

UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS  
UNIDADE ACADÊMICA DE EDUCAÇÃO ONLINE  
ESPECIALIZAÇÃO EM BIG DATA, DATA SCIENCE E DATA ANALYTICS

Roger Denis Vieira

O USO DE DADOS PSICOFISIOLÓGICOS COMO RECURSO DE  
MELHORIA NO CONTEXTO DE ENGENHARIA DE SOFTWARE:  
UM MAPEAMENTO SISTEMÁTICO

São Leopoldo

2019

UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS  
UNIDADE ACADÊMICA DE EDUCAÇÃO ONLINE  
ESPECIALIZAÇÃO EM BIG DATA, DATA SCIENCE E DATA ANALYTICS

Roger Denis Vieira

O USO DE DADOS PSICOFISIOLÓGICOS COMO RECURSO DE  
MELHORIA NO CONTEXTO DE ENGENHARIA DE SOFTWARE:  
UM MAPEAMENTO SISTEMÁTICO

Trabalho de Conclusão de Curso de Especialização apresentado como requisito parcial para obtenção do título de Especialista em Big Data, Data Science e Data Analytics, pelo Curso de Especialização em Big Data, Data Science e Data Analytics da Universidade do Vale do Rio dos Sinos – UNISINOS.

Orientador: Prof. Dr. Kleinner Silva Farias de Oliveira

São Leopoldo

2019

# O Uso de Dados Psicofisiológicos como Recurso de Melhoria no Contexto de Engenharia de Software: Um Mapeamento Sistemático

Roger Denis Vieira

<sup>1</sup>Universidade do Vale do Rio dos Sinos (UNISINOS)  
Av. Unisinos, 950, Bairro Cristo Rei, São Leopoldo, RS – Brasil

**Abstract.** *A systematic mapping had been developed seeking for answering eight research questions regarding using of psychophysiological data in Software Engineering domain. By means of it, was selected 2084 studies from 7 digital databases, reaching a final amount of 27 articles. Mostly of experimental nature, marked by usage the sensing devices and integrated development environments (IDEs), looking for analysis of developers' psychophysiological attributes applied to comprehension and debugging of source code. It was not identified native mechanisms to support psychophysiological data integration and analysis whithin integrated development environment.*

**Resumo.** *Foi desenvolvido um mapeamento sistemático buscando responder oito questões de pesquisa acerca da utilização de dados psicofisiológicos no contexto da Engenharia de Software. Por meio deste, foram pré selecionados 2084 estudos obtidos em 7 bases de dados digitais, chegando-se a um montante final de 27 trabalhos. Em sua maioria, de natureza experimental, caracterizada pelo uso de dispositivos sensores e ambientes integrados de desenvolvimento (IDEs), visando a análise de atributos psicofisiológicos dos desenvolvedores aplicados à compreensão e depuração de código fonte. Não foram identificados mecanismos de suporte à integração e análise dos dados psicofisiológicos de forma nativa nos ambientes integrados de desenvolvimento.*

## 1. Introdução

O desenvolvimento de software é um processo lógico e intelectual, onde cada desenvolvedor aplica o seu conhecimento embasado por premissas lógicas para a produção de soluções voltadas a problemas reais, seja na área corporativa, acadêmica, ou ainda, em problemas corriqueiros do cotidiano. Tal desenvolvimento demanda que o profissional esteja focado e imerso em seu trabalho, atentando a cada detalhe que possa impactar no comportamento da solução que está sendo produzida [Duraes et al. 2016].

Não obstante os investimentos em torno do processo de desenvolvimento de software, suas metodologias e tecnologias, é inevitável a geração de defeitos inerentes ao comportamento do produto a medida que a sua complexidade aumenta. Este fenômeno pode ser atribuído a diversos fatores, como o processo intelectual envolvido na produção de código, o conjunto de ferramentas e técnicas inadequadas, além de problemas intrínsecos às definições de requisitos atrelados ao negócio. Sob a perspectiva em torno dos fatores humanos, autores advogam que o desenvolvedor, enquanto ser humano, é passível de falhas, e que estas são naturais e esperadas durante a execução das suas tarefas [Duraes et al. 2016].

Estudos buscam compreender quais processos cognitivos estão envolvidos durante a produção de software, fazendo o uso de tecnologias como Eletroencefalografia (EEG) e Imagem por Ressonância Magnética Funcional (fMRI) [Siegmund et al. 2014, Duraes et al. 2016, Floyd et al. 2017]. Os autores destes estudos advogam que é possível encontrar relações entre a produção de defeitos em artefatos de software e os processos cognitivos envolvidos no momento da escrita de código, visando criar mecanismos capazes de priorizar a refatoração de código de baixa confiabilidade e até a predição na confecção de possíveis artefatos defeituosos.

Com o advento da microeletrônica, processadores menores, mais velozes e de baixo consumo energético têm surgido, possibilitando que dispositivos móveis e acessíveis para a coleta de sinais biológicos sejam difundidos em experimentos acadêmicos. Esses são os casos do NeuroSky Mindwave Mobile 2 [Rostami et al. 2015] e o Emotiv EPOC e EPOC+ [Radevski et al. 2015] que são utilizados para aumentar a compreensão acerca dos processos cognitivos envolvidos durante o processo de codificação. Aliados à extensibilidade dos atuais Ambientes Integrados de Desenvolvimento, a utilização de tais dispositivos pode permitir o desenvolvimento de novas funcionalidades que permitam a identificação de padrões de comportamento dos desenvolvedores acerca da sua produtividade e qualidade de artefatos produzidos [Fritz e Muller 2016a, Radevski et al. 2015, Zayour e Hajjdiab 2013].

Na última década, dados psicofisiológicos têm sido amplamente utilizados como recurso de melhoria no contexto de engenharia de software. No entanto, a literatura atual carece de uma classificação e um mapeamento sistemático dos trabalhos já realizados e publicados abordando este tema [Floyd et al. 2017, Gonçalves et al. 2019a]. Consequentemente, um entendimento mais detalhado do uso de tais dados permanece ainda limitado. Apoiado por tal argumento, o presente estudo visa responder à seguinte questão de pesquisa: *como dados psicofisiológicos estão sendo utilizados como recurso de melhoria no contexto de Engenharia de Software?*

Este trabalho, portanto, foca em classificar e fornecer uma análise temática dos estudos já publicados sobre o uso de dados psicofisiológicos como recurso de melhoria no contexto de Engenharia de Software. Para isso, um mapeamento sistemático da literatura foi projetado e realizado, seguindo diretrizes bem estabelecidas para responder a oito questões de pesquisa. No total, 2084 estudos foram inicialmente obtidos de 7 bases de dados digitais, sendo 27 estudos selecionados para responder às questões de pesquisa formuladas. Como resultados, observou-se que a maioria dos estudos como sendo de natureza experimental, caracterizada pelo uso de dispositivos para coleta de dados psicofisiológicos e ambientes integrados de desenvolvimento (IDEs), visando a análise de atributos psicofisiológicos dos desenvolvedores aplicados, principalmente, à compreensão e depuração de código fonte. Ademais, não foram identificados mecanismos de suporte à integração e análise de dados psicofisiológicos de forma nativa nos ambientes integrados de desenvolvimento. Por fim, este trabalho identifica alguns desafios, os quais poderão ser explorados pela comunidade científica nos próximos anos.

