha/MG (Centro POP)? *

Marcar apenas uma oval.

- Sim, gostaria de doar o valor de R\$ 10,00 (dez reais).
- Sim, gostaria de doar o valor de R\$ 5,00 (cinco reais).
- Não pretendo doar nenhuma quantia.

6. Autorizo a utilização das minhas respostas para fins de pesquisa, podendo essas serem divulgadas e utilizadas em trabalhos acadêmicos, congressos e outros eventos. *

Marcar apenas uma oval.

- Sim
- Não

Obrigado por participar do estudo!

APÊNDICE 3

Olá,

Primeiramente, agradeço por ter participado da primeira etapa do experimento, respondendo os questionários referentes à pesquisa do Mobiliza 360. Peço que leia o texto abaixo e me retorne até o dia 27/11 (quarta-feira).

Foram 80 pessoas que participaram e R\$ 580,00 reais serão doados para o Centro de Referência Especializada para população em situação de rua em Varginha/MG (Centro POP).

Uma das melhores maneiras de ajudar a mudar a nossa comunidade é através da comunicação com os nossos representantes eleitos. Os vereadores têm interesse em saber o que os cidadãos sentem sobre determinadas questões, especialmente quando essas questões envolvem decisões tomadas por eles.

Uma das formas de fazer com que os representantes do legislativo saibam o que a população deseja ou os princípios que ela considera importantes, é através da comunicação direta.

Pedimos que você escreva uma mensagem para à Câmara de Vereadores de Varginha expressando o que você pensa sobre a situação das pessoas em situação de rua, qual é a sua opinião sobre o assunto e quais iniciativas públicas devem ocorrer para solucionar ou amenizar o problema?

Peço que você me envie essa mensagem para o e-mail mobiliza360.pesquisa@gmail.com que iremos encaminhar para à Câmara de Vereadores de Varginha.

Caso você queira receber o resultado da pesquisa, favor nos informar através deste e-mail.

Agradeço desde já.

Atenciosamente,

Hélio Lemes Costa

Letícia Amaral.

ANEXO 1

TERMO DE AUTORIZAÇÃO DE USO DE IMAGEM

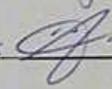
Eu, Luiz Fernando Souza,
brasileiro (nacionalidade), solteiro (estado civil), portador (a)
do RG :° MG-4965 648, inscrito(a) no CPF sob o n.° 002.788.376-03,

AUTORIZO o uso de minha imagem, constante na filmagem de Leticia Almeida Amaral, com o fim específico do seu trabalho de conclusão de curso sem qualquer ônus e em caráter definitivo.

A presente autorização abrangendo o uso da minha imagem na filmagem acima mencionada é concedida à Leticia Almeida Amaral a título gratuito, abrangendo inclusive a licença a terceiros, de forma direta ou indireta, e a inserção em materiais para toda e qualquer finalidade, seja para uso comercial, de publicidade, jornalístico, editorial, didático e outros que existam ou venham a existir no futuro, para veiculação/distribuição em território nacional e internacional, por prazo indeterminado.

Por esta ser a expressão da minha vontade, declaro que autorizo o uso acima descrito, sem que nada haja a ser reclamado a título de direitos conexos à imagem ora autorizada ou a qualquer outro, e assino a presente autorização em 02 (duas) vias de igual teor e forma.

Local e data: Varginha, 06 de setembro de 2013

Assinatura: x  _____

Telefone para contato: () _____