



UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO (UNIRIO)
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA (CCET)
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA (PPGI)

PAULO JOSÉ DE ALCANTARA GIMENEZ

**BUSCA COMO PROCESSO DE APRENDIZAGEM (BPA)
NO CONTEXTO DE COMUNIDADES DE PERGUNTAS E
RESPOSTAS (CQA)**

RIO DE JANEIRO

2022



PAULO JOSÉ DE ALCANTARA GIMENEZ

**BUSCA COMO PROCESSO DE APRENDIZAGEM (BPA)
NO CONTEXTO DE COMUNIDADES DE PERGUNTAS E
RESPOSTAS (CQA)**

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre pelo Programa de Pós-Graduação em Informática da Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Sub-área de concentração 1.03.03.04-9 Sistemas de Informação.

Orientador: Sean Wolfgang Matsui Siqueira

Rio de Janeiro

2022

PAULO JOSÉ DE ALCANTARA GIMENEZ

**BUSCA COMO PROCESSO DE APRENDIZAGEM (BPA)
NO CONTEXTO DE COMUNIDADES DE PERGUNTAS E
RESPOSTAS (CQA)**

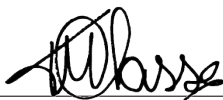
Dissertação apresentada como requisito parcial para a obtenção do título de Mestre pelo Programa de Pós-Graduação em Informática da Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Sub-área de concentração 1.03.03.04-9 Sistemas de Informação.

Aprovado em 01 / 04 / 2022.

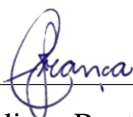
Banca Examinadora:



Prof. Dr. Sean Wolfgang Matsui Siqueira (Orientador)
Universidade Federal do Estado do Rio de Janeiro - UNIRIO.



Prof. Dr. Tadeu Moreira de Classe
Universidade Federal do Estado do Rio de Janeiro - UNIRIO.



Prof. Dr. Juliana Baptista dos Santos França
Universidade Federal do Rio de Janeiro - UFRJ.

G491 Gimenez, Paulo José de Alcantara.
 Busca como Processo de Aprendizagem (BPA) no Contexto de Co-
 munitades de Perguntas e Respostas (CQA)/ Paulo José de Alcantara
 Gimenez. – Rio de Janeiro, 2022.
 185f

 Orientador: Sean Wolfgang Matsui Siqueira.
 Dissertação (Mestrado) — Universidade Federal do Estado do Rio de
 Janeiro, Programa de Pós-Graduação em Informática, 2022.

 1. Sistemas de Informação. 2. Comunidades de Perguntas e Respos-
 tas. 3. Aprendizagem informal. 4. Busca como processo de aprendizagem.
 5. Busca na Web. I. Siqueira, Sean Wolfgang Matsui, orient. II. Título.

Dedico este trabalho à Deus, a minha família, aos meus pais e aos amigos fraternos.

“Não sabendo que era impossível, foi lá e fez.” (Jean Cocteau)

“Tudo parece impossível até que seja feito.” (Nelson Mandela)

AGRADECIMENTOS

Agradeço à Deus a energia e perseverança para completar essa etapa de minha vida, encarando a pandemia Covid-19, com seus reflexos sobre todos os aspectos de nossas vidas, e particularmente neste contexto acadêmico, levar-nos a uma nova realidade de estudo e práticas.

Agradecimento especial à minha querida esposa Alice Maria pelo carinho, apoio, incentivo e compreensão durante essa trajetória excepcionalmente estressante e aflitiva, tumultuada pela pandemia que ainda vivenciamos. Seu equilíbrio e comprometimento permitiram o suporte necessário a esta empreitada, que muitas vezes me afastou das atividades familiares.

Agradeço o carinho e compreensão de meus filhos. Ao Luiz Paulo, que em sua tenra infância e a seu modo amoroso soube entender os momentos de afastamento de seu pai. À minha filha Helena, que no auge de sua adolescência e com sua compreensão de mundo pode me incentivar com o seu exemplo de dedicação, nos momentos mais difíceis dessa trajetória.

À minha mãe Adilsa, pelo amoroso exemplo de força e fibra, que com sua experiência de vida, serviu de referência para mim. Ao meu pai saudoso e querido José, que manifestava sempre satisfação e admiração pelas minhas conquistas, por mais simples que fossem.

Ao meu orientador e amigo, Sean, pelo apoio, incentivo e desafios, bem como o tempo dedicado em reuniões presenciais, ainda quando era possível, e principalmente à distância, investindo na minha qualificação como pesquisador, enxergando um potencial que nem mesmo eu percebia, principalmente quando adentrei no campo de pesquisas qualitativas como no presente trabalho. Professor Sean, suas contribuições, considerações e orientações foram críticas para o sucesso dessa jornada até aqui.

Aos colegas do grupo de pesquisa *SAL — Semantics and Learning* do PPGI/DIA/UNIRIO, que me aturaram e providenciaram *feedbacks* durante toda a minha trajetória de pesquisa, fortalecendo o meu papel de pesquisador e a minha pesquisa. Nossos contatos via *WhatsApp*, *Discord*, *Zoom* e *Google Meet* foram o esteio socio-técnico para depuração dessa pesquisa. Um destaque especial ao Marcelo Machado e ao Cleber pelos nossos artigos publicados, bem como ao Marcelo Tibau Tibau e ao Marcelo Loutfi pelas contribuições à presente pesquisa.

Aos participantes voluntários da pesquisa, alunos do PPGI — ou agora egressos, meu

profundo agradecimento e respeito pela disponibilidade e empenho na participação com naturalidade das atividades e sessões, que formaram a base da pesquisa, mesmo em um cenário pandêmico e de ajustamento de nossas vidas. Também enalteço a confiança depositada na seriedade desta pesquisa, pois muitos de vocês nem sequer me conheciam, mas deram um voto de confiança no empreendimento.

Aos professores membros da banca, que mesmo em um período de férias, estavam disponíveis para participar de minha defesa, apresentando contribuições relevantes para o aprimoramento desta pesquisa.

Aos membros da infraestrutura administrativa e acadêmica do PPGI, da secretaria às coordenações diversas, agradeço a seriedade e retidão no pronto atendimento de todas as necessidades colocadas.

Este estudo foi financiado em parte pelo 'Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) — Brasil' — Processo 315374/2018–7, projeto 'Busca como aprendizagem: a busca de informações como ferramenta de aprendizagem'.

Gimenez, Paulo José de Alcantara. **Busca como Processo de Aprendizagem (BPA) no Contexto de Comunidades de Perguntas e Respostas (CQA)**. 2022. 185 f. Dissertação (Mestrado) - Departamento de Informática Aplicada, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2022.

RESUMO

As Comunidades de perguntas e respostas (CQA) cumprem um importante papel de fonte de informação confiável e possuem conhecimentos *crowdsourced*, validados por pares de especialistas nos assuntos na própria comunidade. Normalmente elas são acionadas a partir de mecanismos de buscas da Web, devido à natureza da busca em formato de pergunta, que aproxima as buscas das postagens em meio a demais fontes como resultados da busca. Em uma busca mais complexa, como nos cenários de busca como um processo de aprendizagem (BPA), de caráter predominantemente exploratório, há uma exigência metacognitiva maior, face à necessidade de navegar, filtrar, selecionar, avaliar, agregar e explorar muitos resultados da busca sobre o conhecimento (nestes casos, geralmente técnico) necessário. Diferente de outros trabalhos relacionados, que fazem uso extensivo de questionários, *logs* e dados massivos de pouca profundidade, neste estudo, adotamos uma abordagem qualitativa através da Teoria Fundamentada em Dados Construtivistas, em que utilizamos o *Think Aloud Protocol*(TAP) e a *Intensive Interviewing* na coleta dos dados, para obter um conjunto de dados mais rico a partir da interação de usuários com o ambiente utilizado, representando a experiência dos participantes na busca por informação através da Web com base em CQA. O ambiente para a aplicação da pesquisa foi construído através da customização do sistema de buscas *Google Search* para utilização nas sessões práticas das buscas. Para avaliar com maior ênfase a relevância da CQA nesse cenário, controlamos o fluxo de informações para os usuários ao restringir os resultados às principais CQA levantadas da literatura e Internet. A análise, considerou a interação e entrevistas com dezessete participantes de pós-graduação em sistemas de informação, representando 10% da população total, que atendiam ao critério de saturação teórica definido. Como resultado, temos principalmente a contribuição de uma análise qualitativa da aprendizagem de usuários a partir da busca Web como processo de aprendizagem utilizando CQA. Além disso, temos contribuições científicas pela extensão da taxonomia de estados de consulta e do modelo de busca exploratória ESKiP, técnicas pelo ambiente customizado e os processos de transcrição automatizada, e teóricas pela associação dos Princípios de Influência de Autoridade e Consenso com a compreensão da CQA através da Teoria de Comunidades de Práticas. Como achados deste estudo, temos que os

participantes com maior familiaridade com a CQA a usam com maior frequência, as perguntas detalhadas e respostas bem elaboradas contribuem para a aprendizagem, existe maior dificuldade no alinhamento da intenção de busca com a avaliação da pergunta na trilha quando envolve a compreensão de conceitos, e que há uma percepção majoritária de inadequação das trilhas em CQA para aprendizagem de indivíduos iniciantes no assunto, requerendo a existência de outros tipos de conteúdos para preencher a lacuna de conhecimento.

Palavras-chave: Busca na Web; Comunidades de Perguntas e Respostas; Aprendizagem Informal; Busca como processo de aprendizagem; Sistemas de Informação.

Gimenez, Paulo José de Alcantara. **Search as a Learning Process (BPA) in the Context of Question and Answer Communities (CQA)**. 2022. 185 f. Dissertação (Mestrado) - Departamento de Informática Aplicada, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2022.

ABSTRACT

Question and Answer Communities (QAC) fulfill an important role as a trusted source of information and have *crowdsourced* knowledge, validated by peers of subject-matter experts in the community itself. They are typically triggered from Web search engines, due to the question-format nature of the search, which brings search queries closer to the posts in the midst of other sources such as search results. In a more complex search, such as in search as learning process (BPA) scenarios, which are predominantly exploratory in nature, there is a greater metacognitive requirement, given the need to navigate, filter, select, evaluate, aggregate, and explore many search results on the knowledge needed (in these cases, usually technical). Unlike other related works, which make extensive use of questionnaires, *logs* and massive data of little depth, in this study, we adopted a qualitative approach through the Constructivist Grounded Theory, in which we used the *Think Aloud Protocol* (TAP) and the *Intensive Interviewing* in data collection, to obtain a richer dataset from the interaction of users with the environment used, representing the participants' experience in the search for information through the Web based on CQA. For this research, we built a program environment through the customization of the search system *Google Search* for use in practical sessions of searches. To further evaluate the relevance of CQA in this setting, we controlled the flow of information to users by restricting the results to the main CQA surveyed from the literature and the Internet. The analysis considered the interaction and interviews with seventeen postgraduate participants in information systems, who represents 10% of the total population and met the defined theoretical saturation criterion. As a result, we have mainly the contribution of a qualitative analysis of users' learning from the Web search as a learning process using CQA. In addition, we have scientific contributions through the extension of the query state taxonomy and the ESKiP exploratory search model, techniques through the customized environment and automated transcription processes, and theoretical contributions through the association of the Principles of Influence of Authority and Consensus with the understanding of CQA through the Theory of Communities of Practice. As findings of this study, we have that the participants with greater familiarity with the CQA use it more frequently, the detailed questions and well-crafted answers contribute to learning, there is greater difficulty in

aligning the search intent with the evaluation of the question on the trail when involves the understanding of concepts, and that there is a majority perception of inadequacy of the CQA trails for the learning of individuals who are new to the subject, requiring the existence of other types of content to fill the knowledge gap.

Keywords: Web searching; Question and Answer Communities; Informal Learning; Searching as Learning; Information systems.

LISTA DE ILUSTRAÇÕES

Figura 1 – Tela da Busca Customizada do <i>Google Search</i> para o Experimento. . .	36
Figura 2 – Exemplo de Codificação Inicial	45
Figura 3 – Codificação Inicial — Utilização da CQA. Participante P21915	54
Figura 4 – Razão ou Motivo para Uso de CQA	54
Figura 5 – Utilização da CQA — Razão x Finalidade x Frequência	56
Figura 6 – Processo de Busca — Concepção da Busca.	61
Figura 7 – Busca Livre — Seleção dos termos e texto da Busca.	62
Figura 8 – Processo de Busca: elementos da formulação da consulta, estendido do <i>ESKiP</i>	63
Figura 9 – Processo de Busca — Critérios de checagem utilizados.	64
Figura 10 – Evidências de critérios de checagem utilizados.	65
Figura 11 – Processo de Busca — Extração da Informação: Seleção e Exibição. . .	65
Figura 12 – Evidências de critérios de checagem utilizados para Tomada de Deci- são. (Atividade Condicionada Direta)	85
Figura 13 – Combinação de Estratégias de Concepção da Busca por Participantes em Atividade de Múltiplas Tarefas.	90
Figura 14 – Combinação de Formato de Seleção dos termos por Participantes em Atividade de Múltiplas tarefas.	91
Figura 15 – Combinação de Artíficos de Direcionamento de Seleção dos termos por Participantes em Atividade de Múltiplas tarefas.	92
Figura 16 – Combinação de Estratégias de Exibição dos Resultados da Busca por Participantes em Atividade de Múltiplas tarefas.	95
Figura 17 – Frequência de foco em elementos da CQA por participante — Ativi- dades da Sessão#3	97
Figura 18 – Correlação de elementos de foco na análise da trilha pelos participan- tes (Atividade Condicionada Múltipla)	100
Figura 19 – Evidências de critérios de checagem utilizados para Tomada de Deci- são. (Atividade Condicionada Múltipla)	101
Figura 20 – Fragmento da Postagem (1. ^a QUESTÃO) sobre conceitos da Lingua- gem C	110
Figura 21 – Fragmento da Postagem (2. ^a QUESTÃO) sobre implementação em lin- guagem C	112
Figura 22 – Impacto das atividades no comportamento ou conhecimento do parti- cipante.	119

Figura 23 – Comparação de Similaridade da Busca e Analítico na CQA nas três sessões	136
---	-----

LISTA DE TABELAS

Tabela 1 – Comparativo de Características entre as principais CQA	38
Tabela 2 – Questões de orientação para as Entrevistas	39
Tabela 3 – Perfil e Heterogeneidade dos Participantes	42
Tabela 4 – Marcadores da Transcrição	43
Tabela 5 – Similaridade Comportamental entre os Participantes no Processo de Busca (Atividade Livre).	67
Tabela 6 – Sumarização do Comportamento dos Participantes no Processo de Busca (Atividade Livre).	67
Tabela 7 – Estratégia de leitura da trilha CQA (Atividade Livre)	69
Tabela 8 – Elementos da trilha CQA de Foco de atenção pelo Participante (Atividade Livre)	70
Tabela 9 – Tomada de Decisão e Evidências na análise da Trilha CQA (Atividade Livre).	71
Tabela 10 – Similaridade Comportamental entre os Participantes na Análise das Trilhas CQA (Atividade Livre)	72
Tabela 11 – Sumarização do Comportamento dos Participantes na Análise das Trilhas CQA (Atividade Livre).	73
Tabela 12 – Comportamento dos Participantes para Construção da Solução (Atividade Livre)	74
Tabela 13 – Similaridade Comportamental entre os Participantes no Processo de Busca (Condicionada Direta).	80
Tabela 14 – Sumarização do Comportamento dos Participantes no Processo de Busca (Condicionada Direta).	80
Tabela 15 – Estratégia de leitura da trilha CQA (Atividade Condicionada Direta). .	83
Tabela 16 – Elementos da trilha CQA de Foco de atenção pelo Participante (Atividade Condicionada Direta)	84
Tabela 17 – Similaridade Comportamental entre os Participantes no Processo analítico na CQA (Condicionada Direta).	86
Tabela 18 – Sumarização do Comportamento dos Participantes na Análise das Trilhas CQA (Atividade Condicionada Direta).	87
Tabela 19 – Comportamento dos Participantes para Construção da Solução (Atividade Condicionada direta).	87
Tabela 20 – Análise Estatística Descritiva da Duração da Sessão #3 (Condicionada Indireta Combinada).	89

Tabela 21 – Similaridade Comportamental entre os Participantes no Processo de Busca (Atividade Condicionada Múltipla).	96
Tabela 22 – Sumarização do Comportamento dos Participantes no Processo de Busca (Atividade Condicionada Múltipla).	96
Tabela 23 – Estratégia de leitura da trilha CQA (Sessão #3)	99
Tabela 24 – Similaridade comportamental entre os participantes na análise das trilhas CQA (Atividade Condicionada Múltipla).	101
Tabela 25 – Sumarização do comportamento dos participantes na análise das trilhas CQA (Atividade Condicionada Múltipla).	102
Tabela 26 – Comportamento cognitivo para construção da solução — Atividade Condicionada direta.	102
Tabela 27 – Comparação de uso de CQA e seus principais aspectos na percepção dos participantes.	105
Tabela 28 – Magnitude de validação e popularidade das perguntas — Entre 2014 e 2021.	109
Tabela 29 – Considerações dos participantes sobre uso das trilhas para aprendizagem	114
Tabela 30 – Considerações dos participantes sobre o uso e limitações da CSE	116
Tabela 31 – Aspectos da CQA que influem na aprendizagem segundo percepção dos participantes	118
Tabela 32 – Consultas formuladas pelos participantes (Atividade Livre)	159
Tabela 33 – Consultas formuladas pelos participantes (Atividade Condicionada Direta)	160
Tabela 34 – Consultas formuladas pelos participantes (Atividade Condicionada Múltiplas)	161
Tabela 35 – Listagem estruturada dos códigos	167

LISTA DE ABREVIATURAS E SIGLAS

ASK	<i>Anomalous states of knowledge</i>
BPA	Busca como um Processo de Aprendizagem
CLIR	<i>Cross-language Information Retrieval</i>
CLQR	<i>Cross-language question retrieval</i>
CQA	<i>Community Question and Answering</i> — Comunidade de Perguntas e Respostas.
CSE	<i>Google Custom Search Engine</i> ou <i>Programmable Google Custom Search</i>
CTAP	<i>Concurrent Think-Aloud Protocol</i>
ESKiP	<i>Exploratory Search Knowledge-intensive Processes</i>
FLB	<i>Formal Learning Behaviors</i> . Comportamentos Formais de Aprendizagem.
IA	Inteligência Artificial
ILB	<i>Informal Learning Behaviors</i> . Comportamentos informais de Aprendizagem.
IIR	<i>Interaction Information Retrieval</i> , Recuperação de Informação Interativa
KiP	<i>Knowledge-intensive Processes</i>
MP3	Formato de arquivo no padrão MPEG-1/2 Audio Layer 3 de compressão de áudio.
MP4	Formato de arquivo no padrão MPEG-4 Part 14 de container de áudio e vídeo.
OIES	<i>Online information evaluative standards</i>
OISS	<i>Online information searching strategies</i>
Q&A	<i>Question and Answers</i>
QDA	<i>Qualitative Data Analysis</i>
RI	Recuperação da Informação
RTAP	<i>Retrospective Think-Aloud Protocol</i>
SAL	<i>Search as Learning</i>
SDT	Teoria da Autodeterminação , <i>Self-determination Theory</i>
SERP	<i>Search Engine Result Page</i> , página de resultados do mecanismo de busca.
SRL	<i>Self-Regulated Learning</i> — aprendizagem informal e auto regulada.
TAP	<i>Thin-Aloud Protocol</i> ou <i>Thinking Aloud Protocol</i>
TFD	Teoria Fundamentada em Dados ou <i>Grounded Theory</i>
WAV	Formato de arquivo de áudio.
WCAG	<i>Web Content Accessibility Guidelines</i>
UX	<i>User Experience</i>

Sumário

1	Introdução	1
1.1	Contextualização	1
1.2	Motivação	2
1.3	Objetivos	6
1.4	Proposta da pesquisa	6
1.5	Resultados alcançados	8
1.6	Organização da pesquisa	9
2	Fundamentação e Trabalhos relacionados	10
2.1	Aprendizagem Informal	11
2.2	Aprendizagem em Comunidades de Q&A	13
2.3	Aprendizagem através do processo de busca na Web	16
2.4	Integrando a busca com CQA para prover Aprendizagem.	21
2.5	Trabalhos relacionados	24
2.6	Conclusão da fundamentação	28
3	Abordagem epistemológica-metodológica de pesquisa	29
3.1	Abordagem Metodológica	30
3.1.1	Teoria fundamentada em dados (TFD)	30
3.1.2	Intensive Interviewing	32
3.1.3	Thinking Aloud Protocol — TAP	33
3.2	Procedimentos metodológicos do estudo	34
3.2.1	Ambiente de busca customizada na Web sobre CQA	35
3.2.2	Entrevistas semiestruturadas	38
3.2.3	Execução do Experimento	39
3.2.4	Caracterização dos Participantes.	40
3.3	Coleta e Tratamento de Dados.	42
3.4	Codificação	44
3.5	Limitações e Ameaças à Validade	46
3.6	Considerações finais sobre a abordagem metodológica	50
4	Codificação e Análise do Comportamento do Usuário no Processo de Busca Web sobre CQA	52
4.1	Contextualização da utilização de CQA no processo de busca segundo a experiência do participante	54
4.2	Processo de Busca na Web com CQA para Atividade Livre (de autointeresse)	60
4.2.1	Análise do Comportamento no Processo de Busca	60

4.2.2	Análise do Comportamento no Processo analítico da trilha CQA .	67
4.2.3	Análise do Processo Cognitivo de construção da solução	73
4.3	Processo de Busca na Web com CQA para Atividade Condicionada Direta	74
4.3.1	Análise do Comportamento no Processo de Busca na Web	75
4.3.2	Análise do Comportamento no Processo analítico da trilha CQA .	80
4.3.3	Análise do Processo Cognitivo de construção da solução	87
4.4	Processo de Busca na Web com CQA para Atividade Condicionada Múltipla	88
4.4.1	Análise do Comportamento no Processo de Busca	89
4.4.2	Análise do Comportamento no Processo analítico da trilha CQA .	96
4.4.3	Análise do Processo Cognitivo de construção da solução	102
4.5	Análise Comparativa entre CQA a partir de Processo de Busca na Web . .	103
4.6	Análise da Percepção dos Participantes sobre a relevância de trilha CQA para suporte à Aprendizagem	108
4.6.1	Análise de Trilha conceitual ('O QUE') para suporte à Aprendizagem	109
4.6.2	Análise de Trilha de implementação ('COMO') para suporte à Aprendizagem	112
4.6.3	Comparação entre duas trilhas conceituais e de implementação para suporte à aprendizagem	113
4.7	Avaliação de percepção de aprendizagem pelos participantes durante as atividades realizadas	115
4.8	Considerações finais sobre codificação e análise	120
5	Discussão e Análise dos Achados	122
5.1	Sobre a Relevância da CQA no provimento das informações necessárias aos usuários	123
5.2	Sobre existência de conteúdo na CQA que promova a aprendizagem do usuário	125
5.3	Características sócio-técnicas da CQA que suportam ou inibem o seu uso para aprendizagem	128
5.4	Integração da busca como processo de aprendizagem no contexto de CQA	131
5.4.1	CQA como recurso de aprendizagem	131
5.4.2	Comportamento na busca como processo de aprendizagem sobre CQA	134
6	Conclusão	137
6.1	Observações finais	137
6.2	Trabalhos futuros	142

REFERÊNCIAS	144
APÊNDICE A APÊNDICE A - Ficha de Proposta da Pesquisa apresentada no Convite aos participantes voluntários	155
APÊNDICE B APÊNDICE B - Consultas formuladas pelos participantes nas sessões	159
B.1 Consultas realizadas pelos participantes na Sessão #1 Livre	159
B.2 Consultas realizadas pelos participantes na Sessão #2 Condicionada Direta	160
B.3 Consultas realizadas pelos participantes na Sessão #3 Condicionada Múltipla	161
APÊNDICE C APÊNDICE C - Listagem estruturada e agrupada de códigos	167
APÊNDICE D APÊNDICE D - Códigos-fonte em Python para processamento dos vídeos de participação	175
D.1 Conversão de Arquivos MP4 > MP3 > WAV	175
D.2 Extrair Texto de Audio e <i>Chunk</i>	176
D.3 Transformar <i>Chunks</i> com Texto Traduzido automaticamente para Arquivo Legendado	179
D.4 Transformar Arquivo Legendado em Arquivo de Transcrição QDA	183

1. Introdução

1.1 Contextualização

Com o desenvolvimento da Web 2.0, que ampliou o alcance de nossas interações para além dos limites geográficos de nossas comunidades, a popularidade de sistemas baseados em conteúdo gerado pelo usuário, como as Comunidades de Perguntas e Respostas (CQA), tem aumentado continuamente (SRBA; BIELIKOVA, 2016).

As CQA são comunidades sociais que servem para a troca de conhecimentos especializados entre seus membros. O propósito é o compartilhamento de conhecimentos específicos que permitam a todos os indivíduos algum aprendizado informal sobre o domínio e assuntos abordados. Para isso é fundamental a existência de usuários dedicados para manter a qualidade das postagens e o ritmo de crescimento sustentável da própria comunidade (PROCACI et al., 2015).

Para se manter atraente, uma CQA deve manter a qualidade do conteúdo refletido pelas postagens, que dependem tanto do domínio da língua escrita quanto dos assuntos abordados pelos usuários ativos. Para isso, CQA contam com mecanismos de reputação sociotécnica. Essa reputação está associada à participação na comunidade, incluindo interações sociais, comportamentos de autoridade e consenso entre os usuários, além da relevância do tema abordado e da contribuição realizada pelo usuário. A reputação sociotécnica decorre do reconhecimento do colaborador como especialista em determinado assunto ou trilha de discussão. A expressão do consenso pode ser por meio da aceitação, compartilhamento ou votação das considerações expressas pelo usuário especialista nas trilhas (GIMENEZ; SIQUEIRA, 2021).

Tais características podem ser descritas pelos princípios de influência Autoridade e Consenso (CIALDINI; CIALDINI, 1993), que fazem parte da Teoria da Comunicação Persuasiva. Enquanto a Autoridade sugere que as pessoas querem seguir ou se orientar

por meio de autoridades em determinadas áreas, o Consenso sugere verificar e confiar no que os pares estão fazendo, quando há incertezas sobre informações e práticas. Quando se trata de disseminação de informações, o fluxo de conhecimentos e a moldagem de opinião estão atreladas ao papel de influência de determinados usuários em relação aos demais integrantes da rede social (AL-GARADI et al., 2018) percebidos pelos demais através dos mecanismos sociotécnicos.

Estudos na área de informática na educação tentam identificar quais discussões (PROCACI et al., 2019) tem uma profundidade de acordo com um limite mínimo (PROCACI et al., 2017), o assunto associado e as conexões com outros assuntos para uma pré-classificação em categorias (PROCACI et al., 2016), de forma que possam ajudar estudantes a encontrarem informações de melhor qualidade e deixá-las em destaque no ambiente (PROCACI; NUNES; SIQUEIRA, 2021)¹. Assim, as CQA representam uma boa fonte de informações mais específicas sobre um assunto. Por isso, as CQA tornaram-se a base digital de conhecimentos na perspectiva de aprendizagem, que envolve os esforços para criar sistemas de aprendizagem em vários setores e em vários níveis de escala, de comunidades locais a organizações únicas, parcerias, cidades, regiões e o mundo inteiro (WENGER, 2009). Isto ao tratar a busca de conhecimentos especializados associados à aprendizagem.

Na perspectiva de aprendizagem colaborativa, a CQA pode ser reconhecida como um caso especial de aprendizado por pares, quando os usuários não têm a participação de um instrutor ou professor, e assim, agem como alunos adquirindo novos conhecimentos uns dos outros. Na CQA, o processo de resposta a perguntas provoca um fluxo de conhecimento dos usuários mais experientes para os menos experientes. Os novos conhecimentos podem ser adquiridos através da leitura, e não apenas pela interação direta na postagem (perguntando e respondendo), bem como pela elaboração de novas soluções — novos conhecimentos práticos, por meio de discussões anexadas à trilha. Tal cenário descreve o potencial de aprendizagem que pode ser caracterizado como aprendizado informal (SRBA; BIELIKOVA, 2016).

1.2 Motivação

Nas plataformas de CQA alcançadas através de mecanismos de busca, os usuários conseguem elaborar soluções para problemas resolvidos por discussões anexadas a perguntas ou respostas (ARITAJATI; NARAYANAN, 2013). O estabelecimento de uma abordagem

¹ <https://ieducacao.ceie-br.org/ambientes-pessoais-aprendizagem/>

de construção do conhecimento nessas comunidades foi motivado pela tendência atual da era do conhecimento, onde o avanço sustentado do conhecimento é visto como essencial para o progresso social (SCARDAMALIA; BEREITER, 2006; SRBA; BIELIKOVA, 2016). Isto, considerando a característica principal que as CQA visam fornecer as respostas mais adequadas às perguntas postadas recentemente, no menor tempo possível, conseguindo aproveitar o conhecimento tácito (incorporado em suas diversas comunidades) ou o conhecimento explícito (incorporado em todas as questões resolvidas) para responder a um grande número de novas questões postadas a cada dia (SRBA; BIELIKOVA, 2016).

Para os usuários da Internet, que não estão ativos ou não são membros de CQA, essas comunidades servem como recurso de conhecimento especializado, com aplicação imediata e pontual segundo a necessidade de informação de cada usuário em particular. Normalmente são acionadas a partir de mecanismos de buscas da Internet, tais como *Google*² ou *Bing*³ devido a natureza da busca em formato de pergunta (KUMAR; DANDAPAT; CHORDIA, 2020), que aproxima as buscas das postagens em meio a demais fontes como resultados da busca.

Como um caso especial, a pesquisa na Web realizada pelos desenvolvedores de sistemas de informação são realizadas através de mecanismos de busca, principalmente o *Google Search*, e geralmente, envolvem como fonte a *CQA Stack Overflow*⁴, que possui mais 356 milhões de visitas únicas por mês⁵ com mais de 17 milhões usuários⁶ e mais de 130 milhões de postagens⁷. Esse tipo de pesquisa na Web envolve um processo complexo de buscas de informações, exigindo habilidade metacognitiva e influi na eficiência da atuação do profissional, que precisa navegar, filtrar e avaliar muitos resultados da busca sobre o conhecimento técnico necessário (LI et al., 2013).

A *CQA Stack Overflow*, que usa a mesma plataforma do *Stack Exchange*, realizou em 2020 um levantamento por questionário com 65 mil desenvolvedores de sistemas de informação de representação demográfica diversificada (STACKOVERFLOW, 2020). Verificaram que 75% dos entrevistados aprendem uma nova tecnologia em menos de um ano, demonstrando que estão em constante aprendizado para manter suas habilidades atualizadas. Através do uso de mecanismos de busca, quase 90% relataram que visitam o *Stack Overflow* para encontrar solução quando o ficam presos em um problema, contra 52,8% que assistem a vídeos de ajuda ou tutoriais, 49,9% que ligam para um colega de

² <https://www.google.com.br/>

³ <https://www.bing.com/>

⁴ <https://stackoverflow.com/>

⁵ <https://mysite.info/analysis/stackoverflow.com/>

⁶ <https://stackexchange.com/sites>

⁷ <https://data.stackexchange.com/stackoverflow/queries>

trabalho ou amigo e 10,3% que visitam outra comunidade de desenvolvedores. Tais valores expressam a credibilidade da CQA *Stack Overflow* junto a comunidade de prática de desenvolvimento de sistemas, a confiança em suas informações e seu papel crescente de Autoridade no assunto em relação à solução de problemas de desenvolvimento.

Segundo [Chen e Xing \(2016\)](#), as pesquisas no *Google Search* e as perguntas no *Stack Overflow*, principal CQA especializada, possuem forte correlação em função dos termos pesquisados e solicitados. Ambas são as principais formas de busca de informação na Web e ponto de partida para a aprendizagem utilizadas pelos desenvolvedores. Os mecanismos de pesquisa consideram os recursos de programação online apenas como recursos de documentos da Web, sem representação e organização de conhecimento eficaz para ajudar os desenvolvedores a recuperar, agregar e explorá-los em torno do conhecimento de que precisam ([CHEN; XING, 2016](#)). Já CQA especializadas, como a *Stack Overflow*, possuem conhecimento *crowdsourced*⁸ vasto e de alta qualidade, que poderiam ser explorados, dada a correlação apontada ([CHEN; XING, 2016](#)), para prover o conhecimento buscado a partir da pesquisa exploratória dos desenvolvedores. Isto, principalmente quando o desenvolvedor não sabe quais palavras-chave usar ou quando não está procurando uma única resposta.

No contexto da busca como processo de aprendizagem (BPA, inglês: SAL — '*Searching as learning*'), segundo [Liu, Liu e Belkin \(2019\)](#), os sistemas (ou mecanismos) de busca são considerados como tecnologias que podem apoiar a aprendizagem, em vez de meras ferramentas de recuperação de informação, e permitem uma compreensão do comportamento de busca de informações dos usuários no contexto mais amplo da aprendizagem humana. Essa interpretação a partir da perspectiva de aprendizagem não é novo (por exemplo, o modelo ASK de [Belkin \(1980\)](#) já trazia alguns elementos desta perspectiva), mas como coletar a mudança de conhecimento ou aprendizado dos usuários explicitamente durante o processo de busca ainda é uma dificuldade e um dos principais desafios dessa nova área BPA ([LIU; LIU; BELKIN, 2019](#)).

Desta forma, a busca como processo de aprendizagem ([VAKKARI, 2016; RIEH et al., 2016; HOPPE et al., 2018](#)), que ressignifica os mecanismos de busca na perspectiva de aprendizagem, tratando a necessidade de conhecimento a partir buscas complexas de caráter exploratório, poderia se beneficiar da relevância conferida à CQA para prover conhecimentos ([STACKOVERFLOW, 2020; CHEN; XING, 2016](#)) técnicos. Este poderia ser o caso de buscas normalmente realizadas pelos desenvolvedores de sistemas de informação ([CHEN; XING, 2016](#)).

⁸ *Crowdsourced* é a coleta de informações, opiniões ou trabalho de um grupo de pessoas, geralmente obtido através da Internet.

O entendimento do conceito de ‘relevância’ da CQA no presente trabalho segue a vertente focada nos usuários (HAIDER; SUNDIN, 2019), conforme a percepção do usuário, pela forma como é vivenciado, como uma relação entre a informação e o usuário. Isto difere da vertente de sistema, que considera uma propriedade da relação entre o conteúdo do documento e os critérios de busca. Desta forma, utilizamos o conceito de relevância com caráter dinâmico e relativo, profundamente situado e dependente do julgamento humano, isto é, depende das pessoas considerarem algo significativo e útil em situações específicas (HAIDER; SUNDIN, 2019).

Quanto aos aspectos motivacionais pertinentes a este pesquisador, explicáveis a partir da Teoria da Autodeterminação (DECI; RYAN, 2012), destaco as minhas trajetórias profissionais no uso de CQA e acadêmica no estudo de CQA e de BPA. Tal teoria percebe o comportamento humano como orientado para o crescimento e proativo, motivado a assimilar e integrar conhecimentos e capacidades em seus ambientes físicos e sociais. Pela minha trajetória profissional, destaco a utilização rotineira, contínua e corrente de plataforma CQA, principalmente a *Stack Overflow*, através dos mecanismos de buscas na Web para angariar novos conhecimentos e aprender novas tecnologias, ao longo de mais de 20 anos atuando no desenvolvimento de sistemas de informação. Pela minha trajetória acadêmica, que culmina no presente estudo dissertativo, comecei com o estudo de CQA através de pesquisa quantitativa sobre análises das interações dos usuários e suas postagens a partir dos *logs* de comunidades especializadas da *Stack Exchange* que resultou no artigo Gimenez e Siqueira (2021). Em outra frente de estudo, explorei as perspectivas de uso da busca na Web num cenário de aprendizagem, em uma nova ‘subárea’ de pesquisa definida como BPA, resultando nas contribuições (MACHADO; GIMENEZ; SIQUEIRA, 2020) e Gimenez et al. (2020), pelo levantamento sistemático da literatura vigente.

Aliamos a trajetória profissional do pesquisador e os estudos anteriores, considerando as buscas complexas de caráter exploratório, normalmente realizadas pelos desenvolvedores de sistemas de informação (LI et al., 2013; CHEN; XING, 2016), bem como a relevância do uso da CQA como a fonte de informação que atende tal necessidade de informação (STACKOVERFLOW, 2020; CHEN; XING, 2016), segundo a perspectiva de construção de conhecimentos de que precisam. Nesse sentido, investigamos os riscos associados e o potencial de aprendizagem existentes em CQA (ARITAJATI; NARAYANAN, 2013) associado à busca como processo de aprendizagem. Neste contexto, o uso da CQA precisa ser ampliado para além de simples fonte de informações necessárias aos usuários e base de conhecimentos compartilhados, demandando a avaliação da relevância e a adequação da própria comunidade e seus conteúdos para apoiar a aprendizagem do usuário.

1.3 Objetivos

Para avaliar a CQA no escopo de BPA, realizamos as análises sobre o seu uso, compartilhamento de conhecimento, relevância e suporte a aprendizagem do usuário. Adotamos para isso, o controle do fluxo de informações para os usuários ao restringir os resultados às principais CQA levantadas da literatura e Internet, para permitir maior foco no objeto de análise.

No presente estudo, nosso objetivo de pesquisa é **entender os comportamentos e decisões do participante, enquanto usuário-aprendiz, ao realizar buscas na Web e a análise das trilhas de CQA resultantes da busca na perspectiva de aprendizagem informal**. Dentro dessa perspectiva, partimos de quatro questões de pesquisa propostas para subsidiar a análise: (RQ1) Qual a utilidade da CQA no provimento de informações relevantes e necessárias aos usuários para aprendizagem?; (RQ2) Como a organização dos conteúdos compartilhados na CQA ajuda a promover aprendizagem dos usuários-aprendiz?; (RQ3) Quais são as características sócio-técnicas que uma CQA deve ter para dar suporte à aprendizagem?; (RQ4) Como integrar o uso de CQA com a busca como processo de aprendizagem na perspectiva de Aprendizagem Informal?.

1.4 Proposta da pesquisa

Consideramos neste estudo a interação individual de usuários através do sistema de busca na Internet mais utilizado — *Google Search*. Tal sistema foi customizado para considerar somente as CQA enquanto fontes para os resultados apresentados, bem como os requisitos mínimos funcionais para realização da tarefa de busca, de modo a termos um fluxo controlado de informações da busca na Web à análise do conteúdo da trilha da CQA. No processo de transformação de informação buscada em conhecimento aplicável, considerando a experiência e a possível aprendizagem imediata do usuário (GIMENEZ et al., 2020), foram propostas sessões apresentando problemas para serem resolvidos. Os usuários participantes tinham que resolvê-los no contexto de CQA na perspectiva de BPA. Isto, através da formulação de consultas (busca Web), seleção dos resultados a partir de CQA, navegação e análise das trilhas e postagens nas CQA e formulação da solução demandada. As sessões foram elaboradas para ampliar gradualmente o grau de dificuldade das tarefas, partindo daquela de escolha do participante que retratasse a naturalidade de seu comportamento, passando por uma característica de uso técnico da CQA com conceito envolvido, por outra mais complexa de múltiplos conceitos e atividades, até chegar a umas mais conceitual e analítica da CQA e do conteúdo para aprendizagem. Foram re-

alizadas ainda duas sessões de entrevistas, uma inicial sobre a percepção e uso de CQA pelos participantes, e outra final sobre *feedback* do estudo e perspectivas de aprendizagem com CQA.

A abordagem metodológica qualitativa utilizada seguiu o paradigma interpretativa/construtivista, dada a natureza epistemológica do fenômeno de interesse. As questões dessa natureza envolvem teorias associadas e uma compreensão do mesmo a partir significados atribuídos pelos indivíduos (SALDAÑA, 2013). A metodologia utilizada foi a Teoria Fundamentada em Dados na perspectiva Construtivista, que entende o papel ativo do pesquisador na compreensão específica do fenômeno e prevê a compilação da literatura após a análise dos dados para fortalecer o potencial argumentativo da pesquisa. O processo de aplicação da abordagem metodológica envolveu etapas de preparação do estudo com arranjo da aplicação e ambiente, execução das entrevistas e sessões observacionais, coleta de dados, e também, análise de dados com o uso do método de comparação constante.

O protocolo definido para execução das sessões dos usuários foi o *Thinking Aloud Protocol* (TAP), proposto por Ericsson e Simon (1980). Trata-se de um método de pensar em voz alta, onde o usuário declara de forma audível toda a sua interação no ambiente montado para a prática. Em virtude da possibilidade de dificuldade de indivíduos poderem articular seus pensamentos simultaneamente e concluir qualquer tarefa de Recuperação de Informação Interativa (IIR) (ZHANG; SOERGEL, 2016; LIU; LIU; BELKIN, 2019), coube ao pesquisador fazer a instigação do participante para verbalizar oralmente todas as suas ações e motivações, quando necessário. Isto foi possível, devido ao uso combinado da técnica *Intensive Interviewing* com a abordagem TAP, tanto na vertente concorrente quanto na vertente de retrospectiva. Tal técnica também foi utilizada nas sessões de entrevistas semiestruturadas anterior e posterior às sessões de atividades. As intervenções do pesquisador acontecem quando há ausência comportamental, falta de esclarecimentos sobre falas ou sobre as ações durante a interação. A IIR é moldada por (a) pesquisa sobre busca de informações e comportamento de busca e (b) pesquisa sobre o desenvolvimento de novos métodos de interação com recursos eletrônicos (RUTHVEN, 2008), ambas com foco na assistência dos usuários de sistemas de recuperação de informação.

O ambiente de execução do estudo foi a realização de reunião online na plataforma *Google Meet* apoiada pelo *chat* interno. Ao longo da interação, as orientações iniciais das tarefas são prestadas para cada sessão, e esclarecimentos adicionais são fornecidos quando demandados pelo participante, de tal forma que os participantes se sintam à vontade para realizar as atividades e dizer em voz alta tudo o que naturalmente vem à mente. Ressaltamos que a realização ocorreu em um cenário de agravamento da pandemia da

Covid-19 no Brasil no início de 2021, que foi um obstáculo significativo para articulação de mais indivíduos para participação. Contudo, conseguimos 17 participantes voluntários qualificados em relação à proposta do estudo, por serem todos pesquisadores em sistemas de informação, que já atuaram em tarefas complexas de busca por informação relativas às atividades de desenvolvimento em algum momento, com o uso de mecanismos de busca e fontes de informação online. Tal esforço de participação consumiu 26 horas 47 minutos e 45 segundos de duração de conteúdo de áudio e vídeo, ao longo de mais de 2 meses de agenda de entrevistas com algumas remarcações. Tal empreendimento resultou em 4.317 segmentos de texto codificados com 290 códigos abertos agrupados em 10 categorias conforme Seção 35, a partir de 102 documentos oriundos dos relatos e respostas das entrevistas que somaram mais de 6 mil segmentos de transcrição.

1.5 Resultados alcançados

A principal contribuição deste trabalho é uma análise qualitativa da aprendizagem de usuários a partir da busca Web como processo de aprendizagem utilizando CQA. Tal análise difere de outros estudos levantados na literatura corrente pela abordagem metodológica Teoria Fundamentada em Dados Construtivista empregada e ao utilizar dados coletados de experiência e percepção dos usuários diretamente de suas interações através do método TAP e da técnica *Intensive Interviewing*, não se restringido à experiência nem de análise na CQA e nem de busca na Web, mas combinando-as na perspectiva de aprendizagem.

Podemos destacar ainda como contribuições científicas adicionais: (i) extensão da taxonomia ESKiP(TIBAU; SIQUEIRA; NUNES, 2020; TIBAU et al., 2019) de estados de consulta, (ii) estruturação em uma taxonomia multifacetada do modelo de busca exploratória ESKiP descrito por Tibau et al. (2018), (iii) listagem de códigos abertos estruturados e agrupados em categorias temáticas (Apêndice C), e (iv) registro da percepção e reflexão do participante a respeito do seu próprio processo de aprendizagem informal durante as sessões.

Como contribuições técnicas: (i) conjunto de questões guia das duas entrevistas (Apêndice A), (ii) ambiente customizado (CSE) usado nas sessões práticas do estudo, e (iii) códigos-fontes comentados de rotinas em *Python* para transformar os relatos de áudio e vídeo em textos processáveis pela maioria das ferramentas de QDA (Apêndice D).

Como contribuições teóricas: (i) expansão da aplicabilidade do Princípio de Influência de Autoridade (CIALDINI; CIALDINI, 1993) para determinadas CQA por suas caracte-

rísticas, (ii) restrição da aplicabilidade dos Princípios de Influência de Consenso (CIALDINI; CIALDINI, 1993) para o processo de análise de conteúdo das trilhas da CQA, e (iii) constatação que a CQA pode ser compreendida através da Teoria de Comunidades de Práticas, associada ao papel de Autoridade, e possivelmente decorrente das Teorias da Inteligência Coletiva e Sabedoria da Multidão (SRBA; BIELIKOVA, 2016).

Estes achados serão apresentados detalhadamente ao longo da dissertação, principalmente no Capítulo 5.

1.6 Organização da pesquisa

O presente trabalho está estruturado da seguinte forma: neste Capítulo 1 apresentamos a contextualização, motivação e os principais aspectos e contribuições do estudo. No Capítulo 2, apresentaremos a fundamentação sobre os tópicos relacionados ao fenômeno de interesse nas seções: 2.1 com aprendizagem informal e autorregulada (SRL) do usuário aprendiz, contextualizando e destacando os elementos envolvidos; 2.2 com caracterização e uso de CQA para aprendizagem; 2.3 com o desenvolvimento e perspectiva de aprendizagem através de SAL; e 2.4 com a integração de BPA com CQA para promover a aprendizagem. Também apresentaremos os principais trabalhos relacionados ao presente estudo e o enfoque recente da literatura na Seção 2.5, com as considerações finais da fundamentação na Seção 2.6. No Capítulo 3 apresentaremos na seção 3.1 a abordagem metodológica geral da presente pesquisa com a descrição da abordagem qualitativa com os métodos, técnicas e protocolos selecionados; as seções 3.2 com os procedimentos metodológicos do estudo, 3.2.4 com a caracterização dos participantes; 3.3 com coleta e tratamento dos dados qualitativos, 3.4 com a abordagem de codificação; a seção de 3.5 sobre limitações e ameaças à validade do presente estudo com os tratamentos dispensados; e finalmente, a Seção 3.6 com as considerações finais sobre aspectos metodológicos do estudo. No Capítulo 4 apresentaremos a elaboração da codificação inicial e focalizada a partir dos relatos transcritos dos participantes nas sessões de entrevista e de atividades, considerando os comportamentos no processo de busca e no processo analítico da trilha CQA, bem como aspectos cognitivos explícitos da construção da solução ao final das atividades, comparação entre as CQA e percepção quanto ao uso para aprendizagem. No Capítulo 5 discutimos os principais achados e indícios frente às questões de pesquisa. Finalmente, no Capítulo 6 apresentaremos as observações finais e os trabalhos futuros.

2. Fundamentação e Trabalhos relacionados

As plataformas CQA tornaram-se uma alternativa viável para a obtenção de respostas diretas e objetivas às nossas necessidades de informação, ao permitir perguntar a outros membros de uma comunidade. Essas respostas podem ser difíceis de serem encontradas diretamente nos documentos da web, além de requererem a síntese de fatos, experiências e conhecimento de mundo dos membros da comunidade (FIGUEROA; NEUMANN, 2014).

Assim, os serviços de CQA tornaram-se parte integrante da aquisição de informações e conhecimento na Web 2.0, acumulando ao longo do tempo um grande conjunto de perguntas e respostas (Q&A) com alta qualidade dedicados para inteligência humana como uma base de conhecimento ampla (ZHANG et al., 2014). Também promovem o aprendizado informal ao estimular a interação sobre a observação, expressa através dos comentários e votação de todos os elementos da trilha. Serviços de CQA especializadas, como *Brainly* e *Stack Overflow*, visam respectivamente ajudar os alunos e os profissionais programadores, a melhorar seu processo de aprendizado por meio de trocas de perguntas e respostas (LE; SHAH; CHOI, 2019).

Segundo Scardamalia e Bereiter (2006), a Web torna-se o primeiro meio realista para os alunos se conectarem com a construção de conhecimento na história da civilização, permitindo que a sala de aula funcione à parte dela. Podemos assim, agregar ao uso da CQA para aprendizagem, a forma natural de busca de informações na Web, com o uso dos mecanismos de busca. Desta forma, expandimos o ambiente de aprendizagem desde a busca na Web até a análise das trilhas da CQA, na perspectiva de busca como processo de aprendizagem tendo como uma fonte relevante a CQA.

2.1 Aprendizagem Informal

Nas últimas duas décadas, a literatura vem mostrando que entre 70% e 90% do aprendizado organizacional ocorre não por treinamento formal, mas informalmente, no trabalho e de maneira contínua acessando conteúdos online (CERASOLI et al., 2018). Funcionários precisam adquirir novos conhecimentos e habilidades de forma contínua por meios informais e não curriculares, e a maior parte do aprendizado ocorre por meio da experiência de trabalho (TANNENBAUM et al., 2010 apud CERASOLI et al., 2018). Conforme levantamento de StackOverflow (2020), percebemos esse fluxo de aprendizagem para adquirir conhecimentos que reforcem as habilidades profissionais. A aprendizagem informal é uma forma de aprendizagem realizada ao longo da vida (*lifelong learning*), que implica que provavelmente a aprendizagem ocorra fora e para além da escola tradicional (CARNEIRO; LEFRERE; STEFFENS, 2007).

Para diferenciar a Aprendizagem Informal da Aprendizagem Formal, Cerasoli et al. (2018) utilizaram uma abordagem baseada em comportamentos de aprendizagem. Na Aprendizagem Formal, conhecimentos e habilidades são adquiridos através de comportamentos formais de aprendizagem (FLB), como assistir a uma aula, por exemplo. Já na Aprendizagem Informal, eles são adquiridos por comportamentos informais de aprendizagem (ILB), como observação, práticas, discutir um problema, ou fazer perguntas durante uma atividade. Segundo Tannenbaum et al. (2010), os ILB podem ser definidos como: (i) não curriculares, (ii) envolvem atividades desenvolvidas a serviço da aquisição de conhecimentos e habilidades, (iii) ocorrem fora dos contextos de aprendizagem formalmente designados, (iv) são predominantemente autodirigidas e intencionais. Convém aqui diferenciar a Aprendizagem Informal da Aprendizagem Incidental (ou Implícita), pela existência da intenção de aprender ou melhorar, e do direcionamento consciente do processo de aprendizagem pelo próprio indivíduo (TANNENBAUM et al., 2010).

Antecedentes pessoais e situacionais, assim como os resultados alcançados (experiências anteriores), podem ser considerados fatores de predição de comportamentos informais de aprendizagem (CERASOLI et al., 2018). Tais como a predisposição individual (consciência, curiosidade, adaptabilidade, etc.) e a necessidade percebida de aprendizagem informal, associadas às questões situacionais, como as características e demandas de trabalho, oportunidades de novos conhecimentos e *feedbacks* positivos organizacionais (trabalho ou academia).

Em seu trabalho, Boekaerts e Minnaert (1999) identificaram dez atributos abrangentes da aprendizagem informal:

- O processo de aprendizagem é descrito como ativo, voluntário, autodescoberta, autodeterminada, aberta, não ameaçadora, agradável e exploratória.
- Os alunos usam uma série de processos de autorregulação espontaneamente, como auto iniciação da aprendizagem e automonitoramento de seu progresso.
- Os processos autorregulatórios demandam a motivação intrínseca ¹; por outro lado, a motivação intrínseca facilita os processos de autorregulação.
- Frequentemente inserida em um contexto social, o que significa que as dicas sociais são altamente relevantes e que os alunos se envolvem em atividades de aprendizagem cooperativa.
- As situações informais de aprendizagem utilizam objetos, materiais ou configurações (realistas) altamente contextualizados.
- A experiência de aprendizagem é mais qualitativa do que quantitativa, mais orientada para o processo do que para o produto, mais sintética do que analítica e mais orientada para o fluxo.
- A alocação de tempo em episódios efetivos de aprendizagem informal é de natureza sem pressa, com ritmo próprio e aberto, com relativamente poucas restrições de tempo.
- É flexível, mesmo quando há um tipo de currículo associado, significando que a estrutura é não linear e de baixo para cima.
- Existe somente um tipo de avaliação ou autoavaliação coletiva e informal com base no *feedback*.
- Metas estabelecidas tendem a ser mais amplas, que pode levar a uma considerável variabilidade no que é aprendido.

Esses atributos da aprendizagem informal produzem juntos uma forma natural de aprendizagem para o aluno (BOEKAERTS; MINNAERT, 1999), dando a impressão que aprendem de forma espontânea e sem muito esforço consciente (NAVARRETE; CALVI, 2009). Dito isto, os paradigmas instrucionais estão mudando, o ensino e o treinamento estão se tornando mais orientados ao aluno e menos orientado ao instrutor, exigindo assim um maior grau de Aprendizagem Autorregulada (SRL) (CARNEIRO; LEFRERE; STEFFENS, 2007).

¹ Também conhecida como motivação interna, relacionada à força interior capaz de se manter ativa mesmo diante de adversidades, inerente aos interesses individuais e que podem ser alterados apenas por escolha da pessoa. Por exemplo: prazer, dor, a busca por crescimento pessoal, segurança e pertencimento social

Na SRL, os alunos raramente perseguem objetivos únicos e, de fato, geralmente tendem a perseguir múltiplos objetivos que fazem parte de sua complexa estrutura de objetivos (NAVARRETE; CALVI, 2009). Podemos compreender a SRL como pensamentos, sentimentos e ações autogerados, sendo sistematicamente orientados para a realização dos próprios objetivos dos alunos (ZIMMERMAN; SCHUNK, 2001 apud NAVARRETE; CALVI, 2009). Três aspectos-chaves são apontados (BOEKAERTS, 1999 apud NAVARRETE; CALVI, 2009): (i) cognitivo, sobre a capacidade dos alunos de selecionar, combinar e coordenar estratégias de maneira eficaz; (ii) metacognitivo, sobre a capacidade dos alunos de direcionar sua própria aprendizagem; (iii) motivacional/afetivo, sobre envolvimento e comprometimento dos alunos com os objetivos auto-escolhidos. A terminologia SRL tem sido usada de (i) forma ampla, o aluno pode escolher o que, quando e onde aprender; e com (ii) sentido estreito, em situações onde os alunos receberam uma tarefa de aprendizagem e autorregulam o processos de aprendizagem envolvidos na execução da tarefa (CARNEIRO; LEFRERE; STEFFENS, 2007).

A avaliação dos alunos sobre o ambiente de aprendizagem específico afeta o valor que eles atribuem aos objetivos de aprendizagem, bem como a qualidade do processo de aprendizagem. Assim, as habilidades de autorregulação são essenciais para a avaliação de ambientes de aprendizagem não tradicionais como poderosos facilitadores da aprendizagem e para o uso de recursos, como ferramentas digitais e a Internet (NAVARRETE; CALVI, 2009). Nesse aspecto, a partir do levantamento da literatura, Urbina, Villatoro e Salinas (2021) reforçam o uso de elementos ou ambientes tecnológicos como fatores que podem potencialmente favorecer a SRL, bem como a adoção da abordagem do modelo de (Zimmerman (2000)) com suas derivações, no que concebe os alunos como agentes ativos em todo o processo de ensino-aprendizagem.

2.2 Aprendizagem em Comunidades de Q&A

As comunidades online de Q&A (CQA) servem como plataforma online para aprendizagem informal ao promover o aprendizado de qualquer pessoa através da observação dos conhecimentos expressos pelas trilhas de discussão de perguntas e respostas. Os repositórios de conteúdo criados por meio de sites CQA podem ser utilizadas como bibliotecas digitais online que satisfaçam necessidades específicas relacionadas à educação (LE; SHAH; CHOI, 2019). As perguntas podem ser categorizadas em subjetiva (opinião), objetiva (conhecimento factual, normalmente do tipo 'WH': quem, quando, onde, o quê e por quê) e social (interação social). Podemos vislumbrar três entidades de domínio cruciais: pergunta, resposta e usuário, este último pode desempenhar dois papéis principais:

perguntador e respondente (SRBA; BIELIKOVA, 2016). Dependendo da plataforma e organização da informação na CQA, podemos ter a distinção da entidade comentário, que pode atuar sobre uma pergunta, uma resposta ou outro comentário. Consideramos postagem os elementos pergunta, resposta e comentário. Para cada postagem pode-se ter votação dos usuários. Uma resposta pode ser marcada como a melhor resposta, a despeito de ter mais votação favorável.

A observação dos conhecimentos, como os expressos pelas trilhas de discussão de perguntas e respostas, leva à construção mental de novos conhecimentos que serão aplicados, conforme a informação está exposta e influencia ao leitor, irrigando o fluxo dos processos cognitivos e mentais: atenção, retenção, reprodução e motivação, conforme nos apresenta a Teoria de Aprendizagem Social (BANDURA; WALTERS, 1977). As motivações, grau de atenção, necessidade de retenção e reprodução são externas ao contexto da plataforma online de Q&A, intrínsecas ao indivíduo e suas necessidades pessoais ou profissionais.

O estímulo à participação através de postagens nas CQA leva ao aprendizado, considerando que o processo de construção do conhecimento é colaborativo por natureza e disparado através do diálogo (p.ex. comentários e respostas) e discussão entre os pares nas trilhas de discussão (PENA-SHAFF; NICHOLLS, 2004). O desenvolvimento social e cognitivo ocorre primeiro no plano sociológico, através da interação entre pares e com os especialistas (esfera interpsicológica), e, somente então, o conteúdo social e o contexto são transformados e internalizados no plano psicológico (esfera intra-psicológica) (VYGOTSKY; HANFMANN, 1967 apud PENASHAFF; NICHOLLS, 2004).

As interações e o processo colaborativo de construção dos conhecimentos dentro da CQA possuem uma perspectiva mais social, dependente da qualidade da comunidade, por conseguinte, da correta identificação dos mais diversos perfis de usuários. O perfil que demanda especial interesse é o usuário que produz conhecimentos, que promove a aprendizagem informal a partir de suas postagens como especialista nos assuntos e engajado na melhoria da comunidade (GIMENEZ; SIQUEIRA, 2021).

A identificação dos fatores que afetam a motivação dos usuários para participar de atividades da comunidade, especialmente na contribuição dos conhecimentos, são relevantes para entender o papel da CQA como plataforma de aprendizagem informal. O estudo de Guan et al. (2018), revelou que o engajamento das pessoas é influenciado por: (i) confiança baseada em identidade, que é a confiança de nível mais alto derivada da compreensão, apreciação e apoio mútuos entre as pessoas; (ii) *feedback* da contribuição anterior do conhecimento e (iii) oportunidades de exposição social. A confiança baseada em identidade traz um fluxo constante de informação e oferece oportunidade de aprendi-

zagem de outros usuários, que podem acumular conhecimentos e habilidades recebendo respostas de outros e observando o comportamento de contribuição dos outros (GUAN et al., 2018). Algumas funcionalidades podem apoiar no engajamento, tais como o papel do usuário de destaque, votação e comentários sobre as contribuições realizadas. A distinção social do usuário de destaque em relação aos demais usuários da CQA pode ser descrita pelos Princípios de Influência (CIALDINI; CIALDINI, 1993), com destaque para os princípios da Autoridade e do Consenso. Enquanto o princípio da Autoridade sugere que as pessoas querem seguir ou se orientar através de autoridades de determinadas áreas, o princípio do Consenso sugere verificar e confiar no que os pares estão realizando ou atuando, quando existem incertezas quanto às informações e práticas.

Os usuários que se destacam em CQA, tem a autoridade porque possuem a reputação social e retêm o consenso expresso pela aceitação voluntária dos seus pares na comunidade. A reputação social provém do reconhecimento como especialista em determinado assunto ou trilha de discussão. A expressão do consenso pode ser através de aceitação, compartilhamento ou votação das considerações manifestadas do especialista nas trilhas. Assim, o destaque de um usuário está ligado à intensidade da demonstração de conhecimentos em determinado momento (GIMENEZ; SIQUEIRA, 2021).

A partir de levantamento da literatura, Srba e Bielikova (2016) apresentam a inserção da CQA na perspectiva de aprendizagem, além da perspectiva de compartilhamento de conhecimentos tácitos dos usuários, onde a CQA pode ser compreendida através da Teoria de Comunidades de Práticas. Dentro dessa perspectiva de aprendizagem, a possibilidade de buscar uma resposta para uma pergunta que se está fazendo é também uma forma específica de aprendizagem informal, aprimorada por tecnologia, mais especificamente a aprendizagem colaborativa apoiada por computador. Nesse contexto, ocorre um fluxo de conhecimento dos usuários mais experientes para os menos experientes, através da leitura, da elaboração da postagem (pergunta ou resposta). Também podemos reconhecer essa interação colaborativa na CQA através de duas Teorias interrelacionadas, a inteligência coletiva e a sabedoria da multidão. A inteligência coletiva é construída sobre uma ideia de que ninguém sabe tudo, mas todo mundo sabe alguma coisa, e assim podemos aproveitar a inteligência de toda a comunidade para resolver tarefas que dificilmente seriam possíveis de serem resolvidas por um indivíduo (SRBA; BIELIKOVA, 2016). O conceito de sabedoria da multidão pode ser descrito como “o processo de considerar a opinião coletiva de um grupo de indivíduos em vez de um único especialista para responder a uma pergunta” (SUROWIECKI, 2005 apud SRBA; BIELIKOVA, 2016)).

Embora Srba e Bielikova (2016) considerem o aprendizado em sistema CQA algo natural, por entender tal sistema como ambiente de aprendizagem colaborativa, destacam

a existência de alguns estudos preocupados com o potencial de aprendizagem dos sistemas CQA e riscos associados (ARITAJATI; NARAYANAN, 2013). Assim, segundo Le, Shah e Choi (2019), um fator crítico é a qualidade das respostas, devido ao risco de não receber respostas de alta qualidade para uma consulta, uma vez que os especialistas não respondam, ou os perguntadores não têm o conhecimento e as habilidades necessárias para avaliar a qualidade das respostas que recebem. Isso poderia fazer com que os alunos construíssem uma base de conhecimento defeituosa ao aplicar informações imprecisas adquiridas de fontes online.

