

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE POS-GRADUAÇÃO
EM CIÊNCIA E TECNOLOGIA DA COMPUTAÇÃO**

**Uma Abordagem Baseada em Jogos Sérios Para o Ensino de
Qualidade de Dados**

Bárbara Flávia Marques Vieira

Itajubá, dezembro de 2016

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE POS-GRADUAÇÃO
EM CIÊNCIA E TECNOLOGIA DA COMPUTAÇÃO**

Bárbara Flávia Marques Vieira

**Uma Abordagem Baseada em Jogos Sérios Para o Ensino de
Qualidade de Dados**

**Dissertação submetida ao Programa de Pós-Graduação em
Ciência e Tecnologia da Computação como parte dos
requisitos para obtenção do Título de Mestre em Ciência e
Tecnologia da Computação.**

Orientador: Prof^a. Dr^a. Melise M. Veiga de Paula

Dezembro de 2016
Itajubá - MG

Agradecimentos

Este trabalho é resultado de contribuições diretas e indiretas de diversas pessoas que estiveram presentes em minha vida.

Agradeço imensamente à minha orientadora Melise M. Veiga de Paula, pela disponibilidade, apoio e, principalmente, paciência e perseverança.

Ao professor Rodrigo Seabra pela disponibilidade e opiniões muito bem-vindas.

Aos meus pais, Lúcia e José Anézio, por serem essa base forte presente em todos os momentos.

Ao meu marido Rafael Braga, por ter tomado para si fardos que eu não conseguia mais carregar.

À minha irmã, Lara, pelo incentivo e auxílio em vários momentos (quem diria).

Ao meu amigo Ederson Vasconcelos, por estar sempre a disposição para ajudar e também pela preocupação constante com este trabalho.

Aos amigos do fórum The Best – RPGMaker, que me aceitaram na rodinha há 13 anos e são, provavelmente, os grandes responsáveis pelo meu início na computação. Obrigada por estarem presentes até hoje.

Aos amigos que mesmo com a minha ausência buscavam estar sempre presentes.

À CAPES, CNPq, pelo apoio financeiro.

Resumo

Atualmente, há uma crescente preocupação com a qualidade dos produtos de informação que são gerados. Contudo, a falta de conhecimento explícito que seja efetivamente prático na área de análise e gestão da qualidade de dados pode dificultar a utilização das abordagens propostas na literatura. Neste trabalho é apresentada a compilação de ferramentas e técnicas estruturada em uma metodologia para análise e gestão de qualidade de dados, tornando-a menos subjetiva para implantação. O aprendizado de uma metodologia para a gestão da qualidade de dados exige prática. Considerando que a aplicação com fins educacionais de uma metodologia nem sempre é possível, a proposta para facilitar este aprendizado é a elaboração de um jogo digital, *Data Quality Simulator*, sendo o objetivo deste trabalho é o desenvolvimento deste jogo. Para isso foram realizados estudos na literatura sobre jogos de aprendizagem e características presentes em jogos. Neste trabalho também se encontra uma pesquisa de opinião sobre jogos para definição de requisitos. As metas educacionais e os requisitos estabelecidos foram mapeados em relação a estrutura do *Data Quality Simulator*. Além das estratégias e técnicas propostas na metodologia e do desenvolvimento do jogo, neste trabalho, é discutida a abordagem a ser utilizada para a avaliação da proposta. É apresentado resultados obtidos com a avaliação de aprendizagem, satisfação do usuário e usabilidade do jogo desenvolvido, sendo que estes positivos para o fortalecimento do uso de jogos para o ensino e prática do conteúdo.

Palavras-chave: metodologia de qualidade de dados, jogos sérios, avaliação de jogos sérios

Abstract

There is a growing concern in the system context about the quality of the information. However, the lack of explicit knowledge that is effectively practical in this area may make it difficult to use the approaches proposed in the literature. This work presents a compilation of tools and techniques structured in a methodology for analysis and management of data quality. Learning a methodology for data quality management requires practice. Considering that, the application for educational purposes of a methodology is not always possible; the proposal to facilitate this learning is the elaboration of a digital game, Data Quality Simulator, being the development of this game the objective of this work. Studies about learning games and features present in games were accomplished. In this work there is also an opinion survey on games for definition of requirements. The educational goals and requirements were mapped on the structure of the Data Quality Simulator. In addition to the strategies and techniques proposed in the methodology and game development, in this work, the approach to be used for the evaluation of the proposal is discussed. This work will discuss the results obtained with the evaluation of learning, user satisfaction and usability of the game developed, and these positive for the strengthening of the use of games for teaching and practice of content.

Keywords: data quality methodology, serious games, evaluation of serious games

SUMÁRIO

1. INTRODUÇÃO	7
1.1. Motivação e Justificativa	7
1.2. Objetivo	9
1.3. Organização da dissertação	10
2. FUNDAMENTAÇÃO TEÓRICA	11
2.1. Qualidade de Dados	11
2.1.1. Definições e problemas de Qualidade de Dados	12
2.1.2. Dimensões de qualidade de dados	14
2.1.3. Metodologias para gestão da qualidade de dados	17
2.1.4. <i>Total Data Quality Management (TDQM)</i>	20
2.2. Aprendizado baseado em jogos	21
2.2.1. Objetivos educacionais e elementos motivacionais	22
2.2.2. Taxonomia de Bloom	23
2.2.3. Taxonomia de motivações intrínsecas	27
3. METODOLOGIA DE AVALIAÇÃO DE QUALIDADE DE DADOS	30
3.1. Preparação	30
3.2. Definição	32
3.2.1. Obter conhecimento sobre o domínio	32
3.2.2. Identificar e selecionar os Temas Estratégicos	33
3.2.3. Detalhar os temas estratégicos definidos	35
3.2.4. Definir as dimensões e métricas que serão analisadas	47
3.3. Avaliação, Análise e Ação	49
4. DATA QUALITY SIMULATOR	51
4.1. Pesquisa de opinião em relação a jogos	51
4.2. Descrição do jogo	61
4.3. Mapeamento utilizado para desenvolvimento do jogo	68
5. METODOLOGIA DE AVALIAÇÃO	78
5.1. Procedimento de avaliação executado	78
5.2. Análise dos resultados obtidos	85

6. CONSIDERAÇÕES FINAIS	98
REFERÊNCIAS BIBLIOGRÁFICAS	100

1. Introdução

1.1. Motivação e Justificativa

A facilidade na obtenção e armazenamento de dados tem resultado em bases de dados extensas e de diversas aplicações e modelos. Porém, uma quantidade massiva de informações disponível não implica, necessariamente, no fator qualidade. Por esse motivo, diferentes tipos de projetos e organizações que utilizam como matéria prima a informação concentram grande parte de seus esforços na limpeza e preparação dos dados (CAO et al., 2010).

Em alguns contextos, os sistemas de suporte a tomada de decisão não utilizam todo o conteúdo do banco de dados por não terem absoluta garantia da validade dos dados (CHENGALUR-SMITH; BALLOU; PAZER, 1999), já que informações incorretas ou inconsistentes tendem a distorcer análises de resultados e causar prejuízo por conta de ações que se baseiam nelas. Assim, a tomada de decisão em um sistema de informação, bem como as definições de regulamentações e regras de uma organização e a visão dos que utilizam os dados são impactadas diretamente pela qualidade dos dados (BAGCHI; BAI; KALAGNANAM, 2006).

Para melhorar a produtividade e prover decisões mais corretas é importante dar ênfase na seleção dos dados (JAQUEIRA, 2009). Porém, examinar apenas o armazenamento dos dados ou sua fonte, apesar de necessário, não é suficiente, pois não são encontrados apenas erros nas características inerentes ao dado. Todo o processo de obtenção do dado até chegar à aplicação proposta também é influenciado por fatores modificadores que podem denegrir a qualidade da informação final.

Os problemas na qualidade dos dados estão se tornando cada vez mais evidentes e as complicações que se manifestam a partir da falta de qualidade das informações podem ter uma propagação alta e de um alto custo (WANG et al., 2008). Logo, uma análise adequada não envolve apenas identificar que a qualidade observada é insuficiente, mas também encontrar a origem dos possíveis problemas e buscar soluções.

Para Batini e Scannapieco (2006), o estudo sobre qualidade de dados não é apenas uma questão tecnológica, mas também de gerenciamento. Logo, pesquisadores de diferentes áreas como, por exemplo, contabilidade, estatística e engenharia de produção têm estudado a qualidade de dados em suas respectivas áreas (WANG et al., 2008).

Apesar do reconhecimento da importância de uma gestão adequada para o controle de qualidade de dados, ainda existem poucos estudos sobre as políticas apropriadas e sobre os

métodos de criação de bons sistemas que alcancem esse objetivo, já que o estudo da qualidade de dados e informação é relativamente recente. Já em 1995, Wang et al. (1995) evidenciaram este fato afirmando que os modelos comumente utilizados para armazenamento de dados apresentam como foco a captura da própria informação, ou seja, quais atributos ou entidades deveriam estar presentes no modelo para que o mundo real seja representado digitalmente, mas não abordavam de forma objetiva se os valores dos dados em si apresentavam boa qualidade.

Este cenário estimulou a elaboração de metodologias e políticas para garantia de qualidade de dados. Porém, para que as organizações optem por aderir à utilização de alguma abordagem para qualidade de dados é necessário um maior entendimento dos processos e técnicas propostas, assim como informações sobre custo, recursos e papéis envolvidos.

Na literatura, são encontradas diferentes abordagens para avaliação e gerenciamento dos dados (WANG; STOREY; FIRTH, 1995) (BATINI; SCANNAPIECO, 2006) (BASTOS et al., 2010) (BAGCHI; BAI; KALAGNANAM, 2006). Um aspecto comum e considerado bastante efetivo é denominado dimensão, que pode ser entendido como as características de informações usadas para definir, mensurar e gerenciar dados (MCGILVRAY, 2008).

A acurácia, completude, consistência e atualidade são exemplos de dimensões comumente citadas e apresentadas como características importantes para os consumidores de informação (BLAKE; MANGIAMELI, 2008) e podem ser utilizadas para que se tenha uma avaliação de quão bem classificada uma informação pode ser nas diversas perspectivas possíveis (CHENGALUR-SMITH; BALLOU; PAZER, 1999). Com esse tipo de medição, se torna mais simples visualizar onde estão os maiores focos e origens de problemas.

A análise da literatura permite observar que, para algumas dimensões descritas, os trabalhos encontrados relacionados ao assunto apresentam diferentes abordagens de medição e definição de métricas que variam em função do domínio considerado, o que torna a seleção de quais dimensões deverão ser usadas e como as usar uma escolha bastante segmentada de um trabalho para outro.

Diante da diversidade e da complexidade das diversas abordagens de avaliação e gerenciamento de qualidade de dados, é importante analisar as metodologias propostas e como as atividades definidas devem ser realizadas na prática.

Tendo em vista esse problema de definição, as estratégias computacionais podem auxiliar os usuários a identificar e tratar problemas, possibilitando que as atividades voltadas para a qualidade dos dados sejam realizadas com menos custo e de forma mais rápida. Além disso, utilizar uma boa estratégia como base para a condução do processo de garantia de

qualidade de dados pode reduzir o tempo de estudo sobre o que é relevante e servir de guia para a criação de métodos de avaliação (BARROS, 2009).

Para a implantação de uma estratégia é necessário que as pessoas envolvidas tenham um bom entendimento de como deve ser sua aplicação. Porém, o aprendizado deste tipo de conteúdo requer prática e, como a prática por meio de casos reais nem sempre é possível, a proposta desta pesquisa é facilitar a absorção de conhecimento com base em simulações, o que motivou a elaboração de um jogo denominado *Data Quality Simulator*.

Utilizar jogos digitais no ensino não é novidade. Desde os anos 80, com o surgimento de computadores pessoais e videogames, pesquisadores já visionavam seus impactos na área de educação, sendo que há mais de dez anos trabalhos relacionados aparecem com frequência crescente em conferências e periódicos (MORENO-GER et al., 2014). Em comparação com métodos mais comuns de ensino, vários estudos mostram que um jogo consegue oferecer uma experiência bastante elevada, podendo ser considerada uma potencial ferramenta de aprendizagem (CARVALHO et al., 2015) (HENRIQUE et al., 2015).

Além da possibilidade de aprender conteúdos de forma mais prática, outro argumento para a criação de um jogo deriva de seu potencial motivacional. A vontade de aprender é considerada uma motivação intrínseca; ou seja, a prática de exercício de aprendizagem é o que aumenta a vontade de aprender. Porém, muitas vezes, o ambiente de aprendizagem é um ambiente confinado e com caminho e currículo muito bem definido, situação que falha em preservar a energia natural que conservaria o aprendizado espontâneo (MALONE; LEPPER, 1987).

Neste caminho, enquanto ambientes escolares lutam para manter seus alunos envolvidos, videogames têm prosperado como uma indústria de entretenimento focada em criar produtos que capturam a atenção de jogadores de todas as idades e origens (MORENO-GER et al., 2014), sendo a indústria dos jogos um dos setores que mais cresce (AHMAD; RAHIM; ARSHAD, 2014).

1.2. Objetivo

Considerando que o aprendizado de uma metodologia de análise de qualidade de dados requer prática e que a prática por meio de casos reais nem sempre é possível, a proposta é facilitar a absorção de conhecimento com base em simulações, mais especificamente um jogo. O objetivo dessa pesquisa, então, consiste em desenvolver um jogo para aprendizado e treinamento e, assim, aumentar as possibilidades de criação do conhecimento nesta área,

considerando tanto cenários acadêmicos quanto profissionais. Para que este objetivo seja alcançado é necessário a realização de objetivos específicos.

O primeiro objetivo específico consiste em analisar as metodologias existentes, em especial a TDQM (*Total Data Quality Management*), sob a perspectiva tecnológica e desenvolver uma metodologia de análise e gestão de qualidade de dados que possa ser considerada de forma mais prática durante a execução de um projeto real. O objetivo seguinte compreende o desenvolvimento do jogo proposto, iniciado a partir do estudo sobre aprendizagem baseada em jogos, seguido pela definição do conteúdo que será abordado e as características de jogos que serão aplicadas e a elaboração da estrutura do jogo e sua implementação. Por fim, o terceiro objetivo específico consiste em executar uma avaliação experimental com usuários para validação do jogo elaborado.

1.3. Organização da dissertação

Este trabalho está estruturado da seguinte maneira: o capítulo 2 apresenta a fundamentação teórica com seções que versam sobre qualidade de dados e aprendizado baseado em jogos. O capítulo 3 descreve o fluxo da metodologia desenvolvida para gestão de qualidade de dados e suas etapas; o capítulo 4 apresenta o jogo desenvolvido *Data Quality Simulator*; o capítulo 5 apresenta a metodologia de avaliação considerada como método de validação e, por fim, o capítulo 6 expõe a conclusão.

2. Fundamentação teórica

Para que os objetivos específicos sejam alcançados é necessário ter conhecimentos fundamentais sobre os temas relacionados. Este capítulo apresenta conceitos, técnicas e abordagens relacionadas a qualidade de dados que serão base para o desenvolvimento da metodologia de análise de qualidade de dados proposta. Também será explanado, neste capítulo, fundamentos teóricos sobre aprendizado em jogo, taxonomias de aprendizado e conceitos relacionados, utilizados como base para o desenvolvimento do jogo *Data Quality Simulator*.

2.1. Qualidade de Dados

A capacidade de armazenar uma quantidade massiva de dados não significa que os dados podem ser considerados como informações úteis. É necessário que os dados armazenados sejam analisados para que a tomada de decisão não seja realizada com base em conhecimentos incorretos (CHENGALUR-SMITH; BALLOU; PAZER, 1999). As organizações necessitam de informações confiáveis para que possam realizar o trabalho de promover as metas e objetivos de seus negócios.

Em geral, as organizações, ao focar a administração da produção de informação numa visão computacional, concentram-se nos aspectos de software e hardware envolvidos no processo, mas não nas propriedades da informação, como o seu processo de produção e ciclo de vida.

Muitas vezes, as dificuldades são provenientes da negligência em relação aos dados utilizados e podem ser confundidas com problemas de outra natureza, o que pode causar gastos desnecessários para a organização. Por exemplo, se um erro resultante de falha humana for atribuído como problema da máquina utilizada, o tratamento realizado aos dados para solucionar a inconformidade pode ser inadequado uma vez que o problema a ser resolvido é diferente do problema real (TORELLI, 2005).

A análise da qualidade da informação é uma estratégia que pode ser usada para isolar quais dados devem ser utilizados e quais devem ser desprezados. Barros (2009) afirma que a diminuição de informações ruins diminui consideravelmente os problemas e suas consequências.

Ao visualizar o processo de transformação dos dados como um processo de produção de um produto manufaturado, pode se identificar uma estrutura de gestão que auxilia na administração desse tipo de recurso. Além disso, é possível que resultados positivos

significativos sejam alcançados a partir da utilização de métodos que facilitem a gestão dos dados considerando o aspecto de qualidade (BARROS, 2009).

2.1.1. Definições e problemas de Qualidade de Dados

Definir o conceito de Qualidade de Dados contribui para delimitar o que deve ser considerado durante uma avaliação e quais metas e métodos serão empregados. Autores utilizam definições com essência de qualidade semelhante, porém com pontuais aspectos de métrica diferentes.

Torelli (2005, p.20) afirma que “convencionalmente, QI (Qualidade de Informação) tem sido definida como quão íntegra a informação está”. Bastos *et al* (2010, p. 412) acreditam que a qualidade dos dados é “a medida de adequação a finalidade. Assim, para um produto ou serviço ter qualidade, as suas características devem satisfazer as expectativas.”

Já para Wang e Strong (1996), um dado com alta qualidade deve ser intrinsecamente bom, ou seja, contextualmente apropriado à tarefa, representado claramente e acessível aos consumidores. Os autores sintetizam a definição de qualidade dos dados como “os dados que estão aptos para o uso (*fit for use*) pelo consumidor de dados” (WANG; STRONG, 1996, p.6).

Outras definições podem ser vistas em:

A qualidade da informação é o grau em que as informações e os dados podem ser uma fonte confiável para qualquer e/ou todos os usos necessários. É ter o conjunto certo de informações corretas, no momento certo, no lugar certo, para as pessoas corretas usarem na tomada de decisões, para executar negócios, servir clientes, e alcançar os objetivos da empresa (MCGILVRAY, 2008, p.5).

A qualidade dos dados é definida como a medida em que uma base de dados representa com precisão as propriedades essenciais da aplicação a que se destina, e possui três propriedades distintas: a confiabilidade dos dados, integridade lógica ou semântica e a integridade física (a exatidão dos detalhes da implementação) (WANG; STOREY; FIRTH, 1995, p. 629)

Do ponto de vista do sistema de controle de *feedback* é realmente fácil de se definir qualidade de dados como a medida de concordância entre os dados apresentados visualmente pelo sistema de informação e os mesmos dados no mundo real (WANG *et al.*, 2008, p. 439)

Os autores concordam que a qualidade de dados indica quão positiva são as características dos dados para o usuário da informação; pontos como confiabilidade, precisão e utilidade são apontados. Apesar de, essencialmente, as definições serem parecidas, qualidade de dados aborda um número considerável de aspectos que podem ter maior ou menor valor dependendo da perspectiva e domínio que será apresentado, o que interfere na escolha da definição mais adequada.

Difícilmente, em um sistema de informação os dados serão de extrema precisão, ao ponto de estarem totalmente de acordo com a realidade (WANG *et al.*, 2008). Porém, o objetivo do estudo da qualidade dos dados no sistema é que o mesmo esteja consistente o suficiente para o contexto e o domínio que serão utilizados, assim como tenham a validade, acurácia, tempo, entre outras características, em grau aceitável. Um problema de qualidade de dados é identificado quando existe a não conformidade do grau de qualidade real do dado e o grau necessário para o uso.

Os problemas da qualidade de dados podem ser originados por diferentes aspectos que podem variar desde questões humanas, fatores relacionados ao sistema, ou ao processo, não sendo restrito a alguma etapa específica. McGilvray (2008) apresenta alguns indicadores que alertam para a possibilidade de existirem alguns problemas:

- Necessidade de atividades corretivas;
- Repetição de trabalho;
- Ordens repetidas;
- Acúmulo de reclamações de clientes.

A avaliação dos problemas deve ser dividida entre as várias perspectivas do processo. As equipes de tecnologia de informação são responsáveis pela qualidade do sistema que armazena e modifica os dados, mas não são totalmente responsáveis pelo conteúdo, sendo também função do time de gestão e gerência do processo (MCGILVRAY, 2008).

Sabendo da existência de problemas, para solucioná-los é necessário descobrir quais são as suas causas. O termo fontes de erros é utilizado para tarefas que ocorrem durante operações com os dados que podem gerar problemas (BAGCHI; KALAGNANAM, 2006). Também são consideradas fontes de erros quando operações que deveriam ser realizadas não ocorrem ou ocorrem de forma ilegítima.

As circunstâncias que modificam os dados e podem ocasionar uma baixa qualidade podem ser vistas em alguns fatores específicos. Wang *et al.* (2008) afirmam que, frequentemente, os estudos acerca desses fatores dentro de linhas de produção gerais apontam seis elementos: Homem, Máquina, Material, Métodos, Medição e Ambiente; também conhecidos como 5M1E (*Man, Machine, Material, Methods, Measurement e Environment*). Tendo esses fatores que interagem com a produção em vista, é importante mensurar o grau de modificação que cada um deles realiza sobre os dados, e quais trajetórias continuam a ter a influência dessas modificações.

2.1.2. Dimensões de qualidade de dados

Para encontrar incoerências e erros nos dados e apontar o quanto o problema é aceitável ou não, é importante observar quais características dos dados melhor se associam ao problema. Um dos primeiros passos na avaliação de qualidade de dados é definir quais aspectos serão utilizados como métrica (BATINI; SCANNAPIECO, 2006).

Não existe uma forma geral para especificar os critérios a serem usados em uma avaliação. Autores variados utilizam diferentes dimensões, em quantidade e grau de profundidade diversas, para definir metodologias utilizadas na mensuração da qualidade de dados. Em alguns casos, utilizam de categorias para classificar as dimensões escolhidas.

McGilvray (2008, p.7) utiliza como definição para dimensões “aspectos da qualidade de dados que você pode utilizar para mensurar e gerenciar seus dados e qualidade da informação”. Wang e Strong (1996, p.6) aceitam dimensão de qualidade de dados como “um conjunto de atributos de qualidade de dados que representam um único aspecto ou construção da qualidade de dados”.

Torelli (2006) acredita que as dimensões sejam divisões realizadas na qualidade da informação. Algumas dimensões são subjetivas, ou seja, dependem de um contexto, e outras objetivas, independem do contexto.

Para se escolher quais dimensões são relevantes é preciso conhecer o contexto para o qual os dados estão sendo processados e para que finalidade serão utilizados. O aspecto mais apropriado de informação a respeito da qualidade de dados que se deve levar em conta durante as avaliações depende da natureza do processo de decisão em que a informação será manipulada (CHENGALUR-SMITH; BALLOU; PAZER, 1999).

No gerenciamento de dados existem várias atribuições entre as entidades que manipulam e utilizam os dados. As entidades podem ser fornecedores, consumidores, gerentes ou processadores das informações do sistema (WANG *et al.*, 2008). Um sistema pode ter como fornecedor de dados a mesma entidade que utiliza as informações após o processamento, pode ter entidades separadas, ou mesmo mais consumidores finais para os mesmos dados processados. Esses fatores devem ser levados em conta ao se definir quais dimensões são mais importantes de serem avaliadas.

Os critérios utilizados para se identificar as melhores dimensões de qualidade de dados são divididos em três bases: a própria qualidade de dados, o sistema de informação específico associado à satisfação do cliente e a contabilidade (WANG; STOREY; FIRTH, 1995).

Em relação ao critério qualidade de dados, a estratégia é ter como base a teoria de qualidade de dados para apontar as diferentes dimensões e utilizar as características próprias de cada uma para criar taxas de erros associadas. Em alguns casos, são utilizadas metodologias específicas que são associadas às opiniões dos consumidores para definir quais dimensões utilizadas são mais úteis na avaliação da qualidade dessas aplicações. Já os critérios associados à contabilidade, consideram os custos financeiros da análise de cada dimensão para definir quais são convenientes para a organização (WANG; STOREY; FIRTH, 1995).

Com o objetivo de identificar dimensões importantes para a análise de qualidade de dados, Wang e Strong (1996) iniciaram sua pesquisa propondo quatro aspectos que acreditavam ser mais comuns na avaliação da qualidade de dados: a capacidade do dado de ser recuperado (acessibilidade), a facilidade de compreensão (interpretação), utilidade para o consumidor (relevância) e quão correto o dado está (acurácia). Além desses aspectos, os autores incorporaram atributos, como disponibilidade e atualidade, fortemente mencionados em suas pesquisas anteriores.

Wang e Strong (1996) elaboraram um questionário para definir quais atributos seriam mais importantes para o consumidor dos dados: foram entrevistados 137 pessoas relacionadas à área de qualidade de dados e, como resultado da primeira fase de pesquisa, foram obtidas 179 características. Em um segundo momento, um novo questionário foi elaborado para definir o grau de importância dos atributos indicados. Wang e Strong (1996) criaram, a partir das notas obtidas, uma ordem dos atributos melhores classificados, onde aspectos semelhantes entre alguns atributos foram observados. Um novo questionário foi realizado utilizando os vinte atributos que melhores se classificaram no resultado da busca anterior; os resultados obtidos foram usados para identificar em quais categorias, consideradas abrangentes o suficiente na avaliação da qualidade de dados, cada atributo estaria. Na sequência, os atributos foram ordenados utilizando os resultados anteriores dentro de cada categoria. Estas categorias podem ser observadas na Tabela 1 e também são utilizadas como base em outros artigos, como Torelli (2005) e Chen *et al.* (2012).

Tabela 1 - Categorias de dimensões na qualidade de dados

Categorias	Definição	Dimensões QI
Intrínseca	A característica observada independente do contexto dentro do ciclo de vida dos dados. Estando relacionada a estrutura da dimensão.	Integridade, objetividade, credibilidade, reputação.

