

Avaliação da Qualidade - Site Detran Digital

Funcionalidade: Consulta Veículo - Nada Consta

Victor Hugo O. Leão¹, Abraão A. Ribeiro¹, Felipe Alef P. Rodrigues¹, Guilherme A. de La Fuente V.¹, Lucas Ramon A. de Oliveira¹, Paulo Henrique A. da Silva¹

¹Faculdade do Gama (FGA) – Universidade de Brasília(UnB)
72.444-240 – Gama – Brasília – DF – Brazil

Resumo. *O documento evidencia a avaliação da qualidade da funcionalidade "Consulta Veículo - Nada Consta" do site "Detran Digital" sob a perspectiva de um usuário idoso (65 anos ou mais). Para isso, foi seguido um modelo de qualidade usando como dimensões "desempenho e eficiência" e "usabilidade". A acessibilidade foi explorada usando a Web Accessibility Initiative (WAI), em que diversos pontos descritos pela iniciativa foram citados. Percebeu-se que nem todas as métricas definidas tiveram seu critério de julgamento ideal atingido e, por isso, os resultados da avaliação compõem oportunidades de melhora e recomendações.*

1. Introdução

Esse documento é o resultado final da elaboração e execução do Plano de Medição para a avaliação da qualidade da funcionalidade de "Consulta veículo - Nada Consta" presente no site do "Detran Digital" sob a perspectiva do usuário idoso, isto é, aquele com 65 anos de idade ou mais. A fim de melhor contextualizar a funcionalidade e o produto a serem avaliados, o Detran é o Departamento Estadual de Trânsito, sendo assim, ele tem a função, de acordo com a plataforma Zul Digital, de exigir e impor que as leis de trânsito sejam cumpridas no âmbito de sua jurisdição. É o órgão que planeja e controla tudo aquilo que é referente à documentação dos condutores e seus veículos. Dessa forma, o Detran possui um website que dispõe de diversas funcionalidades para os condutores para com seus documentos de trânsito e seus veículos.

O termo qualidade possui um conceito complexo e com diferentes entendimentos dependendo do ponto de vista e, em software, a qualidade varia com diferentes visões. Para o escopo desse documento, a visão do usuário será a com maior enfoque, uma vez que busca entender o grau em que um produto atende às expectativas dos usuários (nesse caso, maiores de 65 anos), buscando identificar quais atributos são considerados importante por esses que o usam. Dessa forma, a visão do usuário abrange todos os elementos subjetivos que vão além das funcionalidades esperadas para satisfação do mesmo.

A avaliação a ser realizada é feita levando em consideração duas dimensões/características da qualidade e suas respectivas subcaracterísticas: Desempenho e eficiência (capacidade); Usabilidade (proteção contra erros do usuário, acessibilidade e aprendizagem).

Com isso, o processo de avaliação é feito a partir do estabelecimento dos requisitos de avaliação (seu propósito, tipos de produtos a serem avaliados e o modelo de qualidade). Isso, então, possibilita a especificação da avaliação com as devidas métricas e suas particularidades, além da projeção da avaliação, que consiste no plano em si. Após

esses passos, a avaliação é, finalmente, executada de diversas formas, desde entrevistas até ferramentas automatizadas e testes manuais.

A partir da execução da avaliação e os resultados dela estarem disponíveis, eles são relatados de acordo com cada uma das características e subcaracterísticas definidas. Por fim, os resultados expostos podem ser discutidos e melhorias serem recomendadas.

2. Processo de Avaliação

O processo de avaliação é composto de quatro fases: requisitos da avaliação (fase em que se entende de fato o que será avaliado), especificação da avaliação (definição das métricas e seus níveis de pontuação e critérios), projeção da avaliação (plano da avaliação é de fato exposto) e a execução da avaliação (obtenção das medidas, juntamente com sua comparação e julgamento). A partir desses passos, o processo de avaliação é completo e, então, pode-se passar para os relatos dos seus resultados.

2.1. Requisitos de Avaliação

A fim de estabelecer os requisitos da avaliação a ser desenvolvida, três subtópicos serão explorados: propósito da avaliação, tipos de produtos a serem avaliados e modelo da qualidade. Com isso, entenderá-se de forma inicial e sucinta, mas também essencial, quais são as especificações da avaliação e o que de fato será avaliado.

2.1.1. Propósito da avaliação

O propósito da avaliação é medir a qualidade da funcionalidade de "Consulta veículo - Nada Consta" do site do "Detran Digital" sob a perspectiva do usuário idoso para, então, além de relatar os resultados obtidos, também pontuar oportunidades de melhoria e fazer recomendações.

A avaliação terá como foco as seguintes dimensões da qualidade: desempenho e eficiência (capacidade) e usabilidade (proteção contra erros do usuário, acessibilidade e aprendizagem). Essas características e subcaracterísticas estão fortemente ligadas a como usuários de idade avançada utilizam e entendem um sistema digital (nesse caso, um site) que são, muitas vezes, uma novidade para eles. Além disso, por ser um site governamental e com funcionalidades de acesso por ampla parte da população, sua qualidade deve ser avaliada e, conseqüentemente, garantida.

Com isso, serão utilizadas métricas, ferramentas automatizadas, testes manuais e referências externas para atingir o propósito explicitado, principalmente na visão de um usuário com mais de 65 anos.

2.1.2. Tipos de produtos a serem avaliados

O produto "Detran Digital", em sua aplicação web, possui diversas tarefas que podem ser julgadas principais para os seus usuários. No que convém a veículos, pode-se encontrar funcionalidades para consultas, emissão de documentos (CRLV, ATPV e placa de identificação de veículo), além de transferência e inclusão de novos veículos adquiridos. Em adição, também existem opções de consultas, solicitação e outros relacionados

à Carteira Nacional de Habilitação (CNH). Por fim, agendamentos, infrações e consultas administrativas (para empresas) também se destacam como tarefas importantes para os usuários que acessam o produto.