Inicialmente este estudo apresenta uma revisão acerca do embasamento teórico adotado para o seu desenvolvimento (Seção 2), além de enumerar e sintetizar os principais trabalhos correlatos à temática proposta (Seção 3). Em seguida, são detalhadas as etapas escolhidas para o desenvolvimento do mapeamento sistemático (Seção 4), metodologia

esta adotada para a condução do estudo. Posteriormente, a condução do processo de seleção, filtragem e extração dos dados contidos nos estudos selecionados é apresentada (Seção 5), assim como os resultados encontrados (Seção 6), os quais, juntos aos desafios identificados, são discutidos à luz da literatura atual (Seção 7). Por fim, são explicitadas as considerações finais, assim como uma visão prévia de estudos futuros (Seção 8).

## 2. Referencial Teórico

Neste segmento do trabalho são apresentados todos os conceitos-chaves envolvidos na compreensão dos assuntos abordados e confecção do estudo. Os conceitos utilizados foram identificados e referenciados fazendo o uso de trabalhos já desenvolvidos na área acerca da mesma temática, os quais serão detalhados na Seção 3.

### 2.1. Métricas e Dados Psicofisiológicos

Os dados psicofisiológicos<sup>1</sup> referem-se ao conjunto de métricas relacionadas aos processos fisiológicos do corpo humano que surgem como resposta a estímulos oriundos de situações específicas. Tal conjunto de métricas podem ser exploradas de forma a correlacionar as suas variações a determinados estados mentais ou processos cognitivos de um indivíduo por meio de rótulos informacionais que lhe são atribuídos quando este é exposto à determinadas situações controladas [Li e Jain 2015, Andreassi 2007]. Nas próximas subseções, serão apresentadas as principais métricas utilizadas na identificação de dados psicofisiológicos que podem ser encontradas com determinada frequência literatura especializada.

#### 2.1.1. Métricas Oculares

O conjunto de métricas oculares são obtidos fazendo o uso de dispositivos sensores conhecidos por *Eye-trackers*. Estes dispositivos caracterizam-se por utilizarem feixes de laser ou luz no espectro infravermelho para mensurar uma série de atributos do relacionado aos globos oculares do indivíduo, tais como a trajetória dos movimento dos olhos, grau de dilatação da pupila, duração de ponto focal e detecção do piscar de olhos.

Quando aplicados em determinados contextos, os dados coletados por *Eye-trackers* podem indicar uma série de processos psicofisiológicos, tais como carga cognitiva, excitação, valência, atenção, ansiedade e *stress*. O conhecimento sob tais condições psicofisiológicas tem sido empregado no domínio de Engenharia de Software em processos como o desenvolvimento de programas [Fritz et al. 2014], estratégias de depuração [Lin et al. 2016], usabilidade de Ambientes Integrados de Desenvolvimento (IDEs), dificuldades no aprendizado de novas linguagens de programação, entre outros [Hejmady e Narayanan 2012].

---

<sup>1</sup>Na literatura científica, comumente nos trabalhos escritos em inglês, pode-se encontrar o termo *biometric data* (ou *dados biométricos*, em tradução livre), entretanto, neste trabalho, adotou-se a terminologia *dados psicofisiológicos* a fim de reduzir a ambiguidade com o processo de atribuição de identidade a indivíduos por meio de dados biológicos, ou *biometria*. N. do A.

### 2.1.2. Métricas Cerebrais

A atividade dos neurônios no cérebro durante a execução de tarefas diversas desencadeia uma série de flutuações na diferença de potencial, ou tensão, por toda a superfície craniana, a qual pode ser mensurada utilizando-se a técnica de eletroencefalografia (EEG) [Fritz e Muller 2016a, Andreassi 2007]. Pesquisadores identificaram a possibilidade de traduzir a atividade neural e suas flutuações de tensão em um espectro de ondas divididas em diferentes bandas de frequência, podendo cada banda ser relacionada a um respectivo conjunto de atividades (Tabela 1). A onda alfa ( $\alpha$ ), por exemplo, pode ser observada nos indivíduos em estado de relaxamento ou meditação, atenuando-se a medida que atividades físicas e/ou mentais vão aumentando [Andreassi 2007, Fritz e Muller 2016a]. As medidas advindas da atividade eletrodérmica, assim como da temperatura superficial da pele, já foram correlacionadas na literatura com processos psicológicos, como a excitação e algumas emoções específicas [Fritz e Muller 2016a]

**Tabela 1. Espectro de ondas cerebrais e atividades frequentemente relacionadas.**

Onda	Símbolo	Banda	Frequentemente Ligada a:
Alfa	$\alpha$	8 – 12Hz	relaxamento, meditação, reflexão
Beta	$\beta$	12 – 30Hz	estado de alerta, foco
Gama	$\gamma$	31 – 120Hz	atividade cerebral intensa, aprendizado
Delta	$\delta$	0,4 – 4Hz	sono profundo
Teta	$\theta$	4 – 7Hz	sonolência, meditação

Assim como as flutuações de tensão na superfície craniana, as atividades neurais desencadeiam outros processos biológicos de acordo com as atividades físicas e mentais exercidas, entre eles estão a hemodinâmica cerebral e a oxigenação. Tais processos podem ser medidos por meio de técnicas como a Espectroscopia por Infravermelho Próximo (fNIRS<sup>2</sup>) e Imagem por Ressonância Magnética Funcional (fMRI<sup>3</sup>). Ao contrário da eletroencefalografia, as técnicas citadas são capazes de gerar imagens claras na atividade cerebral em regiões bem definidas, possibilitando assim as suas relações com processos físicos e cognitivos [Fishburn et al. 2014, Li e Lu 2009, Solovey et al. 2015].

### 2.1.3. Métricas Dérmicas

As métricas dérmicas são o conjunto de medidas relacionadas às variações de propriedades presentes na superfície da pele de um indivíduo, tais como a sua temperatura e a atividade eletrodérmica (EDA<sup>4</sup>), frequentemente encontrada na literatura sob o termo de resposta galvânica da pele (GSR<sup>5</sup>). Os sensores EDA usualmente coletam dados oriundos da variação da resistência da pele em resposta a algum estímulo externo, por exemplo, quando o indivíduo apresenta-se em excitação, as suas glândulas sudoríparas aumentam a produção de suor, que conseqüentemente reduzirá a resistência da pele, aumentando por sua vez a condutibilidade elétrica da pele [Fritz e Muller 2016a, Haag et al. 2004].