Contextual	Os aspectos observados que resultam em valores diferentes dependendo do momento em que é realizada sua avaliação.	Relevância, valor agregado, obsolescência, preenchimento, quantidade de informação.
Representacional	São dimensões que dizem respeito à coerência e clareza.	Facilidade de interpretação, facilidade de entendimento, representação concisa, representação consistente.
Acessibilidade	São dimensões que julgam a qualidade considerando disponibilidade e segurança no sistema.	Facilidade de acesso, segurança de acesso.

Fonte: Adaptado de Torelli (2005)

Wang e Strong (1996) consideraram que essas quatro categorias conseguem englobar todas as perspectivas de um dado, realçando o fato de serem baseadas na visão dos consumidores de dados e não apenas de estudo teórico.

Algumas das dimensões identificadas por Wang e Strong (1996) são apontadas por outros autores usando denominações diferentes. A seguir, serão descritas algumas dimensões apresentadas por Batini e Scannapieco (2006) que podem ser associadas às dimensões classificadas por Wang e Strong (1996):

- Acurácia (integridade): a acurácia é definida como a proximidade entre um valor v e um valor real v' , considerado como a representação correta do fenômeno da vida real que v busca representar. Por exemplo, se o nome de uma pessoa é João, o valor $v'=João$ é correto, enquanto o valor $v=Joo$ é incorreto. Dois tipos de acurácia podem ser identificados, a acurácia sintática e a acurácia semântica (BATINI; SCANNAPIECO, 2006);
- Consistência (representação consistente): refere-se à violação de regras semânticas definidas ao longo de um conjunto de itens de dados. Com referência à teoria relacional, restrições de integridade podem ser consideradas como formas de regras semânticas; por exemplo, se idade pode variar de 0 até 120 anos (BATINI *et al.*, 2009);
- Completude (preenchimento): pode ser definida como a medida que determina se os dados são de amplitude, profundidade e escopo adequados para a tarefa em mãos (BATINI; SCANNAPIECO, 2006). Três tipos de completude são identificados pelos autores: completude de esquema, definido como o grau em que os conceitos e suas propriedades estão presentes, ou em falta, no esquema; completude de coluna, definida como medida de valores faltantes em uma propriedade específica ou coluna da tabela;

e completude de população, que avalia valores faltantes com referência a registros de toda uma relação;

- Atualidade (obsolescência): Expressa o quão atual são os dados. Essa dimensão tem a importância observada com o fato de que, para determinados destinos, os dados requerem valores com validade específica (BATINI; SCANNAPIECO, 2006).

As dimensões também são conectadas entre si. Se para determinado propósito uma dimensão é considerada mais importante e seus resultados são levados em maior interesse, isso diminui a importância dada a outros aspectos de avaliação (BATINI; SCANNAPIECO, 2006).

2.1.3. Metodologias para gestão da qualidade de dados

Os dados são objetos que representam os fatos do mundo real em um modo que podem ser armazenados, recuperados e processados por um mecanismo de software (BATINI; SCANNAPIECO, 2006). Para compreender de que maneira as atividades de avaliação da qualidade de dados e as atividades de aprimoramento dos dados serão realizadas é interessante considerar o caminho que os dados seguem, os pontos de mudança e as etapas que apresentam a semelhança com o sistema de produção de um produto (MCGILVRAY, 2008).

Reafirmando o paralelo entre a manipulação de dados e a produção de produtos físicos, assim como o produto tem a matéria bruta como entrada, as etapas de produção e a saída como produto final, o dado entra em sua forma original, é manipulado no sistema e, então, tem como saída a informação útil (TORELLI, 2005).

O modo de produção, que Wang *et al.* (2008) chamam de manufaturada, se inicia nas fontes de dados, seguido pela obtenção, armazenamento, processamento, transformação, até a entrega aos usuários consumidores. A divisão em fases do modo de produção dos dados pode ser apresentada de diversas formas e o conjunto de todas as fases é denominado de ciclo de vida dos dados.

A qualidade de dados se torna um conceito com muitas faces e complexo de se investigar, já que suas dimensões têm diferentes definições que concorrem entre si, existindo diversos tipos de dados possíveis e com estruturas dos sistemas também bem variadas.

Wang e Strong (1996) afirmam que o estudo da qualidade de dados pode ser realizado de três formas. (i) A primeira utiliza das pesquisas e da experiência do analisador e suas noções intuitivas para definir os atributos considerados importantes (intuitiva). A segunda (ii) identifica os problemas durante o processo em andamento, focando nas deficiências que são

encontradas pelo caminho e nas divergências do dado com o mundo real (teórica, ou sistêmica). A última (iii) utiliza as opiniões dos clientes; o que pode causar resultados muito individuais e sem padrão exato (empírica).

Na primeira forma os atributos definidos são mais específicos para o uso no processo em questão. Na segunda tendem a ser mais referentes às características dos próprios dados, enquanto que, na terceira, as métricas de avaliação saem do campo de desenvolvimento e têm enfoque nas propriedades de uso.

A diferença entre uma abordagem com ênfase no consumidor e as demais também consiste na visão que o usuário tem da informação. O usuário da informação não tem conhecimento ou controle da qualidade da informação disponível, sendo provável que apresente suas necessidades sem levar em conta os requisitos dos demais usuários (BARROS, 2009).

O resultado do estudo da qualidade de dados pode apresentar inconformidades. Torelli (2005) aponta algumas: a possibilidade de haver um aspecto do mundo real retratado duas vezes no sistema de formas diferentes (duplicidade), de informações diferentes ser representadas de forma idêntica no sistema (ambiguidade), de não existir dados que descrevam uma ou mais informações reais (representação incompleta) e do sistema ter uma representação do modelo não equivalente a nenhum aspecto do mundo real (representação sem correspondência), além de problemas operacionais, como a associação de determinado acontecimento real a dados que representam outro aspecto dentro do sistema.

É importante realizar melhorias no processo de produção da informação e retirar as não conformidades que são encontradas pelo caminho. Criar melhorias na qualidade de dados significa mensurar, avaliar e aprimorar as operações que integram de forma mais efetiva as características dos dados (CHEN et al., 2012). Para que essas melhorias sejam alcançadas, é importante estruturar como serão conduzidas as atividades exigidas.

A partir da análise da literatura sobre gestão da qualidade de dados, é possível observar que as diversas abordagens apresentadas focam na orientação da estruturação de operações complexas, apresentando, em graus diferentes de detalhamento, os recursos necessários e como obtê-los. Além disso, outro ponto comum é a observação de que é necessária a adaptação ao ambiente em que são aplicadas. Não existe a necessidade de que os guias sejam realizados de forma absolutamente precisa. O ideal é que o projetista concilie da melhor forma possível as propostas de acordo com as circunstâncias e características do domínio.

Metodologias de qualidade de dados são conjuntos de orientações e técnicas para definir um processo racional que se inicia na entrada dos dados e passa por todas as fases do processo junto a seus pontos de decisão com foco na qualidade dos dados em todas as etapas (BATINI; SCANNAPIECO, 2006).

Cada metodologia estrutura as atividades que a compõe em fases diferentes, dependendo da abordagem definida. Batini e Scannapieco (2006) apresentam as fases mais comuns visíveis nas abordagens das metodologias de avaliação de qualidade de dados e não seguem estritamente uma ordem específica:

- Análise de dados, onde é verificada a coleta, o fluxo, a arquitetura e as regras pelas quais os dados passam;
- Análise dos requisitos de qualidade de dados, onde são apuradas as metas de qualidade, e abordadas sugestões de melhorias possíveis;
- Encontrar áreas críticas, fase na qual se encontram as partes mais relevantes dos fluxos de dados para se analisar;
- Modelagem do processo, etapa para descrever, de modo formal, os procedimentos que serão realizados;
- Medições, onde são definidos os critérios que serão utilizados e realizadas as medições;
- Estimação de custo, onde é verificado qual o custo para uma qualidade baixa;
- Estimação de benefícios, fase na qual é identificado o ganho obtido com a melhoria da qualidade de dados;
- Atribuição de responsabilidades de processo, etapa na qual, para cada procedimento, é designado quem será seu responsável;
- Atribuição de responsabilidades de dados, etapa na qual, para cada tipo de dados, é determinado o responsável pelo controle;
- Escolha de ferramentas e técnicas, fase na qual são definidas quais ferramentas e técnicas melhores se encaixam para o sistema proposto.

Em geral, a estrutura das metodologias de avaliação pode ser condensada em três etapas, como aponta Barros (2009):

1. Escolher, classificar e medir as dimensões e métricas para a avaliação;
2. Executar a pesquisa com especialista para os critérios subjetivos;
3. Comparar os resultados objetivos com os subjetivos.

É possível classificar as metodologias com base em diversos parâmetros. Usando o critério de orientação, é possível dividi-las em orientadas a dados, quando apenas o próprio dado é utilizado para melhorar a qualidade das informações, e orientadas a processos, quando os procedimentos durante a produção são analisados para identificar as causas dos problemas (BATINI;SCANNAPIECO, 2006).

Outra propriedade para classificação apontada por Batini e Scannapieco (2006) é o que a metodologia a ser classificada considera como principal objetivo, que pode ser o foco na medição e/ou o foco na melhoria. Apesar de relacionadas, já que é necessário que se tenha métricas para que seja possível criar técnicas de melhoria, existe a divisão entre as metodologias com as atividades para a criação de métricas e as metodologias com os procedimentos para aperfeiçoamento.

Outra classificação possível pode ser baseada no propósito das metodologias, propósito geral ou específico. Enquanto a primeira apresenta atividades que podem ser implantadas em diversos tipos de processos, a segunda é criada com finalidade específica, como identificação de objeto ou aplicação em um banco de dados distinto, ou para uma aplicação final já definida.

Existem várias classificações possíveis para as metodologias que se baseiam em diferentes tipos de parâmetros, que podem ser buscados de acordo com a necessidade. Algumas organizações criam suas próprias metodologias ou modificam metodologias existentes para que se enquadrem de forma efetiva em seu domínio. Exemplos podem ser encontrados em Wang et al. (1995), Bastos et al. (2010) e Bahchi et al. (2006).

2.1.4. Total Data Quality Management (TDQM)

Para gestão da qualidade, a abordagem apresentada pelo TQM (*Total Quality Management*) tem como objetivo atender às necessidades dos clientes de determinado produto físico ou serviço. Para isso, segundo Cao *et al* (2010), o TQM considera os gerentes e funcionários e busca trazer melhorias na organização da produção. Como resultado, há uma estrutura base que organiza os procedimentos de avaliação e prevenção dentro da organização com enfoque na qualidade do produto.

A metodologia TDQM (*Total Data Quality Management*) do Instituto de Tecnologia de Massachusetts (MIT – *Massachusetts Institute of Technology*) utiliza o TQM de forma adaptada, assumindo como produto cada parte da informação que é produzida pelo sistema (WANG, 1998). Devido à natureza dos dados, alguns parâmetros e conceitos da gestão de qualidade de produtos físicos são limitados no sistema de informação e devem ser ajustados.

Uma peculiaridade dos dados, por exemplo, é que, diferente da matéria prima de produtos físicos, os dados não são esgotáveis e podem ser reutilizados diversas vezes e para finalidades diferentes (CAO et al., 2010). Isso implica em perspectivas diferentes de cada uma das utilizações para o mesmo item, sendo que nem sempre as especificações dos usuários são iguais.

A metodologia TDQM foi desenvolvida com o acúmulo de pesquisas e extensas experiências práticas (WANG, 1998). Sua estrutura depende de dois conceitos importantes: a definição do ciclo TDQM e a definição de produto de informação.

Adaptando o ciclo PDCA (*Plan, do, check, act*), também conhecido como ciclo de Deming, foram desenvolvidas as quatro etapas que compõem o ciclo TDQM. As etapas do TDQM são: (i) definir, etapa referente às atividades de definição das dimensões importantes de qualidade de dados; (ii) mensurar, etapa composta pelas atividades responsáveis por definir as métricas das dimensões; (iii) analisar, conjunto de atividades de análise dos resultados obtidos com as métricas, e; (iv) aprimorar, etapa onde atividades de melhorias de dados são colocadas em prática. A Figura 1 ilustra o ciclo TDQM .

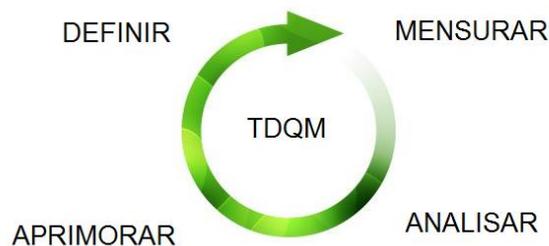


Figura 1 - Ciclo do TDQM

Fonte: Adaptado de Wang (1998)

O conceito de produto de informação é proveniente da semelhança do sistema de informação com o sistema de produtos manufaturados. Isso indica que o resultado do sistema de informação que será entregue para o usuário na posição de consumidor será considerado um produto, denominado produto de informação (WANG, 1998).

2.2. Aprendizado baseado em jogos

Os principais objetivos desse trabalho em relação ao jogo proposto é que o aprendiz assimile a metodologia apresentada no Capítulo 3 corretamente, consiga visualizar seu efeito na prática, e que se mantenha motivado durante a fase de aprendizado. É certo que a efetividade e o potencial educacional de um jogo é variável de acordo com as escolhas

pedagógicas realizadas pelos projetistas (ARNAB et al., 2015). Logo, a fim de desenvolver um jogo que conseguisse atingir os objetivos apresentados, foi realizada uma investigação sobre os componentes essenciais em jogos, abordagens de aprendizado e, principalmente, o relacionamento entre ambos.

Jogos que trazem como objetivo o ensino, aprendizagem ou treinamento são denominados jogos sérios, termo em inglês *serious games* (ROCHA; BITTENCOURT; ISOTANI, 2015). Jogos com outros propósitos sérios como conscientização ou socialização também fazem parte do gênero jogos sérios (RIBEIRO et al., 2015). Um dos problemas ao se projetar um jogo sério é combinar as metas de aprendizagem que se espera atingir com sua utilização e as características típicas de um jogo. Os desenvolvedores de jogos sérios, que apresentam características de entretenimento, consideradas fundamentais, não devem priorizar tais características prejudicando o caráter educacional do jogo (RIBEIRO et al., 2015).

2.2.1. Objetivos educacionais e elementos motivacionais

Os jogos podem apresentar diferentes características, o que se reflete em diferentes tipos de aprendizagem. Logo, é essencial entender o quanto os componentes de um jogo podem interferir na aprendizagem e utilizar medições apropriadas para avaliar os resultados esperados em relação às metas e ao engajamento (ARNAB et al., 2015).

Antes de estudar os componentes de um jogo é interessante considerar quais são as metas que se espera atingir para que os componentes candidatos ao projeto sejam adequadamente selecionados. Jogos sérios para educação e treinamento têm como meta primária a pedagogia, sendo pedagogia a prática da teoria de aprendizagem (ARNAB et al., 2015). A ausência de uma teoria de aprendizagem bem fundamentada no desenvolvimento de um jogo sério impacta de forma direta no sucesso do jogo em seu propósito de ensino (RIBEIRO et al., 2015).

É mais fácil atingir metas quando estas são bem definidas. No contexto educacional, especificar claramente e de forma estruturada os objetivos relacionados ao conhecimento e competências que os aprendizes devem alcançar, direciona corretamente o processo de ensino às estratégias, métodos, limites de conteúdo e instrumentos de avaliação para um ensino efetivo (FERRAZ; BELHOT, 2010). Esses objetivos são denominados objetivos educacionais (COSTA et al., 2014).

Para definir os objetivos educacionais em relação ao nível de aprendizagem que se espera com a aplicação do *Data Quality Simulator* e avaliar seu impacto optou-se pelo uso da taxonomia de Bloom (BLOOM, 1956) que aparece na literatura entre as abordagens mais

utilizadas na avaliação de softwares educacionais (HENRIQUE et al., 2015) e jogos digitais educacionais (RIBEIRO et al., 2015).

2.2.2. Taxonomia de Bloom

A taxonomia de Bloom foi desenvolvida por um grupo de examinadores de diferentes universidades, encabeçados por Benjamin Bloom, e publicada em 1956. A inspiração para sua criação foi organizar os níveis de aprendizagem que os educadores esperam alcançar de forma semelhante às taxonomias de outras áreas como, por exemplo, a classificação dos relacionamentos de animais e plantas na biologia. A taxonomia estabeleceu um padrão na forma de linguagem sobre objetivos educacionais (FERRAZ; BELHOT, 2010). A principal aplicação da taxonomia é ajudar professores, instrutores e avaliadores a especificar objetivos de tal forma que se torne mais fácil planejar experiências de aprendizado e preparar dispositivos de avaliação (BLOOM, 1956).

A taxonomia é dividida em três domínios: cognitivo, afetivo e psicomotor; sendo que os dois últimos, apesar de mencionados por Bloom *et al.* (1956), só foram aprofundados em trabalhos posteriores (BLOOM; KRATHWOHL; MASIA, 1972) (FERRAZ; BELHOT, 2010). O domínio cognitivo inclui metas relacionadas ao reconhecimento e lembrança de informações e desenvolvimento de habilidades e competências intelectuais. O domínio afetivo engloba metas relacionadas ao interesse do aprendiz, suas atitudes e valores; são metas que buscam conduzir o nível de interesse que o aprendiz tem com o conteúdo apresentado. O domínio psicomotor é a área de manipulação e habilidades motoras (FERRAZ; BELHOT, 2010). Neste trabalho, será considerado o domínio cognitivo, a fim de definir objetivos educacionais relacionados ao conteúdo da metodologia na condição de elementos intelectuais; e o domínio afetivo, para direcionar o estudo do que é considerado engajador para o aprendiz, buscando que o mesmo considere o conteúdo ensinado valioso e se sinta motivado a continuar o aprendizado.

Os níveis de aprendizagem em cada domínio da taxonomia de Bloom foram estruturados de forma hierárquica, nos quais metas que se encontram em determinado nível são dependentes das metas de níveis anteriores ou são construídas a partir delas (FERRAZ; BELHOT, 2010).

Para o domínio cognitivo, foram estabelecidos seis níveis de aprendizagem: conhecimento, compreensão, aplicação, análise, síntese e avaliação, nessa ordem. Cada nível está relacionado aos verbos que expressam ações e que podem ser usados ao descrever as atividades que serão desenvolvidas pelos aprendizes. Esses verbos se relacionam com a

definição dos objetivos educacionais (COSTA et al., 2014). A Tabela 2 apresenta os níveis com suas respectivas definições e verbos. No domínio afetivo, são classificados os níveis: receber fenômenos, responder a fenômenos, valorização, organização e caracterização, e suas definições são observadas na Tabela 3.

Tabela 2 - Níveis no domínio cognitivo

Níveis	Definição	Verbos
Conhecimento	O aluno irá recordar ou reconhecer informações, ideias e princípios na forma aproximada em que foram aprendidos.	Escreva, lista, rotule, mostre, tabule, enumere, copie, selecione, nomeie, diga, defina, reproduza, relate, identifique, cite, colete, evoque.
Compreensão	O aluno traduz, compreende ou interpreta informações com base em conhecimento prévio.	Explique, associe, entenda, estimule, agrupe, sumarie, converta, discuta, traduza, ordene, diferencie, resuma, parafraseie, descreva, interprete, ilustre.
Aplicação	O aluno seleciona, transfere e usa dados e princípios para completar um problema ou tarefa com um mínimo de supervisão.	Use, compute, resolva, aplique, calcule, termine, experimente, demonstre, descubra, determine, torne, estabeleça, articule, transfira, ensine, prepare, construa.
Análise	O aluno distingue, classifica e relaciona pressupostos, hipóteses, evidências ou estruturas de uma declaração ou questão.	Analise, classifique, categorize, compare, contraste, deduza, arranje, conecte, divida, priorize, indique, diagrame, discrimine, focalize, separe.
Síntese	O aluno cria, integra e combina ideias num produto, plano ou proposta nova para ele.	Crie, proponha, formule, modifique, planeje, elabore, invente, projete, desenvolva, ligue, componha, generalize, substitua, integre, rearranje, reescreva, adapte, antecipe, compile.
Avaliação	O aluno aprecia, avalia ou critica com base em padrões e critérios específicos.	Julgue, argumente, avalie, recomende, critique, justifique, decida, teste, convença, conclua.

Fonte: (COSTA et al., 2014)

Tabela 3 - Níveis no domínio afetivo

Níveis	Definição
Receber fenômenos	Referente a disposição do aprendiz em atender um fenômeno particular de estímulo (atividades em classe, textos, música...)
Responder a fenômenos	Referente a participação ativa por parte do aprendiz, que não apenas atende a um fenômeno, mas também reage a ele de alguma forma. As metas nesse nível buscam que o aprendiz apresente por si só vontade de responder aos fenômenos percebidos.
Valorização	Referente ao valor ou importância que o aprendiz dá a determinado objeto, fenômeno ou comportamento, desde a mais simples aceitação a um grau mais complexo de compromisso. Estão relacionados a esse nível metas referentes às atitudes e apreciação, que têm seus valores observados de forma mais clara quando considerado os hábitos consistentes e estáveis do aprendiz.
Organização	Referente a união dos diversos valores, resolvendo conflitos entre eles para começar a construir um sistema de valores consistente. As metas aqui envolvem que o aprendiz compare, relacione e organize os valores que criou, construindo para si uma ordem de importância de suas atitudes em relação ao conteúdo.
Caracterização	Neste nível o aprendiz já tem um sistema de valores que controla seu comportamento criando um hábito espontâneo em sua vida. Logo, suas ações se tornam generalizadas e preditivas. As metas presentes nesse nível são relacionadas ao ajuste apropriado dos padrões desenvolvidos pelo aprendiz.

Fonte: (BLOOM; KRATHWOHL; MASIA, 1972)

É possível observar pelas definições dos níveis de aprendizado que para se atingir efetivamente metas de níveis superiores é necessário já ter atingido metas de níveis inferiores de forma plena. Portanto, é preciso ter conhecimento sobre quais serão as atividades instrucionais envolvidas para se atingir cada uma das metas progressivamente.

2.2.3. Taxonomia de motivações intrínsecas

O ideal em uma atividade de aprendizagem é que ela seja intrinsecamente motivadora, ou seja, que a pessoa se envolva por sua própria vontade e não para receber uma recompensa externa ou para evitar uma punição (MALONE; LEPPER, 1987). Os aprendizes devem avaliar a atividade de forma favorável ou não se sentirão motivados a continuar o aprendizado (SAVI et al., 2010). Nesse sentido, Malone e Lepper (1987) estabeleceram uma taxonomia revisada de motivações intrínsecas classificadas em motivações individuais e interpessoais.

As motivações individuais estão relacionadas com componentes presentes na atividade, neste caso, o jogo, que são independentes do relacionamento do aprendiz (na condição de jogador) com outras pessoas: desafio, fantasia, curiosidade e controle.

Para o desafio, a literatura afirma que existe uma faixa ótima de desafio preferível. Tanto atividades de nível trivial como atividades de nível impossível são igualmente de baixo interesse para as pessoas (SAVI et al., 2010). O ideal é que as atividades sejam de um nível de dificuldade intermediário, causando maior estímulo intrínseco.

Dentre os interesses observados dentro do tópico desafio pode-se apontar a percepção do aprendiz de sua competência, um fluxo de estados e a sensação de auto-eficiência. Assim, para que uma atividade seja desafiadora, deve prover metas claras de forma que a realização seja incerta. Também deve-se fornecer um *feedback* de desempenho relacionado à meta. Desse modo, o jogador poderá analisar se foi capaz de realizá-la, se teve uma mudança do estado atual em que ele se encontra para o estado desejado (cumprimento da meta), e se sua forma de realização foi eficaz, buscando manter a autoestima do jogador (MALONE; LEPPER, 1987).

Uma forma de implementação deste artifício é criar uma sequência gradual de níveis de dificuldade, para que jogadores de qualquer nível possam progredir com percepção de sucesso satisfatório. É interessante fornecer ao jogador informações sobre seu estado atual, indicando qual seria o estado desejado e um conjunto de possíveis ações para transformar o estado atual no seguinte, além de um método de pesquisa para que o aprendiz analise qual a melhor ação a ser tomada para atingir o novo estado. Assim, jogos se tornam divertidos

porque tocam o instinto inato das pessoas para a aquisição de novas habilidades e percepção do próprio progresso (MORENO-GER et al., 2014).

O outro elemento, a fantasia, é definido como um ambiente com habilidade de fazer surgir na mente situações físicas que não estão realmente presentes (MALONE; LEPPER, 1987). Um argumento para a utilização da fantasia é a possibilidade de prover a sensação de poder, sucesso e/ou sorte e criar maestria em situações que são confusas ou indisponíveis na vida real.

Nesta questão, há a dificuldade em se identificar quais necessidades que as pessoas têm para conseguir criar um ambiente de fantasia que consiga ser atraente para os diferentes jogadores. O fator identificação é um dos mecanismos que podem ser considerados para alcançar este objetivo, sendo observado quando o jogador percebe uma similaridade entre si e um personagem, cria admiração pelo personagem ou nota relevância do ponto de vista do personagem. Além dos aspectos emotivos da fantasia, também existem os aspectos cognitivos. Metáforas e analogias podem trazer um ganho na aprendizagem relacionando novas informações aos conteúdos que o aprendiz já absorveu anteriormente ou aprimorando a memória de um mesmo conteúdo com imagens mais vívidas (MALONE; LEPPER, 1987).

O terceiro elemento, curiosidade, assim como o desafio, tem relação com uma faixa ideal. Muita ou pouca discrepância entre a expectativa provocada e a informação entregue podem diminuir o interesse do jogador em relação ao novo conteúdo. Malone e Lepper (1987) apresentam duas formas de curiosidade: curiosidade sensorial, despertada pela atração da atenção com variações de som, luz e outros estímulos sensoriais do ambiente; e a curiosidade cognitiva, impulsionada pela expectativa de mudança na estrutura cognitiva, ou seja, quando o ambiente proporciona ao jogador a sensação de não ter o conhecimento certo ou completo e que deve buscar consertar essa falha.