Assim como exposto anteriormente, o escopo da avaliação é a funcionalidade de "Consulta veículo - Nada Consta", sendo, então, a função com maior dedicação para a avaliação realizada. Desde a página inicial para um usuário sem login realizado, apenas duas janelas o separam do acesso à funcionalidade: login e a principal (no qual ele seleciona a funcionalidade desejada).

Em busca de ter uma maior abrangência na avaliação a ser feita, duas telas/janelas, principalmente, serão avaliadas (sendo as únicas com possível interação para com o usuário já com a funcionalidade aberta): nenhum veículo disponível (figura 1) e com veículos disponíveis (figura 2). Com isso, os tipos de produtos a serem avaliados são essas duas vertentes (sem veículos e com veículos disponíveis) da funcionalidade citada.

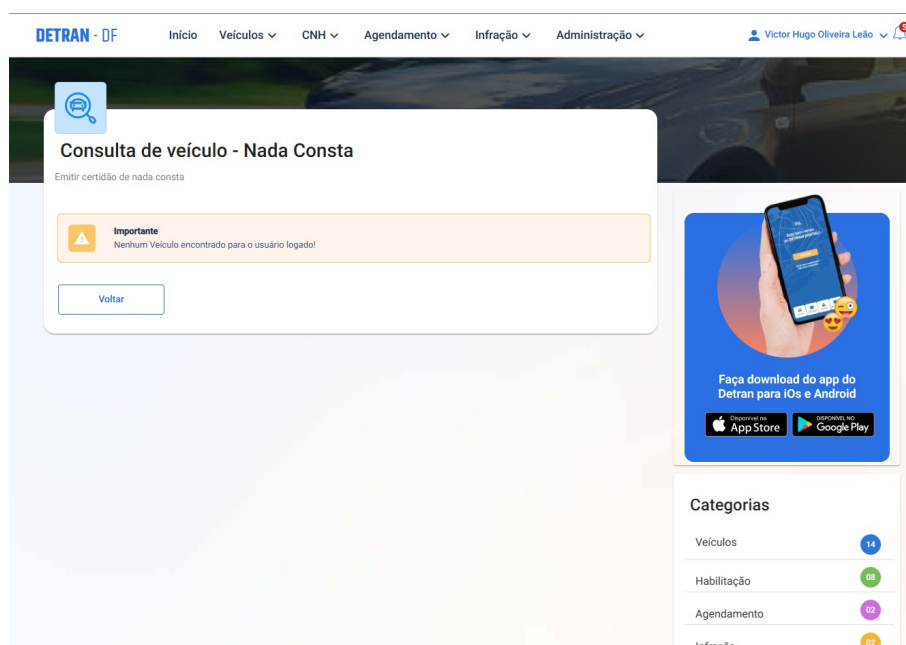


Figure 1. Tela da funcionalidade sem veículos disponíveis

Faz-se importante mencionar que os principais usuários do produto são cidadãos que possuem veículos em seu nome e desejam, periodicamente ou quando necessário, checar se há algo a respeito de débitos vencidos e não quitados em relação ao seu veículo.

Por ser um produto de software que pode receber uma grande gama de usuários, isto é, com diferentes idades, escolaridades e formações, espera-se que o nível de conhecimento exigido necessário em relação à informática é o menor possível, para que todos possam acessar a funcionalidade de forma completa sem muitos problemas. Além disso, em relação ao domínio da aplicação em si, o nível de conhecimento exigido necessário deve ser, pelo menos, básico, já que, para acessar a funcionalidade, deve-se entender que deve haver um veículo em seu nome cadastrado.

Por fim, para o tipo de produto exposto à avaliação (website), os requisitos de hardware e software não são explícitos pelo fornecedor da aplicação (Detran), entretanto,

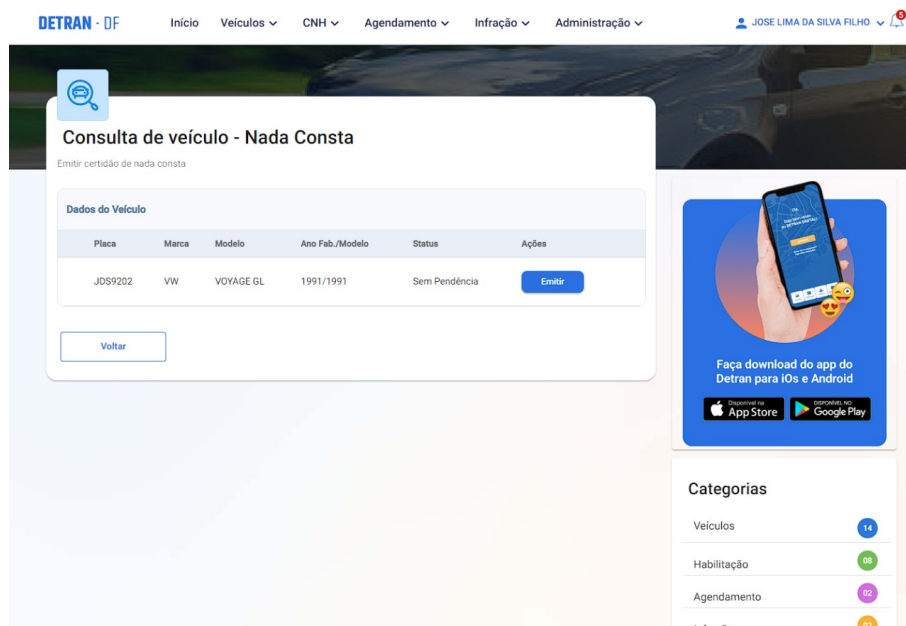


Figure 2. Tela da funcionalidade com veículos disponíveis

pode-se admitir que um computador de pequeno porte com acesso à internet, mouse e teclado seja suficiente como hardware. Já para software, apenas um navegador padrão também é o ideal para acessar a funcionalidade desejada. Foram testados e acertados que os navegadores Microsoft Edge e Google Chrome funcionam para executar o produto de software.