2.3 Aprendizagem através do processo de busca na Web

Busca como processo de aprendizagem (BPA) pode ser vista de acordo com o conjunto de teorias e princípios pedagógicos que orientam o processo de aprendizagem (GIMENEZ et al., 2020). Mesmo tendo uma compreensão antagônica da aprendizagem, uma teoria não invalida a outra, e elas complementam e refinam as práticas educacionais (SCHUNK, 1991). Picciano (2017) examinou referenciais teóricos e modelos focados na educação através de meios online, mapeando três grandes paradigmas que orientam as teorias: (i) Comportamentalismo (ou Behaviorismo) com foco em como as pessoas se comportam; (ii) Cognitivismo que se baseia no conceito onde o processo mental tem um papel importante na aprendizagem, estabelecendo a conexão entre os estímulos ambientais e as respostas do indivíduo; e (iii) Construtivismo com foco no desenvolvimento do indivíduo, onde o conhecimento é construído como resultado da experiência individual ou da interação social (LEFRANCOIS, 2012).

Segundo Urgo e Arguello (2022), os mecanismos de pesquisas atuais não são projetados para apoiar o aprendizado conforme já definido em estudos anteriores sobre busca como processo de aprendizagem (VAKKARI, 2016; RIEH et al., 2016; HOPPE et al., 2018), pois são otimizados para cumprir tarefas de pesquisa simples e não fornecem suporte adequado para trabalhos em tarefas complexas que envolvem aprendizado (GIMENEZ et al., 2020). Isto é, sua finalidade de busca de informações factuais não é tão bem-sucedida em facilitar níveis mais elevados de conhecimento (TAIBI et al., 2017). As tarefas de aprendizagem diferem de acordo com a intenção de busca e apresentam complexidades distintas que promovem divergências no comportamento e no resultado da busca, uma vez que são realizadas buscas para esclarecer as mais variadas dúvidas, auxiliando na construção do conhecimento dos usuários em diferentes domínios (MACHADO; GIMENEZ; SIQUEIRA, 2020).

Pesquisas em BPA tem despontado como uma agenda que se concentra na aprendizagem que ocorre durante o processo de busca, considerando os resultados da busca e da aprendizagem (MACHADO; GIMENEZ; SIQUEIRA, 2020). Trabalhos como o de Rieh et al. (2016) tratam da pesquisa abrangente que facilitaria o aprendizado crítico e criativo além do aprendizado receptivo. Vakkari (2016) traz reflexões sobre o significado de “aprender” e “procurar” sob a mesma perspectiva, relacionando as etapas do processo de busca com os tipos de mudanças na estrutura do conhecimento. Dando-nos uma visão ampla de BPA, Hoppe et al. (2018) enfatizam a necessidade de exploração de recursos multimodais das fontes de aprendizagem e sua integração com recuperação da informação (RI) para desenvolvimento de pesquisas futuras.

Quando o propósito da busca é aprendizagem, segundo Rieh et al. (2016), podemos observar dois comportamentos contraditórios dos usuários: (i) o esmero/esforço por encontrar as palavras-chave certas, formular consultas e entender a busca; (ii) a aceitação dos resultados iniciais de forma acrítica, em vez de reavaliar os resultados ou revisar suas buscas, quando os resultados da busca não satisfazem suas necessidades de informação. Contudo, os usuários podem superestimar suas habilidades de busca, acreditando que analisar (acessar, pesquisar, utilizar e avaliar) fontes de informação são tarefas fáceis, e desta forma, tendem a percorrer, avaliar e selecionar o conteúdo online rapidamente, sem investir muito esforço (MACHADO; GIMENEZ; SIQUEIRA, 2020).

O estudo empírico de Tseng, Liang e Tsai (2014) examinou se o SRL e o padrões de avaliação de informações online (OIES — *online information evaluative standards*) dos aprendizes influenciaram suas estratégias de busca de informações online (OISS — *Online information searching strategies*). Nesse trabalho, quatro fatores decorrentes desses padrões foram verificados: (i) uso de várias fontes para validar se informações estão corretas; (ii) autoridade da fonte que dá confiabilidade à informação prestada; (iii) conteúdo correspondendo com a intenção de pesquisa; e (iv) aspectos técnicos que podem inviabilizar a utilidade da informação, tais como tempo de recuperação ou linguagem. Assim como três domínios das estratégias de busca de informações foram avaliados: (i) comportamental (“controle” e “desorientação”), que descreve as habilidades necessárias para manipulação e navegação na Internet, como as funções de pesquisa avançada fornecidas pelos mecanismos de pesquisa; (ii) processual (“tentativa e erro”, “resolução de problemas”), que descreve os procedimentos adotados durante a busca face aos resultados encontrados, e principalmente quando não há informações suficientes; e (iii) metacognitivo (“pensamento intencional”, “seleção de ideias principais” e “avaliação da informação”), geralmente associado ao planejamento de metas antes de executar ações de busca. No contexto de busca para aprendizagem, os resultados alcançados expressaram que os

níveis básico e avançado de autorregulação de aprendizagem (SRL) dos alunos desempenha papel crucial e o padrão de avaliação avançado de informações tem efeito positivo, sobre as estratégias de busca dos alunos (TSENG; LIANG; TSAI, 2014).

A partir do levantamento da literatura, Machado, Gimenez e Siqueira (2020) propuseram a organização das seguintes variáveis no processo de busca para aprendizagem em três dimensões abaixo:

- Usuário, que envolve as recomendações contextualizadas de RI ou compreensão do comportamento de busca.
 - Conhecimento Prévio do Usuário: representando o atual nível de aprendizagem do usuário, que pode ser rotulado de *Usuário Iniciante* a *Usuário Especialista* (LU; HSIAO, 2017), ou a partir de traços cognitivos (TAIBI et al., 2017).
 - Informações Demográficas: representa a descrição demográfica do usuário, que considera inclusive o idioma (TAIBI et al., 2017) nativo do usuário.
- Interação, que envolve a dinâmica de execução da busca com as informações e recursos habilitados pela ferramenta pelo usuário.
 - Atividades de busca exploratória: relacionadas ao caminho percorrido pelos usuários ao explorar o espaço de busca, bem como elementos de tomada de decisão. Podemos associar ao *Modelo de Busca Exploratória* de Marchionini (2006), assim como a taxonomia do Processo Intensivo de Conhecimento de Busca Exploratória (TIBAU et al., 2019).
 - Atividades de interação: baseadas nas ações de navegação do usuário ao analisar os recursos recuperados por suas pesquisas. Podemos considerar o uso do mouse, bem como o número, diversidade e complexidade de consultas, navegação (páginas visitadas, páginas salvas, tempo de permanência) e sessão de pesquisa (MACHADO; GIMENEZ; SIQUEIRA, 2020).
 - *Feedback* ativo do sistema de busca (SE): relacionadas ao comportamento da ferramenta de busca em relação às ações do usuário, tais como sugestões de consulta, sinônimos de palavras, gramática e verificação ortográfica, além de sinalizações sobre os resultados (*links* já visitados em roxo, por exemplo).
- Domínio do conhecimento, que envolve a base/fonte do domínio do conhecimento onde o usuário está buscando adquirir aprendizado.
 - Representação: representar um conjunto de conceitos dentro de um domínio e os relacionamentos entre eles. Envolve considerar a interligação desses

conceitos entre idiomas com base nos conceitos reconhecidos na consulta do usuário (TAIBI et al., 2017).

- Nível Cognitivo do Recurso: representa o nível cognitivo necessário para entender um recurso a ser retornado pela pesquisa.
- Características dos recursos: envolve os elementos, atributos e aspectos dos recursos retornados por um sistema de busca que podem influir na aprendizagem a curto ou longo prazo.

Em seu estudo dissertativo empírico, Zhang (2020) explorou como os indivíduos usam mecanismos de pesquisa e como estes podem apoiar melhor as necessidades de informação dos usuários ao trabalhar em tarefas que envolvem esforços criativos. Em uma primeira fase usou questionário online com 175 participantes e, depois, realizou um estudo diário de duas semanas com o uso de navegador *Firefox* customizado para registro diário *in-situ* com 17 participantes de um conjunto prévio de 33 estudantes da Universidade da Carolina do Norte.

Comparativamente, o estudo de Tsai et al. (2012) investigou o impacto nas estratégias de busca de informações online em função de contexto de busca de aprendizado de tópicos e de informações da vida cotidiana. Inicialmente, com uso de questionário aplicado a 304 estudantes universitários, depois por entrevistas intensivas em uma subamostra de 20 deles. Como resultado, os desempenhos para buscar informações da vida diária foram significativamente melhores daquelas utilizadas para atividades de aprendizagem de tópicos. Destaque para as estratégias comportamentais e metacognitiva, sugerindo que sejam dependentes do contexto, ou seja, devido ao conhecimento prévio e das experiências dos alunos em relação a um tópico específico. Verificaram também que tarefas difíceis, motivação limitada e confusão instrucional causam retardo do controle, geram ansiedade e desorientação com frequência aos alunos enquanto procuram informações relacionadas ao estudo. Isto é, o contexto da busca, o desenho das tarefas, qualidade das instruções percebidas para as tarefas e a motivação própria do indivíduo podem modificar o desempenho dos usuários durante as buscas para aprendizagem.

Liu, Liu e Belkin (2019) realizaram um estudo com 31 alunos universitários chineses de diversas especialidades de formação, que efetuaram duas tarefas de busca relacionadas ao aprendizado, e durante o experimento, eles preenchiam um questionário de antecedentes e utilizavam o mapa mental para organizar suas ideias e informações. Na caracterização de quando e como os usuários aprendem novos conhecimentos durante o processo de busca, identificaram que os usuários tiveram mudanças de conhecimento em todas as etapas do processo. Também descobriu-se que o tipo de tarefa tem um efeito significativo

nos estilos de mudança de conhecimento dos usuários.

Segundo [Tibau, Siqueira e Nunes \(2022\)](#), na busca exploratória na Web, após aplicar TAP a 6 professores participantes para avaliar o mapeamento de fluxo de conhecimentos utilizando o modelo ESKiP, distinguiu três comportamentos principais dos usuários durante a busca: (i) a familiarização com o assunto reduz a incerteza envolvida na atividade e possibilita selecionar termos de pesquisa úteis; (ii) a capacidade de autocontrole no processo de busca aumenta à medida que o usuário avança nas atividades, ganhando mais entendimento no tema, refletido pela reformulação de consultas, formalizada por melhores definições dos termos de busca; e (iii) a capacidade de avaliar a relevância da informação recuperada, determinando a decisão sobre o que será selecionado e extraído ([TIBAU; SIQUEIRA; NUNES, 2022](#)).

Quando consideramos os alunos utilizando os sistemas de busca, desempenhando o papel de usuário-aprendiz, podemos verificar que existem duas ações principais executadas: (i) o processo de tomada de decisão de formulação e reformulação de consultas ([LIU et al., 2010](#)); e (ii) a navegação onde o aprendiz interage com a informação ([HERDER; JUVINA, 2004](#)). Na primeira, o aluno deve usar adequadamente os termos-chave para recuperar informações relevantes e desenvolver estratégias de busca bem-sucedidas. Enquanto na segunda, o aluno é responsável por selecionar as fontes, identificando quais são confiáveis e estabelecer relações entre conceitos para processar informações para a construção do conhecimento. Assim, a dinâmica dessas duas tarefas contribui para a aprendizagem à medida que o aprendiz está engajado em seu processo. O ganho de conhecimento neste contexto vem do comportamento de busca na web e da análise do uso da informação ([WILSON, 1999](#)).

No levantamento de [Gimenez et al. \(2020\)](#), foram catalogados mecanismos que influenciam o processo de aprendizagem, tais como engajamento dos participantes, a avaliação como motivadora durante as sessões, a intervenção do professor (ou instrutor) durante a sessão, reordenação da página de resultados de pesquisa para apoiar a aprendizagem, julgamentos e intervenções em sistemas de busca para melhorar a experiência de aprendizagem, recursos de documentos da web e a relação com resultados de aprendizagem de curto ou longo prazo. Também verificou-se o projeto de sessão de busca para aprendizagem, que representa os recursos estruturais, ambientais e interativos da sessão usados para apoiar e garantir o aprendizado. Sob a ótica do que é a aprendizagem e como ela ocorre durante uma sessão de busca online, estamos preocupados em como os recursos de design da sessão fornecem informações importantes sobre o que torna os alunos mais eficazes.

2.4 Integrando a busca com CQA para prover Aprendizagem.

As perguntas postadas nas CQA devem refletir as necessidades dos usuários, assim como as consultas enviadas aos mecanismos de pesquisa (LIU et al., 2008). Os mecanismos de pesquisa tradicionalmente esperavam que as necessidades de informação fossem formuladas como palavras-chave, em vez de uma forma natural, como perguntas (WHITE; RICHARDSON; YIH, 2015).

Atualmente, devido às necessidades de informação, as pessoas interessadas em encontrar o conhecimento de que precisam, navegam na Internet para encontrar as respostas para suas perguntas (HAJIAMINSHIRAZI; MOMTAZI, 2020). As buscas são um meio para os usuários procurarem as informações necessárias (YUSUF; YUNUS; WAHID, 2019), encontrando-as em muitos casos nas CQA, que atuam como fonte. A visão de classificação da recuperação (*ranked retrieval*) de informação supõem que (i) as palavras da consulta da busca estejam presentes nos documentos que se deseja ver e (ii) que a classificação dos documentos recuperados será suficiente porque o usuário conseguirá reconhecer aqueles que desejavam encontrar (GALUŠČÁKOVÁ; OARD; NAIR, 2021). Isso não se aplica quando os documentos a serem pesquisados estão em um idioma desconhecido do usuário que realizou a busca, por isso recentes estudos exploram soluções para recuperação da informação multi-linguagem (CLIR: *Cross-language Information Retrieval*).

As CQA possuem questões monolíngues, com predominância da língua inglesa, especialmente em domínio técnicos. Tal característica é um entrave ao reuso das postagens para usuários não falantes de inglês (no caso da maioria das plataformas), restringindo o compartilhamento de conhecimentos existentes, dado que usuários de CQA nem sempre são bem-sucedidos em encontrar suas respostas em seus sistemas CQA nativos (HAJIAMINSHIRAZI; MOMTAZI, 2020). Além disso, os serviços de CQA são baseados em expressões dos usuários no modelo de *Bag-of-Word (BoW)* e possuem limitações para capturar a similaridade semântica das questões em relação à intenção de busca do usuário (FIGUEROA; NEUMANN, 2016). A busca nos serviços CQA pode obter as respostas corretas para as consultas de perguntas, encontrando automaticamente perguntas semelhantes que já foram respondidas por outros usuários, enquanto os mecanismos de busca retornam longas listas de resultados organizados em páginas, que precisam ser analisadas pelo usuário (ZHANG et al., 2014).

Novas abordagens de recuperação de perguntas em vários idiomas (CLQR: *Cross-language question retrieval*) tem sido exploradas recentemente (LI; DU; CHEN, 2020;

HAJIAMINSHIRAZI; MOMTAZI, 2020), indo além dos modelos tradicionais de recuperação de informação (RI), ou do uso de estratégias baseadas em tradução (LEE et al., 2008; XUE; JEON; CROFT, 2008) ou em modelo de tópicos (ZHANG et al., 2014), para superar as limitações de CQA com questões monolíngues. Da mesma forma, novas abordagens tem sido propostas, como o uso de classificação semântica sensível ao contexto de consultas de pesquisa em bases de conhecimento das CQA, contribuindo para uma integração eficaz dos mecanismos de pesquisa Web com tais bases (FIGUEROA; NEUMANN, 2016).

No contexto de buscas online, taxonomias tem sido propostas para descrever a intenção das perguntas feitas em mecanismos de busca na Web. Elas buscam refletir o comportamento de busca recente dos usuários e aprofundam a compreensão das intenções, objetivos e respostas esperadas do usuário, amparadas nas novas necessidades de informação e na constante evolução dos mecanismos de busca e suas novas funcionalidades ofertadas aos usuários (CAMBAZOGLU et al., 2021).

As buscas com intenção de responder a perguntas aparecem na categoria informativa, que pode ser dividida em (WHITE; RICHARDSON; YIH, 2015): (i) direcionada, com objetivo de aprendizado de algo específico sobre um tópico, e se subdividindo em Aberta ou Fechada, quando a pergunta pode ser respondida com uma única resposta inequívoca; e (ii) não direcionada, com objetivo de aprendizado sobre um tópico em geral.

O trabalho de White, Richardson e Yih (2015), baseado em estudos sobre o aumento de buscas formuladas como linguagem natural, investigou a natureza e a prevalência dessas intenções e o impacto desse aumento no desempenho do mecanismo de pesquisa. Como resultados, os autores observaram que 10,3% das consultas emitidas para um mecanismo de busca têm intenção de pergunta direta e 3,2% são formuladas como perguntas em linguagem natural. Assim como, o mesmo bom desempenho foi verificado quando a intenção de pesquisa é declarada como consultas ou perguntas. Já no trabalho de Cambazoglu et al. (2021), um volume pequeno de perguntas do tipo *'Why, Where, When, Is/Are/Do/Does'* foi relatado, podendo ser devido (i) a um desempenho relativamente ruim de mecanismos de buscas em responder a esse tipo de pergunta, ou (ii) aos usuários utilizarem outros meios para satisfazer suas necessidades de informação, como as CQA para as perguntas sobre o *'Why'*. Também foi identificado que uma passagem ou um pedaço de texto seria suficiente para responder adequadamente a quase todas as perguntas. Isto é, há uma demanda crescente por respostas diretas nas buscas na Web.

Segundo Metzler et al. (2021), os usuários desejam o envolvimento com um especialista para tratar de suas necessidades de informação, porém acabam recorrendo a um

mecanismo de busca, que está disponível e que não responde diretamente às necessidades, mas provê referências a respostas. Já os sistemas de Q&A de sucesso possuem *corpus* limitado não escalável criado sob demanda por especialistas humanos.

Segundo [Liu et al. \(2008\)](#), um mecanismo de pesquisa pode exibir perguntas semelhantes ou usar as melhores respostas como *snippets* de resultados de pesquisa quando consultas semelhantes forem realizadas. Nesse cenário ideal, assume-se que as melhores respostas são respostas boas e relevantes para as perguntas de emparelhamento por semelhança ([LIU et al., 2008](#)). Assim, a combinação de ambas (busca Web e CQA) melhoraria muito a eficiência da recuperação de informações online.

No processo de busca por respostas para perguntas relacionadas ao conteúdo escolar, os alunos usam como principal fonte de pesquisa os mecanismos de busca na Web, que podem retornar resultados fora do contexto educacional, ou como segunda escolha em uma tendência crescente, os sites de Q&A da comunidade social, onde podem obter respostas de outros colegas online ([YILMAZ et al., 2019](#)).

Em *survey* encontrado na literatura ([STACKOVERFLOW, 2020](#)), verificou-se que muitas vezes desenvolvedores desejam uma solução de codificação, usam os mecanismos de busca na Web e se deparam como primeiro resultado um *link* na cor roxa (porque já o visitaram antes). As respostas sugerem uma ocorrência frequentemente aceitável para determinadas tarefas (51,6%), enquanto o sentimento relatado variaram entre 18,3% indiferente, 15,9% divertido e 14,3% incomodado. Isto demonstra a expectativa que os desenvolvedores possuem em relação a certas tarefas frente ao uso da busca e de fonte CQA, e a uma provável inadequação de suas estratégias de busca, face aos sentimentos manifestados.

Geralmente, segundo [Srba e Bielikova \(2016\)](#), o processo de construção de uma pergunta na CQA começa por uma busca malsucedida de recuperação de informação na web através de um mecanismo de busca, ou até mesmo na busca dentro da própria CQA. A pergunta formulada não precisa ser reduzida a palavras-chave ou limitada a alguma semântica básica, como ocorre no processo de formulação da consulta nas buscas online na web. Assim, as informações necessárias podem ser melhor descritas, com mais precisão e clareza, possibilitando uma melhor compreensão dos usuários. O mesmo se sucede com as respostas. Isto seria uma contribuição da CQA para mecanismos de busca na concepção de busca como processo de aprendizagem.

2.5 Trabalhos relacionados

A aprendizagem informal para um processo de aprendizagem individual parece se dividir em dois grupos de estudos empíricos, associados ao desenvolvimento profissional ou ao desempenho do aluno (ambiente acadêmico). O trabalho de [Jeong et al. \(2018\)](#) sintetiza a literatura a partir de 465 artigos e apresenta o intenso interesse de pesquisa para o desenvolvimento profissional do indivíduo, propondo uma estrutura conceitual para entender as atividades de aprendizagem informal usando três dimensões: competência de aprendizagem, intencionalidade e relação de desenvolvimento. Nesse estudo, os autores usam a Teoria de Campo de Lewin, baseado na ideia de que o ambiente estimula a mudança de comportamento através das forças sociais ([ROȘCA, 2020](#)), aplicada para identificar antecedentes de aprendizagem nos níveis individual, grupal e organizacional.

Nessa perspectiva profissional, existem muitos trabalhos investigando o uso de CQA como recursos de aprendizagem online para desenvolvimento de software ([MAMYKINA et al., 2011](#)) e recentemente para desenvolvimento de jogos ([KAMIENSKI; BEZEMER, 2021](#)). O foco está na aquisição de novos conhecimentos e na solução de problemas ([MAMYKINA et al., 2011](#); [ZHANG et al., 2019](#)), pela utilização das informações compartilhadas, principalmente os códigos-fontes existentes ([WU et al., 2019](#)). Contudo, à medida que os usuários tornam-se mais experientes tendem a diminuir o uso da CQA ([KAMIENSKI; BEZEMER, 2021](#)). Isto, provavelmente acontece em função da qualidade do conteúdo, com destaque para o código-fonte, devido à necessidade de muita modificação/adaptação para reuso, à incompreensibilidade ou a baixa qualidade de código, que poderia ser mitigado com o uso de um ambiente validador integrado para teste do código ([WU et al., 2019](#)). Assim, temos estudos focados na análise da qualidade do conhecimento (*crowdsourcing*) compartilhado ([ZHANG et al., 2019](#)) e como destacar as contribuições de conhecimento científico profissional através de dispositivos linguísticos específicos utilizados nas postagens ([ZHANG et al., 2019](#)).

Na perspectiva de comunidades de práticas, alguns trabalhos sobre CQA buscam determinar as características que influenciam o comportamento da contribuição do conhecimento em uma CQA. Um exemplo recente é a investigação empírica ([GUAN et al., 2018](#)) que alinha algumas teorias e analisa da interação dos usuários chineses na plataforma Zhihu ², que possui mais de 17 milhões de usuários, 3,5 milhões de questões em mais de 100 mil tópicos. Para identificar os aspectos socio-técnicos foram utilizadas análises estatísticas sobre o conjunto de metadados das relações entre conteúdo e usuários na comunidade. Uma contribuição relevante do trabalho ([GUAN et al., 2018](#)) foi a distin-

² <https://en.wikipedia.org/wiki/Zhihu>, <https://www.zhihu.com/>

ção entre dois perfis de usuários em relação à contribuição contínua de conhecimentos: produtores e consumidores. Os consumidores de conhecimentos procuram respostas de outros usuários fazendo perguntas. Já os produtores, se envolvem nas discussões de tópicos e expressam suas opiniões e informações. A reciprocidade aparenta ser a principal causa da contribuição contínua nessas comunidades de Q&A social, quando os usuários contribuem devido à expectativa de respostas para as próprias perguntas no futuro.

Com foco na aprendizagem informal do aluno, observamos o surgimento de estudos que avaliam as trilhas de Q&A na CQA como recursos que poderiam ser utilizados como material de apoio de aprendizagem (KUMAR; CHAUHAN, 2020). Por exemplo, Kumar e Chauhan (2020) propuseram uma solução tecnológica de processo de enriquecimento de texto baseado em tópicos para conteúdo dos livros didáticos utilizados para aprendizagem. Os alunos tendem a usar as trilhas de CQA para obter mais conhecimento. Estudos anteriores buscavam identificar as características específicas dessas CQA que influenciam positivamente nos resultados da aprendizagem, com uso de teorias sociais cognitivas para classificá-los de acordo com os recursos que eles fornecem e a qualidade da experiência do usuário (LIU; CHENG; LIN, 2013). Como podem existir várias repostas associadas à mesma pergunta, gastam muito tempo coletando as informações necessárias da web para obter uma imagem melhor do conceito de um livro didático (KUMAR; CHAUHAN, 2020).

Possivelmente impulsionado pelo novo cenário digital mundial devido à pandemia da Covid-19, observamos o uso de CQA para área de saúde, como ferramenta de interação paciente-médico. O estudo de Hong et al. (2020) apresenta o uso de CQA para a interação de pacientes crônicos (diabetes e hepatite), que precisam de acompanhamento constante sobre três tópicos: prevenção e exame, diagnósticos, e tratamento. A conclusão após aplicar métodos de análise linear sobre as perguntas e repostas, foi do efeito negativo do tamanho da pergunta e do valor da recompensa no número de respostas recebidas, sugerindo que os pacientes não precisam pagar uma taxa para receber uma resposta, mas precisam refinar e reduzir suas perguntas se quiserem respostas mais oportunas de especialistas.

O estudo de (SONG; LIU; ZHANG, 2021) examinou a seleção de fontes dos estudantes universitários e os comportamentos de uso ao procurar informações relacionadas à saúde, por meio do uso de *log* do mecanismo de busca *Baidu* de realização de quatro tarefas sobre saúde e cinco tipos de fontes. Identificaram que as CQA sociais são usadas com maior frequência, os sites de saúde profissional menos selecionadas e utilizadas, e os de notícias e do tipo *Wiki* com maior probabilidade de uso. Nesse estudo, percebeu-se que o tipo de tarefa tem impacto significativo na seleção e uso da fonte, onde os sites *Wiki* foram mais selecionados e utilizados para as mais simples (receptivas) enquanto os sites

de CQA e notícias para lidar com tarefas críticas.

Poucos trabalhos exploram a contribuição dos comentários para o conhecimento *crowd-sourced*. [Zhang et al. \(2019\)](#) constataram que eles são informativos e ocorrem em sua maioria após a aceitação de uma resposta com pouca integração às respostas associadas. Por isso os usuários inexperientes tendem a levantar limitações e preocupações, enquanto usuários experientes tendem a aprimorar a resposta por meio de comentários. Já [Zhu et al. \(2021\)](#) focou natureza das discussões de questões, verificou que a prevalência das discussões de perguntas ocorre antes da primeira resposta e que as perguntas mais discutidas recebem respostas mais lentamente.

Também, há poucos estudos sobre a navegação dentro das CQA que permita aos usuários encontrar as informações necessárias, mesmo para a *Stack Overflow* que tem sido alvo de muitas pesquisas, ajudando-o como decidir se uma resposta é relevante para sua tarefa e contexto. Isto é devido a dificuldade de descobrir qual é a melhor solução para o problema específico, em função das opções de informações contextuais que os usuários gostariam de usar para restringir a busca ([NADI; TREUDE, 2020](#)). Estudos avaliam o sistema de recompensas como ferramenta para garantir ao máximo soluções para as perguntas, e assim reduzir o quantitativo sem respostas.

No trabalho de [Zhou et al. \(2020\)](#), ao analisar o sistema do *Stack Overflow*, foram verificados que questões com recompensa (por exemplo, questões abertas há mais de 100 dias sem resolução e que foi proposta uma recompensa) levam à maior probabilidade de solução de uma questão na maioria dos casos. Outros estudos avaliam fatores que impactam na contribuição contínua de conhecimentos, como o trabalho de [Chen, Baird e Straub \(2019\)](#), ao aplicar estimativa de regressão de Poisson, verificou que a votação impacta na motivação da contribuição (votos positivos motivam, votos negativos desestimulam) e que comentários podem moderar a relação entre votação e contribuição. Destacamos a existência de pesquisas específicas sobre comunidades sociais acadêmicas com ambiente Q&A embutido (*ResearchGate*), onde o foco está na identificação da motivação de membros para responder, ler, seguir ou recomendar uma pergunta, ou uma resposta. Um exemplo é o trabalho de [Deng et al. \(2019\)](#), que concluiu pela existência de limitar de tamanho da descrição da pergunta (cerca de 150 palavras) a partir do qual os acadêmicos rapidamente perdem o interesse, e assim não leem a descrição.

Estudos que buscam teorizar sobre as CQA, são em sua maioria pesquisas empíricas sobre a contribuição do conhecimento na comunidade, majoritariamente associadas a Teoria da Autodeterminação ([LOU et al., 2013](#); [CHEN; BAIRD; STRAUB, 2019](#)), a Teoria Cognitiva Social ([HSU et al., 2007](#); [CHEN; HUNG, 2010](#)). Esta última por vezes

associada a teorias do capital social e da troca social (JIN et al., 2015). A Teoria da Auto-determinação (SDT — *Self-determination Theory*) é uma teoria derivada empiricamente da motivação humana e da personalidade em contextos sociais que diferencia a motivação em termos de ser autônomo e controlado (DECI; RYAN, 2012). Ela foi relacionada na investigação de fatores motivacionais que afetam distintamente a qualidade e a quantidade da contribuição voluntária de conhecimento (LOU et al., 2013), bem como possibilidades motivacionais nas formas de votação e comentários (CHEN; BAIRD; STRAUB, 2019).

A Teoria Cognitiva Social ressalta os papéis-chave dos processos vicários, simbólicos e autorregulatórios na aprendizagem e no desempenho das ações (BANDURA; WALTERS, 1977). Ela foi considerada ao investigar fatores significativamente influentes no aumento da transferência de conhecimento (CHEN; HUNG, 2010), tais como a norma de reciprocidade, confiança interpessoal, vantagem relativa percebida, e também autoeficácia de compartilhamento e expectativas de resultados para influências pessoais (HSU et al., 2007), assim como a auto-apresentação (*user's self-presentation*) de um usuário, o reconhecimento dos pares e a aprendizagem social (JIN et al., 2015). No estudo de Cavusoglu, Li e Huang (2015) foi proposto um sistema de crachás hierárquicos, um mecanismo de recompensa baseado em princípios de gamificação, na participação e engajamento do usuário na CQA *Stack Overflow*, que apresentou fortes evidências empíricas que confirmam o valor desses princípios. Já Zhao, Detlor e Connelly (2016), abordaram o efeito do prazer em ajudar os outros na atitude do usuário em relação ao compartilhamento do conhecimento. Este efeito é prejudicado pelas recompensas organizacionais virtuais, enquanto o efeito da autoeficácia do conhecimento na atitude em relação ao compartilhamento do conhecimento é prejudicado pela reciprocidade.

Nas abordagens empíricas dos estudos da literatura recente que analisam a CQA na perspectiva de compartilhamento de conhecimento e aprendizagem, preponderam o uso da análise de dados coletados das plataformas (KAMIENSKI; BEZEMER, 2021; WU et al., 2019; CHEN; HUNG, 2010; HSU et al., 2007; JIN et al., 2015; DENG et al., 2019) e pesquisa de respostas dos participantes a partir da aplicação de questionários (KAMIENSKI; BEZEMER, 2021; SRBA et al., 2019; WU et al., 2019; NADI; TREUDE, 2020; LOU et al., 2013). Outros estudos avaliaram os efeitos da CQA para aprendizagem, com separação de grupo de participante por CQA para responder as mesmas perguntas, seguida de pós-testes e mediada por um instrutor (LIU; CHENG; LIN, 2013). Temos ainda, através da metodologia de pesquisa-ação, avaliação de somente perguntas e respostas associadas a vídeos de aula (p.ex. física), estruturados em ciclos com testes de observação e resultados de aprendizagem (PUTRA; JUNAID; SULMAN, 2021). Finalmente, em relação à abordagem metodológica, o artigo de Liu, Li e Yang (2020) aplica a abordagem e

os procedimentos metodológicos similares aos empregados em nosso trabalho. Contudo, não envolve CQA e nem a perspectiva de aprendizagem. Tal artigo foca o processo de construção de conhecimento através de busca exploratória (LIU; LI; YANG, 2020), através de análise qualitativa de dados coletados de 25 participantes de diversos domínios, a partir da gravação de sessões de pensamento em voz alta e entrevistas semiestruturadas.

Desta forma, nosso estudo difere dos demais estudos levantados da literatura corrente pela abordagem metodológica *Grounded Theory* e ao utilizar dados a partir da interação de usuários com as plataformas CQA em cenários simulados, representando a experiência dos participantes na busca por informação através da Web com base em CQA. Assim, levando a busca como processo de aprendizagem para o contexto de CQA. Também optamos por utilizar entrevistas e técnicas observacionais ao uso de questionário, por permitir um maior engajamento dos participantes e possibilitar a exploração de novas informações e perspectivas associadas.

2.6 Conclusão da fundamentação

Neste capítulo apresentamos a fundamentação para o estudo, com base na literatura, cobrindo os tópicos relacionados ao fenômeno de interesse, que envolve a busca como processo de aprendizagem no contexto da CQA na perspectiva de aprendizagem informal. Caracterizamos a aprendizagem autorregulada do usuário-aprendiz como especialização da aprendizagem informal, sendo um tipo de aprendizagem ao longo da vida. Abordamos o uso da CQA para aprendizagem ao compartilhar conhecimentos especializados, destacando características e aspectos sociotécnicos, assim como o desenvolvimento e perspectiva de aprendizagem através de BPA. Com isso, apresentamos a integração de BPA com CQA para promover a aprendizagem. Por fim, apresentamos diversos trabalhos relacionados ao presente estudo, que fornecem muitos pontos de intercessão para serem explorados, apesar de não serem exatamente similares, principalmente devido ao contexto de BPA em CQA, à natureza qualitativa da pesquisa e à abordagem metodológica empregada.

Considerando que uma das contribuições desta pesquisa, e também um diferencial, está na abordagem metodológica, descrevemos no próximo capítulo a abordagem epistemológica-metodológica contemplada nesta pesquisa.

3. Abordagem epistemológica-metodológica de pesquisa

As questões de natureza epistemológicas abordam teorias de conhecimento e uma compreensão do fenômeno de interesse, e ainda, sugerem a exploração de ações, processos e percepções dos participantes encontrados nos dados (SALDAÑA, 2013). Na análise, para catalogar e revelar esses conhecimentos a partir da compreensão do fenômeno, utilizamos uma seleção dos métodos de codificação detalhados em um guia de codificação (SALDAÑA, 2013). Esse guia de codificação é organizado apresentando cada método e sua finalidade, organizado em seções de fontes/origens, descrição, a aplicação, exemplos, análises e notas associadas.

Como o fenômeno de especial interesse é contextualizado pelo processo de busca na *Web* como processo de aprendizagem, restrito às fontes de conhecimentos em CQA, são relevantes para o entendimento do fenômeno, tanto o comportamento quanto os processos analíticos (e cognitivos) do participante durante a realização de atividades envolvidas.

A presente pesquisa segue a abordagem qualitativa, concebida sob o paradigma interpretativa/construtivista, onde os significados atribuídos pelos indivíduos para objetos e situações são mais importantes que o entendimento da dinâmica do mundo social sob a perspectiva natural/física a ela atribuídas. As múltiplas faces da realidade são resultados da construção social proveniente das experiências e interações dos indivíduos (CRESWELL, 2014). Com isso, não é possível estabelecer uma perspectiva única sobre os acontecimentos, visto que depende da interpretação do observador-pesquisador sobre todos os achados oriundos dos procedimentos qualitativos.

3.1 Abordagem Metodológica

A abordagem metodológica *Grounded Theory*¹, combinada com o método *Think-Aloud Protocol* (TAP)², foram escolhidos devido à natureza epistemológica do fenômeno em estudo, à ausência de *dataset* disponível que possibilitasse explorar as diferentes facetas do comportamento do usuário na busca, e à lacuna observada na literatura para o fenômeno em questão, o cenário de busca como processo de aprendizado em CQA. Além da natureza epistemológica, buscávamos compreender esse comportamento não apenas a partir da interação do usuário, mas principalmente por suas reflexões no processo. O método TAP foi utilizado para coleta dos dados, por possibilitar a descrição do comportamento, ações e principalmente do “pensamento” do participante pelo próprio durante sessões simuladas que retratam cenários reais em sistemas de informação. Adicionalmente, a técnica de *Intensive Interviewing* (CHARMAX, 2014) foi aplicada para complementar às informações coletadas, bem como as atividades de observação do pesquisador.

3.1.1 Teoria fundamentada em dados (TFD)

A Teoria Fundamentada em Dados (do inglês “*Grounded Theory*”) é uma abordagem metodológica utilizada em pesquisa qualitativa. Nela os dados são sistematicamente coletados e analisados para fundamentar teorias, onde o foco está na busca por conceitos emergentes dos dados (CORBIN; STRAUSS, 2014). Adicionalmente, explora como cada conceito se desenvolve, muda e dá origem às consequências (CHARMAX, 2014). Convém destacar que na pesquisa qualitativa, os resultados não são provenientes de procedimentos estatísticos, mas referem-se aos estudos sobre a vida do indivíduo, experiências de vida, comportamentos e emoções, considerando os aspectos e interações sociais, organizacionais, culturais (CORBIN; STRAUSS, 2014).

A teoria fundamentada em dados (TFD) tem suas origens interacionistas, sendo Interacionismo Simbólico (NICO et al., 2007) um referencial teórico muito utilizado para essa abordagem. O Interacionismo Simbólico (CHARON, 2010) tem como característica principal a inseparabilidade do indivíduo do contexto dentro do qual ele existe, e baseia-se em três suposições relativamente simples: (i) as pessoas se esforçam e agem em direção ao que representa um significado para elas, (ii) o significado surge da interação social e (iii) o significado está sendo tratado e modificado por processos interpretativos. Nesse contexto, o ser humano tem uma imagem ativa maior, os indivíduos interagem e as sociedades são constituídas de indivíduos interagindo, e assim, o ser humano é compreendido

¹ <https://en.wikipedia.org/wiki/Grounded_theory>

² <https://en.wikipedia.org/wiki/Think_aloud_protocol>

como um ser agindo no presente, influenciado não somente pelo que aconteceu no passado, mas pelo que está acontecendo agora (CHARON, 2010). O que fazemos está ligado a essa interação, que não é somente externa, entre os indivíduos, mas também acontece no próprio indivíduo, em um 'mundo' por ele definido, agindo conforme sua definição para a situação em que está vivenciando (CHARON, 2010). Embora essa definição possa ser influenciada por quem interagimos, ela é também resultado de nossa interpretação da situação, envolvendo uma escolha consciente da direção de nossas ações em face dessa definição, a identificação dessas ações e a de outros e a nossa própria redireção (NICO et al., 2007 apud CHARON, 2010).

Na evolução da TFD (SANTOS et al., 2018), os pesquisadores que desenvolveram o método começaram a divergir sobre os procedimentos metodológicos e adotaram linhas de trabalho independentes. Daí surgiram duas abordagens: (1) perspectiva clássica ou Glaseriana (de Glaser) do método, mantendo-se na proposição original; (2) perspectiva relativista ou Straussiana (de Strauss, em parceria com Juliet Corbin) (CORBIN; STRAUSS, 2014), que incorporou novos procedimentos de análise e etapas para o desenvolvimento da teoria. Na década de 2000, deu-se início a uma nova abordagem segundo a perspectiva construtivista do método, idealizada por Charmaz (CHARMAX, 2014). Todas essas abordagens principais possuem características comuns que podem ser consideradas como princípios inerentes ao método: (a) amostragem teórica, (b) análise comparativa constante dos dados, (c) elaboração de memorandos e (d) diferença entre teoria substantiva e teoria formal (SANTOS et al., 2018).

A diferença entre teoria substantiva, que provém de um contexto específico e produz uma teoria aplicada somente ao campo investigado, e a teoria formal foi estabelecida na origem da TFD, que vê nesta última uma consequência da primeira em um contexto mais amplo e universal (SANTOS et al., 2018)l.

A amostragem teórica refere-se ao processo de coleta de dados com objetivo de encontrar elementos que potencializem a identificação de modificações de conceitos, adensamento de categorias, suas propriedades e dimensões, consoante o desenho da pesquisa. Segundo Santos (SANTOS et al., 2018) uma das estratégias para obtenção da amostragem teórica é a composição de grupos amostrais com participantes diferentes, mas com experiências relevantes em relação ao fenômeno em investigação. A amostra é definida inicialmente, mas é expandida para o aprofundamento de conceitos visando o preenchimento das lacunas na teoria emergente, dado o caráter cíclico do método até se atingir o alcance da saturação teórica. Os dados passam pelo processo dinâmico de dedução (construção de hipóteses), indução (assimilação de implicações advindas das hipóteses) e verificação (qualificação ou negação das hipóteses), tornando a explicação teórica cada

vez mais forte.

A análise dos dados (SANTOS et al., 2018) é eminentemente um processo de comparação constante dos dados, conhecido como análise comparativa constante. Os dados coletados passam pelo escrutínio de análise por palavras, linhas ou incidentes, visando gerar códigos conceituais. Depois tais códigos são agrupados e categorizados em um nível superior conceitualmente. São previstos comparações entre códigos, códigos com categorias emergentes e categorias entre si. À medida que os conceitos surjam, a reflexão do pesquisador são registrados em formato de memorandos, que servem como base de considerações para apoio ao desenvolvimento da teoria.

A perspectiva clássica da TFD aplica o Positivismo como base filosófica, apesar de não estabelecer formalmente um referencial teórico (GLASER, 1992). Na perspectiva dissidente relativista (de Straus e Corbin) temos a aplicação dos princípios do Interacionismo Simbólico e do Pragmatismo permeando a metodologia. Na perspectiva construtivista, temos o resgate de ambas referências teóricas com críticas ao seu antecessor (relativista), na ênfase interacionista pragmática e simbólica no significado, na linguagem, na interpretação e na interação, vinculando a TFD ao paradigma construtivista como uma metodologia interpretativa (SANTOS et al., 2018).

A *Constructivist Grounded Theory*³ (CHARMAX, 2014), descreve o processo de condução da observação do participante, considerando o papel ativo do pesquisador na compreensão específica do fenômeno em estudo (WILLIG, 2013). Distingue-se de outras abordagens que restringem o uso da literatura ao longo de todas as fases da pesquisa. Percebe a necessidade de compilação da literatura após a análise dos dados para fortalecer o potencial argumentativo da pesquisa. Essa abordagem construtivista se desvia da abordagem clássica ao construir mais de uma categoria central e ao identificar mais de um processo social. Assim como, recomendar a descrição da teoria em desenvolvimento segundo um formato de narrativa usando reflexões e interpretações (CHARMAX, 2014).

3.1.2 Intensive Interviewing

A *Intensive Interviewing* (CHARMAX, 2014) permite delinear o contexto e a abrangência do fenômeno sob a perspectiva e as experiências do participante, possibilitando que novos dados emergjam (MILLS; BONNER; FRANCIS, 2006) a partir da interação durante a entrevista, que de outra forma talvez não fosse possível (CHARMAX, 2014). A profundidade e o escopo dos dados fazem a diferença para a qualidade e credibilidade do estudo. Assim, além de sua utilidade para o desenvolvimento de categorias centrais,

³ <https://en.wikipedia.org/wiki/Grounded_theory#Constructivist_grounding_theory>

dois outros critérios para os dados são sua adequação e suficiência para representar eventos empíricos (CHARMAX, 2014). Geralmente usa amostras menores, fornecem riqueza de informações detalhadas sobre as razões pelas quais os entrevistados dão as respostas específicas, além de permitir uma longa observação das respostas não-verbais dos entrevistados. Desta forma, com o uso da *Intensive Interviewing* permite uma exploração aprofundada de um determinado tópico ou experiência, despontando como um método útil para a investigação interpretativa.

No contexto da TFD, as perguntas formuladas pedem ao participante que descreva e reflita sobre suas experiências de maneiras que raramente ocorrem na vida cotidiana. Tais questões são mais amplas e abertas, permitindo que em seguida sejam mais concentradas em discussão detalhada do tópico de interesse. Segundo Charmaz ((CHARMAX, 2014), esse modelo de perguntas incentiva o surgimento de afirmações e histórias imprevistas. Isto permite que o entrevistador: (i) vá além da superfície das experiências descritas; (ii) pare para explorar uma declaração ou tópico; (iii) solicite mais detalhes ou explicação; (iv) pergunte sobre os pensamentos, sentimentos e ações do participante; (v) mantenha o participante no assunto; (vi) volte a um ponto anterior; (vii) reafirme o ponto do participante para verificar a precisão; (viii) diminua ou acelere o ritmo; (ix) mude o assunto imediato; (x) validar a humanidade, perspectiva ou ação do participante; (xi) use habilidades de observação sociais para aprofundar a discussão. O que permite ao participante: (i) quebrar o silêncio e expressar suas opiniões; (ii) contar suas histórias e dar-lhes um quadro coerente; (iii) refletir sobre eventos anteriores; (iv) ser especialista; (v) escolher o que e como contar; (vi) compartilhar experiências significativas e ensinar o entrevistador como interpretá-las; (vii) expressar pensamentos e sentimentos não permitidos em outros relacionamentos e ambientes; e (viii) receber afirmação e compreensão.

3.1.3 Thinking Aloud Protocol — TAP

O método *Think-Aloud Protocol* (ou *Thinking Aloud Protocol* — TAP) foi derivado de uma análise teórica dos processos cognitivos que mediam a conclusão de uma tarefa. Identificou-se o modo de verbalizar pensamentos relacionados à tarefa com menor probabilidade de impactar os processos de mediação da execução da tarefa, vinculando à entrada do pensamento em atenção. Com isso o foco principal dos participantes recai na conclusão da tarefa principal, tornando-se a verbalização uma tarefa secundária subordinada que levaria à incompletude ao invés da reatividade. Foi desenvolvida uma metodologia de instrução, aquecimento e lembrança para apoiar este método para verbalizar pensamentos (ERICSSON; FOX, 2011).

No TAP, os participantes expressam o que estão pensando em voz alta, e externalizando a fala interna — todas as palavras que lhes vêm a mente, durante a realização das atividades (*Concurrent Think-Aloud Protocol* — CTAP) ou após a conclusão destas (*Retrospective Think-Aloud Protocol* — RTAP) (KUUSELA; PAUL, 2000). Isto é, envolve a mera verbalização de pensamentos em atenção enquanto o participante está totalmente engajado em completar uma tarefa, ou imediatamente após. Este mecanismo é fundamentalmente diferente do método introspectivo (ERICSSON; FOX, 2011). Esses pensamentos em voz alta, os relatos verbais, devem ser considerados e tratados como dados. A qualidade e utilidade dos mesmos estão associadas ao quão detalhado foram relatados e a recência em relação à tarefa que lhes deu causa (ERICSSON; SIMON, 1980). Tais relatos verbais são propícios para descrever os processos cognitivos envolvidos nas decisões e nos comportamentos do participante durante as sessões de busca. Contudo, os processos cognitivos observáveis não podem ser profundos ou complexos, segundo o conceito de *inner speech* — Vygotsky's (1962) *Thought and Language* e a sua natural restrição para uso na comunicação entre as pessoas (CHARTERS, 2003), dado que demandariam a simplificação em palavras antes de tornarem-se conhecidos pela audiência (o pesquisador, na sessão), representando um gargalo entre o pensamento abstrato amplo e o pensamento verbalizado, mais estreito, emergencial e temporal, gerando um retardo no processo de pensar (CHARTERS, 2003). Assim o foco é o processo cognitivo que reflete o pensamento em atenção na execução da tarefa. Por isso, os relatos verbais precisam ser combinados com outros dados oriundos de outros comportamentos não verbais em um processo de interpretação (ERICSSON; SIMON, 1980). Convém destacar, que tais outros dados, no caso da presente pesquisa, são anotações do pesquisador e observações dos elementos capturados pelas ações em tela (onde clicou, onde parou o mouse, que trechos ou segmentos foram marcados, etc.).

3.2 Procedimentos metodológicos do estudo

O presente estudo empírico é um experimento qualitativo. Conforme Kleining (1986) (KLEINING, 1986), o experimento qualitativo é um método de pesquisa social empírica que pode ser definido como a intervenção/operação em um objeto (social) para pesquisar sua estrutura e heurística, realizada de acordo com regras científicas. Assim, nosso estudo interpretativo e qualitativo está focado na análise do comportamento do usuário em um processo de busca de informações na Web envolvendo exclusivamente CQA, com consequente avaliação de CQA como fonte de informações factuais relevantes e base de conhecimentos especializados (GIMENEZ; SIQUEIRA, 2021). Partindo das correlações

desses dois eixos: informação factual e base de conhecimento, destacamos as características relevantes, do ponto de vista socio-técnico, para suporte à aprendizagem do usuário a partir da perspectiva dos participantes. Nesse procedimento, foram explorados os aspectos do processo cognitivo e analítico do usuário perante o mecanismo de busca fornecido e a organização da CQA, considerando que elementos e fatores facilitariam ou restringiriam um possível processo de aprendizagem baseado em busca na Web sobre as CQA. Tal procedimento foi dividido em quatro etapas sequenciais: preparação, execução, coleta de dados e análise.

- **PREPARAÇÃO.** Consiste na (i) construção do ambiente controlado, que abrange a busca customizada utilizada no experimento considerando as CQA selecionadas; (ii) elaboração dos tópicos da entrevista e do projeto das sessões propostas; e (iii) definição do grupo de participantes e subsequente seleção dos mesmos.
- **EXECUÇÃO.** Consiste na disponibilização do ambiente para acesso pelo participante. Também foi realizada explicação sobre a dinâmica da participação do estudo como todo, descrevendo previamente as orientações prestadas no convite, conforme Apêndice A, esclarecendo sobre TAP e os procedimentos a serem adotados na execução do experimento com o participante. Também inclui a realização do experimento e gravação de vídeo das interações.
- **COLETA DE DADOS.** Consiste na extração, transformação e transcrição textual das entrevistas e do relato verbal do participante, bem como das observações não verbais anotadas pelo pesquisador associadas às ações e comportamentos dos participantes.
- **ANÁLISE DOS DADOS.** Consiste na análise qualitativa dos dados coletados, partindo da codificação dos dados até a extração de características e possível formulação de temas e teorização, segundo a metodologia TFD.

3.2.1 Ambiente de busca customizada na Web sobre CQA

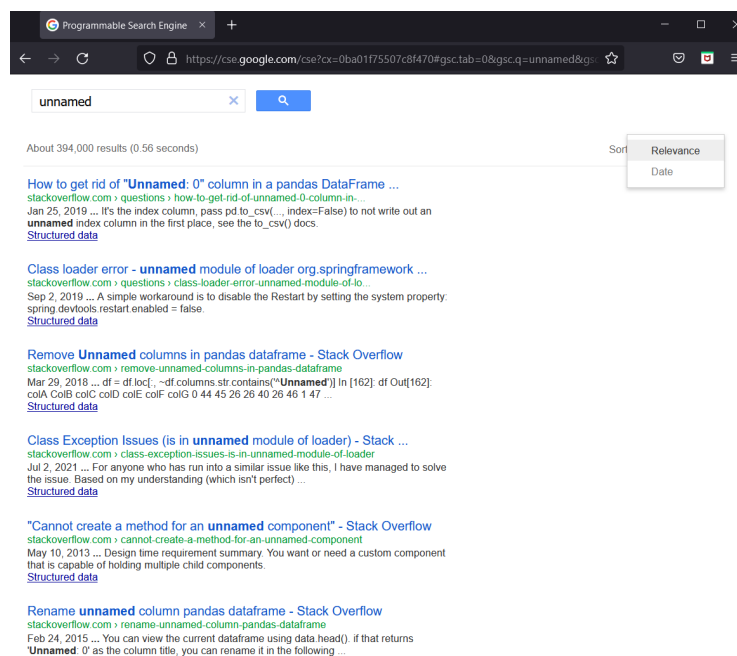
O ambiente de busca de informações na Web foi construído com a customização da ferramenta de busca da Google (*Google Custom Search Engine* — CSE)⁴. Isto permite programar o uso da caixa de pesquisa personalizada que mostre resultados relevantes rapidamente, o que é conhecido como *Programmable Google Custom Search*. O *Google Search Engine* foi escolhido como base devido a sua massiva preponderância de utilização no Brasil (96,75%) e no mundo (92,05%)⁵ (STATCOUNTER, 2021). A customização

⁴ <<https://cse.google.com/>>

⁵ <<https://gs.statcounter.com/search-engine-market-share>>

foi na restrição das fontes para o resultado da busca, mantendo as características básicas conhecidas do mecanismo. Nenhuma restrição ou condicionamento foi estabelecido para navegadores, dispositivos (celular, computador, etc.) ou demais recursos tecnológicos associados. A escolha destes elementos no estudo ficou à critério do próprio participante.

Figura 1 – Tela da Busca Customizada do *Google Search* para o Experimento.



Fonte: do Autor (2022)

Na figura 1, é exibida a aparência e os elementos da busca customizada⁶, a partir de um exemplo. Até mesmo os *Ads* (anúncios de propaganda e links patrocinados) estão presentes, antes dos resultados orgânicos, conferindo o mesmo comportamento da ferramenta de busca.

Os resultados da CSE foram restritos as principais CQA⁷, obtidas a partir de extensa pesquisa livre na Web, considerando o volume de acessos⁸ e a relevância obtida a partir de considerações de especialistas de *Search Engine Optimization* (SEO) disponíveis na Web⁹. Foi considerado o ano 2020 como referência e cruzando diversas listas que classificam as melhores CQA no ano^{10,11}. A partir dessas listas, a seleção das CQA se restringiu àquelas que tivessem o conteúdo majoritariamente textual, e que se enquadrassem como fontes para pesquisas sobre tópicos de Sistemas de Informação. Em complemento, foram

⁶ <https://cse.google.com/cse?cx=0ba01f75507c8f470#gsc.tab=0>

⁷ <https://www.widebinary.com/best-questions-answers-websites-list>

⁸ <https://www.techasoft.com/post/list-of-top-20-question-and-answer-websites-2020>

⁹ <https://www.expert-seo-training-institute.in/blog/question-answer-websites-list/>

¹⁰ <https://thenextweb.com/news/sites-people-opinions-questions-answered>

¹¹ <https://www.99techpost.com/free-question-and-answer-sites-list-best-qa-sites-2020/>

selecionadas duas CQA de destaque no cenário brasileiro que utilizam o idioma português nativo. Isto, em virtude da preponderância de CQA em idioma inglês e o risco disso ter impacto no experimento para os participantes falantes da língua portuguesa. Ao final desse processo de seleção, obtivemos as seguintes CQA:

- **Quora** (<<https://www.quora.com>>). Uma plataforma de perguntas mais genéricas. Recursos disponíveis: votação positiva ou negativa, seguir, agradecer e solicitar Q&A, perfis de membros e arquivamento. Com mais de 54 milhões de visitas únicas em 2021¹². Ocupa a 198.^a posição de site mais acessado no mundo¹³.
- **Reddit** (<<https://www.reddit.com>>). Plataforma social americana organizada em subfóruns. Principais recursos: postagem e votação contra e a favor. Está entre os sites mais acessados no mundo (21.^a posição)¹⁴.
- **Yahoo Answers** (<<https://answers.yahoo.com>>). Uma das mais antigas, conhecida e estudada CQA, descontinuada em 2021, mas com uma base de quase vinte anos que ainda é explorada em estudos acadêmicos (verificável via buscas no Google Scholar). Para perguntar precisava ter conta Yahoo. Tinha um sistema de pontuação e níveis de usuários associados.
- **Stack Overflow** (<<https://stackoverflow.com>>). Uma plataforma de perguntas mais direcionadas ao campo da programação. Recursos disponíveis: perfis de membros e votações. Identificada como a maior e mais confiável plataforma de Q&A. Com mais de 11 milhões de visitas únicas por dia¹⁵. Ocupa a 121.^a posição de site mais acessado no mundo¹⁶.
- **Answers** (<<https://www.answers.com>>). Formalmente conhecida como **WikiAnswers**, possui questões sobre estilo e aspectos da vida humana, de caráter mais genérico. Com mais de 109 milhões de visitas únicas em 2021^{12,17}. Ocupa a 9427.^a posição de site mais acessado no mundo.
- **Answersee** (<<https://www.answersee.com>>). Plataforma de perguntas sob diversos assuntos organizados em tópicos. Como diferencial tem um programa de recompensas, que permite ao membro ganhar dinheiro ao elaborar ou responder uma per-

¹² <https://aelieve.com/rankings/websites/category/reference/best-question-and-answer-sites/>

¹³ <https://mysite.info/analysis/quora.com/>

¹⁴ <https://mysite.info/analysis/reddit.com/>

¹⁵ <https://stackexchange.com/sites?view=list#traffic>

¹⁶ <https://mysite.info/analysis/stackoverflow.com/>

¹⁷ <https://mysite.info/analysis/answers.com/>

gunta, através do sistema de pontuação. Não está listada na base MySite.Info¹⁸ de sites mais acessados.

- **StackExchange** (<<https://stackexchange.com>>). É uma plataforma de Q&A que serve de guarda-chuva para 177 CQA técnicas ou de cunho científico, voltadas para compartilhamento de conhecimentos especializados. Ocupa a posição 493.^a de site mais acessado e tem uma forte presença nos principais sites de busca¹⁹.
- **pt.stackoverflow** (<<https://pt.stackoverflow.com>>). Uma das duas principais CQA em língua portuguesa. É uma versão do Stack Overflow para entusiastas de programação que usam o português ao invés da língua inglesa.
- **QASStack** (<<https://qastack.com.br>>). Uma das duas principais CQA em língua portuguesa. Tem 704 mil visitantes únicos e crescente uso no Brasil²⁰.

No quadro 1 apresentamos um comparativo das principais características entre as mais relevantes CQA selecionadas, que possuem grande volume de acesso e representatividade.

Tabela 1 – Comparativo de Características entre as principais CQA

Características	CQA			
	StackOverflow	Quora	Reddit	Yahoo Answers
Sistema de reputação do usuário	pontuações e <i>badges</i>	Não	<i>karma</i> por Prêmios e por participação	pontuações, níveis e estatísticas
Questões relacionadas	Sim	Sim	Sim	Sim
Atribuição de rótulos	<i>Tags</i>	<i>Topics</i>	Não	Não
Avaliação das perguntas	votação (+) ou (-)	votação (+) ou (-)	votação (+) ou (-)	Marcação de interessante
Avaliação dos comentários	votação (+) ou (-)	votação (+) ou (-)	votação (+) ou (-)	votação (+)
Avaliação das respostas	votação (+) ou (-)	votação (+) ou (-)	votação (+) ou (-)	votação (+)
Identificação da resposta aceita	Sim	Não		Sim, sem sistema de qualidade, filtragem ou revalidação.
Resumo	Nas dicas	Das respostas	Não	Não
Permite seguir?	Questões e tópicos	Questões, tópicos e pessoas	Compartilhamento das postagens	Lista privada de questões
Edição de Pergunta e/ou Respostas	Sujeito a aprovação. Usuários c/reputação 2K+ liberados.	Sim	Com restrições	Não
Estatísticas	mais perguntado, visualizações e última atividade	visualizações, última atividade, seguidores	Não	Não

Fonte: do Autor (2022)

3.2.2 Entrevistas semiestruturadas

A técnica *Intensive Interviewing* utilizada, permite captar não somente informações detalhadas, mas também observar as respostas não-verbais dos entrevistados, como as

¹⁸ <https://mysite.info/analysis/stackoverflow.com/>

¹⁹ <https://mysite.info/analysis/stackexchange.com/>

²⁰ <https://mysite.info/analysis/qastack.com.br/>

motivações e as experiências. Nesse modelo, a partir de cada resposta obtida, novas perguntas podem surgir, enriquecendo o processo de coleta de informações. A aplicação da técnica foi dividida em duas etapas, uma prévia e outra após as sessões. A entrevista imediatamente prévia contextualizou a relação do participante com processos de busca na Web e o uso de CQA, considerando sua percepção do processo e da utilidade da busca de informações em CQA. A entrevista imediatamente após as sessões tratou da obtenção do *feedback* do participante a respeito do experimento e da contribuição percebida considerando o processo de aprendizagem. Na tabela 2 apresentamos o conjunto de questões elaboradas para orientar as entrevistas.

Tabela 2 – Questões de orientação para as Entrevistas

Entrevista PRÉ
Utiliza sites de Comunidades de Perguntas e Respostas (CQA), como o <i>Stackoverflow</i> ou <i>Quora</i> para suas pesquisas de informação? Se sim? Com que finalidade e com qual frequência? Se não? Por quê? Exemplos de finalidade: uso genérico/social, uso técnico/trabalho/acadêmico?
Quando usa, realiza o acesso direto pelas próprias plataformas ou usa algum mecanismo de busca (Google Search p. ex.)?
Considera que sites e plataformas de CQA são úteis? Por exemplo, estarem estruturados em debate (respostas e comentários) a partir de perguntas “reais”?
Se usa mecanismos de buscas, acha mais relevante os resultados que tenham que características de base de dados? CQA / Fórum / blogs / site especializado no assunto / Wikipédia, etc.
Entrevista PÓS
Achou que a descaracterização visual do mecanismo de busca foi um problema? Por quê?
Achou que as bases/sites CQA listadas foram suficientes ou insuficientes? Sentiu falta de alguma CQA específica? Em relação a Fóruns ou outros sites tipo Blog? Vê alguma distinção
A concepção e formulação de sua pergunta foi impactada pela limitação visual da busca? Ou pela restrição dos resultados aos sites de CQA listados?
Considerando o site/plataforma/ambiente de CQA explorados neste experimento, bem como em relação aos temas propostos, alguma consideração ou percepção sobre o formato de apresentação e disposição das perguntas, respostas, comentários e demais elementos nas trilhas?
Na sua percepção, o que destacaria (positivo e/ou negativo) em relação as CQA no uso que normalmente faria: – Para buscar informações. – Para a obtenção de novos conhecimentos ou ampliação de conhecimentos. – Para uso como ferramenta ou plataforma de suporte à aprendizagem.
O que mais te influenciou nas CQA durante o processo de seleção de perguntas, respostas ou comentários para elaboração de tua “própria” solução para as atividades do experimento?
Durante as atividades do experimento, houve novos conhecimentos adquiridos ao longo dos conteúdos das trilhas analisadas na CQA? Por exemplo, alguma nova informação técnica, alguma forma de implementação técnica, novos conceitos, alguma nova abordagem?

Fonte: do Autor (2022)

3.2.3 Execução do Experimento

O experimento foi executado nos meses de fevereiro/21 e março/21, por reunião online na plataforma *Google Meet*, com todas as atividades e falas gravadas com o prévio consentimento do voluntário participante. O consentimento foi registrado no início da própria gravação para evidenciar a materialidade do mesmo. Ao longo da interação, as orientações iniciais das tarefas foram prestadas para cada sessão, e esclarecimentos adicionais foram fornecidos quando demandados pelo participante, garantindo que o participante explicitasse oralmente todas as suas ações e motivações. As intervenções do pesquisador

aconteceram sempre que foi observada ausência comportamental, informacional ou desalinhamento entre suas falas e ações durante a interação. Foi utilizado o método *Thinking Aloud Protocol* para a realização das sessões para a coleta dos relatos verbais dos participantes sobre as atividades (durante ou imediatamente após a execução). Antes e após as sessões experimentais foram realizadas as entrevistas semiestruturadas, baseadas na abordagem *Intense Interviewing*.