O último elemento motivacional apresentado como capaz de atrair as pessoas é a possibilidade de proporcionar aos jogadores uma grande sensação de controle, sendo que o ser humano tem uma tendência a querer controlar o ambiente, buscando encontrar em si mesmo a origem de suas ações e escolhas. Deste modo, uma característica básica para um bom jogo é evidenciar que os resultados obtidos são, de fato, dependentes das respostas dadas por ele (SAVI et al., 2010).

É necessário ter um *feedback* de uma função que quantifique o sucesso e o fracasso e que mostre ao jogador que suas escolhas interferem nesse resultado. Outra possibilidade de apresentar a sensação de controle é permitir que o jogador escolha qual tarefa realizar,

problema estudar ou resolver, qual dificuldade de meta definir, ou em que ordem executar (MALONE; LEPPER, 1987).

Além das motivações individuais, a taxonomia de Malone e Lepper (1987) apresentam as motivações que possuem dependência de outras pessoas, chamadas de interpessoais: cooperação, competição e reconhecimento. Algumas dessas motivações podem ser decompostas em motivações individuais, como a competição, que pode ser entendida como uma forma de motivação por meio do desafio de superar outro jogador.

A motivação por competição pode ser implementada através de comparação de pontuações baseadas em erros e acertos, por exemplo. Assim, é possível comparar diferentes performances dos jogadores ou mesmo a pontuação anterior do mesmo jogador (MALONE; LEPPER, 1987).

A motivação de cooperação pode ser atribuída ao jogo quando os resultados mostrados no *feedback* aos jogadores têm relação com as ações tomadas por todos os participantes. A implementação pode combinar a pontuação dos jogadores individualmente, por exemplo, ao se definir times onde cada jogador jogará sozinho, e as notas de todos os integrantes do time serão somadas para obtenção de uma única nota para a equipe. Outra forma de implementação é quebrar uma atividade em diversas metas menores, fazendo com que cada participante da atividade se ocupe de uma tarefa específica (MALONE; LEPPER, 1987).

A motivação por reconhecimento representa a concordância de que as pessoas gostam de ter seus esforços e conquistas apreciados pelos outros. Para tanto, um ambiente que queira utilizar essa motivação deve fornecer visivelmente resultados das atividades do jogador para que ele consiga apresentar a outras pessoas (SAVI et al., 2010). Esses resultados podem ser o processo e/ou caminho definido pelo jogador, o efeito atingido com suas ações, ou resultado em forma de pontuação ou *ranking* (MALONE; LEPPER, 1987).

Uma conclusão possível ao se analisar todas as motivações apresentadas é a necessidade do aprendiz de obter o *feedback* relacionado com sua performance. Esses resultados serão a confirmação de como sua performance foi aumentando ou diminuindo sua motivação para se manter nesse ambiente de aprendizado.

Para que um jogo apresente elementos motivacionais é necessário implementar recursos que cumpram as funções descritas para cada elemento desejado. Esses recursos são considerados como requisitos do software, sendo que os dez requisitos mais utilizados no desenvolvimento de software educacionais, de acordo com Henrique *et al.* (2015), são *feedback*, pontuação, níveis, tempo, texto, dicas, ajuda, *ranking*, desafio e imagens.

3. Metodologia de avaliação de qualidade de dados

Neste capítulo serão abordadas as etapas da metodologia de avaliação de qualidade de dados desenvolvida a partir dos estudos da fundamentação teórica, com ênfase na metodologia TDQM.

A metodologia descrita para executar um programa de qualidade de dados é composta por cinco etapas que não devem ser entendidas de forma estritamente sequencial, e sim de forma cíclica: preparação, definição, avaliação, análise e ação. A cada iteração de um ciclo de execução do programa de qualidade, o foco pode variar entre as etapas. O ciclo, apresentado na Figura 2, deve ser executado até que a qualidade identificada nos produtos de informação esteja compatível com as expectativas dos atores envolvidos.



Figura 2 - Etapas da metodologia proposta

3.1. Preparação

Esta primeira etapa da metodologia é composta pela elaboração e organização do programa de análise dos dados; ou seja, nessa etapa a organização deve se planejar para que o andamento do programa ocorra de forma estruturada.

É essencial que sejam definidas estratégias que permitam a gestão das atividades executadas. Desta forma, ao elaborar uma metodologia para execução de um programa de qualidade de dados, é pertinente considerar aspectos relacionados à gestão do programa. Nesta etapa, estes aspectos devem ser abordados. Diversas metodologias com objetivos distintos podem ser utilizadas como possíveis recursos para facilitar este processo de gestão.

É importante ressaltar que as discussões apresentadas devem ser consideradas somente como recomendações e a possível utilização em um contexto real deve estar sujeita à análise

de viabilidade de utilização de cada recomendação. Além disso, não faz parte do escopo deste trabalho explorar, em profundidade, os aspectos gerenciais deste processo, o foco está relacionado aos aspectos técnicos da execução de um programa de qualidade de dados

Na 5ª edição do guia PMBOK publicado pelo *Project Management Institute* (2013) é apontado que o gerenciamento de projeto, em geral, apresenta, entre outras, as seguintes tarefas: identificação dos requisitos (refere-se aos requisitos do projeto e não do sistema); identificação das diferentes necessidades, preocupações e expectativas das partes interessadas no planejamento e execução do projeto; estabelecimento, manutenção e execução de comunicações, eficazes e colaborativas entre as partes interessadas; gerenciamento das partes interessadas visando o atendimento aos requisitos do projeto e a criação das suas entregas e manter o equilíbrio das restrições conflitantes do projeto (escopo, qualidade, cronograma, orçamento, recursos e riscos).

Ainda na mesma edição do guia (PMI, 2013), destaca-se que os produtos dessas tarefas estão fortemente entrelaçados. Isso significa que a mudança de um item pode modificar o resultado de outros. Por exemplo, a adição de um novo requisito pode modificar o risco e o tipo de recurso que já foi estabelecido anteriormente.

Um recurso que pode facilitar a delimitação e a gestão das atividades que devem ser exercidas durante toda a execução é elaboração da Estrutura Analítica do Projeto (EAP, do inglês *Work Breakdown Structure*). Essa estrutura é utilizada para definir o escopo do projeto de acordo com artefatos que devem ser entregues (PMI, 2013).

Depois de realizar as tarefas envolvidas na atividade de elaboração do plano para o programa de análise de dados o produto obtido é um plano que pode ser composto por:

- Estrutura Analítica do Projeto, mostrando todas as atividades que serão executadas;
- Estrutura de Gerenciamento do Projeto, apontando quais serão as ferramentas e metodologias que serão utilizadas para gerenciar o programa quando este está em andamento;
- Gerenciamento de responsabilidades, indicando para cada atividade os responsáveis, ou grupo de responsáveis;
- Gerenciamento de riscos, apresentando as possíveis situações indesejadas que podem ocorrer durante execução do programa e as ações que devem ser tomadas para solucioná-las, e
- Plano de comunicação, definindo:

- As ferramentas de comunicação que serão utilizadas, (e-mails, atas, ferramentas eletrônicas, documentos impressos);
- Contatos dos envolvidos (telefones, endereços eletrônicos ou outros meios específicos do usuário), e
- Periodicidade de reuniões, possivelmente já definindo a frequência e o objetivo das futuras reuniões.

3.2. Definição

De acordo com Rezende (2003), um produto de informação é o dado útil, resultante de um sistema de informação, que foi tratado e apresenta um valor e um sentido lógico para quem usa a informação. Neste trabalho, o produto de informação é considerado como referencial da análise. Desta forma, o objetivo desta etapa é identificar os produtos de informação que serão contemplados no programa e as características de cada produto que serão analisadas.

A identificação dos produtos de informação pode ser realizada a partir de diferentes estratégias. De acordo com o *Balanced Scorecard*, ao se tentar definir uma estratégia, a identificação dos temas estratégicos é um elemento importante. Um tema pode ser considerado um conjunto de objetivos com uma finalidade ou um desafio comum. Juntos, os temas representam os eixos da estratégia (KAPLAN; NORTON, 1997).

Neste trabalho, a definição dos temas estratégicos é sugerida como um recurso para facilitar a análise do produto de informação. Enquanto os produtos de informação fornecem as dimensões de análise do produto, os temas devem representar estas dimensões sob a perspectiva do negócio. Esta visão permite apontar subconjuntos de dados e regras existentes no sistema que podem ser abordados separadamente com o objetivo de melhorar a qualidade do produto de informação, sendo esses subconjuntos os temas estratégicos existentes.

Para a execução desta etapa, foram definidas quatro tarefas: (i) obter conhecimento sobre o domínio; (ii) identificar e selecionar os temas estratégicos; (iii) detalhar os temas estratégicos definidos estudando os produtos de informação; e (iv) definir as dimensões e métricas que serão analisadas. Estas tarefas serão detalhadas a seguir.

3.2.1. Obter conhecimento sobre o domínio

Para que os temas estratégicos sejam definidos é necessário obter todo conhecimento, explícito ou tácito, que tenha relevância sobre a situação dos dados no domínio considerado. Dentre as fontes de conhecimento possíveis, pode-se citar:

- *Feedbacks* encaminhados a organização;
- Requisitos do sistema;
- Restrições do sistema e da organização;
- Informações formais de metas, valores, normas e políticas da organização;
- Lições aprendidas: informações sobre antigos problemas, medidas tomadas e resultados alcançados.

A análise das informações obtidas nesta tarefa resulta em uma visão geral, mesmo que ainda pouco detalhada, do que deve ser considerado no projeto em questão, abordando os problemas existentes, as necessidades dos envolvidos e a disponibilidade e restrições da organização para atendê-las.

3.2.2. Identificar e selecionar os Temas Estratégicos

A segunda tarefa desta etapa é identificar os temas estratégicos e descrevê-los considerando o objetivo estratégico da organização em relação a qualidade esperada e o nível de prioridade, como representado na Tabela 4.

Tabela 4 - Definir temas e prioridades

Tema	Objetivo	Prioridade
Apresenta de forma sucinta do que se trata o tema.	Especifica a proposição de valor associada ao tema para os interessados no processo.	Apresenta o grau de prioridade do tema abordado de acordo com a estratégia de priorização definida.

Em um domínio é possível verificar vários temas estratégicos diferentes. Portanto, é importante considerar alguma estratégia de priorização desses temas. A complexidade do tema é um exemplo de medida que pode ser usada para a priorização e seleção dos temas identificados. A priorização também pode relacionar diversas formas de custo, como tempo ou trabalho, além do custo financeiro que o tema envolve e os possíveis riscos ao abordá-lo.

Para ilustrar as atividades propostas na metodologia, será considerado um sistema hipotético de *e-commerce* voltado para venda de livros, que será contextualizado no desenvolvimento do jogo. Todas as informações apresentadas no exemplo são, apesar de possíveis, fictícias, criadas apenas para demonstração do uso da ferramenta. A partir do

estudo das informações obtidas na tarefa anterior, supõe-se que foram encontrados os temas estratégicos apresentados na Tabela 5:

Tabela 5 - Exemplos de temas estratégicos

Tema	Objetivo	Prioridade
Recomendações aos clientes	Este tema refere-se as recomendações de livros aos clientes baseando-se em compras anteriores e avaliações feitas pelos próprios clientes em relação aos livros adquiridos. O sistema permite que o cliente avalie as recomendações tendo sido observado que as notas referentes às recomendações não são satisfatórias.	Média
Precificação de produtos	Este tema é referente aos preços dos livros. Foram identificadas inconsistências relacionadas ao preço de alguns produtos que é exibido com valores diferentes dependendo da etapa de interação com usuário.	Alta
Disponibilidade de produtos	Este tema refere-se à permissão de fechamento de compras verificando a quantidade em estoque. Foram observados problemas relacionados à não identificação do estado do produto fora de estoque e/ou fora de linha até o momento de finalização da compra.	Alta
Informações específicas de livro	Este tema refere-se aos dados específicos de cada livro, como título, autor, número de páginas e descrição. Foram identificados problemas relacionados à disponibilidade de algumas das informações técnicas necessárias para os clientes.	Alta

Seção de comentários de clientes.	Este tema tem relação ao recurso disponível no site correspondente à seção de comentários. O cliente tem permissão para enviar ao sistema opiniões sobre determinado livro. Porém, foi observado problemas relacionados à funcionalidade de exclusão de comentários pelo autor.	Média
-----------------------------------	---	-------

3.2.3. Detalhar os temas estratégicos definidos

Após as atividades de definição dos temas estratégicos é necessário que se realize um detalhamento maior sobre eles. No detalhamento, serão analisados os produtos de informação relacionados a cada tema. Esse detalhamento envolve: (i) identificar o(s) produto(s) de informação, (ii) analisar o ciclo de vida do(s) produto(s) identificado(s), (iii) identificar os atores que interagem com cada produto, como interagem e o que esperam dos dados para que considerem o produto de informação como de qualidade.

Para identificação do produto de informação é necessário relacionar o objetivo dos temas com os dados envolvidos; assim, é possível distinguir os conjuntos de dados que são relacionados ao mesmo produto de informação. Por exemplo, o tema Informações específicas de livros aborda problemas com dados técnicos dos livros. No caso, todos os atributos envolvidos pelo tema (título, autor, número de páginas e descrição) são referentes ao produto de informação Livro; logo, este é o produto de informação a ser analisado.

O objetivo do tema Recomendações aos clientes tem relação com informações sobre as notas de avaliação dos livros, compras anteriores dos clientes e recomendações anteriores avaliadas pelos clientes. Os produtos de informação identificados para esse tema são: produto de informação Livro, produto de informação Compras e produto de informação Cliente. A Tabela 6 apresenta os produtos de informação definidos para o exemplo retratado.

Tabela 6 – Produtos de informação relacionados aos temas estratégicos

Tema	Produto de informação
Recomendações aos clientes	<ul style="list-style-type: none"> • Cliente • Compra • Livro

Precificação de produtos	<ul style="list-style-type: none"> • Livro
Disponibilidade de produtos	<ul style="list-style-type: none"> • Compra • Livro
Informações específicas de livro	<ul style="list-style-type: none"> • Livro
Seção de comentários de clientes.	<ul style="list-style-type: none"> • Cliente • Livro

McGilvray (2008) apresenta em seu livro um *framework* para qualidade de informação que tem como objetivo prover uma estrutura organizada para que seja analisado de forma menos complexa o ambiente onde está o produto de informação. A principal ferramenta utilizada no *framework* de McGilvray (2008), chamada de matriz de interação, utiliza dois conceitos relacionados aos produtos de informação: o ciclo de vida da informação e os componentes chaves.

Para estudar um produto é importante entender todo o caminho que ele percorre desde sua origem até o descarte. O acrônimo POSMAD, apresentado por McGilvray (2008), divide o ciclo de vida da informação em seis fases:

- Planejamento (*plan*): etapa em que são identificados os objetivos de utilização do produto de informação estudado, e ocorre o planejamento da arquitetura em que será implementado e desenvolvimento de normas e definições para esta estrutura;
- Obtenção (*obtain*): etapa relacionada à análise de como os dados são adquiridos; por exemplo, criação de registros, carregamento externo, compras ou carregamento de arquivos externos;
- Armazenamento e compartilhamento (*store and share*): etapa referente a verificar como os dados são armazenados e disponibilizados para uso;
- Manutenção (*maintain*): esta etapa aborda como é feita a atualização, alteração e manipulação dos dados. Como e se é executada a limpeza, a transformação, se existem fusões de registros e/ou interações semelhantes;
- Aplicação (*apply*): é a análise das informações relacionadas ao uso dos dados. É a fase do ciclo em que investiga como são recuperados os dados para tomada de decisão, realização de transações, criação de relatórios e execução de processos automáticos;

- Descarte (*dispose*): etapa relacionada a analisar como as informações são arquivadas e como os dados ou registros são excluídos.

McGilvray (2008) indica em seu *framework* de qualidade de informação a existência de quatro componentes chaves que afetam a informação durante todo seu ciclo de vida. Esses componentes abrangem os elementos 5M1E (Homem, Máquina, Material, Método, Medição e Ambiente) apontados por Wang et al. (2008). McGilvray (2008) orienta que esses componentes sejam observados para cada fase do ciclo de vida da informação, sendo:

- Dados: fatos conhecidos ou outros itens de interesse;
- Processos: funções, atividades, ações, tarefas e procedimentos;
- Pessoas e organizações: organizações, times, indivíduos, atores e funções;
- Tecnologia: formulários, aplicações, base de dados, programas, códigos, e outras ferramentas que vão desde grandes bancos até o papel físico.

A matriz de interação do POSMAD é a organização, em um modelo de tabela, dos componentes chaves e da estrutura apresentada do ciclo de vida das informações. Essa matriz apresenta todo o conhecimento necessário para cada componente presente no ciclo de vida das informações. Na Tabela 7 são apresentadas possíveis questões para o estudo do ciclo de vida de um produto de informação.

Ao analisar a informação, é importante identificar, para cada produto de informação, todos os atores que interagem com o dado durante o ciclo de vida, definir qual o seu papel de atuação: fornecedor, produtor, consumidor ou gerente de informação (WANG, 1998), especificar como é feita essa interação com o produto e o objetivo dessa interação.

Para ter um conhecimento mais exato do papel de cada ator no sistema, é importante criar questões pertinentes que consigam expor, da forma mais clara possível, quais interações ocorrem com os dados, as causas e necessidades dessas ocorrências, quais os objetivos e o que é esperado pelos atores.

As respostas obtidas com a análise da matriz de interação do POSMAD são utilizadas para a elaboração de uma definição formal de como o sistema funciona e se relaciona com o produto de informação estudado em diversos níveis de perspectivas e profundidade. Realizando esse estudo para cada produto de informação que compõe um tema estratégico, a visão que se tem do tema se torna mais abrangente e clara, o que permite que a avaliação e resolução dos problemas associados ao tema sejam realizadas de forma mais objetiva.

Considerando o contexto hipotético de *e-commerce* descrito anteriormente, as Tabelas 8, 9 e 10 abordam um exemplo de utilização da matriz de interação do POSMAD em relação ao produto de informação Livro.

Tabela 7 - Matriz de interação POSMAD e componentes chaves

	Planejamento	Obtenção	Armazenamento e compartilhamento	Manutenção	Aplicação	Descarte
Dados	Quais são os objetivos de negócio? Quais são as regras de negócios? Quais são os padrões de dados?	Quais dados são adquiridos? Quais dados são inseridos no sistema?	Quais dados são guardados? Quais dados são compartilhados? Quais são os dados principais a serem copiados para uma recuperação rápida?	Quais dados são modificados no sistema? Quais dados são modificados antes de serem migrados, integrados ou compartilhados?	Quais informações são disponíveis para uso? Quais informações são necessárias para suportar transações, métricas, conformidades, requisitos e outros objetivos?	Quais dados precisam ser arquivados? Quais dados precisam ser apagados?
Processos	Quais são os treinamentos e estratégias de comunicação?	Como são adquiridos os recursos internos e externos? Como os dados são inseridos no sistema? Quais são os gatilhos (<i>triggers</i>) para se inserir novos registros?	Quais são os processos para armazenamento? Quais são os processos para compartilhamento?	Como os dados são atualizados? Como os dados são monitorados para detectar modificações? Como são os padrões de monitoramento?	Como os dados são utilizados? Como são acessados e assegurados os dados? Como os dados se tornam acessíveis aos usuários?	Como os dados são armazenados? E deletados? Como são gerenciados os processos e locais de armazenamento de arquivos?
Pessoas e organizações	Quem identifica os objetivos da empresa? Quem são os responsáveis pela definição das regras de negócios? Quem gerencia os envolvidos nessa fase?	Quem adquire informações sobre as fontes? Quem insere novos registros no sistema?	Quem faz o papel de suporte no armazenamento da tecnologia? E no compartilhamento?	Quem decide o que deve ser atualizado? Quem realiza modificações nos dados? Quem deve ser informado quando há modificações?	Quem acessa diretamente os dados? Quem usa a informação?	Quem estabelece a política de retenção de informação? Quem armazena os dados? Quem tem permissão para apagar os dados? Quem precisa ser informado?
Tecnologia	Qual é a arquitetura de alto nível e tecnologia que dá suporte ao negócio?	Como a aplicação é usada para criar uma nova informação?	Qual é a tecnologia que armazena o dado? E a que compartilha?	Como o dado é mantido e atualizado no sistema?	Como as regras de negócio são aplicadas na arquitetura da aplicação?	Como as informações são apagadas do sistema? E como são arquivadas?

Fonte: Adaptado de McGilvray (2008)

Tabela 8 - Exemplo de utilização da matriz de interação do POSMAD - Parte 1

	Planejamento	Obtenção
Dados	<p>O objetivo da empresa é atender rigorosamente as necessidades das editoras (vendedores) e leitores (compradores). As regras de negócio estão especificadas no documento de requisitos da aplicação. Os dados devem seguir os padrões estabelecidos pela ficha catalográfica definida em conjunto com os interessados envolvidos.</p>	<p>Dados obtidos por carga de dados:</p> <ul style="list-style-type: none"> • Codisbn, nome, codautor, codeditora, genero, preço de compra, numerodepaginas, idiomas, dimensoes, peso e tipodecapa. <p>Dados obtidos por interação com o sistema:</p> <ul style="list-style-type: none"> • Taxa de acréscimo de preço, aumento no estoque e comentários.
Processos	<p>Os usuários internos da empresa recebem treinamento sempre que uma nova funcionalidade é acrescentada ou quando há alguma mudança de regra. Os responsáveis por determinar quando e como o treinamento será realizado é a equipe de TI.</p>	<p>Informações inseridas através da interação das diferentes classes de usuário com a aplicação utilizam formulários padronizados.</p> <p>Informações por carga são automáticas.</p> <p>Informações sobre quantidade em estoque são obtidas com a atualização automática do valor já existente no banco e o acréscimo de novas entradas pela equipe de logística ou decremento automático que ocorre na finalização de compras por clientes.</p>

Pessoas e organizações	As regras de negócio são definidas pela equipe de TI a partir da interação com os diversos perfis de interessados. Além disso, é a equipe de TI quem é responsável por gerenciar as mudanças solicitadas e evoluções necessárias. Os perfis de usuários são divididos em:	Editoras:
	<ul style="list-style-type: none">• Clientes: quem utiliza o site como consumidor de livros.• Equipe de logística: responsáveis pelo setor de entregas e recebimentos de livros.• Equipe de finanças: relacionados aos valores dos produtos.• Equipe de TI: relacionados ao desenvolvimento e manutenção do sistema.• Editoras: este grupo acessa o sistema apenas para envios de informações sobre os livros vendidos.	<ul style="list-style-type: none">• Informações dos livros são obtidas a partir do envio por parte das editoras de arquivos XML.
		Equipes internas:
		<ul style="list-style-type: none">• A equipe do financeiro incluir valores de taxa de acréscimo nos preços.• A equipe de logística inclui valores referentes a entrada de novos livros no estoque.
		Clientes:
		<ul style="list-style-type: none">• Os clientes têm permissão para adicionar comentários sobre os livros.• As compras que decrementam a quantidade em estoque são realizadas por clientes.

Tecnologia	<p>A empresa utiliza e-mails e a ferramenta Trello para organização e comunicação das equipes de funcionários.</p> <p>A comunicação com os clientes é definida via SAC disponível no site da loja. Os clientes podem reportar problemas e enviar sugestões de melhorias.</p>	<p>As informações técnicas dos livros obtidas pela editora são obtidas em arquivos XML</p>
------------	--	--

Tabela 9 - Exemplo de utilização da matriz de interação do POSMAD - Parte 2

	Armazenamento e compartilhamento	e Manutenção
Dados	<p>Todos os dados são armazenados no banco de dados. Informações técnicas dos livros, de compras, comentários e sugestões de clientes são, também, armazenadas em <i>backup</i>. Os dados são acessados apenas na aplicação <i>web</i> da <i>e-commerce</i>, não sendo compartilhada em outras aplicações.</p>	<p>O preço inicial é inserido pelo funcionário da empresa e depois as modificações são executadas automaticamente pela aplicação sempre que a taxa de preço for modificada. A quantidade em estoque é modificada quando a equipe de logística acrescenta os novos livros recebidos e quando um cliente finaliza compras. Informações técnicas só são modificadas por perfis de usuários específicos para correção de eventuais erros de cadastro.</p>
Processos	<p>A empresa utiliza o site como meio de compartilhamento de dados. Cada grupo de usuários têm níveis de permissão diferentes e acessos a diferentes páginas.</p>	<p>Os dados atualizados via interação com clientes são:</p> <ul style="list-style-type: none"> • Comentários sobre os livros. <p>Os dados atualizados via gatilhos são:</p> <ul style="list-style-type: none"> • Quantidade em estoque e preço. <p>Os eventos que disparam a atualização dos dados via gatilho são:</p> <ul style="list-style-type: none"> • Finalização de compras e modificações na taxa de preço. <p>Em relação aos dados inseridos pelo usuário, a equipe de finanças modifica a taxa de preço utilizada para o cálculo do preço final.</p> <p>Os dados atualizados via carga de dados são:</p> <ul style="list-style-type: none"> • Informações técnicas enviadas pelas editoras em arquivos XML.

Pessoas e organizações	A equipe de TI é a responsável pelo suporte da tecnologia usada para armazenamento.	Novas atualizações de preço são anunciadas pela equipe financeira. Modificações nos dados de forma automática, como decremento de quantidade em estoque a cada compra, são implementadas pela equipe de TI. Apenas perfis de usuários específicos podem alterar informações já inseridas para correções de eventuais erros de cadastro.
Tecnologia	Os dados são armazenados em um banco de dados relacional RELBD e a aplicação foi desenvolvida usando a linguagem de programação Ruby.	Os dados são armazenados em tabelas relacionais. As atualizações são feitas pelo site com permissões diferentes para diferentes grupos de usuários.