2.1.3. Reporte de erro encontrado na aplicação

Durante a fase de experimentação/avaliação do produto, "Detran Digital", foi encontrado um bug por um dos membros da equipe. Ao executar a "Consulta de veículo - Nada Consta", utilizando cadastro próprio e válido, o avaliador percebeu que o sistema exibe um veículo que não pertence, nem nunca pertenceu, ao mesmo. O erro pode ser visualizado na figura 2, onde o veículo VOYAGE GL é exibido. Isso dificultou a realização da avaliação da qualidade da funcionalidade, uma vez que não foi realmente possível constatar informações.

2.1.4. Modelo de qualidade

O modelo de qualidade a ser utilizado foi extraído da norma SQuaRE (Software Quality Requirements and Evaluation, que em português: Requisitos e Avaliação da Qualidade de Software), que é uma norma constituída por um conjunto de normas com finalidades específicas. Ela foi criada pela motivação de que, com o aumento do uso de softwares, produtos de qualidade são fundamentais de serem desenvolvidos e, portanto, avaliar e constatar constantemente a qualidade é uma tarefa necessária. Dessa forma, o objetivo da SQuaRE, é cobrir, de forma logicamente organizada, enriquecida e unificada, os seguintes processos principais: especificação dos requisitos de qualidade de software e avaliação da

qualidade do software, apoiada por um processo de medição da qualidade do software.

A norma SQuaRE, também conhecida como ISO 25000, trabalha com características e subcaracterísticas para a avaliação de software.

A primeira característica a ser incluída no modelo é a de "Desempenho e Eficiência". Ela estabelece o desempenho de um determinado sistema em relação à quantidade dos recursos utilizados sob condições estabelecidas. Para essa característica, a subcaracterística de "Capacidade" será analisada, ou seja, se o produto consegue performar e suportar o que lhe é solicitado.

A outra característica no modelo de qualidade é a "Usabilidade", que diz respeito que um produto ou sistema deve ser usado por um usuário específico para o alcance de metas específicas com eficácia, eficiência e satisfação em um contexto de uso determinado. Assim, três subcaracterísticas serão avaliadas. Primeiramente, a proteção contra erros do usuário, isto é, como o produto consegue ou oferece oportunidades para prevenir erros do usuário.

Segundamente, a acessibilidade será a segunda subcaracterística a ser avaliada para a "Usabilidade" dos produtos. Para ela, será usado como referência o Web Accessibility Initiative: Ageing Education and Harmonization (WAI-AGE) que consiste em um projeto da Comissão Europeia que recomenda boas práticas para o desenvolvimento de interfaces inclusivas para pessoas idosas baseado no WCAG 2.0. De forma geral, os usuários idosos possuem certas "deficiências" que podem afetar seus usos em softwares Web: visão (redução do constraste, percepção de cores e foco, que dificulta a leitura de páginas Web); habilidade física (redução do controle motor, dificultando o uso de mouse e cliques em alvos pequenos); e habilidade cognitiva (redução da memória curta, dificuldade em concentração e distração, o que dificulta seguir a navegação e completar tarefas). Dessa forma, a subcaracterística de acessibilidade será altamente explorada, seguindo as ideias apresentadas acima.

Por fim, a aprendizagem será também avaliada no escopo de "Usabilidade". Como o próprio nome já explica, a aprendizagem é a capacidade de entender o produto que está sendo usado de forma simples e saber como usar suas funcionalidades de forma adequada sem dificuldades.

Assim, o modelo de qualidade do produto é melhor exposto visualmente na figura 3 abaixo, trabalhando com duas características e suas subcaracterísticas.

2.2. Especificação da avaliação

A especificação da avaliação é uma fase em que as métricas que serão usadas no processo são definidas, além de que seus níveis de pontuação e critérios de julgamentos, isto é, como elas serão interpretadas para o relato dos seus resultados, também são expostos.

Para selecionar as métricas da avaliação, o paradigma Goal Question Metric (GQM), que pode ser traduzido para Objetivo Questão Métrica, será utilizado. De forma simples, essa técnica consiste em medir a partir de um propósito especificado. Com isso, é possível entender quais dados e informações são importantes a organização (ou quem está realizando a avaliação) necessita [Victor R. Basili and Rombach 1994]. O GQM trabalha com três níveis que serão melhor explicados e aplicados para o projeto abaixo.

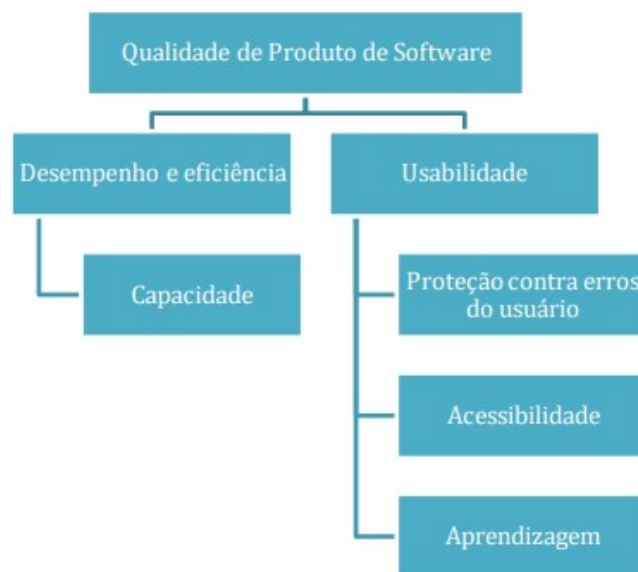


Figure 3. Modelo da qualidade de software a ser usado

Primeiramente, trabalha-se com o nível conceitual, em que um objetivo é definido. Seguindo o que já foi exposto em seções acima, a tabela 1 abaixo apresenta o objetivo como definido pelo GQM:

Table 1. Objetivo da medição GQM

Analisar	Consulta Veículo - Nada Consta no Detran Digital (Site)
para o propósito de	entender e melhorar
Com respeito a	desempenho e eficiência e usabilidade
Do ponto de vista do/a	usuário idoso (65 anos ou mais de idade)
No contexto da	disciplina de qualidade de software

Em seguida, o nível operacional busca encontrar questões que são usadas para caracterizar como o objetivo vai ser performed. Além disso, as questões tentam caracterizar o objeto (nesse caso, a funcionalidade) com respeito às dimensões da qualidade e o ponto de vista exposto. Dessa forma, a partir do objetivo definido, as seguintes questões foram levantadas:

1. Um usuário idoso consegue personalizar os principais elementos na tela da funcionalidade?
2. A funcionalidade possui um tempo de resposta hábil para que os idosos não tenham dificuldade para esperar por respostas longas?
3. A funcionalidade de login (necessária para acessar a consulta de veículo - nada consta) consegue identificar se os dados inseridos estão corretos e avisar de forma intuitiva caso não estejam?
4. A ferramenta apresenta elementos para ajudar o usuário a entender e aprender

a como usar a funcionalidade?

Cada uma das questões busca focar em uma das subcaracterísticas do modelo de qualidade apresentado acima. A primeira está relacionada à acessibilidade, em que, em conjunto com a WAI-AGE, busca-se entender como a aplicação se apresenta para o usuário. Em seguida, a segunda questão trabalha em conforme com a subcaracterística de capacidade (questões mais técnicas). A terceira questão, por sua vez, pensa na proteção contra erros do usuário, subcaracterística de usabilidade. Finalmente, a última pergunta busca explorar a subcaracterística de aprendizagem.

Por fim, o nível quantitativo é desenvolvido, em que as métricas em si são definidas para responder as perguntas feitas. Com isso, os dados levantados pelas métricas devem ser capazes de criar relatos precisos respondendo as perguntas a partir do objetivo. É importante que as métricas tenham as seguintes características para que sejam bem aplicadas: simplicidade, consistência e objetividade, fácil de ser coletada, robustez, validade e independente da linguagem de programação. Dessa maneira, as métricas são expostas na seção abaixo.

2.2.1. Métricas

Para responder a questão 1, as seguintes métricas foram definidas:

1. Porcentagem de zoom disponível com responsividade.
2. Flexibilidade das opções de cores para o plano de fundo (background).

Para responder a questão 2, estabelecemos as seguintes métricas:

3. Tempo de resposta médio (ART - Average Response Time): A média de tempo que o software leva para responder a uma solicitação do usuário.

4. Tempo de resposta máximo (MRT - Maximum Response Time): O tempo máximo que o software leva para responder a uma solicitação do usuário.

Para responder a questão 3, temos as seguintes métricas:

5. Disponibilização da descrição para campos incompletos que são obrigatórios.
6. Disponibilização da descrição quando a informação preenchida pelo usuário não é permitida.

Para responder a questão 4, a seguinte métrica foi escolhida:

7. Quantidade de seções de ajuda para botões na tela da funcionalidade.

2.2.2. Níveis de pontuação para as métricas

Os níveis de pontuação para as métricas se referem aos domínios que as métricas podem ter. Para as métricas definidas, os níveis de pontuação específicos são os seguintes:

Para a métrica 1, os níveis de pontuação estão relacionados à porcentagem de zoom com responsividade aceita pela funcionalidade. Assim, a tabela 2 apresenta a porcentagem e seus respectivos níveis de pontuação.

Para a métrica 2, o nível de pontuação está mais relacionado a como a aplicação permite a flexibilidade e mudança nas opções de cores para o plano de fundo. A tabela 3 mostra o nível de pontuação para essa métrica.

Para as métricas 3 e 4 os níveis de pontuação estão relacionadas com o tempo para a resposta do software. A tabela 4 apresenta os níveis de pontuação para a métrica 3 e a tabela 5 para a métrica 4.

Já as métricas 5 e 6 têm seus níveis de pontuação ligados com a quantidade inteira de descrições (podem ser no formato de aviso) para os campos não preenchidos ou incorretamente preenchidos. A tabela 6 possui o nível de pontuação tanto para a métrica 5, quanto para a 6.

Por fim, a métrica 7 possui um nível de pontuação também relacionado a uma quantidade inteira, mas dessa vez, com a quantidade de seções de ajuda na tela da funcionalidade. Assim, a tabela 7 apresenta esse nível de pontuação.

As tabelas estão disponíveis no apêndice D - Especificação operacional das métricas.

2.2.3. Critérios para julgamento das métricas

Os critérios para julgamento das métricas diz respeito, em essência, como interpretar os resultados das medições.

Para as métricas 1 e 2 de acessibilidade, foi usado como referência o Web Accessibility Initiative: Ageing Education and Harmonization (WAI-AGE) para a elaboração dos níveis de pontuação e os critérios de julgamento. Para a métrica 1, julga-se, pelo ponto G142 do WAI-AGE que o ideal para a porcentagem de zoom é de 200 por cento, logo, ideal seria o critério de julgamento dessa métrica. Para a segunda métrica, o ponto G175 do WAI-AGE explicita que é necessário uma funcionalidade que permite, pelo menos, mudar as cores de background de preto ou branco, já que por problemas de visão, essas cores podem ser ruins de visualização, assim, poder selecionar cores pré-definidas (mediano) é o critério para esse nível.