<sup>2</sup>functional Near Infrared Spectroscopy

<sup>3</sup>functional Magnetic Resonance Imaging

<sup>4</sup>Electrodermal Activity

<sup>5</sup>Galvanic Skin Response

#### 2.1.4. Métricas Cardiorrespiratórias

No processo de revisão dos estudos utilizados no referencial teórico, foram identificados quatro tipos de métricas relacionadas às características cardiorrespiratórias dos indivíduos analisados nos experimentos: pulso do volume sanguíneo (BVP<sup>6</sup>), a frequência cardíaca (HR<sup>7</sup>), a variabilidade da frequência cardíaca (HRV<sup>8</sup>) e a frequência respiratória (RR<sup>9</sup>). A frequência cardíaca refere-se ao número de vezes que o músculo cardíaco realiza um ciclo cardíaco (movimento sistólico e diastólico) no período de um minuto, enquanto a variabilidade da frequência cardíaca indica a variação de tempo entre dois ciclos consecutivos. O pulso do volume sanguíneo, também mencionado na literatura sob o termo "pulso", mede o fluxo sanguíneo em partes específicas do corpo de acordo com mudanças comandadas pelo sistema nervoso simpático, especialmente em situações estressoras [Andreassi 2007, Fritz e Muller 2016a]. As métricas apresentadas nesta categoria estão fortemente ligadas à carga cognitiva e diversas emoções [Fritz e Muller 2016b].

#### 2.2. Teoria da Carga Cognitiva

O corpo humano além de aportar recursos físicos constantemente, também com faz com os chamados recursos mentais – como memória, atenção e raciocínio – objetivando a manutenção das funções vitais do indivíduo. Quando necessários, tais recursos podem ser também alocados de forma específica com a finalidade de possibilitar a execução de tarefas iniciadas pelo indivíduo, como por exemplo, a resolução de problemas lógicos envolvidos no processo de desenvolvimento de software.

Ao montante de recursos mentais, em especial a memória de trabalho, que são demandados durante o esforço mental necessário ao desempenhar um conjunto de atividades cognitivas simples e complexas, atribui-se o termo *Carga Cognitiva* [Sweller et al. 2011]. Dada a sua natureza, a Carga Cognitiva ainda pode ser dividida em três tipos especializados:

1. **Carga Cognitiva Intrínseca:** está atrelada ao nível de dificuldade associado a um tópico instrucional simples. Todas as instruções possuem um grau inerente de dificuldade associada, sendo que o nível de dificuldade não pode ser alterado pelo agente instrucional. Porém, o esquema mental que é utilizado para representar uma determinada tarefa pode ser fracionado em esquemas menores, de forma a serem tratados isoladamente e posteriormente reagrupados sem comprometer semanticamente o esquema original.
2. **Carga Cognitiva Extrínseca:** este tipo de carga cognitiva pode ser gerado de acordo com a forma em que a informação é apresentada por parte do agente instrucional àquele que executará determinada atividade. Diferentes formas de apresentação de um tópico podem gerar níveis diferentes de carga cognitiva extrínseca.
3. **Carga Cognitiva Pertinente:** diferente dos demais tipos de cargas cognitivas apresentadas, a carga cognitiva pertinente pode ser considerada como um tipo *artificial* de carga cognitiva. Ela é resultante a criação de esquemas mentais para

---

<sup>6</sup>*Blood Volume Pulse*

<sup>7</sup>*Heart Rate*

<sup>8</sup>*Heart Rate Variability*

<sup>9</sup>*Respiratory Rate*

representação de determinado conjunto instrucional, seja por meio de automação, analogias ou desconstrução e reconstrução de conceitos, a fim de facilitar o processo de compreensão e aprendizagem em torno de determinado tópico.

Em suma, a carga cognitiva de um desenvolvedor, por exemplo, pode ser afetada por fatores como a apresentação do código em um Ambiente Integrado de Desenvolvimento (IDEs), o seu domínio sobre determinada linguagem de programação, o seu conhecimento em torno de erros gerados e a sua compreensão da solução de software como um todo [Fritz e Muller 2016a].

### **2.3. Ambientes Integrados de Desenvolvimento**

Os Ambientes Integrados de Desenvolvimento (IDEs<sup>10</sup>) podem ser considerados como o principal conjunto de ferramentas que os desenvolvedores utilizam, tanto para tarefas de criação como para manutenção de código-fonte. Este conjunto de ferramentas engloba editores de código-fonte, navegação por arquivos, compiladores e depuradores, assistentes para criação e refatoração de artefatos, ferramentas destaque de sintaxe (*Syntax Highlighting*), entre outras funcionalidades [Snipes et al. 2015]. Para o desenvolvedor é essencial, além do domínio sobre a linguagem de programação em questão e os aspectos lógicos envolvidos no problema, possuir conhecimento sobre as ferramentas que serão utilizadas [Astromskis et al. 2017].

Nos últimos anos, autores têm realizado pesquisas a respeito do uso de IDEs por parte dos desenvolvedores, tanto no processo de desenvolvimento [Astromskis et al. 2017, Zayour e Hajjdiab 2013], como na sua eficácia enquanto ferramenta de depuração [Zayour e Hamdar 2016]. Usualmente empregando *Eye-trackers* para coleta de dados quantitativos, os autores avaliam aspectos como a facilidade de uso, a organização da interface gráfica, o conjunto de ferramentas disponíveis, entre outros recursos variantes de acordo com o fornecedor. Visando a evolução das ferramentas e facilidade no desenvolvimento de estudos, alguns autores têm produzido trabalhos com o objetivo de facilitar a integração de dados biométricos, como o oriundo de *Eye-trackers*, dentro dos Ambientes Integrados de Desenvolvimento, de forme que estes dados possam ser relacionados ao código-fonte gerado, ou ainda, à interação do desenvolvedor com a ferramenta [Guarnera et al. 2018].

### **2.4. Dados Psicofisiológicos na Engenharia de Software**

Os dados psicofisiológicos possuem grande aplicabilidade no que se refere a compreensão do que os desenvolvedores experienciam enquanto estão realizando o seu trabalho, seja durante um ciclo intenso de trabalho, enquanto apresentam pleno foco no desenvolvimento, quanto nos momentos em que estes enfrentam algum tipo de dificuldade no que tange a compreensão de algum trecho de código ou ainda no processo lógico de concepção de soluções [Fritz e Muller 2016a].

No contexto da Engenharia de Software, a utilização de dados psicofisiológicos tem se mostrado emergente que se refere ao entendimento dos processos cognitivos envolvidos no durante o desenvolvimento de software e seus correlatos [Fritz e Muller 2016a]. Estudos nesta área têm explorado desde a idealização de novos dispositivos sensores de baixo custo que sejam capazes de realizar medições fidedignas de sinais psicofisiológicos,

---

<sup>10</sup>do inglês, *Integrated Development Environment*

até a análise de como desenvolvedores novatos e experientes compreendem a escrita do código-fonte [Lee et al. 2016, Szu et al. 2013] e como este aspecto pode auxiliar na melhoria da produção de artefatos de software [Radevski et al. 2015].