Tabela 10 - Exemplo de utilização da matriz de interação do POSMAD - Parte 3

	Aplicação	Descarte
Dados	<p>Os clientes acessam informações técnicas, como título, autor, descrição e preço.</p> <p>Os funcionários da empresa utilizam as informações do sistema para relatórios e têm acesso a dados relacionados a</p>	<p>Mesmo que não exista mais quantidade em estoque todas as informações técnicas do livro devem ser mantidas.</p> <p>Comentários sobre o livro feitos pelos clientes podem ser deletados apenas pelo autor do comentário.</p>
Processos	<p>Os clientes acessam os dados para efetuar compras de livros, comparar produtos e preços e avaliar livros adquiridos.</p> <p>As equipes internas acessam os dados para criação de relatórios e consultas.</p> <p>A permissão de acesso é diferente para cada perfil de usuário, dependendo do objetivo de aplicação do mesmo.</p>	<p>Para deleção de comentários o cliente cadastrado que inseriu o comentário utiliza o sistema, que fará a busca na relação de comentários a partir do ISBN do livro, do código do cliente e uma identificação do comentário.</p> <p>Apenas um perfil de administrador específico pode remover informações técnicas dos livros já inseridas para correções de eventuais erros de cadastro.</p>

Pessoas e organizações	<p>O cliente, que utiliza os dados técnicos para consultas, só tem acesso à essas informações como leitura. O cliente também acessa informações sobre avaliações e comentários.</p> <p>A equipe de logística tem acesso a informações técnicas dos livros e quantidade em estoque.</p> <p>A equipe de finanças tem acesso a informações técnicas dos livros, preço e taxa de acréscimo no preço.</p> <p>As editoras têm acesso ao sistema para envio de arquivos XML dos livros comprados pela <i>e-commerce</i>.</p>	<p>O cliente pode modificar e apagar avaliações e comentários de sua autoria. Informações técnicas só podem ser removidas por usuários com perfil de administrador.</p> <p>As equipes internas da empresa têm permissão para modificar os valores dos dados com que se relacionam, mas não têm permissão para deleção.</p>
Tecnologia	<p>O sistema solicita nome de usuário e senha para seções que permitam alteração nos dados ou acessem informações específicas para a empresa. O usuário não cadastrado tem acesso limitado.</p>	<p>É utilizado sistemas de segurança para bloquear acessos de usuários não cadastrados e gerenciar permissões.</p>

3.2.4. Definir as dimensões e métricas que serão analisadas

Durante a atividade de detalhamento dos temas estratégicos selecionados são observadas as necessidades de negócios que o tema deve compreender. Para analisar os dados de acordo com essas necessidades devem ser selecionadas dimensões específicas dos dados que consigam abranger características relevantes para a análise das necessidades de negócio solicitadas.

As avaliações fornecem uma imagem da qualidade real dos dados e informações, sendo que um dos maiores benefícios da avaliação das dimensões é obter uma evidência concreta dos problemas encontrados anteriormente, além de trazer informações básicas para investigação de causas, erros nos dados e prevenção de erros de dados futuros (MCGILVRAY, 2008).

Esta atividade tem como objetivo selecionar quais dimensões serão avaliadas, definir quais serão as métricas utilizadas para essa avaliação, definir quais valores obtidos pelas métricas serão considerados satisfatórios para que o dado seja considerado de qualidade e elaborar um plano para aplicação das métricas.

Primeiramente, é preciso compreender quais características dos dados representa cada dimensão. Assim, será possível relacionar tais dimensões aos problemas e oportunidades identificados durante o estudo do ciclo de vida das informações realizado anteriormente.

Na literatura existem autores que destacam várias dimensões diferentes, sendo que cada uma das dimensões de qualidade de dados requer diferentes ferramentas, técnicas e processos para que sejam mensuradas. Isso significa que necessitam de níveis diferentes de recursos como tempo, recursos financeiros e trabalho humano. Portanto, é importante avaliar quais dimensões são adequadas para as necessidades que os temas abordam e compreender o que será obtido na avaliação de cada (MCGILVRAY, 2008).

McGilvray (2008) aponta que ao se escolher as dimensões que serão avaliadas deve-se considerar apenas dimensões que são possíveis ou práticas de serem avaliadas, já que muitas dimensões podem requerer custos proibitivos para a organização.

Definidas as dimensões que serão avaliadas, devem ser estabelecidas métricas que devolvam resultados coerentes que conduzirão as ações de melhoria (MCGILVRAY, 2008). Logo, é possível definir o que significa qualidade de dados para seus produtos de informação no contexto considerado, produzindo regras de integridade de dados. Essas regras geralmente denotam as condições de erro durante as análises (BATINI et al., 2009).

Desta forma, é necessário definir estratégias para mensurar a qualidade dos dados em relação às regras de integridade definidas. Para tanto, uma ou mais métricas são associadas às dimensões separadamente. Para cada métrica, um ou mais métodos de medição devem ser realizados para fornecer informações de (i) onde a medição ocorre, (ii) quais dados são incluídos, (iii) o dispositivo de medição e (iv) a escala em que os resultados são relatados (BATINI; SCANNAPIECO, 2006).

Cada organização define quais são as dimensões importantes para seus objetivos e deve definir também quais componentes dos dados (atributos, relações, tuplas, etc.) devem ser estudados para retornar valores para essas dimensões. Muitas vezes, identificar quais componentes serão analisados se torna uma tarefa mais difícil do que selecionar a métrica, já que geralmente as métricas se resumem às fórmulas de proporção; ou seja, são derivadas da verificação da razão entre os dados armazenados que correspondem à regra de integridade desejada e o número total de dados armazenados (LEE et al., 2009). Três formas de medição de dimensões de qualidade foram identificadas por Pipino *et al* (2002).

- *Simple ratio* (razão simples), que mede a relação entre os registros de um componente selecionado que respeitam ou não uma regra sobre o número total de registros;
- Operador de mínimo e máximo, que lida com dimensões que requerem a agregação de múltiplas variáveis de qualidade de dados;
- Média ponderada, que é utilizada quando uma organização, por exemplo, tem um bom entendimento da importância de cada avaliação que compõe a avaliação global de uma dimensão, então, uma média ponderada dessas avaliações é apropriada.

Em geral, as metodologias consideram as atividades de análise dos dados e a medição como fase de avaliação sendo realizadas a partir de diferentes abordagens. Considerando as possíveis dificuldades para traduzir as dimensões escolhidas em métricas que representem qualidade, as diretrizes na Tabela 11, representadas pelo acrônimo RUMBA (*Reasonable, Understandable, Measurable, Believable, Achievable*), podem ser utilizadas para avaliar as métricas escolhidas (KOVAC, LEE; PIPINO, 1997):

Tabela 11 - Diretrizes de seleção de métrica

R	A métrica é razoável? (<i>Reasonable</i>)
U	A métrica é compreensível? (<i>Understandable</i>)
M	A métrica é mensurável? (<i>Measurable</i>)
B	A métrica é acreditável? (<i>Believable</i>)

Estabelecida a métrica para cada dimensão, é necessário planejar como será realizada a captura dos dados que serão avaliados, qual será o método de acesso e as ferramentas necessárias, o formato de saída dos dados, o método de amostragem, a estratégia para realização da medição e as pessoas envolvidas na extração dos dados e avaliação (MCGILVRAY, 2008).

3.3. Avaliação, Análise e Ação

Na etapa de avaliação são realizados todos os procedimentos que foram definidos na etapa anterior, como aplicação de questionários, formulários, relatórios, implementação de códigos e demais métodos usados para obtenção de valores utilizados nas métricas. Nesta etapa, são obtidos os valores definidos nas métricas que ficam aptos a serem analisados.

Na análise, são investigadas as causas que justifiquem os resultados obtidos. Para isso, as etapas do ciclo de vida assim como os componentes relacionados são analisados de modo que seja possível identificar os pontos que necessitam de melhorias.

É importante considerar que algumas etapas do ciclo de vida dos dados, mesmo não tendo relação direta com o grau de qualidade das dimensões, podem representar oportunidades de melhorias.

Com base nesta análise são coletadas sugestões de atividades e modificações no processo que possam aprimorar as dimensões dos dados considerados relevantes. As propostas são analisadas buscando entender se o grau de melhoria que promovem compensam os custos que exigem; sempre relacionando-as com as limitações da organização e as exigências dos usuários consumidores dos dados. O resultado da análise será a base para o planejamento da aplicação dessas atividades de melhorias no processo sendo necessário estabelecer o modo de aplicação e os responsáveis pela execução.

Na ação, tendo em vista os resultados dos procedimentos da avaliação e análise, é necessário realizar as atividades de melhorias. Para esta etapa, Batini e Scannapieco (2006) apontam as atividades de aprimoramento mais comuns:

- Aquisição de novos dados: atualizar o banco com dados de qualidade superior àqueles já existentes;
- Normalização e reformulação dos dados: analisar dados já existentes de acordo com os padrões e formatações previamente definidos;

- Identificação de objetos: verificar entre os arquivos de uma mesma tabela quais representam o mesmo objeto no mundo real;
- Integração de dados: definir operações que buscam apresentar os dados não heterogêneos de forma unificada;
- Confiabilidade da fonte: classificar as fontes dos dados de acordo com a qualidade dos dados providos por elas;
- Localização de erro: localizar arquivos que não atendem às regras semânticas estabelecidas;
- Correção de erro: depois de identificar as tabelas a serem conferidas, definir as regras de semântica e as localizações dos erros, reparar os valores incorretos para que respeitem as regras;
- Otimização de custo: aperfeiçoar uma meta de qualidade de dados de acordo com o objetivo de custo.

Algumas outras atividades, como cruzamento e limpeza de esquemas, e análise de perfis podem ser menos frequentes, mas também fazem parte das atividades relacionadas à qualidade de dados.

Depois de realizadas as ações de acordo com os resultados obtidos na análise e avaliação, as dimensões devem ser reavaliadas de maneira que seja possível verificar se as ações trouxeram, de fato, alguma melhoria. Após esta análise, é possível que o grau de qualidade exigido inicialmente tenha sido alcançado de forma satisfatória. Isso implica que as ações de melhoria propostas modificaram o sistema e devem ser documentadas. É importante registrar quais foram os resultados, as situações nas quais as ações de melhoria foram definidas e estabelecer mecanismos para que estas ações se sustentem de forma eficiente.

4. Data Quality Simulator

O jogo *Data Quality Simulator* foi desenvolvido utilizando a ferramenta RPGMaker MV, disponível pela empresa japonesa Degica (2016), em versão WEB e *desktop* para o sistema operacional Windows.

O primeiro passo e primordial foi definir qual conteúdo seria abordado no jogo. Devido à extensão de tópicos e definições que a metodologia de avaliação de qualidade de dados abrange, o jogo mantém o foco nos aspectos técnicos da metodologia; logo, a etapa de preparação, por considerar questões gerenciais, não foi considerada. Esta versão do jogo aborda as atividades: estudo do ciclo de vida da informação, o uso da matriz de interação do POSMAD e a identificação de dimensões.

A síntese do jogo, que será detalhada posteriormente neste capítulo, é a apresentação de um ambiente virtual que simula um ambiente de trabalho de uma organização atuante no setor da informação, sendo seu plano de fundo uma *e-commerce* de venda de livros. No cenário, além do jogador, são representados NPC (Personagens não jogáveis, do inglês *Non-Player Character*) e objetos característicos. A primeira meta do jogador é preencher a matriz de interação do POSMAD com informações coletadas a partir da interação com os NPC e objetos do cenário. Conforme é preenchida a matriz do POSMAD, se torna possível identificar problemas de qualidade de dados e assim definir as dimensões mais apropriadas de acordo com os conhecimentos obtidos durante as interações anteriores.

O desenvolvimento do jogo teve início com a pesquisa na literatura buscado aprofundar o conhecimento sobre jogos sérios, teorias de aprendizagem, características motivacionais e estruturais de jogos que pudessem ser aplicadas. Previamente a implementação do jogo, foi realizada também uma pesquisa de opinião de maneira que os requisitos do jogo pudessem ser definidos de forma mais apropriada.

Após a análise de opinião em relação a jogos foi idealizada a estrutura do jogo. Foram definidas as metas de aprendizagem que se espera alcançar com o jogo e relacionadas aos estados presentes na estrutura temporal do jogo planejado. Após a adequação das metas o jogo foi implementado utilizando a ferramenta definida.

4.1. Pesquisa de opinião em relação a jogos

Além das motivações intrínsecas apresentadas na literatura, a fim de encontrar outros componentes presentes em jogos que pudessem ser considerados no desenvolvimento de um jogo sério engajador foi realizado uma pesquisa sobre o interesse das pessoas na área. O

instrumento de coleta definido foi um questionário com objetivo de avaliar a opinião da amostra sobre a utilização de jogos educacionais, além de obter detalhes sobre características de jogos que pudessem ser implementadas na aplicação a ser desenvolvida.

O questionário foi definido e dividido em seções, detalhadas no decorrer deste capítulo em conjunto com a análise das respostas. Foi utilizado os formulários disponibilizados pelo Google. Após ser disponibilizado e divulgado por duas semanas através de redes sociais em grupos da Universidade Federal de Itajubá, grupos relacionados a jogos e e-mails de divulgação acadêmicos, foram obtidas 160 respostas.

A primeira seção, composta por questões de aspecto pessoal para identificação do perfil da amostra, é explicitada na Tabela 12.

Tabela 12 – Questões pessoais

Questões	Alternativas possíveis
Questão A - qual a sua idade?	Resposta única: <ul style="list-style-type: none"> ● Menos de 15 anos, ● Entre 16 e 20 anos ● Entre 21 e 25 anos ● Entre 26 e 30 anos ● Mais de 31 anos
Questão B - Qual a sua ocupação?	Possível selecionar múltiplas respostas: <ul style="list-style-type: none"> ● Desenvolvedor de jogos, ● Estudante, ● Pesquisador, ● Professor. ● Profissional no mercado de trabalho, ● Outro.
Questão C - Você é da área de TI?	Resposta única: <ul style="list-style-type: none"> ● Sim ● Não.
Questão D - Já teve contato com jogos educacionais ou de treinamento?	Resposta única: <ul style="list-style-type: none"> ● Sim ● Não.

Dos entrevistados 43,1% estão na faixa etária de 21 e 25 anos, 75% são estudantes e 61,3% são da área de TI, como pode ser observado nos gráficos das Figura 3, 4 e 5,

respectivamente. Isso demonstra que as respostas obtidas são relevantes para este trabalho na busca por componentes que capturem a atenção do jogador, já que o jogo tem como proposta o ensino de conteúdos relacionados à qualidade de dados para profissionais da área de tecnologia da informação, colocando-os na posição de aprendizes de uma nova metodologia.

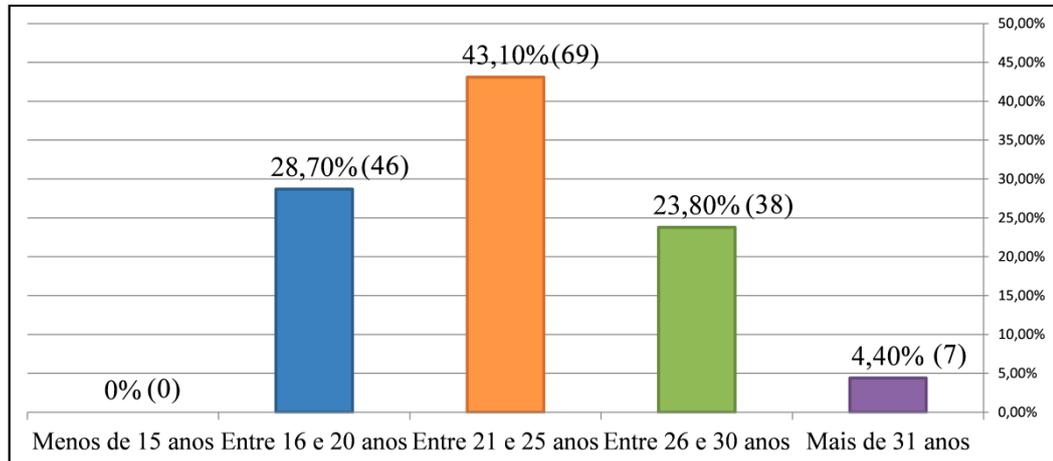


Figura 3 – Questão A - Faixa etária

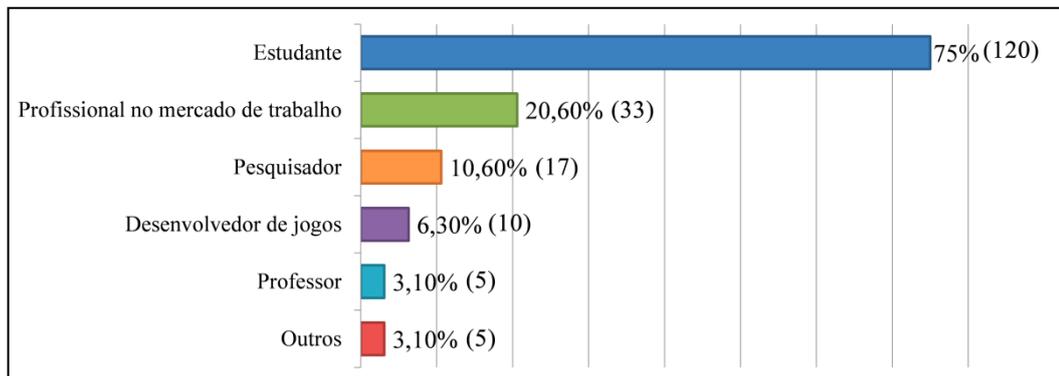


Figura 4 – Questão B - Área de atuação

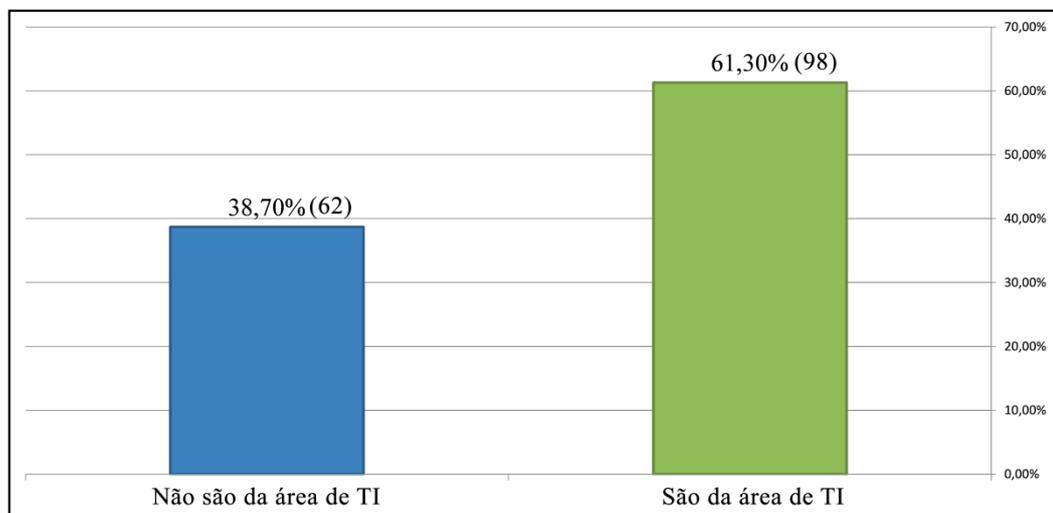


Figura 5 – Questão C - Porcentagem de entrevistados da área de TI

Em relação à última questão da seção de análise de perfil (Tabela 12), 57,5% (92 participantes) já tiveram contato com jogos educacionais ou de treinamento. Para essa parcela de participantes, foi dado acesso à segunda seção da pesquisa na qual foram apresentadas questões específicas sobre jogos educacionais ou de treinamento, descritas na Tabela 13.

Tabela 13 – Questões sobre jogos educacionais ou de treinamento

Questões	Alternativas possíveis
Questão E - Qual a finalidade do jogo educacional utilizado?	Possível selecionar múltiplas respostas: <ul style="list-style-type: none"> • Ensino de um novo conteúdo, • Material extra para estudo, • Motivacional, • Revisão de conteúdo já apresentado, • Simulação e treinamento, • Outro (Qual?).
Questão F - Qual foi a duração em que o jogo educacional em questão foi utilizado?	Resposta única: <ul style="list-style-type: none"> • Um dia, • De um a sete dias, • De sete dias a quinze dias, • De quinze dias a um mês, • Mais de um mês.
Questão G - Você acha que o tempo foi suficiente?	Resposta em escala linear de 1 a 5, com: <ul style="list-style-type: none"> • 1 – Não foi suficiente, • 5 – Foi mais tempo que o necessário.
Questão H - Além do jogo, você usou material para complementar o estudo?	Resposta única: <ul style="list-style-type: none"> • Sim, • Não.

Questão I - Se sim, que tipo de material? Possível selecionar múltiplas respostas:

- Não houve outro material,
- Apresentações de slides,
- Material textual (Livros, apostilas...)
- Listas de exercícios,
- Vídeos,
- Outro (Qual?).

Questão J - O jogo educacional que usou foi avaliado? Se sim, de quais formas? Possível selecionar múltiplas respostas:

- Não houve avaliações,
 - Foi avaliado se o jogo conseguiu facilitar seu aprendizado,
 - Foi avaliado se o jogo funcionava,
 - Foi avaliado se você gostou do jogo,
 - Foi avaliado se você teve algum problema com a interface do jogo,
 - Outro (Qual?)
-

Na segunda seção, a questão E sobre a finalidade dos jogos educacionais e de treinamento foi feita de maneira a analisar se a proposta do jogo *Data Quality Simulator* está alinhada a experiência dos participantes. Foi observado que 54,3% das 92 pessoas que já tiveram contato com jogos educacionais utilizaram o jogo para simulações e treinamento, 46,7% como material extra para estudo e 45,7% para ensino de um novo conteúdo, como apresentado no gráfico da Figura 6.

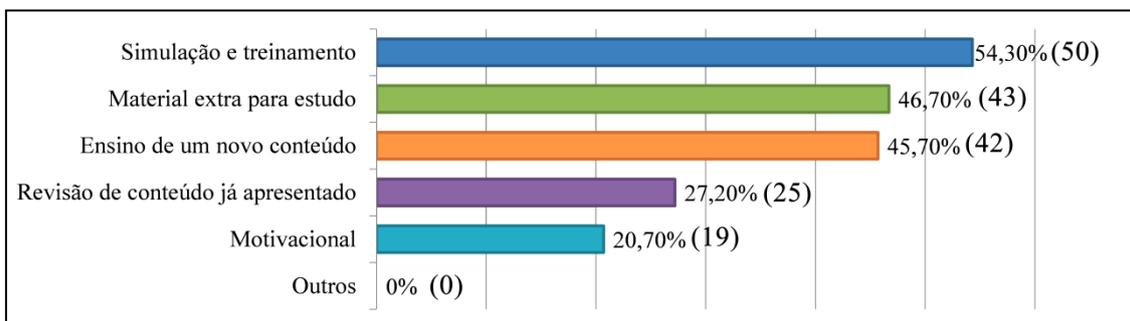


Figura 6 – Questão E - Utilização dos jogos educacionais e de treinamento

O jogo *Data Quality Simulator* tem como principal objetivo ser um ambiente onde o aprendiz pode simular os conteúdos da metodologia de análise de qualidade de dados e absorver o conteúdo de maneira mais prática, o que se enquadra nas experiências dos

entrevistados. Porém, a utilização do jogo de treinamento como material extra pode ser impactante na forma de apresentação do jogo ao público-alvo.

Entre os entrevistados que utilizaram jogos de treinamento, 68,5% utilizaram o jogo como material complementar para estudo, o que pode indicar uma falta de confiança no jogo como único método de aprendizagem.

Nesta seção do questionário também foi identificada a satisfação dos entrevistados em relação ao tempo de uso do jogo educacional. O gráfico da Figura 7 apresenta as faixas de tempo em que os jogos educacionais e de treinamento foram utilizadas pelas pessoas entrevistadas. Na Figura 8 o gráfico relaciona, para cada faixa de tempo em que o jogo foi utilizado, a porcentagem de pessoas que consideraram o respectivo período não suficiente, pouco suficiente, suficiente, pouco mais que o necessário ou mais tempo que o necessário.

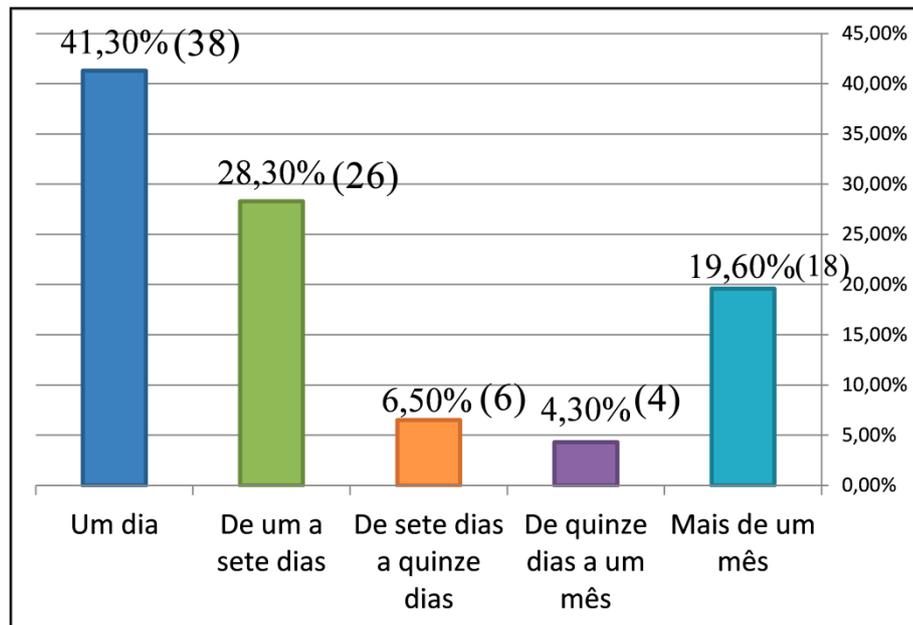


Figura 7 – Questão F - Faixa de tempo em que os jogos foram utilizados

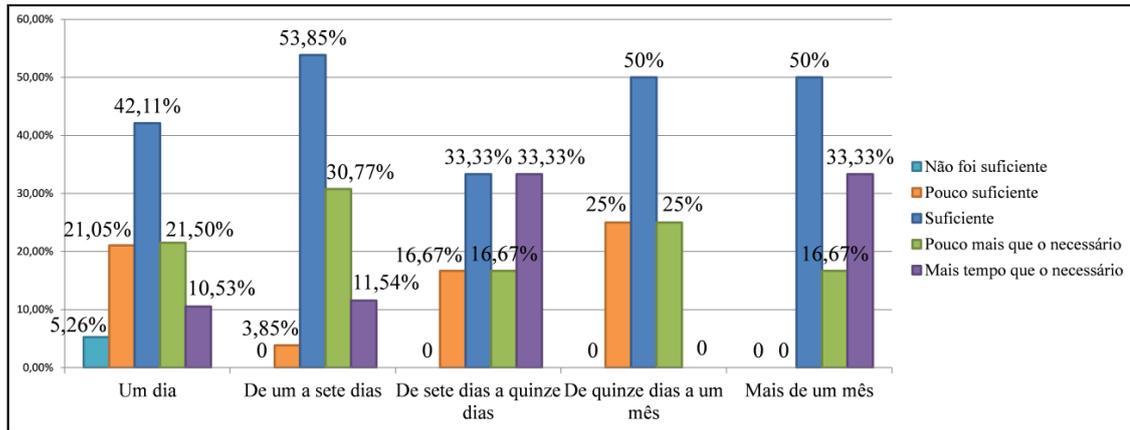


Figura 8 – Questões F e G – Relação tempo de uso e satisfação do aprendiz

Nota-se que, para os entrevistados que tiveram o tempo de utilização de um jogo no período de apenas um dia, o tempo foi insatisfatório para alguns aprendizes 26,31%. Por outro lado, 50% disseram que mais de um mês pode ser considerado um tempo excessivo. Baseado na observação do gráfico da Figura 8, o jogo foi idealizado para ser utilizado por sete dias; sendo que disponível para os aprendizes pelo período de três semanas para que melhor adequem essa faixa de tempo dentro do período disponível. O critério utilizado para a escolha do tempo de uso foi encontrar a faixa de tempo com maior número de respostas em suficiente, seguido de menor número de respostas extremas (não foi suficiente e mais do que o necessário).