Para as métricas 3 e 4, o critério de julgamento é o nível de pontuação ideal, uma vez que a resposta deve ser efetiva em tempo hábil, para que os usuários de 65 anos ou mais não tenham que esperar e imaginam que há algum erro com a aplicação.

De semelhante forma, para as métricas 5 e 6, o critério de julgamento é o nível de pontuação ideal, porque o sistema deve ser reportar campos não preenchidos ou preenchidos incorretamente para ambos os casos (login e senha) sendo, então, o usuário possível realizar o login ou ser protegido de cometer erros. Essas definições foram feitas seguindo os pontos G83 e G84 da WAI-AGE.

Enfim, para a métrica 7, o critério de julgamento varia caso o usuário possua um veículo cadastrado ou não. Caso não, somente há um botão disponível (voltar), então essa opção seria ideal. Caso sim, os dois botões disponíveis devem haver uma seção de ajuda, sendo assim, ideal, também.

2.3. Projeção da avaliação

A fase de projeção de avaliação ocorre após a definição dos requisitos de avaliação, isto é, o seu propósito, os tipos de produtos a serem avaliados e o modelo de qualidade a ser seguido, e a especificação das métricas. O objetivo da projeção é produzir o plano de avaliação, especificando o método a ser utilizado e as instruções seguidas para sua utilização. Os recursos necessários para a realização da avaliação também são especificados no plano de avaliação.

2.3.1. Plano de avaliação

Inicialmente, é necessário definir o método de avaliação para a obtenção de todas as métricas definidas para a avaliação da qualidade. A intenção inicial dos autores e realizadores da avaliação era obter dados a partir de usuários idosos, seja por uma entrevista, formulário e/ou simulação, porém, por falta de disponibilidade, uma vez o único usuário com a funcionalidade acessível encontrado foi, como reportado na seção 2.1.3., afetado por um erro do sistema, não foi possível utilizar nenhum desses métodos de avaliação.

Dessa forma, o método de avaliação diferiu para cada uma das métricas, mas todas foram obtidas de forma manual pelos realizadores da avaliação ou com ferramentas de automatização. Nos parágrafos abaixo é possível encontrar a definição e explicação do método de avaliação especificando como ele se aplica para cada uma das métricas.

Primeiramente, deve-se realizar o login na plataforma web do "Detran Digital" no navegador Microsoft Edge para ter acesso à funcionalidade "Consulta Veículo - Nada Consta", objeto da avaliação. Assim, ao acessar a funcionalidade, para medir a métrica 1, confira se o zoom da página está em 100 por cento (padrão) e, então, aplica-se zoom de 150 por cento na tela, faz-se a avaliação da responsividade e, posteriormente, aplica-se zoom de 200 por cento para, também, analisar a responsividade da aplicação. Em seguida, para a métrica 2, deve-se, também na tela da funcionalidade, buscar alguma opção de alteração de cor do background (plano de fundo) e, caso exista, é importante analisar como se dá a alteração das cores (são pré-definidas ou de escolha do usuário).

Para as métricas 3 e 4, em que é necessário obter o tempo de resposta do sistema, nesse caso o site do "Detran Digital" na funcionalidade "Consulta Veículo - Nada Consta", é necessário usar uma ferramenta/software automatizadas para calcular esse tempo. Assim, recomenda-se o uso da ferramenta Pingdom Website Speed Test[], em que é possível obter o tempo de carregamento de um site em um servidor no Brasil. Com a ferramenta disponível, deve-se obter o tempo 5 vezes, para então, calcular a média de tempo de resposta e, também, definir o tempo máximo de resposta, conseguindo, com isso, as medidas.

Como as métricas 5 e 6 estudam as caixas de entrada de texto da funcionalidade de login, caso a aplicação esteja autenticada em algum usuário, é necessário a realização do logout. Assim, na tela de login, primeiramente para a métrica 5, tente avançar sem preencher a área de login (CPF) e veja se alguma descrição ou aviso é dado, em seguida, faça o mesmo teste para a mesma análise na área de senha. Para a métrica 6, reinicie o fluxo (logout), assim, é necessário tentar inserir caracteres não numéricos no campo de CPF e analisa-se se algum aviso ou descrição foi indicado, assim, para a senha, tenta-se

a inserção de caracteres não relacionados à língua portuguesa e analisa se houve alguma descrição pelo sistema.

Por fim, para a métrica 7 é necessário estar com login realizado, primeiramente, em uma conta sem veículos cadastrados, e na tela da funcionalidade. Assim, é necessário analisar se o botão de 'voltar' possui alguma seção de ajuda para explicar seu funcionamento. Para a segunda obtenção dessa métrica, deve-se, em uma conta com veículos cadastrados, analisar o botão de 'voltar', mas também o de 'emitir' para ver se alguma seção ou ferramenta traz aprendizagem para alguma dessas ações.

Para realizar os métodos de avaliação especificados acima são necessários recursos (alguns já podem ser extraídos pelos fluxos dos métodos). Estes são:

1. Conexão estável à internet;
2. Dispositivo Desktop com acesso a navegador web, de preferência Microsoft Edge (utilizado para obtenção das métricas);
3. Conta com veículo não cadastrado;
4. Conta com veículo cadastrado;
5. Mouse padrão;
6. Teclado padrão.

2.4. Execução da avaliação

Com o plano de avaliação traçado, ele é executado e, então, as medidas são obtidas e é possível compará-las com os critérios de julgamento selecionados e explicitados anteriormente. Desse modo, pode-se verificar o software de acordo com os requisitos de avaliação, como descrito na especificação da avaliação.