3.2.4 Caracterização dos Participantes.

Os participantes foram convidados a partir de uma lista de alunos de pós-graduação *stricto-sensu* de sistemas de informação, que estavam cursando ou recém-egressos, que tiveram algum contato com os pesquisadores. Tal familiaridade com a área de sistema de informação favorece a execução do experimento e a realização da entrevista, devido à relação de confiança entre participante-pesquisador, que facilita as técnicas de *Intensive Interviewing* e de *Thinking Aloud Protocol*.

Os participantes foram contatados por mensagem eletrônica ou contato telefônico. O regime de participação foi voluntário. O momento da participação coube majoritariamente ao participante definir em sua agenda. O convite foi feito com a apresentação da proposta do experimento qualitativo, disponível no Apêndice A, contendo: (i) objetivo, expectativas, descrição e técnicas a serem empregadas; (ii) descrição do ambiente com URL da sessão da reunião e do mecanismo CSE acompanhada da lista de CQA, orientações sobre Ads; (iii) considerações sobre a privacidade e aspectos éticos sobre a coleta e processamento dos dados coletados; (iv) lista de questões de orientação da entrevista; (v) descritivo das sessões a serem elaborada; e (vi) possível agenda com tempos estimados das atividades. Desta forma, tomamos as medidas necessárias para garantir que o participante tivesse condições adequadas para tomar sua decisão, de maneira consciente, quanto à participação no experimento.

Em virtude do impacto da pandemia Covid-19 na realização das atividades acadêmicas, somente foi possível completar a realização de 20 convites, após muitas tentativas de contatos (43 pós-graduandos), dentre um quantitativo de cerca de 170 alunos do curso de mestrado e de doutorado, discentes ou recém-egressos naquele período. Buscou-se a heterogeneidade do grupo de participantes, com representação étnica, gênero, nível acadêmico (mestrado, doutorado), situação de matrícula (discente, recém-egressos), perfil profissional (desenvolvedor, cientista de dados, *Web Designer* ou analista de UX, gerente e suporte de sistemas), e se possuem ou não bolsa de estudos.

Por se tratar de um estudo epistemológico interpretativista, com questões abertas

(*open-ended*), sobre comportamento humano num cenário simulado com situações reais afeitas ao grupo selecionado, o tamanho de amostra foi considerado satisfatório. Isto desde que as características heterogêneas da população estivessem representadas. Como se trata de uma população do estudo de tamanho finito, estabeleceu-se nível de confiança de 90%, condicionando margem de erro máxima de 12%, com desvio padrão de 0,9 (90%) relativo à familiaridade de conhecimentos e uniformidade do grupo em relação ao objeto da pesquisa. Com isso, o tamanho da amostra mínima seria de 16 participantes.

Ao final do processo voluntário de recrutamento incremental, ocorrido ao longo da realização do estudo para satisfazer critério de saturação teórica, obtivemos 17 participações (85% de aceite), que representam cerca de 10% do grupo de alunos do curso utilizado como população. Destacamos que a não participação ocorreu motivada pela falta de agenda envolvendo a carga de atividades curriculares do participante durante momento de pico pandêmico regional da Covid-19. Da mesma forma, houve 20% de reagendamentos e um prolongamento adicional de 1 mês. Ressaltamos que as perguntas e seus desdobramentos, bem como as sessões de atividades eram facultativas, dando a opção aos participantes de não responder, não realizar ou fazê-los parcialmente.

A complexidade envolvida na nossa pesquisa qualitativa, tal qual outros estudos citados na seção 2.5 que não envolviam questionários, não permite tratar dados de muitos participantes. A profundidade e riqueza de detalhes providos pela *Intensive Interviewing* e pela TAP trazem um volume de informações processadas (mais de 6 mil segmentos e de 26,5 horas de transcrição), as elaborações de diversos processos para transcrição (102 documentos de relatos e respostas), e muita energia despendida para codificação e análise (4317 segmentos de texto codificados com 290 códigos abertos agrupados em 10 categorias). Esse processo complexo não busca simplesmente quantificar o fenômeno em estudo, mas qualificá-lo a partir de aspectos cognitivos e subjetivos, de capturar as decisões dos participantes no processo de busca (Web + CQA) e suas reflexões sobre o processo de aprendizagem.

Adicionalmente, consideramos que com esse conjunto de participantes atingimos a saturação teórica, a partir da saturação da coleta de dados, quando nenhum novo elemento/conceito foi encontrado e não há acréscimo de novas informações para o desenvolvimento e categorização teórica (SAUNDERS et al., 2018), a partir do qual não altera a compreensão do fenômeno estudado. Seguimos os princípios de saturação de dados propostos por Francis et al. (2010), da seguinte forma: 1.º partindo de tamanho da amostra inicial de ao menos 10 (dez) participantes com a apropriada diversidade e qualificação dos participantes frente ao objeto de estudo; 2.º com critério de parada após 10 participantes, quando mais 3 (três) participações forem realizadas sem novos temas emergentes,

considerado assim como o ponto de saturação dos dados. O critério de parada é testado após cada participação sucessiva a partir da décima primeira. De qualquer forma, tornamos a enfrentar tal esse aspecto de saturação teórica da pesquisa nas seções de ameaça à validade e limitações da (seção 3.5).

A tabela 3 apresenta a heterogeneidade da amostra. Destacamos que cruzamentos não foram realizados para não comprometer o anonimato, dado que permitiria identificar o participante pelos seus pares no grupo e na instituição.

Tabela 3 – Perfil e Heterogeneidade dos Participantes

Classificação		Qtde	Pct
Gênero	Feminino	7	41,18 %
	Masculino	10	58,82 %
Etnia/Raça	Branca	10	58,82 %
	Preto ou Pardo	7	41,18 %
Faixa etária	Até 30 anos	3	17,65 %
	De 31 a 45 anos	10	58,82 %
	Acima de 45 anos	4	23,53 %
Perfil Acadêmico	Mestrando	9	52,94%
	Doutorando	8	47,06 %
Situação Acadêmica	Discente	12	70,59 %
	Recém-egresso	5	29,41 %
Bolsista Acadêmico	Sim	4	23,53 %
	Não	13	76,47 %
Tópico de Pesquisa	Sistemas Colaborativos	2	11,76 %
	Inteligência Artificial	3	17,65 %
	Interface Humano-Computador	4	23,53 %
	Processos e Negócios	2	11,76 %
	Análise de Redes Sociais	2	11,76 %
	Semântica na Web	4	23,53 %
Perfil Profissional	Programador/Desenvolvedor	8	47,06 %
	Web designer / analista UX	3	17,65 %
	Gerente / Supervisor	2	11,76 %
	Suporte / Operações	2	11,76 %
	Professor / Instrutor informática	2	11,76 %
Experiência Profissional	Até 5 anos	3	17,65 %
	Entre 6 e 10 anos	5	29,41 %
	Acima de 10 anos	9	52,94 %

Fonte: do Autor (2022)

3.3 Coleta e Tratamento de Dados.

Os dados não verbais coletados a partir da aplicação da *Intense Interviewing* e do *Thinking Aloud Protocol* foram transcritos em segmentos de texto padronizados conforme os marcadores disponíveis na tabela 4. Adicionalmente, parêntesis foram utilizados para incluir o que o ator estava falando ou descrevendo, que fora anteriormente apresentado, para correlacionar o objeto e contexto. Também foram utilizados colchetes para explicitar uma ação não verbalizada pelo ator, ou então, para designar elementos visuais e textuais

observado pelo pesquisador, tais como seleção de *links*, abertura de páginas, movimentos de *mouse*, texto da busca, etc.

Tabela 4 – Marcadores da Transcrição

Marcador	Descrição	Classificação
Ator	Quem está atuando ou falando.	Participante
		Pesquisador
Efeito	Qualificação do efeito ou ação majoritariamente realizada por um ator.	Ignorar: ruídos ou conversas amistosas.
		Comportamento: ação não verbal, por observação.
		Contexto: fala que contextualiza ou descreve atividades do experimento.
		Fala: ação verbal do Participante explicitando ações ou a partir da entrevista.
		Pergunta: ação verbal do Pesquisador, para arguir ou esclarecer algo com o Participante.
		Intervenção: ação verbal do Pesquisador para reestabelecer ações ou falas do Participante.
		Protocolo: ações ou falas dos atores, para alinhar o entendimento dos protocolos utilizados.
Estilo	O que se está tratando, qual parte das atividades ou entrevistas.	Entrevista Prévia
		Entrevista Pós
		Sessão Autointeresse
Anotação	A partir da fala ou ação do ator, ou anotação do pesquisador.	Comentário
		Diálogo

Fonte: do Autor (2022)

Os vídeos das reuniões individuais de interação de cada participante durante as atividades foram coletados através do *Google Drive* no formato MP4. Algumas gravações, devido a falhas de comunicação, ficaram fragmentadas e precisaram passar por um processo de *merge* através da ferramenta *MP4Joiner*. A partir desse tratamento inicial de uniformização do conjunto, as seguintes etapas foram realizadas:

1. Extração do Áudio do Vídeo. Uso de *script* em *Python 3* aplicando a ferramenta *ffmpeg-python* para a conversão do vídeo MP4 para áudio em formato WAV. Código-fonte disponível na Seção D.1 do Apêndice D.
2. Extração do Texto do Áudio. Uso de *script* em *Python 3* aplicando o mecanismo de transcrição *Google Speech-to-Text*. Tal mecanismo, por parametrização quebra o áudio em segmentos (os *chunks*) com o texto de transcrição correspondente. Os *chunks* idealmente representam as falas segmentadas em função de tempo (milissegundos) e intensidade (decibéis) do áudio no formato WAV em processamento dos falantes (pesquisador e participante). Código-fonte disponível na Seção D.2 do Apêndice D.
3. Formatação do Texto em legenda. Uso de *script* em *Python 3* convertendo os segmentos de textos e os segmentos de áudio identificados (os *chunks*) da etapa anterior, em um formato de apresentação de legenda comum associado à imagem (vídeo), combinando as marcações temporais com o vídeo original. Código-fonte disponível na Seção D.3 do Apêndice D.
4. Revisão da Legenda em relação ao Vídeo. Utilização do *software Aegisub32* pelo pesquisador para tratar o conteúdo legendado aplicando os marcadores apresentados na tabela 4, e revisando e ajustando conteúdos na grafia correta.

5. Transcrição rotulada para análise. Utilização de *script* em *Python 3* para transformar o texto legendado com os marcadores em transcrição estruturada e rotulada para análise posterior, aderente aos programas analisadores qualitativos, de modo genérico, que preveem um formato textual mais nos moldes de falas individualizadas ou de perguntas e respostas (entrevista). Código-fonte disponível na Seção D.4 do Apêndice D.

Para os dois primeiros participantes, tratados como caso piloto, validação e controle do processo de coleta e tratamento de dados, o processamento das etapas citadas foi único para todo o vídeo da interação. Para os demais casos, foi utilizada a função pesquisar da ferramenta *MP4Splitter*, a partir das marcações temporais de início e fim de cada parte do experimento obtidas do arquivo texto *sbv* gerado pelo *Google Meet*, para gerar MP4 e texto de legenda separado. Isto para adequar às abordagens de análise da fase posterior.

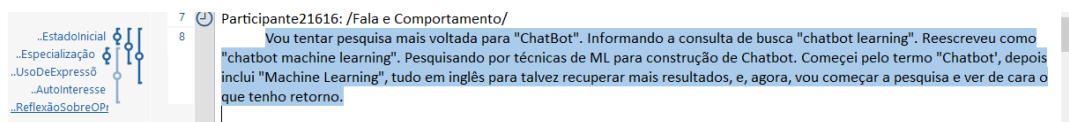
3.4 Codificação

Os códigos são os conceitos que denotam um fenômeno de interesse, representando a abstração de algo que tenha significado para o pesquisador, tais como um evento, objeto, ação ou alguma interação observada (CORBIN; STRAUSS, 2014). As categorias podem ser compreendidas como agrupamentos de conceitos em um nível mais alto de abstração. Na abordagem relativística da TFD (CORBIN; STRAUSS, 2014), o processo de codificação pode ser organizado em: (i) codificação aberta, (ii) axial e (iii) seletiva. Já na abordagem construtivista (CHARMAX, 2014), duas etapas principais de codificação são preconizadas: (i) codificação inicial, (ii) codificação focalizada. Adotamos no presente trabalho a abordagem construtivista, por tal abordagem melhor compreender a dinâmica da interação do participante e do pesquisador atuante, além de fazer uma correlação direta com o Guia de Métodos de Codificação (SALDAÑA, 2013) através dos dois ciclos ou fases previstos nesse guia.

Na codificação inicial (CHARMAX, 2014) (ou (*Open Coding*, termo mais conhecido oriundo da abordagem relativista), os dados são explorados minuciosamente, estabelecendo a distinção, análise, comparação, conceituação e a categorização dos dados. Nesse processo, os códigos podem ser reagrupados a partir da análise comparativa um-a-um, gerando as unidades de análise — categorias conceituais. Os códigos gerados podem referir-se a uma palavra, uma frase, um parágrafo/segmento de texto ou mesmo uma fala completa do participante. Ressalte-se a fragmentação dos dados e a microanálise para descobrir dimensões relevantes das categorias e suas subcategorias, e a relação entre es-

tas numa relação de causalidade. Nesse estágio, podem ser utilizados perguntas sensíveis (quais os elementos envolvidos no fenômeno caros ao participante), orientadores (guiadores da entrevista e mutáveis), teóricas (inerentes ao processo e conexões de conceitos), e estruturais (apoiam o desenvolvimento da teoria) (SANTOS et al., 2018). A codificação realimenta a própria codificação no caminho, refinando os dados e direcionando a investigação posterior através dos *insights* sobre conjunto deve ser analisado.

Figura 2 – Exemplo de Codificação Inicial



Fonte: do Autor (2022)

Na codificação focalizada (correspondente às Axial e Seletiva, mais conhecidas, da abordagem relativista), temos a classificação, seleção e sintetização de grandes conjuntos de dados. A reflexão associada nesse processo deve ser suportada pelas anotações (*memos*) teóricas, mitigando a subjetividade inerente ao processo. Os códigos elaborados nessa etapa são mais direcionados, seletivos e conceituais, para sintetizar e explicar segmentos maiores de dados (SANTOS et al., 2018). Podem ser utilizados os códigos ou categorias mais significativos e/ou frequentes identificados na codificação inicial, para uma melhor compreensão analítica dos dados. São examinados os relacionamentos entre os elementos codificados dos dados, considerando novos conceitos que surjam a partir de semelhanças e diferenças nessas relações, quer seja entre as categorias ou com suas propriedades relacionadas (CORBIN; STRAUSS, 2014). Isso leva à identificação do fenômeno ou categoria central da pesquisa. A categoria central representa o conceito organizador central mais potente analiticamente, e explica a maior parte da variação em um padrão de comportamento no fenômeno em estudo (GLASER, 1992 apud GASSON, 2004)[pág. 75].

A implementação da abordagem TFD segue o Guia de Codificação de Saldaña (2013) (SALDAÑA, 2013), onde são detalhados diversos métodos de codificação com objetivos e finalidades específicas para explorar os dados. Os dois ciclos (*First Cycle Coding* e *Second Cycle Coding*) dessa abordagem podem ser correlacionados às fases de codificação inicial e focalizada descritas por Charmax (2014), respectivamente. A vantagem desse uso está no formato de apresentação dos métodos, como um guia, e nas orientações dos processos de transição após cada ciclo, compreendendo técnicas para organizar os códigos e estratégias. Essas estratégias permitem focar e teorizar sobre os códigos e categorias. Os métodos do primeiro ciclo preveem a revisão e re-codificação, à medida que novos dados são processados. Os processos do segundo ciclo requerem habilidades analíticas

como classificação, priorização, integração, síntese, abstração, conceituação e construção de teoria.

Considerando as questões iniciais da pesquisa, na perspectiva de aprendizagem, quatro tópicos iniciais foram categorizados para tratar a percepção e relevância da CQA no processo de busca na Web, a qualidade e estrutura das trilhas dentro da CQA para o interesse de busca, e os aspectos cognitivos do processo de construção da solução após a análise do conteúdo da CQA dos resultados selecionados:

- Contextualização e Utilização de CQA.
- Processo de Busca na Web.
- Processamento analítico na CQA.
- Processo de reflexão e construção da solução.

Para cada uma dessas categorias iniciais foram explorados conceitos a partir da codificação dos dados, considerando as transcrições das entrevistas pré e pós sessões, bem como das próprias sessões.

Quanto a categorização do Processo de Busca na Web, consideramos a taxonomia *ESKiP* de estado de consulta, baseada no modelo *KiP* de Busca Exploratória, que apoia a representação da estrutura e comportamento da reformulação da consulta em sistemas de buscas (TIBAU; SIQUEIRA; NUNES, 2020). Esse modelo permite compreender as motivações da busca e o tratamento decisório com base nas informações recuperadas.

3.5 Limitações e Ameaças à Validade

Estabelecer a confiabilidade e considerar as limitações do estudo são fatores importantes para refletir com precisão a integridade do projeto de pesquisa (GLESNE; PESHKIN, 1992 apud KOLB, 2012). Todo estudo apresenta ameaças à validade (RUNESON et al., 2012), isto é, eventos que podem impactar ou limitar o resultado do estudo (WOHLIN et al., 2000). Cabe ao pesquisador a responsabilidade em tomar medidas de precaução para confirmar áreas de validade dentro de sua pesquisa (CORBIN; STRAUSS, 2014). A seguir discutimos as principais ameaças associadas a este estudo e suas formas de tratamento.

Quanto à validade de construção, que diz respeito a eventos que podem prejudicar a medição correta do estudo, podemos considerar como ameaça: (1) o uso de questões-guias inadequadamente elaboradas; (2) aplicação do TAP sem seguir os procedimentos

prescritos nele; e (3) uso incorreto ou inadequada de ferramentas (*Software*) para coleta, tratamento e análise dos dados. Para mitigá-las, foi efetuado um levantamento preliminar da literatura sobre aspectos terminológicos, fatores, classificações e frases de designação e definição associado ao tema em estudo. Além disso, toda a proposta, conforme Apêndice A, foi apresentada e validada por pares de pesquisadores²¹, que contribuíram para o seu aprimoramento, cerca de 30 dias antes da fase de execução. Ferramentas técnicas para apoiar os procedimentos da coleta à análise foram avaliados e selecionados, permitindo iniciar a fase de Preparação com segurança.

As ameaças à validade interna abordam a precisão dos dados, incorporando os procedimentos de triangulação, verificações de membros e envolvimento dos participantes (CRESWELL, 2014), e, diz respeito a eventos não controlados pelo pesquisador que podem produzir distorções no resultado. Consideramos no presente estudo: (1) a comunicação entre os participantes voluntários que já participaram e os que ainda não o realizaram, compartilhando informações sobre o experimento também poderia influenciar nos resultados; (2) os participantes voluntários podem se esforçar para que seu desempenho no estudo difira do que seria em outras condições, seja ou pelo fato de estarem sendo observados, ou por pertencerem à mesma comunidade de alunos de pós-graduação e temerem comparações; (3) indisponibilidade de tempo dos participantes voluntários necessário à execução de todas as atividades das sessões e entrevistas, principalmente em meio a um cenário pandêmico agravado da Covid-19 (Fev–Abr/2021); e (4) codificação aberta realizada desestruturadamente ou sem alinhamento à pesquisa. Para mitigar tais ameaças, algumas ações foram tomadas: (a) os convites e negociação de participação eram individuais e reforçado o sigilo da participação e privacidade; (b) participantes que poderiam ter mais proximidade foram agendados para data e horários mais próximos, reduzindo o intervalo que poderia ser utilizado para a comunicação lateral sobre o experimento; (c) cada participante recebeu um código na pesquisa que o anonimaliza e descaracteriza inclusive a ordem de participação; (d) a agenda foi montada com definição prioritária do próprio participante, permitindo remarcações, retomada em horários distintos no mesmo dia; (e) uso do diálogo ativo, intervenção, quando observado ocorrências sobre dificuldades associadas à simulação ou a duração, promovendo o bem-estar do participante para a continuidade ou de desistência.

Para ameaças à validade e à confiabilidade devido a problemas procedimentais de aplicação do TAP, realizou-se piloto com o participante voluntário de maior experiência no uso do protocolo em suas pesquisas, bem como participante em outros estudos. Isto permitiu ao pesquisador adequar seu comportamento e abordagens durante o expe-

²¹ Grupo *Semantics and Learning* (SaL) — <https://sal.uniriotec.br/>

rimento. Por exemplo, no início da interação com o participante e a cada nova sessão, a adoção da prática de efetuar orientações sobre a metodologia e protocolos utilizados no experimento. Isto porque algumas sessões poderiam ser longas e exigirem muito dos participantes. Assim, durante os procedimentos de coleta para mitigar riscos inerentes a aplicação *Think-Aloud Protocol*, o pesquisador assume o papel de observador atento e atuante, durante o desempenho das atividades com verbalização concomitante (conforme o *Concurrent Think-Aloud Protocol*) pelo participante, e também após as atividades com a verbalização de retrospectiva (conforme o *Retrospective Think-Aloud Protocol*), através da instigação em caso de insuficiência ou ausência de informações prestadas. Além disso, realizamos anotações correlacionando a verbalização em relação aos demais comportamentos não verbais verificáveis durante o experimento. Cabe também ao pesquisador manter a adequada aplicação do método, provendo a instrução e a preparação quando a observância dos procedimentos pelo participante, e a lembrança de tempos em tempos dos procedimentos pelo participante. Da mesma forma, o pesquisador deve garantir que haja fluidez da verbalização do pensamento em relação ao comportamento observável do participante.

Para ameaças à validade e à confiabilidade devido a problemas no uso da metodologia *Grounded Theory* e a técnica de comparação constantes, considerando o trabalho de (KOLB, 2012), podemos destacar os seguintes riscos: (1) uso de perspectiva mista de abordagens TFD durante as observações dos participantes.

Para mitigar esses riscos, foi utilizado a técnica de triangulação, usada para aumentar com precisão a fidelidade da interpretação dos dados usando vários métodos de coleta de dados (GLESNE; PESHKIN, 1992), tais como o principal *TAP*, o *Intensive Interviewing*, apoiados pelas anotações de observações do pesquisador. Também foi estabelecido foco no significado, ação e processo segundo a abordagem metodológica da TFD construtivista, conforme aconselha Charmax para manter a qualidade do estudo e permitir observar o critério de originalidade (CHARMAX, 2014 apud BERTHELTSEN; LINDHARDT; FREDERIKSEN, 2017), avaliada pela significância social da teoria.

Quanto as ameaças à validade externa, que prejudicam a generalização dos resultados do estudo, destacamos que o foco da pesquisa qualitativa é formar impressões e entendimentos únicos de eventos, em vez de generalizar as descobertas (CRESWELL, 2014). Não temos o objetivo nesse estudo para a generalização dos resultados para outros contextos. Os achados e constatações são circunscritos ao grupo de participantes qualificados e ao contexto específico, pois do contrário, seria necessário a execução com grupos mais diversificados e a realização de múltiplos estudos, considerando, por exemplo, regionalismos e níveis de escolaridades distintos. Entretanto, podemos citar como ameaças: (1)

impossibilidade de replicação ou aplicação do desenvolvimento da teoria fundamentada pode ser aplicado a outros casos; (2) representatividade limita dos participantes; (3) contexto organizacional único restrito a uma área acadêmica; e (4) população participante escolhida por conveniência do contato e da voluntariedade. A replicação ou aplicação em outros contextos é possível, uma vez que descrevemos os procedimentos gerais e específicos detalhadamente, bem como fornecemos nos Apêndices os recursos tecnológicos, projeto da pesquisa e a lista compilada da codificação aberta.

Quanto as ameaças à validade de conclusão, que prejudicam o estabelecimento de relacionamentos estatísticos, não se aplicam ao presente estudo, interpretativo com análise qualitativa, buscando indícios interpretativos sobre um determinado assunto e contexto. Embora tenhamos utilizados técnicas e métricas da estatística descritiva, bem como algumas análises quantitativas, ambas foram para reforçar a análise qualitativa.

Segundo Kolb (2012), outra grande ameaça à validade pode ser o próprio pesquisador. Duas ameaças à validade que comumente ocorrem em estudos qualitativos são o viés do pesquisador e a reatividade, o efeito que o pesquisador tem no cenário ou no estudo (BICKMAN; ROG, 2008). Para minimizar os efeitos do viés do pesquisador sobre o estudo, deve-se desenvolver a reflexividade. Ao ser reflexivo, o pesquisador deve incorporar a consciência contínua de refletir, examinar e explorar seu relacionamento em todas as etapas do processo de pesquisa. Ele deve manter uma constante reflexão ao longo da duração do estudo (BICKMAN; ROG, 2008), numa atitude de atendimento sistemático ao contexto de construção do conhecimento. Desta forma, para mitigar ameaças à validade dos indícios interpretativos alusivos à conclusão, aprimoramos cuidadosamente o exame do desenvolvimento das codificações, adicionando uma revisão a partir de cada uma das sessões de atividades com todos os participantes. Com isso, agregamos uma perspectiva focada nos objetivos de cada sessão, diferente da tradicional de processamento por participante. Em todos os casos foram aplicadas a técnica de comparação constante. Para minimizar os efeitos da reatividade do pesquisador e da inclusão dos preconceitos pessoais na análise, examinamos detidamente todos os casos negativos para aprofundar a compreensão dos participantes que estamos estudando. Os casos negativos se referem a dados que parecem estar distantes dos demais dados coletados e não coincidem com a teoria emergente. Ao investigar casos negativos, os pesquisadores podem obter informações valiosas a partir do menor dado que pode parecer insignificante em um momento, mas prova o contrário na análise final (TAYLOR; BOGDAN; DEVAULT, 2015 apud KOLB, 2012). Desta forma, sempre que um caso negativo surgia, investigávamos suas relações com demais casos e conceitos, além de estabelecer nova perspectiva de análise.

Consideramos ter atingido a saturação teórica da amostra, conforme exposto na seção

3.2.4, balizado pela composição de grupos amostrais representativos e com experiência relevante e diversa para a busca como processo de aprendizagem sobre CQA na perspectiva de aprendizagem informal. Constatamos pelo uso dessa estratégia (SANTOS et al., 2018) e a medida em que ao processar novas participações nenhum novos conceitos surgiram para preencher possíveis lacunas teóricas emergentes dos dados. Isto tomando-se o cuidado de manter a composição por participantes de perfis distintos e com estratificação representativa e diversificada conforme visto no 3, com diferentes níveis de experiência, formação e área de atuação, bem como caracterização demográfica distinta da amostra. Eles fornecem indicadores iniciais válidos sobre as questões analisadas. O conjunto de códigos, categorias e análises qualitativas e quantitativas descritivas podem ser consideradas atuais e representativo para a elaboração de uma interpretação para o fenômeno em estudo.

3.6 Considerações finais sobre a abordagem metodológica

Neste capítulo apresentamos a abordagem de pesquisa qualitativa, segundo o paradigma interpretativa/construtivista e com finalidade exploratória, através do método de pesquisa Teoria Fundamentada em Dados (TFD) Construtivista, que entende o papel ativo do pesquisador e prevê a compilação da literatura após a análise dos dados para fortalecer o potencial argumentativo da pesquisa. Utilizamos o *Think Aloud Protocol*(TAP) e a *Intensive Interviewing* como técnica de coleta de dados para obter um conjunto de dados mais rico a partir da interação de usuários com o ambiente customizado, possibilitando a exploração de novas informações e perspectivas associadas. Para a análise dos dados, utilizamos o *Constant Comparative Method*, e também, o Guia de Métodos de Codificação em dois ciclos de Saldaña (2013) para organizar a representação da codificação.

O processo de aplicação da metodologia envolveu a construção de ambiente para uso, sessões de atividades com o uso da TAP e sessões de entrevistas, seguindo os procedimentos previstos na TFD Construtivista. O ambiente foi uma customização do sistema de buscas *Google Search* restringindo os resultados às principais CQA levantadas da literatura e Internet. Foram três sessões com desafios práticos que envolviam de busca na Web e análise de trilhas CQA, elaboradas para ampliar gradualmente o grau de dificuldade das tarefas, uma sessão de análise comparativa de CQA para uma pergunta conceitual e de análises de trilhas 'QUE' e 'COMO' para suporte à aprendizagem específica. Também realizamos duas entrevistas, uma antes para diagnóstico da relação do participante com o uso de CQA, e outra ao final para obter o *feedback* da participação e perspectiva de ocorrência de aprendizagem durante as atividades. Finalmente tratamos as ameaças à validade

relativas aos aspectos metodológicos da presente pesquisa.

4. Codificação e Análise do Comportamento do Usuário no Processo de Busca Web sobre CQA

Para avaliar a CQA como fonte de informações factuais relevantes e base de conhecimentos especializados, a abordagem de análise empregada no presente trabalho, considerou o princípio do código inicial (CHARMAX, 2014), onde dados são codificados a partir de uma compreensão vaga dos tipos que podem ser relevantes, através da classificação dos seus elementos em temas ou categorias, na busca por padrões entre as categorias, tais como comunalidade, associação e causalidade implícita (GASSON, 2004). Utilizou-se da construção de guia de tópicos (LOWE, 1995 apud GASSON, 2004) para orientar a codificação inicial fundamentada nos elementos associados às questões iniciais da pesquisa:

- Qual a utilidade da CQA no provimento de informações relevantes e necessárias aos usuários para aprendizagem?
- Como a organização dos conteúdos compartilhados na CQA ajuda a promover aprendizagem dos usuários-aprendiz?
- Quais são as características sócio-técnicas que uma CQA deve ter para dar suporte à aprendizagem?
- Como integrar o uso de CQA com a busca como processo de aprendizagem na perspectiva de Aprendizagem Informal?

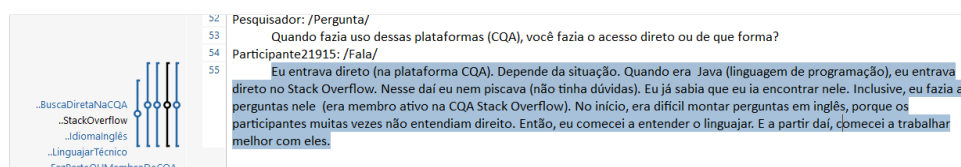
Para o entendimento dessas questões, foram propostas duas entrevistas, que tiveram como base as questões listadas na tabela 2, e quatro sessões de atividades relativas ao processo de busca na Web utilizando CQA como fonte:

- Entrevista Contextual (pré-atividades). Realizada imediatamente antes das sessões de busca. Tal entrevista focou na experiência dos participantes quanto ao uso de CQA como fonte de informação em um processo de busca na Web.
- Sessão #1 — Busca livre, de autointeresse. Proposta de atividades de busca de livre escolha do participante, manifestando o seu autointeresse em algum assunto relacionado a sistemas de informação. Nessa sessão de atividades, buscou-se analisar o processo natural de busca do participante em um cenário de familiaridade e/ou domínio do assunto da pergunta — a necessidade de informação.
- Sessão #2 — Pergunta Condicionada Direta. Proposta de atividades de busca associada a uma pergunta direta de implementação de solução técnica na área de Computação, associada a um assunto conhecido (série de Fibonacci). O aspecto de pergunta direta está relacionada à possibilidade de encontrar uma solução diretamente a partir dos resultados da busca, isto é, intenção de busca informacional direta fechada.
- Sessão #3 — Pergunta Condicionada Indireta Combinada. Proposta de múltiplas atividades de buscas, que precisem ser combinadas entre si para definir uma solução para a pergunta enunciada. O Contexto da pergunta está associado a análise de conjunto de dados. Aqui exploramos a intenção de busca informacional direta aberta.
- Sessão #4 — Análise Comparativa da organização e postagens de diferentes CQA em relação ao suporte à aprendizagem. Considerando três perguntas características existentes em alguma CQA, no padrão 'WH'. A primeira pergunta explora a análise comparativa entre as CQA, bem como o comportamento da intenção de busca informacional não direta. As demais tratam da aderência das postagens para suporte à aprendizagem do assunto envolvido nas postagens.
- Entrevista de *feedback* e de relato de aprendizagem (pós-atividades). Realizada imediatamente após as sessões de busca. A entrevista tratou da percepção dos participantes sobre as sessões realizadas, considerando a contribuição de CQA no contexto de Busca como um Processo de Aprendizagem (BPA). Também tratou do *feedback* do participante a respeito do ambiente tecnológico disponibilizado para as sessões práticas, e da contribuição dessas sessões para novos conhecimentos e o suporte para aprendizagem dos assuntos.

4.1 Contextualização da utilização de CQA no processo de busca segundo a experiência do participante

Para a contextualização do uso de CQA como fonte num processo de busca por informação na Web, utilizamos uma entrevista semiestruturada e com uso da técnica de *Intense Interviewing*, realizada imediatamente antes das sessões práticas de busca. Consideramos o relato extraído dessa entrevista, baseado nas questões-guia apresentadas na Tabela 2, como a fonte principal dessa codificação. Como pode ser observado na Figura 3, temos alguns elementos visuais¹ para delinear as codificações com o texto, bem como a separação das atuações dos envolvidos, pesquisador perguntando e o participante respondendo.

Figura 3 – Codificação Inicial — Utilização da CQA. Participante P21915

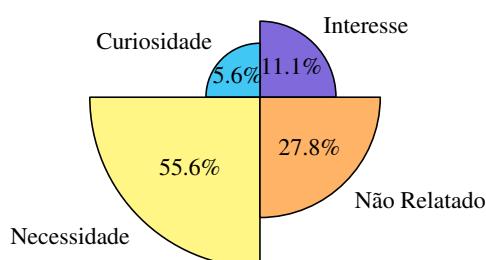


Fonte: Fonte: do Autor (2022)

A primeira pergunta, conforme Tabela 2, explorava a familiaridade e uso de CQA para as pesquisas de informações na Web, tentando identificar qual a finalidade e a frequência com as possíveis razões associadas.

Quanto à razão ou motivo para uso de CQA, obtivemos 13 respostas qualificáveis, onde apenas o participante P21818 se disse tanto motivado pela Necessidade, utilizando a CQA *Stack Overflow*, quanto pelo Interesse, utilizando a CQA *Reddit*. Para essa análise, diferenciamos o conceito de Necessidade (significado: o que não se pode evitar; inevitável), de Interesse (significado: o que é importante, útil ou vantajoso, moral, social ou materialmente) e de Curiosidade (significado: desejo intenso de ver, ouvir, conhecer, experimentar algo novo, original, desconhecido) a partir de seus significados. A distribuição percentual pode ser constatada na Figura 4.

Figura 4 – Razão ou Motivo para Uso de CQA



Fonte: do Autor (2022)

¹ Ferramenta MaxQDA, <<https://www.maxqda.com/>>

Na motivação por Necessidade, os principais relatos são sobre “busca de solução de problemas”: participantes P21610, P21715, P21818, P21915, P22218, P31718; “busca de informações”: P21818; “para resolver dúvidas”: P21710, P22423; “para trabalho” (de desenvolvedor): P21720, P30217. Por exemplo:

“Na minha profissão trabalho com programação, então quando estou fazendo algo novo ou usando algum componente, durante a programação, para resolver as dúvidas.” (Participante P21710).

“(...) preciso resolver um problema de forma rápida.” (Participante P21715).

“É mais para busca de informações mesmo.” (Participante P21818).

Na motivação por Interesse, quando não há nenhum agente externo de pressão, verificamos a ocorrência da “busca de informações”: P21818; e a “busca de conhecimento” sobre algo, como uma dúvida ou respostas diretas, ou compreensão de algoritmos: P22015. Destacamos alguns relatos a seguir:

“Recentemente, eu tenho usado *Reddit* para buscar informações sobre mercado financeiro... tem algumas comunidades lá que tenho interesse nesse sentido.” (Participante P21818).

“Busco por artigos. Busco por entendimento de determinadas dúvidas. Busco muitas vezes respostas de determinadas perguntas diretas. Busco eventualmente algoritmos. Esse tipo de coisa acontece dependendo do interesse.” (Participante P22015).

Na motivação por Curiosidade, parece se destacar a vontade (“gosto muito”) do indivíduo e o aumento do engajamento (“participar da comunidade”), levando a oportunidades de aprendizado, conforme observamos no relato a seguir:

“(...) uso o *Quora* mais para curiosidades. Por exemplo, eu gosto muito de participar da comunidade do *Quora* em espanhol. Então é mais para uso cultural e treinar a língua espanhola.” (Participante P21510).

Observamos que os relatos associados à Necessidade, ao citarem a CQA sempre mencionam a *Stack Overflow*, ao passo que para Interesse ou Curiosidade, temos a citação nominal da *Reddit* e da *Quora* respectivamente. Isto é condizente com a percepção que *Stack Overflow* é uma CQA concentrada em um campo específico (desenvolvimento e programação) enquanto *Reddit* e *Quora* fornecem uma plataforma geral para perguntas de todos os tipos².

² https://en.wikipedia.org/wiki/Social_information_seeking

Quanto à finalidade explícita para uso da CQA, destacamos que a maioria (88,23%) utiliza para fins técnicos, como trabalho profissional ou acadêmico. Desse conjunto, 60% relata uso para demandas pontuais ou específicas, enquanto 20% também relata uso para fins genéricos ou sociais. Apenas um participante relata o uso pontual ou específico, mas não destacou se é mais genérico ou mais técnico. Outro relata não usar CQA, ao menos conscientemente, conforme relato a seguir:

“Para falar a verdade, quando disse muito pouco, o certo era falar nunca. Vou direto no *Google (Search)* e faço a pergunta que quero, pesquiso por uma ‘string’.” (Participante P31718).

Na Figura 5, que apresenta a visualização das conexões entre os códigos por segmentos de textos marcados nos documentos de transcrição, destacamos que a Necessidade está intimamente ligada à demanda técnica para uso predominantemente diário, assim como o uso pontual está associada à demanda técnica para uso sazonal ou raro.

Figura 5 – Utilização da CQA — Razão x Finalidade x Frequência



Fonte: do Autor (2022)

O foco técnico dado para o uso de CQA parece estar realmente associado às atividades técnicas desempenhadas no dia-a-dia profissional ou acadêmico do participante, conforme os relatos:

“Desde que fui transferido da área de desenvolvimento para área de suporte (não uso mais tais CQA). Faz uns três anos.” (Participante P21915).

“Muito raro (agora). Eu usava o *Stack Overflow*, principalmente, na época que era programador. Agora não preciso.” (Participante P21610).

“Eu diria que não tem uma frequência muito certa não. É sazonal, talvez uma vez a cada dois meses, mas depende muito. Tem vezes que está mais ‘pegado’ (grande demanda), aí é uma vez por semana. Dependendo... não tem uma frequência muito certa.” (Participante P22519).

A segunda pergunta, conforme Tabela 2, explorou a forma pela qual o participante acessa a CQA e identificou se havia a percepção consciente que estava acessando uma CQA. Observamos que 95,1% dos relatos indicam que a usam sempre por meio de mecanismos de busca, notadamente o *Google Search*. Desse total, 6,25% declararam fazer uso eventual de busca direta na CQA através do recurso de navegação entre perguntas similares, denotando a irrelevância do recurso de busca fornecido pela plataforma.

“Eu geralmente uso pelo *Google (Search)*. Para entrar nessas plataformas CQA vou através do *Google (Search)*. Na plataforma, aí às vezes, olha a resposta e julgo que não é bem aquilo ali. Muitas vezes tem dúvidas similares, então navego na plataforma por *hiperlinks*. Vou clicando nos *hiperlinks* e chegando em outras páginas.” (Participante P21510).

“Raramente é direto na plataforma. Na maioria das vezes vou pelo *Google*, mas já esperando que ele vá me dar nos primeiros resultados respondidas dessas plataformas.” (Participante P21820).

“Eu usava pelo *Google Search* e acabava caindo na própria plataforma.” (Participante P21610).

“Diretamente não. Pesquiso no *Google (Search)* e normalmente quando tenho alguma dúvida técnica, ele me direciona para o *Stack Overflow*, o *Quora*.” (Participante P21616).

Interessante notar que somente 1 (um) participante disse não reconhecer ou ter familiaridade com CQA, isto é, distingui-las de outras bases nos resultados da busca:

“Na verdade, nem as conhecia (CQA listadas).” (Participante P22613).

Destacamos que apenas outro participante relatou que já utilizou o recurso de busca interno da plataforma CQA como rotina, provavelmente devido ao seu próprio processo de engajamento na comunidade CQA em questão, por ser membro e buscar ampliar sua interação:

“Eu entrava direto (na CQA). Depende da situação. Quando era Java (linguagem de programação), eu entrava direto no *Stack Overflow*. Nesse daí eu nem piscava (não tinha dúvidas). Eu já sabia que eu ia encontrar nele. Inclusive, eu fazia as perguntas nele. No início, era difícil montar perguntas em inglês, porque os participantes muitas vezes não entendiam direito. Então, eu comecei a entender o linguajar. A partir daí, comecei a trabalhar melhor com eles.” (Participante P21915).

Na perspectiva de reconhecimento da CQA e atuação como membro de alguma delas, quatro participantes (23,53%) se declararam membros de umas ou mais CQA: *Stack Overflow* com 2, *Quora* com 2, *Reddit* com 1 e no passado *Answer Yahoo!* com 1. Como preferências, os relatos citam o *Stack Overflow* (73,93%) para demandas técnicas e o *Quora* (17,65%) para as culturais.

Através da terceira pergunta, conforme Tabela 2, exploramos o quanto a estrutura e forma de apresentação do conteúdo impacta na utilidade e relevância da CQA face à necessidade de informação. Não foi possível identificar relatos sobre a perspectiva sentimental em relação à CQA, que não estejam associados à percepção de utilidade. Percebemos uma dicotomia quando a qualidade dessa organização e também ao conteúdo, considerando como uma força ou como uma fraqueza (restrição). Por isso exploramos melhor esse aspecto de estrutura na Seção 4.5, investigando tais elementos à luz da percepção prática quanto ao uso para aprender ou compartilhar conhecimento, Abaixo destacamos alguns relatos:

“No momento que se precisa resolver problemas, como está estruturada nas plataformas de CQA, fica muito melhor para consultar.” (Participante P21710).

“Acredito que elas são úteis. Não acho que são tão estruturadas assim para dar exatamente as respostas que estamos querendo. Percebo uma necessidade de fazer o filtro do filtro.” (Participante P22015).

“Era útil, mas às vezes, mais atrapalhava do que ajudava. Porque tinham um monte de respostas, onde a partir de um problema, cada uma apresentava uma solução, mas que não funcionava em algum aspecto. Aí tinha que passar por várias respostas até encontrar uma que atendesse. Valeria mais apenas, muitas vezes, ler da última resposta para a primeira. Você tinha que ler a pergunta e ir para o final para ver a última resposta.” (Participante P21610).

Com a quarta pergunta, presente na Tabela 2, exploramos a relevância da CQA em relação a outras fontes, tais como “fóruns” ou “blogs”, para a necessidade do usuário participante. Podemos observar que para algumas intenções de busca, as CQA são as fontes preferidas nos resultados retornados, particularmente devido à confiança da informação percebida em usos anteriores e a autoridade creditada àquela CQA, a *Stack Overflow*. Notadamente percebe-se a associação da CQA à resolução de problemas mais imediatos. Também se associa os sites de documentação ou do provedor tecnológico, ou de “blogs” a um entendimento mais aprofundado. Destacamos a seguir alguns relatos:

“Em geral, eu faço a busca pelo *Google (Search)*, e acesso o que vem. Aí sim, no *Google (Search)* vou dando preferência a algumas comunidades que parecem assim mais confiáveis, mais respostas. Principalmente o *Stack Overflow*, sendo a melhor para mim, para os problemas que eu procuro resolver.” (Participante P21915).

“Na verdade, eu uso muito o *Google (Search)*, por exemplo, e também o *DuckDuckGo* que tenho usado bastante. Nos resultados que eles trouxeram, geralmente procuro ou fórum, ou algo que tenha cara de tutorial. Uso essas duas coisas, eu não fico só preso a essas ferramentas de CQA não.” (Participante P31818).

Normalmente utilizo (CQA). Principalmente para pesquisa sobre programação. Por exemplo, o *Stack Overflow* é bem popular, então quando estou programando e tenho uma determinada dúvida ali, ou o código não funcionou, pego e jogo no Google e as respostas, por ser relacionado a código-fonte, normalmente as primeiras respostas são relacionados a esse tipo de site. (...) Normalmente uso o Google como ponto de partida. (Participante P21710).

“Vou pelo *Google (Search)* (...) acho que esse tipo de mecanismo como o *Stack Overflow* relevante, eu vou direto nele (a partir dos resultados da busca Web). Quando quero resolver um problema, eu vou direto nele. Mas quando quero entender, eu começo superficialmente pelo *Wikipedia* e depois parto para o site do fornecedor, da documentação, etc.” (Participante P21715).

Além disso, quando se trata de conhecimentos ou aprendizado, sobretudo quando se conhece muito pouco sobre o assunto, as CQA são preteridas e, as demais fontes são preferidas em uma combinação, a depender da estratégia de busca individual de cada participante:

“Quando quero aprender sobre um assunto que eu não conheço nada, que ele é complexo, que ele é longo, em geral, eu não procuro essas comunidades. Procuro uma biblioteca ou uma documentação mais ‘par-ruda’ porque eu não sei nada. Então terei que aprender muita coisa. (...) Se estiver buscando um conhecimento mais abrangente, mais amplo de algo que não sei nada, normalmente busco ou num ‘blog’, ou numa documentação, ou algo do tipo.” (Participante P22218).

Como já destacado, quando se trata de alinhar as CQA a um objetivo ou finalidade de busca por necessidade de informação (em inglês *‘information need’*), verificamos a tentativa de classificar tais comunidades segundo a aplicabilidade a um domínio ou a partir da interação existente na mesma, como destacamos no relato de exemplificação a seguir:

“O *Stack Overflow* é para o domínio de programação. O *Yahoo! Answers* é muito comum para perguntas mais genéricas, normalmente as respostas são de pessoas discutindo ou dando opiniões mais pessoais. O *Stack Exchange* para código também. O *Quora*, normalmente quando faço alguma busca de termo genérico também, às vezes encontro.” (Participante P21710).

Com base no relato a seguir, podemos vislumbrar o uso massivo da CQA como um mecanismo social que organiza as informações, com um processo de validade da construção do conhecimento através do consenso entre pares (os membros e especialistas da

própria comunidade), que leva à confiabilidade da informação e estabelece sua relevância sócio-técnica, e, que consequentemente, atrai mais acessos e usuários.

“Agrega, primeiro porque tem muita gente usando a plataforma. É necessário esses mecanismos sociais, na minha opinião, para organizar melhor as informações. Você entende até sobre a confiabilidade daquela informação. Por exemplo, no *Stack Overflow*, você tem uma página, está em frente a alguma pergunta, e você tem ali listando algumas respostas. Para cada resposta você tem aquele comentário sobre aquela resposta, que provavelmente fornecerá argumentos. Se aquela resposta é válida. Pergunta sobre aquela resposta (nos comentários). Se aquilo faz sentido, ou então, pessoas refutando aquela resposta, dizendo que está errado. Além disso, tem a votação. Procurará pela resposta mais votada que outras pessoas viram e votaram.” (Participante P21818).

4.2 Processo de Busca na Web com CQA para Atividade Livre (de autointeresse)

Para compreender o comportamento natural do usuário durante o processo de busca na Web, ainda que os resultados estivessem restritos às CQA, foi proposta uma pesquisa sobre algum tópico de interesse e de livre escolha do participante, desde que no contexto de Sistemas de Informação. A seguir apresentamos o enunciado proposto:

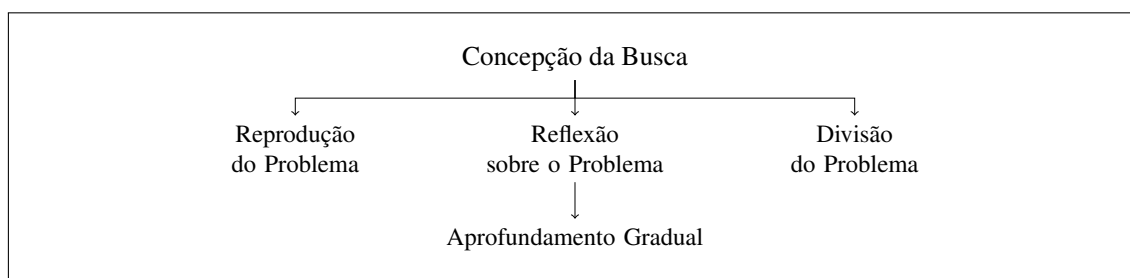
“Enunciado Sessão #1: formule uma pergunta qualquer no contexto de Sistemas de Informação, que: (i) seja parte do seu dia a dia profissional ou acadêmico, e (ii) retrate a busca por alguma informação ou implementação, considerando que seja aplicável a CQA.” (Pesquisador).

4.2.1 Análise do Comportamento no Processo de Busca

Utilizamos a taxonomia proposta por [Tibau, Siqueira e Nunes \(2020\)](#), como base inicial para a construção, onde os elementos da taxonomia foram considerados como códigos iniciais, e organizados nas seguintes subcategorias: concepção da busca, seleção dos termos e texto da busca, checagem dos resultados, e extração da informação. Tais elementos categorizados foram utilizados para elaboração da codificação a partir da análise dos dados coletados:

- **Concepção da busca.** Identificamos na codificação o comportamento de reprodução de um problema real experimentado ou atual do participante em 41,18% dos casos, e a reflexão sobre o problema no 58,82% restante. A divisão do problema em partes não foi observada nos relatos desta sessão, de acordo com todas as alternativas apresentada na Figura 6.

Figura 6 – Processo de Busca — Concepção da Busca.



Fonte: do Autor (2022)

Os casos observados de reflexão estão relacionados aos fundamentos da busca exploratória, quando não se está familiarizado com o que precisa ou não tem o domínio do assunto envolvido. Assim, o participante em seu relato ia construindo uma visão do que achava necessário e o que seria possível, explorando as possibilidades de resultados a partir da primeira construção de consulta, que geralmente era mais genérica ou simplificada. Eis alguns exemplos:

“Comecei pelo termo ‘*Chatbot*’, depois incluí ‘*Machine Learning*’, tudo em inglês para talvez recuperar mais resultados, e, agora, começarei a pesquisa e ver de cara o que tenho de retorno.” (Participante P21616).

“(…) Implementação do Teorema de Bayes no Python, diremos assim ‘*bayes theorem python*’. Começarei com uma bem simples para ver o que vem. (Participante P21510).

Os casos de reprodução de um problema estão relacionados à busca direta sobre um cenário ou situação muito específico. Envolvem geralmente uma demanda técnica, relativa a uma determinada informação ou a uma solução de implementação de código, de familiaridade de domínio do participante. A seguir apresentamos alguns relatos de exemplos:

“Procurarei ‘erro de integração oracle sqoop’ [Digitando e falando].” (Participante P21720).

“Uma busca que faço sempre e relacionada a erros que temos no Oracle. Trabalho muito com *plsql*. Então, por exemplo, hoje fiz uma *query* boba e eu já até sei a resposta desse erro, mas que no passado eu já realizei essa busca. Que é o código de erro do Oracle e a descrição do erro, e geralmente, as comunidades costumam trazer muitas respostas.” (Participante P22519).

- **Seleção dos termos e texto da busca.** Com base na estrutura de códigos da Figura 7, observamos um uso significativo de construção textual livre (33,3%). Embora, ao serem confrontados sobre os termos e textos utilizados, a maioria manifestasse que normalmente o fazem através de palavras-chave (ou termos-chave). Parece ser um processo inconsciente/involuntário de comportamento habitual na busca. O uso de estruturação de perguntas ocorreu em 14,3%, quando se tratava de desconhecimento

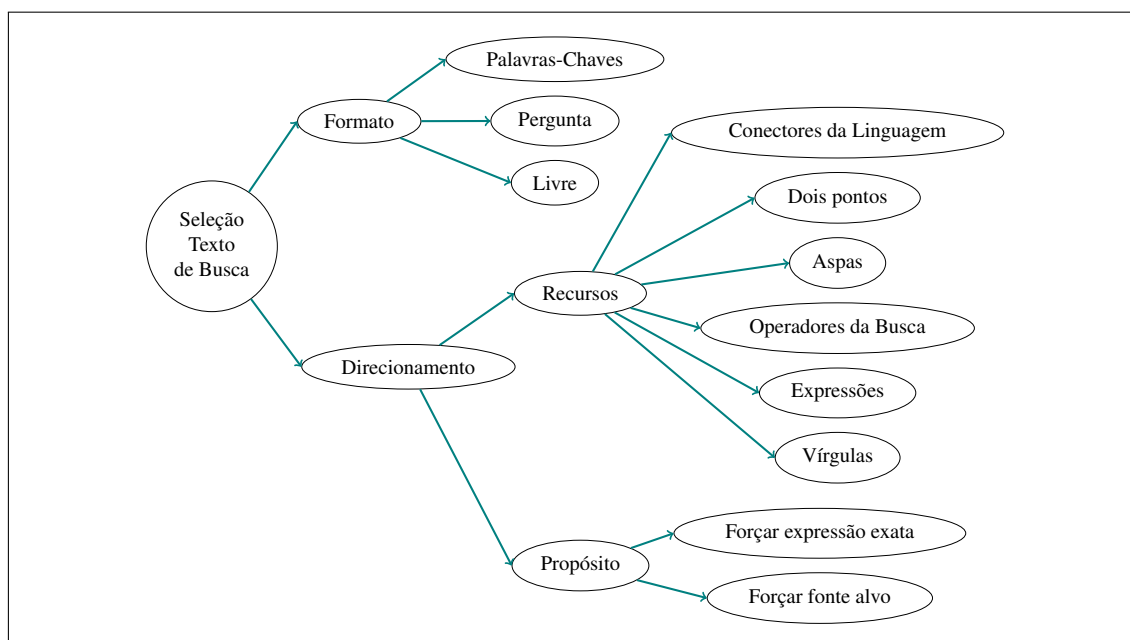
do contexto ou conceito envolvido, e o uso de palavras-chave aconteceu em 52,4% dos casos.

Destacamos que o ato de formular consulta está ligado a característica da intenção de busca do usuário. Por isso, o formato utilizado pode variar com o mesmo participante, como podemos observar nos relatos:

“Às vezes coloco pergunta mesmo. “O quê que é isso?”. Se é algum assunto que eu não conheço nada. Às vezes não, às vezes coloco tópico mesmo, tentando restringir.” (Participante P31718).

“Eu geralmente uso palavras-chave, a não ser que queira fazer uma pergunta muito específica. Do tipo quero resposta para essa pergunta, o que é tal coisa. Às vezes faço isso, e geralmente e ele (mecanismo de busca) traz respostas relacionadas.” (Participante P22015).

Figura 7 – Busca Livre — Seleção dos termos e texto da Busca.



Fonte: do Autor (2022)

Para a necessidade de informação de autointeresse dos participantes, verificamos que 35,29% deles não tentaram nenhuma forma de direcionamento ao elaborar o texto de busca. Os casos de tentativa de direcionamento do mecanismo de busca envolviam ou forçar a fonte alvo (emprego do termo “*stack*”) ou garantir que o termo/texto exato (emprego de texto entre aspas ou uso de dois pontos) esteja presente no resultado. Neste último caso, o participante parece ter domínio do contexto envolvido. O uso do termo “*stack*”, relacionado a CQA *Stack Overflow*, está associado à confiança dos participantes na credibilidade dessa plataforma e seu conteúdo em função do hábito de uso. Eis alguns relatos de exemplo:

“Coloco como frase. Apenas quando tenho um erro do Oracle que possui um número de erro. Ai boto “Oracle” mais dois pontos e o número do erro na busca.” (Participante P21720).

“[Digitando ORA-01722: “invalid number”].” (Participante P22519).

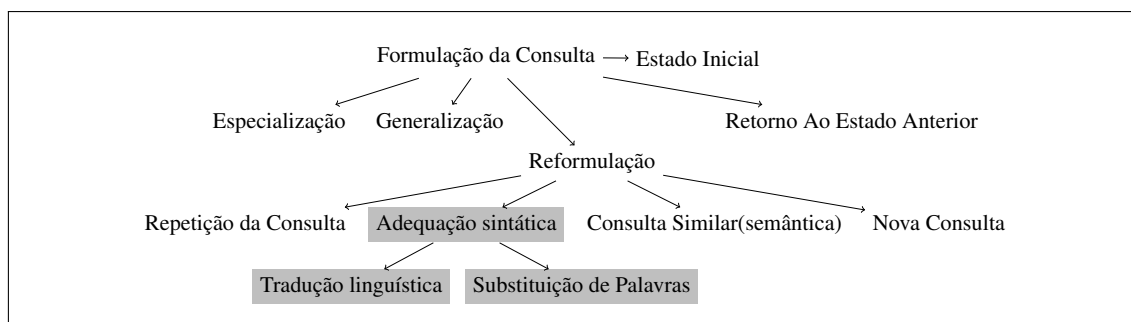
“Eu voltaria para o *Google (Search)* e refaria a minha busca aqui. Em alguns casos, eu ainda escrevo aqui o ‘stack’. [Modificou a *query* para ‘round number python stack’], assim, para garantir que ele vá lá no *Stack Overflow*, mas não busco por dentro do *Stack Overflow*.” (Participante P21715).

Os recursos para tentar direcionar o mecanismo de busca mais utilizados foram: 66,7% expressões, 26,7% conectores da linguagem/idioma (do inglês: ‘with’, ‘to’, ‘in’, ‘and’; do português: ‘de’, ‘que’), e 6,7% aspas. O uso desses recursos combinados foram observados em alguns poucos casos, 17,65% com expressões e conectores, e 5,88% com expressões e aspas.

- **Formulação da consulta.** Na figura 8 apresentamos os elementos codificados a partir do *ESKiP*, estendido (em destaque cinza) após a codificação dos relatos. As consultas similares observadas tiveram como modificação a troca da linguagem de programação, ou seja, a mudança de domínio ([STACKOVERFLOW, 2020](#); [NIVALA et al., 2020](#)). A substituição de palavras ocorreu pela troca do verbo empregado ou da retirada de conectores do idioma. Diferenciamos a substituição de palavras de consultas similares, porque a primeira trata-se de ajustes sintáticos na consulta formulada, enquanto a segunda trata-se de ajustes semânticos. A substituição de palavras tem a ver também com a correção dos termos que estejam inadequados na linguagem ou idioma. A especialização ocorreu com a adição de novos termos, para restringir o resultado da busca. Os participantes que usam o idioma nativo (PT-BR), continuam a fazê-lo mesmo sabendo que a maioria das bases está no idioma inglês (EN), e reformulam através da tradução dos termos entre os idiomas (PT-BR → EN) para garantir que o resultado contemplasse mais itens.

Nessa extensão do *ESKiP*, destacamos a separação da abordagem sintática da semântica pelo participante quando trabalha a reformulação de sua consulta, resultando na ‘Tradução Linguística’ ou na ‘Substituição’ (ou correção) de palavras.

Figura 8 – Processo de Busca: elementos da formulação da consulta, estendido do *ESKiP*

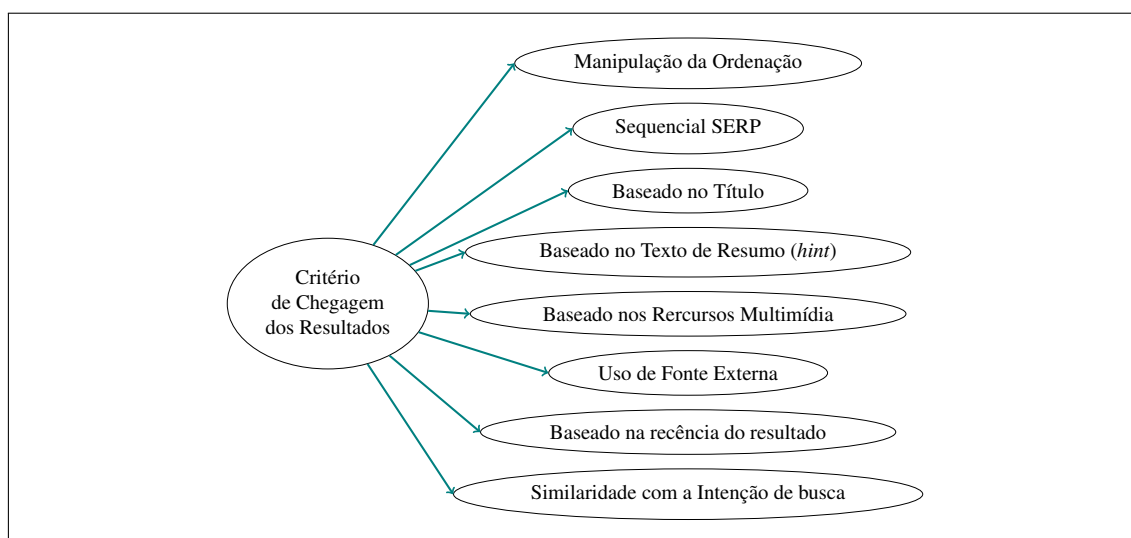


Fonte: do Autor (2022)

Conforme Tabela 32, disponível no Apêndice B, foram realizadas 30 (trinta) consultas nessa sessão, sendo 13 (treze) consultas adicionais ao 'Estado Inicial' (primeira consulta).

- **Checagem dos resultados.** Na Figura 9 apresentamos os códigos gerados a partir dos relatos. O uso de fonte externa, no caso *Wikipedia*, ocorreu uma única vez, devido à incerteza do participante em relação ao conceito associado à sua intenção de busca.

Figura 9 – Processo de Busca — Critérios de checagem utilizados.



Fonte: do Autor (2022)

Conforme figura 10 com casos combinados em diversos critérios, partindo da leitura sequencial dos resultados de busca (58,8%) na página de resultados do mecanismo de busca (SERP), alinhado à análise da fonte (47,1%), título (58,8%) e texto de resumo (47,1%). Assim, no que foi relatado e observado do comportamento dos participantes, há um respeito à ordem dos resultados na SERP, à caracterização do resultado principalmente pelo título e, secundariamente, pelo texto de resumo, assim como à origem da informação.