Apesar do aumento na produção de estudos nesta área, alguns autores advogam a baixa consistência e falta de consenso apresentado nestes, sendo assim, buscam realizar levantamentos que tragam uma visão atual e consolidada em torno da utilização de dados psicofisiológicos no contexto da Engenharia de Software [Gonçales et al. 2019a, Obaidellah et al. 2018]. Mediante o desenvolvimento de mapeamento e revisões sistemáticas, os autores almejam identificar lacunas na temática que possam servir de insumo para a confecção de novos estudos capazes de atender às problemáticas apresentadas.

### 3. Trabalhos Relacionados

Esta seção apresenta uma análise comparativa deste artigo com a literatura atual. Para isso, vários artigos foram identificados após aplicar a string de busca “Psychophysiological Data AND Software Engineering” na Google Scholar<sup>11</sup>, sendo sete selecionados pela similaridade em relação à temática explorada neste estudo. Tais estudos são brevemente analisados a seguir.

#### 3.1. Análise dos Trabalhos Relacionados

A seguir estão dispostas uma relação de análises sintéticas de cada trabalho relacionado.

- **TR01:** investigou-se uma nova abordagem para a classificação da dificuldade na compreensão de código-fonte, utilizando dados biométricos associados a estados psicofisiológicos. Por meio de um experimento envolvendo 15 desenvolvedores, utilizou-se *Eye-trackers*, sensores de atividade eletrodermal (EDA) e eletroencefalógrafos (EEG) a fim de verificar como os dados coletados poderiam prover uma forma de predição quando um desenvolvedor consideraria uma tarefa como sendo difícil.
- **TR02:** dissertou-se em torno de uma série de experimentos controlados realizados em estudos anteriores, os quais aplicavam sensores diversos para a coleta de dados. Discorreram também sobre as estratégias adotadas para coleta, preparação, transformação, carga e análise dos dados, além de elucidarem os algoritmos de aprendizado de máquina aplicados para a extração de *features* relevantes aos objetivos de cada estudo.
- **TR03:** os autores replicaram experimentos previamente apresentados na literatura [Floyd et al. 2017], mediante o uso de dispositivos sensores de baixo custo. Para tanto, realizaram medições utilizando EEG, EDA e BVP (*Blood Volume Pulse*) em uma amostra de 28 estudantes de Ciência da Computação enquanto estes executavam tarefas envolvendo a compreensão de textos escritos em linguagem natural e código-fonte.
- **TR04:** foi desenvolvido um mapeamento sistemático com o objetivo de identificar lacunas nos trabalhos que abordavam o estudo de carga cognitiva no contexto de Engenharia de Software. Foram analisado 33 artigos relacionados que interseccionavam as temáticas de carga cognitiva, engenharia de software, uso de dispositivos sensores para coleta de dados psicofisiológicos.

---

<sup>11</sup><https://scholar.google.com/>

- **TR05:** realizou-se um levantamento literário entre os anos de 2007 e 2017 no que diz respeito ao estado da arte acerca da utilização de dados biométricos cerebrais, os tipos de sensores utilizados, como esses dados eram coletados, os seus contextos de aplicação e as abordagens estatísticas e de aprendizado de máquina para a extração de *features* em torno da utilização de tais dados.
- **TR06:** propôs-se um modelo conceitual de arquitetura objetificando a aquisição de dados em tempo real provenientes de EEGs de baixo custo. Realizou-se também um teste piloto em que foram utilizados os dispositivos Emotiv EPOC e EPOC+ a fim de avaliar características como usabilidade e confiabilidade durante a aplicação em experimentos. Ademais, os autores dissertaram em torno das implicações éticas envolvidas na coleta de dados psicofisiológicos em tempo real para a avaliação da produtividade de desenvolvedores.
- **TR07:** por meio de um estudo experimental, foi mensurada a carga cognitiva e a familiaridade de um grupo de estudantes do curso de Ciência da Computação no que se refere à interação destes com Ambientes Integrados de Desenvolvimento, utilizando um dispositivo EEG de baixo custo para a coleta de dados psicofisiológicos. Os autores apontaram ainda, as vantagens do EEG em relação à técnicas similares, e as limitações técnicas dos dispositivos sensores de baixo custo.

### 3.2. Análise Comparativa e Oportunidades

Foram definidos cinco Critérios de Comparação (CC) para auxiliar no processo de identificação das similaridades e diferenças entre o trabalho proposto e os artigos selecionados. Esta comparação é crucial para tornar o processo de identificação de oportunidades de pesquisa utilizando critérios objetivos, ao invés de subjetivos. Os critérios são descritos a seguir:

- **Mapeamento Sistemático (CC01):** estudos que desenvolveram um mapeamento sistemático da literatura visando adquirir uma visão do atual estado da arte.
- **Integração de Dados (CC02):** trabalhos que pesquisaram como os dados psicofisiológicos estão sendo suportados pelos ambientes integrados de desenvolvimento.
- **Métricas Psicofisiológicas (CC03):** nesta categoria estão os estudos que pesquisaram em torno de métricas biológicas que estivessem ligadas a processos cognitivos.
- **Aplicações na Engenharia de Software (CC04):** estudos que buscaram avaliar como dados psicofisiológicos são aplicados na Engenharia de Software.
- **Ambientes Integrados de Desenvolvimento (CC05):** pesquisas relacionadas a ambientes integrados de desenvolvimento, tanto em torno do seu uso como na sua melhoria.

A Tabela 2 apresenta a comparação dos trabalhos selecionados, contrastando os mesmos com o trabalho proposto. Algumas lacunas e oportunidades de pesquisas são observadas:

1. apenas o trabalho proposto foi o único a atender completamente todos os critérios de comparação;

**Tabela 2. Análise comparativa dos Trabalhos Relacionados selecionados**

Trabalho Relacionado	Critério de Comparação				
	CC1	CC2	CC3	CC4	CC5
Trabalho proposto	●	●	●	●	●
TR01	○	○	●	◐	○
TR02	○	○	●	●	○
TR03	○	○	●	●	○
TR04	●	○	●	●	○
TR05	○	○	◐	○	○
TR06	○	◐	◐	●	○
TR07	○	◐	◐	●	●

● Atende Completamente   ◐ Atende Parcialmente   ○ Não Atende

2. métricas psicofisiológicas e aplicações em Engenharia de Software foram os temas mais explorados, enquanto mapeamento sistemático e ambientes integrados de desenvolvimento os menos explorados;
3. nenhum estudos explorou detalhadamente o uso de dados psicofisiológicos no contexto de ambiente integrados de desenvolvimento.

Portanto, a seguinte oportunidade de pesquisa foi identificada: uma análise holística da literatura que contemple os cinco critérios de comparação estabelecidos. Esta oportunidade é explorada nas próximas seções.

## 4. Metodologia

Conceitualmente, a metodologia pode ser definida como o estudo da organização e dos meios pelos quais o autor conduzirá a sua pesquisa visando a convergência desta em direção aos objetivos por ele estipulados [Fonseca 2002]. Sendo assim, para o desenvolvimento deste estudo e almejando a sua concordância com os objetivos definidos, foi adotado o mapeamento sistemático como metodologia de pesquisa.