2.4.1. Obtenção das medidas

As medidas obtidas a partir da execução da avaliada estão presentes na planilha da imagem 4, presente no apêndice D.

2.4.2. Comparação com critérios

A partir das medidas obtidas, é possível comparar o que foi realmente medido com os critérios de julgamento previamente decididos:

1. Métrica 1: A medida obtida para a métrica 1 atinge o critério ideal definido (200 por cento).
2. Métrica 2: A medida obtida para a métrica 2 não atinge o critério ideal definido (usuário pode definir a cor), sendo, na verdade, nenhuma disponibilidade para alteração de cores (ruim).
3. Métrica 3: A medida obtida para a métrica 3 não atinge o critério ideal definido (menos de 100ms), sendo, na verdade, 1.09s (ruim).

4. Métrica 4: A medida obtida para a métrica 4 não atinge o critério ideal definido (menos de 500ms), sendo, na verdade, 1.15s (mediano).

5. Métrica 5: A medida obtida para a métrica 5 atinge o critério ideal definido (2 descrições).

6. Métrica 6: A medida obtida para a métrica 6 não atinge o critério ideal definido (2 descrições), sendo, na verdade, 0 (ruim).

7. Métrica 7: A medida obtida para a métrica 7 não atinge o critério ideal definido (2 seções de ajuda), sendo, na verdade, 0 (ruim).

2.4.3. Julgamento dos resultados

Os resultados esperados eram que, pelo menos para maioria (se não todas) das medidas, o nível de pontuação atingido fosse o ideal, uma vez que o sistema analisado, como já explorado anteriormente, é altamente utilizado pela parte da população brasileira que possui algum veículo. Com isso em mente, as funcionalidades (em especial analisada a de Consulta Veículo - Nada Consta) deviam possuir uma implementação que atingisse níveis satisfatórios para métricas relacionadas às características de desempenho e eficiência e usabilidade e suas subcaracterísticas estudadas durante a avaliação da qualidade.

3. Resultados da avaliação

Nessa seção os resultados da avaliação serão relatados seguindo a visão das características e subcaracterísticas do modelo de qualidade já exposto. Assim, a partir do processo de avaliação e obtenção das medidas, estas serão relatadas no contexto específico de cada uma das dimensões que elas fazem referência.

3.1. Desempenho e Eficiência

A primeira característica a ter seus resultados relatados é a de desempenho e eficiência. A sua única subcaracterística avaliada durante todo o processo foi a de capacidade.

3.1.1. Capacidade

Para analisar a capacidade do sistema, duas métricas foram definidas para fazer a avaliação dessa subcaracterística: tempo de resposta médio (ART - Average Response Time), que é a média de tempo que o software leva para responder a uma solicitação do usuário; e tempo de resposta máximo (MRT - Maximum Response Time), que é o tempo máximo que o software leva para responder a uma solicitação do usuário. Com essas métricas, então, seria possível entender se o sistema possui um tempo de resposta agradável para usuários de 65 anos ou mais.

Dessa forma, foram avaliados que o ideal para essas métricas seriam um ART de menos de 100ms e um MRT de menos de 500ms como ideais, uma vez que, assim, as respostas às solicitações de um usuário seriam em tempo hábil e, portanto, não haveriam, principalmente para os idosos, nenhum problema ou erro envolvendo a subcaracterística de capacidade.

Entretanto, a partir da obtenção das medidas, percebeu-se que os critérios de julgamento ideais não foram atingidos para essas métricas, o que traz o resultado de que, para a capacidade no quesito de tempos de resposta, a aplicação não possui a qualidade desejada.

3.2. Usabilidade

A segunda e última característica com resultados para serem relatados é a de usabilidade. Diferentemente da característica de desempenho e eficiência, a usabilidade foi estudada em três subcaracterísticas diferentes: proteção contra erros do usuário, acessibilidade e aprendizagem.

3.2.1. Proteção contra erros do usuário

Para a subcaracterística de proteção contra erros do usuário foram definidas duas métricas relacionadas à funcionalidade de realizar login, mais especificamente no preenchimento dos dados de autenticação (CPF e senha). Essas métricas foram: disponibilização da descrição para campos incompletos que são obrigatórios; e disponibilização da descrição quando a informação preenchida pelo usuário não é permitida. Dessa maneira, com essas métricas, seria possível avaliar se a aplicação previne os erros de usuários ao tentarem acessar a funcionalidade de Consulta Veículo - Nada Consta, já que ela necessita da autenticação de login.

Assim, os critérios de julgamento considerados ideais tinham a disponibilização de uma descrição ou aviso tanto para o campo de login (CPF), quanto para o de senha, caso o usuário inserisse uma informação incompleta em um campo obrigatório ou caso a informação não fosse permitida. Dessa forma, mesmo usuários idosos (65 anos ou mais) que pudessem ser suscetíveis a cometer erros ao tentar preencher essas informações para realizarem a sua autenticação seriam amparados (ou pelo menos alertados) pelo sistema de seus erros.

É importante mencionar que essas duas métricas seguem, respectivamente, os pontos G83 e G84 da WAI-WCAG, que descrevem como devem ser

A primeira métrica dessa subcaracterística conseguiu atingir o nível de satisfação desejado. Por outro lado, a segunda métrica teve um mau desempenho. Com isso, pode-se concluir que para o aviso de tentar realizar um login sem preencher os campos necessários, o sistema protege contra erros do usuário com sucesso. Já para o preenchimento de campos com informação não permitida, como, por exemplo, letras em um campo de CPF (apenas números), o sistema não fornece nenhum tipo de auxílio, não prevenindo o erro do usuário nesse quesito.