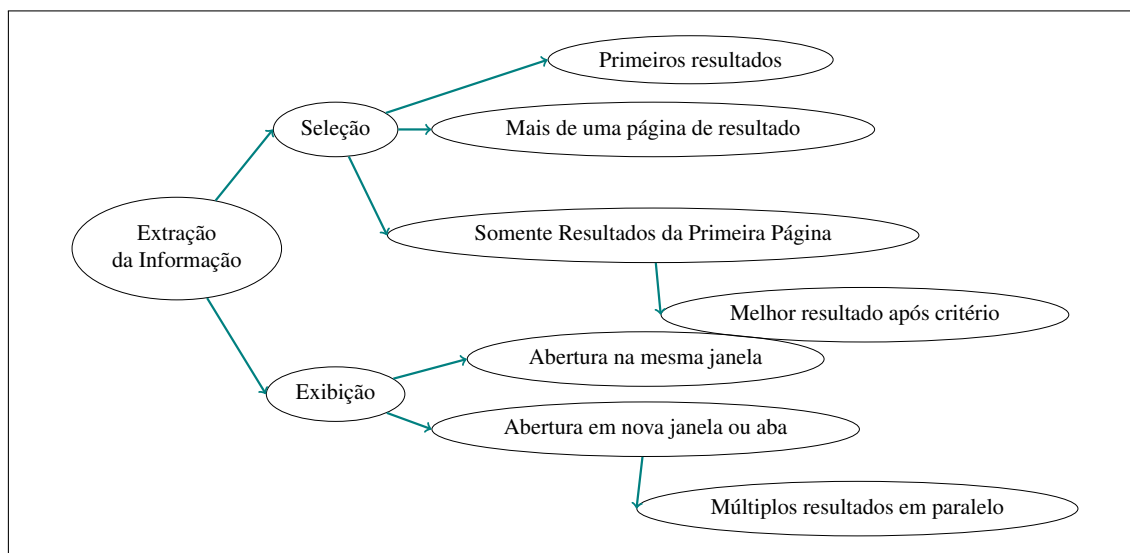
Figura 10 – Evidências de critérios de checagem utilizados.

. BaseadoNoTítulo "Lendo os títulos dos resultados da busca. E descartando em função do que sugere o título." "Eu vejo primeiro pelo título, e se ele me interessa, aí eu tento ver se nessas poucas linhas (resumo) tem algo interessante." "Normalmente o título é o que mais me influencia." "[Deu olha olhada nos títulos dos resultados]" "Eu procuro olhar o título para ver se aparece as palavras que eu acho mais relacionados ao tema. Porque essa ferramenta de busca, geralmente ... elas não necessariamente trazem um resultado com todas as palavras." . SequencialSEO "O segundo é o "...classifier". O terceiro é "Understanding..." Eu quero algo bem simples." "[Lendo um a um na sequência os resultados]. [Selecionou o primeiro]" "Eu costumo olhar mais resultados. Eu vou clicar no segundo link (dos resultados)." "No segundo link, o Resultado está mais próximo do que acontece comigo, na realidade." "[Selecionou o primeiro resultado "Devo logo procurar um primeiro emprego ...]" "[Selecionando o primeiro resultado "Best Calendar/Blind ...archers?"]" "O volume de informação que retorna é muito grande. O que te interessa é uma parte específica. Então eu vou dar uma olhada rápida, ver mais ou menos o que tem haver comigo, aí vou entrar naquele primeiro." "Na verdade IA tá puxando muito esse tema. Então vou pular o primeiro (resultado)___, porque não vi nada falando de SI. O segundo, também não___. O terceiro ___, talvez me interesse, porque está falando de técnicas, pode estar relacionado a alguma linguagem de programação ou ferramenta."	. BaseadoNaFonte "O primeiro veio ___ do Stackoverflow. Como veio Stackoverflow, e é uma comunidade que estou acostumado, então eu cliço nela" "Foi o primeiro resultado porque a base é a Stackoverflow e estou acostumado com ela." "Geralmente, eu vejo o primeiro resultado, que é o stackoverflow, já é um site que eu conheço, então já vou clicar no primeiro resultado." "eu dou uma olhada na fonte de onde está vindo isso. Eu valorizo o StackOverflow. O título e algum comentário (resumo que aparece no resultado), eu tento entender se aquilo que está no comentário tem haver com o que eu estava pensando. Eu olho o resumo que aparece aqui. Eu sempre olho, não só o título, mas a origem, de onde está vindo (Stackoverflow, Yahoo...), se é do próprio site do fornecedor/sistema." . BaseadoNoTextoResumo "Estou vendo aqui o resumo que já dá uma explicação. Esse primeiro aqui pode me dar a resposta que eu quero, então vou clicar nele." "aí eu tento ver se nessas poucas linhas (resumo) tem algo interessante." "Eu vou muito pelo título. Só se o texto, as duas linhas apresentadas (resumo), mudarem a minha idéia, mas é muito difícil." . UsoDeFonteExterna "[Foi na Wikipédia para validar os termos e conceitos, revisando o entendimento]."
--	---

Fonte: do Autor (2022)

- **Extração da informação.** Na Figura 11 apresentamos os códigos gerados a partir dos relatos. Para a extração da informação sempre foi utilizado a primeira página de resultados.

Figura 11 – Processo de Busca — Extração da Informação: Seleção e Exibição.



Fonte: do Autor (2022)

Contudo, nem sempre se fez o processamento de todos os resultados dessa página, restringindo-se aos primeiros resultados (58,82%), que eram visíveis em tela, onde 90% desses casos o participante chegou ao melhor resultado após aplicação dos critérios de checagem dos resultados, com 52,94% do total de ocorrências. Isto se deve

nessa sessão de autointeresse, à prevalência de intenções para resolver um problema pontual, com poucos casos de exploração de informações ou conhecimentos.

“Normalmente abro todos os *links* interessantes da primeira página de resultados.” (Participante P21616).

“Costumo olhar ali os primeiros resultados (da busca).” (Participante P21710).

“Então pularei o primeiro (resultado), porque não vi nada falando de SI. O segundo, também não. O terceiro talvez me interesse, porque está falando de técnicas, pode estar relacionado a alguma linguagem de programação ou ferramenta.” (Participante P31818).

O comportamento para exibição da informação extraída da busca foi majoritariamente a abertura em uma nova janela ou aba (66,66%), enquanto apenas 33,34% conduziram na mesma aba. Vale destacar que 2 (dois) participantes alternaram entre esses dois métodos de exibição. Mesmo empregando a abertura em nova aba, a maioria o conduziu de forma sequencial, isto é, um resultado sendo analisado por vez. Tivemos 2 (dois) casos de participantes que abriram múltiplos resultados em paralelo, e que declaradamente o fizeram por estarem explorando os conteúdos em busca de novas informações de seu interesse. Convém destacar que a abertura na mesma janela foram comportamentos naturais do participante, e que houve uma minoria (2 casos) que o realizaram por descuido, conforme o próprio relato.

“Abro em duas janelas (em outras janelas).[Abriu o primeiro resultado e o terceiro resultado em duas novas janelas, ambas trilhas do *Stack Overflow*].” (Participante P21710).

“O que farei é o seguinte. Clico no botão *BACK* e ao voltar, dou uma checada em todos os resultados da busca.” (Participante P21510).

No relato a seguir podemos verificar que os primeiros resultados e as primeiras páginas são as mais acessadas pelos participantes, e a partir daí, optam em refazer a consulta do que tentar novos conjuntos de resultados em outras páginas.

“Prefiro refazer a *query* do que descer até um terceiro ou quarto nível (página de resultado) do *Google (Search)*, eu não faço isso.” (Participante P30217).

Uma vez codificado em um documento por participante a experiência de cada um, sob a mesma base de códigos apresentados acima, estabelecemos a correlação do comportamento entre os participantes através da análise de similaridade para cada um, que verifica a semelhança ou a diferença em função da frequência de código, onde temos 0 (zero) para nenhuma e 1 (um) para total similaridade. Na Tabela 5 verificamos que o máximo alcançado foi 0,71 entre os participantes P21610 e P21720, enquanto o mínimo alcançado foi 0,00 entre P21710 e P21915.

Tabela 5 – Similaridade Comportamental entre os Participantes no Processo de Busca (Atividade Livre).

Participantes	P21510	P21610	P21616	P21710	P21715	P21720	P21818	P21820	P21915	P22015	P22218	P22423	P22519	P22613	P30217	P31718	P31818
P21510																	
P21610	0,32																
P21616	0,44	0,35															
P21710	0,33	0,21	0,38														
P21715	0,50	0,50	0,47	0,23													
P21720	0,29	0,71	0,47	0,19	0,35												
P21818	0,47	0,29	0,53	0,31	0,40	0,41											
P21820	0,44	0,33	0,24	0,36	0,19	0,22	0,33										
P21915	0,39	0,47	0,28	0,00	0,50	0,41	0,29	0,11									
P22015	0,41	0,31	0,47	0,33	0,33	0,28	0,62	0,58	0,24								
P22218	0,41	0,31	0,38	0,07	0,54	0,35	0,50	0,12	0,50	0,33							
P22423	0,29	0,19	0,11	0,27	0,29	0,17	0,19	0,21	0,12	0,13	0,29						
P22519	0,28	0,54	0,17	0,15	0,27	0,38	0,18	0,38	0,33	0,27	0,27	0,21					
P22613	0,44	0,44	0,33	0,38	0,22	0,39	0,35	0,50	0,15	0,38	0,22	0,25	0,40				
P30217	0,53	0,53	0,50	0,38	0,38	0,47	0,44	0,40	0,35	0,47	0,29	0,11	0,40	0,50			
P31718	0,41	0,31	0,29	0,45	0,25	0,28	0,50	0,58	0,17	0,54	0,25	0,29	0,36	0,57	0,47		
P31818	0,33	0,40	0,47	0,33	0,33	0,35	0,62	0,36	0,17	0,54	0,33	0,13	0,19	0,47	0,38	0,43	

Fonte: do Autor (2022)

Aplicando estatística descritiva sobre essa matriz de similaridade entre os participantes, obtivemos a Tabela 6

Tabela 6 – Sumarização do Comportamento dos Participantes no Processo de Busca (Atividade Livre).

Métrica	Valor
Desvio Padrão	0,130590435819573
Variância	0,017053861927546
Mínimo	0,000
1º Quartil	0,270
Mediana	0,350
3º Quartil	0,440
Máximo	0,710

Fonte: do Autor (2022)

4.2.2 Análise do Comportamento no Processo analítico da trilha CQA

Dado que o participante acessa a página com a trilha da CQA que foi selecionada no processo de busca, nesta seção traçamos o comportamento dentro da CQA. Avaliamos quais elementos são essenciais para a tomada de decisão quanto a aceitação ou não da postagem como atendendo a sua necessidade de informação. Assim como, qual o processamento analítico na CQA, isto é, qual a forma de leitura, quais são os elementos que são focos de atenção e que evidências são consideradas para a tomada de decisão.

A tomada de decisão está principalmente associada à resposta, isto é, o quão a resposta está elucidando (qualidade e alinhamento) a necessidade de informação, que em muitos casos tem o objetivo de implementação de algo. O fato de nas plataformas de CQA existir uma resposta validada como melhor resposta e também haver o consenso entre os pares, influi nessa percepção.

“Vou muito mais pela resposta. Por isso que me desanima ver poucas respostas. Ter uma resposta destacada como a melhor para mim, é importante.” (Participante P21616)

A navegação por *hiperlinks*, externos ou não à CQA, está associada a necessidade de informações ou esclarecimentos adicionais, orientadas pelo desejo e não pela imposição de ter que acessar aquele *link* (por exemplo, por pouca informação prestada). Cinco participantes manifestaram interesse em acessar três *links* externos e dois internos (para outra postagem na CQA), mas sempre como complementação a que se estava buscando na própria trilha. A seguir alguns exemplos de relatos:

“Tem um *link* para ver mais, é um site que conheço. Cliquei nele, mas sairá da CQA.” (Participante P21510).

“Entro nos *links*, mas só se for algum *link* que explica algo relevante sobre aquele assunto na totalidade. Eu não entro no *link* para saber o que é a pandemia, por exemplo.” (Participante P22015).

“Mas, muitas vezes, isso não é suficiente, ele resolve alguns problemas, mas traz outros. . . deixa alguns detalhes em aberto. Então você precisa fazer uma pesquisa mais abrangente, nesses casos vou nos *links* e acabo futucando.” (Participante P22218).

“Legal, tem indicação de *links* de códigos, de *papers*. Nesse momento aqui prefiro ler. Prefiro ver até o final comentário dele. Eu não clicaria em nenhum desses links nesse momento.” (Participante P31818).

Para o processamento cognitivo da postagem, alguns participantes optaram por traduzir para o idioma português para tornar o entendimento mais ágil, mesmo no caso de terem efetuados buscas já em inglês. Isto remete ao esforço cognitivo que o participante teria que fazer para entender a postagem, por não ser nativo no idioma. Para reduzir tal esforço, foi utilizado o recurso do Navegador Web para a tradução para o idioma nativo do participante. Contudo, a maioria manteve no idioma inglês os resultados originados nesse idioma.

“Eu geralmente, quando faço esse tipo de busca, eu faço em inglês. Porque eu já espero, e tenho conhecimento, que estas ferramentas(mecanismos de busca) são mais utilizadas nesse assunto utilizando o idioma inglês e, aí eu já costumo colocar em inglês.” (Participante P21820).

“Em geral. escrevo em inglês, porque tem muito mais resultado.” (Participante P22218).

Quanto ao objetivo da análise da trilha, observamos que 52,94% estavam atrás de alguma solução de implementação, 11,76% estavam em busca de conhecimentos sobre algo, e 35,30% estavam querendo informação para uso eminente.

“Mas quando quero entender um assunto. Quero saber o que está se falando, geralmente, sobre um determinado assunto. Atuo desta forma, geralmente passo longe desses *Ads*, vou para aquilo que realmente é aparentemente relevante, que o mundo está falando.” (Participante P22015).

“Você vê, por exemplo, a pessoa colocou uma dúvida parecida com a minha necessidade.” (Participante P21915).

Quanto ao comportamento da leitura da trilha na CQA, observamos conforme a Tabela 7 um leve prevalência da leitura seletiva, quando o participante vai direto para algum elemento/segmento da trilha, principalmente para a primeira resposta ou a resposta aceita, que no caso de algumas CQA se coincidem (a resposta aceita é a melhor resposta, que teve mais pontuação). Bem como, no processo de leitura sequencial, a leitura Top-Down, quando o participante começa pela pergunta e vai descendo pelas da análise até atingir um ponto, uma resposta, que faça ele desistir ou dar-se por satisfeito.

Tabela 7 – Estratégia de leitura da trilha CQA (Atividade Livre)

Leitura	Quantitativo	Percentual
Inversa(Pergunta+inversaoRespostas)	0	0%
Transversal	3	10%
Seletiva	9	30%
Sequencial	18	60%
....TopDown	9	30%
....EmProfundidade	3	10%

Fonte: do Autor (2022)

A leitura em profundidade está associada à necessidade de entendimento sobre o assunto e compreensão do contexto levando a mais esforço cognitivo, quer seja por interesse ou baixo conhecimento a respeito. Vejamos alguns relatos:

“[Lendo detalhadamente as postagens]. Elaborei uma pergunta bem genérica, as respostas estão assim também, mas aqui fala também em técnicas. (...) Está indo ao encontro do que vi em artigos. (...) Estou lendo as respostas e vendo se dá para tirar algo. Mas é que as palavras que usei são bem genéricas. Normalmente pesquiso coisas mais específicas. (...) Julgo que sim, serve.” (Participante P21616).

“Quando eu faço uma pesquisa que em CQA. Eu normalmente vejo tudo. Por isso que eu lhe disse que passei anos fazendo isso para tentar resolver uns problemas em java, então eu não lia só a resposta, eu lia tudo. Porque se eu não lesse tudo, eu não ia entender como a pessoa conseguiu chegar.” (Participante P21915).

Já a leitura transversal está associada à tentativa de capturar alguma informação que se sobressaia na trilha, e então, partir daquele ponto e fazer a leitura. Geralmente associado à busca visual de termos específicos esperados pelo usuário leitor.

“Normalmente, quando você entra em algo, você não sai lendo tudo. Você vai escanear (leitura transversal) e ver mais ou menos o que te interessa.” (Participante P31718).

Quanto ao foco, destacamos que a primeira resposta demandou mais atenção inicial dos participantes (94,12%), que muitas vezes é entendida como a melhor resposta (maior

votação) ou a resposta aceita. A identificação dos elementos da CQA que respaldam isso influi no foco natural inicial. O único comportamento não alinhado a isso, foi do participante que declarou não conhecer ou usar tais plataformas.

“Para essa pergunta achei a resposta. Fui direto para essa resposta por conta desse ‘tíquezinho’ aqui (resposta aceita), eu leria direto esta resposta, eu não vou nem olhar as outras.” (Participante P30217).

“Não costumo olhar. A data de postagem só se for algo muito antigo, uma tecnologia que mudou, mas também não costumo olhar.” (Participante P22218).

“Confesso que inicialmente eu não olho a reputação de quem perguntou, mas muitas das vezes costumo olhar a reputação de quem respondeu.” (Participante P21820).

“Pela primeira (primeira resposta, mais votada, mas não aceita). Comecei com o primeiro, que teve 128 votos. Tem 13 respostas (na trilha).” (Participante P22519).

Na Tabela 8 destacamos os elementos de foco em uma trilha CQA durante o processo analítico e cognitivo do participante.

Tabela 8 – Elementos da trilha CQA de Foco de atenção pelo Participante (Atividade Livre)

Elemento em Foco	Ocorrências	Participantes	Características e finalidade
Todas as respostas	0	0	
Título da Pergunta	8	4	Participante também processa o corpo da pergunta, total ou parcialmente.
Corpo da Pergunta	10	7	Na sequência o foco muda para primeira resposta.
Comentário da Pergunta	1	1	Raramente averiguado.
Tag usada na Pergunta	1	1	Raramente averiguado, mas utilizado para confirmar domínio ou assunto.
Pontuação da Pergunta	2	1	Raramente averiguado, mas utilizado para confirmar a própria qualidade da pergunta.
Recência ou postagem mais atual	5	4	Utilizado para averiguar possível desuso do conteúdo frente a dinâmica de evolução do assunto.
Reputação	6	4	Raramente averiguado, mas utilizado para confirmar confiabilidade da informação.
Melhor resposta ou destacada	14	9	Reforça a confiabilidade da informação.
Quantidade de Respostas	8	5	Quantitativo baixo gera sentimento de desconfiança.
Três Primeiras de Respostas	3	3	Geralmente um limitador do quantitativo de respostas serem processadas.
Primeira resposta	9	7	Sempre tem a atenção do participante no foco inicial ou subsequente.
Pontuação da Resposta	3	2	Raramente averiguado, mas utilizado para confirmar confiabilidade da informação.
Comentários da Resposta	3	2	Raramente averiguado, mas utilizado para eliminar dúvidas quanto a aplicabilidade da resposta.

Fonte: do Autor (2022)

Uma vez realizada a análise da trilha, quanto a tomada de decisão sobre o aceite e uso dessa trilha, a partir do modelo ESKIP (TIBAU; SIQUEIRA; NUNES, 2022) verificamos a existência de três possibilidades: (i) Aceitação, (ii) Dúvida e (iii) Recusa. Em todos os casos, o participante era instado a declarar (falar) qual a decisão, além das evidências, que eram por vezes observáveis através do comportamento em tela.

Na Tabela 9 apresentamos o quantitativo de ocorrências (tomadas de decisão), de participantes e as principais evidências. Os comentários da pergunta são raramente utilizados, devido à baixa acreditação pelos participantes. Já os comentários da resposta, são mais utilizados que os da pergunta, e se prestam ao propósito de reforçar a aplicabilidade da resposta associada, ou, quando o objetivo é para conhecimento sobre algo, onde restam dúvidas que podem ser sanadas por material mais aprofundado, como documentação oficial fornecida por *links* nesses comentários, ou novas informações.

“Buscaria saber sobre esse caso aqui [selecionando o trecho de comando de código no texto da resposta], mas diz que tem potencial problema de desempenho. Já estou vendo aqui nos comentários com uma sugestão de outro comando que poderia fazer mais sentido, e ele indica a documentação do *Redis*. Posso entrar aqui?” (Participante P21218).

Tabela 9 – Tomada de Decisão e Evidências na análise da Trilha CQA (Atividade Livre).

Decisão	Ocorrências	Participantes	Principais evidências
Dúvida	4	4	Poucas Interações ou Respostas. Conteúdo de Página Externa Acessada.
Recusa	8	5	Conteúdo (Trecho Textual do Corpo da pergunta ou da resposta).
Aceitação	18	12	Trecho de Código. Trecho textual da Resposta.

Fonte: do Autor (2022)

Observamos que as poucas ocorrências de dúvidas estavam associados à pouca quantidade de respostas, à existência de outra trilha aberta em outra aba para ainda ser analisada ou ao entendimento obtido a partir do acesso a um conteúdo do *link* externo disponível na trilha. Em todos os casos, elas evoluíram para recusa daquela trilha. Eis um exemplo:

“De cara, observo a quantidade de respostas. Essa (trilha) teve apenas uma resposta. O que já me desanima.” (Participante P21616).

Para a recusa da trilha como útil para a necessidade de informação, ocorre a análise de conteúdos textuais do corpo da pergunta e da resposta em relação à intenção do participante na busca. As ações subsequentes observadas são de ver demais trilhas abertas no caso de múltiplos resultados da busca em paralelo, e principalmente refazer a consulta no processo de busca. Também existe o comportamento de revisitar os resultados da busca para avaliar outros resultados ainda não visitados.

“Parte sim, (...). Mas digamos que eu tô querendo algo ainda antes disso. (...) No caso ele não me atendeu.” (Participante P21915).

“Caso não resolvesse o meu problema eu tentaria outras estratégias. Eu continuaria olhando os outros resultados ou ia tentar reformular a consulta.” (Participante P21820).

Para o aceite da trilha, observamos um processo analítico mais intenso, com muita reflexão sobre o conteúdo, que despendia um tempo maior na fala e explicação, com a preponderância da análise de trechos de texto da resposta e de trechos de códigos existentes tanto nas respostas (principalmente) quanto na pergunta. Outras evidências também eram sinalizadas, mas vinham sempre associadas ao trecho de código ou trecho textual da resposta.

“Está indo ao encontro do que vi em artigos. (...) Estou lendo as respostas e vendo se dá para tirar algo. Mas é que as palavras que usei são bem genéricas. Normalmente pesquiso coisas mais específicas. (...) Eu acho que sim, serve. Acho que foi.” (Participante P21616).

“Para mim, a primeira respondeu, e ainda deu um *link* externo, que seria um caminho para saber mais sobre o assunto.” (Participante P22218).

Aplicando a análise de similaridade do comportamento dos participantes no processo analítico na CQA, conforme já efetuado anteriormente para o processo de busca, verificamos através da Tabela 10 que o máximo alcançado foi 0,71 entre os participantes P21710 e P30217, enquanto o mínimo alcançado foi 0,00 entre P31718 e a maioria dos participantes, bem como entre P22423 e diversos outros. Em ambos os casos, houve maior incidência de incompatibilidade de comportamento com os demais do grupo, caracterizado provavelmente pelo perfil de ambos participantes serem no espectro não desenvolvedor, diferente do perfil declarado dos demais, agravado pelo fato de não serem usuários de CQA.

Tabela 10 – Similaridade Comportamental entre os Participantes na Análise das Trilhas CQA (Atividade Livre)

Participantes	P21510	P21610	P21616	P21710	P21715	P21720	P21818	P21820	P21915	P22015	P22218	P22423	P22519	P22613	P30217	P31718	P31818
P21510																	
P21610	0,24																
P21616	0,22	0,29															
P21710	0,23	0,15	0,22														
P21715	0,29	0,13	0,26	0,33													
P21720	0,16	0,17	0,40	0,23	0,13												
P21818	0,35	0,18	0,32	0,24	0,28	0,29											
P21820	0,35	0,16	0,32	0,31	0,36	0,21	0,33										
P21915	0,28	0,38	0,45	0,13	0,12	0,35	0,33	0,14									
P22015	0,27	0,13	0,14	0,18	0,25	0,19	0,26	0,11	0,33								
P22218	0,28	0,22	0,26	0,21	0,27	0,21	0,47	0,20	0,41	0,43							
P22423	0,15	0,08	0,05	0,00	0,22	0,07	0,05	0,07	0,00	0,00	0,00						
P22519	0,09	0,10	0,25	0,20	0,18	0,20	0,21	0,25	0,09	0,11	0,19	0,00					
P22613	0,06	0,14	0,15	0,22	0,18	0,13	0,10	0,06	0,13	0,27	0,06	0,00	0,12				
P30217	0,20	0,13	0,20	0,71	0,27	0,20	0,21	0,36	0,12	0,25	0,27	0,00	0,18	0,18			
P31718	0,08	0,08	0,00	0,00	0,00	0,00	0,00	0,00	0,07	0,10	0,07	0,17	0,07	0,00	0,00		
P31818	0,21	0,33	0,21	0,10	0,08	0,21	0,10	0,06	0,38	0,17	0,29	0,11	0,06	0,00	0,08	0,13	

Fonte: do Autor (2022)

Na Tabela 11 observamos a análise estatística do grupo que demonstra a heterogeneidade comportamental dos participantes, com uma baixa similaridade.

Tabela 11 – Sumarização do Comportamento dos Participantes na Análise das Trilhas CQA (Atividade Livre).

Métrica	Valor
Desvio Padrão	0,122501144897055
Variância	0,015006530501089
Mínimo	0,000
1º Quartil	0,090
Mediana	0,180
3º Quartil	0,260
Máximo	0,710

Fonte: do Autor (2022)

Novamente observamos um comportamento heterogêneo do grupo de participantes no processo analítico e cognitivo em CQA, tal qual já observados para o processo de busca. Ressaltamos que nesta seção o objetivo lançado para busca e análise das trilhas foram definidos pelos próprios participantes, conferindo uma perspectiva particular de interesse. Na próxima seção, uniformizamos tal objetivo a partir de uma proposição comum para avaliar se os comportamentos e estratégias se modificam ou não.

4.2.3 Análise do Processo Cognitivo de construção da solução

Para a avaliação do processo de construção da solução, dado que as atividades da sessão envolvem ou a busca por uma solução, implementação e informação para uso imediato, ou a busca por um conhecimento sobre algo, consideramos as perspectivas de aprendizado a partir de teorias de *Problem-Solving* (FREDERIKSEN, 1983) e da *Unified Theories of Cognition* (NEWELL, 1994) destacando o processo de análise e definição do espaço do problema, baseado na proposição de Schoenfeld (2014) (SCHOENFELD, 2014), que envolve quatro aspectos: (i) recursos, um conjunto de operadores; (ii) heurística, um conjunto de métodos para atingir os objetivos; (iii) controle, um conjunto de regras de seleção; e (iv) crenças, um conjunto de objetivos. Conforme essas abordagens teóricas, a aprendizagem surge de atividades direcionadas a objetivos, isto é, as necessidades de informação. Conhecimentos específicos precisam ser adquiridos para satisfazer essas necessidades. Estruturas e processos cognitivos dos participantes são ativados para tratar tais conhecimentos, e propiciar o aprendizado. Não nos aprofundamos nos processos e estruturas cognitivas, focamos nas ações e resultados percebidos ou declarados pelos participantes. A comparação dos resultados do processo de busca na Web sobre bases CQA com processamento analítico das trilhas da CQA realizados pelos participantes, seguindo o uso da análise de TAP (Ericsson & Simon, 1984) onde os relatos verbais de uma pessoa resolvendo uma tarefa são usados como indicadores de processos cognitivos, temos os elementos conforme a Tabela 12 a seguir.

Tabela 12 – Comportamento dos Participantes para Construção da Solução (Atividade Livre)

Comportamento	Ocorrência	Aspectos cognitivos para solução
Uso de Mouse para marcar partes de atenção na tela	0,0%	Controle e Recursos
Armazenar solução para uso posterior	0,0%	Controle e Recursos
Usar favoritos do navegador para guardar a trilha CQA	10,0%	Controle e Recursos
Validação da solução com material externo	10,0%	Heurística e Recursos
Múltiplas avaliações e uso de várias respostas da trilha	20,0%	Heurística e Recursos
Uso direto da solução encontrada na trilha CQA	30,0%	Heurística e Crença
Uso preliminar ou inicial da informação da trilha CQA	15,0%	Heurística e Crença
Adaptação da solução encontrada na trilha CQA	15,0%	Heurística e Crença

Fonte: do Autor (2022)

A categorização do comportamento em função do aspecto cognitivo para construção da solução do problema na Tabela 12, pode ser melhor compreendida a partir de alguns relatos a seguir:

“(...) Isso daqui já ficaria nos meus favoritos. Por que em algum momento ia precisar de utilizar um CSS que pudesse fazer isso.” (Participante P21915).

“Usarei esse aqui e resolver minha questão. [Editando no *Notepad*]. Mas, eu ajustaria.” (Participante P21510).

“Basicamente o que mais me atrai é ter um código prático [selecionando o texto/segmento de código com o mouse] que eu possa usar e adaptar para a minha realidade.” (Participante P30217).

“Aqui no caso não é um ‘post’ sobre a solução de um problema. É mais conceitual mesmo, pelo que estou percebendo. Então é mais o que estou procurando, um ponto de partida.” (Participante P31818).

“Nesse caso aqui foi fácil porque na própria pergunta já tem a resposta que quero. Não precisei ver as respostas.” (Participante P21610).

“(...) Vou na primeira (resposta) e de cara já tem o que eu quero. A partir daqui eu já iria talvez de novo buscar na documentação (...), ver como que aplico esse comando apesar de estar aqui. [Marcando parte do trecho de código na primeira resposta e aceita].” (Participante P21818).

4.3 Processo de Busca na Web com CQA para Atividade Condicionada Direta

Diferente da seção 4.2 anterior, onde exploramos o comportamento natural de busca na Web sobre CQA, nesta, condicionamos os participantes a uma necessidade informação de implementação de solução técnica na área de computação, associada a um assunto

conhecido (série de Fibonacci). O aspecto de pergunta direta está relacionada à possibilidade de encontrar uma solução diretamente a partir dos resultados da busca. Na proposta de tarefa existem dois conceitos envolvidos, a série de Fibonacci e a complexidade algorítmica, expandindo a análise para o entendimento do quão o participante tem conhecimento ou não, e se isso influencia e de que forma nas estratégias e comportamentos adotados. O enunciado da pesquisa:

“Enunciado Sessão #2: implemente um algoritmo com complexidade $O(n)$ que compute a sequência de Fibonacci até número X dado, tal que X pertence aos Números Naturais. Considere qualquer linguagem de programação que desejar. A tarefa única seria 'Computar o somatório dos números que façam parte da sequência de Fibonacci(...)’”.

4.3.1 Análise do Comportamento no Processo de Busca na Web

Conforme a estrutura de conceitos adotadas na Subseção 4.2.1, mantivemos a organização nas seguintes subcategorias para elaboração da codificação a partir da análise dos dados coletados:

- **Concepção da busca.** A Figura 6 esboça as possibilidades verificadas através da codificação. Para as atividades da sessão corrente, identificamos a ocorrência da divisão do problema em partes em 18,5%, enquanto a reprodução do problema, por ser proposto, caiu para 29,6%, assim como a reflexão sobre o problema, que caiu para 51,9% dos casos. Isto se deve ao fato de além de terem que lembrar ou conhecer a série de Fibonacci, também identificaram terem que avaliar uma solução de implementação de tal série segundo a proposta de tarefa que atendesse ao critério de complexidade estabelecido. Assim o comportamento de reflexão do participante foi alterado para dividir o problema em partes, da de menor conhecimento para a maior, como no relato a seguir:

“Primeiro faria uma pesquisa sobre a ordem de complexidade do algoritmo de Fibonacci. Como não sei se é a complexidade dada, pode ser redundante na questão (...) Fui reformulando a consulta. A linha de raciocínio foi a seguinte: ver se qualquer algoritmo de Fibonacci é $O(n)$, e vi que dependendo de como se implementa pode ser de outra complexidade.” (Participante P21610).

Novamente observamos comportamentos de busca exploratória, desta vez para se familiarizar com o contexto e necessidade de informação imposta na sessão num processo de aprofundamento gradual. Eis alguns exemplos:

“Eu começaria de uma busca mais genérica. Costumo dar uma olhada. Aí, por exemplo, como um dos requisitos aqui do exercício aqui é $O(n)$, se eu não encontrar aqui algo relacionada ao

O(n), aí buscarei mais uma vez aqui (fazer nova consulta na página de busca).” (Participante P21715).

“Primeiro eu fiz essa pesquisa mais geral. Nessa altura dou uma olhadinha no que tem. Aí depois [Entrando na trilha aberta primeiramente na outra aba] incluirei o termo (...) e eu vou rever a minha *query* para ver se encontro alguma coisa mais específica.” (Participante P21616).

Os casos de reprodução estão relacionados à busca direta, com o uso direto da questão-pergunta da tarefa ou a sua versão interpretada traduzida para o idioma inglês, colocados como a consulta, como no relato:

[Vendo a atividade posta no *chat*] [Selecionou a questão da atividade proposta e colocou no campo da busca 'Computar o somatório dos números que façam parte da sequência de Fibonacci'] (Participante P21720).

- **Seleção dos termos e texto da busca.** Com base na codificação da Figura 7, observamos o uso da sugestão do mecanismo de busca durante a digitação da consulta (4 participantes). Houve a predominância de construção textual livre (52,6%), e uso de estruturação de perguntas (10,6%) quando se tratava de desconhecimento do contexto ou conceito envolvido. No sentido contrário, tivemos a percepção de uso de palavras-chave em 36,8% das ocorrências, sempre associado a compreensão do domínio do problema e assunto a partir de reflexão sobre a questão.

Exemplos de formato de pergunta:

“[Clicou em pesquisar, *query* 'Fibonacci how to implement in python']. (...) Aí tem a questão do *Loop*. ... [Refez a consulta como 'Fibonacci function sum how to implement in python'].” (Participante P21616).

“Para entender um pouco melhor sobre o quê que é complexidade, eu ia ver qual é a melhor. Na realidade, está exigindo O(n), então eu teria que lembrar. Aí talvez, quando quero aprender alguma coisa nova ou relembrar bota uma frase. Eu não uso palavras-chave, eu colocaria uma frase [Digitou na busca 'what's O(n)'].” (Participante P30217).

Exemplos de formato de texto livre:

“Termo usado na busca 'computing Fibonacci numbers in R'. ” (Participante P21510).

“Copiei e cole aqui 'computar o somatório dos números de Fibonacci' [o participante colocou a frase de busca proposta na pesquisa].” (Participante P22519).

“Procurarei aqui Fibonacci em javascript. [Digitou 'Fibonacci using for in javascript']. Não coloquei O(n). Coloquei lá uma interpretação minha com relação a O(n).” (Participante P21820).

Exemplo de formato a partir de termos-chave:

“Eu tô pensando aqui em começar justamente nessa linha de um algoritmo que tenha por base aí a questão de Fibonacci. [Digitando “algoritmo, Fibonacci”]. Independente de linguagem de programação. Aí dependendo do que ele me trouxer, eu já faço uma avaliação. Por exemplo, solução simples para algoritmo de Fibonacci utilizando inclusive o *Stack Overflow*. Como se eu não soubesse como fazer uma implementação.” (Participante P22015).

Diferente do que ocorreu durante a atividade de busca de autointeresse, não houve direcionamento de fonte alvo, somente o uso de expressões para garantir o termo/texto exato (emprego de texto entre aspas) com apenas um participante (“*natural numbers*”), ou cobrir um aspecto da tarefa (“ $O(n)$ ”) com dois participantes. O uso do recurso de aspas também foi utilizado para separar os termos. Eis alguns relatos de exemplo:

“[Refez a consulta com “fibonacci algorithm” $O(n)$ “natural numbers”]. Utilizando aspas para segmentar os termos compostos para tentar um resultado mais preciso.” (Participante P21610).

Os recursos para tentar direcionar o mecanismo de busca mais utilizados foram: 41,2% expressões, 41,2% conectores da linguagem/idioma, 5,9% aspas e 5,9% vírgula. Já a combinação desses recursos, registramos 17,65% com expressões e conectores, e ainda 5,88% com expressões e aspas. Tais recursos parecem servir ao propósito de refinamento dos resultados, evitando analisar múltiplas páginas de resultados:

É muito mais fácil para mim, refinar a minha pesquisa de alguma forma, e usar... eu não sei se aqui posso usar o termo + (mais) entre aspas... Então, eu refinaria usando esses comandos (“expressões” e conectores de busca: +). Mais do que ir para uma segunda página.” (Participante P21616).

- **Formulação da consulta.** Os elementos estão codificados a partir dos relatos conforme a figura 8 com base no *ESKiP*. Destacamos a maior incidência de especialização (23,3%) da consulta do estado inicial. Houve a substituição de palavras em 5,9% a partir da sugestão do CSE, e a tradução linguística (PT-BR para EN) em 5,9% mas previamente à reelaboração de nova consulta (Participante P21616). Observamos que 17,6% dos participantes não se preocuparam com a correta grafia (“Fibonacitti”, “fibronacci”, “specfic”) das palavras, acatando a sugestão do CSE ou apenas verificando os resultados que retornaram adequadamente à grafia correta. Os casos de **’Nova Consulta’** (do *ESKiP*) foram gerados pelos participantes a partir do que foi observado em algum dos resultados retornados, copiando-os como a nova *query* (duas transições a partir do Estado Inicial) ou utilizando seus conceitos para reformulação (duas transições a partir da Nova Consulta anterior, recorrente em um único participante). Na Tabela 33 podemos observar essas situações, onde temos 30

(trinta) consultas dessa sessão de busca condicionada direta, com 13 (treze) consultas adicionais ao Estado Inicial (primeira consulta).

- **Checagem dos resultados.** Com base na figura 9, observamos um aumento no comportamento de leitura sequencial dos resultados de busca (70,6%) na página de resultados do mecanismo de busca (SERP), em função da insegurança de alguns participantes a respeito da própria intenção imposta pela demanda da sessão de atividades. A importância da análise da fonte regrediu para metade (23,5%) do que se observou quando da busca de autointeresse do participante, embora o reconhecimento da fonte se apresente como contribuição secundária para os participantes com perfil de desenvolvimento. Já a checagem baseada no título e no texto de resumo praticamente permaneceram iguais (ambos com 52,9%). Desta forma, confirmamos através desse segundo cenário de busca na Web sob CQA que há um respeito à ordem dos resultados na SERP, bem com à caracterização pelo título e pelo texto de resumo. Eis um exemplo do uso da sequência e da identificação da fonte:

“O primeiro é do *Stack Overflow*, eu gostei. Porque como programador, encontro muita coisa de programação lá e facilita bastante.” (Participante P31818).

“(...) Bom aí temos aqui o *Stack Overflow* novamente. Ah sim, está filtrado, né? Mas vamos lá. Iterando a sequência de Fibonacci. Tem o *Javascript* aqui no texto. Esse aqui eu abriria e abrirei inclusive. [Abrindo o 1 resultado da consulta que é do *Stack Overflow*]. Abrirei em uma nova aba, ver se me satisfaz.” (Participante P21820).

“Analisando resumo do primeiro resultado ‘Dúvida na extração da equação de recorrência do Fibonacci’ (fonte *pt.StackOverflow*). Analisando resumo do segundo resultado ‘Complexidade computacional da sequência do Fibonacci’ (fonte *Quora*). Hum! O de 2 elevado a N. [Clicou nesse 2.º resultado, por ficar intrigado com a complexidade informada no resumo]. [Abriu em uma nova janela].” (Participante P21610).

- **Extração da informação.** A partir da figura 11 com a codificação completa dos relatos, onde tivemos um único caso de ir além da primeira página de resultados. Novamente, tal qual na busca de autointeresse, houve predominância de processamento dos primeiros resultados (47,06%), onde 87,5% desses casos o participante chegou ao melhor resultado após aplicação dos critérios de checagem (41,18% do total de ocorrências). Destacamos ainda, o uso de mais de uma página de resultado foi relatado quando se trata de entendimento de algum assunto, conforme o relato a seguir:

“Quanto estou fazendo uma busca que é mais genérica... no sentido de entender um domínio de conhecimento que irá ter diversos conceitos diferentes... significa que eventualmente entrarei em diferentes sites para que aquelas informações se complementem. Então nesse caso aí é que eu acessaria mais de uma (página de resultados).” (Participante P21710).

Interessante notar que, embora fosse uma demanda imposta aos participantes, que não se enquadrava no espectro de autointeresse, ainda assim, por se tratar de um assunto convencional de sistemas de informação e de conhecimento comum aos perfis dos participantes, observamos geralmente o mesmo comportamento em ambas as seções ('Fibonacci': condicionada direta; autointeresse). Isto se deve à prevalência de intenções para resolver um problema pontual, de implementação, com poucos casos de exploração de informações ou conhecimentos. O diferencial observado foi para tratar a questão de complexidade estabelecida na demanda da tarefa, onde dois comportamentos foram observados: (i) ignorar a complexidade a priori e focar na extração da informação de implementação de 'Fibonacci'; (ii) considerar de início a complexidade na extração da informação. No primeiro caso, os participantes optaram pela interpretação do que seria a solução adequada para a demanda. No segundo caso, foi utilizada a expressão ' $O()$ ' direto na extração. Eis alguns exemplos de relatos de comportamento:

"Normalmente eu pego o primeiro como eu abri aqui, que eu já vi que ele tem. O segundo ele usa programação dinâmica. Achei interessante, abrirei também em outra aba. Olha esse aqui, por exemplo, já está me dando inclusive a relação de 1,62 elevado a N quer dizer $O(N)$. Esse aqui está com uma resposta boa, usarei esse aqui também. Aqui esse aqui tem várias implementações de Fibonacci inclusive com $O(\log.N)$ e $O(N)$, caso queira realizar alguma comparação. Se por acaso esse me parecer mais completo abro uma aba. Então tenho várias abas e começarei pela primeira. [Sempre avaliando o texto resumo dos resultados]." (Participante P21915).

"A primeira coisa que eu faria é buscar a fórmula de Fibonacci. Eu até acho que sei, mas não tenho certeza. ... Selecionei o primeiro (resultado da busca)." (Participante P22218).

O comportamento para exibição da informação extraída da busca foi novamente majoritariamente a abertura em uma nova janela ou aba (88,23%), enquanto apenas 11,77% conduziram na mesma aba. Contudo, nesta sessão de atividades houve um aumento de mais de 30% nesse comportamento em relação ao caso de busca de autointeresse. Também aumentaram o número de casos de participantes que abriram múltiplos resultados em paralelo, agora com 6 (seis) ocorrências. Da mesma forma que na sessão anterior, estavam explorando os conteúdos para entender o assunto (relembrar ou compreender: 'fórmula de Fibonacci', 'complexidade'), principalmente sanando dúvidas a respeito. Eis alguns exemplos de relatos:

"Às vezes abro várias janelas. Como, por exemplo, eu vi ali no resultado de cara duas partes que me interessem, normalmente sendo complementares, eu abriria a duas, uma em cada aba." (Participante P21710).

"Aí encontro aqui o que estava em mente no *Stack Overflow*, a melhor sequência. Olharei isso aqui primeiro. Vejo que tem outra aqui (resultado) também. Abrirei ambas. Na dúvida é melhor assim." (Participante P21915).

“[Observando os títulos e corpo de cada pergunta das trilhas abertas]. Então assim, seria mais janelas de auxílio para não ter que ficar procurando. Eu meio que dou uma geral em tudo aquilo que nesse instante, meio que me importa para tentar entender esse algoritmo. Dificilmente fico só em uma janela e me concentro horas em cima dela. Porque, de repente, eu estou em uma linha de leitura que possa surgir alguma dúvida, aí eu já vou para outra janela que pode me dar a resposta. Procuro geralmente trabalhar dessa forma.” (Participante P22015).

Na análise de similaridade de comportamento na busca, conforme a Tabela 13, observamos ainda uma grande diferença de comportamento, com pico de similaridade máxima de 0,64 (entre P21510 e P21720) e mínima de 0,13 (entre P22519 e P31718), apesar de haver uma suave melhora na similaridade geral do conjunto em relação à busca livre de autointeresse.

Tabela 13 – Similaridade Comportamental entre os Participantes no Processo de Busca (Condicionada Direta).

Participantes	P21510	P21610	P21616	P21710	P21715	P21720	P21818	P21820	P21915	P22015	P22218	P22423	P22519	P22613	P30217	P31718	P31818
P21510																	
P21610	0,45																
P21616	0,39	0,33															
P21710	0,43	0,29	0,44														
P21715	0,60	0,48	0,35	0,38													
P21720	0,64	0,43	0,44	0,40	0,39												
P21818	0,21	0,26	0,18	0,36	0,27	0,29											
P21820	0,43	0,35	0,21	0,20	0,29	0,50	0,50										
P21915	0,32	0,39	0,27	0,21	0,42	0,44	0,33	0,53									
P22015	0,53	0,36	0,37	0,62	0,39	0,50	0,29	0,31	0,37								
P22218	0,21	0,33	0,25	0,36	0,27	0,20	0,33	0,15	0,18	0,29							
P22423	0,29	0,32	0,24	0,45	0,33	0,27	0,63	0,33	0,31	0,36	0,63						
P22519	0,31	0,40	0,26	0,27	0,53	0,29	0,33	0,36	0,41	0,22	0,33	0,42					
P22613	0,54	0,42	0,44	0,38	0,57	0,40	0,25	0,38	0,35	0,31	0,36	0,45	0,46				
P30217	0,47	0,46	0,35	0,37	0,43	0,45	0,35	0,44	0,48	0,38	0,28	0,41	0,50	0,37			
P31718	0,50	0,25	0,31	0,33	0,33	0,46	0,30	0,33	0,40	0,46	0,18	0,27	0,13	0,33	0,20		
P31818	0,47	0,26	0,39	0,33	0,33	0,44	0,31	0,54	0,47	0,44	0,21	0,38	0,31	0,43	0,33	0,50	

Fonte: do Autor (2022)

Na Tabela 14 observamos algumas informações estatísticas do conjunto representado por essa matriz de similaridade entre os participantes.

Tabela 14 – Sumarização do Comportamento dos Participantes no Processo de Busca (Condicionada Direta).

Métrica	Valor
Desvio Padrão	0,105493523224252
Variância	0,011128883442266
Mínimo	0,130
1º Quartil	0,290
Mediana	0,360
3º Quartil	0,440
Máximo	0,640

Fonte: do Autor (2022)

4.3.2 Análise do Comportamento no Processo analítico da trilha CQA

Visto que o participante tenha selecionado um ou mais resultados da busca, e notadamente o tenha feito a partir da análise do título ou do texto de resumo, ambos presentes

na parte de pergunta da trilha da CQA em análise, observamos que a leitura do título da pergunta é minorada na maioria das vezes.

Novamente identificamos que a tomada de decisão na análise da trilha na CQA, a respeito de atender à necessidade de informação para implementação de algo, está associada à resposta e calcada na qualidade da postagem e na organização que permita analisar e recuperar a informação desejada. A existência de boa votação ou 'pontuação positiva' para a resposta e sinalizada como validada pelos pares, de eventuais comentários sobre tal resposta para sanar dúvidas de entendimento, bem como trechos de código que possam ser verificados e testados, são cruciais para a tomada de decisão. A seguir, destacamos alguns relatos:

“Essas características de ter lá algum código que possa olhar e analisar. A clareza como a postagem é feita e organizada facilita e chama atenção para a análise e a recuperação da informação. Se o conteúdo está mal ou bem escrito, está organizado, está bem votado. O usuário que postou é aquele em quem mais gosta ou aprecia (reconhecimento).” (Participante P21610).

“Eu acho que a ferramenta (*Stack Overflow*) é bem interessante nesse sentido, de tentar nos poupar tempo a partir da maneira que ela foi implementada. Eu nem li aqui direito (textos da pergunta e da resposta aceita). Na verdade, aqui está a resposta (marcando o código e resultado de execução em tela com mouse). Já sabia do tipo de resposta que precisava, então nem li o texto, já fui mais para a resposta.” (Participante P21820).

Nesse caso de comunidades, avalio também as respostas que estão sendo dadas em relação aquilo. Aqui ele está indo numa linha um pouco mais detalhada [lendo o texto da resposta e o código apresentado](...) Então, aqui ele dá uma explicada melhor.” (Participante P22015).

“Eu nem sei se verei esse, nenhum deles está validado (resposta aceita).” (Participante P30217).

“Quando fico na dúvida com relação ao conteúdo daqui da resposta aceita. Geralmente é mais quando conheço (assunto) e quero tirar uma dúvida. Olho e vejo que está estranho, aí costumo olhar aqui também (comentário da resposta) o que o pessoal está falando.” (Participante P21715).

A navegação por hiperlinks, externos ou não à CQA, novamente foi definida pelo desejo de mais informação e principalmente pelo hábito do participante. Somente três participantes manifestaram interesse em acessar links internos à trilha, dois por hábito e um por desejo de ampliar o conhecimento sobre o assunto.

Nesse caso específico de algoritmo, eu avalio, faço uma análise mínima do que ele está chamando algoritmo aqui (no corpo da pergunta). (...) Então no caso aqui (da primeira resposta), ele está meio que dando uma resposta em relação a esse assunto. Observo o que pode estar relacionado a isto. Então dependendo dos links relacionados a esse, aquilo que estou vendo, se não me satisfaz, eu avalio também os links relacionados. Feito isso, eu ainda dou uma verificada no quê está relacionado (links de perguntas relacionadas), e se for o caso, eu entro lá também para tentar ampliar o meu conhecimento a respeito aquele assunto.” (Participante P22015).

“Tenho o hábito de seguir links. Tenho o hábito de criar aquela “tripa”: um (link) que manda para um que manda para outro, assim por diante. Vou seguindo desde que esteja me levando para algum lugar.” (Participante P31718).

Para o processamento cognitivo da postagem, cerca de 29,4% dos participantes ou optaram por conteúdos em português, ou por traduzir para o idioma português para tornar o entendimento mais ágil, mesmo no caso de terem efetuados buscas já em inglês. O destaque está na necessidade de compreensão do conceito de ‘complexidade $O(n)$ ’ a partir da análise das trilhas.

“[Começou a analisar o texto da pergunta]. [Optou por traduzir o conteúdo da trilha na aba do navegador de inglês para português]. [Continuou a ler]. Essa aqui não está boa.” (Participante P21610).

“[Visualizando a trilha] [Rolagem vertical rápida, parando na melhor resposta] [Analisando o algoritmo proposto] Pela minha análise é um algoritmo $O(n)$ [referindo-se a questão da complexidade que evitou explorar na *query*].” (Participante P21710).

Embora o objetivo da presente sessão de atividade fosse a implementação de solução, observamos que 29,4% (cinco participantes) buscaram informação para uso associado à implementação, e dois (11,7%) tentaram obter conhecimento a respeito. Em todos os casos, o alvo era a compreensão de ‘complexidade algorítmica’, seja para lembrar ou entender do que se trata. Assim, a inexistência de ‘familiaridade’ sobre determinados tópicos ou assuntos envolvidos em uma demanda levam à condição imperativa de sanar tal hiato de conhecimento antes de atacar o objetivo da demanda.

“Para entender um pouco melhor sobre o que é complexidade, eu ia ver qual é a melhor. No enunciado está exigindo $O(N)$, então eu teria que lembrar. (...) [Entrou na trilha]. Eu faria diferente a minha busca. Na verdade, para buscar um conceito, lembrar um conceito, eu não usaria *Stack Overflow* (isto é, CQA).” (Participante P30217).

“Agora que eu me reorientei aqui, buscando a linguagem que eu conheço. Eu vou ler aqui [Lendo o título e corpo da pergunta]. Vejo a resposta aceita. Nessa aceita, verifico a questão de complexidade para ver se atende.” (Participante P21715).

“[Lendo a trilha] [Pulou para os comentários da postagem] [Entendeu os aspectos relativos à parte da complexidade na questão] Realizarei nova pesquisa.” (Participante P21610).

Quanto ao comportamento da leitura da trilha na CQA, observamos conforme a Tabela 15 um aumento na leitura seletiva, quando o participante vai direto para algum elemento/segmento da trilha, principalmente para trecho de código na primeira resposta ou a resposta aceita, bem como no detalhe da pergunta.

“Olhei somente o conteúdo, principalmente as expressões e fórmulas nas respostas. Vi também o código sugerido na pergunta e vi que não era adequado para $O(n)$. Já automaticamente descarto. Também tem o conhecimento prévio que auxilia na interpretação do conteúdo.” (Participante P21610).

A leitura em profundidade foi novamente detectada como associada à necessidade de entendimento aprofundado sobre um assunto, que deveria ser de maior complexidade. Nesse tipo de leitura em uma CQA, não só as respostas são acessadas, mas também os comentários associados. Eis um relato pontuando isso:

“Teria que ser um problema mais complexo (...) Nesses casos, pararia para ler os comentários explicando como o algoritmo funciona. Aqui é um exemplo muito simples e não vejo essa necessidade. Também tem o fato da experiência.” (Participante P21710).

Tabela 15 – Estratégia de leitura da trilha CQA (Atividade Condicionada Direta).

Leitura	Quantitativo	Percentual
Inversa(Pergunta+inversaoRespostas)	0	0%
Transversal	3	14,3%
Seletiva	8	38,1%
Sequencial	9	47,6%
....TopDown	4	19,0%
....EmProfundidade	1	4,8%

Fonte: do Autor (2022)

Os casos de leitura transversal estão realmente associados à captura visual de algo que se sobressaia na trilha, e a partir desse ponto, faz-se a leitura complementar associada. Tal comportamento foi o mesmo da sessão de autointeresse anterior, embora os participantes não tenham se repetido no todo. Apenas um participante parece ter esse comportamento (‘escanear’ a página) como hábito. Curioso notar que isso possa acontecer já no início da trilha, o que pode denotar um entendimento equivocado de leitura sequencial quaisquer, por levar o participante a ler muitos elementos da trilha. Eis relatos de exemplo:

“Eu pegaria esse aqui (marcando o código do corpo da pergunta). Agora começarei a olhar. [rolou a trilha e começou a olhar comentários da pergunta]. Tenho que olhar com mais calma.” (Participante P31718).

“Eu geralmente sou visual. Então vi um código, se fosse, por exemplo, nesse caso aqui, ele me deu um código simples. O que me chamou a atenção foi o código. Eu pegaria isso aqui (marcando o trecho de código).” (Participante P22423).

Quanto ao foco, destacamos que o corpo da pergunta demandou mais atenção inicial (70,59%) dos participantes nesta sessão do que na anterior de autointeresse. Na Tabela 16 apresentamos os elementos de foco do participante ao processar analiticamente a trilha da CQA.

Tabela 16 – Elementos da trilha CQA de Foco de atenção pelo Participante (Atividade Condicionada Direta)

Elemento em Foco	Ocorrências	Participantes	Características e finalidade
Todas as respostas	0	0	
Título da Pergunta	32	11	Participante também processa o corpo da pergunta, total ou parcialmente.
Corpo da Pergunta	57	15	Na sequência o foco muda para primeira resposta. Eventualmente muda-se de trilha, faz-se tradução do idioma ou consulta-se anotação externa.
Comentário da Pergunta	8	7	Raramente averiguado, mas utilizado para tirar dúvida da postagem.
Tag usada na Pergunta	0	0	Não ocorreu.
Pontuação da Pergunta	4	2	Raramente averiguado, mas utilizado para confirmar a própria qualidade da pergunta.
Recência ou postagem mais atual	2	2	Utilizado para averiguar possível desuso do conteúdo frente a dinâmica de evolução do assunto.
Reputação	7	4	Raramente averiguado, mas utilizado para confirmar confiabilidade da informação.
Melhor resposta ou destacada	51	14	Reforça a confiabilidade da informação.
Quantidade de Respostas	3	3	Quantitativo baixo gera sentimento de desconfiança.
Três Primeiras de Respostas	15	10	Geralmente um limitador do quantitativo de respostas a serem processadas.
Primeira resposta	23	10	Sempre tem a atenção do participante no foco inicial ou subsequente.
Pontuação da Resposta	11	8	Raramente averiguado, mas utilizado para confirmar confiabilidade da informação.
Comentários da Resposta	10	6	Raramente averiguado, mas utilizado para eliminar dúvidas quanto a aplicabilidade da resposta.

Fonte: do Autor (2022)

A primeira resposta, ou a melhor resposta, parece reter a atenção principal nesse tipo de comunidade pelos participantes (88,24%). Interessante notar que o participante que havia alegado não conhecer ou usar tais plataformas já demonstrou um comportamento mais alinhado ao demais, embora ainda não detectando alguns elementos visuais e iconográficos fornecidos pela plataforma (votação, marcador de melhor resposta, etc.).

“[Lendo o corpo da pergunta. Ficou em dúvida se servia. Questionou-se. Voltou a ver a anotação no bloco de notas. Voltou a aba da CQA, navegou até a 1.ª resposta e realizou leitura transversal].” (Participante P21510).

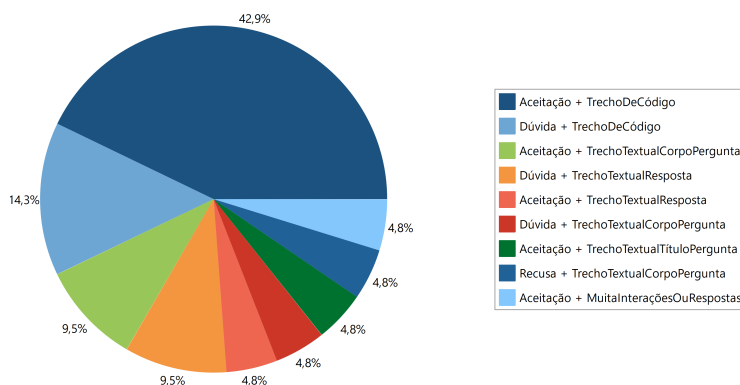
“O usuário que está fazendo a pergunta recebeu várias pontuações sobre isso. Então acredito que seja a pergunta mais relevante sobre esse assunto aqui no *Stack Overflow*. Mesmo porque, eles não deixam perguntas repetidas. . . Vou direto para a resposta. Essa aqui já me chama a atenção por vários motivos. Muita pontuação positiva, aceita como certa, tem um código aqui e você poder rodar o código para verificar (tem um botão de execução). Isso geralmente é o que espero de resposta de programação.” (Participante P21820).

“Ele aqui está dizendo olha tem as minhas 16 últimas respostas foram fechadas cinco anos atrás, ou seja, ninguém mais está usando essa resposta.” (Participante P21915).

“Aí verei a próxima resposta. A próxima resposta começo vendo a quantidade de votos. Porque para mim a quantidade de votos é um indicativo que as pessoas que passaram por aqui antes, deram maior importância àquela resposta do que para esta.” (Participante P31818).

Quanto a tomada de decisão em relação ao aceite e uso da trilha, exibimos na Figura 12 a correlação entre a decisão e os elementos de evidência dessa definição.

Figura 12 – Evidências de critérios de checagem utilizados para Tomada de Decisão. (Atividade Condicionada Direta)



Fonte: do Autor (2022)

Observamos que os trechos textuais tanto da pergunta quanto da resposta são essenciais para a decisão, principalmente quanto à existência de trechos de código. Devido à questão da 'complexidade $O(n)$ ', observamos o uso de comentários da resposta para elucidar esse entendimento para o trecho de código existente na resposta associada.

“Lendo a trilha. Pulou para os comentários da postagem.” (Participante P21610).

“Das respostas. . . tem que ler os comentários, porque o algoritmo pode estar errado. Alguém pode ter dado um comentário. Assim, o algoritmo parece certo, mas não está seguindo o método $O(n)$.” (Participante P21915).

Observamos que as poucas ocorrências (7 em 5 participantes) de dúvidas estavam associados ou a qualidade do conteúdo das respostas, ou existência de outra trilha aberta em outra aba para ainda ser analisada. Em todos os casos, elas evoluíram para recusa daquela trilha. Eis um exemplo:

“Ele usa o fib aqui, mas ele não fala qual que é e até quanto. Ele fala o número de vezes, mas não até qual número.” (Participante P22519).

As recursas (1 ocorrência em 1 participante) ou não aceitação da trilha como solução do problema está associado a análise do conteúdo do corpo da pergunta ou do não alinhamento da trilha com a necessidade de informação:

“[Lendo o corpo da pergunta sem muita atenção ao fragmento de código existente. Ignorou os comentários da pergunta. Passou para primeira resposta (aceita) e ficou analisando]. Ah! Deixa eu fazer outra coisa aqui. [Desistiu da trilha].” (Participante P22519).

Para o aceite (15 ocorrências em 13 participantes) da trilha, observamos um processo analítico mais intenso, com as mesmas características da sessão de busca de autointeresse anterior. Os participantes tendem a olhar e analisar o trecho de código com base em suas experiências, complementando com o entendimento a partir de trechos textuais da resposta e do corpo da pergunta, ou até mesmo confrontando mais de uma trilha, sobretudo para atender ao quesito de complexidade.

“No Laço, é sim. No outro aqui já é recursivo [mudou para a aba do outra trilha].” (Participante P21820).

Aplicando a análise de similaridade do comportamento dos participantes no processo analítico na CQA, conforme já efetuado anteriormente para o processo de busca, verificamos através da Tabela 17 que o máximo alcançado foi 0,59 entre os participantes P21715 e P31718, enquanto o mínimo alcançado foi 0,00 entre P22519 e quatro participantes (P21715, P21820, P21915, P22423).

Tabela 17 – Similaridade Comportamental entre os Participantes no Processo analítico na CQA (Condicionada Direta).

Participantes	P21510	P21610	P21616	P21710	P21715	P21720	P21818	P21820	P21915	P22015	P22218	P22423	P22519	P22613	P30217	P31718	P31818
P21510																	
P21610	0,29																
P21616	0,38	0,38															
P21710	0,33	0,35	0,22														
P21715	0,29	0,50	0,47	0,38													
P21720	0,33	0,35	0,29	0,45	0,47												
P21818	0,33	0,15	0,20	0,20	0,13	0,09											
P21820	0,31	0,33	0,21	0,55	0,53	0,55	0,18										
P21915	0,08	0,09	0,18	0,08	0,11	0,08	0,11	0,25									
P22015	0,50	0,26	0,33	0,20	0,33	0,29	0,17	0,27	0,14								
P22218	0,25	0,24	0,31	0,25	0,40	0,36	0,10	0,23	0,08	0,13							
P22423	0,30	0,26	0,19	0,44	0,27	0,30	0,29	0,40	0,10	0,15	0,20						
P22519	0,27	0,14	0,25	0,00	0,18	0,17	0,11	0,00	0,00	0,33	0,18	0,00					
P22613	0,36	0,25	0,25	0,19	0,32	0,19	0,07	0,25	0,31	0,31	0,29	0,23	0,06				
P30217	0,19	0,43	0,32	0,36	0,47	0,27	0,15	0,33	0,13	0,17	0,38	0,14	0,13	0,22			
P31718	0,31	0,33	0,21	0,31	0,35	0,31	0,18	0,38	0,15	0,27	0,33	0,56	0,07	0,43	0,33		
P31818	0,24	0,33	0,50	0,40	0,59	0,40	0,13	0,38	0,19	0,21	0,33	0,29	0,06	0,33	0,26	0,22	

Fonte: do Autor (2022)

Novamente observamos um comportamento heterogêneo do grupo de participantes no processo análise e cognitivo em CQA como verificado na sessão de atividade anterior. Isto confirma, que para uma sessão de atividade comum temos preservados os comportamentos e estratégias. Na próxima seção, abordamos uma sessão de múltiplas atividades que devem ser combinadas para atingir um objetivo, onde ampliamos a complexidade da demanda e assim, averiguar se em um cenário mais complexo os comportamentos e estratégias se modificam ou não. Na Tabela 18 observamos a análise estatística do grupo.