Ainda que o mapeamento sistemático possua uma metodologia bem consolidada para a sua condução, os seus tópicos são extensíveis às necessidades do pesquisador [Petersen et al. 2015]. Sendo assim, as etapas adotados para o desenvolvimento do presente trabalho ficaram estruturadas conforme elucidadas pelas seções subsequentes.

### 4.1. Questões de Pesquisa

Além dos objetivos, as Questões de Pesquisa (ou QPs) guiarão o pesquisador no desenvolvimento do seu mapeamento sistemático, auxiliando na identificação de possíveis estudos que venham a convergir com a temático do estudo em desenvolvimento [Petersen et al. 2015]. As Questões de Pesquisa utilizadas neste estudo são apresentadas na Tabela 3, trazendo consigo os seus propósitos e as variáveis associadas a estas.

### 4.2. Estratégia de Busca

A segunda etapa na metodologia do Mapeamento Sistemático é a definição da estratégia de busca, a qual tem como objetivo alcançar uma amostra representativa de estudos que possam responder de forma generalizada às Questões de Pesquisa definidas anteriormente.

**Tabela 3. Questões de Pesquisa (QPs), suas descrições e variáveis relacionadas**

Questão de Pesquisa	Motivação	Variável
<b>QP01:</b> Como seria uma taxonomia para classificar estudos envolvendo IDEs e dados psicofisiológicos?	Construir uma taxonomia capaz de classificar estudos envolvendo dados psicofisiológicos e IDEs.	Taxonomia
<b>QP02:</b> Que tipos de dispositivos têm sido utilizados na coleta de dados psicofisiológicos?	Listar os dispositivos utilizados nos estudos selecionados para a coleta de dados.	Dispositivos
<b>QP03:</b> Que tipos de dados são extraídos pelos dispositivos utilizados?	Elucidar os tipos de dados extraídos e o que representam no escopo dos estudos analisados.	Atributos
<b>QP04:</b> Como os dados extraídos são suportados pelas IDEs?	Explicar como os dados extraídos podem ser integrados nas IDEs.	Integração
<b>QP05:</b> Com quais propósitos os dados extraídos têm sido utilizados dentro das IDEs?	Trazer o que os dados extraídos visam explicar dentro das IDEs no contexto de Engenharia de Software.	Propósitos
<b>QP06:</b> Quais são as principais contribuições dos estudos?	Explicitar as contribuições de cada estudo no que tange a utilização de dados psicofisiológicos no contexto de Engenharia de Software.	Contribuições
<b>QP07:</b> Qual a natureza dos estudos analisados?	Compreender a natureza do estudo quanto ao seu objetivo.	Natureza
<b>QP08:</b> Onde os estudos selecionados foram publicados?	Identificar os veículos de publicação onde os estudos foram publicados entre os anos de 2009 e 2019.	Veículos de Publicação

A Estratégia de Busca utilizada neste estudo foi concebida em duas etapas: (1) seleção das bases de dados digitais relacionadas à temática do estudo, e (2) construção de uma *string* de busca e suas variantes, respeitando as peculiaridades dos mecanismos de busca presentes em cada base de dados.

#### 4.2.1. Bases de Dados

A Tabela 4 apresenta bases de dados eletrônicas utilizadas. Elas foram escolhidas por serem amplamente utilizadas e por estudos anteriores de mapeamento da literatura terem demonstrado efetividade das mesmas [Bischoff et al. 2019, Gonçales et al. 2019b]. Ademais, a seleção de bases de dados foi embasada na cobertura dos seus mecanismo de busca, os quais retornaram artigos dos principais periódicos de Engenharia de Software.

#### 4.2.2. *String* de Busca

A *string* de busca é um construto oriundo da combinação de termos primários (*major terms*) e seus sinônimos (*minor terms*), sendo utilizada como insumo para os mecanis-

**Tabela 4. Bases de dados e seus endereços digitais**

Base de Dados	Endereço Eletrônico
ACM Digital Library	<a href="http://portal.acm.org/">http://portal.acm.org/</a>
DBLP	<a href="https://dblp.uni-trier.de/">https://dblp.uni-trier.de/</a>
IEEE Xplore	<a href="http://ieeexplore.ieee.org">http://ieeexplore.ieee.org</a>
Science Direct	<a href="https://www.sciencedirect.com/">https://www.sciencedirect.com/</a>
Scopus	<a href="https://www.scopus.com/">https://www.scopus.com/</a>
Semantic Scholar	<a href="https://www.semanticscholar.org/">https://www.semanticscholar.org/</a>
Springer Link	<a href="http://www.springerlink.com/">http://www.springerlink.com/</a>

mos de busca das bases de dados selecionadas. Tal insumo permite que os mecanismos de busca retornem uma coleção de estudos potencialmente relevantes contributivos para o desenvolvimento do Mapeamento Sistemático. A Tabela 5 traz a relação de termos principais e seus sinônimos que foram utilizados no desenvolvimento deste estudo.

Com o objetivo de produzir uma *string* de busca funcional, foram adotados os seguintes passos para o desenvolvimento da mesma: (1) definir os termos principais; (2) pesquisar pelos sinônimos relacionados ou ainda, palavras com equivalência semântica, e (3) combinar os termos principais aos seus sinônimos utilizando operadores lógicos, tais como “AND” e “OR”. Aplicando-se as características citadas, chegou-se na seguinte *string* de busca:

*(eeg OR electroencephalography OR neural OR brain OR cognitive OR psychometric OR bioinformatic) AND (software OR developer OR programmer OR professional OR “software engineering”) AND (ide OR code OR editor OR “integrated development environment”)*

**Tabela 5. Termos primários da *string* de busca e os seus sinônimos**

Termo Principal	Sinônimos
Neuroscience	eeg, fmri, electroencephalography, “functional magnetic resonance imaging”, neural, brain, cognitive, psychometric, bioinformatic
Software Engineering	software, developer, programmer, professional, software engineering
IDE	ide, code, editor, integrated development environment
Data Analysis	“data analysis”, “signal analysis”, analytic, “signal processing”

Devido às peculiaridades sintáticas provenientes de cada mecanismo de busca contido nas bases de dados, a *string* de busca inicial necessitou ser adaptada individualmente de forma a prover um melhor resultado sem perder o seu valor semântico (Apêndice C).

#### 4.2.3. Critérios de Inclusão e Exclusão

Uma vez selecionadas as bases de dados e construída a *string* de busca, a etapa subsequente no processo de desenvolvimento do Mapeamento Sistemático foi composta pela

definição dos Critérios de Inclusão e os Critérios de Exclusão (CI e CE, respectivamente).