3.2.2. Acessibilidade

A subcaracterística de acessibilidade foi avaliada seguindo o Web Accessibility Initiative: Ageing Education and Harmonization (WAI-AGE) e, assim, as seguintes métricas foram definidas para analisar de forma mais específica essa dimensão: porcentagem de zoom disponível com responsividade; e flexibilidade das opções de cores para o plano de fundo

(background). Essas métricas seguem, respectivamente, os pontos G142 e G175 da WAI-AGE.

Os critérios de julgamento ideais para essas métricas definiam um zoom de 200 por cento com responsabilidade e a possibilidade do usuário escolher a cor do plano de fundo que desejasse. Assim, usuário com problemas de visão poderiam utilizar a funcionalidade com maior acessibilidade e melhor experiência.

Dessa maneira, a primeira métrica dessa subcaracterística cumpriu o critério de julgamento delineado, uma vez que com um zoom de 200 por cento, o usuário ainda conseguia mexer na aplicação com responsividade e completo acesso. Todavia, a segunda métrica não conseguiu atingir nem o nível mediano, já que não é possível alterar o background de maneira alguma no sistema. Assim, usuários idosos, principalmente aqueles com problemas no sentido da visão, poderiam ter problemas ao acessar o software.

3.2.3. Aprendizagem

Por fim, a última subcaracterística explorada durante o processo de avaliação foi a aprendizagem. Para ela foi definida uma métrica: quantidade de seções de ajuda para botões na tela da funcionalidade. O objetivo dessa métrica era entender se a aplicação possuía algum tipo de suporte para que usuário que não conhecesse a funcionalidade fossem capaz de aprender e utilizar os elementos que realizam ações (botões).

Assim, o critério de julgamento satisfatório para essa métrica seria que todos os botões presentes na tela da funcionalidade tivessem uma seção de ajuda que pudessem, nem que de forma breve, explicar o que aconteceria caso acionados. Assim, os usuários de 65 anos ou mais que não fossem familiarizados com aplicações web e, a do "Detran Digital", principalmente, seriam capazes de aprender acerca da funcionalidade e ter uma melhor experiência de uso.

Contudo, a medida obtida não cumpria o desejado, já que nenhum botão na tela da funcionalidade possuía uma seção de ajuda. Desse modo, a aprendizagem na visão dessa métrica falhou.

3.3. Discussão dos resultados

Por meio dos resultados obtidos e relatados acima, os autores conseguem fazer algumas observações, como: pontos fortes, oportunidades de melhoria e recomendações.

3.3.1. Pontos fortes

A partir de toda a avaliação feita e dos resultados, 2 pontos fortes podem ser observados no sistema:

1. A preocupação com a responsividade é evidenciada e é de suma importância, já que, cada vez mais, usuários usam diferentes tipos de dispositivos e tamanhos de telas para acessar aplicações web. Dessa maneira, esse ponto forte pode ser destacado e traz grande credibilidade para a funcionalidade Consulta Veículo - Nada Consta do site "Detran Digital".

2. Adicionalmente, o processo de login integrado com o gov.br funciona de forma intuitiva e um usuário com acesso a seu login (CPF) e senha conseguem, caso preenchendo tudo de forma correta, acessar de forma simples sua conta. O fornecimento de avisos sobre erros de campos não preenchidos funcionam de forma excelente.

3.3.2. Oportunidades de melhoria

Assim como todo sistema, a funcionalidade avaliada no site possui pontos que estão no caminho correto, mas que poderiam trabalhar com certas melhorias. Assim, é destacado abaixo 1 oportunidade de melhoria:

Um usuário que tenta realizar seu login na funcionalidade, mas tenta preencher caracteres inválidos no campo de CPF não consegue preencher, o que é correto. Entretanto, seria interessante caso uma descrição, como pontuado no G84 do WAI-WCAG, que explicasse a necessidade do preenchimento somente com números seria de suma importância para proteger contra erros do usuário ainda mais.

3.3.3. Recomendações

Por fim, serão feitas recomendações para que o sistema possa atingir critérios ideais nas dimensões estudadas:

1. Deve-se implementar uma funcionalidade que garanta que o usuário possa alterar a cor do background da aplicação, uma vez que cores padrões (branco e preto) podem atrapalhar e dificultar a visão e leitura por parte de alguns usuários.

2. A capacidade, principalmente no que convém ao tempo de resposta, deve ser melhorada, talvez pela contratação de um servidor de hospedagem melhor alocado. Isso melhoraria a percepção de desempenho e eficiência do sistema para usuários e, principalmente para os idosos, eles não teriam menor percepção de erros em relação à aplicação.

3. A implementação de seções de ajuda atreladas a ações que o programa executa é de suma importância para que usuários que não entendem como executar tais ações possam aprender acerca das funcionalidades presentes.

4. Considerações Finais

Em conclusão, pode-se dizer que o propósito inicial da avaliação de medir a qualidade da funcionalidade de "Consulta Veículo - Nada Consta" do site do "Detran Digital" sob a perspectiva do usuário idoso foi atingida com sucesso. Após explicitar os tipos de produtos a serem avaliados e qual o modelo de qualidade seria utilizado, foi possível especificar as métricas seguindo as características e subcaracterísticas definidas. Assim, apesar de nenhum método de avaliação que usasse diretamente um usuário final da aplicação ser utilizado, foi possível obter medidas que possibilitaram as análises das métricas e responder as perguntas propostas no GQM do projeto.

5. Referências

- DETRAN, DENATRAN, CONTRAN: O QUE FAZ CADA ÓRGÃO DO SISTEMA NACIONAL DE TRÂNSITO?. Zul Digital, 2019. Disponível

em: <https://www.zuldigital.com.br/blog/detran-denatran-contran-orgao-transito/>. Acesso em: 26 de jan. de 2023.