Tabela 18 – Sumarização do Comportamento dos Participantes na Análise das Trilhas CQA (Atividade Condicionada Direta).

Métrica	Valor
Desvio Padrão	0,127898635654094
Variância	0,016358061002179
Mínimo	0,000
1º Quartil	0,178
Mediana	0,265
3º Quartil	0,330
Máximo	0,590

Fonte: do Autor (2022)

4.3.3 Análise do Processo Cognitivo de construção da solução

A partir dos relatos verbais do participante declarando como irá construir a solução a partir dos resultados do processo de busca na Web sobre bases CQA com processamento analítico das trilhas da CQA, e considerando os aspectos cognitivos, temos os elementos conforme a Tabela 19 a seguir. Podemos observar que, por ser atividade simulada, os participantes têm a crença que já reuniriam as condições necessárias, através de suas próprias heurísticas, para simplificar a solução do problema, com pouco ou nenhum uso de recursos, ou aplicação de controle.

Tabela 19 – Comportamento dos Participantes para Construção da Solução (Atividade Condicionada direta).

Comportamento	Ocorrência	Aspectos cognitivos para solução
Uso de Mouse para marcar partes de atenção na tela	0,0%	Controle e Recursos
Armazenar solução para uso posterior	0,0%	Controle e Recursos
Usar favoritos do navegador para guardar a trilha CQA	8,7%	Controle e Recursos
Validação da solução com material externo	4,3%	Heurística e Recursos
Múltiplas avaliações e uso de várias respostas da trilha	13,0%	Heurística e Recursos
Uso direto da solução encontrada na trilha CQA	8,7%	Heurística e Crença
Uso preliminar ou inicial da informação da trilha CQA	4,3%	Heurística e Crença
Adaptação da solução encontrada na trilha CQA	60,9%	Heurística e Crença

Fonte: do Autor (2022)

A categorização do comportamento em função do aspecto cognitivo para construção da solução do problema na Tabela 19, pode ser melhor compreendida a partir de alguns relatos a seguir:

“[Retornou para a tela de busca, após rever a atividade no ‘Chat’ do Google Meet].” (Participante P21616).

“Aí, está aqui a resolução e tem os comentários do código. Fica mais fácil de entender. Aqui a pessoa dá uma resolução, uma explicação. Está explicando cada linha do código.” (Participante P22613).

“Se eu conseguir fazer isso, rodar esse programa e testar a complexidade dele, aí eu partiria para outra solução, para fazer a comparação de complexidade.” (Participante P31818).

“Aparentemente, parece que ele faz isso (o objetivo da pesquisa). Aí eu ia ter que levar um tempo para assimilar.” (Participante P31718).

“Nesse caso aqui já faz. [Racionalizando as linhas de códigos]. Posso adaptar o algoritmo.” (Participante P21710).

“Pegarei esse código aqui e vou adaptar para o que eu preciso.” (Participante P21715).

“Eu iria adaptar. Tentar adaptar talvez o mínimo né.” (Participante P21720).

“Isso sem sombra de dúvida, faria a adaptação do código. Geralmente eu pego e guardo isso como um código mesmo, numa pasta temporária até fazer o uso efetivo disso para aquilo que me interessa.” (Participante P22015).

“Eu poderia aproveitar e já guardar essa página nos meus arquivos. Normalmente, eu guardo aqui no histórico, guardo nos meus *bookmarks*. Então guardo isso e direi o que achei. . .” (Participante P21915).

4.4 Processo de Busca na Web com CQA para Atividade Condicionada Múltipla

Diferente das duas sessões de atividades anteriores, a primeira de livre escolha e de autointeresse do participante, e a segunda uma atividade simples condicionada direta, nesta sessão foi proposta uma necessidade de informação mais analítica, que demanda a combinação de múltiplas buscas e análise de trilhas CQA para a construção de uma solução resultante. Isto é, trata-se de uma atividade com pelo menos três tarefas distinguíveis, conforme o enunciado a seguir:

“ENUNCIADO SESSÃO #3: considere que tenha um conjunto de dados (*dataset*) de seu experimento em formato CSV armazenado no seu Google Drive, bem como um conjunto de dados de referência (obtido de outro trabalho/pesquisa para comparação).

O *dataset* de referência já está normalizado (*Gaussian distribution*). Mas será necessário avaliar se o *dataset* do seu experimento também está. Após isso, faz-se necessário avaliar a correlação entre os dois conjuntos.

Pode-se considerar implementação em R ou Python. Desta forma as tarefas necessárias são:

- 1) Acesso e leitura aos dois *datasets* no Google Drive.
- 2) Realizar o teste de distribuição normal no seu *dataset*.
- 3) Realizar o teste de correlação entre os *datasets*.”

Esse contexto de análise de conjunto de dados foi escolhido por requerer mais conhecimento e uma combinação de informações e técnicas de implementação, elevando o nível de complexidade em relação às sessões de atividades anteriores. Com isso, retiramos os participantes de suas zonas de conforto, de programação ou conhecimentos gerais de sistemas de informação. Desta forma, nivelamos os participantes frente ao problema numa perspectiva mais alinhada a exploração de informações.

Na proposta existem alguns conceitos que requerem o entendimento de seu significado e de como operacionalizá-los para uma solução: leitura no Google Drive, (teste de) distribuição normal, (teste de) correlação entre datasets.

Dada a complexidade da sessão, foi estabelecido o limite máximo de referência para a realização de 30 minutos. Apenas dois participantes não realizam a atividade até o fim. Um por estar utilizando o dispositivo celular para realizar todo o experimento, e outro por alegada limitação de tempo associada a problemas de conexão. Acima de 30 minutos somente o participante que fez parte do piloto do estudo (o primeiro) e o participante que teve a conexão do experimento interrompida múltiplas vezes. Na Tabela 20 podemos verificar a análise estatística da duração.

Tabela 20 – Análise Estatística Descritiva da Duração da Sessão #3 (Condicionada Indireta Combinada).

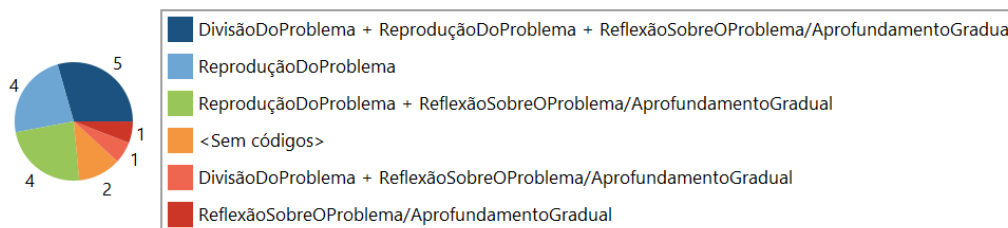
Métrica	Duração (H:M:S)
Mínimo	00:10:09
1º Quartil	00:18:02
Mediana	00:23:17
Média	00:25:10
3º Quartil	00:28:07
Máximo	00:49:52

Fonte: do Autor (2022)

4.4.1 Análise do Comportamento no Processo de Busca

- **Concepção da busca.** Os elementos de codificação da concepção da busca podem ser vistos por completo na Figura 6. Para as atividades da sessão corrente, identificamos a ocorrência da divisão do problema em partes em 12,37%, uma redução de 33%. Enquanto a reprodução do problema ampliou-se para 39,18%, e a reflexão sobre o problema regrediu mais para 48,45%. Isto se deve ao fato da atividade já conter tarefas definidas. Destacamos que 2 participantes optaram por não realizar as tarefas dessa sessão. Ao longo da realização da atividade, verificamos a combinação de estratégias pelos participantes de acordo com cada tarefa, como pode ser constatado na Figura 13:

Figura 13 – Combinação de Estratégias de Concepção da Busca por Participantes em Atividade de Múltiplas Tarefas.



Fonte: do Autor (2022)

A estratégia de divisão do problema foi utilizada para seguir a proposta de atividade e segundo o hábito do participante:

“De alguma forma eu procuro trabalhar com problemas menores e depois tentar juntar. Eu acho que busco conteúdo e ferramentas que me ajudem a montar essas partes. Essas páginas que salvei também são interessantes, e depois posso voltar e ler com um pouco mais de calma, para tentar entender melhor.” (Participante P31818).

Novamente observamos comportamentos de busca exploratória, desta vez para se familiarizar com o contexto e os conceitos envolvidos na necessidade de informação:

“[Rolando os resultados verticalmente] Tanto que aqui ele fala ‘normality test’. Então, digitei as palavras soltas, porque não lembro exatamente em inglês se o termo é ‘normal distribution’, e aí eu vi aqui que ele tem (...) Nessa imagem (do resultado da busca) mostra a curva então tem a ver com o que eu quero. Então está aqui, ele usou outro termo, então pego esse outro termo (‘normal’) e substituo (no lugar ‘normality’) aqui (query) para tentar ter algo mais apurado.” (Participante P21616).

“[Digitando ‘read dataset Google Drive’] [Clicou na busca] [Declarando que usará a linguagem R]. Fiz uma busca mais genérica primeiro. [Lendo os títulos dos resultados]. Já apareceu um resultado com R...” (Participante P21710).

Os casos de reprodução estão relacionados à busca direta, com o uso direto da pergunta da tarefa ou a sua versão interpretada traduzida no idioma inglês, como no relato:

“[Selecionando o texto da tarefa, copiou e colou na caixa de busca ‘Acesso e leitura aos datasets python’].” (Participante P21720).

- **Seleção dos termos e texto da busca.** Com base na codificação da Figura 7, observamos o uso da sugestão do mecanismo de busca durante a digitação da consulta (3 participantes), bem como a utilização de parte do título de resultados anteriores da busca (4 participantes) e a utilização de informação obtida na análise anterior de trilha da CQA (4 participantes):

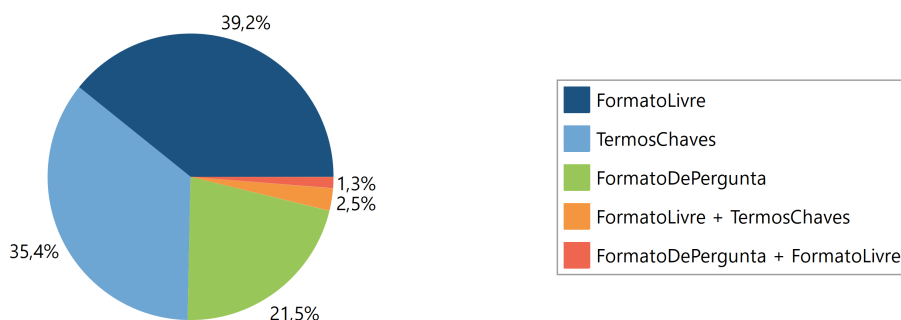
“Refazendo nova consulta 'mount csv from google drive 'r script'”. Não consegui. [Nova consulta sugestionada 'read csv file google-drive in R script'].” (Participante P21610).

“Copiou o título do 1º resultado "pearson correlation and significance" e acrescentou "in R programming". Agora percebeu que veio mais resultados na linguagem R desejada.” (Participante P21510).

“Então, essa aqui já sei que já posso usar o Shapiro test aqui. Se não me engano, faz a distribuição normal. Aí eu precisava ver como é o parâmetro que devo passar. Faria uma segunda pesquisa em cima de Shapiro aqui (marcando trecho do código). [Iniciando nova busca] Agora eu já sei qual é o comando que verifico se é uma distribuição normal. É esse shapiro test. [Digitando 'shapiro test how to use'] [Efetuou a busca] Aí eu veria como usar o shapiro test.” (Participante P21715).

Conforme observamos na Figura 14, houve ligeira predominância de construção textual livre (43,02% se combinadas com outras formas por participante) seguida pelo uso de palavras-chave. Destacamos que nenhum participante durante suas atividades alternou entre palavras-chave e o formato de pergunta nas tarefas.

Figura 14 – Combinação de Formato de Seleção dos termos por Participantes em Atividade de Múltiplas tarefas.



Fonte: do Autor (2022)

Curiosamente, tivemos 6 participantes que utilizaram um formato diferente para cada tarefa, mas também 2 (dois) outros que persistiram no mesmo formato (palavras-chave ou textual livre), que sugere um comportamento habitual:

[Analisando a demanda] [Digitando 'python, dataset, read'] É o que estou querendo fazer. (...) Terei que refazer a query. [Digitando 'python, dataset, gaussian distribution, colab'] (...) “[Lendo a demanda] [Digitando 'pyhon, dataset, correlation test, colab'].” (Participante P22015).

Exemplos de formato de pergunta:

“Escrevendo a query de busca 'how to access dataset in google driver in R'.” (Participante P21510).

“Então voltarei (no buscador) e incluirei um 'how to' [a consulta ficou 'correlation test in r how to']. Só para voltar uma resposta mais direta para o que quero.” (Participante P21710).

Exemplos de formato de texto livre:

“Escrevendo a query ‘normal distribution test in R’. Executando a busca.” (Participante P21510).

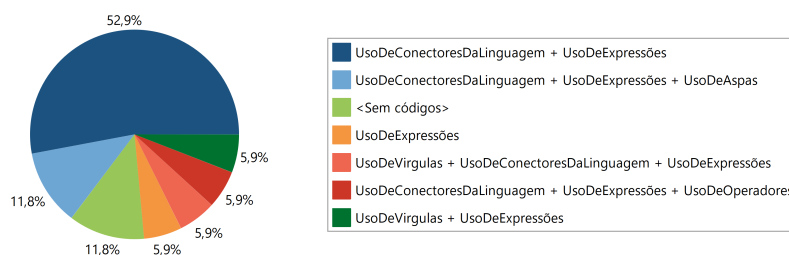
“Vou fazer a busca aqui [Digitando ‘correlation test in R’].” (Participante P22218).

Exemplo de formato de uso de palavras-chave:

“Isso já pesquisaria assim ‘Python correlation pandas’. Para tentar filtrar esses resultados.” (Participante P31818).

Os recursos para tentar direcionar o mecanismo de busca mais utilizados foram: 88,2% expressões, 76,5% conectores da linguagem/idioma, 11,8% aspas, 11,8% vírgula e 5,9% operadores/filtros de busca. Na Figura 15 verificamos as combinações utilizadas.

Figura 15 – Combinação de Artíficos de Direcionamento de Seleção dos termos por Participantes em Atividade de Múltiplas tarefas.



Fonte: do Autor (2022)

O uso da vírgula está associado ao hábito de um participante, enquanto o uso de conectores da linguagem está associado ao direcionamento da linguagem de programação (‘in python’, ‘in R’), recurso (‘in/from google drive’), formato de pergunta (‘how to’). Nesta sessão confirmamos o uso de tais recursos para refinamento dos resultados, restringindo o máximo possível à intenção de busca. As expressões foram massivamente utilizadas devido à natureza de algumas tarefas (‘distribution test’, ‘correlation test’, ‘read file’):

“(...) Quando o meu foco é alcançar uma plataforma de CQA. Embora aqui (CSE) você já tenha filtrado para mim. Para tentar alcançar elas (CQA), digamos que tivesse usando o Google sem os seus filtros (CSE) faço esse tipo de consulta, mais frasal. Isto porque quero alcançar (CQA). Se fosse outro tipo (de base), eu ia por AND/OR, dependendo da minha intenção e do tipo de resultado que estou esperando. Para forçar costume usar mais frasal.” (Participante P21820).

- **Formulação da consulta.** Os elementos estão codificados a partir dos relatos conforme a figura 8 com base no *ESKiP*. Na Tabela 34 apresentamos as consultas formuladas pelos participantes na sequência em que executaram, inclusive a ordem

de atuação nas tarefas. Relembrando que 2 (dois) participantes (P22423 e P22613) não realizaram tal sessão. Outros 2 (dois) participantes realizaram as tarefas fora da sequência proposta no enunciado. Apenas um participante explicitou dúvida quanto ter conseguido atingir a solução de uma das tarefas, relatando falta de conhecimento sobre o assunto.

Identificamos no processo exploratório da busca, o uso acentuado de realimentação da busca a partir dos resultados obtidos na formulação anterior, principalmente pela análise prévia da trilha CQA com coleta de algum conceito ou terminologia. Tais elementos eram inseridos na consulta, refinando-a. Também verificamos um uso acentuado de esclarecimentos adicionais a partir da releitura do enunciado durante a execução das tarefas, bem como do uso das sugestões do mecanismo de busca. Em alguns casos, sobretudo envolvendo as tarefas similares 2 e 3, os participantes se apropriavam de conhecimentos obtidos na tarefa anterior (#2) para formular a consulta na nova tarefa (#3).

- **Checagem dos resultados.** A figura 9 apresenta o quadro completo de códigos gerados a partir dos relatos. Tal qual na sessão anterior sobre 'Fibonacci' e 'complexidade algorítmica', novamente identificamos um aumento significativo de leitura sequencial de resultados da busca na página de resultados (93,3%), reforçando que tal comportamento está associado ao desconhecimento ou falta de domínio sobre o assunto. Nesse cenário, não se verificou a importância de análise da fonte (6,7%), a não ser pelo hábito de algum participante. Em todos os casos, houve 100% de análise do título, com aumento significativo da análise do texto resumo (46,7%) e do surgimento da análise dos recursos multimídias (imagens e figuras) visíveis nos resultados (26,7%). Também observamos a análise de similaridade com a intenção de busca (20,0%) explicitada pelos relatos após a análise dos elementos citados. Eis alguns relatos exemplificando tais comportamentos:

“[Digitando 'R read file in google drive']. Usando uma *string* em inglês, acredito que voltará alguma coisa. Voltou uma primeira aqui [Lendo o título]. Faz sentido porque não voltou muita coisa aqui (resumo), além disso, tem um código. Não ficou muito claro, mas parece que pelo título é o que estou procurando. [Selecionou para abrir]. Abrirei e tentarei entender aqui.” (Participante P22218).

“Nessa imagem (do resultado da busca) mostra a curva, então tem a ver com o que eu quero.” (Participante P22218).

“Peguei o segundo aqui porque ele já plota.(...) Estou vendo o segundo (resultado) primeiro. Você vê que o gráfico já me interessou. (...) Influencia na escolha (ter o ícone de gráfico no resultado) porque no caso do R especificamente, o que ele apresenta para ajudar na sua pesquisa é gráfico e tabela.” (Participante P21915).

“O título e depois a fonte aqui. A base é o *Stack Exchange*. O que faço aqui, como ele apareceu em primeiro, eu acredito que ele é um dos resultados mais relevantes, mesmo não sendo do *Stack Overflow* e eu preferindo o *Stack Overflow*, eu entraria para olhar aqui e encontrar uma resposta.” (Participante P21715).

- **Extração da informação.** Com base na figura 11, destacamos que não houve nenhum acesso aos resultados que não estivessem na primeira página de resultados. Isto se deve à realimentação da consulta, com já dito, a partir da qual se obteve os resultados anteriores de busca. Em 33,3% das vezes foram processados somente os primeiros resultados, enquanto em 40,0% os participantes avaliaram toda a primeira página. Em todos os casos (100,0%), o participante chegou ao melhor resultado após aplicar algum dos critérios de checagem. Cabe ressaltar que alguns participantes modificavam o entendimento somente após análise da trilha CQA resultante da busca.

Diferente das sessões de atividades anteriores, apesar de serem ainda tarefas de implementação, envolviam conhecimentos que não eram de domínio da maioria dos participantes (46,7% declaram sem conhecimento e 66,6% com pouco conhecimento sobre algum conceito associado às tarefas: (i) leitura de arquivo no Google Drive, (ii) distribuição normal, (iii) teste de correlação entre conjunto de dados). Por isso, observamos que o comportamento do participante no processo de busca mudou, por estarem fora da 'zona de conforto' (de conhecimento) deles. Alguns começaram pela extração de informação associado ao entendimento de tais conhecimentos previamente à extração da solução de implementação que era alvo das tarefas. Eis alguns exemplos de relatos de comportamento:

“Porque é um domínio/assunto meio incerto para mim. Não está na minha zona de conforto. (...) Prefiro refazer a consulta.” (Participante P21818).

“Com certeza. Me 'amarraria' em fazer isso (testes estatísticos de distribuição e de correlação). De fato me tiraria da zona de conforto.” (Participante P22015).

“No Google, seria a primeira coisa só para entender o que quê é. Eu acho que esse código pode servir. Tenho que entender melhor.” (Participante P31718).

“Já estou preso, não sei como usar CQA para isso. [Retornou ao CSE]. É (dificuldade) de conceito. Como saber como a resposta ali me indica.” (Participante P21820).

Com esta sessão de múltiplas atividades, confirmamos que o comportamento majoritário é de abertura em uma nova janela ou aba (93,3%). A estratégia exibição de múltiplos resultados concomitantemente em diversas abas ocorreu em 46,7% das vezes, com 7 (sete) participantes. A abertura na mesma janela, com o eventual uso do botão 'BACK' do navegador foi observado em cerca de 25% das ações, marcada

pela simplicidade da tarefa percebida pelo participante ou pelo hábito do participante (P21715).

“Lendo pelos títulos... [Abriu em outra aba]. Esse aqui parece interessante. [Abriu em outra aba]. Vou ler de trás para frente (ordem das abas abertas, resultados selecionados), porque esse aqui das páginas que abri pareceu interessante. Esse já é do *Stack Exchange*, que me lembre até agora não tinha aberto nenhuma página com ele.” (Participante P21820).

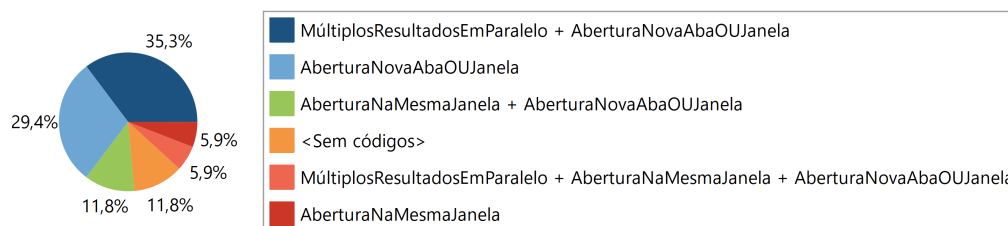
“[Lendo os resultados da busca pelo título]. Eu entraria nesse terceiro. Como eu disse, eu geralmente abro uns três (resultados). [Foi até o final da página um de resultados]. Bom, abri esses dois aqui.” (Participante P30217).

“Abrirei aqui esse resultado. Abrirei mais dois (resultados). O primeiro parece ser bem... a solução exata. Mas prefiro abrir mais algumas abas, para tentar pesquisar ou complementar informação. Tem 'dataframe', uma palavra também importante. Estou tentando ver as palavras-chave que tem no título. Eu não leio as informações abaixo (texto resumo).” (Participante P31818).

“[Lendo o título do 1.º resultado] Vou dar uma olhada nesse. [interpretando o conteúdo do título]. [Lendo o título do 2.º resultado, interpretando][Lendo e marcando trechos do hint]... De cara, eu acho que veria quase todos. Primeiro eu leria um por um para determinar qual que seria o que melhor me atenderia. Não o que é o melhor, mas o que melhor me atenderia. Nem sempre o que te atende é o melhor código.” (Participante P21915).

Na figura 16 apresentamos as combinações utilizadas pelos participantes ao longo da execução das atividades.

Figura 16 – Combinação de Estratégias de Exibição dos Resultados da Busca por Participantes em Atividade de Múltiplas tarefas.



Fonte: do Autor (2022)

Aplicando a análise de similaridade do comportamento dos participantes no processo analítico da busca obtivemos a Tabela 21.

Tabela 21 – Similaridade Comportamental entre os Participantes no Processo de Busca (Atividade Condicionada Múltipla).

Participantes	P21510	P21610	P21616	P21710	P21715	P21720	P21818	P21820	P21915	P22015	P22218	P22519	P30217	P31718	P31818
P21510															
P21610	0,67														
P21616	0,58	0,55													
P21710	0,71	0,53	0,62												
P21715	0,44	0,44	0,62	0,60											
P21720	0,50	0,44	0,48	0,60	0,45										
P21818	0,63	0,58	0,54	0,59	0,52	0,52									
P21820	0,56	0,48	0,61	0,67	0,59	0,46	0,46								
P21915	0,39	0,48	0,42	0,40	0,40	0,46	0,41	0,41							
P22015	0,42	0,47	0,46	0,57	0,50	0,50	0,50	0,38	0,33						
P22218	0,50	0,44	0,55	0,60	0,52	0,52	0,46	0,52	0,46	0,57					
P22519	0,57	0,43	0,48	0,71	0,45	0,61	0,45	0,52	0,52	0,43	0,61				
P30217	0,52	0,45	0,57	0,63	0,55	0,55	0,55	0,62	0,42	0,45	0,55	0,47			
P31718	0,64	0,53	0,48	0,68	0,45	0,60	0,52	0,46	0,46	0,50	0,52	0,61	0,55		
P31818	0,50	0,53	0,68	0,52	0,52	0,40	0,52	0,46	0,58	0,44	0,52	0,52	0,55	0,52	

Fonte: do Autor (2022)

Através do uso da estatística descritiva, conforme a Tabela 22 verificamos que o máximo alcançado foi 0,71 entre os pares de participantes P21510 com P21710 e P22519, enquanto o mínimo alcançado foi 0,33 entre P21915 e P22015.

Tabela 22 – Sumarização do Comportamento dos Participantes no Processo de Busca (Atividade Condicionada Múltipla).

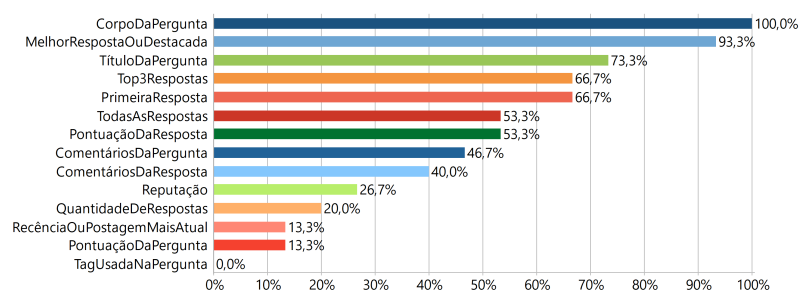
Métrica	Valor
Desvio Padrão	0,077873523833751
Variância	0,006064285714286
Mínimo	0,330
1º Quartil	0,460
Mediana	0,520
3º Quartil	0,570
Máximo	0,710

Fonte: do Autor (2022)

4.4.2 Análise do Comportamento no Processo analítico da trilha CQA

Após ter realizada a seleção dos resultados da busca que apontavam para postagens CQA candidatos à solução da demanda da sessão de atividades, os participantes analisavam as postagens da CQA que foram abertas em sua maioria em novas abas. Verificamos o foco de atenção dos participantes e notamos principalmente a análise do corpo/detalhe da pergunta com 24,7%, superando inclusive a análise do título da pergunta com 13,9% e da melhor resposta (ou aceita) com 22,1%. Assim, os participantes por não terem domínio do assunto, utilizaram a descrição da pergunta para obter novos conhecimentos e avaliar o alinhamento da trilha à necessidade de informação imposta pelas tarefas. Destacamos que todos os participantes avaliaram o corpo da pergunta, conforme a frequência apresentada na Figura 17:

Figura 17 – Frequência de foco em elementos da CQA por participante — Atividades da Sessão#3



Fonte: do Autor (2022)

Conforme já identificado na sessão de atividade anterior, a tomada de decisão na análise da trilha na CQA está associada à qualidade da postagem e na organização que permita analisar e recuperar a informação desejada. A existência de votação ou ‘pontuação positiva’ para a resposta e sinalizada com validade pelos pares, muitas interações e múltiplas respostas, e a reputação do respondente influem na percepção de qualidade da postagem. Destacamos alguns relatos:

“As imagens ajudam sim, mas elas não são o bastante para eu tomar ou não a decisão. Tenho que olhar o título. Faço aquela olhada, que diferente da outra (trilha), este aqui já tem código. Vendo aqui que já tem mais votos, tem mais resposta. Parece-me um pouco mais interessante. Então eu voltaria para verificar a pergunta da pessoa.” (Participante P21820)

“Olhei a reputação após ler a resposta. A resposta olhei primeiro e depois a reputação. A reputação foi para checar que a resposta foi dada por alguém que realmente conhece.” (Participante P21510)

“Estou rolando para baixo para ver se tem mais alguma coisa interessante, tem outras respostas, mas eu não vou agora não. Deixarei só esse primeiro aqui (resposta aceita), ele me pareceu muito bem organizado e tem formas de representar esses dados, que vale a pena investir nesse aqui também, para ver se posso aprender.” (Participante P31818).

Quanto à navegação por hiperlinks, externos ou não à CQA, novamente identificamos o uso pelo hábito ou pela necessidade de mais informação, ou validação do entendimento. Desta vez, 8 (oito) participantes manifestaram interesse em acessar links internos à trilha, dois deles por hábito. Os demais usaram esse recurso para ampliar o entendimento sobre o assunto da trilha.

“Vou direto para a resposta primeira (*marcada com aceita). [Observou o exemplo e o código]. Poderia usar esse exemplo. [Acessou o link da documentação com as informações sobre o método usado para ver os parâmetros].” (Participante P21510).

“Então, eu tentaria verificar se teria algum outro (questões relacionadas) que poderia auxiliar. [Abriu a página '*linked questions*'] [Lendo o título e resumos das questões relacionadas]. É que para poder efetivamente testar ou validar, é necessário ter um treinamento (do *dataset*). (...) Não tem nenhum deles aqui que seja específico apenas para teste.” (Participante P22015).

“Esse aqui (campo de texto da primeira resposta) forneceu um (link de referência) e eu vou abrir aqui (outra aba). (...) E ele já recomendou uma biblioteca, e como é de estatística e correlação tem tudo a ver com isso, queria dar uma olhada para ver outras possibilidades (para aprofundar o conhecimento). Para ver talvez o formato dessa sugestão, dessa função, o que me traria outras possibilidades, as entradas dessa função, o que poderia ter mais.” (Participante P21616).

Quanto ao processamento cognitivo da postagem, cerca de 20% dos participantes ou optaram por conteúdos em português, ou por traduzir para o idioma português para tornar o entendimento mais ágil e melhor do conteúdo.

“[Analisando a trilha *QAS* de conteúdo em português] Como ler os dados do meu Google Drive... [Lendo o corpo da pergunta].” (Participante P21720).

“Pera um pouco que deu uma travada aqui (*leitura em inglês). (...) Posso traduzir aqui. [Uso recurso de tradução de página do Navegador: EN-PT] Facilita bastante.” (Participante P21616).

“Isso aqui (trecho de código do resultado) é exatamente o que estou usando. Então esse aqui é mais interessante, porque ele está tratando como eu também. Claro que eu vou ler com calma o que ele está falando, se eu não souber ou entender o inglês dele, eu jogarei lá no Google Translate para me ajudar.” (Participante P30217).

Nesta sessão, a atividade era também sobre implementação de solução, onde 79,7% da análise na CQA tiveram esse propósito. Contudo, observamos 4,1% para busca por informação para uso e 16,2% para obtenção ou ampliação de conhecimentos. Novamente prepondera a necessidade de compreensão mínima do assunto antes de tratar o objetivo da demanda.

“Se é um assunto que eu não conheço, então tenho que entender melhor. Pelo menos minimamente o que é. Aqui está (resposta) falando como calcula e dando um exemplo. Eu começaria por esse. Uma vez entendendo esse, mais ou menos vendo como é o código, como está estruturado, eu pegaria algum outro exemplo, para contrapor, para ver mais ou menos como o outro fez. A partir daí, eu ia seguir e tentar juntar os dois. Eu precisaria trazer para minha realidade. Eu já sei como eu leio. Agora sei como calculo a correlação. Seria a mesma coisa que eu teria realizado no anterior.” (Participante P31718).

“[Olhando a trilha] É algo interessante, quero testar a distribuição normal, e entendi aqui é para plotar o gráfico e na visualização do gráfico [vendo código e gráfico da resposta aceita] ver se é distribuição normal. Nesse sentido, eu queria confirmar. . . Eu deixo essa busca pausada. . . Só para confirmar eu buscarei ‘gaussian distribution’ para ver o que é.” (Participante P21818).

Quanto ao comportamento da leitura da trilha na CQA, observamos conforme a Tabela 23 a predominância da leitura seletiva, com o participante buscando principalmente trecho de código na primeira resposta ou a resposta aceita, bem como no detalhe da pergunta, com o objetivo de encontrar correspondência com a sua intenção de busca.

“[Abriu a postagem] [Fez rolagem rápida até a melhor resposta]. Bom, aqui não está dando ‘match’ para a minha intenção de busca.” (Participante P21710).

“[Abriu a trilha]. Ah! aqui está mais direcionado.[Analisando o código da resposta] Nossa, esse aqui está muito melhor(. . .) Sim, adaptação, porque ele está falando (marcando segmentos do código). Aí faz o teste de Shapiro. Pronto. Então esse seria o teste.” (Participante P22519).

Tabela 23 – Estratégia de leitura da trilha CQA (Sessão #3)

Leitura	Quantitativo	Percentual
Inversa(Pergunta+inversaoRespostas)	0	0%
Transversal	10	14,9%
Seletiva	44	65,7%
Sequencial	13	19,4%
....TopDown	7	10,4%
....EmProfundidade	1	1,5%

Fonte: do Autor (2022)

Cabe ressaltar que esse comportamento parece ser influenciado pela combinação da base de resultados em CQA, perfil de desenvolvedor de sistemas majoritário dos participantes e o tipo de atividade proposto na sessão:

“Quando foco em comunidades técnicas, o meu objetivo principal é ver um código que seja confiável, que vá funcionar. A análise das respostas na comunidade, no meu caso, é a análise do código disponibilizado.” (Participante P21510).

Foi observado, um comportamento de leitura seletiva não visual, que se utilizou de recurso de procura de termos na trilha, minimizando o processo cognitivo associado:

“Faço busca do termo ‘Google drive’ na trilha. [Não achou. Retornou aos resultados da busca].” (Participante P21610).

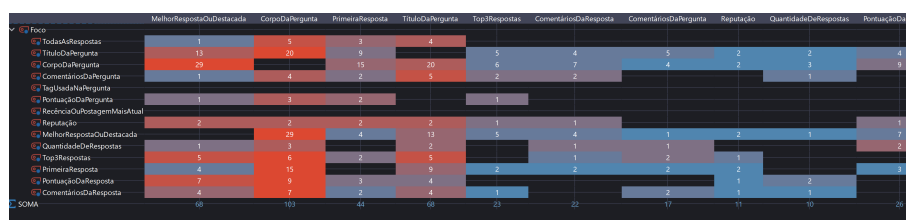
Observamos que nessa sessão as leituras em profundidade e sequencial estão associadas ao comportamento do participante em relação a um assunto de pouco conhecimento. Parece ser um comportamento de segurança, isto é, tentar extrair o máximo de informação da trilha antes de descartá-la.

“São os exemplos que as pessoas estão me dando. Além do código, os exemplos (imagens visuais). Essa discussão está 'muito boa'. É outro que iria favoritar, para depois voltar. Tem bastante coisa aí.” (Participante P31818).

“[Na trilha] Acho que ele vai me ajudar(...) Leio a trilha toda, verei o código (da pergunta) para comparar com o que faria. Aí leio a resposta que tem o 'tique' (resposta aceita). [Lendo a resposta aceita] Ele dá exemplo aqui. Darei mais atenção a essa resposta, mas, na verdade, já tem uma.(...) Antes de olhar os comentários, tentarei o que ele está fazendo aqui (resposta). Se não der certo, volto aqui e olho...” (Participante P30217).

Quanto ao foco, destacamos que o corpo da pergunta demandou mais atenção (88,2%) dos participantes nesta sessão que a melhor resposta (82,4%) e o título (64,7%). Na Figura 18 podemos ver as maiores combinações de foco para a análise da trilha.

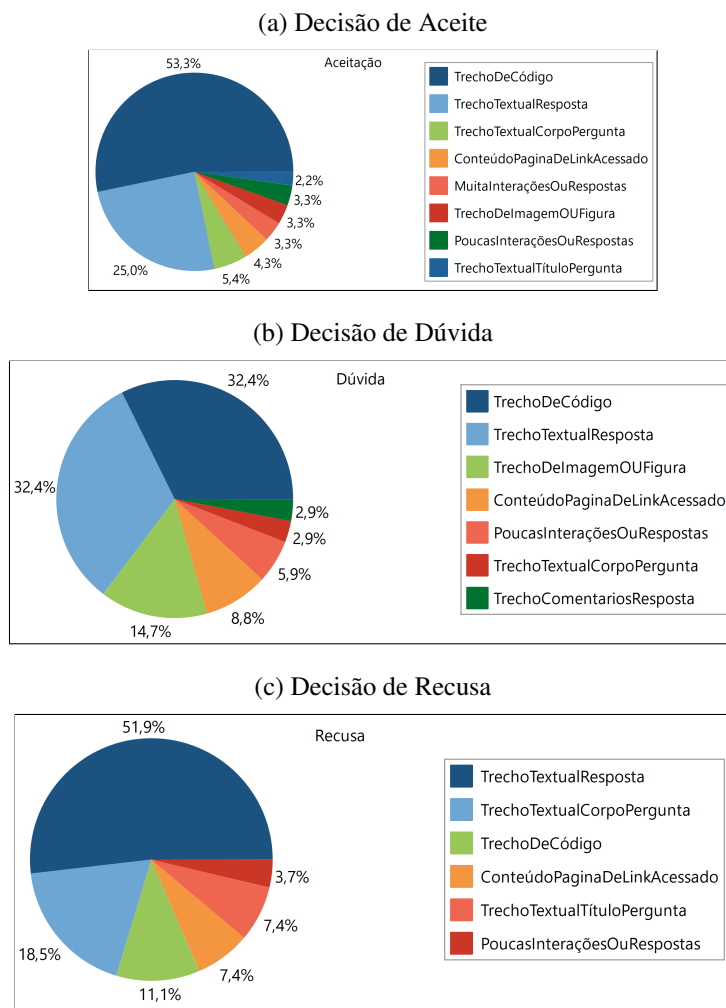
Figura 18 – Correlação de elementos de foco na análise da trilha pelos participantes (Atividade Condicionada Múltipla)



Fonte: do Autor (2022)

Quanto à tomada de decisão em relação ao aceite e uso da trilha, exibimos na Figura 19 a correlação entre a decisão e os elementos de evidência dessa definição. Observamos que os trechos textuais tanto da pergunta quanto da resposta são essenciais para a decisão, principalmente quanto à existência de trechos de código.

Figura 19 – Evidências de critérios de checagem utilizados para Tomada de Decisão. (Atividade Condicionada Múltipla)



Fonte: do Autor (2022)

Aplicando a análise de similaridade do comportamento dos participantes no processo analítico na CQA, conforme já efetuado anteriormente para o processo de busca, obtivemos a Tabela 24.

Tabela 24 – Similaridade comportamental entre os participantes na análise das trilhas CQA (Atividade Condicionada Múltipla).

Participantes	P21510	P21610	P21616	P21710	P21715	P21720	P21818	P21820	P21915	P22015	P22218	P22519	P30217	P31718	P31818
P21510															
P21610	0,53														
P21616	0,55	0,53													
P21710	0,48	0,37	0,48												
P21715	0,54	0,46	0,43	0,45											
P21720	0,52	0,55	0,52	0,44	0,50										
P21818	0,56	0,48	0,44	0,63	0,63	0,46									
P21820	0,72	0,60	0,67	0,42	0,52	0,50	0,44								
P21915	0,48	0,37	0,38	0,52	0,60	0,38	0,63	0,47							
P22015	0,59	0,38	0,39	0,41	0,46	0,50	0,42	0,51	0,36						
P22218	0,54	0,52	0,43	0,39	0,45	0,50	0,48	0,42	0,33	0,46					
P22519	0,56	0,43	0,39	0,48	0,41	0,46	0,58	0,48	0,55	0,48	0,55				
P30217	0,45	0,44	0,45	0,32	0,48	0,58	0,44	0,49	0,48	0,43	0,54	0,50			
P31718	0,37	0,41	0,37	0,45	0,61	0,57	0,47	0,33	0,45	0,35	0,45	0,40	0,55		
P31818	0,50	0,48	0,40	0,48	0,55	0,52	0,50	0,53	0,48	0,38	0,48	0,57	0,44	0,48	

Fonte: do Autor (2022)

Com o uso da estatística descritiva, verificamos pela Tabela 25 que o máximo alcançado foi 0,72 entre os participantes P21510 e P21820, enquanto o mínimo alcançado foi 0,32 entre P21710 e P30217. Verificamos que tivemos um aumento da similaridade no comportamento analítico nas trilhas CQA entre os participantes.

Tabela 25 – Sumarização do comportamento dos participantes na análise das trilhas CQA (Atividade Condicionada Múltipla).

Métrica	Valor
Desvio Padrão	0,076589146985049
Variância	0,005865897435897
Mínimo	0,320
1º Quartil	0,430
Mediana	0,480
3º Quartil	0,530
Máximo	0,720

Fonte: do Autor (2022)

4.4.3 Análise do Processo Cognitivo de construção da solução

A Tabela 26 apresenta os elementos extraídos dos relatos verbais dos participantes declarando o processo de construção da solução a partir dos resultados do processo de busca na Web sobre bases CQA. A organização da tabela se baseia no processamento analítico das trilhas da CQA e nos aspectos cognitivos para a solução.

Tabela 26 – Comportamento cognitivo para construção da solução — Atividade Condicionada direta.

Comportamento	Ocorrência	Aspectos cognitivos para solução
Uso de <i>mouse</i> para marcar partes de atenção na tela	16,7%	Controle e Recursos
Armazenar solucao para uso posterior	6,7%	Controle e Recursos
Usar favoritos do navegador para guardar a trilha CQA	4,4%	Controle e Recursos
Validação da solução com material externo	17,8%	Heurística e Recursos
Múltiplas avaliações e uso de várias respostas da trilha	15,6%	Heurística e Recursos
Uso direto da solução encontrada na trilha CQA	11,1%	Heurística e Crença
Uso preliminar ou inicial da informação da trilha CQA	20,0%	Heurística e Crença
Adaptação da solução encontrada na trilha CQA	24,4%	Heurística e Crença

Fonte: do Autor (2022)

A categorização do comportamento em função do aspecto cognitivo para construção da solução do problema apresentada na Tabela 26 pode ser melhor compreendida a partir de alguns relatos a seguir:

“Eu continuaria nesse aqui (trilha) mais um pouquinho. [Lendo as respostas seguintes] Eu tô tentando procurar aqui, eu não tô encontrando. Achei aqui (em uma das respostas seguintes). Esse é um que eu salvaria também nos meus favoritos, para depois voltar.” (Participante P31818).

“[Vendo a trilha] Aí volta a estratégia de leitura [marcou trecho de código da pergunta] (...) [Olhou a resposta aceita, direto no trecho de código] [Analisando e falando do problema relatado e da correção apresentada]. (...) Isso (leio título e descritivo da pergunta). Vou direto ao

trecho de código, porque nesse caso específico, eu já sei (conhecimento prévio da solução). Ele resolveu dessa forma (trecho da resposta). Então já [vou] guardar esse código em algum lugar e testar depois.” (Participante P31818).

“Me parece que tem a ver. (...) Primeiro acabo de dar uma olhada aqui, porque estava na minha frente, fui descendo e rolando a tela e vi que tinha a ver com leitura... Mas quando descí e vi isso daqui, me parece bem completo. Óbvio que a primeira coisa a fazer seria copiar (trecho de código da resposta) e jogar para uma área, abriria bloco de notas e colocar ele lá. Verificarei testando as bibliotecas dos elementos que tem se está coerente. Então vou de repente entrar em um lugar e pesquisar...” (Participante P31818).

“Usaria o código da pergunta, na verdade. O teste está bem claro aqui. Cruzaria com o resultado da resposta aqui (2.^a), para ter uma certeza do que é uma boa correlação ou não.” (Participante P22218).

“[Abriu a trilha]. [Lendo o detalhe da pergunta]. A pergunta está bem direcionada. Ele está buscando uma função... [Passou para resposta aceita] A primeira resposta parece que já servirá, porque já tem 202 votos. [Analisando o código apresentado] Já efetua de cara, apenas com uma linha. Atendeu né. (...) Dá (para adaptar) com certeza. A segunda resposta já é um método que já usei (ANOVA). Essa foi fácil.” (Participante P22519).

“Mas como já tenho conhecimento prévio, eu acho que essas respostas já cabem. Existem outras maneiras que eu sei. Então talvez eu voltasse para procurar essas outras maneiras. Eu ia fazer igual ao que está aqui, nessas respostas. Talvez eu utilizasse alguns elementos que ele colocou aqui, não necessariamente a resposta inteira. Estou unindo, utilizando mesmo mais de uma resposta. São os recursos que a ferramenta está promovendo e que está me permitindo fazer isso. Pego um pouco desse (resposta), pego um pouquinho daquele (outra resposta). Mas precisa ter conhecimento prévio para isso.” (Participante P21820).

4.5 Análise Comparativa entre CQA a partir de Processo de Busca na Web

Como as CQA possuem propósitos distintos, seja pelos assuntos abordados ou o formato utilizado, nesta seção exploramos comparativamente as CQA na percepção dos participantes em relação à estrutura, organização e conteúdo da CQA, para a finalidade de utilização e se servem para aprendizagem. Foi proposta a necessidade de informação a seguir, através de uma pergunta conceitual mais genérica proposta em inglês devido à existência de mais CQA nesse idioma, e para uso direto no mecanismo de busca:

“ENUNCIADO:

What are the benefits of 5G and how it will benefit IOT?” (Pesquisador).

Nesta atividade os participantes partiram da pergunta proposta pelo enunciado acima no processo de busca até a seleção das trilhas, com a obrigatoriedade de seleção de trilhas de CQA distintas. Percebemos que a ordem (ranqueamento) dos resultados do mecanismo de busca influencia diretamente na percepção inicial de relevância dos conteúdos pelos participantes. A mesma influência do mecanismo de busca foi detectada quanto à sugestão de consulta. Esse papel de influência do mecanismo de busca são destacados nos exemplos de relatos:

“A primeira resposta veio melhor ranqueada, então checarei essa primeiro. Se não me atender, checarei as duas respostas seguintes e o restante.” (Participante P21510).

“[Começou a digitar ‘what is’ e o veio a sugestão de pergunta] Olha aqui... já apareceu... está influenciando, hein!” (Participante P21820).

Destacamos que, diferente das atividades anteriores onde tínhamos questões de implementação de solução, a CQA *Stack Overflow* foi menos selecionada. Isto se deve à percepção majoritária que sua finalidade é para questões mais técnicas e menos conceituais. Em contrapartida, por ser a mais utilizada pelos participantes no dia-a-dia, as outras CQA eram avaliadas segundo o hábito e familiaridade que os participantes tinham com a interface visual do *Stack Overflow*, inclusive pontuando a semelhança que existente com a *Stack Exchange*, conforme alguns relatos destacados a seguir:

“Basicamente, essa estrutura que estou vendo do *Stack Exchange* e a do *Reddit*, que é um pouco parecido, eu usaria para coisas bem pontuais e para assuntos que conheço razoavelmente.” (Participante P22218).

“Se você está numa busca mais genérica, do tipo ‘quero entender o conceito’, você não entrará em um site CQA. Mas no caso do *Quora*, ele teve essa característica de mostrar aqui numa postagem justamente a parte conceitual. Primeiro falou ali um pouco das vantagens do 5G. Agora está fazendo uma correlação aqui com IoT. Eu acho que a própria intenção do usuário já não é buscar um site CQA, ele quer entender conceitualmente.” (Participante P21710).

“Vi três links (resultados) aqui, do *Stack Exchange*, do *Quora* e do *Reddit*. (...) Aqui no *Reddit* tem mais conversa, mais discussão. (...) Na *Quora* já apresenta perguntas que requer um conhecimento mais técnico, está indo mais para teoria também.” (Participante P21510).

“Para essas coisas mais informativas, que não envolvam código, eu já acho que esse tipo de estrutura visual (*Quora*) é melhor.” (Participante P21610).

“Aqui ele (*Reddit*) já mostrou de forma mais objetiva.” (Participante P21616).

“Já usei, mas não tenho costume de entrar (*Reddit*). Aí me perco e não lembro de cabeça. O *Stack Exchange* eu já sei direitinho onde fica tudo. Aqui eu não sei exatamente.” (Participante P21820).

“Esse aqui (*Reddit*) chama mais a atenção porque ele está bem aqui bem na vista. No outro (*Stack Exchange*) teria que olhar mais para cá que geralmente tem mais propaganda (lado direito da tela). Esse aqui (*Stack Exchange*) tinha isso (lado direito com informações relacionadas) mais eu nem me atentei para isso.” (Participante P21720).

Na tabela 27 verificamos as CQA selecionadas pelos participantes, o quantitativo de vezes relatadas e as principais considerações.

Tabela 27 – Comparação de uso de CQA e seus principais aspectos na percepção dos participantes.

CQA	Partic	Citações	Principais aspectos
<i>Reddit</i>	15	35	• Mais debate. Mais objetiva. Respostas mais superficiais. • Aspecto visual menos poluído, melhor ordenação dos conteúdos associados. • Não está nítida informações do usuário que está respondendo. • Layout com muito ícone, muita cor e dificulta a leitura. • Parecem apresentar assuntos desconexos. • Maior variedade de formas de ordenação. • É fórum. • Existência de propaganda no meio da trilha.
<i>Quora</i>	16	42	• Pergunta com descrição pouco elaborada. • Layout ajuda na leitura x ruim visualmente (gosto). • As respostas são mais trabalhadas. • Tem aspecto de ‘artigo’. • Existência de propaganda no meio da trilha.
<i>Stack Exchange</i>	13	37	• Similaridade de estrutura e organização do conteúdo com o <i>Stack Overflow</i>. • Prioriza a votação, deixa bem explícito.
<i>Stack Overflow</i>	8	14	• Respostas mais elaboradas. • Pergunta objetiva e destacada. • Prioriza a votação, deixa bem explícito.

Fonte: do Autor (2022)

Verificamos a partir de alguns relatos, a relevância da data da postagem em relação ao assunto, quando esse assunto na percepção do participante é sensível ao tempo, bem como a validação das repostas pelos pares através da votação, que dá confiança na informação. A qualidade do conteúdo é o mais relevante frente à intenção do participante quanto à necessidade de informação, apesar de ser impactada pelos elementos visuais e

a organização da informação na CQA. Podemos observar tais situações conforme alguns relatos:

“[Lendo a trilha do *Quora*] O que me chama atenção é que tem 313 respostas. Tem um número alto de *views* também. Tem 3 anos (a resposta) e abre até o que ele é (ficha de qualificação do usuário respondente), isso aumenta a credibilidade do que está escrito.” (Participante P21720).

“Basicamente, essa estrutura que estou vendo do *Stack Exchange* e a do *Reddit*, sendo um pouco parecido, eu usaria para coisas bem pontuais e para assuntos que conheço razoavelmente.(...) O *Stack Exchange* prioriza a votação, ele deixa bem explícito o valor, o que agregou às pessoas, tanto na pergunta quanto na resposta. Isso para mim, é positivo e usaria isso. No *Reddit*, isso não é tão explícito quanto no *Stack Exchange*. Isso, para mim, é o que teria mais valor. Essa interface do *Reddit* é mais colorida, um pouco mais visual, mas ela também tem um encadeamento de respostas. Mas prefiro trabalhar com a interface do *Stack Exchange*.” (Participante P22218).

“[Trilha da *Reddit*] [Traduziu para português] [Lendo a pergunta] Um ano atrás, então isso aqui é uma coisa mais recente. Achei que nessa comunidade as respostas são mais superficiais.(...) Já aqui (*Stack Overflow*) falou de uma forma mais completa. . . está tudo autocontido aqui. No máximo ele mandou olhar esse link externo aqui, mas ele já me deu respostas.” (Participante P30217).

“Ela (*Reddit*) é muito limpa (visual). Mas aí vi o texto relacionado, parece que me atraiu no primeiro momento pela limpeza, mas aí depois quando fixei mais o olhar, não aparece nada que me interessa. (...) Aqui está me parecendo assuntos soltos.” (Participante P31718).

“Visualmente todos os dois (*Stack Exchange* e *Reddit*) tem a informação distribuída verticalmente e isso ajuda muito. As explicações poderiam ser melhores. Costumo ver as explicações melhores no *Stack Overflow*, que tem a estrutura mais parecida com o *Stack Exchange*.” (Participante P31818).

Na perspectiva de aprendizagem, alguns relatos destacam aspectos relevantes na CQA, como a indicação do *feedback* de aceite da resposta, a qualidade da postagem que permite uma rápida assimilação e a existência de *links* para aprofundamento, que devem ser ponderados em função do assunto e objetivo da postagem:

“O *Stack* eu gosto. Como busca para aprendizado, eu gosto do *stack* (referência ao *Stack Overflow*), com essa possibilidade de lhe indicar ... olha isso aqui é o que o pessoal mais gostou de todo esse assunto. Gosto dessa proposta do *stack*.” (Participante P22015).

“É a resposta certa porque certamente as pessoas preferiram essa resposta, quem está navegando aqui. Então eu vou ler a resposta. . . talvez eu se eu tivesse muito interessado em saber o que é isso, eu entraria nesse link. Mas como é um artigo de opinião, eu não seria influenciado pela maioria (resposta aceita), certificado como positivo.” (Participante P30217).

“Esse aqui está mais interessante, ele contextualizou, mostrou mais aplicações. Isso facilita o entendimento de quem está começando. Ele deu a definição de alguns termos, isso dá vontade de clicar e ver mais. Esse seria um que eu salvaria como favorito, apesar de ser uma linha só tem as definições, então já é um ponto de partida.” (Participante P31818).

“Para esse tipo de pergunta, eu tenderia a ficar entre o *Stack Exchange* e o *Quora*. Tenderia a tratar esse tipo para respostas mais teóricas. Daria mais atenção a respostas mais teóricas. Buscando informações sobre o tema do que demonstração.” (Participante P21510).

“[Trilha *Stack Overflow*] Essa gostei porque a pergunta foi bem objetiva e destacada. Então me chamou a atenção. Aqui vem a resposta e embaixo os comentários. Então essa seria uma que exploraria bem. Se eu quisesse, por exemplo, montar um texto para falar sobre o assunto, então eu iria nesse caso aqui. Mas depende do objetivo, se for para entender rápido eu iria no outro que tem mesmo informação. Se for para escrever sobre o assunto, eu viria nesse.” (Participante P22423).

Quando avaliamos qual o perfil mais adequado do usuário em um cenário de aprendizagem com base em trilhas de CQA, os relatos indicam que seria restritivo para iniciantes no assunto, mesmo que estejam interessados no assunto. Ainda, reforçam a necessidade de outro tipo de base para prover o conhecimento prévio necessário antes de se utilizar a CQA.

“Olha, eu não sei se esse tipo de Q&A me ajudaria desde o início (iniciante). Porque estou estudando lá um assunto que eu não conheço nada. Talvez pesquisarei uma coisa mais robusta sobre esse assunto em outros caminhos. Depois voltaria para alguma coisa mais detalhada. Talvez seria esse tipo de busca, como o *Stack Exchange*, por exemplo.” (Participante P22218).

“[Vendo os resultados] Por ser um tema que eu não tenho muito conhecimento o que possivelmente eu buscaria não seria necessariamente começar por CQA. Talvez eu entraria um vídeo do YouTube para primeiro entender um pouco de 5G e aplicação dele para IoT. Depois, se tivesse uma dúvida específica entraria em um site CQA. Teria que ser outro tipo de base. Acho que para usar a comunidade você já tem que ter um conhecimento prévio e apenas dúvidas pontuais.” (Participante P22218).

“[Trilha do *Reddit*] Esse aqui eu já li a pergunta. Interessante que ele está perguntando referências sólidas, que foi exatamente o problema que encontrei lá no outro. No outro eu não encontrei referências. É para o pessoal mais especializado. Ficou muito vago. Deixa eu voltar no título. Esse link (de uma das respostas) me parece interessante, mas o restante (não). (Participante P31818).”

Quanto ao alinhamento das CQA à intenção dos participantes, há uma preponderância quanto à busca de como fazer em contraponto ao que fazer, com o entendimento mais voltado ao uso para resolver problemas, principalmente de implementação técnica:

“*Stack Overflow* eu acho que é interessante para resolução de problema, enquanto o *Quora* pode ser mais nessa linha de entender conceitualmente.” (Participante P21510).

“Quando estamos falando de linguagem de programação, eu uso muito. Porque eu sempre busco ali um comando, uma estrutura de alguma coisa. Algo mais pontual. No *Stack Overflow* eu não busco ‘o quê’. Aqui estou buscando ‘o que’ (benefícios 5G). Numa estrutura como essa busco ‘como fazer’. Eu já sei o quê que é, mas como faço isso? Mas quando eu não sei o quê que é, ou não domino o assunto, eu não busco nesse tipo de base (CQA).” (Participante P21715).

4.6 Análise da Percepção dos Participantes sobre a relevância de trilha CQA para suporte à Aprendizagem

Nas seções anteriores, analisamos aspectos comportamentais do participante durante o processo de busca na Web e o processo analítico na CQA. Passamos também pela análise de características entre CQA na percepção pelos participantes de suas finalidades e usos. Além de alguns elementos existentes nas CQA, cada uma com sua forma própria de organização, que conferem maior credibilidade ao conteúdo, a noção de qualidade da postagem (pergunta e resposta) é crítica para a percepção dos participantes quanto ao possível uso das trilhas para suporte à aprendizagem, conforme os últimos relatos citados na seção anterior. Com base nisso, podemos divisar dois enfoques quando à necessidade de informação: (i) ‘o que’, do espectro conceitual ligado à absorção de informações e ampliação de conhecimento; e (ii) ‘como’, de espectro de aplicação ou técnica para implementação e uso.

Para avaliar as trilhas de uma CQA como candidatas a material e seu uso no contexto de aprendizagem, considerando que o aspecto de qualidade da pergunta é crucial, escolhemos a CQA *Stack Overflow*, de maior familiaridade ao grupo de participantes que são da área de sistemas de informação. Nessa CQA selecionamos duas trilhas de perguntas sobre a linguagem de programação C, a primeira (distinção de operadores) envolvendo entendimento de conceito e segunda sobre implementação (tratar caracteres em matriz). Essa seleção das perguntas foi baseada no trabalho de [Ravi et al. \(2014\)](#), que explora a qualidade da pergunta e introduz essas perguntas de 2009 como parâmetro. Como apresentado naquele trabalho que concluiu que as perguntas de maior qualidade continuam atrativas ao longo do tempo, a primeira pergunta tem menos visualização, mas tem mais votação que a segunda. Isto denota que um indicador da popularidade da pergunta na comunidade pode não ser um reflexo fiel da qualidade da pergunta, que a votação positiva para uma pergunta significa que a questão “mostra esforço de pesquisa; é útil e claro.” ([RAVI et al., 2014](#)).

Na Tabela 28 apresentamos a magnitude das métricas votação e visualização, com os dados de 2014 (RAVI et al., 2014) atualizados para Dez/2021 apurados no presente trabalho, com a evolução durante esse período. Ao considerarmos tais métricas com nossas análises anteriores, observamos que raramente os participantes mencionaram o quantitativo de visualizações. Por outro lado, com frequência relataram o impacto e relevância da votação tanto da pergunta quanto da resposta, sendo essa última a principal.

Tabela 28 – Magnitude de validação e popularidade das perguntas — Entre 2014 e 2021.

Documento/Trilha	Visualização			Votação		
	Em 2014	Em 2021	Crescimento	Em 2014	Em 2021	Crescimento
Why does C have a distinction between -> and .?	23,0k	70,0k	204,05%	1	3	200,00%
Putting user input into char array (C Programming).	1,5k	2,0k	33,33%	44	55	25,00%

Fonte: do Autor (2022)

Quanto ao domínio ou familiaridade com a linguagem C, 47,08% dos participantes disseram ter pouco ('enferrujado', 'só estudaram', 'faz muito tempo'), 23,51% nenhum conhecimento. Somente 29,41% declararam ter conhecimento razoável a bom.

4.6.1 Análise de Trilha conceitual ('O QUE') para suporte à Aprendizagem

Para avaliarmos se tópicos mais conceituais apresentados em formato de perguntas em CQA são relevantes para os participantes como fonte para aprendizagem, reapresentamos a primeira questão conforme a seguir e o fragmento de sua postagem na CQA *Stack Overflow* apresentada na Figura 20.

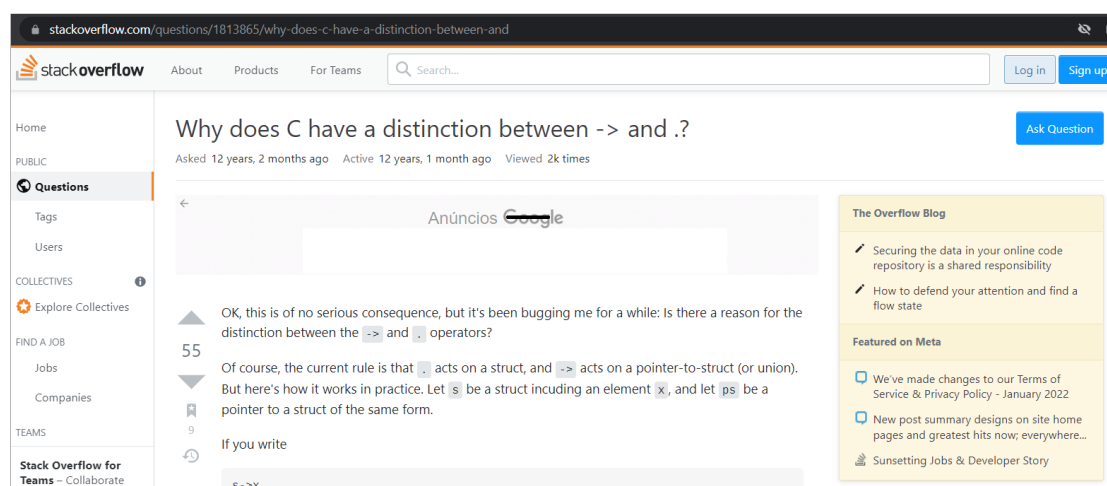
“1.ª QUESTÃO: Why does C have a distinction between -> and .?”³
(Pesquisador).