Os Critérios de Inclusão foram aplicados com o objetivo de incluir estudos no processo de filtragem após a sua coleta inicial nas bases de dados [Petersen et al. 2015], em contrapartida, os Critérios de Exclusão agiram como um mecanismo de descarte de estudos que não representassem respostas qualificadas para as Questões de Pesquisa [Petersen et al. 2015]. Sendo assim, foram definidos como Critérios de Inclusão os seguintes itens:

- **CI01:** Publicação entre os anos de 2009 e 2019;
- **CI02:** Escrita na língua inglesa;
- **CI03:** Disponível de forma completa nas bases de dados digitais;
- **CI04:** Relação com a *string* de busca e Questões de Pesquisa;

Ademais, para a filtragem dos estudos previamente selecionados por meio dos Critérios de Inclusão, foram aplicados os seguintes Critérios de Exclusão:

- **CE01:** Coerência com a *string* de busca mas fora do contexto da pesquisa;
- **CE02:** Inexistência de resumo;
- **CE03:** Apenas um resumo ou *Call for Paper* de conferências;
- **CE04:** Sem relação com o domínio de Engenharia de Software;
- **CE05:** Cópias/duplicatas ou versões antigas de estudos já selecionados;
- **CE06:** Impossibilidade de acesso ao artigo completo;
- **CE07:** Em desacordo com a motivação da pesquisa;

### 4.3. Estratégia para a Extração de Dados

Estabelecidas as etapas para a seleção dos estudos utilizados, fez-se necessária a definição de dois mecanismos, sendo eles: (1) formulário para a extração de dados e (2) esquemas de classificação. Tais mecanismos são de suma importância para o desenvolvimento de trabalhos no formato de Mapeamento Sistemático, uma vez que, de forma padronizada, definirão como se dará a extração e classificação dos dados contidos nos trabalhos selecionados, possibilitando assim uma identificação confiável de possíveis respostas para as Questões de Pesquisa previamente definidas [Petersen et al. 2015].

#### 4.3.1. Formulário de Extração de Dados

A fim de padronizar o processo de extração dos dados contidos nos artigos e uniformizar a classificação das respostas para as Questões de Pesquisa, foi desenvolvido um formulário de extração de dados, conforme pode ser observado na Figura 4.3.1.

#### 4.3.2. Esquemas de Classificação

Objetivando a confecção de um mecanismo efetivo para a classificação de estudos, entre as possíveis abordagens para a extração de categorias utilizadas no esquema de classificação, neste estudo foram adotadas as seguintes:

- **Esquema Independente de Tópico:** categorias genéricas são utilizadas de forma a prover uma visão geral em torno dos estudos selecionados e a sua temática. Usualmente os autores recomendam a utilização de categorias já adotadas em outros mapeamentos sistemáticos;

Dados Bibliográficos	Questões de Pesquisa (RQs)
<p>Esta seção tem por objetivo levantar os dados bibliográficos referentes aos estudos selecionados.</p> <p><b>Título do Trabalho *</b></p> <p>Using cognitive easiness metric for prog</p> <p><b>Referência *</b></p> <p>Yin, M., Li, B., &amp; Tao, C. (2010, June). Using cognitive easiness metric for program comprehension. In The 2nd International Conference on Software Engineering and Data Mining (pp. 134-139). IEEE.</p> <p><b>Resumo *</b></p> <p>*Program comprehension is one of the most critical coarse-grained slice computed based on simplified system phases in software maintenance. During program designing, the dependence graph [8]. The method is regarded as a basic unit, codes having related function and behavior are often (...)</p> <p><b>Ano de Publicação *</b></p> <p>2010</p> <p>VOLTAR PRÓXIMA</p>	<p>Nesta seção serão elucidadas todas as Questões de Pesquisa com as suas categorias previamente definidas.</p> <p><b>Que tipos de dispositivos têm sido utilizados na coleta de dados psicofisiológicos? (RQ02) *</b></p> <p>Imagem por Ressonância Magnética Funcional (fMRI)</p> <p>Escolher</p> <p>Fixação/Atenção</p> <p>Estados Mentais</p> <p>Atividade Neural</p> <p>Stress/Carga Cognitiva</p> <p>Esforço Mental</p> <p>VOLTAR ENVIAR</p>

**Figura 1. Trecho do formulário utilizado para extração de dados**

- **Esquema Dependente de Tópico:** questões específicas emergem enquanto os estudos são filtrados, trazendo assim um visão detalhada do seu conteúdo, sendo uma peculiaridade presente em cada mapeamento sistemático;

Tendo como base os esquemas de classificação apresentadas, foram extraídas as seguintes categorias como possíveis respostas para cada Questão de Pesquisa:

- **Taxonomia (QP01):** esta categoria tem por objetivo guiar o estudo na construção de uma taxonomia capaz de auxiliar na identificação de estudos que utilizem dados psicofisiológicos integrados às IDEs no âmbito de Engenharia de Software. Ao contrário das demais Questões de Pesquisa, esta foi construída a medida que as respostas das demais emergiam.
- **Dispositivos (QP02):** representa quais tipos de dispositivos foram utilizados nos estudos selecionados para a coleta de dados psicofisiológicos e outras métricas biomédicas. Os itens a seguir representam as categorias extraídas utilizadas para a classificação dos estudos neste aspecto:
  - Eletroencefalograma (EEG): tipo de sensoriamento utilizado para medir e registrar a atividade elétrica do cérebro;
  - *Eye-tracker*: dispositivo utilizado para medir a posição ocular do indivíduo, assim como os seus movimentos;
  - Imagem por Ressonância Magnética Funcional (fMRI<sup>12</sup>): mede a atividade cerebral por associação às alterações no fluxo sanguíneo e variações na saturação de oxigênio;
  - Espectroscopia Funcional de Infravermelho Próximo (fNIRS<sup>13</sup>): assim como a fMRI, a fNIRS coleta medidas a atividade cerebral por meio das respostas hemodinâmicas;