Ainda na segunda seção foi questionado se o jogo educacional utilizado foi avaliado seja avaliação funcional, de interface ou de aprendizado. Muitos trabalhos encontrados na literatura pressupõe a capacidade dos jogos como instrumento de aprendizagem; porém, não realizam avaliações (MORENO-GER et al., 2014). 51,1% dos entrevistados responderam que o jogo utilizado não foi avaliado, como apresentado no gráfico da Figura 9, o que corrobora para a afirmação citada.

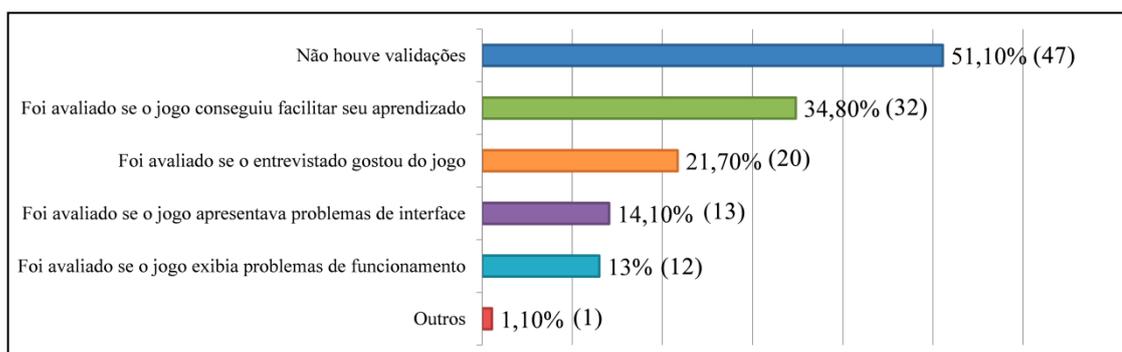


Figura 9 – Questão J - Validações do jogo

A terceira seção do questionário, descrita na Tabela 14, é composta por questões relacionadas aos hábitos e preferências dos entrevistados em relação a jogos de modo geral, não necessariamente jogos considerados jogos sérios. Esta seção é aberta para todos os entrevistados.

Tabela 14 – Questões gerais sobre jogos

Questões	Alternativas possíveis
Questão K - Quais gêneros de jogos você tem preferência?	Possível selecionar múltiplas respostas: <ul style="list-style-type: none"> • Ação, • Aventura, • Estratégia, • RPG (<i>Role-Playing Game</i>), • Esportes, • Corrida, • MMO (<i>Massively Multiplayer Online</i>), • MOBA (<i>Multiplayer Online Battle Arena</i>), • Simulação, • Outro (Qual?)
Questão L - Quais elementos dessas categorias mais lhe agradam?	Questão aberta.
Questão M - Quantas horas semanais/mensais que costuma jogar (frequência)?	Resposta única: <ul style="list-style-type: none"> • Menos de uma hora por semana, • Uma a três horas por semana, • Três a cinco horas por semana, • Mais de cinco horas por semana.
Questão N - Quanto tempo você costuma jogar direto?	Resposta única: <ul style="list-style-type: none"> • Menos de uma hora seguida, • Uma a duas horas seguidas, • Duas a três horas seguidas, • Mais de três horas seguidas

A partir das 160 respostas obtidas na terceira seção foi possível observar que 73,8% dos entrevistados têm preferência por jogos do gênero estratégia, seguido pelo gênero RPG (Jogo de Interpretação de Papéis, do inglês *Role-Playing Game*) com 69,4% dos entrevistados, como apresentado no gráfico da Figura 10 (lembrando que o entrevistado pode selecionar

mais de uma opção de preferência). O jogo proposto neste trabalho pode ser classificado como sendo um jogo de simulação, estratégia e RPG que são categorias compatíveis com as preferências dos entrevistados.

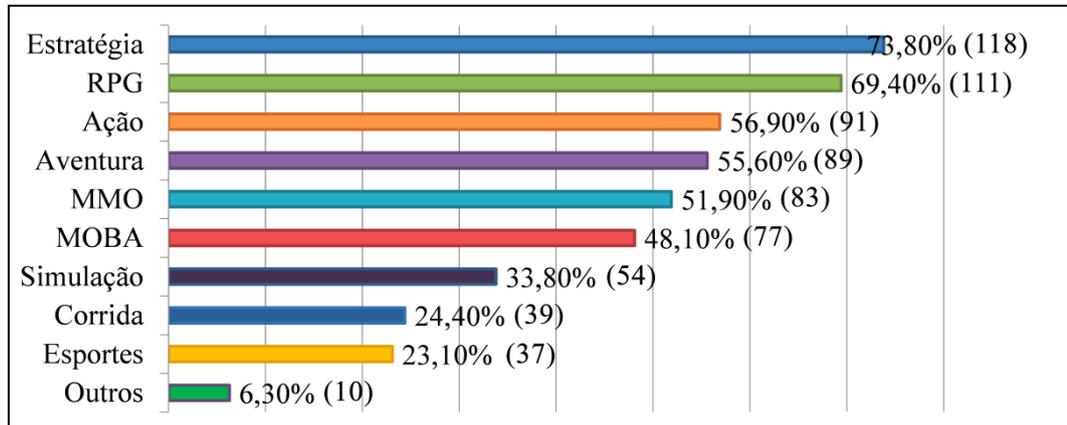


Figura 10 – Questão K - Preferência de gêneros de jogos

Ainda relacionado à questão de preferência de jogos foi perguntado aos entrevistados quais elementos das categorias selecionadas mais lhe agradam. Entre as respostas obtidas dos dois gêneros com maior porcentagem de votos (Estratégia e RPG) se encontram ‘interatividade’, ‘possibilidade de exploração’, ‘ambiente e jogabilidade’, ‘possibilidade de criar estratégias’, ‘sensação de desafio’, e principalmente ‘enredo’, também chamado pelos participantes de história e narrativa. As respostas obtidas reforçam a necessidade de se definir motivações intrínsecas, conforme defendido na literatura.

Outra consideração que o questionário apontou é o tempo médio que um jogador costuma se manter no jogo sem intervalos. O gráfico da Figura 11 mostra que 35% dos entrevistados costumam jogar de uma a duas horas consecutivas, e 29,4% de duas a três horas consecutivas. Desta forma, o tempo médio de uma partida do *Data Quality Simulator* será de uma a duas horas de jogo, sendo possível ao jogador interagir em velocidade normal com todo o conteúdo do jogo nesse tempo. Não foi definido tempo limite para que o jogador finalize o jogo, mas é esperado que todo conteúdo seja acessado em, no máximo, três horas.

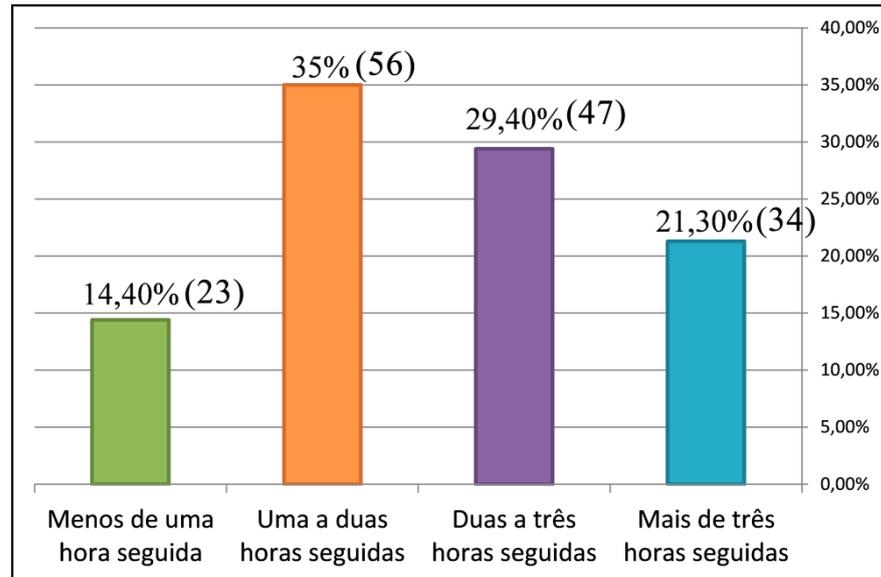


Figura 11 – Questão N - Faixa de tempo em jogo

A quarta seção aborda questões sobre o interesse na aplicação de jogos sérios, sem distinção de áreas de conhecimento, essas questões são indicadas na Tabela 15.

Tabela 15 – Questões sobre interesse em jogos educacionais

Questões	Alternativas possíveis
Questão O - Você acha que o uso de jogos educacionais facilitariam a absorção de conteúdo?	Resposta em escala linear de 1 a 5, com: <ul style="list-style-type: none"> • 1 - Nenhuma ajuda, • 5 - Muita ajuda.
Questão P - Você acha que a utilização de jogos educacionais faz com que os alunos se interessem mais no conteúdo?	Resposta em escala linear de 1 a 5, com: <ul style="list-style-type: none"> • 1 - Nenhuma motivação, • 5 - Muita motivação.

A partir das respostas obtidas na quarta seção do questionário foi possível observar um interesse positivo em relação ao uso de jogos educacionais, sendo que 65,7%% das 160 pessoas entrevistadas acreditam que o uso de jogos educacionais facilita a absorção de conteúdos e 72,6%% acreditam que os jogos aumentem o interesse dos aprendizes ao conteúdo, o que incentiva a criação de jogos sérios. Esses dados são apresentados no gráfico da Figura 12 e 13, respectivamente.

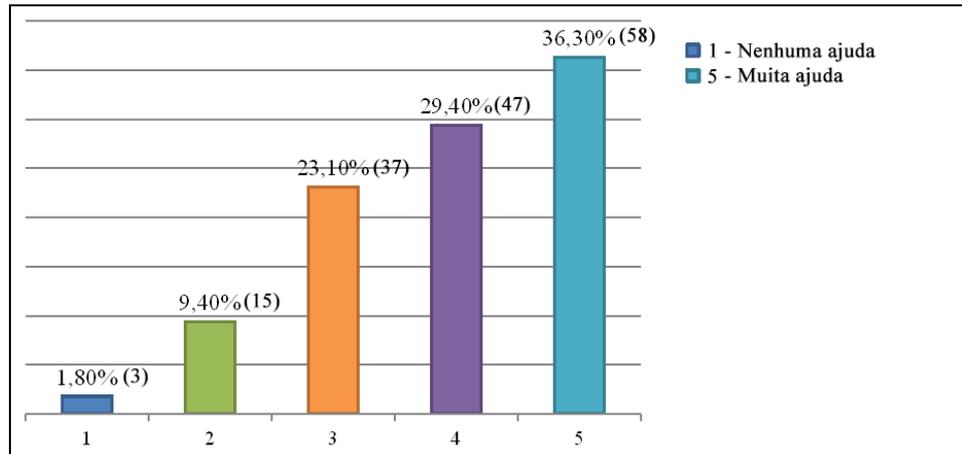


Figura 12 – Questão O - Expectativa de ajuda dos jogos sérios

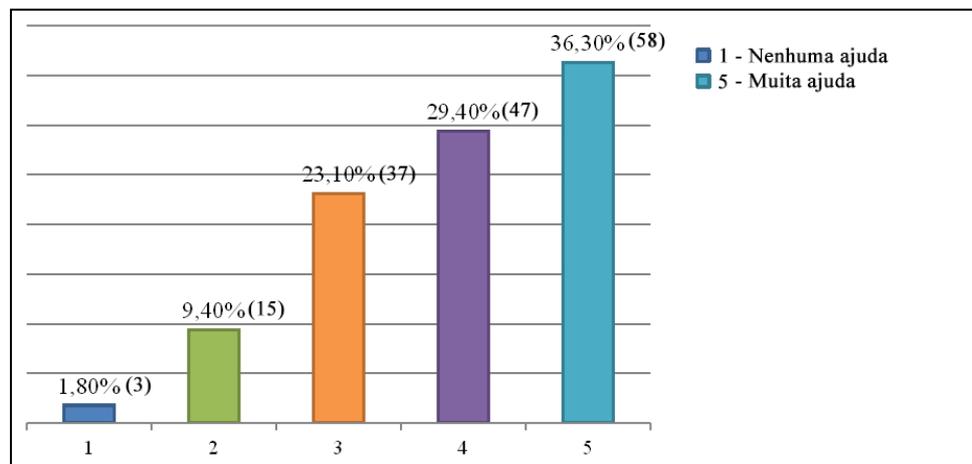


Figura 13 – Questão P - Expectativa de aumento da motivação através de jogos

4.2. Descrição do jogo

Ao iniciar o jogo *Data Quality Simulator* o personagem interpretado pelo jogador, destacado na Figura 14, está em uma sala em frente ao NPC que representa o diretor da empresa fictícia. Esta primeira interação não afetará a pontuação do jogador e será apenas de caráter introdutório. Neste momento, o aprendiz receberá informações sobre seu objetivo no jogo e como alcançá-lo; além de uma breve apresentação dos elementos da metodologia. É permitido que o jogador interrompa a introdução e avance para a próxima fase do jogo, supondo que o jogador já tenha passado por este estágio anteriormente. O NPC informa ao jogador que estará disponível durante todo o jogo com informações mais detalhadas sobre o conteúdo obtido nesse diálogo. O jogador também obtém, nesta interação, uma quantidade de moedas fictícias que serão utilizadas no decorrer do jogo.

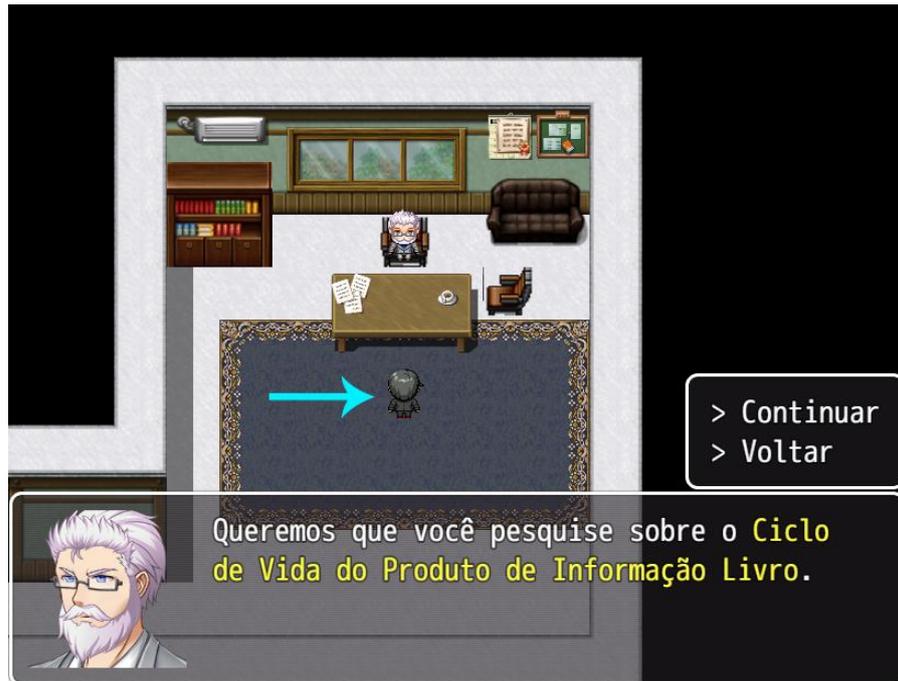


Figura 14 - Introdução do *Data Quality Simulator*

Ao fim da introdução, o jogador aparece em um cenário maior, visualizado na Figura 15, que representa as várias repartições de uma empresa. Neste ambiente o jogador pode se mover livremente e interagir com diferentes NPC através de diálogos. O objetivo é encontrar e resolver os problemas que são identificados a partir das pistas disponibilizadas para preenchimento da matriz POSMAD (canto superior esquerdo na tela, como mostra a Figura 15). Cada célula da matriz está associada a uma pista. Desta forma, ao se obter uma pista, uma célula é preenchida. As pistas são mecanismos para ajudar o jogador a identificar um problema.

Para obter as pistas, são apresentadas uma sequência de pontos de decisão nos quais o jogador deve escolher a pergunta mais apropriada para fazer ao NPC relacionado à etapa e ao componente da matriz do POSMAD.



Figura 15 - Ambiente do *Data Quality Simulator*

No cenário, existem vários NPC distribuídos para ambientação. Foram definidos seis NPC principais, cada um representando uma etapa do POSMAD (planejamento, obtenção, armazenamento e compartilhamento, manutenção, aplicação e descarte). Ao iniciar um diálogo, o jogador é informado se está interagindo com um dos NPC principais e qual etapa do POSMAD o NPC está relacionado. O diálogo é estruturado com quatro pontos de decisão do jogador, cada ponto associado a um componente chave (dados, processos, pessoas e organização e tecnologia). Nesses pontos, são apresentadas quatro questões, o jogador deve selecionar aquela mais apropriada à etapa do POSMAD e ao componente chave específico. Na Figura 16 é ilustrado um exemplo no qual, a etapa do POSMAD relacionada é Armazenamento e Compartilhamento e o componente é Dados, o jogador deve selecionar a questão “Quais dados precisam ser guardados?”.

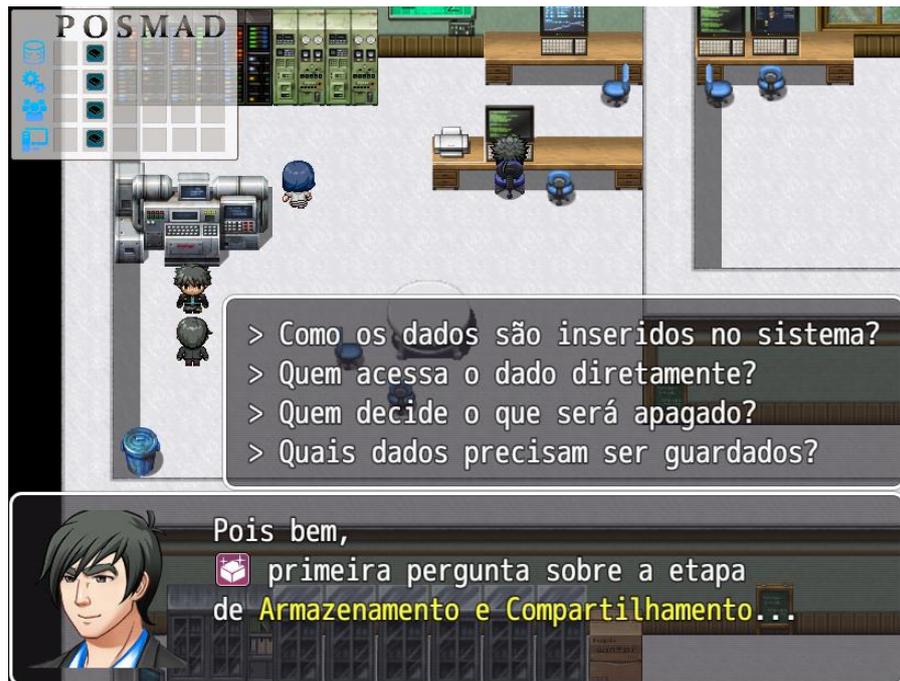


Figura 16 – Exemplo de diálogo com NPC

As opções de questões apresentadas ao jogador foram elaboradas com base nas sugestões de perguntas da tabela do POSMAD apresentadas no capítulo 3. A escolha correta resultará na obtenção de uma pista, representada pelo preenchimento da célula da matriz equivalente a interseção do componente chave e etapa do POSMAD. Depois de obtida uma pista, no decorrer do diálogo, são apresentadas ao jogador informações sobre um determinado problema fictício. Por outro lado, a escolha incorreta deixará uma lacuna na matriz e o NPC indicará apenas que não tem conhecimento sobre o que foi questionado pelo jogador; seguindo, então, para o próximo nível do diálogo.

A estrutura dos diálogos pode ser sistematizada em um grafo, exemplificado na Figura 17. Cada nó do grafo representa um estado do diálogo em que o jogador deve selecionar uma das quatro questões apresentadas. Ao selecionar a opção correta (nó em azul escuro), o jogo segue para o estado no qual o jogador encontra uma pista. Nota-se que se o aprendiz não selecionar a questão correta, ainda assim, terá caminhos possíveis para que obtenha as pistas seguintes relacionadas aos outros componentes chaves e as respectivas etapas.

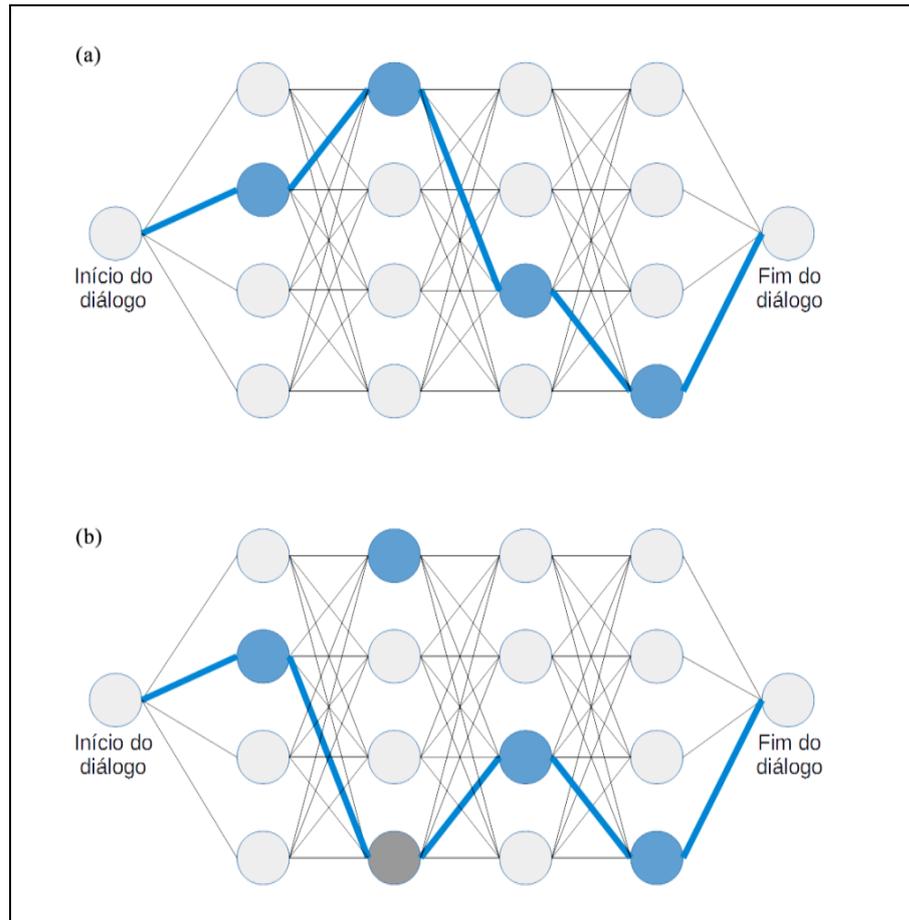


Figura 17 – (a) Caminho com obtenção de todas as pistas (b) Caminho com pista faltante

O objetivo do preenchimento da matriz do POSMAD é observar o ciclo de vida dos dados relacionando todos os fatores que interagem com eles, classificados nos grupos dos componentes chaves, e que podem causar problemas de qualidade de dados. Sendo assim, foi distribuído pelo ambiente NPC, inicialmente invisíveis ao jogador, relacionados a cinco problemas contextualizados com o produto de informação Livro e diferenciados utilizando a fisionomia de um fantasma.

As pistas encontradas no decorrer dos diálogos foram implementadas de modo a associar cada célula da matriz a um ou mais problemas, essa associação é apresentada na Tabela 16. A intenção é que o jogador busque informações suficientes para que os problemas sejam identificados, ficando visíveis no ambiente virtual, como é apresentado na Figura 18. Não é necessário que o jogador obtenha todas as pistas relacionadas a um determinado problema para que o NPC apareça na tela, apenas uma quantidade mínima; porém, o jogador é incentivado a fazer uma busca mais completa, preenchendo toda a matriz, antes de interagir com o problema.