Alguns elementos presentes no site da CQA são inadequados para o aprendizado, tais como propaganda, vagas de emprego, trabalhos remotos e '*blog do Overflow*'. Eles não conferem contribuição para o melhor entendimento do assunto tratado na trilha. Devido a essa existência, os elementos mais relevantes para o aprendizado não possuem posição de destaque na visualização da trilha. Esse é o principal elemento visual compartilhado pelos participantes através do relato, conforme o exemplo a seguir:

“Com relação a isso aqui (quadro direito superior), não tem nada a ver com que estou querendo aprender sobre C. Acho desnecessário ter um destaque tão grande na interface, para falar do *Blog do Overflow*, de trabalhos remotos. Isso aqui (quadro direito inferior, 'linked' e 'related') para um aprendizado é mais relevante. Esse tipo de conteúdo para aprendizado deveria ter um destaque maior. Então essa área direita do site chama muito a atenção. Então para aprendizado, isso deveria ficar para baixo, se esse fosse o foco realmente do *Stack Overflow*, e essas

³ <https://stackoverflow.com/questions/1813865/>

Figura 20 – Fragmento da Postagem (1.^a QUESTÃO) sobre conceitos da Linguagem C



Fonte: do Autor (2022)

coisas relacionadas ('linked' e 'related'), bem legal essa diferença, de pessoas que estão ligando outros conteúdos para te ajudar a entender e os relacionados também.” (Participante P30217).

Quanto à confiabilidade da informação existente na trilha para utilização para aprendizado, sobretudo para os iniciantes no assunto, a existência de respostas aceitas e o ranqueamento das respostas por pontuação parece suportar isso. Eis alguns relatos:

“Pode influenciar, porque que outras pessoas que conhecem, e eu não conheço, não trabalho com isso e estou querendo me aperfeiçoar, e escreveram a melhor resposta, então eu olharei. Quando eu já tiver maturidade no assunto, pode ser até que discorde, mas enquanto não tenho maturidade no assunto, eu olharei, porque aquilo me ajudar.” (Participante P31718).

“Isso para mim já é positivo. Quando se chega ao ponto de votarem, de indicar que está legal, isso já está de bom tamanho. Passa (confiança), sem sombra de dúvida.” (Participante P22015).

“A pontuação ajuda. Não está fazendo juízo de qualidade. É interessante ter esse juízo de quantidade. Essa ordenação do *Stack Overflow* para mim, é perfeita. Porque me dá um juízo de qualidade, através da marcação da melhor resposta e como a primeira, que não necessariamente tem a maior quantidade (pontuação).” (Participante P21510).

Enquanto a pergunta detalhada e as respostas bem elaboradas parecem contribuir para a aprendizagem do conteúdo existente na trilha, os comentários parecem ser irrelevantes e dispersam o entendimento sobre o assunto:

“Os comentários aqui (*Stack Exchange*) nem sempre são para agregar. Por isso acho legal ter a distinção entre resposta e comentários, que não é o que vi em outra ferramenta.” (Participante P21820).

“Sinceramente, eu acho que isso (comentário da pergunta) polui mais que outra coisa. Porque eu acho que tinha que ter um botão dizendo assim, quer ver a resposta, os comentários. Isso incomoda, porque se eu quiser ver comentários, eu clico e ele abre.” (Participante P31718).

“Já é um comentário relacionado à pergunta. Agrega alguma coisa. Penso que as duas respostas são complementares, apesar da outra (a primeira) está melhor estruturada. [Analisando os comentários da segunda resposta] São comentários que não acrescentam muito. Engraçado que essa segunda resposta teve mais engajamento (37 votos) que a primeira (aceita, 10 votos)” (Participante P21610).

“Ela é muito mais discursiva, trazendo para a discussão do que simplesmente uma resposta direta. Ele explica o que é. Isso é legal quando tem pessoas discutindo, debatendo, é exatamente essa ideia de poder não concordar... e ver isso numa linha só, eu acho que é muito importante para o aprendizado. Isso (nesta trilha) proporcionou que a discussão fosse mais elaborada, que tivesse mais discussões. Como é algo mais conceitual, que demanda mais pensamento, mais análise, teve mais discussão, mais comentários. Quando se apresenta dessa forma, eu acho mais interessante para o aprendizado de conceitos, de algumas questões que se tem que não é uma resposta tão direta.” (Participante P21820).

A postagem foi considerada inadequada para aprendizado para iniciantes no assunto. Isto sugere que faltam elementos na trilha que categorizassem o nível de conhecimento requerido para aquele conteúdo, não somente a categorização de tópicos (*tags*).

“Esse aqui eu acho que não... Eu acho que não serve para *newbie* aprender, teria que ter mais conhecimento.” (Participante P21616).

“Antes mesmo dele perguntar... ele discute tudo... quando que usa cada um. Já serve para efeito de aprendizado sobre o assunto. Contudo, depende do usuário que estiver acessando isso aqui, porque se ele seguir, tem umas discussões que para um usuário primário, que está aprendendo, que talvez fizesse perguntas, talvez não vá compreender algumas coisas relacionadas aos padrões (C).” (Participante P21820).

“O que tá me ajudando a compreender o conceito são as respostas. Então essa primeira resposta aqui, mais qualificada do que as outras... Assim, para aprendizado talvez eu lesse essa discussão sim, mas não os comentários.” (Participante P30217).

A dificuldade de alinhamento da intenção de busca com a avaliação da pergunta parece ser mais agravada quando se trata de perguntas mais conceituais, que envolvem a necessidade de conhecimentos prévios:

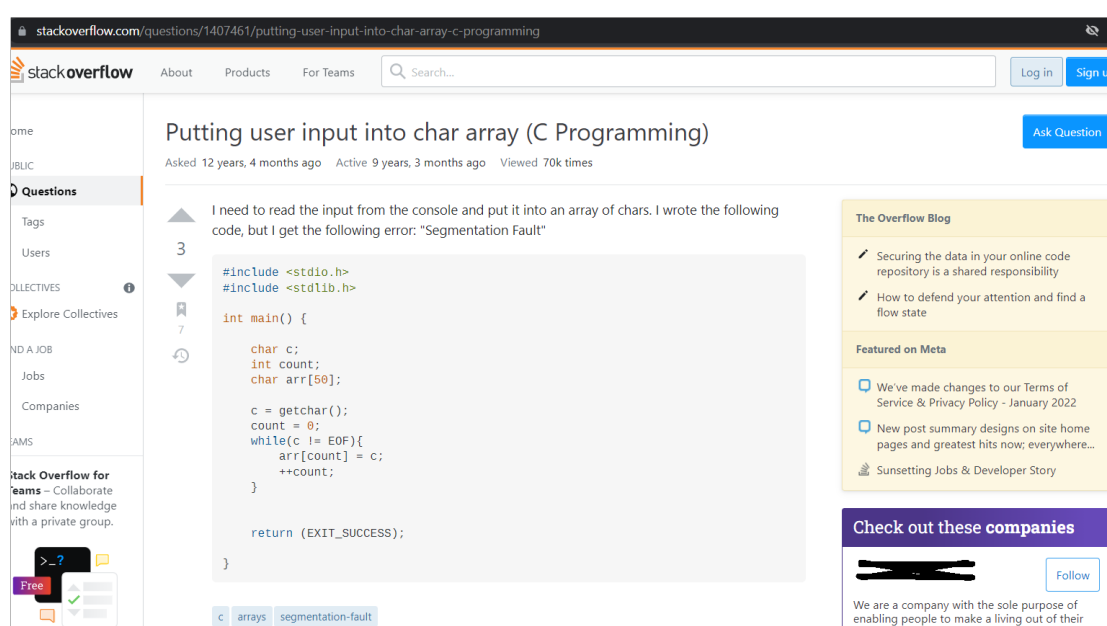
“Nas outras perguntas quem formulou a pergunta fui eu, e eu estava buscando perguntas similares a minha. Nesse caso aqui como eu não formulei a pergunta eu preciso entender melhor o que ele estava pensando quando formulou a pergunta. Nesses casos eu abriria outra aba para começar outra pesquisa para tentar entender a estrutura em C. Para mim, particularmente falta um conhecimento anterior, mas a explicação está boa.” (Participante P31818).

4.6.2 Análise de Trilha de implementação (‘COMO’) para suporte à Aprendizagem

Para verificarmos se tópicos mais técnicos e pontuais apresentados em formato de perguntas em CQA são relevantes para os participantes como fonte para aprendizagem, reapresentamos a primeira questão conforme a seguir e o fragmento de sua postagem na CQA *Stack Overflow* apresentada na Figura 21.

“2.^a QUESTÃO: Putting user input into char array (C Programming).⁴”
(Pesquisador).

Figura 21 – Fragmento da Postagem (2.^a QUESTÃO) sobre implementação em linguagem C



Fonte: do Autor (2022)

Como o conteúdo da resposta foi elaborado e organizado, com código e explicação, além de abordar alguns conceitos associados, despertou interesse e satisfação nos participantes que são assíduos usuários da CQA. Isto denota a percepção geral de como seria uma postagem e a finalidade da CQA, além de sinalizar possíveis caminhos (organização do conteúdo, atenção aos conceitos envolvidos, possibilidade de saber mais através de *links*) para uma trilha mais alinhada à questão de aprendizagem dos usuários:

“Nós queremos encontrar uma resposta, que nos dê literalmente um caminho, que permita fazer um teste, já nos dê o resultado. Então fazemos avaliações e análises sobre aquilo ali. Gosto do jeito que está sendo colocado aqui, porque ele é bastante objetivo.” (Participante P22015).

⁴ <https://stackoverflow.com/questions/1407461>

“Normalmente quem recorre ao *Stack Overflow* quer resolver um problema específico. Não está ali para aprender. Quer resolver um problema prático. Uma vez resolvido o problema, ele não falará, nossa, deixa eu ver aqui e dar uma olhada aqui, e ver se esse item deve ser legal. Normalmente o usuário está no trabalho ou fazendo um trabalho de faculdade. Tentando resolver aquele problema, não está ali para ampliar aqueles conhecimentos.” (Participante P21610).

“Ele é mais didático e mais rápido. Aqui inclusive, eu não costumo ver dessa forma não, ele colocou o código primeiro e depois deu a explicação.” (Participante P31818).

“Aí a pessoa já respondeu corrigindo código dele, o que é maravilhoso. Agora ele colocou a resposta certa. Legal que colocou em cada linha o que a pessoa fez errado. Está bem didático esse aqui. Elaborou o código certo. . . Está dando bastante conhecimento técnico.” (Participante P22519).

“Teve uma coisa que ele editou aqui. Isso é legal. Basicamente essa pessoa dá a resposta que a outra pessoa pede, mas explica alguns conceitos que talvez tenha passado. Isso ajuda a pessoa a compreender o porquê dela estar errando, não só simplesmente dar a resposta. Pode ser utilizado para aprendizado também. É mais dar resposta e falar de alguns conceitos. Talvez fosse interessante a pessoa procurar mais, além do que está só aqui, para entender os conceitos.” (Participante P21820).

4.6.3 Comparação entre duas trilhas conceituais e de implementação para suporte à aprendizagem

Ao final da sessão de análise das trilhas, os participantes foram estimulados a refletir sobre o uso de ambas como material de apoio para aprendizagem, se era possível usar com material de base para aulas em qual seria a finalidade, tais como laboratorial/técnica, implementação ou conceito. Conforme a Tabela 29 e alguns exemplos de relatos, podemos verificar distinções.

“Acho muita postagem assim, perguntas conceituais, que tem alguém respondendo detalhes. (...) Depende do propósito de quem está buscando. Às vezes no sentido de buscar uma boa solução para um problema e você encontrar postagens como a anterior (Questão B) com mais um debate e soluções arquiteturais, grandes argumentos que tem os trade-off entre uma situação e outra.” (Participante P21818).

“Aqui (Questão B) um júnior, programador júnior, por assim dizer, ele já sairia do outro lado. É a cara de quem está desenvolvendo software, código. (...) Isso aqui passaria conhecimento para ele indiscutivelmente. Aqui (Questão A) esse conteúdo, complicaria a vida dele, porque não é o que ele está procurando. Ele já precisaria de um pouco mais de know-how. Ele teria que quebrar muito a cabeça com código, para inclusive ter vontade de chegar e fazer uma análise conceitual do que está por trás de algo aparentemente simples.” (Participante P22015).

“Essa primeira resposta (da 1.^a pergunta) serve como material de apoio. Se tiver que aprender um tema... Se bem que até fala de array e do limite. Mas não trouxe nenhum conceito mais robusto. Eu acho que ele entraria como material complementar. Esse, ao contrário do outro não precisa de um conhecimento mais aprofundado, pode ser usado em um público mais novo no assunto. Ele explica o código, mas não explica nada do conceito. Está bem prática mesmo, e o outro é mais conceitual.” (Participante P31818).

Tabela 29 – Considerações dos participantes sobre uso das trilhas para aprendizagem

Pergunta/Trilha	Considerações dos participantes
Why does C have a distinction between <code>-></code> and <code>.</code> ?	• Para alguém com mais experiência, um usuário programado ‘senior’, um especialista. • Conceitual: conceitos mais avançados, aspectos e teoria de linguagem de programação, incluindo conceitos de compilação. • É um debate: conceitos, sintaxe da linguagem versus semântica. • Falta de exemplos.
Putting user input into char array (C Programming).	• Para alguém com nível mais básico, usuário programador júnior. • Prática: parece aula de programação 1 e 2 de graduação. • Essa resposta com certeza agrega. Isso dá uma aula de C, colocou o código primeiro e depois deu a explicação. • O conteúdo está bem estruturado, organizado. Mais didático e rápido. • É o tipo de pergunta mais comum em CQA (<i>Stack Overflow</i>): abordagem direta com código pronto e explicado.

Fonte: do Autor (2022)

A percepção majoritária capturada a partir de todos os relatos coletados e codificados, é que a CQA teria relevância de uso para busca por assuntos mais técnicos e diretos, como o caso de implementações pontuais de solução técnica, algo de aplicação mais prática.

“Dentro do *Stack Overflow*, eu busco por coisas técnicas. Eu com certeza faria uma busca desse tipo da 2.^a questão. Não só sobre esse assunto, mais sobre coisas assim, diretas e técnicas. É pontual e direto. A outra (1.^a questão) eu não faria no *Stack Overflow*, efetuaria no Google (Search) e procuraria alguma outra base para entender. Para coisas teóricas complexas e grandes, de assuntos de discussões grandes, eu acho que o *Stack Overflow* não cairia bem. Ele é juntamente para pessoas que tem o conhecimento médio para baixo.” (Participante P22218).

“A 2.^a está melhor aqui, para novatos mesmo, mais básica. A 1.^a pergunta ficou muita discussão filosófica. Por exemplo, para a pessoa que tem uma dúvida de como usar na prática, não ficou muito claro. Mas a 2.^a está bem estruturada. A pessoa colocou mesmo para iniciante.” (Participante P22218).

Visto que as CQA permitem a interação e debate através dos comentários, para ajudar a compreender e melhorar tanto a pergunta quanto a resposta, verificamos com os partici-

pantes o impacto disso e a relevância no contexto da aprendizagem do usuário, bem como que outros recursos poderiam contribuir:

“Eu acho que (os comentários ajudam) sim, porque pode complementar, te dar outra opção de fazer. Na 2.^a e não na 1.^a pergunta. Tanto que os comentários que estão aqui (1.^a) não ajudaram em nada. Já os links poderiam ajudar se estivessem melhor posicionados em tela. Na 2.^a pergunta não influiria.” (Participante P30217).

“Acho importante sim, a existência de links para saber mais. Em geral, o que faço é isso, vou abrindo em várias abas e vou lendo gradualmente(...) Os links do lado direito eu nem li. (...) Não vi nenhuma questão que me ajudasse com isso aqui não. Ele (2.^a) serve como material de apoio, mas não entrou em nenhum conceito mais robusto.” (Participante P31818).

4.7 Avaliação de percepção de aprendizagem pelos participantes durante as atividades realizadas

Imediatamente após a realização de todas as sessões de atividades pelos participantes, realizamos uma entrevista com o uso da técnica de *Intense Interviewing* com base nas questões-guia apresentadas na Tabela 2. Na entrevista, tratamos a percepção dos participantes quanto a: (i) limitações e impactos do ambiente tecnológico disponibilizado para as sessões práticas, (ii) a organização e finalidade das trilhas CQA, (iii) possível utilização de CQA para suporte à aprendizagem a partir da análise de três trilhas de exemplo, e (iv) contribuição das sessões para ampliação de conhecimento dos participantes.

Quanto ao ambiente simplificado e customizado de busca CSE, a partir do *Google Search*, apresentamos na Tabela 30 as considerações dos participantes frente ao que está habituado a utilizar ou que percebeu a falta ou inadequação no ambiente. Ressaltamos que nenhum participante se sentiu impactado no comportamento em função das características do mecanismo de busca oferecido.

“Não foi. Porque o Google começou assim. A primeira versão do Google era um quadrado e uma lupa. Não era nem a lupa, eram duas perguntas, estou com sorte e procurar.” (Participante P21915).

“Nem um pouco. Aliás, assim está menos poluído. Ele está super objetivo. É óbvio que tem essa questão dos Ads, que não tem jeito. Se pudesse eliminar isso, seria ótima essa questão da propaganda. Excetuando-se isso, está joia, bastante objetivo, não é infundável, deve trazer aqui umas 15 respostas, as mais relevantes. Eu gostei.” (Participante P22015).

“Achei (CSE) estranho no início, porque não tem a figura, apenas a caixinha e a lupa, mas não atrapalhou não.” (Participante P22613).

“Os elementos que utilizo as palavras-chave, olhando o título, tudo tinha lá. Continua com a forma vertical e continuei rolando, tem a paginação, então não me prejudicou. Teve a correção da palavra que escrevi errado, principais ferramentas que utilizo estavam presentes.” (Participante P31818).

Tabela 30 – Considerações dos participantes sobre o uso e limitações da CSE

Limitações da CSE	Ocorrências	Considerações dos Participantes
Visualização funcional básica	9	• Interface limpa. • Tem a caixa para buscar, o ranqueamento e as informações que o Google já dá normalmente. • Tem a sugestão do mecanismo de busca. • Informações em tela estão em inglês.
Caixa de digitação da consulta limita a visualização, pequena	7	O campo de busca deveria ser maior para ver toda a informação na tela.
Somente permite visualizar resultados textuais	1	Não apresenta resultados de imagem, vídeo ou outros tipos de conteúdo.
Ausência dos ‘cards’ de resposta direta	1	Que repondem direto, sem que você tenha que acessar a plataforma em si, que o próprio Google te responde.
Sem painel Wikipedia	1	A ficha técnica que tem uma descrição rápida
Sem pesquisa por período ou informação de data	1	Solução de implementação depende da versão da linguagem.
Sem marcação do link de resultado já visitado	1	É importante para controlar o que já visitei.
Sem integração com a barra de navegação	1	Uso a pesquisa na barra de navegação que facilita bastante.

Fonte: do Autor (2022)

Um ponto de destaque, mencionado por apenas dois participantes, é que a consulta e o resultado está sempre no mesmo idioma.

“Eu acho que tem, por exemplo, no meu caso, aquela primeira busca que fiz, que não apareceu nada e quando que mudei a ‘string’ para inglês para ter mais material (resultados).” (Participante P22613).

Em relação à limitação dos resultados da busca às bases CQA, apenas um participante declarou que teve que modificar seu comportamento, no modo como elabora a consulta:

“A visual não, mas saber que eram somente CQA sim. Tive que montar minhas queries mais ‘frasais’.” (Participante P21820).

“Não me influenciou para montar a query de busca, porque fiz como eu faço normalmente.” (Participante P22519).

Quanto a listagem das CQA como fonte da busca, apenas 2 participantes declaram que não foram suficientes, que sentiram falta de outras bases.

“Achei suficiente. Depende da tarefa. Na verdade, senti falta de outros resultados, de outros sites que não fossem só CQA. Mas de outras CQAs eu não senti falta não.” (Participante P21610).

“Eu acho que para a pesquisa do Dataset não (foram suficientes). Porque se estive no Google Search normal, ele provavelmente me apontaria para um site do R, oficial, que me explicaria como faria. (Participante P21610).

Na avaliação das CQA em relação a outras bases, foram destacados o uso do YouTube e a Wikipedia para aprendizado em detrimento da CQA por alguns participantes:

“Estão usando para tirar dúvida, para aprendizado não. Para aprendizado você tem cursos no YouTube. Eu mesmo quando tive que aprender o objectiveC da Apple, eu aprendi no YouTube, eu não aprendi na comunidade. Eu ia à comunidade para tirar minhas dúvidas, mas o aprendizado foi tudo dentro do YouTube.” (Participante P21915).

“Talvez eu entraria num vídeo do YouTube para primeiro entender um pouco de 5G e aplicação dele para IoT. Depois, se tivesse uma dúvida específica, entraria em um site CQA.” (Participante P21710).

“No máximo eu usaria o *Quora* para aprendizado ou procuraria outras ferramentas, tipo Wikipédia porque acho que é mais detalhado. Para pegar um assunto do básico ao avançado, vai para o fim como se estivesse lendo uma apostila.” (Participante P22218).

“Se eu quisesse saber em relação a 5G, eu iria para Wikipedia e leria sobre isso. Lá encontraria referências mais aprofundadas. Eu iria por esse caminho. Não por aqui (CQA). Aqui vou quando imagino qual seria a resposta que quero, aí vou nesse tipo de base. Agora, quando não tenho nem ideia, é difícil vir aqui (CQA).” (Participante P21715).

Além disso, os participantes reconhecem que os Blogs são naturalmente mais didáticos que as trilhas da CQA, além de mais conceituais, genéricos e detalhando um assunto do início ao fim. Contudo, ressaltam não haver troca de informação ou discussão, o que ocorre em uma CQA:

“Blog geralmente é mais didático e falará de um assunto de uma forma mais ampla. Pergunta e resposta é mais direcionada.” (Participante P22519).

“O CQA está normalmente voltado para responder perguntas específicas. Os conteúdos de blogs em si, são geralmente conceituais, falando mais genericamente determinados conceitos. Isso interfere na minha busca também.” (Participante P21820).

“Eu acho que são coisas completamente diferentes. Gosto muito de fóruns de discussão, porque você observa as diferenças. A linha de raciocínio sobre um determinado assunto. Ali você vê divergências efetivamente. Num blog não necessariamente.” (Participante P22519).

No Quadro 31 apresentamos os elementos codificados coletados a partir dos relatos dos participantes que são facilitadores ou restringidores para o uso das trilhas de CQA como material de aprendizagem, considerando a busca de informação e obtenção de conhecimento através das mesma.

Tabela 31 – Aspectos da CQA que influem na aprendizagem segundo percepção dos participantes

Facilitadores	Restringidores
• Demonstração prática com fornecimento de explicações mais técnicas. (9 participantes) • Debate de opiniões. (4 participantes) • Texto mais objetivo, maior objetividade. (4 participantes) • Maior enfoque na discussão e conversa. (7 participantes) • Apresentação de pontos de vistas distintos. (2 participantes) • Respostas se complementando. (9 participantes) • Respostas mais elaboradas. (3 participantes) • Variedade de possibilidade de ordenação das respostas. (1 participante) • Clareza do conteúdo permitindo rápida compreensão. (1 participante)	• Conteúdo com enfoque dado mais à sintaxe. (2 participantes) • Pergunta inadequada. (1 participante) • Linguagem inadequada. (3 participantes) • Fornece solução ao invés de conhecimento. (1 participante) • Resposta complexa demanda mais tempo e vezes para análise, devido a organização. (3 participantes) • Texto mais subjetivo quando aborda questões conceituais. (1 participante) • Não é estruturado para aprendizagem. (1 participante) • Falta de referências consistentes. (1 participante)

Fonte: do Autor (2022)

Todavia, quando avaliado como material complementar de aprendizagem, os participantes tendem a considerar adequado, sobretudo para conhecimento mais técnico (13 participantes). Mas destacam alguma necessidade de adequação, visto que consideram que 'não é completo', 'demanda conhecimento prévio do assunto' e alguns tópicos são multiplataformas (requer acessos a diversas fontes).

“Essa pergunta (sobre 5G e IoT), considerando as respostas que obtive, é uma pergunta de 'trans plataforma', que fará com que o usuário explore várias plataformas para obter o conhecimento em cima dessa pergunta.” (Participante P21510).

Quanto ao público-alvo para aprendizagem, todos os participantes destacam a dependência da forma como o conteúdo é abordado e apresentado na trilha, podendo exigir conhecimento prévio do assunto. Também destacam a necessidade de o usuário estar interessado realmente em aprender aquele assunto (11 participantes), dado que não seriam a fonte primária para buscar novos conhecimentos.

“Caso eu não conheça o assunto acho que não agregaria valor (para aprendizagem).” (Participante P22218).

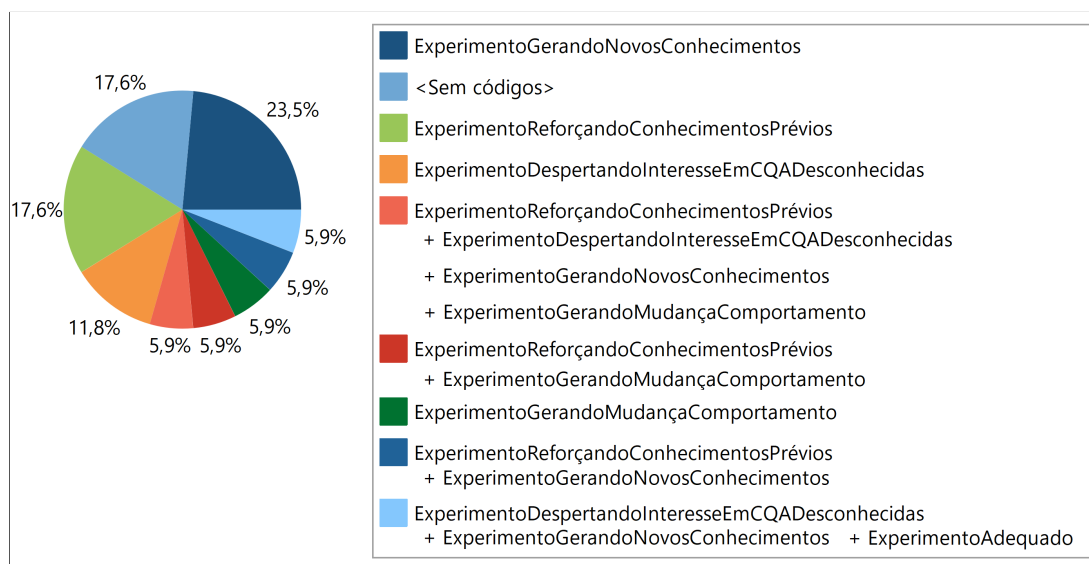
“Eu buscaria em outras fontes para coisas novas.” (Participante P22519).

“Eu acho que é mais no sentido de você tá trabalhando no dia a dia. Aprendendo sobre alguma coisa, na prática. Pode recorrer à comunidade, porque é uma tentativa, aprendendo por tentativa e erro. Você está tentando uma situação e deu erro, então você vai à comunidade para poder encontrar. Não devido à comunidade, mas pela tua experiência prática em meter a mão na massa. Naquilo novo que está fazendo e aprendendo.” (Participante P21818).

“Para mim, é para resolver problemas. Não para obter conhecimentos. Talvez conhecimentos específicos sobre algo que eu já tenho uma certa noção do que é. Não para aprender do zero.” (Participante P21720).

Considerando a própria experiência dos participantes nas atividades propostas no presente trabalho como um processo de busca como processo de aprendizagem, obtivemos a Figura 22. Podemos verificar que somente 17,6% (‘sem códigos’) não manifestaram nenhum impacto no seu comportamento ou aprendizado, em se tratando de novas informações ou conhecimentos adquiridos durante as sessões.

Figura 22 – Impacto das atividades no comportamento ou conhecimento do participante.



Fonte: do Autor (2022)

“Aquele primeiro, que tem a ver com o que estou pesquisando. Verei ainda o vídeo, mas acredito que vá ajudar muito.” (Participante P22613).

“Sim, por exemplo, aquelas coisas que não lembrava e aprendi algumas coisas... não sabia daquela distinção de ponteiros... apesar de serem dois operadores distintos. Mas também não está no meu dia a dia. Não é tão relevante, mas eu aprendi.” (Participante P21818).

“Sim, principalmente ali no início. Quando estava olhando a questão dos *datasets*, utilizando o próprio *Colab* (*Google Colab*). Assim, eu gostei tanto, que eu até marquei como favorito. Porque depois quero dar uma analisada, como faz isso, na prática, utilizando o Google Drive, que eu utilizo muito. Assim, algo muito particular. Gostei de rever a navegação ali no *Reddit* e no *Stack Overflow*. Tenho utilizado, mas é engraçado que esse experimento me fez repensar que tenho que utilizar mais essas comunidades.” (Participante P22015).

“Sim. Na busca do 5G, aprendi o que é IoT e relação entre eles.” (Participante P21720).

“Eu não costumava mesmo ver os comentários em cima das respostas e das perguntas. E ali pode ter mais informação ainda. Isso é uma coisa que eu não explorava. Os comentários em cima das perguntas e em cima das respostas.” (Participante P21616).

4.8 Considerações finais sobre codificação e análise

Neste capítulo apresentamos a elaboração da codificação inicial e focalizada a partir dos relatos transcritos dos participantes nas sessões de entrevista e de atividades, considerando os comportamentos no processo de busca e no processo analítico da trilha CQA, bem como aspectos cognitivos explícitos da construção da solução ao final das atividades, comparação entre as CQA e percepção quanto ao uso para aprendizagem. O processo de aplicação envolveu a construção de ambiente para uso para as sessões práticas, as quais envolviam desafios de busca na Web e análise de trilhas CQA, uma sessão de análise comparativa de CQA para uma pergunta conceitual e de análises de trilhas 'O QUE' e 'COMO' para suporte à aprendizagem específica. Também realizamos duas entrevistas, uma antes das tarefas, para contextualização da utilização de CQA no processo de busca segundo a experiência do participante, e outra, ao final, para obter o *feedback* da participação e perspectiva de ocorrência de aprendizagem durante as atividades. Analisamos o comportamento da busca através dos aspectos: concepção da busca, seleção dos termos e texto da busca, formulação da consulta, checagem dos resultados a partir de critérios, e a extração da informação. Analisamos o comporta analítico da trilha CQA bem como o processo cognitivo e a estratégia de construção da solução pelo participante para cada sessão prática.

Destacamos como achados deste capítulo, o uso de CQA associado às atividades técnicas e de solução de problemas desempenhadas no dia-a-dia profissional ou acadêmico do participante, com acesso predominantemente via mecanismos de buscas na Web, e preferidas em relação a outros tipos de fontes quando se trata de buscar conhecimentos ou

aprendizado. Estendemos a taxonomia proposta por [Tibau, Siqueira e Nunes \(2020\)](#) com a identificação de novos elementos como 'Substituição de Palavras' e 'Tradução Linguística'. Na concepção da busca temos a reflexão associada aos fundamentos da busca exploratória e a aprendizagem de conceitos complexos. Na seleção dos termos temos que o ato de formular consulta está ligado a intenção de busca do usuário, e a predominância de uso de palavras-chave está associada a maior compreensão do domínio. O uso de recursos na formulação está associado ao refinamento ou restrição dos resultados, tais como a expressões para termos característicos da demanda e à intenção de busca. A busca exploratória foi caracterizada pela realimentação da própria busca a partir dos resultados obtidos na formulação anterior, pela análise prévia da trilha CQA com coleta de algum conceito e elementos eram inseridos na consulta para refiná-la. A checagem dos resultados obedeceu à ordem apresentada pelo mecanismo de busca e o uso pontual de fonte externa (Wikipedia), para tratar incertezas conceituais, assim como o uso combinado de múltiplos critérios. Houve mudança de comportamento à medida que saem da 'zona de conforto' (de conhecimento), levando a extrair primeiro informação associada ao 'conceito' desconhecido. Nas trilhas CQA, o uso do corpo/descrição da pergunta para dar suporte ao entendimento quando não possuem domínio do assunto. O conhecimento existente na trilha está intimamente ligada à qualidade da resposta, durante o processamento cognitivo da trilha da CQA, bem como a clareza do título da pergunta e do detalhe da pergunta está intimamente ligada ao alinhamento da intenção, durante o processo de busca na Web. Nesse processo cognitivo que envolve estratégias de construção de solução, a complexidade da tarefa compreendida pelo participante e a seu não domínio do assunto, impuseram uso de mais recursos e combinações de trilhas e postagens, assim como o processamento de mais repostas.

5. Discussão e Análise dos Achados

O estudo da presente pesquisa visou entender os comportamentos e decisões do participante, enquanto usuário-aprendiz, ao realizar buscas na Web e a análise das trilhas de CQA resultantes da busca na perspectiva de aprendizagem informal. O enfoque foi estabelecido sobre quatro questões principais para avaliar o contexto de uso da CQA em BPA, a seguir:

- **RQ1** — Qual a utilidade da CQA no provimento de informações relevantes e necessárias aos usuários para aprendizagem?
- **RQ2** — Como a organização dos conteúdos compartilhados na CQA ajuda a promover aprendizagem dos usuários-aprendiz?
- **RQ3** — Quais são as características sócio-técnicas que uma CQA deve ter para dar suporte à aprendizagem?
- **RQ4** — Como integrar o uso de CQA com a busca como processo de aprendizagem na perspectiva de Aprendizagem Informal?

Discutimos nossas descobertas em relação a cada uma dessas questões e tópicos. No final, analisamos o uso combinado de busca na Web com fonte de informação na CQA pode compor uma busca como um processo de aprendizagem, bem como se seriam aplicáveis a possibilidade de aprendizagem informal ou mesmo aprendizagem auto-regulada, conforme questionamos através da **RQ4**.

5.1 Sobre a Relevância da CQA no provimento das informações necessárias aos usuários

Na questão **RQ1** (Qual a utilidade da CQA no provimento de informações relevantes e necessárias aos usuários para aprendizagem?), exploramos como ocorre a busca na Web com base em CQA, quando elas são suficientes ou preferidas em relação às outras fontes para as necessidades de informação do usuário-aprendiz. Para melhor tratar a relevância para o usuário-aprendiz, restringimos o estudo à busca utilizando um mecanismo de busca customizado com resultados somente de fontes CQA.

Percebemos através dos nossos experimentos qualitativos com os participantes, quando se tratava de atividade prática alinhado ao perfil de desenvolvedor do participante, a expectativa dos mesmos era que sua intenção de busca fosse respondida pelas comunidades de práticas específicas do domínio associado à necessidade de informação. Isto é, os principais resultados da busca Web deveriam corresponder a determinadas CQA como fonte. Alguns participantes formularam suas consultas com o termo ‘*stack*’ ou em formato de pergunta característico de CQA para garantir isso. De fato, 73,93% manifestaram preferência e expectativa de resultados da busca da CQA *Stack Overflow* para demandas técnicas, enquanto 17,65% tiveram esse comportamento para CQA *Quora* quando se tratava de assuntos culturais, ou mais sociais/genéricos. Isto é um indicativo do Princípio de Autoridade (CIALDINI; CIALDINI, 1993) conferida à CQA naquele tópico, baseada na confiabilidade das informações de resultados anteriormente alcançados em questões situacionais similares (CERASOLI et al., 2018). Reciprocamente, a autoridade da CQA fornece confiabilidade à informação prestada, bem como o conteúdo correspondente à intenção de pesquisa (TSENG; LIANG; TSAI, 2014).

Constatamos que a razão para o uso das CQA, na maioria das vezes (66,7%), é um imperativo de necessidade (o que não se pode evitar) ou interesse (o que é importante, útil ou vantajoso), explicitamente declarado pelos usuários participantes. A finalidade técnica é predominante (88,23%). Ambos — razão e finalidade, são devidos às demandas de trabalho profissional ou acadêmico, para sanar dúvidas técnicas, resolver um problema de forma rápida e buscar informações pontuais. Nesses casos, associadas à intenção de busca informacional. A intenção de busca informacional direta é o comportamento predominante observado, que atinge 60% dos participantes, contra 20% relativo à busca informacional não direta, de uso genérico ou social. Tais comportamentos concordam com a percepção de que a experiência profissional gera aprendizado e a intenção autodirigida de aprender para melhorar (TANNENBAUM et al., 2010) que ocorre durante a vida toda (CARNEIRO; LEFRERE; STEFFENS, 2007).

Ao avaliarmos a qualificação e classificação da CQA quanto aos valores da CQA como importante fonte de informações relevantes, foram relatados como aspectos críticos e determinantes: ter um usuário especialista respondendo (35,29%), ter uma resposta aceita (23,53%), o consenso entre os pares (41,18%), a confiabilidade da informação (64,71%), ser um mecanismo de interação social (5,88%). Podemos perceber que os conteúdos de Q&A são utilizados como bibliotecas digitais online especializadas (LE; SHAH; CHOI, 2019) com informações confiáveis, verdadeiras bases de conhecimento de alta qualidade (ZHANG et al., 2014), asseguradas pelos mecanismos de validação (GIMENEZ; SIQUEIRA, 2021) qualitativa (consenso entre pares e especialistas respondendo) e de preferência quantitativa (resposta aceita e votação).

Os participantes que são (ou foram) membros de alguma CQA (23,53% do total), avaliaram mais elementos disponíveis na plataforma, tais como perfil e reputação dos usuários envolvidos nas trilhas, bem como os comentários. Os participantes familiarizados com essas plataformas de CQA, consideraram prioritariamente os elementos de votação e resposta aceita. Já os participantes não familiarizados com a plataforma, parte minoritária do grupo, prolongaram a leitura da trilha de forma sequencial *Top-Down*. Tais comportamentos podem ser associados ao engajamento do indivíduo a determinadas CQA, devido à confiança baseada em identidade e ao *feedback* obtido anteriormente (GUAN et al., 2018) em suas necessidades por essas CQA, bem como às limitações da falta de hábito de uso (NIVALA et al., 2020).

Assim como HajiAminShirazi e Momtazi (2020), Yusuf, Yunus e Wahid (2019), constatamos que o acesso às CQA é (95,1%) através dos mecanismos de buscas, mesmo quando se espera resultados que venham de alguma CQA. As buscas diretas na CQA através dos recursos de pesquisa da plataforma são irrelevantes para usuários não membros da comunidade, mas para os membros, possuem relevância devido ao uso rotineiro no engajamento na comunidade como respondentes. Contudo, há uso eventual da navegação entre perguntas similares através dos *hiperlinks*. Evidenciamos a preferência de retorno ao mecanismo de busca na Web, customizado a partir do *Google Search*, do que seguir na análise das trilhas, para refazer a pesquisa anterior, mesmo que isso possa levar novamente à mesma CQA como resultado. Tal aspecto comportamental, que reflete controle e desorientação segundo Tseng, Liang e Tsai (2014), remete ao papel de Autoridade (CIALDINI; CIALDINI, 1993) do mecanismo de busca. Assim como a habilidade de autorregulação do usuário, principalmente no aspecto cognitivo, ao combinar estratégias de maneira eficaz (BOEKAERTS; MINNAERT, 1999) com uso otimizado de recursos tecnológicos (URBINA; VILLATORO; SALINAS, 2021) na abordagem SRL. De fato, percebemos pelos relatos, que os participantes quando tinham dúvidas acerca do alinha-

mento do conteúdo da trilha à intenção de busca, preferiam retornar ao mecanismo de busca e iniciar o processo ao invés de exaurir a análise de todas as respostas.

5.2 Sobre existência de conteúdo na CQA que promova a aprendizagem do usuário

Na questão **RQ2** (Como a organização dos conteúdos compartilhados na CQA ajuda a promover aprendizagem dos usuários-aprendiz?), exploramos a existência de conteúdo compartilhado na CQA que permita aos usuários aprenderem. Consideramos como tais conteúdos são organizados incorporando o conhecimento tácito dos usuários através das postagens, segundo a perspectiva de [Wenger \(1999\)](#) quanto a grupos de pessoas com um interesse comum em um assunto, colaboram para compartilhar ideias ou encontrar soluções. Nesse sentido, podemos perceber as CQA através da Teoria de Comunidades de Práticas ([SRBA; BIELIKOVA, 2016](#)), já descrita no Capítulo 2, como algumas CQA mais técnicas e focadas em domínios específicos, como a *Stack Overflow* e as da *Stack Exchange*. Também podemos reconhecer essa interação colaborativa na CQA através de duas Teorias inter-relacionadas, a inteligência coletiva e a sabedoria da multidão, como as CQA mais genéricas e de foco mais abrangente, por exemplo, a *Reddit* e a *Quora*. Isto é particularmente notado através dos relatos da maioria dos participantes com perfil de desenvolvimento de sistemas de informação, conforme apresentado no Capítulo 4, que desenvolveram a preferência e confiança nas informações existentes nos conteúdos das trilhas da comunidade *Stack Overflow*. Nesse ponto, podemos entender a Autoridade ([CIALDINI; CIALDINI, 1993](#)) que determinadas CQA possuem para os usuários acerca de determinados domínios, como o caso da *Stack Overflow* descrito.

Verificamos pelos relatos que a tomada de decisão quanto ao uso do conhecimento existente na trilha está intimamente ligada à resposta, durante o processamento cognitivo de análise da trilha da CQA, externalizado. Bem como a seleção da trilha da CQA está intimamente ligada ao título da pergunta e detalhe da pergunta, durante o processo de busca na Web. Isto é, a necessidade de conhecimento do usuário-aprendiz está sendo elucidada por alguma resposta presente, através de sua qualidade que permite entendimento por parte do usuário, a partir do alinhamento da pergunta com a intenção de pesquisa, impactada pela organização do conteúdo que permita analisar e recuperar a informação desejada. Em função disso, o título da pergunta somente é avaliado em 53,25% das ações dos participantes, porque já foi verificado no resultado da busca antes de acessar a trilha CQA. Já o detalhe/corpo da pergunta reteve a atenção em 64,94%, motivado pela falta de domínio do assunto da tarefa por alguns participantes, que a utilizaram para obter novos conhecimentos associados e avaliar o alinhamento da trilha à intenção de busca.

De acordo com [Le, Shah e Choi \(2019\)](#), a qualidade das respostas associada a alguma sinalização de validade é um fator crítico para o uso do conteúdo da CQA para aprendizagem, em função da dependência da existência de especialistas respondendo e da capacidade de compreensão dos perguntadores para avaliar, comprometendo a correção do conhecimento aprendido pelo usuário. Por isso, nos relatos colhidos e descritos no Capítulo 4 anterior, avaliamos a existência de mecanismos de validação e ordenação das respostas e o processo analítico de leitura do conteúdo frente à necessidade de informação.

No presente estudo, percebemos que a contribuição do conhecimento recai fundamentalmente sobre as respostas, com destaque para influência da existência de uma apontada como melhor ou destacada (74,03%), de ordenada em primeiro com maior votação (36,36%), de pontuação significativa (31,17%), ou ter sido postada em data recente (23,38%). Importante ressaltar que normalmente a primeira resposta listada é ou a melhor resposta, ou a de maior pontuação, ou ambas. Os comentários surgem como elucidadores eventuais, quando agregam à discussão na trilha, para o alinhamento da pergunta (36,36%) à intenção da busca ou para ajudar no entendimento do escopo da resposta (37,66%). Nesse contexto, a existência de um sistemática de votação das contribuições que sinalize a validação pelos pares, complementado pelos comentários para tirar dúvidas sobre respostas existentes, fortalecem a percepção de conhecimento válido e correto sendo compartilhado. Isto pode ser descrito pelo conceito de sabedoria das multidões, caracterizada como “o processo de considerar a opinião coletiva de um grupo de indivíduos em vez de um único especialista para responder a uma pergunta” ([SUROWIECKI, 2005](#) apud [SRBA; BIELIKOVA, 2016](#))). Observamos tal conceito através do consenso entre os pares, refletido pela votação nas postagens pelos usuários membros. Nos relatos dos participantes, percebemos a importância desse aspecto tanto para a seleção do conteúdo para uso quanto para conferir confiança à informação. Assim, mesmo com a percepção por alguns participantes que a solução encontrada fosse imperfeita, requerendo alguma adaptação, ela já seria a melhor opção para uso ao invés de reiniciar nova busca por informação.

Ao considerarmos que os participantes manifestaram a necessidade de informação para uso direto (8,62%), para conhecimento sobre algo de seu interesse (24,14%) ou implementação de solução (67,24%), com destaque para o fato das atividades do presente estudo serem majoritariamente de solução de problemas, as CQA cumprem papel de compartilhar conhecimentos conforme os objetivos pessoais e situacionais ([CERASOLI et al., 2018](#)). A necessidade do usuário, associada ao domínio do assunto e a dinâmica de interação ([MACHADO; GIMENEZ; SIQUEIRA, 2020](#)), descrevem o comportamento do usuário ao analisar as trilhas, seu conteúdo e os elementos existentes.

O processo de leitura, um desses comportamentos observados no estudo, acontece em conformidade com o conhecimento prévio do usuário no assunto e de acordo com sua estratégia particular (hábito de 'escanear' o conteúdo, por exemplo) de busca de informação. A partir dos relatos, podemos estabelecer um comportamento de leitura indo da seletiva (43,9%) para a em profundidade (4,7%), que envolve também os comentários, conforme aumenta o interesse do usuário ou reduz-se o conhecimento prévio do assunto da trilha. Assim, a leitura em profundidade sugere uma maior necessidade de um aprofundamento ou compreensão do assunto. A leitura *Top-Down* (16,9%) está associada a um comportamento cognitivo de dúvida, que permite ao usuário investigar as respostas uma a uma até que se dê por satisfeito ou desista da trilha. Uma variação da leitura seletiva é a leitura transversal (14,9%), sendo fortemente apoiada nos aspectos visuais, onde é feito um 'escaneamento' visual buscando por algum termo ou elemento gráfico esperado pelo usuário para o contexto de busca. Nesse processo, a primeira resposta ou a resposta aceita são alvos da leitura seletiva. Novamente, a ordem como as respostas estão apresentadas influi na forma como o conhecimento se apresentará para o usuário leitor. Cabe ressaltar a influência da maioria das atividades demandarem implementação de solução, o perfil de TI e as bases de CQA técnicas predominantes, o que faculta ao usuário leitor a oportunidade de vasculhar seletiva ou visualmente por trechos de códigos, expressões ou fórmulas.

Quando consideramos o conhecimento promovido pela trilha da CQA no processo cognitivo de construção da solução pelo usuário, para os assuntos de maior interesse, seja despertado pelas atividades ou interesse original do participante, sempre foram utilizados recursos tecnológicos, tais como favoritos do Navegador (10,3%) ou salvar em algum recurso externo (2,1%), para permitir retorno à trilha para aprender algo novo ou complementar. Contudo, o uso imediato do conhecimento percebido pelo participante, dado o tipo das tarefas serem de solução de implementação, foi predominante a adaptação (34,5%), e uso preliminar ou direto (24,2%).

Detectamos que o conhecimento presente na resposta precisou ser validado e confrontado com outras respostas (11,7%) ou com material/site externo (12,4%). Nesse aspecto, percebemos o uso da navegação por *hiperlinks* (22,4%), com predominância para os externos (75%) à CQA, como recurso para ampliar o entendimento sobre o assunto e assim angariar mais conhecimentos associados, como complementação a que se estava buscando na própria trilha. Notamos que dois participantes possuem o hábito de usar *links* se acharem necessário.

Resumindo, para que a CQA promova o compartilhamento de conhecimentos adequadamente à aprendizagem, pelo aqui exposto e conforme Capítulo 4, precisamos considerar a aplicação da Teoria de Comunidades de Práticas, associada à Teoria da Inteligência cole-

tiva e da Sabedoria da multidão. O mesmo vale para os aspectos importantes para seleção de conteúdo e confiança na informação existente na trilha, que devem ser tratados através da melhoria da qualidade dos conteúdos da trilha, uso de navegação e os links para encadear os conteúdos, dar destaque ao que é realmente essencial com melhoria de foco e organização.

5.3 Características sócio-técnicas da CQA que suportam ou inibem o seu uso para aprendizagem

Na questão **RQ3** (Quais são as características sócio-técnicas que uma CQA deve ter para dar suporte à aprendizagem?), exploramos a existência de características sociais e técnicas que permitam a CQA atuar no suporte à aprendizagem informal, considerando inclusive aquelas que possam ser inibidoras dessa atuação.

Conforme nos apresenta [Guan et al. \(2018\)](#), há influência de fatores sociais que afetam a motivação dos usuários em utilizar e participar de CQA em um cenário de aprendizagem, tais como a confiança baseada em identidade, e o *feedback* positivo de participação ou uso anterior. Outra característica social é a possibilidade de aproveitar a inteligência coletiva da comunidade especializada para a resolução de tarefas difíceis, que envolvem conhecimentos diversos, associado à existência de debate coletivo sobre a pergunta ([SRBA; BIELIKOVA, 2016](#)), possibilitando enriquecimento da aprendizagem do usuário a respeito do assunto. De fato, percebemos que nas tarefas que envolviam algum conhecimento que o participante não tinha ou que não possuía domínio do assunto, seu comportamento mudava, utilizando os comentários das perguntas para enriquecer seu próprio conhecimento a respeito do assunto, bem como múltiplas respostas para poder formar o conhecimento necessário para prover sua própria resposta para a solução do problema. Isto é, os comentários da pergunta representavam o debate coletivo necessário à compreensão do assunto, enquanto o uso de múltiplas respostas com auxílio de comentários eventuais associados representavam a inteligência coletiva especializada que fundamentaria seu próprio conhecimento adquirido.

O reconhecimento da CQA como Autoridade em determinados assuntos e tópicos é o principal ativo sócio-técnico do seu uso para aprendizagem. Tal característica mais geral está relacionado ao histórico de utilização com sucesso para determinados objetivos, mais técnicos no caso do *Stack Overflow* e mais genérico no caso do *Reddit*. Isto está apoiado na perspectiva que nessas comunidades existem os especialistas do assunto, que há o consenso da melhor resposta (aceitação) viabilizando aquele conhecimento através da

análise de similaridade da pergunta com a intenção da busca do usuário. Também, há percepção da consistência das respostas às perguntas por existir a avaliação dos pares da própria comunidade sobre o assunto explorado.

Assim, a existência de mecanismos socio-técnicos como um sistema de qualificação (sinalização de resposta aceita, a melhor resposta) e de preferência quantificada (votação, mais votada, ordenação por quantitativo de votos) dos conteúdos contribui para aprendizagem do usuário, ao aumentar a colaboração na comunidade e reduzir o número de perguntas de baixa qualidade (ARITAJATI; NARAYANAN, 2013), promovendo uma sinalização de aceitação do que é mais adequado pelo perguntador à sua necessidade (LE; SHAH; CHOI, 2019). Pelos resultados obtidos no estudo, essa sistemática de avaliação das respostas, bem como da pergunta, conferem confiança ao conteúdo das trilhas CQA para uso na aprendizagem de novos conhecimentos, sobretudo daqueles considerados técnicos.

Segundo a categorização da CQA proposta por Srba e Bielikova (2016), podemos considerar os recursos do sistema e dos usuários presentes. Os recursos do sistema são características técnicas: (i) textuais, que estão relacionadas ao corpo textual de uma pergunta analisada ou à própria resposta, como o comprimento da pergunta (quantidade de palavras do título ou detalhamento da pergunta), a estrutura (a presença de um trecho de código, ambiente de execução, *hyperlinks*) e o estilo/legibilidade (vários erros de digitação, densidade de pontuação, média de palavras por comprimento, itemizadores); (ii) não textuais, que capturam metadados importantes sobre uma pergunta/resposta, como o *feedback* da comunidade (votação, favoritos, relatos e a seleção de melhor resposta) e a identificação temporal (hora em que a pergunta foi postada, hora da primeira resposta); (iii) recursos do fluxo que descrevem o contexto de uma pergunta/resposta: relevância/-semelhança número de palavras que se sobrepõem entre uma pergunta e uma resposta e estatísticas (número de respostas/comentários, posição da resposta); (iv) tópicos, que capturam o significado do conteúdo (*tags*) e modelo de idioma/tópico (pacote de palavras). Os recursos dos usuários, que destacamos, são de controle de qualidade provido pelo próprio processo de respostas a perguntas, tais como o nível de atividade e interação, a qualificação do usuário da postagem (reputação, seguidores).

Com base nessa categorização (SRBA; BIELIKOVA, 2016) descrita, podemos destacar as seguintes características identificadas no processo de análise cognitiva pelos participantes sobre as trilhas que impactam na percepção de uso para aprendizagem:

- Postagens com poucas respostas são desmotivadoras e desacreditadas, embora somente as primeiras são geralmente “visitadas”/analisadas.

- Comentários parecem ser irrelevantes, podendo ser eventualmente utilizados para esclarecer as respostas a que estão associados, por exemplo.
- Recência da postagem tem pouco impacto na decisão.
- A Reputação é relevante para os usuários que já utilizam CQA, principalmente para os que são membros, enquanto parece ser irrelevante para aqueles que não utilizam CQA.
- A crença na relevância da postagem está associada à percepção do nexos da pergunta em relação ao objetivo da busca (título e corpo da pergunta), e a existência de melhor resposta/destacada, validada pelos “especialistas” da comunidade.
- Os títulos e corpo da pergunta estão atrelados à avaliação de nexos entre a postagem e o objetivo da busca.
- A evidência de aceitação recai sobre trechos de textos, quando se trata de busca por informação, ou trechos de códigos, quando se trata de implementação de solução técnica computacional.
- Elementos multimídias da trilha (gráficos, figuras ou vídeos) parecem estar associados a demandas de conhecimentos e não à busca imediata por informação ou implementação.

Ao verificarmos essas características, indicadores de impacto no uso da trilha CQA para aprendizagem do usuário, em relação à literatura, constatamos o aspecto informativo dos comentários com baixa efetividade (ZHANG et al., 2019), o efeito negativo de poucas respostas ou tamanho da pergunta (HONG et al., 2020), reputação subsidiada pela confiança baseada na identidade (GUAN et al., 2018), o nexos com a intenção de pesquisa e crença na relevância da postagem podem ser descritas em função das três dimensões: usuário (nível de aprendizado, informações demográficas), interações e domínio do conhecimento (MACHADO; GIMENEZ; SIQUEIRA, 2020).

Em resumo, para responder quais características socio-técnicas uma CQA deve ter para suporte à aprendizagem, temos que refletir sobre a relevância conquistada pela CQA como fonte naquele domínio do assunto primeiramente. O que a faz ser reconhecida nesse cenário, pela existência de mecanismos tecno-sociais, tais como haver especialistas no assunto para responder tal qual um professor ou instrutor, haver consenso dos pares tal qual a troca de experiências e níveis de entendimentos de alunos mais ou menos graduados no assunto, e haver a consistência e clareza das respostas validadas pela própria comunidade. Ainda e principalmente, devemos considerar o tipo de busca nesse processo de busca para aprendizagem na Web, como, por exemplo, a necessidade de informação de uso direto ou

mesmo de implementação, inerente a conhecimento factual, constituídas por perguntas normalmente do tipo "WH" (FIGUEROA; NEUMANN, 2014). Nesse caso, por se tratar de busca por informação objetiva, normalmente com a intenção de busca informacional direta, o que influencia é a qualidade dos trechos de textos das respostas e similaridade da pergunta (título e detalhe) com a intenção de busca. Quando se trata de solução de implementação, o que realmente influencia é a existência de trechos de códigos, com maior impacto se existir ambiente para execução acoplado à resposta que permita simular e validar na prática a resposta técnica. Já quando se trata de busca para conhecimento, há uma percepção geral dos participantes a inadequação de CQA para esse fim, quando se prefere outras fontes para aprender que possibilitem um aprofundamento de conhecimentos mais complexos, citando fontes como *Wikipedia*, *blogs* temáticos, sites de documentação, *YouTube*. Contudo, quando os elementos multimídias estiveram presentes na trilha da CQA, tais como gráficos, figuras ou vídeos, isto pareceu preencher tal lacuna.

5.4 Integração da busca como processo de aprendizagem no contexto de CQA

Finalmente, na questão **RQ4** (Como integrar o uso de CQA com a busca como processo de aprendizagem na perspectiva de Aprendizagem Informal?), restringido a perspectiva de aprendizagem informal, e considerando os aspectos da aprendizagem auto-regulada, discutimos o comportamento ao longo da busca na Web e da análise das trilhas CQA retornadas como resultado da busca. Para isso, a sequência de sessões de atividades cumprem esse papel, variando os tipos de intenção de busca e oferecendo a perspectiva de busca exploratória. Também traçamos um panorama sobre o papel de plataforma online para aprendizagem informal da CQA, dado que no estudo restringimos os resultados da busca às mesmas, associado à qualificação dos participantes da comunidade acadêmica de sistemas de informações, com experiência profissional associada às atividades de desenvolvimento de sistemas.

5.4.1 CQA como recurso de aprendizagem

Quanto à caracterização da CQA como recurso de aprendizagem, ao compartilhar conhecimentos colaborativos tácitos dos membros especializados em determinados assuntos, os usuários-aprendizes através da observação desses conhecimentos desenvolvem novos conhecimentos a partir de suas próprias construções mentais. Isso pode ser percebido através do comportamento dos participantes do estudo pela construção das soluções das demandas e o encadeamento das tarefas de mesma sessão. Além disso, pode ser influenciada pela forma como tais conhecimentos estão organizados na CQA ou acessíveis

pelos mecanismos de busca, quando as ações observadas dos participantes davam conta da mudança de comportamento pela exposição de algum termo ou pela ordem dos resultados. Isto é condizente com a perspectiva da Teoria de Aprendizagem Social (BAN-DURA; WALTERS, 1977). Assim, no processo de busca na Web, é importante ter o resultado adequado à necessidade de aprendizagem, e não só ser informacional. Também deve-se permitir que interações de recuperação de informação do usuário possam levar os usuários-aprendizes a novos conhecimentos, aprofundando gradualmente sua pesquisa a partir dos resultados intermediários observados nesse processo de busca. De fato, observamos a realimentação da busca com alguns resultados anteriormente alcançados, o que pode ser fortalecido pelo processo de análise da trilha, com a existência de discussões enriquecendo e oferecendo novas fontes de conhecimento (*hiperlinks*).

Segundo Wenger (2009), a aprendizagem ocorre principalmente na relação entre aprendizes e outros mais avançados, através de um conjunto complexo de relações sociais, sendo a CQA um currículo vivo para o aprendiz, que não se limita ao perfil de iniciante. Contudo, percebemos pelos resultados desse estudo, apresentados no Capítulo 4, que as perguntas e respostas precisam ser adequadas para aprendizagem, considerando que isto depende do interesse e motivação do usuário-aprendiz, e principalmente do nível de conhecimento pretérito a respeito do assunto. Isto porque essas perguntas podem se apresentar: (i) em níveis de compreensão e conhecimento distintos; (ii) com uso de linguagem mais ou menos especializada; (iii) com descrição mais conceitual ou mais técnica dos conteúdos; (iv) existência ou não de elementos não textuais (imagens, figuras, vídeos) mais complexos. Percebemos que a existência de elementos e quadros sociais (propagandas, vagas de emprego, etc.) inadequados para o aprendizado, que servem como fonte de dispersão e ocupam destaque visual nas trilhas, são um obstáculo visual significativo.

O tipo de conhecimento desejado, mais conceitual ou mais técnico, explorados sobre aspectos da linguagem C no Capítulo 4 respectivamente às seções 4.6.1 e 4.6.2, influi na percepção inicial da trilha da CQA para suporte à aprendizagem. Quanto à aprendizagem de conceitos, a pergunta mais detalhada e objetiva não muito extensa, corroborando a literatura existente sobre tamanho da postagem, bem como respostas mais bem elaboradas são fundamentais. Quando se trata de aprendizagem técnica, a resposta mais elaborada e organizada com código exemplificando, idealmente com ambiente de execução acoplado, complementada com explicações conceituais associadas são fundamentais.

Assim, a partir da nossa análise dos relatos dos participantes, constatamos que as trilhas da CQA são valiosos recursos de aprendizagem secundários, utilizados em complementação a outros materiais ou fontes de conhecimentos primários, geralmente mais conceituais e didáticos. Muitos dos relatos colocam o uso de CQA para sanar dúvidas. De

fato, esse nosso achado corrobora a literatura recente. [Srba et al. \(2019\)](#) exploraram, através de questionários, a perspectiva dos estudantes sobre os efeitos das estruturas de CQA em ambientes educacionais. Os autores propuseram características dos sistemas CQA que descrevem sua adequação para mediar discussões on-line em diferentes configurações de implantação e para diferentes perfis de alunos, com destaque para variabilidade de uso em cursos e perfis de aluno, flexibilidade nos tipos de perguntas e tópicos, apoio a diferentes níveis de participação, e conteúdo bem organizado e compreensível para o aluno.

Em nossa análise, detectamos que em CQA as perguntas e as respostas associadas podem estar em níveis de compreensão e conhecimentos distintos, com uso de linguagem mais ou menos especializada, com conteúdo mais conceitual ou mais técnico, ou apresentando elementos não textuais mais complexos, de acordo com o foco da CQA. Dependendo do interesse do usuário-aprendiz e nível de conhecimento a respeito, isso pode facilitar ou dificultar o uso para aprendizagem do mesmo.

Na literatura recente, já há alguns estudos que tratam da necessidade de informações adicionais, usadas justamente para compensar o desequilíbrio de conhecimento entre as perguntas e as respostas ([KUMAR; CHAUHAN, 2020](#)). As abordagens são ainda focadas no tratamento da correlação entre pares de perguntas e respostas. Elas vão desde o uso de modelagem de pares de textos de perguntas aliados ao uso das informações de usuários e respostas para melhor seleção de resposta, tentando estabelecer similaridades semânticas entre os usuários pela forma de escrita ([WEN et al., 2018](#)), até o uso de conhecimento externo do contexto associado às sentenças presentes nas postagens para classificar os pares/correlação das perguntas e repostas ([SHEN et al., 2018](#)).

Desta forma, para que as trilhas da CQA possam ser utilizadas como material complementar de aprendizagem, deve ser considerada a necessidade de adequação frente à finalidade de uso e ao material primário de aprendizagem. Isto é, os pares de perguntas e respostas selecionados devem ser organizados em relação ao conteúdo didático de referência, mais completo. Além disso, precisa ser aderente ao nível de compreensão e conhecimento do usuário-aprendiz, considerando quais conhecimentos prévios ou noção sobre o assunto possui. Para essa adaptação ao usuário-aprendiz, permitindo que ele navegue entre as postagens, obtivemos um resultado positivo sobre a possibilidade de uso dos *links* e questões relacionadas, desde que ambos estejam alinhados à perspectiva de aprendizagem e mais consistentes com a postagem.