<sup>12</sup>Functional Magnetic Resonance Imaging

<sup>13</sup>Functional Near-Infrared Spectroscopy

- Atividade Eletrodermal (EDA<sup>14</sup>): realiza séries de medição em torno da variação galvânica da pele;
  - Combinação de Dispositivos: aplicável quando mais de um dispositivo foi utilizado no experimento, também aplicado à dispositivos de sensoriamento complementares, como pulseiras de medição e faixas peitorais equipadas com sensores de monitoramento cardíaco e respiratório;
  - Sem Dispositivos: quando a natureza do estudo não demandou que dispositivos sensores fossem utilizados.
- **Atributos (QP03):** refere-se às características baseadas no contexto associadas aos dados coletados por meio dos dispositivos sensores. Os atributos identificados nos estudos selecionados foram classificados em:
    - Fixação/Atenção: visa indicar por quanto tempo o foco visual de um indivíduo permanece em uma Área de Interesse, geralmente associada com a atenção [Beelders e du Plessis 2016], entretanto, sem medidas complementares de Estados Mentais, não é possível inferir de forma clara a relação deste atributo com a atenção focada.
    - Estados Mentais: representa um conjunto de atributos responsáveis pela relação entre processos cognitivos (atenção, aprendizado, meditação, sonolência, etc...) e a atividade elétrica dos neurônios, geralmente expressados por meio de ondas cerebrais, tais como Alfa, Beta, Gama, Delta e Teta [Szu et al. 2013].
    - Atividade Neural: esta categoria abrange os conjuntos de padrões de ativações neurais e as suas semânticas associadas [Lane et al. 1997].
    - *Stress/Carga Cognitiva*: frequentemente associado com a memória de trabalho de um indivíduo enquanto este executa ou tarefa ou mesmo um conjunto de tarefas que demandem sua troca de contexto. [Sweller 1988].
    - Esforço: tem por objetivo explicar a coleção de métricas biológicas ligadas ao aporte de recursos providos pelo corpo durante da execução de tarefas cognitivas [Veltman e Gaillard 1998].
    - Atributos Combinados: aplicável aos estudos que realizaram a análise três ou mais atributos combinados.
    - Sem Atributos: categoria adotada nos estudos em que não foram avaliados atributos específicos em torno dos dados psicofisiológicos.
  - **Integração (QP04):** esta categoria indica se o estudo analisado implementou algum mecanismo capaz de integrar os dados coletados às IDEs utilizadas.
    - Suportado: aplicável quando o estudo analisado utilizou alguma ferramenta capaz de realizar a integração entre os dados psicofisiológicos coletados e a IDE utilizada;
    - Não Suportado: quando um estudo fez uso de dados psicológicos, assim como de IDEs, mas a análise dos dados deu-se de forma independente em relação à IDE;
    - Não Aplicável: categoria adotada em estudos analisados que fazem uso de dados psicofisiológicos mas não possuem em seu contexto o uso de IDEs;

---

<sup>14</sup>Electrodermal Activity

- **Propósitos (QP05):** traz um ponto de vista geral em torno dos principais objetivos estipulados pelos autores em cada estudo, no que tange a utilização dos dados coletados e os seus correlatos atributos. Sendo assim, as categorias identificadas foram enumeradas da seguinte forma:
  - Compreensão de Código-fonte e Estratégias de Depuração: esta categoria de propósitos tenta quantificar a facilidade na compreensão de um programa, do ponto de vista do desenvolvedor, ademais analisa as estratégias adotadas por este durante o processo de depuração;
  - Interfaces Cérebro Computador (BCI<sup>15</sup>): categoria aplicada quando o principal propósito do estudo é o de desenvolver uma interface capaz de possibilitar a interação entre humanos e computadores por meio do uso de sinais oriundos das atividades neurais;
  - Carga Cognitiva: relacionados aos estudos que conduzem medições ligadas ao esforço mental de um desenvolvedor enquanto este executa tarefas complexas;
  - Melhoria na Produtividade: visa encontrar formas de prover melhores ferramentas e melhoras processos no contexto de Engenharia de Software e na qualidade de vida do desenvolvedor;
  - Fora do Contexto: aplicável nos casos em que o propósito do estudo não era ligado diretamente à utilização de dados psicofisiológicos no contexto de Engenharia de Software.
- **Contribuições (QP06):** diferenciando-se das demais, esta categoria adotará duas abordagens: a primeira de caráter subjetivo, apresentará um texto sintético com as principais contribuições identificadas pelos autores para o desenvolvimento deste trabalho. Sob a mesma óptica, a segunda abordagem buscará generalizar as contribuições identificadas por meio das seguintes categorias, de acordo com o valor retornado para o desenvolvimento do presente estudo:
  - Métrica: o estudo apresenta, mesmo que indiretamente, alguma métrica oriunda dos dados psicofisiológicos que pudesse ser correlacionada a algum dos atributos identificados e com algum dos propósitos apresentados anteriormente;
  - Ferramenta: trouxe alguma ferramenta tanto para a coleta de dados, como para processamento/análise destes;
  - Modelo: quando o estudo contribui para a apresentação de um modelo conceitual ou prático de arquitetura ou de experimentação envolvendo o uso dos dados psicofisiológicos;
  - Método: apresenta um método isolado que possa ser aproveitado nos aspectos de coleta, processamento ou análise de dados psicofisiológicos;
  - Processo: similar ao item *método*, entretanto esta categoria engloba uma sequência de métodos para coleta, processamento e ou análise dos dados coletados;
  - Integração: possibilita a identificação de alguma abordagem de integração dos dados psicofisiológicos coletados com os ambientes integrados de desenvolvimento.

---

<sup>15</sup>Brain Computer Interface

- **Natureza (QP07):** a natureza refere-se ao tipo de pesquisa desenvolvida quanto ao seu objetivo principal e desenvolvimento. O estilo de classificação frequentemente adotado na literatura é o de [Wieringa et al. 2006], o qual determinam as seguintes categorias:
  - Proposta de Solução: relativo aos estudos que propõem novas técnicas para solução de problemas existentes, ou ainda, a revisão de técnicas já adotadas;
  - Pesquisa de Avaliação: nesta categoria aplicam-se os estudos que avaliam técnicas e práticas aplicadas na indústria em problemas específicos;
  - Pesquisa de Validação: estudos destinados à avaliação de técnicas inovadoras, mas que não foram ainda validadas quanto à sua aplicação na indústria;
  - Artigo Filosófico: estudos os quais visam trazer uma nova perspectiva no que se refere às estruturas de conhecimentos já consolidados;
  - Artigo de Experiência: este tipo de trabalho discute como outro autor resolveu um problema aplicando uma técnica na prática, sem a necessidade de também utilizar a técnica durante o desenvolvimento do estudo.
  - Artigo de Opinião: engloba todos os estudos que apresentam opiniões embasadas em torno de um determinado tópico de pesquisa.
  - Experimento Controlado: nesta natureza de estudos, os autores realizaram experimentos em ambiente controlado seguindo metodologias bem consolidadas e reproduzíveis.
- **Veículos de Publicação (QP08):** identifica em quais veículos de publicação os estudos selecionados tem sido publicados, como conferências, oficinas, periódicos científicos, etc..., como o objetivo de inferir o nível de maturidade do tópico na comunidade científica.

#### 4.4. Ameaças à Validade e Contramedidas

No que se refere ao desenvolvimento de Mapeamentos Sistemáticos, mesmo que estes possuam uma metodologia bem definida e passível de reprodução, as ameaças à validade do estudo são inerentes ao mesmo [Petersen et al. 2015]. Neste quesito, foram enumeradas três naturezas de ameaças: *construção*, *condução* e *conclusão*.

A primeira natureza de ameaças se refere às definições presentes na metodologia, em especial às *strings* de busca adaptadas para cada base de dados, as quais podem não representar um retorno de mesmo valor para cada uma das bases. A fim de mitigar esta ameaça, foram construídas *strings* de busca individuais utilizando as peculiaridades de cada motor de busca.

A natureza subsequente de ameaças implica na condução do processo de filtragem, seleção e classificação dos estudos, o qual pode ter adotado uma classificação muito genérica ou, em contrapartida, muito específica, comprometendo a formação de grupos de estudos. Neste aspecto buscou-se realizar uma classificação iterativa, ou seja, no decorrer na classificação dos estudos, novas categorias foram adicionadas, em detrimento à exclusão de outras, demandando novos ciclos de classificação de estudos já classificados previamente.