Tabela 16 - Associação de células da matriz e problemas

	Planejamento	Obtenção	Armazenamento e compartilhamento	Manutenção	Aplicação	Descarte
Dados	Problema 4	Problema 3 e 4	Problema 1	Problema 1	Problema 4	Problema 5
Processos	Problema 3	Problema 3	Problema 2	Problema 1	Problema 1 e 2	Problema 5
Pessoas e organização	Problema 4	Problema 3	Problema 2	Problema 1	Problema 2	Problema 5
Tecnologia	Problema 4	Problema 3	Problema 2 e 4	Problema 2	Problema 2	Problema 5

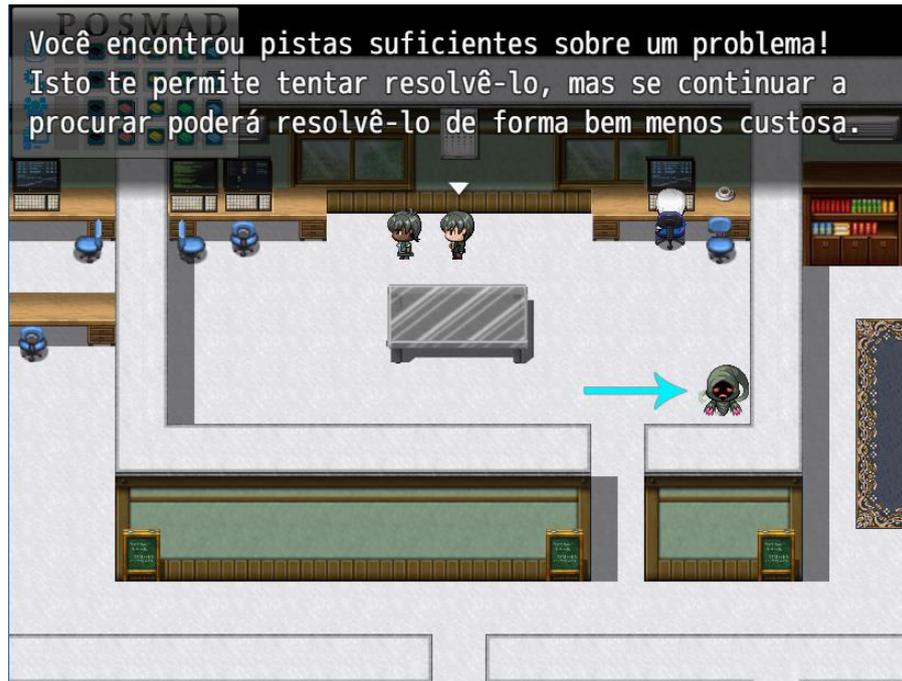


Figura 18 – Problema encontrado

Depois de identificado o problema, o jogador pode dialogar com o NPC que o representa e definir as dimensões apropriadas, como é demonstrado na Figura 19. As informações que permitiram que o problema fosse identificado são relevantes para a definição correta das dimensões.

Podemos considerar que o jogo é dividido em duas etapas. A primeira etapa é definida pela identificação dos problemas relacionados à qualidade de dados, para isso é necessário o preenchimento da a matriz do POSMAD. Como resultado dessa etapa é observado que quanto maior o número de pistas, maior é o número de problemas encontrados. A segunda etapa é a solução destes problemas, que deve ser realizada selecionando corretamente a dimensão correspondente as informações obtidas para cada problema.

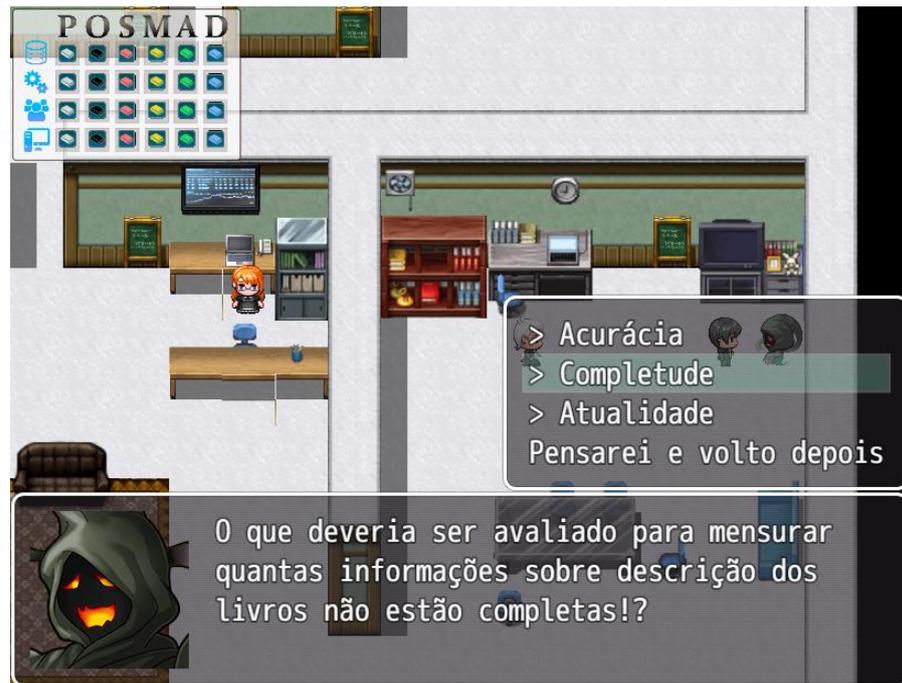


Figura 19 – Interação com problema para definição de dimensão

A partir da resolução dessa etapa, a opção de ação para resolução do problema é disponibilizada; porém, para que possa ser executada, é necessário utilizar um valor de moeda. O valor necessário para executar uma ação que solucionaria um problema é reduzido quanto maior o número de pistas encontradas relacionadas a este problema.

O jogo encerra nas seguintes situações:

- O jogador encontra e consegue solucionar todos os problemas;
- O jogador já interagiu com todos os NPC associados às etapas do POSMAD, porém, ainda existem problemas não visíveis; e,
- O dinheiro fictício acaba e existem problemas não solucionados.

A pontuação final é calculada utilizando o número de pistas encontradas, a quantidade de moedas restantes e o numero de problemas solucionados.

4.3. Mapeamento utilizado no desenvolvimento do jogo

Baseado no trabalho de Airasian e Miranda (2002) foi realizada uma adequação do conteúdo abordado no *Data Quality Simulator* e os níveis de aprendizado apontados na taxonomia de Bloom. Nas Tabelas 17, 18 e 19, está descrito o mapeamento dos tópicos relacionados ao conteúdo proposto do jogo nos níveis de aprendizado definidos no domínio cognitivo. As tabelas se referem a tópicos diferentes, respectivamente: ciclo de vida dos

dados, matriz do POSMAD e dimensões. Para cada nível, é definida a meta de aprendizagem e a estratégia adotada no jogo para atingir esta meta.

Tabela 17 - Relação da Taxonomia de Bloom e as metas e atividades para o Ciclo de Vida da Informação

	Conhecimento	Compreensão	Aplicação	Análise	Síntese	Avaliação
Metas	Saber as fases do ciclo de vida da informação e suas definições.	Entender o que ocorre em cada fase do ciclo.	Perceber como é aplicado a divisão das etapas do ciclo de vida dos dados em um cenário hipotético.	Analisar quais atividades que ocorrem durante o processamento de dados têm relacionamento com determinada etapa do ciclo de vida dos dados.	Relacionar diversas informações do ciclo de vida dos dados para compreender uma informação maior.	Ser capaz de analisar se uma informação sobre ciclo de vida dos dados é relevante ou não.
Atividades	Essa informação é apresentada em forma de texto durante interações com NPC relacionados a cada etapa do ciclo de vida da informação.	Essa informação é apresentada em forma de texto durante interações com NPC relacionados a cada etapa do ciclo de vida da informação.	O ambiente virtual apresenta uma situação hipotética para que o usuário associe às etapas do ciclo de vida dos dados.	Durante diálogos com NPC que representam as etapas do ciclo de vida dos dados, o jogador deve selecionar questões coerentes à etapa definida.	O jogador deve associar as informações de cada etapa do ciclo de vida da informação obtidas durante a busca por pistas para entender o contexto de um problema e solucioná-lo.	O jogador deve avaliar se as informações obtidas nos diálogos são relevantes para entender o ciclo de vida da informação relacionada ao problema e; assim, solucioná-lo.

Tabela 18 - Relação da Taxonomia de Bloom e as metas e atividades para a Matriz do POSMAD

	Conhecimento	Compreensão	Aplicação	Análise	Síntese	Avaliação
Metas	Saber o que representam as células que formam a matriz do POSMAD e como é estruturada.	Compreender como a matriz funciona e é preenchida.	Ser capaz de aplicar a matriz do POSMAD em uma situação hipotética.	Conseguir identificar o papel de cada elemento que compõe a matriz.	Organizar diferentes informações da matriz para obter um conhecimento maior.	Ser capaz de analisar se uma informação sobre a matriz é relevante ou não.
Atividades	A estrutura da matriz é visível na tela e seus componentes são explicitados textualmente.	É solicitado ao aprendiz que busque informações interagindo com o ambiente para preencher a matriz.	O ambiente virtual apresenta uma situação hipotética para que o usuário associe à matriz.	O aprendiz deve selecionar as questões durante os diálogos associando todas as partes que compõem a matriz.	Os elementos que compõem a matriz são apresentados durante a interação com os NPC que representam as etapas, apontando também os componentes chaves. O jogador deve associar as informações de cada elemento da matriz do POSMAD para selecionar corretamente as questões.	Informações fora do contexto do ciclo de vida dos dados são apresentadas durante as interações, porém não são associadas a nenhuma parte da matriz.

Tabela 19 - Relação da Taxonomia de Bloom e as metas e atividades para Dimensões

	Conhecimento	Compreensão	Aplicação	Análise	Síntese	Avaliação
Metas	Saber as dimensões de qualidade apresentadas e suas definições.	Ter um entendimento sobre as dimensões de qualidade.	Ser capaz de identificar uma dimensão através da análise do cenário.	Ser capaz de analisar se uma dimensão tem relação com um problema apresentado.	Conseguir compreender a função das dimensões na análise de qualidade de dados.	Conseguir a selecionar entre diversas dimensões a que melhor tenha relação com um problema apresentado.
Atividades	Essa informação é apresentada em forma de texto durante interações que o jogador deve realizar com NPC associados aos problemas.	A dimensão é associada com informações de um problema para que o jogador contextualize e compreenda sua definição.	O diálogo entre o NPC associado ao problema e o jogador informa as definições das dimensões. O jogador deve analisá-las e selecionar a que melhor se aplica.	O diálogo entre o NPC associado ao problema e o jogador informa as definições das dimensões. O jogador deve analisá-las e selecionar a que melhor se aplica.	Após a identificação da dimensão correta é abordado o motivo para seu uso.	Dimensões fora do contexto são apresentadas para que o jogador avalie se são relevantes ou não para a definição de dimensões.

Em relação ao domínio afetivo e ao engajamento do aprendiz com o jogo, foram consideradas as motivações intrínsecas identificadas na literatura e os resultados do questionário utilizado na pesquisa de opinião.

Seguindo com o planejamento para o desenvolvimento do jogo, foi utilizado o modelo ATMSG (*Activity Theory-based model for serious games*) (CARVALHO et al., 2015) como base para detalhar o modo de execução de cada meta estabelecida nas tabelas anteriores. São definidas, para cada meta, as ações que o jogador deve realizar para que a meta definida seja alcançada e as ferramentas do jogo utilizadas. O fluxograma exibido na Figura 20 representa, de forma simplificada, os estados do jogo durante uma partida.

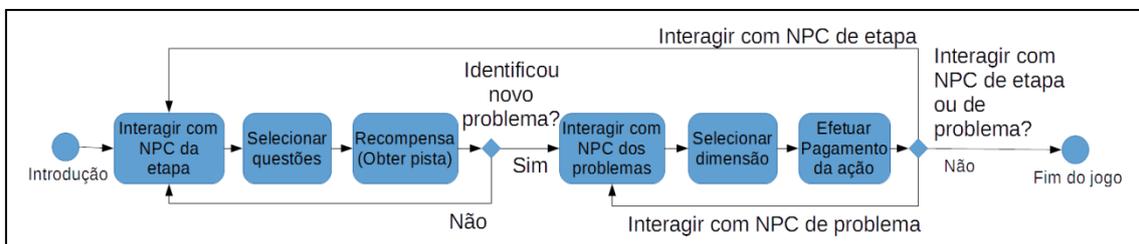


Figura 20 - Fluxograma das etapas do *Data Quality Simulator*

Este fluxograma é utilizado para adequação das metas no decorrer da linha temporal do jogo. Desse modo, se torna mais fácil observar se todas as metas definidas estão presentes. Como resultado é obtido uma tabela que estrutura o jogo e exibe a visão educacional intencionada. A fim de organizar as tabelas resultantes do mapeamento definido, os estados do fluxograma foram estruturados nas células do eixo superior de cada tabela. As Tabelas 20, 21 e 22 apresentam a visão educacional proposta para o *Data Quality Simulator* para cada conteúdo abordado (ciclo de vida, matriz do POSMAD e dimensões).

Tabela 20 - Visão educacional proposta para *Data Quality Simulator* para o Ciclo de Vida da Informação

Interagir com NPC da etapa	Selecionar questões	Recompensa (obter pista)	Interagir com NPC do problema	Selecionar dimensão	Efetuar pagamento da ação
Metas: Conhecimento, Compreensão.	Metas: Análise.	Aplicação,		Metas: Síntese e Avaliação.	
Ferramentas: NPC, objetos, enredo, mensagens.	Ferramentas: NPC, objetos, mensagens.	NPC, enredo,		Ferramentas: NPC, objetos, enredo, mensagens.	
Ação: O jogador deve andar e se aproximar dos NPC e objetos do cenário para obter informações.	Ação: O jogador deve selecionar as questões verificando quais das opções oferecidas têm relação com a etapa definida ou não.			Ação: O jogador deve relacionar as informações que julga relevante para entender o problema e definir a dimensão correta.	

Tabela 21 - Visão educacional proposta para *Data Quality Simulator* para a Matriz de interação do POSMAD

Interagir com NPC da etapa	Selecionar questões	Recompensa (obter pista)	Interagir com NPC do problema	Selecionar dimensão	Efetuar pagamento da ação
Metas de: Conhecimento, Compreensão.	Metas de: Análise.	Aplicação, Metas de: Síntese.	Metas de: Análise e Síntese.	Metas: Síntese e Avaliação.	Metas: Síntese e Avaliação.
Ferramentas: NPC, objetos, mensagens, enredo, imagens.	Ferramentas: objetos, mensagens.	NPC, enredo, pontuação, advertência/parabenização, imagens.	Ferramenta: sistema de avisos de advertência/parabenização, imagens.	Ferramentas: NPC, objetos, enredo, mensagens.	Ferramentas: NPC, objetos, enredo, mensagens.
Ação: A estrutura da matriz é visível na tela e seus componentes são explicitados textualmente.	Ação: O jogador deve selecionar as questões verificando quais das opções oferecidas têm relação com o componente chave ou etapa definida.	Ação: no final da ação o jogador obtém <i>feedback</i> de resposta correta. O jogador então deve observar as informações obtidas com a pista encontrada e relacioná-las com a lacuna preenchida na matriz.	Ação: O jogador deve relacionar as informações que julga relevante para entender o problema e definir a dimensão correta.	Ação: O jogador deve relacionar as informações que julga relevante para entender o problema e definir a dimensão correta.	Ação: O jogador deve relacionar as informações que julga relevante para entender o problema e definir a dimensão correta.

Tabela 22 - Visão educacional proposta para *Data Quality Simulator* para as Dimensões

Interagir com NPC da etapa	Selecionar questões	Recompensa (obter pista)	Interagir com NPC do problema	Selecionar dimensão	Efetuar pagamento da ação
			Metas de: Conhecimento, Compreensão.	Metas de: Aplicação, Análise, Síntese, Avaliação.	
			Ferramentas: NPC, objetos, mensagens.	Ferramenta: NPC, objetos, enredo, mensagens, questionário, sistema de pontuação, <i>feedback</i> .	
			Ação: O conteúdo é incorporado ao texto do diálogo. O jogador deve entender o contexto do problema e as definições de dimensões apresentadas.	Ação: O jogador deve analisar as opções disponíveis e selecionar a dimensão que melhor se aplique ao problema. No final da ação o jogador obtém <i>feedback</i> de resposta corretas/erradas e a necessidade do uso da dimensão.	

Com esse modelo, é possível observar a distribuição das metas de aprendizagem definidas em relação aos níveis de aprendizagem que o jogo se propõe a alcançar. Além disso, pode ser constatado também que todas as metas foram incluídas no projeto de implantação. Assim, o jogo foi desenvolvido tendo uma potencial base para ensino.

5. Metodologia de Avaliação

Na literatura, é possível encontrar diversos trabalhos relacionados aos vários propósitos dos jogos, tanto jogos sérios até jogos de entretenimento. Entretanto, grande parte desses trabalhos apenas especula sobre o potencial quanto à aprendizagem e o engajamento (MORENO-GER et al., 2014) (RIBEIRO et al., 2015) (SAVI et al., 2010).

Moreno-Ger *et al* (2014) apontam que a validação completa de um jogo sério é uma tarefa bastante complicada, que inclui validações de requisitos, verificação da transferência adequada de conhecimento, usabilidade, entre outros. Porém, a utilização de avaliações a respeito de jogos educacionais é o que permite que instrutores identifiquem a qualidade de um jogo e percebam a relevância de seu uso para, então, aplicá-lo em seu projeto de ensino (RIBEIRO et al., 2015). Os resultados dessas avaliações também são relevantes para os desenvolvedores, pois contribuem na identificação de oportunidades de melhoria (SAVI et al., 2010).

Logo, é necessário ter definido “o que” e “como” o jogo será validado para garantir a eficácia de seu uso (ROCHA; BITTENCOURT; ISOTANI, 2015). Neste trabalho, foram avaliados três aspectos: (i) o jogo como instrumento de aprendizagem, presente na maioria dos trabalhos que possuem estratégias avaliativas encontrados na literatura (HENRIQUE et al., 2015); (ii) a satisfação do usuário e (iii) a usabilidade, consideradas avaliações pertinentes a serem realizadas para a melhoria de um software (DIAS et al., 2013).

5.1. Procedimento de avaliação executado

O procedimento de avaliação foi definido a partir de um experimento no qual o jogo seria apresentado, utilizado e avaliado por uma amostra de potenciais usuários.

O processo de avaliação foi executado em três etapas: pré-jogo, utilização do jogo e pós-jogo. Durante a etapa de pré-jogo, foi apresentada uma explanação sobre os conceitos de qualidade de dados e sobre a metodologia descrita no Capítulo 3. Essa apresentação ocorreu em sala de aula permitindo que os alunos levantassem dúvidas e discussões. Imediatamente após a apresentação, foi aplicado um questionário como instrumento de coleta de dados para elaboração de um diagnóstico do conhecimento dos participantes antes da utilização do jogo.

A aplicação de um questionário como ferramenta de obtenção de dados é motivada por ser um método de fácil aplicação e de baixo custo (HENRIQUE et al., 2015).

As questões que compõem o questionário foram analisadas por um especialista, de forma a garantir que fossem suficientemente específicas e abrangentes ao conteúdo didático

definido no jogo. O questionário foi composto de 8 questões. Para as questões 1 à 4, o participante deveria escolher uma única resposta (Tabela 23). Já para as questões 5 à 8, o participante deveria associar as colunas apresentadas em cada questão (Tabela 24). A opção “Não sei a resposta” foi incluída na tentativa de reduzir distorções nos resultados nos casos em que o participante realmente não soubesse a resposta.

Tabela 23 – Questões específicas de conteúdo

Questões	Alternativas possíveis
Questão 1 - Quais são as etapas que dividem o ciclo de vida dos dados e dão origem ao acrônimo POSMAD?	<p>Resposta única:</p> <ul style="list-style-type: none"> ● Planejamento, Obtenção, Serviços, Manutenção, Atualização e Distribuição. ● Planejamento, Obtenção, Armazenamento e Compartilhamento, Manutenção, Aplicação e Descarte. ● Preparação, Obtenção, Serviços, Manutenção, Armazenamento e Compartilhamento e Descarte. ● Planejamento, Obtenção, Serviços, Armazenamento e Compartilhamento, Atualização e Descarte. ● Preparação, Obtenção, Serviços, Manutenção, Aplicação e Distribuição. ● Não sei a resposta.

Questão 2 - O que significa Produto de Informação?	<p>Resposta única:</p> <ul style="list-style-type: none">• É o conjunto de agentes que manipulam os dados durante a trajetória desde sua obtenção até descarte.• O produto resultado do processamento, manipulação e organização de dados que será entregue para o usuário.• É o processo pelo qual os dados passam até serem entregues ao usuário final na posição de consumidor.• São representações do mundo real em um modo que podem ser armazenados, recuperados e processados por um mecanismo de software.• Não sei a resposta.
Questão 3 - Quais são os componentes chaves que você precisa avaliar em cada etapa do POSMAD?	<p>Resposta única:</p> <ul style="list-style-type: none">• Dados, Processos, Pessoas e organização, e Tecnologia.• Processos, Estrutura, Pessoas e organizações, e Máquinas.• Dados, Estrutura, Processos, e Pessoas e organizações.• Dados, Máquinas, Pessoas e organizações, e Tecnologia.• Não sei a resposta.
Questão 4 - O que significa a dimensão acurácia em um projeto de qualidade de dados?	<p>Resposta única:</p> <ul style="list-style-type: none">• Representa o quão suficiente é o dado para a tarefa que ele será usado.• Representa a proximidade do valor representado e o valor verdadeiro.• Avalia se o dado observado é importante ou não para aquele contexto.• Informa se a informação ainda é relevante ou se já perdeu a validade.• Não sei a resposta.

Tabela 24 – Questões de associação de conteúdo

Questões	Termos a ser associados	Opções de associação
Questão 5 – Associe cada definição com sua respectiva dimensão, se houver.	<ul style="list-style-type: none"> ● Representa o quão suficiente é o dado para a tarefa que será usado. ● Representa a proximidade do valor representado e o valor verdadeiro. ● Avalia se o dado observado é importante ou não para aquele contexto. ● Informa se a informação ainda é relevante ou se já perdeu a validade. ● Grau de violação dos valores de um dado para as regras semânticas definidas. ● Grau em que os dados podem ser uma fonte confiável para os usos necessários. 	<ul style="list-style-type: none"> ● Acessibilidade. ● Acurácia. ● Atualidade. ● Completude. ● Consistência. ● Confiabilidade. ● Relevância. ● Não sei a resposta.
Questão 6 - Estão listadas algumas perguntas que podem ser realizadas durante a análise de um produto de informação. Associe cada pergunta a uma única Etapa do POSMAD.	<ul style="list-style-type: none"> ● Como os usuários acessam os dados? ● Quais dados podem ser apagados? ● Como os dados são deletados? ● Quem insere os registros no sistema? ● Como os dados são compartilhados no sistema? ● Existe algum treinamento para usuários? ● Quais dados precisam ser guardados? ● Que ferramentas são utilizadas para deleção de dados? 	<ul style="list-style-type: none"> ● Planejamento. ● Obtenção. ● Armazenamento e Compartilhamento. ● Manutenção. ● Aplicação. ● Descarte. ● Não sei a resposta.

<p>Questão 7 – Estão listadas algumas perguntas que podem ser realizadas durante a análise de um produto de informação. Associe cada pergunta a um componente chave.</p>	<ul style="list-style-type: none"> • Como os usuários acessam os dados? • Quais dados podem ser apagados? • Como os dados são deletados? • Quem insere os registros no sistema? • Como os dados são compartilhados no sistema? • Existe algum treinamento para usuários? • Quais dados precisam ser guardados? • Que ferramentas são utilizadas para deleção de dados? 	<ul style="list-style-type: none"> • Dados. • Processos. • Pessoas e organização. • Tecnologia. • Não sei a resposta.
<p>Questão 8 – A partir da lista de problemas associe com a dimensão mais apropriada.</p>	<ul style="list-style-type: none"> • Algumas informações são apresentadas sem todos os campos preenchidos. • Os dados que são apresentados aos usuários não são importantes para o contexto. • As informações apresentadas não equivalem ao dado real. • As informações apresentadas estão defasadas ou perderam a validade. • O usuário não consegue manusear o dado para utilizá-lo. 	<ul style="list-style-type: none"> • Acessibilidade. • Acurácia. • Atualidade. • Completude. • Consistência. • Confiabilidade. • Relevância. • Não sei a resposta.

No experimento, foi formada uma amostra de participantes com 30 alunos da disciplina de Banco de Dados dos cursos de Ciência da Computação e Sistema de Informação da Universidade Federal de Itajubá.

Após a aplicação do questionário na etapa pré-jogo, os participantes tiveram acesso ao jogo em sala de aula. Além disso, o jogo também foi disponibilizado para download e foi solicitado aos participantes que o utilizassem durante um período de três semanas.

Ao fim do período estipulado, teve início a etapa de pós-jogo do processo de avaliação. Nesta etapa, além das questões de conteúdo abordadas anteriormente, foram acrescentadas, em um novo questionário, questões referentes ao tempo de uso, satisfação e usabilidade. As questões referentes ao tempo de utilização foram inseridas com propósito de verificar se o tempo de uso influenciou no aspecto de aprendizagem.

A estratégia definida para avaliar o aprendizado foi comparar os dados coletados na etapa pré-jogo com os dados obtidos na etapa pós-jogo a partir da aplicação do questionário de conteúdo em ambas as etapas. Já os aspectos de usabilidade e satisfação foram investigados tendo como base modelos já existentes na literatura: o modelo *eGameFlow* (RIBEIRO et al., 2015) (NEVES et al., 2013) (SANTOS et al., 2015), o modelo apresentado por Savi et al (2010) e a abordagem de Santos et al (2015). Estes modelos apresentam critérios para avaliação, como imersão, nível de reação e experiência do usuário, que estão atrelados aos elementos de usabilidade e satisfação. A partir do estudo desses modelos foram estabelecidas as questões apresentadas na Tabela 25.