5.4.2 Comportamento na busca como processo de aprendizagem sobre CQA

Executamos três sessões de busca técnica por solução de problema, apresentados nas seções 4.2, 4.3 e 4.4 e uma busca para conhecimento mais geral sobre um assunto, na seção 4.5. Percebemos que o comportamento de reflexão sobre a tarefa está associado à falta de conhecimento suficiente para o que é demandando, relacionado aos fundamentos da busca exploratória, quando o usuário-aprendiz gradualmente vai construindo seu conhecimento sobre a necessidade de informação da tarefa, seja pela reformulação das consultas, seja pela realimentação dessa busca a partir de análises de conteúdos da CQA em consultas anteriores.

Notadamente, o mecanismo de busca desempenhou um papel de Autoridade devido à naturalidade e costume comportamental de uso pelo participante, que na maioria das vezes respeitava seu ranqueamento, considerando que a falta de resultados satisfatórios eram de sua responsabilidade pela inadequada construção da consulta. Os poucos casos de tentativa de modificar o comportamento do mecanismo de busca estavam atrelados à inclusão de termos que figurassem no resultado da busca (exemplo 'sqlcode') ou a forçar uma determinada fonte, base de resultados (exemplo 'stack'). Nesse último caso, como já descrito, destacamos o papel de Autoridade de determinadas CQA a depender do domínio do conhecimento.

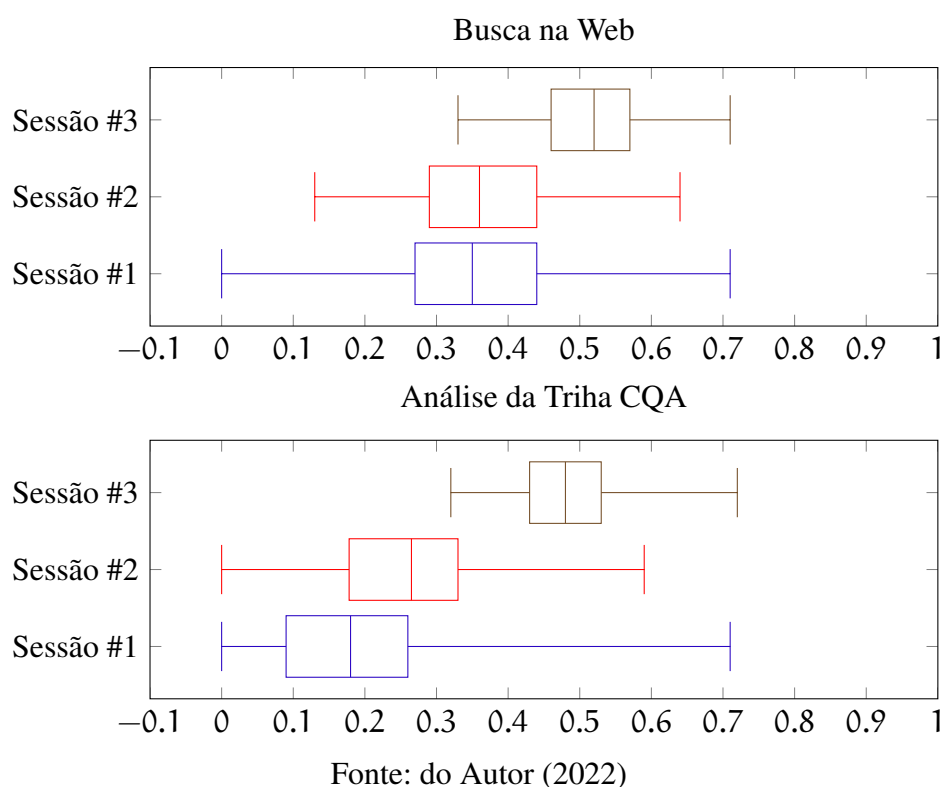
A análise de títulos e fragmentos de texto nos resultados, bem como imagens, eram primariamente utilizadas para analisar se o conteúdo daquele resultado se enquadraria na necessidade de informação. Posteriormente, na trilha CQA, o detalhamento/corpo da pergunta servia para reorientar a análise de similaridade da intenção de busca com a pergunta feita, enquanto a resposta configurava o conhecimento pretendido. Quando o usuário não percebia o alinhamento da pergunta à intenção, geralmente retornava e refazia a busca. Quando o usuário não percebia na resposta o conhecimento pretendido, geralmente utilizava outras respostas na trilha, ou mesmo, outras respostas em outras trilhas já abertas previamente em paralelo. De fato, isso corrobora o trabalho de [Tseng, Liang e Tsai \(2014\)](#), sobre os fatores decorrentes dos padrões de avaliação de informações, que podemos citar correlacionando com nossos achados: (i) várias fontes para validar a correção da informação que gerará o conhecimento pretendido; (ii) autoridade que já confere a confiança na informação por reconhecimento da comunidade de prática dentro daquele domínio; (iii) conteúdo útil quando a resposta se enquadrar no que estou esperando compreendido como o objetivo da pesquisa; e (iv) rapidamente acessível quando o conhecimento pretendido está claro nas respostas ou são rapidamente acessíveis na mesma sem confusão, ou complexidade de entendimento.

As diferentes tarefas de aprendizagem com níveis de complexidade e intenção de busca distintos, propostas pelas três primeiras sessões de atividades do estudo, possibilitaram a realização de buscas para esclarecer as mais variadas dúvidas e compreensão da necessidade de informação e exploração de conceitos, na expectativa que promovessem divergências de comportamento e no resultado da busca auxiliando na construção do conhecimento dos usuários em diferentes domínios (MACHADO; GIMENEZ; SIQUEIRA, 2020). Constatamos essa divergência ao analisarmos tanto a similaridade no processo de busca quanto no processo de análise da trilha da CQA. Percebemos que a qualificação do grupo de participantes representou uma heterogeneidade grande comportamental em ambos os processos, ao aplicar análise de similaridade para cada participante em cada sessão. A técnica de Análise de Similaridade para Documentos, que representa o par participante-sessão, foi aplicada com a medida de distância euclidiana quadrada (soma dos desvios quadrados, precedido da normalização dos valores) pela frequência de código, onde temos 0 (zero) para nenhuma e 1 (um) para total similaridade. Visualizamos na Figura 23 valores de similaridade baixos, caracterizando a heterogeneidade de comportamento do grupo. Também podemos perceber que a medida que a tarefa se desloca da zona de conforto dos participantes, isto é, fora do domínio de conhecimento deles, conseguimos ver simetria comportamental na busca e na análise da trilha da CQA, bem como a homogeneização de comportamento do grupo todo, conforme apresentado para a sessão #3, que envolvia múltiplas atividades que demandaram muitas buscas exploratórias para que os participantes obtivessem novos conhecimentos a respeito dos conceitos envolvidos nas tarefas. As estratégias e comportamentos da sessão #1 de autointeresse não mudaram significativamente para a sessão #2 sobre Fibonacci e Complexidade quando consideramos o grupo todo, apesar de termos observado mudanças significativas em alguns indivíduos, sobretudo quando consideramos o perfil de desenvolvedor, bem como ter consciência sobre uso de CQA previamente ao presente estudo.

No processo de busca, obtivemos os picos entre participantes P21610 e P21720 (0.71) na primeira sessão, entre P21510 e P21720 (0.64) na segunda sessão, e entre P21710 com P21710 e P22519 na terceira sessão (0.71). Registramos as mínimas entre os participantes P21710 e P21915 (0.00, nenhuma similaridade) na primeira sessão, entre P22519 e P31718 (0.13) na segunda sessão, e entre P21915 e P22015 (0.33) na terceira sessão.

No processo de análise das trilhas de CQA, obtivemos os picos entre os participantes P21710 e P30217 (0.71) na primeira sessão, entre P21715 e P31718 (0.59) na segunda sessão, e entre P21510 e P21820 (0.72) na terceira sessão. Registramos nenhuma similaridade (0.00) entre o participante P31718 e quase todos os outros, assim como entre P22423 e diversos outros na primeira sessão. Novamente observamos ocorrência de

Figura 23 – Comparação de Similaridade da Busca e Analítico na CQA nas três sessões



nenhuma similaridade, mas agora entre P22519 e quatro outros participantes (P21715, P21820, P21915, P22423) na segunda sessão. Curiosamente, na terceira sessão a mínima foi entre os participantes P21710 e P30217 (0.32), um valor muito elevado em relação ao alcançado nas sessões anteriores.

6. Conclusão

6.1 Observações finais

Comunidades técnicas de perguntas e respostas (CQA técnicas) têm sido amplamente utilizadas pelos desenvolvedores para preencher lacunas de conhecimentos e fornecer informações que precisam pontualmente, sobretudo para reutilização de códigos-fontes, que não estão disponíveis rapidamente através de documentação oficial ou por outros meios de compartilhamento de conhecimento na Web, tais como fóruns, Wikipédia, *YouTube*, blogs ou sites especializados. O *Stack Overflow* é o expoente principal nesse cenário, alvo crescente de diversos estudos científicos sobre a dinâmica de funcionamento e da organização das informações (JEONG et al., 2018).

Conforme apresentado na Seção 2.5, houve um movimento de mudança de foco nos estudos sobre CQA. Até recentemente, eles se concentravam em examinar os tópicos principais das perguntas dos usuários, as interações, comportamentos e como melhorar as participações na CQA, fatores de qualidade das postagens, previsão de discussões relevantes, encontrar o especialista e achar respostas para perguntas similares. Contudo, notamos nos últimos anos na literatura, que houve uma mudança gradual de foco, surgindo mais estudos para examinar os aspectos de uso, compartilhamento de conhecimentos e principalmente a inclusão da CQA na perspectiva de aprendizagem, como, por exemplo, a necessidade de superar a barreira do idioma/linguagem. Na esteira desse movimento, influenciado pela expansão do uso das redes sociais, apareceram análises e usos de CQA para outros domínios fora da área de tecnologia da informação, como a de saúde (HONG et al., 2020; SONG; LIU; ZHANG, 2021) e de educação (LE; SHAH; CHOI, 2019).

Dois cenários de aplicação da CQA no nosso dia a dia parecem se destacar a partir do que os estudos recentes apontam (JEONG et al., 2018), com base nos trabalhos relacionados e na literatura de fundamentação (Seção 2.5). O primeiro cenário e historicamente usado, trata do compartilhamento de conhecimentos técnicos e especializados, agora dis-

persos para diversos domínios do conhecimento. O segundo cenário diz respeito ao uso da CQA para apoiar alunos na aprendizagem de novos conhecimentos ou para o desenvolvimento profissional de habilidades para as demandas de trabalho. Neste último cenário, a CQA ainda tem sido explorada como recursos de aprendizagem complementar aos currículos mais técnicos, como os da área de tecnologia da informação, tanto para as organizações como para as instituições de ensino (CQA educacionais, integração a MOOC — *Massive Open On-line Course*, complemento a livros textos). Através do presente estudo, também verificamos uma maior identidade dos usuários que buscam informações na Web com as CQA, em função do reconhecimento do valor agregado às suas intenções de pesquisa de informações obtidas no passado. Contudo, ainda há uma percepção de limitação de uso para informações factuais/pontuais de baixa complexidade de compreensão, alinhando a CQA para responder a intenções de pesquisa do tipo 'WH' (pergunta), para o que os mecanismos de buscas tem priorizado como resultado as CQA de referência em cada assunto postulado na busca.

Iniciamos nosso estudo qualitativo empírico fazendo uma avaliação de percepção dos participantes quanto às próprias experiências de uso da CQA. Verificamos a preferência do uso da *Stack Overflow* para questões técnicas e que outras CQA aparecem quando se trata de conhecimentos gerais (*Quora*) ou engajamento consciente em algum assunto e grupo de discussão (*Reddit*). Os participantes desenvolvedores, mais familiarizados com o uso da CQA, usam com mais frequência e atribuem maior confiabilidade às informações existentes.

Mesmo durante o agravamento da pandemia da Covid-19 no Brasil no início de 2021, conseguimos 17 participantes, voluntários, profissionais da área de tecnologia da informação e estudantes de pós-graduação de sistemas de informação. Tal engajamento voluntário resultou em 26 horas 47 minutos e 45 segundos de duração de conteúdo de áudio e vídeo, ao longo de 2 meses de agenda de entrevistas e marcações, com 4.317 segmentos de texto codificados em 290 códigos abertos agrupados em 10 categorias conforme Seção 35, a partir de 102 documentos oriundos dos relatos e respostas das entrevistas que somaram mais de 6 mil segmentos de transcrição. Com essa amostra seletiva e bem representativa, com experiência relevante e diversa para a busca como processo de aprendizagem sobre CQA, consideramos ter atingido a saturação teórica da amostra, que cobriu 10% da população-alvo, confirmado pela inexistência de surgimentos de novos conceitos à medida que as participações são processadas, isto considerando a métrica de três seguidas sem novos conceitos.

Na sequência do estudo, realizamos três sessões de busca na Web restrita às fontes CQA, explorando o comportamento e estratégia do participante na intenção de busca in-

formacional direta. Iniciamos com a questão proposta pelo próprio participante, onde pudemos estabelecer o parâmetro comportamental natural. Escalamos para uma questão de implementação (cálculo de Fibonacci) do tipo de intenção de busca informacional direta fechada, onde se espera uma única resposta sem precisar combinar resultados, mas colocamos um obstáculo conceitual (complexidade algorítmica) para avaliar qualitativamente o quanto isso impacta no comportamento e decisões do participante. Por fim, propusemos uma questão (análise de *dataset*) de múltiplas atividades com igualmente múltiplos conceitos envolvidos. Isto permitiu observar os participantes fora da zona de conforto, igualando o nível de conhecimento entre eles em relação ao assunto. Assim, a partir da análise de similaridade do comportamento no processo de busca na Web e de análise na trilha da CQA, percebemos uma heterogeneidade do grupo em ambos os processos, com alguns casos de nenhuma similaridade (0.000 na primeira sessão) até o máximo pouco acima de 70% de similaridade, com uma ligeira homogeneização do grupo através das sessões ao considerarmos a mediana, conforme discutido no Capítulo 5. Nessas sessões percebemos a autoridade da CQA, principalmente a *Stack Overflow*, devido à confiança nas informações e seu reconhecimento como relevante no histórico de busca de cada participante.

Na sessão seguinte, exploramos o uso de busca na Web combinada com base CQA para uma intenção de busca informacional direta aberta, para compreensão de conceitos e relações entre 5G e IoT, não ligada à implementação ou solução de problema, comparando as CQA na percepção dos participantes em relação à estrutura, organização e conteúdo para a finalidade de utilização e se servem para aprendizagem. Nesse caso, percebemos a autoridade do mecanismo de busca, explicitamente o *Google Search* utilizado, quando foi respeitado o ranqueamento dos resultados e a sugestão de consulta. Quando examinamos o uso de trilhas de CQA em duas atividades de análise de trilhas 'O QUE' e 'COMO' da linguagem C, percebemos que o ambiente e organização da CQA são inadequados para o aprendizado, com muitos elementos e conteúdos de dispersão para a atenção, enquanto os mais relevantes para o aprendizado não possui posição de destaque na visualização da trilha. Em compensação, a confiabilidade da informação é assegurada através da existência de respostas aceitas e o ranqueamento das respostas por pontuação. Também constatamos que pergunta detalhada e as respostas bem elaboradas contribuem para a aprendizagem do conteúdo existente na trilha, enquanto os comentários parecem ser irrelevantes e dispersam o entendimento sobre o assunto. Contudo, quando se trata de compreensão de conceitos, há uma dificuldade de alinhamento da intenção de busca com a avaliação da pergunta na trilha. Por fim, a percepção majoritária é que as trilhas em CQA são inadequadas para aprendizagem de indivíduos iniciantes no assunto, requerendo a existência de outros tipos de conteúdos para preencher a lacuna de conhecimento.

No presente estudo, objetivamos entender os comportamentos e decisões do participante, enquanto usuário-aprendiz, ao realizar buscas na Web e a análise das trilhas de CQA resultantes da busca na perspectiva de aprendizagem informal. Dentro dessa perspectiva, partimos de quatro questões de pesquisa propostas para análise: (RQ1) Qual a utilidade da CQA no provimento de informações relevantes e necessárias aos usuários para aprendizagem?; (RQ2) Como a organização dos conteúdos compartilhados na CQA ajuda a promover aprendizagem dos usuários-aprendiz?; (RQ3) Quais são as características sócio-técnicas que uma CQA deve ter para dar suporte à aprendizagem?; e (RQ4) Como integrar o uso de CQA com a busca como processo de aprendizagem na perspectiva de Aprendizagem Informal?. As atividades propostas foram do contexto da área de formação de Computação e Sistemas de Informação. Os achados, apresentados no Capítulo 4 e discutidos no Capítulo 5, mostram-se especialmente relevante para a aprendizagem nessa área, que envolve um tipo complexo de busca, fundamentalmente exploratória, para prover o desenvolvimento intensivo de novas habilidades em um curto espaço de tempo. Tais achados fornecem interessante perspectiva de melhorias no processo de busca na Web, pois os mecanismos de recuperação de informação poderiam utilizá-los para organizar e melhorar o acesso aos elementos de CQA (como avaliações, melhor resposta como parte dos *snippets*) e específicos do conteúdo (como o código-fonte), além das otimizações dos resultados (indexação/ranking) considerando mais conteúdos da trilha e não só a pergunta. As CQA também seriam beneficiadas com melhor organização das trilhas na CQA, conforme achados, privilegiando os conhecimentos dispostos nas trilhas com navegação entre os itens e questões relacionadas, que também poderia ser disponibilizada para os mecanismos de busca utilizá-los para navegar e utilizar como resultados complementares à intenção da busca.

Previamente ao presente estudo, tivemos as contribuições científicas dos artigos publicados sobre interação e perfis na CQA (GIMENEZ; SIQUEIRA, 2021), dimensões e variáveis de BPA (MACHADO; GIMENEZ; SIQUEIRA, 2020), e características de aprendizagem em BPA (GIMENEZ et al., 2020). Como resultado deste estudo dissertativo, a partir da codificação e análises subsequentes, discutidas em relação à literatura atual, podemos destacar as seguintes principais contribuições:

- Científicas:
 - A extensão da taxonomia ESKiP (TIBAU; SIQUEIRA; NUNES, 2020; TIBAU et al., 2019) de estados de consulta, que possibilita a comparação entre os comportamentos de usuários em busca de conhecimento (usuário-aprendiz) de outros usuários que realizam buscas transacionais ou factuais através de siste-

mas de busca genérico na Web. Apresentamos na Figura 8 um novo elemento-filho da '**Reformulação**', denominado '**Adequação sintática**' subdividido em '**Tradução linguística**' e a '**Substituição de Palavras**' que foi relocado na estrutura, assim como a reconceituação de '**Consulta Similar**' em um contexto de adequação semântica.

- A estruturação em uma taxonomia multifacetada do modelo de busca exploratória ESKiP descrito por Tibau et al. (2018), utilizado para visualização do padrão de busca e do processo de aprendizagem, contemplando o processo de tomada de decisão e os filtros para pesquisa. Tal taxonomia contempla todas as atividades intensivas de conhecimento relacionadas, tais como a concepção da busca (Figura 6), a seleção dos termos e texto da busca (Figura 7), os critérios de checagem dos resultados (Figura 9), extração da informação (Figura 11).
 - A listagem de códigos abertos estruturados e agrupados em categorias temáticas, disponível no Apêndice C, que representa o processo de busca na Web, análise das trilhas CQA e cognição sobre a resolução, bem como aspectos facilitadores ou restridores para aprendizagem, e também qualificação da relevância da CQA para aprendizagem. Tal conjunto de códigos permite em estudos futuros avaliações multifacetadas e subsidia a elaboração de uma taxonomia que abranja busca como processo de aprendizagem sobre sistema de perguntas e respostas.
 - Análise qualitativa sobre o uso de comunidades de práticas online (p.ex. CQA) na perspectiva de aprendizagem a partir do processo de busca para aprendizagem realizado via um sistema de busca genérico (*Google Search*), apresentando indícios e achados resultantes com suporte de estatística descritiva.
 - Registro de percepção e reflexão do participante a respeito do seu próprio processo de aprendizagem informal, intenção de pesquisa associada à necessidade informação. O próprio estudo possibilitou novos conhecimentos, reforço de conhecimentos existentes e novas perspectivas de uso de tecnologias para apoiar a aprendizagem (CQA e sistema de busca).
- Técnicas:
- Existência de conjunto de questões guia das duas entrevistas, disponível no Apêndice A, que permite replicação para outros tipos de bases e análises.
 - Uso de ambiente customizado (CSE) sobre o mais representativo sistema de busca (*Google Search*, acima de 90% *marketshare*) existente atualmente. Tal ambiente amplia possibilidades de análises nas buscas a partir de ajustes e configurações conforme o propósito desejado.

- Códigos-fontes comentados de rotinas em *Python*, disponível no Apêndice D para transformar relatos de áudio e vídeo em textos processáveis pela maioria das ferramentas de QDA (*qualitative data analysis*) existente.
- Teóricas:
 - Identificamos a aplicabilidade do Princípio de Influência de Autoridade (CIALDINI; CIALDINI, 1993) para determinadas CQA, a depender do histórico de utilização com sucesso para determinados objetivos pelos participantes, suportada pela existência de especialistas, o consenso dos pares e a objetividade das respostas. Bem como para o mecanismo de busca, que está relacionado a naturalidade e hábito do uso retratada pela confiança nos resultados e sua ordem. Tal constatação em relação a CQA, expande o entendimento anterior em estudos anteriores na literatura (GIMENEZ; SIQUEIRA, 2021; GUAN et al., 2018) e converge com a verificação de Tseng, Liang e Tsai (2014) sobre confiabilidade na informação prestada derivada da autoridade da fonte.
 - Confirmamos com restrição a aplicabilidade dos Princípios de Influência de Consenso (CIALDINI; CIALDINI, 1993) para o processo de análise de conteúdo das trilhas da CQA, onde a existência do consenso entre pares manifestada pelo sistema de votação e pela resposta aceita. Tal qual em trabalhos anteriores (GIMENEZ; SIQUEIRA, 2021), porém restrito aos indivíduos que tenham familiaridade com a CQA (p.ex. membros), conforme os relatos.
 - Constatamos que a CQA pode ser compreendida através da Teoria de Comunidades de Práticas, associada ao papel de Autoridade, decorrente das Teorias da Inteligência Coletiva e Sabedoria da Multidão, conforme esboçado no trabalho de Srba e Bielikova (2016) e ratificado em nosso estudo pelos relatos dos participantes de perfil de desenvolvimento em relação a CQA *Stack Overflow*, considerando sistema de votação e múltiplas respostas.

6.2 Trabalhos futuros

Nesta pesquisa nos concentramos na análise qualitativa do uso de CQA no contexto de aprendizagem a partir da busca por informações na Web. Avaliamos o papel de relevância das CQA no provimento de informação e compartilhamento de conhecimentos nesse contexto. Estabelecemos as características relevantes do ponto de vista socio-técnico para suporte à aprendizagem informal, considerando o uso da busca na Web como um processo de aprendizagem.

Como possibilidade de trabalho futuro, temos o desenvolvimento de um *framework* que permita incorporar as trilhas CQA como trilhas de aprendizagem de maneira complementar, e também suplementar, aos materiais de aprendizagem convencionais. Isto, tanto para o desenvolvimento de habilidades profissionais quanto para o desenvolvimento dos alunos frente aos currículos escolares, a partir da busca como processo de aprendizagem.

Uma perspectiva futura envolve o repensar a busca como processo de aprendizagem contemplando CQA para prover informação factual e conhecimentos da comunidade de prática. Isto implica no reagrupamento de informações, no rearranqueamento e reclassificação dos resultados de busca e na promoção de informações adicionais a cada resultado (p.ex. através de *links*), apresentando questões/respostas relacionadas em CQA. Assim, explorar o uso de CQA em categorias específicas de busca, considerando tipos específicos de usuários e situação de necessidade de informação em relação a outros tipos de bases.

Na perspectiva de alinhamento dos conhecimentos compartilhados pelas trilhas de CQA para prover aprendizagem, consideramos a necessidade de tratamento dos conteúdos presente nessas trilhas, para além da qualidade, em função dos tipos de conhecimentos utilizados na cognição: factual, conceitual, processual e metacognitivo. Como no estudo de tese de [Syed \(2020\)](#), abordando detalhamento dos tipos de aprendizagem ao longo das dimensões da complexidade cognitiva da tarefa de aprendizagem usando a revisão da taxonomia de *Bloom*. Desta forma, vislumbramos o uso da taxonomia para Ensino, Aprendizagem e Avaliação de *Bloom* revisada (Lembrar: reconhecimento e recuperação; compreender: interpretação, exemplificação, classificação, resumo, inferência, comparação e explicação; aplicar: execução e implementação; analisar: diferenciação, organização e atribuição; avaliar: verificação e crítica; criar: geração, planejamento e produção) para classificar os conteúdos, por segmentos e elementos existentes (comentários, pergunta, corpo da pergunta, respostas, questões relacionadas, etc.).

Outro ponto a aprofundar, diz respeito aos métodos e mecanismos para avaliar a aprendizagem, que propiciem e reconhecer o ganho de conhecimento a partir do processo de busca enquanto processo de aprendizagem ([GIMENEZ et al., 2020](#)). Devem considerar objetivos claros de aprendizagem, materiais de avaliação confiáveis e alinhados ao contexto de aprendizagem pretendido. O que na proposição de [Urgo e Arguello \(2022\)](#), deve considerar duas dimensões ainda pouco estudadas na busca como aprendizagem, a retenção de longo prazo (conseguir usar o que foi aprendido a longo prazo) e a transferência de aprendizado (conseguir usar o que foi aprendido em um novo contexto).

REFERÊNCIAS

- AL-GARADI, M. A.; VARATHAN, K. D.; RAVANA, S. D.; AHMED, E.; MUJTABA, G.; KHAN, M. U. S.; KHAN, S. U. Analysis of online social network connections for identification of influential users: Survey and open research issues. **ACM Computing Surveys (CSUR)**, ACM, v. 51, n. 1, p. 16, 2018.
- ARITAJATI, C.; NARAYANAN, N. H. Facilitating students' collaboration and learning in a question and answer system. In: **Proceedings of the 2013 conference on Computer supported cooperative work companion**. [S.l.: s.n.], 2013. p. 101–106.
- BANDURA, A.; WALTERS, R. H. **Social learning theory**. : Prentice-hall Englewood Cliffs, NJ, 1977. v. 1. 0 p.
- BELKIN, N. J. Anomalous states of knowledge as a basis for information retrieval. **Canadian journal of information science**, v. 5, n. 1, p. 133–143, 1980.
- BERTHELSEN, C. B.; LINDHARDT, T.; FREDERIKSEN, K. A discussion of differences in preparation, performance and postreflections in participant observations within two grounded theory approaches. **Scandinavian journal of caring sciences**, Wiley Online Library, v. 31, n. 2, p. 413–420, 2017.
- BICKMAN, L.; ROG, D. J. **Handbook of applied social research methods**. [S.l.]: Sage publications, 2008.
- BOEKAERTS, M. Self-regulated learning: Where we are today. **International journal of educational research**, Elsevier, v. 31, n. 6, p. 445–457, 1999.
- BOEKAERTS, M.; MINNAERT, A. Self-regulation with respect to informal learning. **International journal of educational research**, Elsevier, v. 31, n. 6, p. 533–544, 1999.
- CAMBAZOGLU, B. B.; TAVAKOLI, L.; SCHOLER, F.; SANDERSON, M.; CROFT, B. An intent taxonomy for questions asked in web search. In: **Proceedings of the 2021 Conference on Human Information Interaction and Retrieval**. [S.l.: s.n.], 2021. p. 85–94.
- CARNEIRO, R.; LEFRERE, P.; STEFFENS, K. Self-regulated learning in technology enhanced learning environments: A european review. **Higher Education**, v. 17, n. 3, 2007.

- CAVUSOGLU, H.; LI, Z.; HUANG, K.-W. Can gamification motivate voluntary contributions? the case of stackoverflow q&a community. In: **Proceedings of the 18th ACM conference companion on computer supported cooperative work & social computing**. [S.l.: s.n.], 2015. p. 171–174.
- CERASOLI, C. P.; ALLIGER, G. M.; DONSBACH, J. S.; MATHIEU, J. E.; TANNENBAUM, S. I.; ORVIS, K. A. Antecedents and outcomes of informal learning behaviors: A meta-analysis. **Journal of Business and Psychology**, Springer, v. 33, n. 2, p. 203–230, 2018.
- CHARMAX, K. **Constructing Grounded Theory**. 2nd. ed. [S.l.]: Sage publications, 2014.
- CHARON, J. M. **Symbolic interactionism: An introduction, an interpretation, an integration**. [S.l.]: Pearson College Division, 2010.
- CHARTERS, E. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. **Brock Education Journal**, v. 12, n. 2, 2003.
- CHEN, C.; XING, Z. Towards correlating search on google and asking on stack overflow. In: IEEE. **2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)**. [S.l.], 2016. v. 1, p. 83–92.
- CHEN, C.-J.; HUNG, S.-W. To give or to receive? factors influencing members' knowledge sharing and community promotion in professional virtual communities. **Information & management**, Elsevier, v. 47, n. 4, p. 226–236, 2010.
- CHEN, L.; BAIRD, A.; STRAUB, D. Why do participants continue to contribute? evaluation of usefulness voting and commenting motivational affordances within an online knowledge community. **Decision Support Systems**, Elsevier, v. 118, p. 21–32, 2019.
- CIALDINI, R. B.; CIALDINI, R. B. **Influence: The psychology of persuasion**. , Morrow New York, v. 0, n. 0, p. 0, 1993.
- CORBIN, J.; STRAUSS, A. **Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory**. 4th. ed. [S.l.]: Sage publications, 2014.
- CRESWELL, J. W. **Investigação qualitativa e projeto de pesquisa**. 3rd. ed. [S.l.]: Penso, Porto Alegre, 2014.
- DECI, E. L.; RYAN, R. M. **Self-determination theory**. sage publications ltd, 2012.
- DENG, S.; TONG, J.; LIN, Y.; LI, H.; LIU, Y. Motivating scholars' responses in academic social networking sites: An empirical study on researchgate q&a behavior. **Information Processing & Management**, Elsevier, v. 56, n. 6, p. 102082, 2019.
- ERICSSON, K.; FOX, M. Thinking aloud is not a form of introspection but a qualitatively different methodology: Reply to schooler (2011). **Psychological bulletin**, v. 137, p. 351–4, 03 2011.
- ERICSSON, K. A.; SIMON, H. A. Verbal reports as data. **Psychological review**, American Psychological Association, v. 87, n. 3, p. 215, 1980.

- FIGUEROA, A.; NEUMANN, G. Category-specific models for ranking effective paraphrases in community question answering. **Expert Systems with Applications**, Elsevier, v. 41, n. 10, p. 4730–4742, 2014.
- _____. Context-aware semantic classification of search queries for browsing community question–answering archives. **Knowledge-Based Systems**, Elsevier, v. 96, p. 1–13, 2016.
- FRANCIS, J. J.; JOHNSTON, M.; ROBERTSON, C.; GLIDEWELL, L.; ENTWISTLE, V.; ECCLES, M. P.; GRIMSHAW, J. M. What is an adequate sample size? operationalising data saturation for theory-based interview studies. **Psychology and health**, Taylor & Francis, v. 25, n. 10, p. 1229–1245, 2010.
- FREDERIKSEN, N. Implications of cognitive theory for instruction in problem solving. **ETS Research Report Series**, Wiley Online Library, v. 1983, n. 1, p. 363–407, 1983.
- GALUŠČÁKOVÁ, P.; OARD, D. W.; NAIR, S. Cross-language information retrieval. **arXiv preprint arXiv:2111.05988**, 2021.
- GASSON, S. Rigor in grounded theory research: An interpretive perspective on generating theory from qualitative field studies. In: **The handbook of information systems research**. [S.l.]: IGI Global, 2004. p. 79–102.
- GIMENEZ, P. J. d. A.; MACHADO, M. d. O. C.; PINELLI, C. P.; SIQUEIRA, S. W. M. Investigating the learning perspective of searching as learning, a review of the state of the art. In: SBC. **Anais do XXXI Simpósio Brasileiro de Informática na Educação**. [S.l.], 2020. p. 302–311.
- GIMENEZ, P. J. d. A.; SIQUEIRA, S. W. M. How much do i stand out in communities q&a? an analysis of user interactions based on graph embedding. In: **XVII Brazilian Symposium on Information Systems**. [S.l.: s.n.], 2021. p. 1–8.
- GLASER, B. G. **Basics of grounded theory analysis: Emergence vs forcing**. [S.l.]: Sociology press, 1992.
- GLESNE, C.; PESHKIN, A. Becoming qualitative researchers: An introduction. **Pearson**, Longman, 1992.
- GUAN, T.; WANG, L.; JIN, J.; SONG, X. Knowledge contribution behavior in online q&a communities: An empirical investigation. **Computers in Human Behavior**, Elsevier, v. 81, p. 137–147, 2018.
- HAIDER, J.; SUNDIN, O. **Invisible search and online search engines: The ubiquity of search in everyday life**. [S.l.]: Taylor & Francis, 2019.
- HAJIAMINSHIRAZI, S.; MOMTAZI, S. Cross-lingual embedding for cross-lingual question retrieval in low-resource community question answering. **Machine Translation**, Springer, v. 34, n. 4, p. 287–303, 2020.
- HERDER, E.; JUVINA, I. Discovery of individual user navigation styles. **Proc. Workshop on Individual Differences in Adaptive Hypermedia, Held at AH2004**, 01 2004.

HONG, Z.; DENG, Z.; EVANS, R.; WU, H. Patient questions and physician responses in a chinese health q&a website: Content analysis. **Journal of medical Internet research**, JMIR Publications Inc., Toronto, Canada, v. 22, n. 4, p. e13071, 2020.

HOPPE, A.; HOLTZ, P.; KAMMERER, Y.; YU, R.; DIETZE, S.; EWERTH, R. Current challenges for studying search as learning processes. In: **7th Workshop on Learning & Education with Web Data (LILE2018), in conjunction with ACM Web Science 2018, Amsterdam, NL, 27 May, 2018**. [S.l.: s.n.], 2018.

HSU, M.-H.; JU, T. L.; YEN, C.-H.; CHANG, C.-M. Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations. **International journal of human-computer studies**, Elsevier, v. 65, n. 2, p. 153–169, 2007.

JEONG, S.; HAN, S. J.; LEE, J.; SUNALAI, S.; YOON, S. W. Integrative literature review on informal learning: Antecedents, conceptualizations, and future directions. **Human Resource Development Review**, Sage Publications Sage CA: Los Angeles, CA, v. 17, n. 2, p. 128–152, 2018.

JIN, J.; LI, Y.; ZHONG, X.; ZHAI, L. Why users contribute knowledge to online communities: An empirical study of an online social q&a community. **Information & management**, Elsevier, v. 52, n. 7, p. 840–849, 2015.

KAMIENSKI, A.; BEZEMER, C.-P. An empirical study of q&a websites for game developers. **Empirical Software Engineering**, Springer, v. 26, n. 6, p. 1–39, 2021.

KLEINING, G. Das qualitative experiment. **Kölner Zeitschrift für Soziologie und Sozialpsychologie**, v. 38, n. 4, p. 724–750, 1986.

KOLB, S. M. Grounded theory and the constant comparative method: Valid research strategies for educators. **Journal of emerging trends in educational research and policy studies**, Scholarlink Research Institute, v. 3, n. 1, p. 83–86, 2012.

KUMAR, A.; DANDAPAT, S.; CHORDIA, S. Translating web search queries into natural language questions. **arXiv preprint arXiv:2002.02631**, 2020.

KUMAR, S.; CHAUHAN, A. Recommending question-answers for enriching textbooks. In: SPRINGER. **International Conference on Big Data Analytics**. [S.l.], 2020. p. 308–328.

KUUSELA, H.; PAUL, P. A comparison of concurrent and retrospective verbal protocol analysis. **American Journal of Psychology**, University of Illinois Press, v. 113, n. 3, p. 387–404, 2000.

LE, L. T.; SHAH, C.; CHOI, E. Assessing the quality of answers autonomously in community question–answering. **International Journal on Digital Libraries**, Springer, v. 20, n. 4, p. 351–367, 2019.

LEE, J.-T.; KIM, S.-B.; SONG, Y.-I.; RIM, H. C. Bridging lexical gaps between queries and questions on large online q&a collections with compact translation models. In: **Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing**. [S.l.: s.n.], 2008. p. 410–418.

LEFRANCOIS, G. **Theories of Human Learning: What the Professor Said**. Cengage Learning, 2012. (PSY 361 Learning Series). ISBN 9781111829742. Disponível em: <https://books.google.com.br/books?id=SH8DIhiY8QMC>.

LI, B.; DU, X.; CHEN, M. Cross-language question retrieval with multi-layer representation and layer-wise adversary. **Information Sciences**, Elsevier, v. 527, p. 241–252, 2020.

LI, H.; XING, Z.; PENG, X.; ZHAO, W. What help do developers seek, when and how? In: IEEE. **2013 20th working conference on reverse engineering (WCRE)**. [S.l.], 2013. p. 142–151.

LIU, C.; GWIZDKA, J.; LIU, J.; XU, T.; BELKIN, N. J. Analysis and evaluation of query reformulations in different task types. In: **Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem**. USA: American Society for Information Science, 2010. (ASIS&T '10, 17).

LIU, E. Z.-F.; CHENG, S.-S.; LIN, C. H. The effects of using online q&a discussion forums with different characteristics as a learning resource. **The Asia-Pacific Education Researcher**, Springer, v. 22, n. 4, p. 667–675, 2013.

LIU, H.; LIU, C.; BELKIN, N. J. Investigation of users' knowledge change process in learning-related search tasks. **Proceedings of the Association for Information Science and Technology**, Wiley Online Library, v. 56, n. 1, p. 166–175, 2019.

LIU, P.; LI, F.; YANG, Z. A grounded theory approach for modelling the knowledge construction process in exploratory search. In: **Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020**. [S.l.: s.n.], 2020. p. 409–412.

LIU, Y.; LI, S.; CAO, Y.; LIN, C.-Y.; HAN, D.; YU, Y. Understanding and summarizing answers in community-based question answering services. In: **Proceedings of the 22nd International Conference on Computational Linguistics (COLING 2008)**. [S.l.: s.n.], 2008. p. 497–504.

LOU, J.; FANG, Y.; LIM, K. H.; PENG, J. Z. Contributing high quantity and quality knowledge to online q & a communities. **Journal of the American Society for Information Science and Technology**, Wiley Online Library, v. 64, n. 2, p. 356–371, 2013.

LOWE, A. The basic social processes of entrepreneurial innovation. **International Journal of Entrepreneurial Behavior & Research**, MCB UP Ltd, 1995.

LU, Y.; HSIAO, I.-H. Personalized information seeking assistant (pisa): from programming information seeking to learning. **Information Retrieval Journal**, Springer, v. 20, n. 5, p. 433–455, 2017.

MACHADO, M. d. O. C.; GIMENEZ, P. J. de A.; SIQUEIRA, S. W. M. Raising the dimensions and variables for searching as a learning process: A systematic mapping of the literature. In: SBC. **Anais do XXXI Simpósio Brasileiro de Informática na Educação**. [S.l.], 2020. p. 1393–1402.

- MAMYKINA, L.; MANOIM, B.; MITTAL, M.; HRIPCSAK, G.; HARTMANN, B. Design lessons from the fastest q&a site in the west. In: **Proceedings of the SIGCHI conference on Human factors in computing systems**. [S.l.: s.n.], 2011. p. 2857–2866.
- MARCHIONINI, G. Exploratory search: from finding to understanding. **Communications of the ACM**, ACM, v. 49, n. 4, p. 41–46, 2006.
- METZLER, D.; TAY, Y.; BAHRI, D.; NAJORK, M. Rethinking search: Making experts out of dilettantes. **SIGIR Forum** 55, n. 13, p. 27, 2021.
- MILLS, J.; BONNER, A.; FRANCIS, K. The development of constructivist grounded theory. **International journal of qualitative methods**, SAGE Publications Sage CA: Los Angeles, CA, v. 5, n. 1, p. 25–35, 2006.
- NADI, S.; TREUDE, C. Essential sentences for navigating stack overflow answers. In: IEEE. **2020 IEEE 27th International Conference on Software Analysis, Evolution and Reengineering (SANER)**. [S.l.], 2020. p. 229–239.
- NAVARRETE, V. D.; CALVI, L. Self-regulated and informal learning: Understanding how new digital contexts support autonomous learning processes. **Interaction Design and Architecture (s)**, TT Press, n. 7-8, p. 30–34, 2009.
- NEWELL, A. **Unified theories of cognition**. [S.l.]: Harvard University Press, 1994.
- NICO, L. S.; BOCCHI, S. C. M.; RUIZ, T.; MOREIRA, R. d. S. A grounded theory como abordagem metodológica para pesquisas qualitativas em odontologia. **Ciência & Saúde Coletiva**, SciELO Public Health, v. 12, p. 789–797, 2007.
- NIVALA, M.; SEREDKO, A.; OSBORNE, T.; HILLMAN, T. Stack overflow–informal learning and the global expansion of professional development and opportunities in programming? In: IEEE. **2020 IEEE Global Engineering Education Conference (EDUCON)**. [S.l.], 2020. p. 402–408.
- PENA-SHAFF, J. B.; NICHOLLS, C. Analyzing student interactions and meaning construction in computer bulletin board discussions. **Computers & Education**, Elsevier, v. 42, n. 3, p. 243–265, 2004.
- PICCIANO, A. G. Theories and frameworks for online education: Seeking an integrated model. **Online Learning**, ERIC, v. 21, n. 3, p. 166–190, 2017.
- PROCACI, T. B.; NUNES, B. P.; NURMIKKO-FULLER, T.; SIQUEIRA, S. W. Finding topical experts in question & answer communities. In: IEEE. **2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT)**. : , 2016. p. 407–411.
- PROCACI, T. B.; NUNES, B. P.; SIQUEIRA, S. W. **Aprender ao longo da vida através de ambientes pessoais de aprendizagem**. 2021. <<https://ieducacao.ceie-br.org/ambientes-pessoais-aprendizagem/>>. Acessado: 19-Fev-2022.
- PROCACI, T. B.; SIQUEIRA, S. W.; NUNES, B. P.; NURMIKKO-FULLER, T. Modelling experts behaviour in q&a communities to predict worthy discussions. In: IEEE. **2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)**. : , 2017. p. 291–295.

PROCACI, T. B.; SIQUEIRA, S. W. M.; BRAZ, M. H. L. B.; ANDRADE, L. C. V. de. How to find people who can help to answer a question?—analyses of metrics and machine learning in online communities. **Computers in Human Behavior**, Elsevier, v. 51, p. 664–673, 2015.

PROCACI, T. B.; SIQUEIRA, S. W. M.; NUNES, B. P.; NURMIKKO-FULLER, T. Experts and likely to be closed discussions in question and answer communities: An analytical overview. **Computers in Human Behavior**, Elsevier, , v. 92, p. 519–535, 2019.

PUTRA, M. I. J.; JUNAID, M.; SULMAN, F. The ability of the question and answer (q&a) method with the help of learnin videos against student learning outcomes amid the covid-19 pandemic. **EDUKATIF: Jurnal Ilmu Pendidikan**, v. 3, n. 5, p. 2160–2169, 2021.

RAVI, S.; PANG, B.; RASTOGI, V.; KUMAR, R. Great question! question quality in community q&a. In: **Eighth International AAAI Conference on Weblogs and Social Media**. [S.l.: s.n.], 2014.

RIEH, S. Y.; COLLINS-THOMPSON, K.; HANSEN, P.; LEE, H.-J. Towards searching as a learning process: A review of current perspectives and future directions. **Journal of Information Science**, v. 42, n. 1, p. 19–34, 2016.

ROŞCA, V. I. Implications of lewin’s field theory on social change. In: **Proceedings of the International Conference on Business Excellence**. [S.l.: s.n.], 2020. v. 14, n. 1, p. 617–625.

RUNESON, P.; HOST, M.; RAINER, A.; REGNELL, B. **Case study research in software engineering: Guidelines and examples**. [S.l.]: John Wiley & Sons, 2012.

RUTHVEN, I. Interactive information retrieval. **Annual review of information science and technology**, v. 42, p. 43–92, 2008.

SALDAÑA, J. **The coding manual for qualitative researchers**. 2nd. ed. [S.l.]: Sage publications, 2013.

SANTOS, J. L. G. d.; CUNHA, K. S. d.; ADAMY, E. K.; BACKES, M. T. S.; LEITE, J. L.; SOUSA, F. G. M. d. Análise de dados: comparação entre as diferentes perspectivas metodológicas da teoria fundamentada nos dados. **Revista da Escola de Enfermagem da USP**, SciELO Brasil, v. 52, 2018.

SAUNDERS, B.; SIM, J.; KINGSTONE, T.; BAKER, S.; WATERFIELD, J.; BARTLAM, B.; BURROUGHS, H.; JINKS, C. Saturation in qualitative research: exploring its conceptualization and operationalization. **Quality & quantity**, Springer, v. 52, n. 4, p. 1893–1907, 2018.

SCARDAMALIA, M.; BEREITER, C. **Knowledge building: Theory, pedagogy, and technology**. [S.l.]: na, 2006.

SCHOENFELD, A. H. **Mathematical problem solving**. [S.l.]: Elsevier, 2014.

SCHUNK, D. H. **Learning Theories: An Educational Perspective**. New York: Merrill. 1991.

SHEN, Y.; DENG, Y.; YANG, M.; LI, Y.; DU, N.; FAN, W.; LEI, K. Knowledge-aware attentive neural network for ranking question answer pairs. In: **The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval**. [S.l.: s.n.], 2018. p. 901–904.

SONG, X.; LIU, C.; ZHANG, Y. Chinese college students' source selection and use in searching for health-related information online. **Information Processing & Management**, Elsevier, v. 58, n. 3, p. 102489, 2021.

SRBA, I.; BIELIKOVA, M. A comprehensive survey and classification of approaches for community question answering. **ACM Transactions on the Web (TWEB)**, ACM New York, NY, USA, v. 10, n. 3, p. 1–63, 2016.

SRBA, I.; SAVIC, M.; BIELIKOVA, M.; IVANOVIC, M.; PAUTASSO, C. Employing community question answering for online discussions in university courses: Students' perspective. **Computers & Education**, Elsevier, v. 135, p. 75–90, 2019.

STACKOVERFLOW. **Stack Overflow's 2020 developer survey (2020)**. 2020. <https://insights.stackoverflow.com/survey/2020>.

STATCOUNTER. **Search Engine Market Share Worldwide / Brazil**. 2021. <https://gs.statcounter.com/search-engine-market-share>. Accessed: 2010-Out-10.

SUROWIECKI, J. **The wisdom of crowds**. [S.l.]: Anchor, 2005.

SYED, R. **Models and Algorithms for Understanding and Supporting Learning Goals in Information Retrieval**. Tese (Doutorado) — University of Michigan, 2020.

TAIBI, D.; FULANTELLI, G.; MARENZI, I.; NEJDL, W.; ROGERS, R.; IJAZ, A. Sar-web: a semantic web tool to support search as learning practices and cross-language results on the web. In: IEEE. **2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)**. [S.l.], 2017. p. 522–524.

TANNENBAUM, S. I.; BEARD, R. L.; MCNALL, L. A.; SALAS, E. Informal learning and development in organizations. **Learning, training, and development in organizations**, p. 303–332, 2010.

TAYLOR, S. J.; BOGDAN, R.; DEVAULT, M. **Introduction to qualitative research methods: A guidebook and resource**. [S.l.]: John Wiley & Sons, 2015.

TIBAU, M.; SIQUEIRA, S. W.; NUNES, B. P. Think-aloud exploratory search: Understanding search behaviors and knowledge flows. In: SPRINGER. **The International Research & Innovation Forum**. [S.l.], 2020. p. 303–315.

_____. Mapping knowledge flows in exploratory web searches. In: HICSS. **Proceedings of the 55th Hawaii International Conference on System Sciences**. 2022. Disponível em: <http://hdl.handle.net/10125/80007>.

TIBAU, M.; SIQUEIRA, S. W.; NUNES, B. P.; BORTOLUZZI, M.; MARENZI, I. Modeling exploratory search as a knowledge-intensive process. In: IEEE. **2018 IEEE 18th international conference on advanced learning technologies (icalt)**. [S.l.], 2018. p. 34–38.

- TIBAU, M.; SIQUEIRA, S. W. M.; NUNES, B. P.; NURMIKKO-FULLER, T.; MANRIQUE, R. F. Using query reformulation to compare learning behaviors in web search engines. In: **2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)**. [S.l.: s.n.], 2019. v. 2161-377X, p. 219–223.
- TSAI, M.-J.; LIANG, J.-C.; HOU, H.-T.; TSAI, C.-C. University students' online information searching strategies in different search contexts. **Australasian Journal of Educational Technology**, v. 28, n. 5, 2012.
- TSENG, S.-C.; LIANG, J.-C.; TSAI, C.-C. Students' self-regulated learning, online information evaluative standards and online academic searching strategies. **Australasian Journal of Educational Technology**, v. 30, n. 1, 2014.
- URBINA, S.; VILLATORO, S.; SALINAS, J. Self-regulated learning and technology-enhanced learning environments in higher education: A scoping review. **Sustainability**, Multidisciplinary Digital Publishing Institute, v. 13, n. 13, p. 7281, 2021.
- URGO, K.; ARGUELLO, J. Learning assessments in search-as-learning: A survey of prior work and opportunities for future research. **Information Processing & Management**, Elsevier, v. 59, n. 2, p. 102821, 2022.
- VAKKARI, P. Searching as learning: A systematization based on literature. **Journal of Information Science**, v. 42, n. 1, p. 7–18, 2016.
- VYGOTSKY, L.; HANFMANN, E. **Thought and language (3. pbk. print. ed.)**. [S.l.]: Cambridge, Mass.: MIT Press, 1967.
- WEN, J.; MA, J.; FENG, Y.; ZHONG, M. Hybrid attentive answer selection in cqa with deep users modelling. In: **Proceedings of the AAAI Conference on Artificial Intelligence**. [S.l.: s.n.], 2018. v. 32, n. 1.
- WENGER, E. **Communities of practice: Learning, meaning, and identity**. [S.l.]: Cambridge university press, 1999.
- WENGER, E. C. Communities of practice: A brief introduction. In: . [S.l.: s.n.], 2009.
- WHITE, R. W.; RICHARDSON, M.; YIH, W.-t. Questions vs. queries in informational search tasks. In: **Proceedings of the 24th international conference on World Wide Web**. [S.l.: s.n.], 2015. p. 135–136.
- WILLIG, C. **Introducing qualitative research in psychology**. 3rd. ed. [S.l.]: McGraw-hill education (UK), 2013.
- WILSON, T. D. Models in information behavior research. **Journal of Documentation**, v. 55, p. 249–270, 1999.
- WOHLIN, C.; RUNESON, P.; HÖST, M.; OHLSSON, M. C.; REGNELL, B.; WESSLÉN, A. **Experimentation in software engineering**. [S.l.]: The Kluwer International Series in Software Engineering, Norwell, USA, Kluwer Academic Publishers., 2000.
- WU, Y.; WANG, S.; BEZEMER, C.-P.; INOUE, K. How do developers utilize source code from stack overflow? **Empirical Software Engineering**, Springer, v. 24, n. 2, p. 637–673, 2019.

XUE, X.; JEON, J.; CROFT, W. B. Retrieval models for question and answer archives. In: **Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval**. [S.l.: s.n.], 2008. p. 475–482.

YILMAZ, T.; OZCAN, R.; ALTINGOVDE, I. S.; ULUSOY Özgür. Improving educational web search for question-like queries through subject classification. **Information Processing & Management**, v. 56, n. 1, p. 228–246, 2019. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0306457318301663>>.

YUSUF, N.; YUNUS, M.; WAHID, N. A comparative analysis of web search query: informational vs navigational queries. **Int. J. Adv. Sci. Eng. Inf. Technol**, v. 9, n. 1, p. 136–141, 2019.

ZHANG, H.; WANG, S.; CHEN, T.-H.; HASSAN, A. E. Reading answers on stack overflow: Not enough! **IEEE Transactions on Software Engineering**, IEEE, v. 47, n. 11, p. 2520–2533, 2019.

ZHANG, K.; WU, W.; WU, H.; LI, Z.; ZHOU, M. Question retrieval with high quality answers in community question answering. In: **Proceedings of the 23rd ACM international conference on conference on information and knowledge management**. [S.l.: s.n.], 2014. p. 371–380.

ZHANG, P.; SOERGEL, D. Process patterns and conceptual changes in knowledge representations during information seeking and sensemaking: A qualitative user study. **Journal of Information Science**, SAGE Publications Sage UK: London, England, v. 42, n. 1, p. 59–78, 2016.

ZHANG, W.-N.; LIU, T.; YANG, Y.; CAO, L.; ZHANG, Y.; JI, R. A topic clustering approach to finding similar questions from large question and answer archives. **PloS one**, Public Library of Science San Francisco, USA, v. 9, n. 3, p. e71511, 2014.

ZHANG, Y. Beyond searching: Understanding how people use search to support their creative endeavors. 2020.

ZHANG, Y.; LU, T.; PHANG, C. W.; ZHANG, C. Scientific knowledge communication in online q&a communities: Linguistic devices as a tool to increase the popularity and perceived professionalism of knowledge contribution. **Journal of the Association for Information Systems**, v. 20, n. 8, p. 3, 2019.

ZHAO, L.; DETLOR, B.; CONNELLY, C. E. Sharing knowledge in social q&a sites: The unintended consequences of extrinsic motivation. **Journal of Management Information Systems**, Taylor & Francis, v. 33, n. 1, p. 70–100, 2016.

ZHOU, J.; WANG, S.; BEZEMER, C.-P.; HASSAN, A. E. Bounties on technical q&a sites: a case study of stack overflow bounties. **Empirical Software Engineering**, Springer, v. 25, n. 1, p. 139–177, 2020.

ZHU, W.; ZHANG, H.; HASSAN, A. E.; GODFREY, M. W. An empirical study of question discussions on stack overflow. **arXiv preprint arXiv:2109.13172**, 2021.

ZIMMERMAN, B. J. Attaining self-regulation: A social cognitive perspective. In: **Handbook of self-regulation**. [S.l.]: Elsevier, 2000. p. 13–39.

ZIMMERMAN, B. J.; SCHUNK, D. H. **Self-regulated learning and academic achievement: Theoretical perspectives**. [S.l.]: Routledge, 2001.

A. APÊNDICE A - Ficha de Proposta da Pesquisa apresentada no Convite aos participantes voluntários

Título provisório: relevância das Comunidades de Perguntas e Repostas para Usuários de ferramentas de busca na Web como recursos de informação para aprendizagem.

Objetivo: Entender o processo de Busca em Comunidade de Perguntas e Respostas, bem como processo cognitivo do “usuário” realizado durante atividades de busca com foco na aprendizagem.

Descrição: Avaliar os recursos de comunidades de perguntas e respostas (CQA — *Communities of Question-Answering*) enquanto fonte de informação especializada relevante e ambiente de compartilhamento de conhecimentos que propiciem a aprendizagem do usuário, durante um processo de busca na Web. O foco é no uso, a qualidade estrutural (apresentação e disposição dos conteúdos) e facilidade para encontrar a informação desejada nas trilhas, a partir dos elementos existentes (posicionamento das postagens, organização visual, anotações, votação, resposta de destaque, links, etc.), bem como a qualidade das postagens como recurso de aprendizagem. Também averiguar se há alguma restrição para uso como “material” de aprendizagem e para que tipo de aprendizagem pode ser mais adequado.

Técnica empregada:

- Entrevista prévia (situacional) e após (feedback) experimento.
- Aplicação do Protocolo *Thinking Aloud* (o participante verbaliza oralmente tudo que está fazendo no ambiente/sistema durante a sessão e tarefas, com destaque para o que está buscando (a query), quais resultados está selecionado, nas trilhas da CQA o que está chamando mais atenção para propor a própria resposta, etc. Serão 4 sessões de 10 a 15 minutos.

Ambiente:

- Sessão montada na Sala do *Google Meet*.
 - URL: <https://meet.google.com/pjr-suzv-suh?pli=1&authuser=0>
 - Recurso de gravação áudio e vídeo (somente da tela de busca), e *chat* da sala.

- Mecanismo de busca (Google Search) customizado para somente acessar bases/sites CQA.
 - <https://cse.google.com/cse?cx=0ba01f75507c8f470#gsc.tab=0>
 - As CQA restringidas são:
 - * <https://pt.stackoverflow.com>
 - * <https://stackoverflow.com>
 - * <https://stackexchange.com>
 - * <https://answers.yahoo.com>
 - * <https://www.answersee.com>
 - * <https://www.answers.com>
 - * <https://www.quora.com>
 - * <https://www.reddit.com>
 - * <https://qastack.com.br/>
 - O resultado da busca, pode vir depois da seção de Google Ads (resultados com o botão de link “» Visit WebSite”).

Privacidade do participante

- O participante deve concordar com a gravação somente do seu áudio (*Thinking Aloud Protocol*), somente da tela de navegação e do *chat* da Sala.
- O pesquisador garante o **anonimato** do participante durante toda a pesquisa, além do experimento.
- Todos os elementos da sessão que forem gravados (imagem da navegação e digitação no navegador, texto no *chat*, e o áudio do participante) serão tabulados e codificados (catalogação, classificação, taxonomia/ontologia, entidade/atributo, etc.).
- Eventual necessidade de transcrição de áudio em texto, será garantido o anonimato do participante, com desidentificação se preciso for, e aplicado o uso da linguagem formal apropriada, sob inteira responsabilidade do pesquisador.

Sessão de Atividades.

1. Entrevista — perguntas Situacionais.
 - E1P1. Utiliza sites de Comunidades de Perguntas e Respostas (CQA) tipo *Stack Overflow* ou Quora para suas pesquisas de informação?
 - Se sim? Com que finalidade e com qual frequência?
 - Se não? Porquê?
 - Por exemplo, com uso genérico/social ou uso técnico/trabalho/acadêmico?
 - E1P2. Quando fez uso (se fez), realizou o acesso direto pelas próprias plataformas ou usa algum mecanismo de busca (Google Search p. ex.)
 - E1P3. Considera que sites e plataformas de CQA são úteis?
 - Por exemplo, estarem estruturados em debate (respostas e comentários) a partir de perguntas “reais”?
 - E1P4. Se usa mecanismos de buscas. Acha mais relevante os resultados que tenham que características de base de dados? CQA / Fórum / blogs / site especializado no assunto / Wikipédia, etc.

2. Seção de questões de busca e seleção de trilhas e respostas em CQA.

Sessão 1 — Pesquisa livre.

Enunciado: Faça uma formulação de uma pergunta qualquer, que faça parte do seu dia a dia (profissional ou acadêmico), no contexto de Sistemas de Informação, que retrate a busca por alguma informação ou implementação que acredite que seja aplicável a CQA.

Sessão 2 — Pergunta Condicionada direta.

Enunciado: Implemente um algoritmo com complexidade $O(n)$ que compute a sequência de Fibonacci até número X dado, tal que X pertence aos Naturais. Considere qualquer linguagem de programação.

Tarefa 1) Computar o somatório dos números que façam parte da sequência de Fibonacci.

Sessão 3 — Pergunta Condicionada indireta (combinada)

Enunciado: Considere que tenha um conjunto de dados de seu experimento em CSV armazenado no seu *Google Driver*, bem como um conjunto de dados de referência. O *dataset* de referência já está normalizado (*Gaussian distribution*). Necessita avaliar se o *dataset* do seu experimento também está. Após isso, faz-se necessário avaliar a correlação entre os dois conjuntos.

Pode-se considerar implementação em R ou Python.

Tarefa 1) Acesso e leitura aos *datasets* (o seu e o de referência) no *Google Driver*.

Tarefa 2) Realizar o teste de distribuição normal no seu *dataset*.

Tarefa 3) Realizar o teste de correlação entre os *datasets*.

Sessão 4 — Pergunta Comparativas (+trabalhos anteriores)

Qa) *What is benefits of 5g and how will benefit iot?*

Qb) *Why does C have a distinction between -> and .?*

Qc) *Putting user input into char array (C Programming).*

3. Entrevista — perguntas guias de *Feedback* do experimento.

E2P1. Julgou que a descaracterização visual do mecanismo de busca foi um problema? Por quê?

E2P2. Julgou que as bases/sites CQA listadas foram suficientes ou insuficientes? Sentiu falta de alguma CQA específica?

Em relação a Fóruns ou outros sites tipo Blog? Vê alguma distinção?

E2P3. Na concepção e formulação de sua pergunta foi impactada pela limitação visual da busca? Ou pela restrição dos resultados aos sites de CQA listados?

E2P4. Considerando o site/plataforma/ambiente de CQA explorados neste experimento, bem como em relação aos temas propostos, alguma consideração ou percepção sobre o formato de apresentação e disposição das perguntas, respostas, comentários e demais elementos nas trilhas.

E2P5. Na sua percepção o que destacaria (positivo e/ou negativo) em relação as CQA no uso que normalmente faria:

— Para buscar informações.

— Para a obtenção de novos ou ampliação de conhecimentos.

— Para uso como ferramenta ou plataforma de suporte à aprendizagem.

E2P6. O que mais te influenciou nas CQA durante o processo de seleção de perguntas, respostas ou comentários para elaboração de tua “própria” solução para as atividades do experimento?

E2P7. Durante as atividades do experimento, houve novos conhecimentos adquiridos ao longo dos conteúdos das trilhas analisadas na CQA? Por exemplo, alguma nova informação técnica, alguma forma de implementação técnica, novos conceitos, alguma nova abordagem?