- Victor R. Basili, G. C. and Rombach, H. D. (1994). Goal Question Metric Paradigm. John Wiley Sons, Inc.
- How to Meet WCAG (Quick Reference). Web Accessibility Initiative- WAI, 2019. Disponível em: <https://www.w3.org/WAI/WCAG21/quickref/>. Acesso em: 26 de jan. de 2023.
- ISO/IEC 25000:2005(en) Software Engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Guide to SQuaRE. ISO, 2005. Disponível em: <https://www.iso.org/obp/ui/iso:std:iso-iec:25000:ed-1:v1:en>. Acesso em: 26 de jan. de 2023.

6. Apêndice A - Tabela de contribuição da equipe

Nome	Contribuição
Victor Hugo O. Leão	30%
Paulo Henrique A. da Silva	14%
Felipe Alef P. Rodrigues	14%
Guilherme A. de La Fuente V.	14%
Lucas Ramon A. de Oliveira	14%
Abraão A. Ribeiro	14%

Figure 4. Contribuição Membros

7. Apêndice B - Relato de experiência dos estudantes

Victor Hugo O. Leão: Atuei em todas as seções do projeto e consegui tirar muito aprendizado de como uma avaliação da qualidade deve ser feita, tanto na teoria, quanto na prática. Além disso, gostei muito de aplicar o GQM em um projeto verdadeiro. Diria que nosso grupo teve certas dificuldades em realmente usar a funcionalidade e encontrar usuários que encaixassem no escopo da avaliação para poder coletar métricas melhor utilizáveis no trabalho, então, talvez, a funcionalidade possa ser algo a ser pensado na especificação do trabalho. Entretanto, de qualquer forma, acredito que o intuito e propósito foram cumpridos.

Abraão A. Ribeiro: Participei da construção dos subtópicos da projeção da avaliação especificando os métodos a serem utilizados e também seus resultados. Houve dificuldade em escolher quais métodos seriam utilizados no projeto já que se refere a uma atividade que requer a opinião de todos os integrantes do grupo.

Felipe Alef P. Rodrigues: Atuei em todos os tópicos do trabalho e gostei muito de utilizar, na prática, os conceitos da disciplina e fazer uma avaliação de qualidade. Uma dificuldade que enfrentamos foi encontrar pessoas para coletar as métricas. Algo que poderia ser mudado seria a escolha da funcionalidade, que poderia ser feita pelo grupo e aprovada pela professora, para garantir que o grupo consiga utilizar a funcionalidade corretamente.

Guilherme A. de La Fuente V.: Participei da construção do documento utilizando dos conhecimentos aprendidos em sala. Foi muito interessante poder colocar em prática os assuntos que vimos e avaliar a qualidade de um projeto real. Sugiro a possibilidade de

mudar como é feita a escolha da funcionalidade para cada equipe, deixando com que o próprio grupo possa ter essa decisão.

Lucas Ramon A. de Oliveira: Eu participei de todos os aspectos do trabalho e foi muito gratificante poder aplicar os conceitos aprendidos na disciplina e realizar uma análise de qualidade. Uma dificuldade que encontramos foi a falta de dados de métricas precisos e confiáveis. Devido a falta de ferramentas adequadas para coletar e analisar esses dados, ou a dificuldade em encontrar usuários dispostos a participar da avaliação.

Paulo Henrique A. da Silva: Participei do desenvolvimento de todos os tópicos deste trabalho, com isso pude aplicar de fato os conhecimentos adquiridos em sala e os demais necessários para a conclusão do mesmo. Houve certa dificuldade em conseguir utilizar o sistema do Detran Digital pois não possuíamos cadastro, entretanto, solucionamos o problema e cumprimos o proposto.

8. Apêndice C - Coleta de Dados

Abaixo pode ser encontrada a tabela com a coleta das métricas referenciadas na seção de obtenção das medidas (2.4.1.).

Métrica 1	200%					
Métrica 2	Nenhuma					
Métrica 3	1.09seg	1.15seg	1.15seg	1.02seg	1.02seg	1.11seg
Métrica 4	1.15seg					
Métrica 5	2					
Métrica 6	0					
Métrica 7 (carro cadastrado)	0					
Métrica 7 (semcarro cadastrado)	0					

Figure 5. Métricas obtidas

9. Apêndice D - Especificação operacional das métricas

Abaixo estão presentes as tabelas referenciadas na seção de nível de pontuação (2.2.2.).

Table 2. Nível de pontuação métrica 1

Porcentagem	Nível de Pontuação
100% (padrão)	Ruim
150%	Mediano
200%	Ideal

Table 3. Nível de pontuação métrica 2

Disponibilidade para alteração de cores	Nível de Pontuação
Nenhuma	Ruim
Cores pré-definidas	Mediano
Usuário pode definir a cor	Ideal

Table 4. Nível de pontuação métrica 3

Tempo	Nível de Pontuação
500ms	Ruim
200ms a 500ms	Mediano
100ms a 200ms	Bom
Menos de 100ms	Ideal

Table 5. Nível de pontuação métrica 4

Tempo	Nível de Pontuação
Mais de 2seg	Ruim
1seg a 2seg	Mediano
500ms a 1seg	Bom
Menos de 500ms	Ideal

Table 6. Nível de pontuação métricas 5 e 6

Quantidade de descrições	Nível de Pontuação
0 (nem login, nem senha)	Ruim
1 (para login OU senha)	Mediano
2 (para login E senha)	Ideal

Table 7. Nível de pontuação métrica 7

Quantidade de seções de ajuda	Nível de Pontuação
0 (para nenhum botão)	Ruim
1 (para botão de voltar)	Mediano/Ideal
2 (para botão de voltar e emitir)	Ideal