Tabela 25 – Questões específicas para o questionário pós-jogo

Questões	Alternativas
Questão 9 - Quantas vezes jogou?	Resposta única: <ul style="list-style-type: none"> ● Uma a duas vezes. ● Duas a três vezes. ● Quatro a cinco vezes. ● Mais de cinco vezes.
Questão 10 - Em cada jogada, qual foi sua média de tempo de <i>gameplay</i> ?	Resposta única: <ul style="list-style-type: none"> ● Menos de uma hora. ● Uma a duas horas. ● Duas a três horas. ● Mais de três horas.
Questão 11 - Compreendi quais eram meus objetivos no jogo.	Escala de Likert de cinco níveis: <ul style="list-style-type: none"> 1 - Não compreendi de forma alguma. 5 - Compreendi facilmente.

Questão 12 - Consegui realizar as ações necessárias para atingir os objetivos.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Não consegui realizá-las.</p> <p>5 - Consegui facilmente.</p>
Questão 13 - Consegui completar o jogo.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Não consegui completar o jogo.</p> <p>5 - Consegui facilmente.</p>
Questão 14 - Consegui encontrar todos os cinco problemas definidos no jogo.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Não consegui encontrar todos os problemas.</p> <p>5 - Consegui facilmente.</p>
Questão 15 - Entendi o que os ícones representavam.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Não entendi nenhum ícone.</p> <p>5 - Entendi facilmente.</p>
Questão 16 - Foi apresentado <i>feedback</i> suficiente sobre o estágio em que me encontrava no jogo.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Nenhum <i>feedback</i>.</p> <p>5 - Houve <i>feedback</i> suficiente.</p>
Questão 17 - Foi apresentado <i>feedback</i> sobre meus erros e acertos.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Nenhum <i>feedback</i>.</p> <p>5 - Houve <i>feedback</i> suficiente.</p>
Questão 18 - O jogo respondia corretamente às minhas interações.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Não respondia corretamente a nenhuma interação.</p> <p>5 - Respondia corretamente a todas as minhas interações.</p>
Questão 19 - A interface do jogo é agradável.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Muito desagradável.</p> <p>5 - Muito agradável.</p>
Questão 20 - Percebi que minhas ações no jogo interferiam nos resultados obtidos.	<p>Escala de Likert de cinco níveis:</p> <p>1 - Nenhuma de minhas escolhas interferiam nos resultados.</p> <p>5 - Minhas ações interferiam bastante nos resultados.</p>

Questão 21 - Você jogaria novamente?	Escala de Likert de cinco níveis: 1 – Não jogaria com certeza. 5 – Jogaria com certeza.
Questão 22 - Você indicaria o jogo para alguém que quer praticar o conteúdo?	Escala de Likert de cinco níveis: 1 – Não indicaria com certeza. 5 – Indicaria com certeza.

Também foi incorporado ao questionário um campo aberto para sugestões e comentários. As questões utilizadas para coleta dos dados podem ser organizadas de acordo com três critérios de avaliação investigados, além das questões definidas para análise da forma de uso do jogo, como mostra a Tabela 26.

Tabela 26 – Questões específicas para o questionário pós-jogo

Critério de avaliação	Questões
Aprendizagem	Questões 1 a 8
Modo de uso	Questões 9 e 10
Usabilidade	Questões 11 a 19
Satisfação do usuário	Questões 20 a 22

5.2. Análise dos resultados obtidos

As questões 9 e 10, exibidas na Tabela 25, foram definidas para um melhor entendimento sobre o modo de uso do jogo fora do ambiente de sala de aula. É possível observar no gráfico da Figura 21 que nenhuma das partidas teve tempo de *gameplay* superior a duas horas, o que corrobora o tempo idealizado durante o desenvolvimento do jogo.

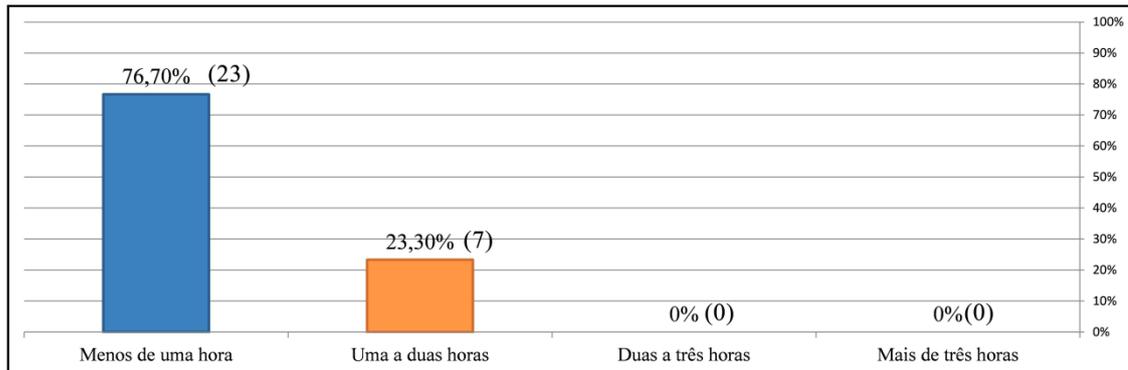


Figura 21 – Questão 10 – Quanto tempo durou cada partida

As questões 11 a 15, associadas aos critérios de usabilidade, estão relacionadas à capacidade do jogo de apresentar ao jogador informações suficientes para que saiba quais ações deve realizar e como realizá-las sem dificuldades. Para responder essas questões o participante deveria selecionar o quanto se sente representado pela expressão na escala de Likert (MAURER; PIERCE, 1998), variando de 1 (discordo fortemente) até 5 (concordo fortemente). Os resultados podem ser observados na Figura 22.

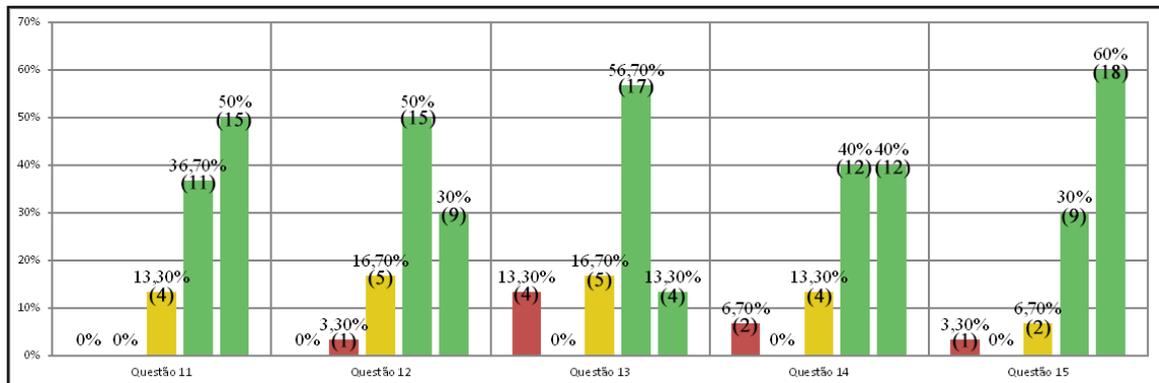


Figura 22 – Questões 11 a 15 – Entendimento e facilidade de utilização do jogo

Dois possíveis métricas que podem ser derivadas dos resultados é a soma da quantidade de respostas com maiores graus de concordância (concordo e concordo fortemente), para indicar o grau de aceitação, e a soma da quantidade de respostas discordo e discordo fortemente, para indicar o grau de rejeição (CAETANO et al, 2016).

Usando essa estratégia, as respostas para a questão 11 - Compreendi quais eram meus objetivos no jogo, indicam que 86,7% dos participantes responderam positivamente sobre a capacidade de compreensão dos objetivos, não tendo sido identificadas reações negativas. Isso sugere que as metas foram definidas e apresentadas de forma clara.

Porém, as respostas para a questão 12 – Consegui realizar as ações necessárias para atingir os objetivos, indicam alguma dificuldade. Nesta versão do jogo, a cada rodada não é possível repetir uma interação com os NPC que representam as etapas do ciclo de vida da

informação. Este fato pode ter dificultado a obtenção de pistas e a identificação dos problemas que é um dos objetivos do jogo.

Esta restrição quanto à interação com NPC também se refere à questão 13 - Consegui completar o jogo, que obteve respostas menos favoráveis do que as demais questões desta categoria (13,3% de discordância e 70% de concordância em algum grau). Além disso, o fato de 6,7% dos participantes terem respondido de forma negativa à Questão 14 – Consegui encontrar todos os problemas definidos no jogo também indicam uma necessidade de evolução do jogo de forma a minimizar esse efeito.

Considerando as dificuldades observadas, é possível que jogadores com um número alto de erros tenham sido desmotivados a completarem o jogo. Para isso podem ser definidas estratégias que considerem os erros cometidos como, por exemplo, penalidade a partir de redução de moedas ou perda de pontos, mas que ofereçam uma nova oportunidade ao jogador de repetir a interação com os NPC sem que diálogos anteriores tenham que ser revistos.

A questão 15 – Entendi o que os ícones representavam mostra a presença de um usuário que não entendeu os ícones utilizados, apresentados durante a introdução. Durante a utilização do jogo, após a introdução, o jogador pode rever as informações sobre ícones interagindo com o NPC que representa o diretor da empresa. Como a interação com este NPC se distancia das ações relacionadas diretamente a concretização dos objetivos do jogo, é possível que o jogador não se sinta instigado em desempenhá-la. Pequenos tutoriais ou guias distribuídos durante a rodada ou no ambiente para que o jogador possa rever essas informações enquanto realiza as ações primordiais podem auxiliar no entendimento dos ícones. Outra possibilidade seria investigar representações mais adequadas para definição dos ícones.

As questões 16 a 19 têm relação com percepção do participante quanto a resposta do jogo aos seus comandos, como *feedbacks* eficientes, resposta correta a suas ações e sensação de controle. As respostas obtidas são apresentadas no gráfico da Figura 23.

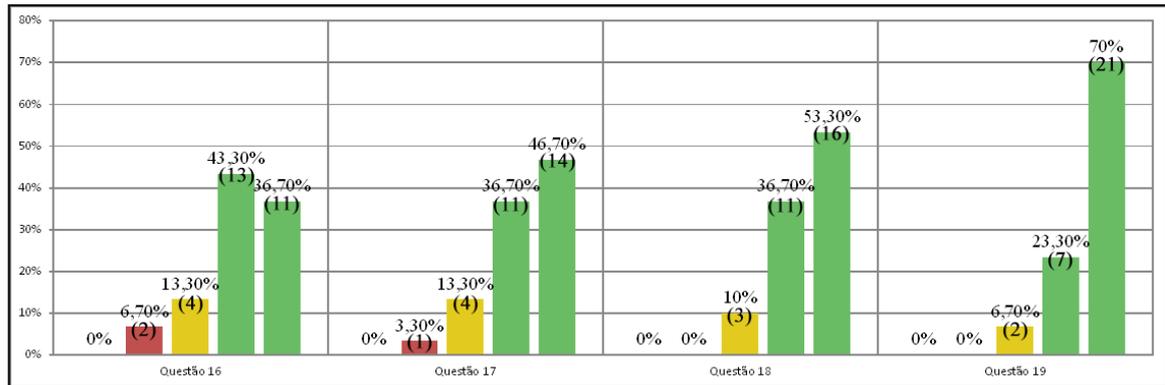


Figura 23 – Questão 16 à 19 – Capacidade de resposta do jogo à instruções.

Apesar do indicador positivo de 80% e 83,4%, respectivamente, as questões 16 – Foi apresentado *feedback* suficiente sobre o estágio em que me encontrava no jogo e 17 – Foi apresentado *feedback* sobre os meus erros e acertos, tiveram algumas respostas abaixo do esperado. *Feedbacks* relacionados ao erro cometido durante as interações para completar a matriz de interação do POSMAD são incorporados ao diálogo durante a própria interação. Após a interação com um NPC, caso o jogador não tenha encontrado uma pista, o espaço referente àquele elemento da matriz é deixado em branco. Contudo, na interface não há um elemento que diferencie um espaço em branco deixado por conta de um erro ao tentar obter uma pista e o espaço em branco que ainda não foi preenchido porque o jogador ainda não interagiu com o NPC correspondente. Esta característica da interface pode confundir o jogador quanto à possibilidade de continuar a busca para completar a matriz. Logo, esse resultado sugere a necessidade de alteração da interface para eliminar esse efeito.

O jogo ainda permite que a pontuação e a quantidade de moedas seja acessada clicando com o botão direito do mouse em qualquer lugar da tela, ou interagindo com o NPC que representa o diretor da empresa. Um recurso possível para aumentar o nível de *feedback* do jogo quanto à apresentação destas informações é manter esses dados na tela durante toda utilização.

Em relação à questão 18 – O jogo respondia corretamente às minhas interações, 90% dos participantes responderam positivamente quanto ao caráter responsivo do jogo. Este resultado pode indicar que as implementações das ações necessárias para execução dos objetivos foram definidas e realizadas corretamente.

A questão 19 – A interface do jogo é agradável, obteve 93,3% como indicador positivo. A interface desenvolvida é baseada na interface clássica de RPG, o que pode ter contribuído para a familiarização do ambiente por participantes que já tiveram contato com jogos anteriores. Contudo, esse aspecto não foi analisado.

Para uma visão geral, foi elaborado um gráfico a partir das respostas de todas as questões de usabilidade (Figura 24). Ao analisar a Figura 24, é possível perceber que na distribuição dos valores de resposta (1: discordo fortemente – 5: concordo fortemente), há uma ocorrência maior de respostas positivas o que sugere que este aspecto do jogo está adequado.

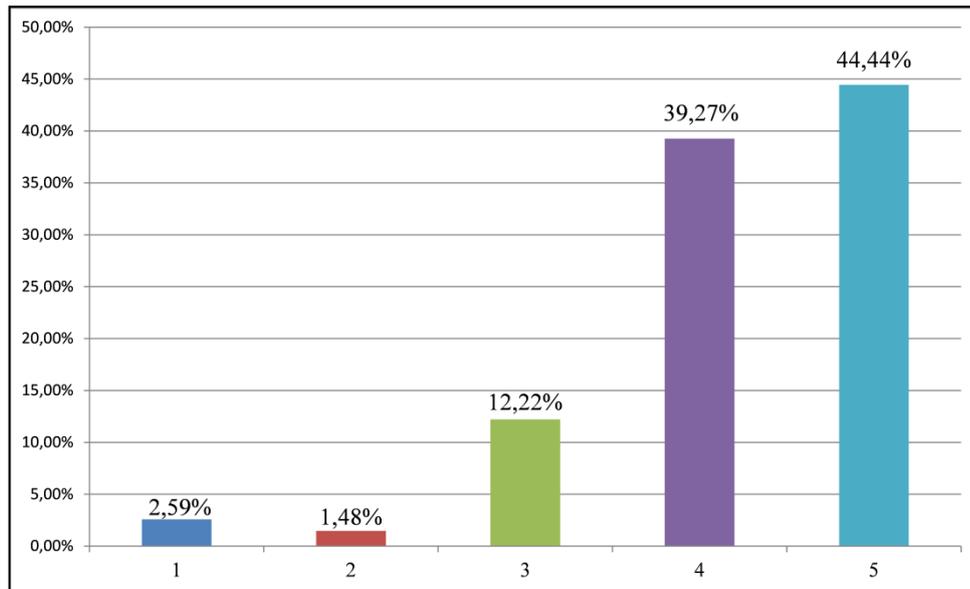


Figura 24 – Frequência de respostas referentes a questões de usabilidade.

As questões 20 a 22 são relacionadas à satisfação do participante. Os resultados podem ser observados nos gráficos das Figuras 25, 26 e 27, respectivamente.

A questão 20 questiona o participante quanto a sensação de interferência no jogo causada por suas ações. O resultado, com 93,3% de reações favoráveis, contribui para a avaliação positiva do jogo no aspecto motivacional, sendo que esta percepção tem grande relação com o elemento controle, um dos elementos definidos entre as motivações intrínsecas na Seção 2.2.3 (SAVI et al., 2010). Uma das possíveis causas para a sensação de controle é a percepção, pelo jogador, que seus erros e acertos interferem nos resultados, como na sequência do jogo ou na pontuação.

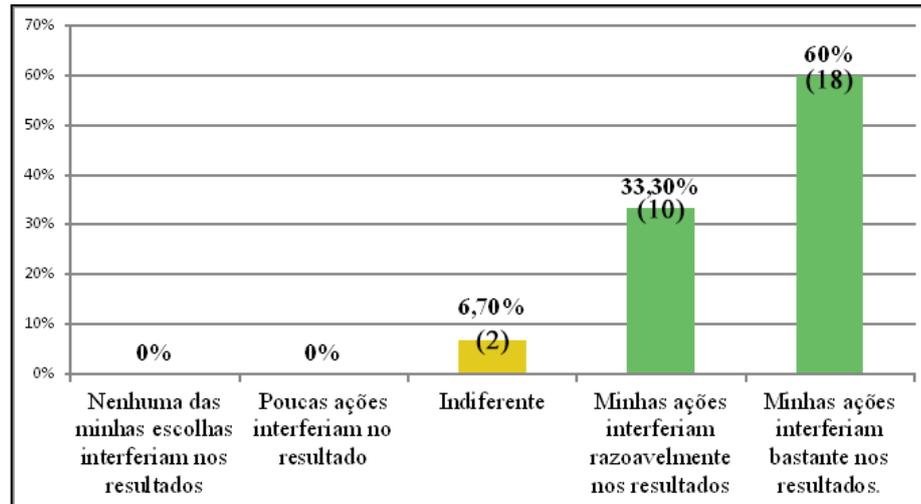


Figura 25 – Questão 20 – Referente a sensação de controle.

Os resultados obtidos com a questão 22 (Figura 27), que verifica a possibilidade de indicação do jogo, indicam que os participantes acreditam no potencial do jogo como instrumento para prática. Porém, o gráfico observado na Figura 26, apesar de apresentar 70% dos resultados favoráveis ao uso novamente do jogo, ainda obteve 6,6% de respostas negativas e 23,3% indiferentes. Uma das possíveis justificativas para esses resultados é a apresentação de um conteúdo longo em um número reduzido de diálogos, que podem ser considerados cansativos, assim como textos bastante extensos, não motivando o jogador a repetir as lições abordadas. Uma forma de reduzir a sensação causada pela leitura dos diálogos é a quebra dessas informações em um número maior de interações, adicionando outros personagens. Também há a possível relação entre os jogadores que não conseguiram completar o jogo e a desmotivação para tentar novamente. Há de se considerar o fato do jogo ser pequeno e muitos usuários já tê-lo finalizado, diminuindo a motivação.

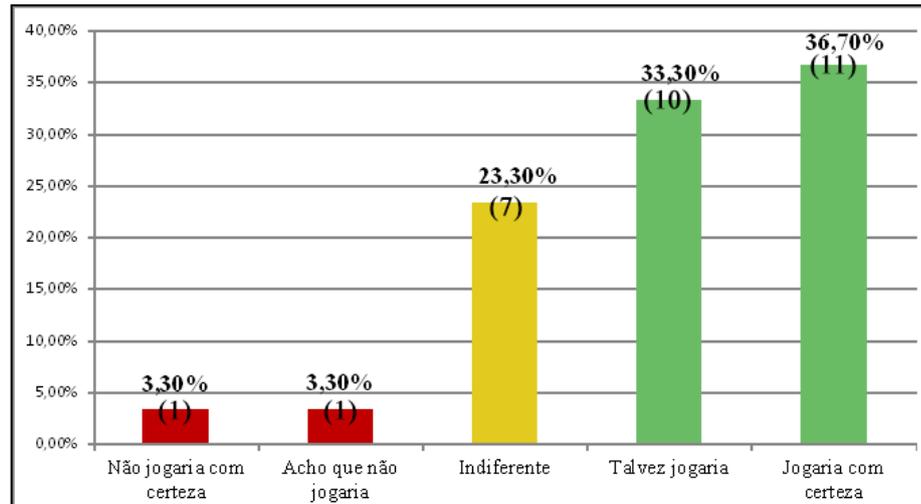


Figura 26 – Questão 21 – Referente a vontade do participante de jogar novamente.

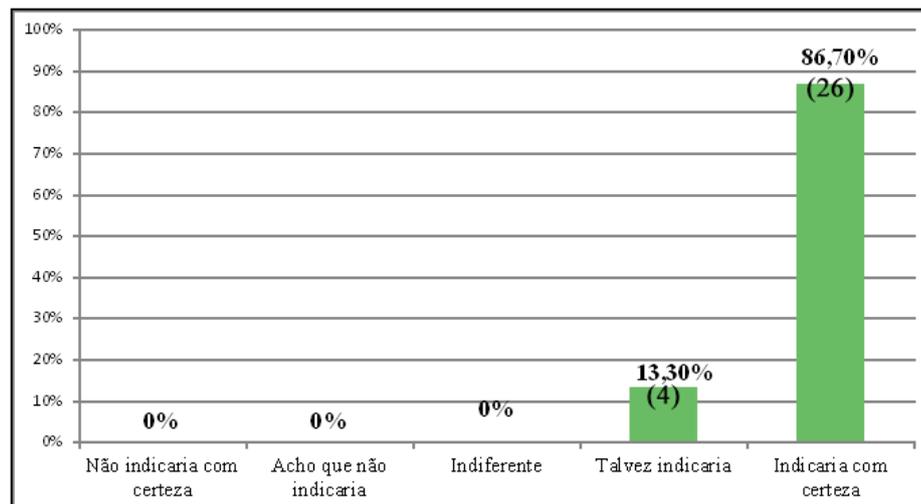


Figura 27 – Questão 22 – Referente a indicação para prática do conteúdo abordado.

O gráfico que representa as respostas de todas as questões de satisfação do usuário é apresentado na Figura 28. Assim como na usabilidade, a distribuição dos resultados indica que quanto ao aspecto de satisfação, a proposta do jogo obteve boa aceitação.

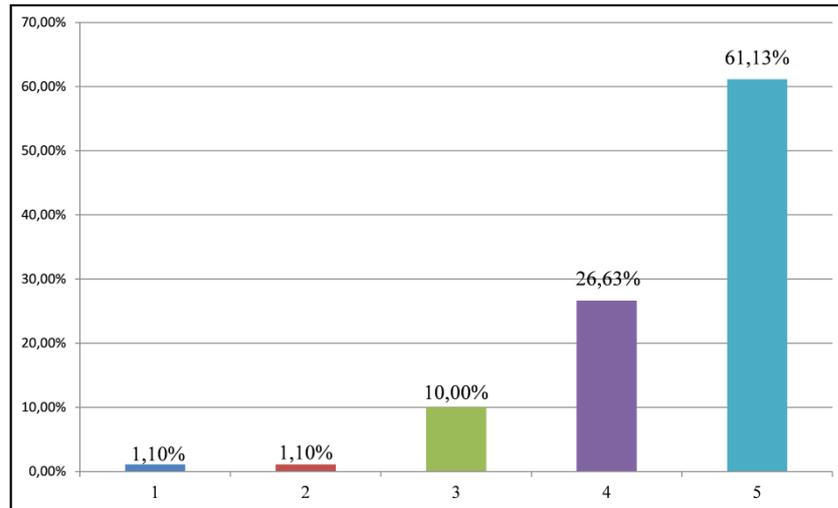


Figura 28 – Frequencia de respostas referentes à questões de satisfação do usuário.

A análise relacionada ao aspecto de aprendizagem foi realizada comparando a porcentagem de acertos obtida através da aplicação do questionário (Tabelas 23 e 24) na etapa pré-jogo e a porcentagem de acerto para as mesmas questões na etapa pós-jogo.

O gráfico da Figura 29 apresenta as porcentagens de acerto antes e depois do uso do jogo para as questões objetivas de 1 a 4. A questão 1 teve um aumento significativo na porcentagem de acertos. Esta pergunta aborda a nomenclatura de cada etapa do ciclo de vida dos dados, conteúdo apresentado nas interações com NPC durante toda a interação associada ao primeiro objetivo do jogador. A apresentação repetitiva desses dados pode ser uma das causas do aumento de acertos na etapa pós-jogo.

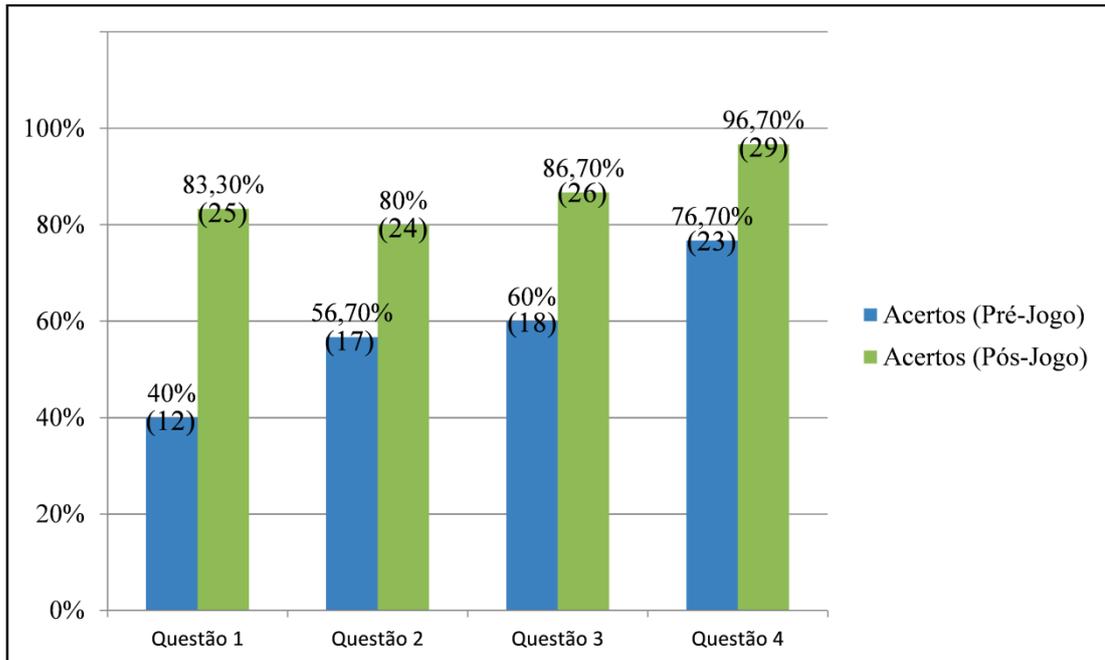


Figura 29 – Questões 1 à 4 - Porcentagem de acerto na etapa pré-jogo e pós-jogo.