B. APÊNDICE B - Consultas formuladas pelos participantes nas sessões

B.1 Consultas realizadas pelos participantes na Sessão #1 Livre

Tabela 32 – Consultas formuladas pelos participantes (Atividade Livre)

Participante	Código	Consulta (<i>query</i>)	Relato associado
P21510	Estado Inicial Especialização	bayes theorem python implementing bayes theorem python script	Aqui eu vou fazer o seguinte, eu vou reformular a <i>query</i> . Vou manter a mesma questão mas com...
	Consulta Similar	implementing bayes theorem R	
P21610	Estado Inicial	tool for conceptual diagram database	
P21616	Estado Inicial	chatbot learning	
	Especialização	chatbot machine learning	
P21710	Estado Inicial	angular cli	
	Especialização	angular cli version mismatch	
P21715	Estado Inicial	round number python	Eu voltaria para o <i>Google (Search)</i> e refaria a minha busca aqui. E em alguns casos, eu ainda escrevo aqui o "stack". [Modificou a <i>query</i>], assim, para garantir que ele vá lá no <i>Stack Overflow</i> .
	Especialização	round number python stack	
P21720	Estado Inicial	erro de integração oracle sqoop	Ao invés talvez de usar eu botaria "oracle data integrator sql to hdfs". [Marcando primeira parte do título do primeiro resultado da busca]. Usou o retorno da busca para reformular a consulta.
	Nova Consulta	oracle data integrator sql to hdfs	
P21818	Estado Inicial	Redis startswith	Então eu vou buscar ... porque "startswith" já é meio que um prefixo de <i>query</i> que se costuma fazer no mundo de desenvolvimento de software uma <i>query</i> por um prefixo de uma palavra então vou buscar dessa forma. [Tomou como base um resultado da busca anterior para refinar e reformular a nova busca.] Então eu vou fazer uma nova busca... que é o novo comando que eu vi.
	Nova Consulta	Python Redis scan match	
P21820	Estado Inicial Consulta Similar	How to use couchdb with docker?! couchdb in docker	Agora eu estou colocando uma outra consulta diferente se é uma pergunta vamos ver os tipos de resultados vão vir. Os resultados vieram bem parecidos.
P22015	Estado Inicial	transformação digital	
P22218	Estado Inicial	Mapping entity jpa	Eu refaria.
	Generalização	entity jpa	
P22423	Estado Inicial	Moodle comunidade rotulo ferramenta	Usou tradução de idioma Eu iria voltar (na <i>query</i> de busca), e iria tentar ser mais específica.
	Tradução Linguística Consulta Similar Generalização	Moodle tools target Moodle tools label Moodle tools h5p	
P22613	Estado Inicial	softwares de ensino de música	Reformulou antes de analisar os resultados Usou tradução de idioma
	Consulta Similar Tradução Linguística	softwares que apoiam o ensino de música softwares teaching music	
P30217	Estado Inicial	função php gerar número aleatório	Prefiro refazer a <i>query</i> do que descer até um terceiro ou quarto nível do <i>Google (Search)</i> , eu não faço isso.
P31718	Estado Inicial	WCAG 2.1	
P31818	Estado Inicial Substituição de Palavras	Explainability and information systems explainability information systems	Eu até coloquei o "AND" aqui, mas para ser bem sincero, eu não costumo usar. Vou até tirar.

B.2 Consultas realizadas pelos participantes na Sessão #2 Condicionada Direta

Tabela 33 – Consultas formuladas pelos participantes (Atividade Condicionada Direta)

Participante	Código	Consulta (<i>query</i>)	Relato associado
P21510	Estado Inicial	computing fibonacci numbers in R	
P21610	Estado Inicial Nova Consulta Especialização Consulta Similar	ordem de complexidade fibonacci algorithm fibonacci O(n) algorithm fibonacci O(n) "natural numbers" "fibonacci algorithm"O(n) "natural numbers"	
P21616	Estado Inicial Especialização	Fibonacci how to implement in python Fibonacci function sum how to implement in python	
P21710	Estado Inicial	fibonacci java	
P21715	Estado Inicial Especialização	fibonacci function o(n) fibonacci function o(n) javascript	
P21720	Estado Inicial	Computar o somatório dos números que façam parte da sequência de Fibonacci	
P21818	Estado Inicial	fibonacci python	
P21820	Estado Inicial	Fibonacci using for in javascript	Eu vou procurar aqui Fibonacci em javascript. [Digitou "Fibonacci using for in javascript"]. Não coloquei O(n). Eu coloquei lá uma interpretação minha com relação a O(n)
P21915	Estado Inicial Especialização	fibonacitti Best algoritmo java fibonacitti best algorith java o(n)	Ah, você quer O(n). [Redigindo "fibonacitti best algorith java o(n)"]. Está aí.
P22015	Estado Inicial	algoritmo, fibonacci	
P22218	Estado Inicial	formula fibonacci	
P22423	Estado Inicial	Fibonacci soma	
P22519	Estado Inicial Especialização Nova Consulta	computar o somatório dos números de fibonacci python computar o somatório dos números de fibonacci Fibonacci sequence linear	
P22613	Estado Inicial Especialização	algoritmo para somar algoritmo para somar sequência fibonacci	
P30217	Estado Inicial Nova Consulta Nova Consulta Nova Consulta	Print Fibonacci series from x to y generating Fibonacci from natural number o(n) Getting Fibonacci sequence optimized Fibonacci Sequence O(n)	e (...) ele me retornou coisas mais interessantes, aí eu abro aqui aí eu vou e mudo a minha string. Aí eu vou usar o que ele usou ali. [Copia um dos títulos que retornou no resultado e acrescenta a complexidade: "generating Fibonacci from natural number o(n)"].
P31718	Estado Inicial	fibonacci function space complexity	Vou começar por Fibonacci.. [Digitou "Fibonacci" e acabou selecionando a primeira sugestão combo da CSE: "fibonacci function space complexity"] Não tem nada em português, né.
P31818	Estado Inicial Especialização	Fibonacci sum Fibonacci sum specific number	Aqui no caso da ferramenta ... ah tá, tá, tá também ela sugere (CSE: 'Showing result for) uma correção pra mim (sobre o texto de consulta), vou clicar, porque não escrevi direito (termo "specific")

B.3 Consultas realizadas pelos participantes na Sessão #3 Condicionada Múltipla

Tabela 34 – Consultas formuladas pelos participantes
(Atividade Condicionada Múltiplas)

Código	Consulta (query)	Relato associado
Participante P21510		
Tarefa #1		
Estado Inicial	how to access dataset in google driver in R	
Tarefa #2		
Estado Inicial	normal distribution test in R	
Tarefa #3		
Estado Inicial	correlation test in R between datasets	
Especialização	correlation test in R programming between datasets	
Especialização	correlation test in R programming between not raster datasets	Abriu uma nova aba e usou parte da solução (biblioteca) para buscar como query
Especialização	correlation test in R programming between normal distributed datasets	Reiniciou a pesquisa incrementando palavras na consulta.
Especialização	pearson correlation test in R programming between normal distributed datasets	
Consulta Similar	pearson correlation and significance in R programming	Copiou o título da 1º resultado "pearson correlation and significance" e acrescentou "in R programming"
Participante P21610		
Tarefa #2		
Estado Inicial	r language "normalization test" dataset	
Repetição da Consulta	r language "normalization test" dataset	[Usou o botão BACK do navegador para retornar a CSE. Reexecutou a consulta.]
Tarefa #3		
Estado Inicial	correlation datasets normal in r language	
Nova Consulta	person test "r language"	Fazendo a busca person test "r language"
Substituição de Palavras	pearson test "r language"	Viu o resultado e refez a consulta.
Tarefa #1		
Estado Inicial	how to access dataset in google drive with "r language"	
Repetição da Consulta	how to access dataset in google drive with "r language"	Retornou novamente ao resultado da consulta [usou o botão BACK]
Nova Consulta	csv "google drive" "r language"	Voltou ao resultado e analisou. Refez a consulta após esclarecimentos adicionais.
Especialização	read csv "google drive" "r language"	Refez a consulta 'read csv "google drive" "r language"

Especialização	read csv from "google drive""r language"	Refez a consulta 'read csv from "google drive" "r language"'.
Substituição de Palavras	read csv from "google drive" "r script"	Refez novamente a busca 'read csv from "google drive""r script"'.
Consulta Similar	mount csv from google drive "r script"	Refazendo nova consulta 'mount csv from google drive "r script"'
Consulta Similar	read csv file googledrive in R script	Nova consulta a partir da sugestão do Mecanismo de busca
Participante P21616		
Tarefa #1		
Estado Inicial	how to read an csv file in python	
Especialização	how to read an csv file from Google drive in python	Então pensando aqui, faltou justamente acrescentar aqui 'from Google Drive'
Tarefa #2		
Estado Inicial	normal distribution	
Especialização + Substituição de Palavras	normality test distribution python	[Rolando os resultados verticalmente] Tanto que aqui ele fala 'normality test'. Então, digitei as palavras soltas, porque não lembro exatamente em inglês se o termo é 'normal distribution', e aí eu vi aqui que ele tem ... Nessa imagem (do resultado da busca) mostra a curva então tem a ver com o que eu quero. Então tá aqui, ele usou outro termo, então eu pego esse outro termo ('normal') e substituo (no lugar 'normality') aqui (query) para tentar ter algo mais apurado.
Tarefa #3		
Estado Inicial	correlation test between two datasets python	
Participante P21710		
Tarefa #1		
Estado Inicial	read dataset google drive	
Tarefa #2		
Estado Inicial	gaussian distribution test in r	
Tarefa #3		
Estado Inicial	correlation test in r	
Especialização	correlation test in r how to	Então vou voltar (CSE) e incluir um "how to"(a query passou a ser "correlation test in r how to").
Participante P21715		
Tarefa #2		
Estado Inicial	normal test r language	

Nova Consulta	shapiro test how to use	Então, essa aqui já sei que já posso usar o Shapiro test aqui. Se não me engano ... faz a distribuição normal. E aí eu precisava ver como é o parâmetro que eu passo para ele. Aí eu faria uma segunda pesquisa em cima de Shapiro aqui (marcando trecho do código)
Especialização	shapiro test how to use r language	
Tarefa #1		
Estado Inicial	how to read dataset google drive r language	
Tarefa #3		
Estado Inicial	correlation test r language	
Consulta Similar	pearson r language	E quando bati o olho no 5o., que é o Pearson... ai ... beleza. O Pearson já me deu mais uma dica. Há, o Pearson é um teste de correlação... Eu não entraria aqui não, procuraria algum do <i>Stack Overflow</i> falando do Pearson. Talvez refaria a busca. Até para escrever aqui (CSE), se não sei eu volto no resultado e copio o termo
Participante P21720		
Tarefa #1		
Estado Inicial	Acesso e leitura aos datasets python	
Especialização	Acesso googledrive e leitura aos datasets python	
Substituição de Palavras	Acesso google drive e leitura aos datasets python	
Tarefa #2		
Estado Inicial	teste distribuição normal no seu dataset python	
Generalização	distribuição normal no seu dataset python	
Nova Consulta	test dataset python gaussiana	
Tarefa #3		
Estado Inicial	correlation datasets python	
Especialização	correlation between two datasets python	
Participante P21818		
Tarefa #1		
Estado Inicial	read csv colab	
Tarefa #2		
Estado Inicial	normal distribution	
Nova Consulta	gaussian distribution dataframe	Digitei gausseana sem os dois 's', ai utilizei um recurso da busca que é completar automaticamente para mim

Generalização	gaussian distribution	Eu vou deixar essa busca pausada... Só para confirmar eu vou buscar 'gaussian distribution' para ver o que é.
Nova Consulta	test if dataframe is normal distribution	
Especialização	test if panda dataframe is normal distribution	
Tarefa #3		
Estado Inicial	Correlation panda dataframe	
Participante P21820		
Tarefa #1		
Estado Inicial	how to read files from Google drive using python	
Tarefa #2		
Estado Inicial	check if a dataset is a normal distribution	
Especialização	check if a dataset is a normal distribution in python	
Nova Consulta	normallity test in python	Na verdade o meu próximo passo seria continuar pesquisando, para ver se eu já não achava uma resposta direto. Por exemplo sabendo novo termos
Nova Consulta	test hypotesis using python	Vamos ver o que volta aqui. Escrevi hipotese errado.
Nova Consulta	chi-square to check if distribution is normalized	Eu ia ter que ir para o conceito mesmo. Eu sei quais funções posso usar. Mas para saber se tá normalizado ou não, eu não vi ali nenhuma resposta falando.
Especialização	chi-square to check if distribution is normalized in python	Já estou preso, não sei como usar CQA para isso. [Retornou à busca]. É (dificuldade) de conceito. Como saber como a resposta alí me indica, daquelas funções que eu vi (citando alguns métodos). Como aquelas respostas me indicam se está normalizado ou não. Aquele link me levou para o 'qui-square', que também é uma que já tava lá. Só que não sei o resultado dela o que significa para mim. Se está normalizado ou não. Na verdade, eu posso continuar perguntando aqui (CSE), posso ter perdido algo. Mas acho que voltaria a fazer as mesmas coisas.
Tarefa #3		
Estado Inicial	correlation test using python	
Participante P21915		
Tarefa #1		

Estado Inicial	R how-to open csv file in google drive	
Tarefa #2		
Estado Inicial	R how-to normal distribution comparative two tables	
Substituição de Palavras	'R how-to normal distribution comparative two datasets	
Tarefa #3		
Estado Inicial	normal distribution comparative matrix	
Especialização	normal distribution comparative matrix two datasets ANOVA	
Nova Consulta	R how-to ANOVA method	
Participante P22015		
Tarefa #1		
Estado Inicial	python, dataset, read	
Tarefa #2		
Estado Inicial	python, dataset, distribution test, colab	
Consulta Similar	python, dataset, gaussian distribution, colab	
Tarefa #3		
Estado Inicial	python, dataset, correlation test, colab	
Participante P22218		
Tarefa #1		
Estado Inicial	R read file in google drive	
Tarefa #2		
Estado Inicial	normality test in R	
Tarefa #3		
Estado Inicial	correlation test in R	
Participante P22519		
Tarefa #1		
Estado Inicial	python how to read dataset from Google Drive	
Tarefa #2		
Estado Inicial	python how to test dataset gaussian distribution	
Tarefa #3		
Estado Inicial	python correlation test dataset	
Participante P30217		
Tarefa #1		
Estado Inicial	python load from csv "Google Drive"	
Tarefa #2		
Estado Inicial	gaussian distribution python function	
Tarefa #3		
Estado Inicial	correlation test in python	
Especialização	correlation test with panda in python	

Participante P31718		
Tarefa #1		
Estado Inicial	Read a file on google drive with python	
Tarefa #2		
Estado Inicial	python Gaussian distribution	
Tarefa #3		
Estado Inicial	correlation test panda dataset	
Participante P31818		
Tarefa #1		
Estado Inicial	how to access csv dataset on google drive	
Tarefa #2		
Estado Inicial	normal test pandas	
Substituição de Palavras	normality test pandas	Mas aqui nesse ponto, eu vejo que escrevi 'normal' e 'normality' talvez seja melhor. Vou dar uma olhada até o final e se não achar muita coisa legal, eu vou voltar e escrever 'normality' pra ver se vem os resultados
Tarefa #3		
Estado Inicial	Python correlation pandas	

C. APÊNDICE C - Listagem estruturada e agrupada de códigos

Tabela 35 – Listagem estruturada dos códigos

Lista de Códigos	Frequência
(Todos)	4317
Qc_TrilhaCOMO	24
Qb_TrilhaOQUE	24
Qa_Comparacao	17
CSE+CQA_Impactos	2
CSE+CQA_Impactos \CSE+CQA_ImpactouNoComportamentoDeBusca	1
CSE+CQA_Impactos \CSE+CQA_NaoImpactouNoComportamentoDeBusca	1
CSE_ÉLimitada	38
CSE_ÉLimitada \CSE_SemInformacaoData	1
CSE_ÉLimitada \SE_ComoAutoridade	3
CSE_ÉLimitada \SE_ComoAutoridade \RanqueamentoDoMecanismoDeBusca	1
CSE_ÉLimitada \Google_É_AggregatedSearch	1
CSE_ÉLimitada \CSE_SomenteParaBuscaTextual	1
CSE_ÉLimitada \CSE_BuscaExigeQueryNoIdiomaDosResultados	2
CSE_ÉLimitada \CSE_SugerePorInfluênciaDeOutrasSessoesBusca	1
CSE_ÉLimitada \CSE_ComVisualFuncionalBásico	10
CSE_ÉLimitada \CSE_InterfaceMuitoPobre	4
CSE_ÉLimitada \CSE_SemCardsRespostaDireta	1
CSE_ÉLimitada \CSE_SemPainelWikipedia	1
CSE_ÉLimitada \CSE_SemRecursoBuscaImagem	1
CSE_ÉLimitada \CSE_SemPesquisaPorPeriodo	1
CSE_ÉLimitada \CSE_SemMarcacaoLinkResultadoJáVisitado	1
CSE_ÉLimitada \CSE_SemPermitirVerCompletamenteQueryDigitada	8
CSE_ÉLimitada \CSE_SemBuscaIntegradaNaBarraNavegacao	1
CQA_EmRelaçãoOutrasBases	0
CQA_EmRelaçãoOutrasBases \Youtube_CrescenteUsoParaAprendizagem	2
CQA_EmRelaçãoOutrasBases \WikipediaOUOutrosSitesDeConceitos_SaoMaisDetalhados	2
CQA_EmRelaçãoOutrasBases \CQA_BlogsSaoMaisDidaticos	3

CQA_EmRelaçãoOutrasBases \CQA_BlogsSãoMaisConceituaisGenéricos	4
CQA_EmRelaçãoOutrasBases \CQA_BlogsSãoMaisDetalhadosDiscutemDeInicioAoFim	2
CQA_EmRelaçãoOutrasBases \CQA_BlogFocaNoConteúdoSemTrocassDiscussao	1
CQA_Listagem	265
CQA_Listagem \CQA_ListaInsuficiente	2
CQA_Listagem \CQA_ListaSuficiente	17
CQA_Listagem \QASStack	3
CQA_Listagem \Reddit	48
CQA_Listagem \Answeree	1
CQA_Listagem \AnswersYahoo	4
CQA_Listagem \Quora	64
CQA_Listagem \StackExchange	46
CQA_Listagem \StackOverflow	61
CQA_Listagem \Outras	1
CQA_QualificacaoEClassificacao	0
CQA_QualificacaoEClassificacao \ValoresDaCQA	0
CQA_QualificacaoEClassificacao \ValoresDaCQA \UsuáriosEspecialistasRespondendo	6
CQA_QualificacaoEClassificacao \ValoresDaCQA \TerUmaRespostaAceita	4
CQA_QualificacaoEClassificacao \ValoresDaCQA \ConsensoEntrePares	7
CQA_QualificacaoEClassificacao \ValoresDaCQA \ConfiabilidadeDaInformação	11
CQA_QualificacaoEClassificacao \ValoresDaCQA \UmMecanismoSocial	1
CQA_QualificacaoEClassificacao \FazParteOUMembroDeCQA	7
CQA_QualificacaoEClassificacao \NãoFazParteOUMembroDeCQA	2
CQA_QualificacaoEClassificacao \RestricaoParticipacaoNovatos	3
CQA_QualificacaoEClassificacao \OrganizaçãoDaInformação	64
CQA_QualificacaoEClassificacao \OrganizaçãoDaInformação \ForneceExplicaçõesBásicas	9
CQA_QualificacaoEClassificacao \OrganizaçãoDaInformação \DistinçãoEntreComentáriosE-Respostas	1
CQA_QualificacaoEClassificacao \OrganizaçãoDaInformação \NãoEstáNosMoldesDeQ&A	4
CQA_QualificacaoEClassificacao \OrganizaçãoDaInformação \FormatoDeFórum	3
CQA_QualificacaoEClassificacao \OrganizaçãoDaInformação \ParteOuElementoEstrutura_Relevante	27
CQA_QualificacaoEClassificacao \OrganizaçãoDaInformação \ParteOuElementoEstrutura_Irrelevante	15
CQA_QualificacaoEClassificacao \OrganizaçãoDaInformação \ParteOUElementoEstrutura_MaisRelevante	3
CQA_QualificacaoEClassificacao \ProblemaAssociado	5
CQA_QualificacaoEClassificacao \ProblemaAssociado \IdiomaInglês	2
CQA_QualificacaoEClassificacao \ProblemaAssociado \LinguajarTécnico	4
CQA_QualificacaoEClassificacao \PreferênciasDeCQA	9
CQA_QualificacaoEClassificacao \PostagensTeoricas	2
CQA_QualificacaoEClassificacao \PerguntaQA_SaoDuasEmUma	20

CQA_QualificacaoEClassificacao \PerguntaQA_SaoDuasEmUma \PerguntaQA_NãoÉSuperficial	3
CQA_QualificacaoEClassificacao \PerguntaQA_SaoDuasEmUma \PerguntaQA_NãoÉProfunda	4
CQA_QualificacaoEClassificacao \PerguntaQA_SaoDuasEmUma \PerguntaQA_DemandaRespostasMaisTeórica	1
CQA_QualificacaoEClassificacao \PerguntaQB	25
CQA_QualificacaoEClassificacao \PerguntaQC	21
CQA_QualificacaoEClassificacao \BuscaRespostaDireta	15
CQA_QualificacaoEClassificacao \BuscaCOMOFazer	3
CQA_QualificacaoEClassificacao \EnfoqueMaisTeorico	11
CQA_Avaliação	0
CQA_Avaliação \CQA_ComoAutoridade	5
CQA_Avaliação \CQA_UsoParaTirarDuvidas	6
CQA_Avaliação \CQA_FocoAprendizagem_RequerFomentarDiscussao+LinksExternos	13
CQA_Avaliação \CQA_NaoServeParaAprendizagemMaisComplexa	3
CQA_Avaliação \CQA_ComoPortaEntradaParaAprenderMais	2
CQA_Avaliação \CQA_UsoParaResolverProblemas	19
CQA_Avaliação \CQA_UsoBuscaInformaçãoDeOpinioes	2
CQA_Avaliação \CQA_NaoServePara_InformaçãoConceitual	11
CQA_Avaliação \CQA_ServePara_InformaçãoPontual	13
CQA_Avaliação \CQA_ParaAmpliarConhecimento_ComRestricoes	19
CQA_Avaliação \CQA_FonteSecundariaInformacao_Apenas	9
CQA_Avaliação \CQA_ParaAprendizagemTecnica_Atende	15
CQA_Avaliação \CQA_ParaAprendizagemGeralPrecisaMaisDiversidadeBasesBusca	13
CQA_Avaliação \CQA_SendoDescontinuadaParaUsoAprendizado	1
CQA_Avaliação \CQA_PrimeiraExperienciaDeUso	3
CQA_Avaliação \CQA_ComparaçãoEntreElas	28
CQA_Avaliação \CQA_AprendizadoBaseadoNaPratica	4
CQA_Avaliação \CQA_ParaAprendizagemInformal	1
CQA_Avaliação \CQA_PontuaçãoÉJuízoDeQuantidade	3
CQA_Avaliação \CQA_ConsultaInterna_FaltaRelevancia	1
CQA_Avaliação \CQA_FórunsSãoParaDebates_MenosDiretos	10
CQA_Avaliação \CQA_MelhorRespostaÉJuízoDeQualidade	2
CQA_Avaliação \CQA_ImpactoPelaCulturaAprendizadoPrevio	1
CQA_Utilização	89
CQA_Utilização \MotivadoPorNecessidade	11
CQA_Utilização \MotivadoPorInteresse	5
CQA_Utilização \MotivadoPorCuriosidade	2
CQA_Utilização \FinalidadeGenéricaOUSocial	6
CQA_Utilização \FinalidadeTécnica(AcadêmicaOUTrabalho)	26
CQA_Utilização \FinalidadePontualOUEspecífico	16
CQA_Utilização \UsoFrequênciaDiária	12
CQA_Utilização \UsoFrequênciaSemanal	2

CQA_Utilização \UsoFrequênciaSazonalOURaramente	7
CQA_Utilização \UsoFrequênciaNenhuma	2
UsuárioParticipante	0
UsuárioParticipante \DomínioDoAssunto	0
UsuárioParticipante \DomínioDoAssunto \SemConhecimentoSobreOAssunto	43
UsuárioParticipante \DomínioDoAssunto \ConhecimentoSobreOAssunto	40
UsuárioParticipante \AcessoÀInformação	0
UsuárioParticipante \AcessoÀInformação \UsoFerramentaDeBusca	27
UsuárioParticipante \AcessoÀInformação \UsoFerramentaDeBusca \BuscaSecundáriaNaCQA	2
UsuárioParticipante \AcessoÀInformação \BuscaDiretaNaCQA	1
UsuárioParticipante \AcessoÀInformação \BuscaDiretaNaCQA \ComoMembroDaCQAPara-Responder	1
UsuárioParticipante \Sentimento	0
UsuárioParticipante \Sentimento \SentimentoDeUtilidade	6
UsuárioParticipante \Sentimento \SentimentoDeFrustração	7
UsuárioParticipante \Sentimento \SentimentoDeSatisfação	14
UsuárioParticipante \Sentimento \SentimentoDeIndiferença	2
ProcessoDeBusca	0
ProcessoDeBusca \BuscaMaisConceitual	15
ProcessoDeBusca \EstratégiaDeBusca	7
ProcessoDeBusca \EstratégiaDeBusca \MaisInformacaoNaBuscaMenosAbrangenciaNosResultados	1
ProcessoDeBusca \EstratégiaDeBusca \ProcuraPorSimilaridadeDeTermosChaves	1
ProcessoDeBusca \EstratégiaDeBusca \MinimizarOCustoDaBusca	1
ProcessoDeBusca \EstratégiaDeBusca \EscolhaPorIntuição	1
ProcessoDeBusca \UsoDeRecursoExterno	12
ProcessoDeBusca \Problemas	9
ProcessoDeBusca \Intenção	0
ProcessoDeBusca \Intenção \Imposição	2
ProcessoDeBusca \Intenção \Necessidade	18
ProcessoDeBusca \Intenção \AutoInteresse	16
ProcessoDeBusca \ConcepçãoDaBusca	1
ProcessoDeBusca \ConcepçãoDaBusca \DivisãoDoProblema	12
ProcessoDeBusca \ConcepçãoDaBusca \ReproduçãoDoProblema	38
ProcessoDeBusca \ConcepçãoDaBusca \ReflexãoSobreOProblema/AprofundamentoGradual	47
ProcessoDeBusca \SeleçãoDoTermoBusca	0
ProcessoDeBusca \SeleçãoDoTermoBusca \UsoDeInformacaoObtidaCQANaBuscaAnterior	5
ProcessoDeBusca \SeleçãoDoTermoBusca \UsoDeParteTituloResultadoAnterior	4
ProcessoDeBusca \SeleçãoDoTermoBusca \UsoDeSugestaoSE	8
ProcessoDeBusca \SeleçãoDoTermoBusca \DirecionamentoSEO	1
ProcessoDeBusca \SeleçãoDoTermoBusca \DirecionamentoSEO \UsoDeVirgulas	6
ProcessoDeBusca \SeleçãoDoTermoBusca \DirecionamentoSEO \UsoDeConectoresDaLinguagem	52

ProcessoDeBusca \SeleçãoDoTermoBusca \DirecionamentoSEO \UsoDeExpressões	87
ProcessoDeBusca \SeleçãoDoTermoBusca \DirecionamentoSEO \UsoDeOperadores	1
ProcessoDeBusca \SeleçãoDoTermoBusca \DirecionamentoSEO \UsoDeAspas	9
ProcessoDeBusca \SeleçãoDoTermoBusca \Formato	0
ProcessoDeBusca \SeleçãoDoTermoBusca \Formato \FormatoDePergunta	24
ProcessoDeBusca \SeleçãoDoTermoBusca \Formato \FormatoLivre	59
ProcessoDeBusca \SeleçãoDoTermoBusca \Formato \TermosChaves	51
ProcessoDeBusca \FormulaçãoDaConsulta	1
ProcessoDeBusca \FormulaçãoDaConsulta \Reformulação	3
ProcessoDeBusca \FormulaçãoDaConsulta \Reformulação \TraduçãoIdioma	4
ProcessoDeBusca \FormulaçãoDaConsulta \Reformulação \SubstituiçãoDePalavras	8
ProcessoDeBusca \FormulaçãoDaConsulta \Reformulação \ConsultaSimilar	9
ProcessoDeBusca \FormulaçãoDaConsulta \Reformulação \RepetiçãoDaConsulta	4
ProcessoDeBusca \FormulaçãoDaConsulta \Reformulação \NovaConsulta	36
ProcessoDeBusca \FormulaçãoDaConsulta \Especialização	25
ProcessoDeBusca \FormulaçãoDaConsulta \Generalização	4
ProcessoDeBusca \FormulaçãoDaConsulta \RetornoEstadoAnterior	1
ProcessoDeBusca \FormulaçãoDaConsulta \EstadoInicial	58
ProcessoDeBusca \ChecagemDeResultados	0
ProcessoDeBusca \ChecagemDeResultados \Critério	0
ProcessoDeBusca \ChecagemDeResultados \Critério \SimilaridadeComAIntençãoDeBusca	9
ProcessoDeBusca \ChecagemDeResultados \Critério \BaseadoNaRecênciaData	1
ProcessoDeBusca \ChecagemDeResultados \Critério \UsoDeFonteExterna	1
ProcessoDeBusca \ChecagemDeResultados \Critério \BaseadoRecursoMultimidia	6
ProcessoDeBusca \ChecagemDeResultados \Critério \BaseadoNoTextoResumo	35
ProcessoDeBusca \ChecagemDeResultados \Critério \BaseadoNoTítulo	127
ProcessoDeBusca \ChecagemDeResultados \Critério \SequencialSERP	116
ProcessoDeBusca \ChecagemDeResultados \Critério \BaseadoNaFonte	18
ProcessoDeBusca \ChecagemDeResultados \Critério \ManipulaçãoDaOrdenação	1
ProcessoDeBusca \ExtraçãoDaInformação	0
ProcessoDeBusca \ExtraçãoDaInformação \Seleção	0
ProcessoDeBusca \ExtraçãoDaInformação \Seleção \MaisDeUmaPaginaResultados	1
ProcessoDeBusca \ExtraçãoDaInformação \Seleção \SomenteResultadosDaPrimeiraPagina	11
ProcessoDeBusca \ExtraçãoDaInformação \Seleção \MelhorResultadoApósCritério	85
ProcessoDeBusca \ExtraçãoDaInformação \Seleção \PrimeirosResultados	41
ProcessoDeBusca \ExtraçãoDaInformação \Exibição	0
ProcessoDeBusca \ExtraçãoDaInformação \Exibição \MúltiplosResultadosEmParalelo	34
ProcessoDeBusca \ExtraçãoDaInformação \Exibição \AberturaNaMesmaJanela	17
ProcessoDeBusca \ExtraçãoDaInformação \Exibição \AberturaNovaAbaOUJanela	110
ProcessoCognitivoNaCQA	2
ProcessoCognitivoNaCQA \QuestoesRelacionadas	19
ProcessoCognitivoNaCQA \QuestoesRelacionadas \QuestõesLinked	4
ProcessoCognitivoNaCQA \OutraTrilhaAbertaMaisInteressante	4

ProcessoCognitivoNaCQA \BuscaDePalavraNaTrilha	2
ProcessoCognitivoNaCQA \ElementosDeTomadaDecisao	20
ProcessoCognitivoNaCQA \NavegaçãoPorHiperlinks	20
ProcessoCognitivoNaCQA \NavegaçãoPorHiperlinks \Externo à CQA	28
ProcessoCognitivoNaCQA \NavegaçãoPorHiperlinks \Interno à CQA	8
ProcessoCognitivoNaCQA \Linguagem	0
ProcessoCognitivoNaCQA \Linguagem \Português	12
ProcessoCognitivoNaCQA \Linguagem \Original	98
ProcessoCognitivoNaCQA \Linguagem \TraduçãoDaTela	26
ProcessoCognitivoNaCQA \Objetivo	0
ProcessoCognitivoNaCQA \Objetivo \InformaçãoParaUso	10
ProcessoCognitivoNaCQA \Objetivo \ConhecimentoSobreAlgo	28
ProcessoCognitivoNaCQA \Objetivo \ImplementaçãoDeSolução	78
ProcessoCognitivoNaCQA \Foco	6
ProcessoCognitivoNaCQA \Foco \TodasAsRespostas	12
ProcessoCognitivoNaCQA \Foco \TítuloDaPergunta	75
ProcessoCognitivoNaCQA \Foco \CorpoDaPergunta	128
ProcessoCognitivoNaCQA \Foco \ComentáriosDaPergunta	43
ProcessoCognitivoNaCQA \Foco \TagUsadaNaPergunta	2
ProcessoCognitivoNaCQA \Foco \PontuaçãoDaPergunta	13
ProcessoCognitivoNaCQA \Foco \RecênciaOuPostagemMaisAtual	23
ProcessoCognitivoNaCQA \Foco \Reputação	27
ProcessoCognitivoNaCQA \Foco \MelhorRespostaOuDestacada	140
ProcessoCognitivoNaCQA \Foco \QuantidadeDeRespostas	20
ProcessoCognitivoNaCQA \Foco \Top3Respostas	32
ProcessoCognitivoNaCQA \Foco \PrimeiraResposta	48
ProcessoCognitivoNaCQA \Foco \PontuaçãoDaResposta	34
ProcessoCognitivoNaCQA \Foco \ComentáriosDaResposta	49
ProcessoCognitivoNaCQA \Decisão	2
ProcessoCognitivoNaCQA \Decisão \Dúvida	35
ProcessoCognitivoNaCQA \Decisão \Recusa	32
ProcessoCognitivoNaCQA \Decisão \Aceitação	79
ProcessoCognitivoNaCQA \Evidência	0
ProcessoCognitivoNaCQA \Evidência \TrechoComentariosResposta	2
ProcessoCognitivoNaCQA \Evidência \MuitaInteraçõesOuRespostas	11
ProcessoCognitivoNaCQA \Evidência \PoucasInteraçõesOuRespostas	8
ProcessoCognitivoNaCQA \Evidência \ConteúdoPaginaDeLinkAcessado	12
ProcessoCognitivoNaCQA \Evidência \TrechoDeImagemOUFigura	20
ProcessoCognitivoNaCQA \Evidência \TrechoTextualTítuloPergunta	8
ProcessoCognitivoNaCQA \Evidência \TrechoTextualCorpoPergunta	17
ProcessoCognitivoNaCQA \Evidência \TrechoTextualResposta	70
ProcessoCognitivoNaCQA \Evidência \TrechoDeCódigo	124
ProcessoCognitivoNaCQA \Leitura	2

ProcessoCognitivoNaCQA \Leitura \Inversa(Pergunta+inversaoRespostas)	1
ProcessoCognitivoNaCQA \Leitura \TopDown	25
ProcessoCognitivoNaCQA \Leitura \Transversal	22
ProcessoCognitivoNaCQA \Leitura \Seletiva	65
ProcessoCognitivoNaCQA \Leitura \Sequencial	24
ProcessoCognitivoNaCQA \Leitura \EmProfundidade	7
ProcessoCognitivoConstruçãoDaSolução	0
ProcessoCognitivoConstruçãoDaSolução \UsoMouseParaMarcarPartesDeAtençãoNaTela	7
ProcessoCognitivoConstruçãoDaSolução \ArmazenarSolucaoParaUsoPosterior	3
ProcessoCognitivoConstruçãoDaSolução \ValidaçãoDaSoluçãoComMaterialExterno	18
ProcessoCognitivoConstruçãoDaSolução \UsoDiretoDaSoluçãoEncontradaNaCQA	14
ProcessoCognitivoConstruçãoDaSolução \MúltiplaAvaliaçãoEUsoDeVáriasRespostas	17
ProcessoCognitivoConstruçãoDaSolução \UsoPreliminarInicialDaInformacaoNaCQA	21
ProcessoCognitivoConstruçãoDaSolução \AdaptaçãoDaSoluçãoEncontradaNaCQA	50
ProcessoCognitivoConstruçãoDaSolução \UsarFavoritosDoNavegador	15
Experimento_Feedback	0
Experimento_Feedback \ExperimentoReforçandoConhecimentosPrévios	7
Experimento_Feedback \ExperimentoDespertandoInteresseEmCQADesconhecidas	4
Experimento_Feedback \ExperimentoGerandoNovosConhecimentos	7
Experimento_Feedback \ExperimentoGerandoMudançaComportamento	3
Experimento_Feedback \ExperimentoAdequado	1
Experimento_Ocorrências	0
Experimento_Ocorrências \IncapidadeConclusãoTarefaNoTempo	3
Experimento_Ocorrências \IncapidadeConclusãoTarefaNoTempo \IntervençãoDoObservador	2
Experimento_Ocorrências \DesistênciaTarefa	2
Aprendizagem	1
Aprendizagem \FacilitadoresDeAPrendizagem	27
Aprendizagem \FacilitadoresDeAPrendizagem \DemonstraçãoPráticaxperimental	11
Aprendizagem \FacilitadoresDeAPrendizagem \DemonstraçãoPráticaxperimental \ForneceExplicaçõesMaisTécnicas	16
Aprendizagem \FacilitadoresDeAPrendizagem \DebateDeOpiniões	7
Aprendizagem \FacilitadoresDeAPrendizagem \TextoMaisObjetivo	1
Aprendizagem \FacilitadoresDeAPrendizagem \EnfoqueMaisDiscussãoConversa	9
Aprendizagem \FacilitadoresDeAPrendizagem \ApresentaçãoDePontosDeVistasDistintosFortaleceAprendizado	5
Aprendizagem \FacilitadoresDeAPrendizagem \RespostasComplementares	9
Aprendizagem \FacilitadoresDeAPrendizagem \MaisObjetividade	4
Aprendizagem \FacilitadoresDeAPrendizagem \RespostasMaisElaboradas	4
Aprendizagem \FacilitadoresDeAPrendizagem \VariedadeDeTipoDeOrdenacaoRespostas	1
Aprendizagem \FacilitadoresDeAPrendizagem \PermiteRápidaCompreensão	1
Aprendizagem \RestringidoresDeAprendizagem	24
Aprendizagem \RestringidoresDeAprendizagem \EnfoqueNaSintaxe	2

Aprendizagem \RestringidoresDeAprendizagem \EnfoqueNaSintaxe \PerguntaInadequadaPara-Aprendizado	1
Aprendizagem \RestringidoresDeAprendizagem \LinguagemInadequada	3
Aprendizagem \RestringidoresDeAprendizagem \EstruturadaPorSintaxeAoInvésDaSemântica	2
Aprendizagem \RestringidoresDeAprendizagem \ForneceSoluçãoAoInvésDeConhecimento	2
Aprendizagem \RestringidoresDeAprendizagem \RespostaComplexaRequerMaisTempoEVeze-sAnálise	5
Aprendizagem \RestringidoresDeAprendizagem \TemPropagandasNoMeio	2
Aprendizagem \RestringidoresDeAprendizagem \TextoMaisSubjetivo	1
Aprendizagem \RestringidoresDeAprendizagem \InteresseNoMenorTamanhoDeTexto	2
Aprendizagem \RestringidoresDeAprendizagem \PoucoEstruturado	1
Aprendizagem \RestringidoresDeAprendizagem \FaltaDeReferência	3
Aprendizagem \MaterialComplementarDeAprendizado	5
Aprendizagem \MaterialComplementarDeAprendizado \PerguntaExploratóriaDeConhecimento	1
Aprendizagem \MaterialComplementarDeAprendizado \PerguntaExploratóriaDeConhecimento \PerguntaTransplataformaOUBases	1
Aprendizagem \MaterialComplementarDeAprendizado \PerguntaCaracterísticaDeCQA	4
Aprendizagem \MaterialComplementarDeAprendizado \PerguntaTécnicaEspecífica	11
Aprendizagem \PerfilUsuarioAprendiz	0
Aprendizagem \PerfilUsuarioAprendiz \Newbie	13
Aprendizagem \PerfilUsuarioAprendiz \InteressadoNoAssunto	13
Aprendizagem \PerfilUsuarioAprendiz \ComConhecimentoDoContextoOUPrevio	40

D. APÊNDICE D - Códigos-fonte em Python para processamento dos vídeos de participação

D.1 Conversão de Arquivos MP4 > MP3 > WAV

```

1  !pip3 install ffmpeg-python
2  # importing libraries
3  import os
4  import ffmpeg
5  import datetime
6  from __future__ import print_function
7
8  ffmpegCmd = 'C:/Apps/ffmpeg-2021-07-06-full_build/bin/ffmpeg_y_l_i_'
9
10 def converter(orig, dest, regera):
11     existe = os.path.exists(dest)
12     if existe and regera:
13         print('\t\tExcluindo_' + dest)
14         os.system('del_' + dest + ' "_/F_/Q')
15         comando = ffmpegCmd + ' _' + orig + ' _' + dest + ' "'
16         #print('\t\tExecutando ' + ffmpegCmd + orig + ' ' + dest)
17         res = os.system(comando)
18
19 def converterMP4toMP3(filefullMP4, filefulMP3, regera):
20     #filemp4 = pastaOrig + filename
21     #filemp3 = pastaTmp + filename.replace(".mp4", ".mp3")
22     converter(filefullMP4, filefulMP3, regera)
23
24 def converterMP3toWAV(filefullMP3, filefullWAV, regera):
25     #filemp3 = pastaTmp + filename.replace(".mp4", ".mp3")
26     #filewav = pastaDest + filename.replace(".mp4", ".wav")
27
28 ##### *****
29 ##### P R O C E S S A M E N T O           P R I N C I P A L
30 ##### *****
31 pathBase = ".\\Coleta\\"
32 participantes =
33     ["21510", "21610", "21616", "21710", "21715", "21720", "21818", "21820", "21915",
34     "22015", "22218", "22423", "22519", "22613", "30217", "31718", "31818"]
35 tipos = ["INI", "PRE", "S1", "S2", "S3", "S4", "POS"]
36 # Gerar para cada çacombinao de Participante e Tipo
37 for p in participantes:

```

```

38     for tipo in tipos:
39         participante= "Participante"+p
40         arq = participante+"_"+tipo+".mp4"
41         print('Arq.Selecionado:_', arq)
42         start = datetime.datetime.now()
43         print('\tlo_Mapear_Path:_)
44         filemp4 = pathBase+participante+"/"+arq
45         filemp3 = pathBase + "work/" + arq.replace(".mp4", ".mp3")
46         filewav = pathBase + "work/" + arq.replace(".mp4", ".wav")
47         if (not os.path.exists(filemp4)):
48             print('\tNo_existe_arquivo_MP4_!')
49             continue
50         start2 = datetime.datetime.now()
51         print('\t2o_Converter_MP4->MP3:_) + str(start2))
52         try:
53             if (not os.path.exists(filemp3) and not os.path.exists(filewav)):
54                 converterMP4toMP3(filemp4, filemp3, False)
55             else:
56                 print('\tArquivo_MP3/WAV_á_j_foi_gerado_antes_!')
57         except Exception as e:
58             print(arq + '\tErro_..._) + str(e))
59             continue
60         start3 = datetime.datetime.now()
61         print('\t3o_Converter_MP3->WAV:_) + str(start3))
62         try:
63             if (not os.path.exists(filewav)):
64                 converterMP3toWAV(filemp3, filewav, False)
65             else:
66                 print('\tArquivo_WAV_á_j_foi_gerado_antes_!')
67         except Exception as e:
68             print(arq + '\tErro_..._) + str(e))
69             continue
70         end = datetime.datetime.now()
71         elapsed = end - start
72         if (os.path.exists(filewav) and os.path.exists(filemp3)):
73             os.remove(filemp3)
74         print('Convertido_..._) + str(end) )
75         print('çãDurao:_) + str(elapsed) + "\n\n")

```

D.2 Extrair Texto de Audio e *Chunk*

```

1  !pip3 install playsound
2  !pip3 install pydub
3  !pip3 install SpeechRecognition
4
5  # importing libraries
6  import speech_recognition as sr
7  import os
8  import subprocess
9  from pydub import AudioSegment
10 from pydub.silence import split_on_silence, detect_silence, detect_nonsilent,
    detect_leading_silence
11 import datetime
12 from __future__ import print_function

```

```

13 # https://stackoverflow.com/questions/55805350/
    \when-why-does-split-on-silence-return-an-empty-list
14 import shutil
15
16 # create a speech recognition object
17 r = sr.Recognizer()
18
19 # https://cloud.google.com/speech-to-text/docs/languages
20 # https://www.analyticssteps.com/blogs/python-essentials-voice-control-part-1
21 def get_recognize_impl(impl, audio):
22     if impl == 'google':
23         return r.recognize_google(audio, language="pt-BR")
24     elif impl == 'bing':
25         return r.recognize_bing(audio, language="pt-BR")
26     elif impl == 'google_cloud':
27         return r.recognize_google_cloud(audio, language="pt-BR")
28     elif impl == 'hondify':
29         return r.recognize_hondify(audio, language="pt-BR")
30     elif impl == 'ibm':
31         return r.recognize_ibm(audio, language="pt-BR")
32     elif impl == 'wit':
33         return r.recognize_wit(audio, language="pt-BR")
34     else :
35         return r.recognize_sphinx(audio, language="pt-BR")
36
37 def get_audio_transcription(filewav, impl):
38     with sr.AudioFile(filewav) as source:
39         audio = r.record(source, duration=120)
40         texto = get_recognize_impl(impl, audio)
41         #r.recognize_google(audio, language="pt-BR"
42         return texto
43
44 # a function that splits the audio file into chunks # and applies speech recognition
45 def get_large_audio_transcription(path, impl, repera):
46     folder_name = path.replace(".wav", "-chunks")+impl
47     #print("folder_name:", folder_name)
48     print("\tValidando_êexistncia_de_processamento_anterior..._ " +
          str(datetime.datetime.now()))
49     if (os.path.isdir(folder_name)):
50         if (repera):
51             shutil.rmtree(folder_name)
52         else:
53             print("\t\t_áJ_existe_chunks..._")
54             return ""
55     if not os.path.isdir(folder_name):
56         # create a directory to store the audio chunks
57         os.mkdir(folder_name)
58     """
59     Splitting the large audio file into chunks
60     and apply speech recognition on each of these chunks
61     """
62     #print("path:", path)
63     #print("impl:", impl)
64     # open the audio file using pydub
65     print("\tQuebrando_AudioFile_em_Chunks..._ " + str(datetime.datetime.now()))
66     sound = AudioSegment.from_wav(path)

```

```

67     # split audio sound where silence is 700 miliseconds or more and get chunks
68     chunks = split_on_silence(sound,
69         # experiment with this value for your target audio file
70         min_silence_len = 1000,
71         # adjust this per requirement
72         silence_thresh = sound.dBFS-16, #14,
73         # keep the silence for 1 second, adjustable as well
74         keep_silence=500,
75     )
76
77     print("\tQuantidade_de_Chunks:_" + str(len(chunks)) + "_" +
78         str(datetime.datetime.now()))
79
80     todoTexto = ""
81     # process each chunk
82     for i, audio_chunk in enumerate(chunks, start=1):
83         # export audio chunk and save it in the 'folder_name' directory.
84         chunk_filename = os.path.join(folder_name, f"chunk{i}.wav")
85         audio_chunk.export(chunk_filename, format="wav")
86         # recognize the chunk
87         with sr.AudioFile(chunk_filename) as source:
88             audio_listened = r.record(source)
89             # try converting it to text
90             texto = ""
91             try:
92                 texto = get_recognize_impl(impl, audio_listened)
93                 #r.recognize_google(audio_listened, language="pt-BR")
94             except sr.UnknownValueError as e:
95                 todoTexto = todoTexto + "Error_chunk"+str(i)+".wav:_\n"
96                 print("\t\tTranscription_Error:_chunk"+str(i)+".wav_\n")
97             else:
98                 todoTexto = todoTexto+ "chunk"+str(i)+".wav:__" + texto.capitalize()+
99                     "\n"
100
101     # return the text for all chunks detected
102     # print (todoTexto)
103     return todoTexto
104
105 ##### *****
106 #####          P R O C E S S A M E N T O          P R I N C I P A L
107 ##### *****
108 pathBase = ".\\Coleta\\"
109 pathWork = pathBase + "work\\"
110 participantes =
111     ["21510", "21610", "21616", "21710", "21715", "21720", "21818", "21820", "21915",
112         "22015", "22218", "22423", "22519", "22613", "30217", "31718", "31818"]
113 tipos = ["INI", "PRE", "S1", "S2", "S3", "S4", "POS"]
114
115 #thislist = ["hondify", "ibm", "wit", "sphinx", "google_cloud", "bing", "google"]
116 impl = "google"
117 global multitexto
118
119 for p in participantes:
120     for tipo in tipos:
121         multitexto = ""
122         participante= "Participante"+p
123         arg = participante+"_"+tipo+".wav"

```

```

120 # -----
121 filewav = pathWork + arq
122 pastaChunkWork = filewav.replace(".wav","-chunks")+impl
123 arqGtranscr = pathBase + participante + "\\\" +
    arq.replace(".wav","."+impl+".txt")
124 # -----
125 print('Arq.Selecionado:_',arq)
126 if (not os.path.exists(filewav)):
127     print('\t_ãNo_existe_arquivo_WAV_!')
128     continue
129 start = datetime.datetime.now()
130 # -----
131 try:
132     print('°4_Quebrar_o_WAV_em_chunks_e_transcrever_com_[ ' + impl + ' ]_...')
133     multitexto = get_large_audio_transcription(filewav,impl, False)
134     if (len(multitexto) < 2):
135         print('\t_Transcricao_ãno_realizada_!')
136         continue
137     print('°5_Salvando_a_çãtranscrio_...')
138     with open(arqGtranscr, 'w') as outfile:
139         #for texto in multitexto:
140             outfile.write(multitexto + "\n")
141             ##print(texto)
142 except Exception as e:
143     print('\t_Ocorreu_um_Erro_..._' + str(e))
144     continue
145 # -----
146 end = datetime.datetime.now()
147 elapsed = end - start
148 print('Transcrito..._' + str(end) )
149 print('çãDurao:' + str(elapsed) + "\n\n")
150 #if (os.path.exists(pastaChunkWork)):
151     # os.rmdir(pastaChunkWork)

```

D.3 Transformar *Chunks* com Texto Traduzido automaticamente para Arquivo Legendado

Transforma o texto transcrito automaticamente pelo *Google Translator*, considerando os segmentos *Chunk* associados, em um formato para *Software Aegisub*, que permitirá a edição dos textos pelo operador, corrigindo possíveis problemas e identificando os atores.

```

1 !pip3 install librosa
2
3 import librosa
4 import time
5 # import OS module
6 import os
7 import os.path
8
9 def getSessao(tipo):
10     if(tipo=='S1'):

```

```

11         return 'SESSAO#1'
12     elif(tipo=='S2'):
13         return 'SESSAO#2'
14     elif(tipo=='S3'):
15         return 'SESSAO#3'
16     elif(tipo=='S4'):
17         return 'SESSAO#4'
18     else:
19         return tipo
20
21 def obterTermoChunk(linha):
22     pChunk = linha.find("chunk")
23     pWav = linha.find(".wav")
24     if (pChunk == -1 or pWav == -1):
25         return "chunk0.wav"
26     else:
27         ptam = pWav - pChunk + 4
28         #print( pChunk, pWav,ptam, linha[pChunk:pChunk+ptam])
29         return linha[pChunk:pChunk+ptam]
30
31 # Montador de Cabecalho
32 def salvarCabecalho(tipo, videoFile, outfile):
33     outfile.write(f'{"[Script_Info]"}\n')
34     outfile.write(f'{"[_Script_generated_by_Aegisub_3.2.2]"}\n')
35     outfile.write(f'{"[_http://www.aegisub.org/"]"}\n')
36     outfile.write(f'{"Title:_Default_Aegisub_file"}\n')
37     outfile.write(f'{"ScriptType:_v4.00+"}\n')
38     outfile.write(f'{"WrapStyle:_0"}\n')
39     outfile.write(f'{"ScaledBorderAndShadow:_yes"}\n')
40     outfile.write(f'{"YCbCr_Matrix:_TV.601"}\n')
41     outfile.write(f'{"PlayResX:_1280"}\n')
42     outfile.write(f'{"PlayResY:_720"}\n')
43     outfile.write(f'{""}\n')
44     outfile.write(f'{"[Aegisub_Project_Garbage]"}\n')
45     outfile.write(f'{"Last_Style_Storage:_+_getSessao(tipo)_"}\n')
46     outfile.write(f'{"Audio_File:_+_videoFile"}\n')
47     outfile.write(f'{"Video_File:_+_videoFile"}\n')
48     outfile.write(f'{"Video_AR_Mode:_4"}\n')
49     outfile.write(f'{"Video_AR_Value:_1.777778"}\n')
50     outfile.write(f'{"Video_Zoom_Percent:_0.500000"}\n')
51     outfile.write(f'{"Scroll_Position:_38"}\n')
52     outfile.write(f'{"Active_Line:_46"}\n')
53     outfile.write(f'{"Video_Position:_75312"}\n')
54     outfile.write(f'{""}\n')
55     outfile.write(f'{"[V4+_Styles]"}\n')
56     outfile.write(f'{"Format:_Name,_Fontname,_Fontsize,_PrimaryColour,_
        SecondaryColour,_OutlineColour,_BackColour,_Bold,_Italic,_Underline,_
        StrikeOut,_ScaleX,_ScaleY,_Spacing,_Angle,_BorderStyle,_Outline,_Shadow,_
        Alignment,_MarginL,_MarginR,_MarginV,_Encoding"}\n')
57     outfile.write(f'{"Style:_
        Default,20,&H00FFFFFF,&H000000FF,&H00000000,&H00000000,0,0,0,0,100,100,
        0,0,1,2,2,2,10,10,10,1"}\n')
58     outfile.write(f'{"Style:_
        PRE,Arial,20,&H00F844F,&H000000FF,&H00000000,&H00000000,0,0,0,0,100,100,
        0,0,1,2,2,2,10,10,10,1"}\n')
59     outfile.write(f'{"Style:_

```

```

        SESSAO#1,Arial,20,&H00FFFFFF,&H000000FF,&H00000000,&H00000000,0,0,0,0,100,100,
        ,0,0,1,2,2,2,10,10,10,1"}\n')
60 outfile.write(f'{"Style:
        SESSAO#2,Arial,14,&H006E6F70,&H000000FF,&H00000000,&H00000000,0,0,0,0,100,100,
        ,0,0,1,2,2,2,10,10,10,1"}\n')
61 outfile.write(f'{"Style:
        SESSAO#3,Arial,20,&H00F1E612,&H000000FF,&H00000000,&H00000000,0,0,0,0,100,100,
        ,0,0,1,2,2,2,10,10,10,1"}\n')
62 outfile.write(f'{"Style:
        SESSAO#4,Arial,20,&H00F115E0,&H000000FF,&H00000000,&H00000000,0,0,0,0,100,100,
        ,0,0,1,2,2,2,10,10,10,1"}\n')
63 outfile.write(f'{"Style:
        POS,Arial,20,&H00FFFFFF,&H000000FF,&H00000000,&H00000000,0,0,0,0,100,100,
        ,0,0,1,2,2,2,10,10,10,1"}\n')
64 outfile.write(f'{""}\n')
65 outfile.write(f'{"[Events]"}\n')
66 outfile.write(f'{"Format:
        Layer,
        Start,
        End,
        Style,
        Name,
        MarginL,
        MarginR,
        MarginV,
        Effect,
        Text"}\n')
67
68 def gerarDicionarioChunks(pathTemp, pathTipo):
69     dictChunk = {}
70     segChunk = 0;
71     # Processando pastas com google Chunks
72     for (root, dirs, file) in os.walk(pathTemp):
73         for f in file:
74             if '.wav' in f:
75                 segChunk = librosa.get_duration(filename=pathTemp + "/" + f)
76                 dictChunk[f]=
77                     [int(f.replace("chunk","").replace(".wav","")), [f, segChunk, ',', ',']]
78
79     ## ---
80     ##print(dictChunk)
81     # Ordenando os Chunks
82     lista = [None] * len(dictChunk)
83     for item in dictChunk:
84         val = dictChunk.get(item)
85         lista[val[0]-1] = val[1]
86     ## ---
87     # Computando inicio e fim dos Chunks
88     segAcum = 0;
89     for chunk in lista:
90         ini = (time.strftime('%H:%M:%S.',
91                             time.gmtime(segAcum))+str('%03f'%segAcum).split('.')[1])
92         segAcum = segAcum + chunk[1]
93         end = (time.strftime('%H:%M:%S.',
94                             time.gmtime(segAcum))+str('%03f'%segAcum).split('.')[1])
95         texto = 'Dialogue:
96         ,0,' + ini + ',' + end + ',' + pathTipo + ',ALL,0,0,0,TODO,'
97         chunk[2] = ini
98         chunk[3] = end
99         dictChunk[chunk[0]] = texto
100     return dictChunk

```

```

def montarArquivoLegendaASS(tipo, arqFullOrig, arqFullDest, retirar, videoFile,
dictChunk):
    ultChunk = 0
    conta = 0
    with open(arqFullOrig) as infile, open(arqFullDest, 'w') as outfile:

```

```

101     salvarCabecalho(tipo, videoFile, outfile)
102     for line in infile:
103         src = obterTermoChunk(line)
104         target = dictChunk.get(src)
105
106         if(target):
107             print("EH:",ultChunk,src, target)
108         else:
109             print("NOT:[" +src+"]_", line)
110             continue
111
112         line = line.replace(src, target)
113         if(line.startswith("Error")):
114             #print("Erro")
115             line = line.replace("TODO","Error_" + src)
116             line = line[6:]
117         else:
118             line = line.replace("TODO", src)
119             line = line.replace("_:_","")
120             #print(line)
121             if(len(line)>1 and ("chunk" in line)):
122                 outfile.write(line)
123
124     ##### *****
125     #####          P R O C E S S A M E N T O          P R I N C I P A L
126     ##### *****
127     pathBase = ".\\Coleta\\"
128     participantes =
129         ["21510","21610","21616","21710","21715","21720","21818","21820","21915",
130             "22015","22218","22423","22519","22613","30217","31718","31818"]
131     tipos =["INI","PRE","S1","S2","S3","S4","POS"]
132
133     for p in participantes:
134         # áVariveis de trabalho
135         participante= "Participante"+p
136
137         for tipo in tipos:
138             pathTemp = pathBase + "work\\"+participante + "_" + tipo+"-chunksgoogle\\"
139             arquivoOrigem = pathBase + participante + "\\"+participante + "_" + tipo +
140                 ".google.txt"
141             arquivoDestino = pathBase + participante + "\\"+participante + "_" + tipo +
142                 ".google.ass"
143
144             video = participante + "_" + tipo + ".mp4"
145             textoExcluir = participante + "_" + tipo+"-chunksgoogle\\"
146
147             if (os.path.exists(pathTemp) and os.path.exists(arquivoOrigem)):
148                 dictoSegmentosAudio = gerarDicionarioChunks(pathTemp, tipo)
149                 #print(dictoSegmentosAudio)
150                 montarArquivoLegendaASS(tipo,arquivoOrigem, arquivoDestino, textoExcluir,
151                     video, dictoSegmentosAudio)
152                 print("OK:_", len(dictoSegmentosAudio), arquivoOrigem, arquivoDestino,
153                     os.stat(arquivoDestino))
154             else:
155                 print('ãNo_existe:_\n\t' + pathTemp + "_\n\t_ou_" + arquivoOrigem +"\n\n")

```

D.4 Transformar Arquivo Legendado em Arquivo de Transcrição QDA

```

1 import time
2 import os
3
4 def tratarAtor(ator):
5     if(ator=="ALL"):
6         ator="Todos"
7     return ator
8
9 def tratarAcao(acao):
10     #if(acao=="ALL"):
11     #    acao="Todos"
12     return acao
13
14 def tratarGrupo(tipo, tipoAnt):
15     if(tipo=="S1"):
16         tipo="SESSAO_#01"
17     elif(tipo=="S2"):
18         tipo="SESSAO_#02"
19     elif(tipo=="S3"):
20         tipo="SESSAO_#03"
21     elif(tipo=="S4"):
22         tipo="SESSAO_#04"
23     elif(tipo=="PRE"):
24         tipo="ENTREVISTA_ÉPR"
25     elif(tipo=="POS"):
26         tipo="ENTREVISTA_ÓPS"
27     elif(tipo=="Default"):
28         tipo= tipoAnt
29     return tipo
30
31 def tratarLinha(linha, tipoAnt):
32     '''
33     0      ,      1,      2,      3,      4, 5, 6, 7,      8,      9
34     Rotulo, Start, finish, grupo, ator, x, y, z, Tipo, Texto
35     Dialogue: 0,0:03:10.39,0:03:10.43,PRE,Participante,0,0,0,Fala,Tem duas ou êtrs
36               comunidades que ão çconheo.
37     Comment: 0,0:04:04.43,0:04:34.60,SESSAO#1,Pesquisador,0,0,0,Protocolo,çãApresentao
38               da ãSesso 1 - Livre
39     '''
40     texto= ""
41     tipo=""
42     ator=""
43     end=""
44     start=""
45     acao=""
46     te=[]
47     temConteudo = False
48     if (linha.__contains__("Comment:")):
49         linha = linha.replace("Comment:", "")
50         linha = linha.strip()
51         temConteudo = True
52     elif (linha.__contains__("Dialogue:")):
53         linha = linha.replace("Dialogue:", "")
54         linha = linha.strip()

```

```

53         temConteudo = True
54     if (temConteudo):
55         partes = linha.split(',')
56         i = 0
57         for parte in partes:
58             if (i==1):
59                 start = parte.strip()
60                 start = start[:len(start)-1]
61             elif (i==2):
62                 end = parte.strip()
63                 end = end[:len(end)-1]
64             elif (i ==3):
65                 tipo = tratarGrupo(parte.strip(), tipoAnt)
66             elif(i==4):
67                 ator = tratarAtor(parte.strip())
68             elif(i==8):
69                 acao = tratarAcao(parte.strip())
70             elif(i >=9):
71                 if(texto==""):
72                     texto = texto + parte.strip()
73                 else:
74                     texto = texto + ",_" + parte.strip()
75                     i=i+1
76             ##Fim do For
77             te = [start, end, tipo, ator, acao, texto]
78             return te
79
80     ##### *****
81     #####          P R O C E S S A M E N T O          P R I N C I P A L
82     ##### *****
83     pathPart= "Participante22519"
84     segmento = "S4"
85
86     pathBase= ".\\Coleta\\"
87     if (segmento == ""):
88         pathOrig = pathBase + pathPart + "\\ "+pathPart
89     else:
90         pathOrig = pathBase + pathPart + "\\ "+pathPart + "_" + segmento
91     print(pathOrig)
92
93     arqFullOrig = pathOrig + ".ass"
94     arqFullDest = pathOrig + ".transcript.txt"
95     correnteAtor = ""
96     correnteGrupo = ""
97     linhas = 0
98     tipoAnt = ""
99
100     print(arqFullOrig)
101     print(arqFullDest)
102     with open(arqFullOrig, encoding="utf-8") as infile, open(arqFullDest, 'w',
103         encoding="utf-8") as outfile:
104         for line in infile:
105             conteudo = tratarLinha(line, tipoAnt)
106             if(conteudo== None):
107                 continue
108             novaLinha = conteudo[0] ##+ "," + conteudo[1]

```

```
108     if (correnteGrupo != conteudo[2]):
109         novaLinha = novaLinha + "\n_____ " + conteudo[2] + " _____\n"
110         correnteAtor = ""
111         correnteGrupo = conteudo[2]
112     if (correnteAtor != conteudo[3]):
113         novaLinha= novaLinha + "\n" + conteudo[3] + " :_/ " + conteudo[4] + " /_\n\t"
114         + conteudo[5]
115         #print(novaLinha)
116         #break
117         outfile.write(f'{novaLinha}\n')
117 print ("Fim_->", arqFullDest )
```