A definição de Produto de Informação, requerida na questão 2, não é explicitada de forma direta no jogo. Apesar de mencionada durante a introdução, o entendimento desta definição é realizado implicitamente durante a simulação de um caso hipotético. Logo, é razoável inferir que alguns participantes foram capazes de construir a definição a partir da apresentação indireta durante a simulação aumentando a quantidade de acertos.

A questão 3 pergunta ao participante quais componentes chaves são avaliados em cada etapa do POSMAD. Assim como as etapas do ciclo de vida dos dados, informações sobre os componentes chaves são abordados por todo o percurso de realização do primeiro objetivo do jogador, que é completar a matriz.

A questão 4 é referente a definição de uma dimensão em específico, a acurácia. As dimensões são abordadas durante a execução do segundo objetivo do jogo, quando o jogador interage com os NPC que representam os problemas. Nesta interação, as definições formais de diversas dimensões são apresentadas claramente junto ao nome da dimensão associada. Desse modo, o jogador pode descobrir o que a dimensão representa e selecionar a que se adequa com as informações sobre o problema.

Para cada problema, várias dimensões são oferecidas como opções, o que possibilita ao jogador rever as definições das dimensões mais de uma vez, fato que pode auxiliar na aprendizagem. Apesar do aumento de 20% no número de acertos, as informações obtidas com a análise de usabilidade apontam que alguns participantes podem não ter completado a matriz

do POSMAD de forma a alcançar esse estágio do jogo. Além das modificações sugeridas anteriormente para que o jogador alcance a segunda etapa do jogo, também pode ser interessante disponibilizar as informações sobre dimensões durante a etapa na qual o jogador faz a busca por pistas.

A Figura 30 ilustra os resultados obtidos para a questão 5. Nesta questão, o participante deveria associar o nome da dimensão a sua correta descrição. No gráfico, estão representadas as quantidades de acertos antes e depois da utilização do jogo para cada dimensão: acurácia, atualidade, completude, consistência, confiabilidade e relevância. O gráfico é estruturado apresentando a resposta correta para as definições apresentadas na ordem de suas aparições no questionário.

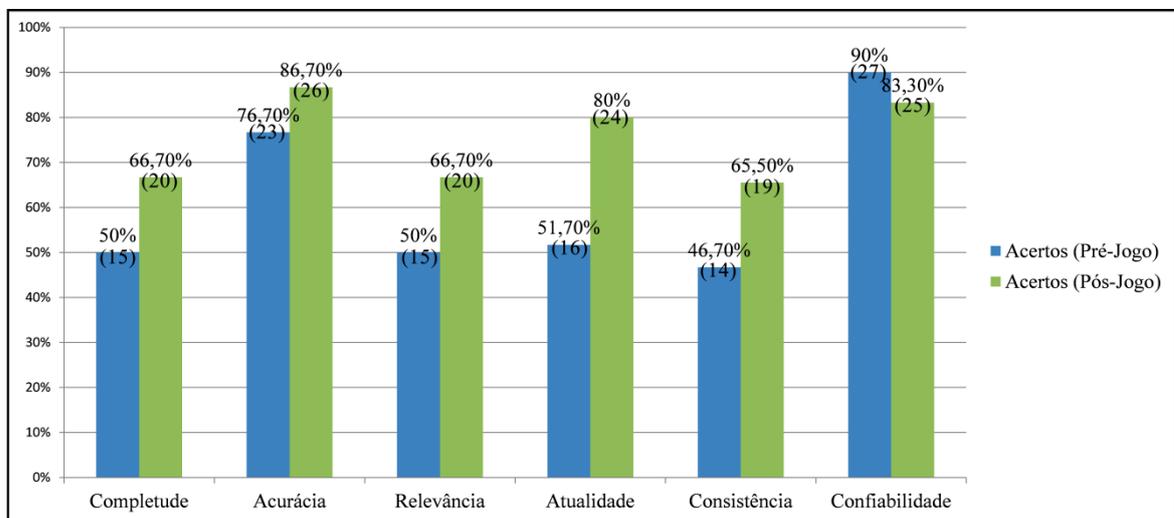


Figura 30 – Questão 5 - Porcentagem de acerto na etapa pré-jogo e pós-jogo.

Ao analisar o gráfico, é possível observar que para uma dimensão, a confiabilidade, houve uma queda da quantidade de acertos. Este resultado pode ser justificado pelo fato de ser esta a única dimensão não considerada no jogo.

Na questão 6 o participante deve associar uma sequência de perguntas referentes a investigação realizada durante o preenchimento da matriz do POSMAD com a etapa do ciclo de vida dos dados que cada uma é relacionada. Para a questão 7 foi solicitado que a associação das mesmas perguntas fossem feitas com um dos quatro componentes chaves. As perguntas que deveriam ser associadas podem ser vistas na Tabela 24. As Figuras 31 e 32 apresentam, para as questões 6 e 7 respectivamente, as respostas corretas na ordem em que as questões solicitadas apareciam no questionário.

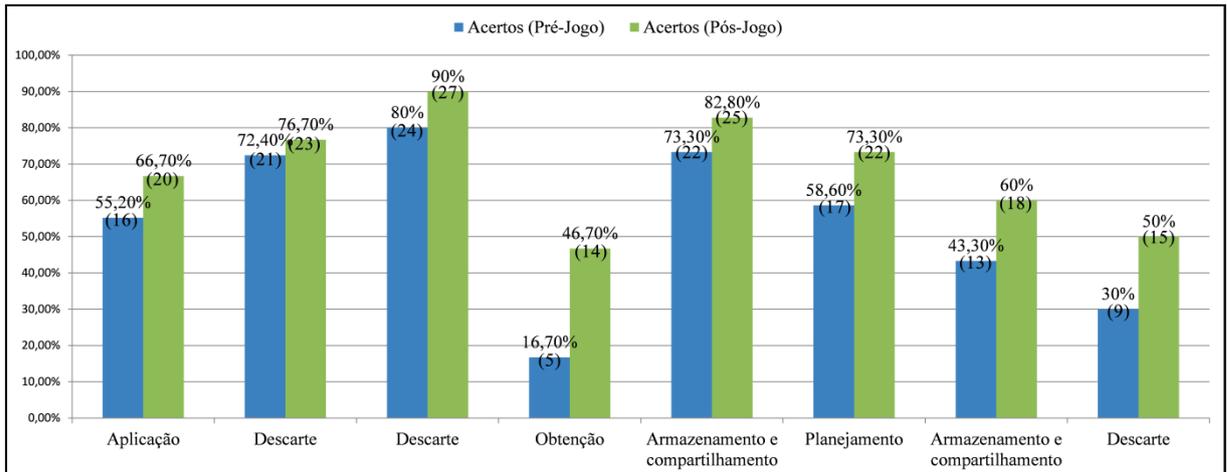


Figura 31 – Questão 6 - Porcentagem de acerto na etapa pré-jogo e pós-jogo.

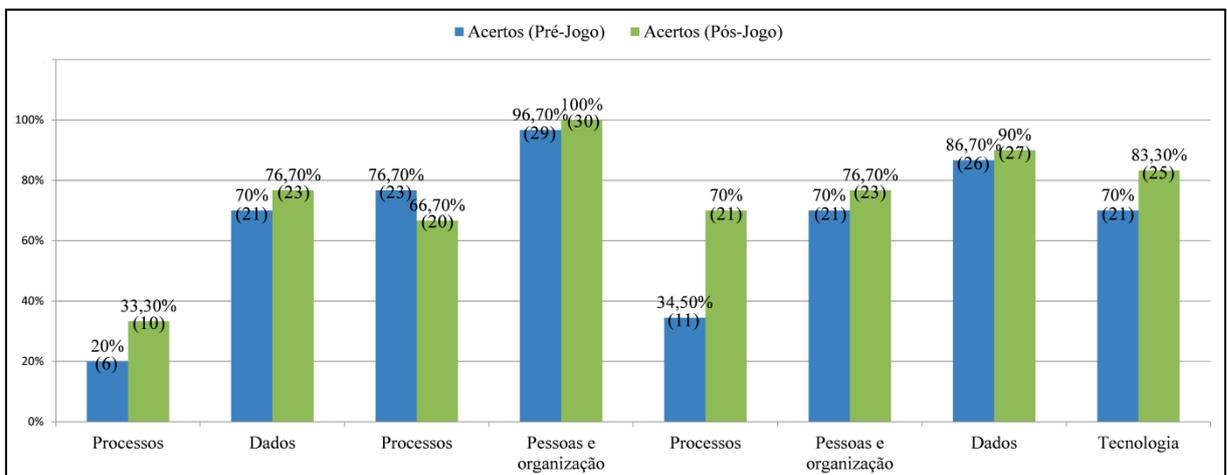


Figura 32 – Questão 7 - Porcentagem de acerto na etapa pré-jogo e pós-jogo.

Apesar do aumento na porcentagem de acertos para a maioria das associações realizadas nas questões 6 e 7, a diferença desta porcentagem pode ser considerada baixa em comparação com a diferença encontrada nas questões anteriores 1 à 5. A associação abordada nas questões 6 e 7 são apresentadas no jogo durante o preenchimento da matriz do POSMAD, sendo que o participante deve associar a pergunta correta à etapa do ciclo de vida dos dados e ao componente chave em que se encontra. Nesta versão do jogo, é informado que jogador errou a associação, porém não é informada a resposta correta. Considerando que a interação não pode ser realizada novamente nesta mesma rodada, o jogador que não reiniciar o *gameplay* não terá certeza absoluta de qual associação seria a adequada. Uma solução é permitir que o jogador refaça a interação na mesma rodada, e que aprenda utilizando o método de tentativa e erro, ou que o jogo apresente a resposta correta como *feedback* de seu erro.

A questão 8 apresenta várias descrições sucintas sobre problemas de qualidade de dados e solicita ao participante que relacione com a dimensão mais adequada. A Figura 33 apresenta a relação de acertos na etapa pré-jogo e pós-jogo para cada dimensão correta.

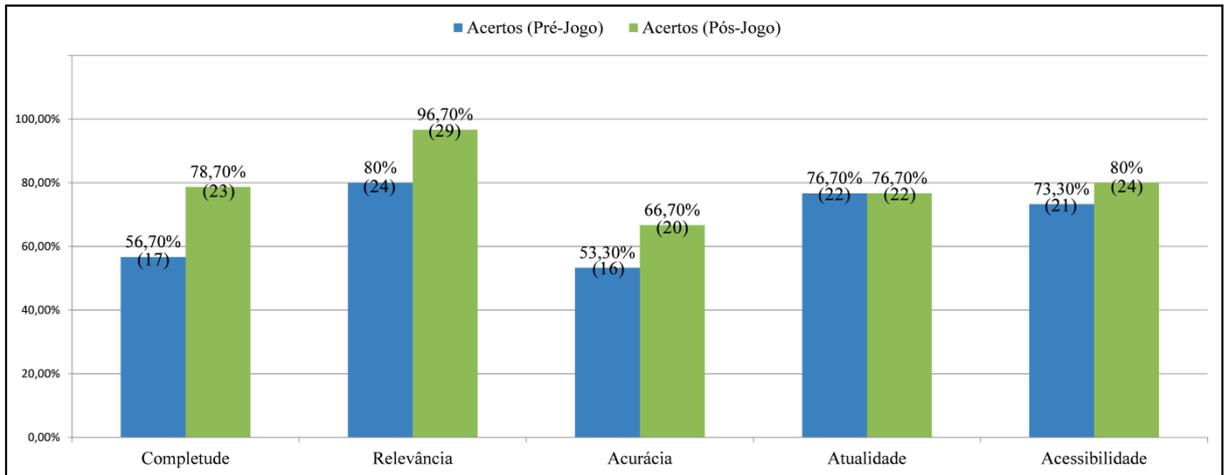


Figura 33 – Questão 8 - Porcentagem de acerto na etapa pré-jogo e pós-jogo.

O jogo apresenta cinco problemas diferentes para serem associados com as dimensões. Apesar dos problemas apontados no questionário não serem os mesmos abordados no jogo, durante a seleção da dimensão relacionada o jogador tem acesso à definição das dimensões. A partir do conhecimento do significado de uma dimensão e da prática em associar problemas a uma dimensão específica, existe a possibilidade do aumento da capacidade de relacionar os dois elementos.

Além das análises individuais das questões é possível verificar a relação do uso contínuo do jogo ao suporte à aprendizagem. A questão 9 (Tabela 25) foi utilizada para categorizar a amostra em: participantes que utilizaram o jogo uma a três vezes e participantes que utilizaram o jogo mais de quatro vezes.

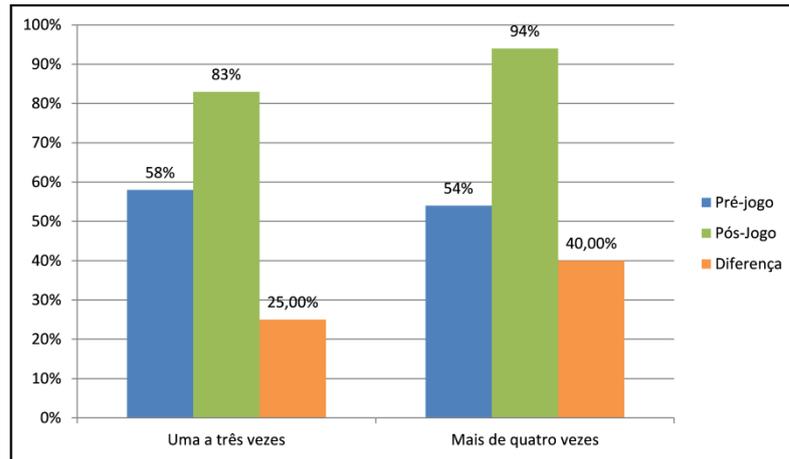


Figura 34 – Diferença de número de acertos relacionado ao tempo de uso

É possível notar que os participantes que utilizaram o jogo um número maior que quatro vezes teve um aproveitamento melhor. Uma das possíveis causas para este resultado é a repetição dos conceitos já visualizados, reforçando o conteúdo aprendido em rodadas anteriores. Outro argumento é que, ao iniciar uma nova rodada, o jogador pode refazer interações que antes havia respondido de forma errada, possibilitando que atinja estágios do jogo que antes não poderia ter ido, no qual estão presentes novos conceitos, como é o caso do conteúdo dimensões.

A validade da avaliação pode ser ameaçada considerando o possível conhecimento prévio dos participantes. Novas avaliações com diferentes nichos podem atenuar possíveis inconsistências e prover resultados mais acurados.

6. Considerações Finais

É possível observar a necessidade de metodologias práticas para a avaliação de qualidade de dados. Nesse contexto, este trabalho apresentou como uma de suas contribuições uma sequência de etapas baseadas, em especial, no modelo TDQM em conjunto com sugestões de ferramentas que pudessem guiar de forma menos subjetiva as organizações em processo de avaliação da qualidade de seus dados.

Diante da dificuldade de encontrar situações reais em que as metodologias de avaliação pudessem ser ensinadas, a criação de um jogo educacional, pertencente ao gênero de jogos sérios, se tornou o objetivo deste trabalho, sendo este objetivo alcançado com o desenvolvimento do *Data Quality Simulator*, considerado também uma das contribuições deste trabalho. Este trabalho também contribuiu à literatura com a descrição da metodologia utilizada na elaboração e avaliação do *Data Quality Simulator*.

Percebe-se que o crescente uso de tecnologias de jogos digitais é uma forma relevante de melhoria para as abordagens educacionais, em especial pela oportunidade apelativa que os jogos têm a grande população. Neste quesito, este trabalho corrobora para a literatura a partir das pesquisas de opinião realizadas.

Abordagens educacionais foram utilizadas e metodologias de validação foram estudadas para criar uma base concreta no desenvolvimento do projeto, objetivo principal deste trabalho.

Foi observado que a concepção de um jogo sério necessita de fundamentos proveniente de diversas áreas além do conteúdo a ser abordado, como teorias de aprendizagem e elementos motivacionais. Também é oportuno a validação dos jogos desenvolvidos, contribuindo para o aprimoramento do jogo proposto e para a validação de uso de jogos no aspecto educacional.

Ao se avaliar o jogo *Data Quality Simulator* utilizando a estratégia apresentada neste trabalho é possível identificar elementos passíveis de melhoria. Além dos critérios discutidos, um elemento que pode ser trabalhado de forma menos superficial é a disponibilidade de recursos para interação social, considerado relevante no método de avaliação *GameFlow* (NEVES et al, 2014). O jogo ainda pode ser estendido para que outras etapas e ferramentas da metodologia de análise de qualidade de dados sejam incorporadas; considerando ainda a implementação das melhorias apontadas.

O estudo realizado neste trabalho, as estratégias empregadas e as informações obtidas a partir da análise das pesquisas podem ser também utilizadas para o desenvolvimento de jogos relacionados à outros temas.

A análise geral das avaliações de satisfação do usuário, usabilidade e aprendizagem se mostraram positivas, de acordo com os participantes. Neste sentido, concluímos que o desenvolvimento do jogo *Data Quality Simulator* fortalece o uso de jogos para o ensino e prática de conteúdo.

Como trabalhos futuros, além da implementação das sugestões e aprimoramentos identificados durante o processo de desenvolvimento do jogo e análise da avaliação, é possível realizar a extensão do *Data Quality Simulator* para outros conteúdos e ferramentas da metodologia. Também é possível utilizar as técnicas apresentadas para a elaboração de um novo jogo referente à outro tema além da qualidade de dados.

REFERÊNCIAS BIBLIOGRÁFICAS

- AHMAD, M.; RAHIM, L.; ARSHAD, N. **A review of educational games design frameworks: An analysis from software engineering.** In: IEEE. [S.l.]: Computer and Information Sciences (ICCOINS), 2014.
- AIRASIAN, P. W.; MIRANDA, H. **The role of assessment in the revised taxonomy. Theory into practice,** Taylor & Francis, v. 41, n. 4, p. 249–254, 2002.
- ARNAB, S. et al. **Mapping learning and game mechanics for serious games analysis.** British Journal of Educational Technology, Wiley Online Library, v. 46, n. 2, p. 391–411, 2015.
- BAGCHI, S.; BAI, X.; KALAGNANAM, J. **Data quality management using business process modeling.** In: IEEE. 2006 IEEE International Conference on Services Computing (SCC'06). [S.l.], 2006. p. 398–405.
- BARROS, R. **Qualidade de informação na web : um prognóstico fuzzy baseado em metadados.** Tese (Doutorado) — Programa de Pós-graduação em Engenharia de Sistemas e Computação (COPPE), Universidade Federal do Rio de Janeiro, 2009.
- BASTOS, M. et al. **Data integration: Quality aspects.** In: IEEE. Transmission and Distribution Conference and Exposition: Latin America (T&D-LA), 2010 IEEE/PES. [S.l.], 2010. p. 411–416.
- BATINI, C. et al. **Methodologies for data quality assessment and improvement.** ACM computing surveys (CSUR), ACM, v. 41, n. 3, p. 16, 2009.
- BATINI, C.; SCANNAPIECO, M. **Data Quality: Concepts, Methodologies and Techniques** (Data-Centric Systems and Applications). Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006. ISBN 3540331727.
- BLAKE, R. H.; MANGIAMELI, P. **The effects and interactions of data quality and problem complexity on data mining.** In: ICIQ. [S.l.: s.n.], 2008. p. 160–175.
- BLOOM, B. S. **Taxonomy of educational objectives: The classification of educational goals.** Longman Group, 1956.
- BLOOM, B. S.; KRATHWOHL, D. R.; MASIA, B. B. **Taxionomia de Objetivos Educacionais: Dominio Afetivo.** Porto Alegre: Globo, 1972. 203p.
- CAETANO, B. P., de OLIVEIRA, G. W., de PAULA, M. M., & de SOUZA, J. M. **Democracia digital: uma análise sobre recursos e aceitação.** XII Brazilian Symposium on Information Systems, Florianópolis, SC, 2016.
- CAO, J. et al. **Data lifecycle process model and quality improving framework for tqm practices.** In: IEEE. E-Product E-Service and E-Entertainment (ICEEE), 2010 International Conference on. [S.l.], 2010. p. 1–6.
- CARVALHO, M. B. et al. **An activity theory-based model for serious games analysis and conceptual design.** Computers & Education, Elsevier, v. 87, p. 166–181, 2015.
- CHEN, M. et al. **Survey on data quality.** In: IEEE. [S.l.], 2012.
- CHENGALUR-SMITH, I. N.; BALLOU, D. P.; PAZER, H. L. **The impact of data quality information on decision making: an exploratory analysis.** IEEE Transactions on Knowledge and Data Engineering, IEEE, v. 11, n. 6, p. 853–864, 1999.

- COSTA, R. et al. **Classificação cognitiva das atividades avaliativas utilizadas nos ambientes virtuais de aprendizagem com base na taxonomia de bloom.** In: . [S.l.]: Revista de Informática Aplicada - USCS/UFABC, 2014.
- DIAS, J., BRANDÃO, I., NASCIMENTO, F., HETKOWSKI, T., & PEREIRA, T. **Avaliação de jogos educacionais digitais baseada em Perspetivas: Uma experiência através do jogo-simulador Kimera.** XII SBGames, São Paulo – SP, 2013.
- FERRAZ, A. P.; BELHOT, R. **Taxonomia de bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais.** In: . [S.l.]: The Scientific Electronic Library Online, 2010.
- HENRIQUE, M. S. et al. **Uma revisão sistemática da literatura sobre o uso de teorias de aprendizagem em softwares educacionais.** RENOTE, v. 13, n. 2, 2015.
- JAQUEIRA, A. **Data Warehouse: Informação com qualidade para facilitar a geração de estratégias.** Dissertação (Mestrado) — Universidade Candido Mendes, 2009.
- KAPLAN, R. S.; NORTON, D. P. **A estratégia em ação: balanced scorecard.** [S.l.]: Gulf Professional Publishing, 1997.
- KOVAC, R.; LEE, Y. W.; PIPINO, L. **Total data quality management: The case of iri.** In: IQ. [S.l.: s.n.], 1997. p. 63–79.
- LEE, Y. W. et al. **Journey to data quality.** [S.l.]: The MIT Press, 2009.
- LESYUK, A. **Mastering Redmine.** Packt Publishing Ltd, 2013
- MALONE, T.; LEPPER, M. **Making learning fun: A taxonomy of intrinsic motivations for learning.** v. 3, 1987.
- MAURER, T. J.; PIERCE, H. R. **A comparison of Likert scale and traditional measures of self-efficacy.** Journal of applied psychology, v. 83, n. 2, p. 324, 1998.
- MCGILVRAY, D. **Executing data quality projects: Ten steps to quality data and trusted information (TM).** [S.l.]: Elsevier, 2008.
- MORENO-GER, P. et al. **Serious games: A journey from research to application.** In: IEEE. 2014 IEEE Frontiers in Education Conference (FIE) Proceedings. [S.l.], 2014 p.1–4.
- NEVES, D. E., SANTOS, L. G., SANTANA, R. C., & ISHITANI, L. **Avaliação de jogos sérios casuais usando o método GameFlow.** Revista Brasileira de Computação Aplicada, 6(1), 45-59, 2013.
- NORMAN, E. S.; BROTHERTON, S. A.; FRIED, R. T. **Work breakdown structures: The foundation for project management excellence.** [S.l.]: John Wiley & Sons, 2010.
- PIPINO, L. L.; LEE, Y. W.; WANG, R. Y. **Data quality assessment.** Communications of the ACM, ACM, v. 45, n. 4, p. 211–218, 2002.
- Project Management Institute. **Practice standard for work breakdown structures.** In: [S.l.]: Project Management Institute, Inc., 2013.
- REZENDE, D. A. **Planejamento de sistemas de informação e informática.** São Paulo: Atlas, 2003.
- RIBEIRO, R. et al. **Teorias de aprendizagem em jogos digitais educacionais: um panorama brasileiro.** In: . [S.l.]: RENOTE - Revista Novas Tecnologias na Educação, 2015.

- ROCHA, R.; BITTENCOURT, I.; ISOTANI, S. **Análise, projeto, desenvolvimento e avaliação de jogos sérios e afins: uma revisão de desafios e oportunidades**. Anais do XVI Simpósio Brasileiro de Informática na Educação (SBIE 2015), 2015.
- SANTOS, W., NETO, S., da SILVA JUNIOR, C. G., & BITTENCOURT, I. I. **Avaliação de Jogos Educativos: Uma Abordagem no Ensino de Matemática**. In Anais do Simpósio Brasileiro de Informática na Educação (Vol. 26, No. 1, p. 657). 2015.
- SAVI, R. et al. **Proposta de um modelo de avaliação de jogos educacionais**. In: . [S.l.]: RENOTE - Revista Novas Tecnologias na Educação, 2010.
- TORELLI, M. **Qualidade da Informação sob a Perspectiva de Produto**. Dissertação (Mestrado) — Pontifícia Universidade Católica de São Paulo (PUC-SP), 2005.
- WANG, K. et al. **Analysis of data quality and information quality problems in digital manufacturing**. In: IEEE. Management of Innovation and Technology, 2008. ICMIT 2008. 4th IEEE International Conference on. [S.l.], 2008. p. 439–443.
- WANG, R. Y. **A product perspective on total data quality management**. Communications of the ACM, ACM, v. 41, n. 2, p. 58–65, 1998.
- WANG, R. Y.; STOREY, V. C.; FIRTH, C. P. **A framework for analysis of data quality research**. IEEE transactions on knowledge and data engineering, IEEE, v. 7, n. 4, p.623–640, 1995.
- WANG, R. Y.; STRONG, D. M. **Beyond accuracy: What data quality means to data consumers**. Journal of management information systems, v. 12, n. 4, p. 5-33, 1996.
- WOLFF, D. Get things done with Trello: Your quick access to productivity and success. Organized Living Press, 2014.