Por último, a natureza de conclusão traz consigo vieses dos autores aplicados aos resultados, podendo gerar inclinações à conclusões conveniente. Para tanto, buscou-se

utilizar o embasamento teórico de estudos similares a fim de corroborar com os resultados, reforçando assim as conclusões obtidas.

## 5. Execução do Processo de Seleção

Esta seção tem por finalidade descrever como foi conduzido o processo de filtragem dos estudos potencialmente relevantes obtidos nas bases de dados eletrônicas apresentadas na Tabela 4 da Seção 4.2.1.

O processo de filtragem utilizado neste trabalho deu-se em oito etapas distintas, apresentadas por meio da Figura 2. Excetuando-se a etapa de Pesquisa Inicial, a qual utilizou apenas os Critérios de Inclusão na triagem, as demais utilizaram agrupamentos dos Critérios de Exclusão como diretivas para a remoção dos estudos inicialmente selecionados.

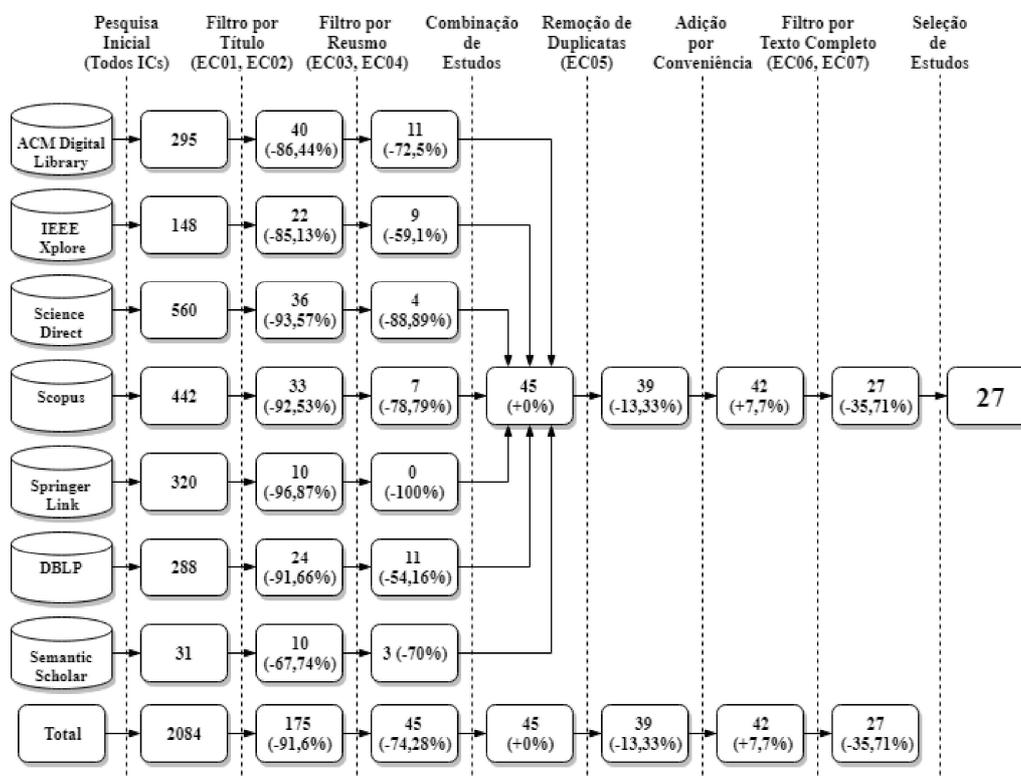


Figura 2. Processo de seleção dos estudos

A seguir estão descritas as etapas adotadas, assim como as suas diretrizes de condução.

- **Pesquisa Inicial:** esta etapa caracterizou-se pela inserção da *string* de busca (Seção 4.2.2) nas bases de dados escolhidas (Seção 4.2.1), combinada aos filtros compostos parametrizados segundo os Critérios de Inclusão (Seção 4.2.3), com isto, a busca retornou um montante de 2084 artigos a serem filtrados, agrupados segundo suas respectivas bases de dados, conforme ilustrado na Figura 2.
- **Filtro por Título:** nesta etapa do processo de filtragem foram analisados os títulos dos artigos resultantes da Pesquisa Inicial. Foram utilizados para filtragem os Critérios de Exclusão CE01 e CE02 (vide Seção 4.2.3), desta forma, o

montante inicial de 2084 artigos foi reduzido para um total de 175, representando um redução de 91,6% do volume de estudos inicialmente selecionados (Figura 2).

- **Filtro por Resumo:** responsável pela precisão na escolha de estudos relevantes para o desenvolvimento do Mapeamento Sistemático da Literatura, esta etapa consistiu na leitura dos resumos de todos os 175 artigos resultantes. Foram aplicados como filtros os Critérios de Exclusão CE03 e CE04 (vide Seção 4.2.3), reduzindo o volume de estudos em 74,28%, chegando assim a um total de 45 artigos (Figura 2).
- **Combinação de Estudos:** conforme ilustrado pela Figura 2, as etapas de filtragem anteriores aplicaram-se de forma individual à cada uma das bases de dados eletrônicas. Foram removidos os rótulos que ligavam os artigos às suas bases de dados de origem, fazendo com que estes compusessem um único *corpus* literário, permitido assim a execução da etapa subsequente.
- **Remoção de Duplicatas:** com os artigos já combinados, aplicou-se o Critério de Exclusão CE05, responsável pela remoção de duplicatas de estudos e/ou versões antigas dos mesmos. Tal filtragem acarretou na redução de 13,39% do total de estudos, resultando assim em um *corpus* literário de 39 artigos (Figura 2).
- **Adição por Conveniência:** visando um aumento no volume de estudos analisados, foram adicionados por conveniência 3 artigos relacionados com a temática e dentro dos Critérios de Inclusão e Exclusão utilizados anteriormente, implicando em um aumento de 7,7% do total, ou ainda, 42 artigos (Figura 2). Tal adição deu-se por conhecimento dos autores e por meio de análises encadeada das referências presentes nos artigos já analisados (*snowballing*), seguindo assim as recomendações de [Petersen et al. 2015] para seleção.
- **Filtro pelo Texto Completo:** nesta fase, os artigos provenientes da etapa anterior foram lidos de forma integral, aplicando-se os Critérios de Exclusão CE06 e CE07 (Seção 4.2.3) acarretando na remoção de 35,71% dos artigos e gerando um montante final de 27 trabalhos (Figura 2).
- **Seleção dos Estudos:** a fase final do processo resultou na seleção de 27 artigos, nomeados neste trabalho como Estudos Primários, os quais estão dispostos no Apêndice A. Ademais, por meio das estratégias descritas na Seção 4.3, foram extraídos os dados de forma a possibilitar a obtenção das respostas para as Questões de Pesquisa.

## 6. Resultados

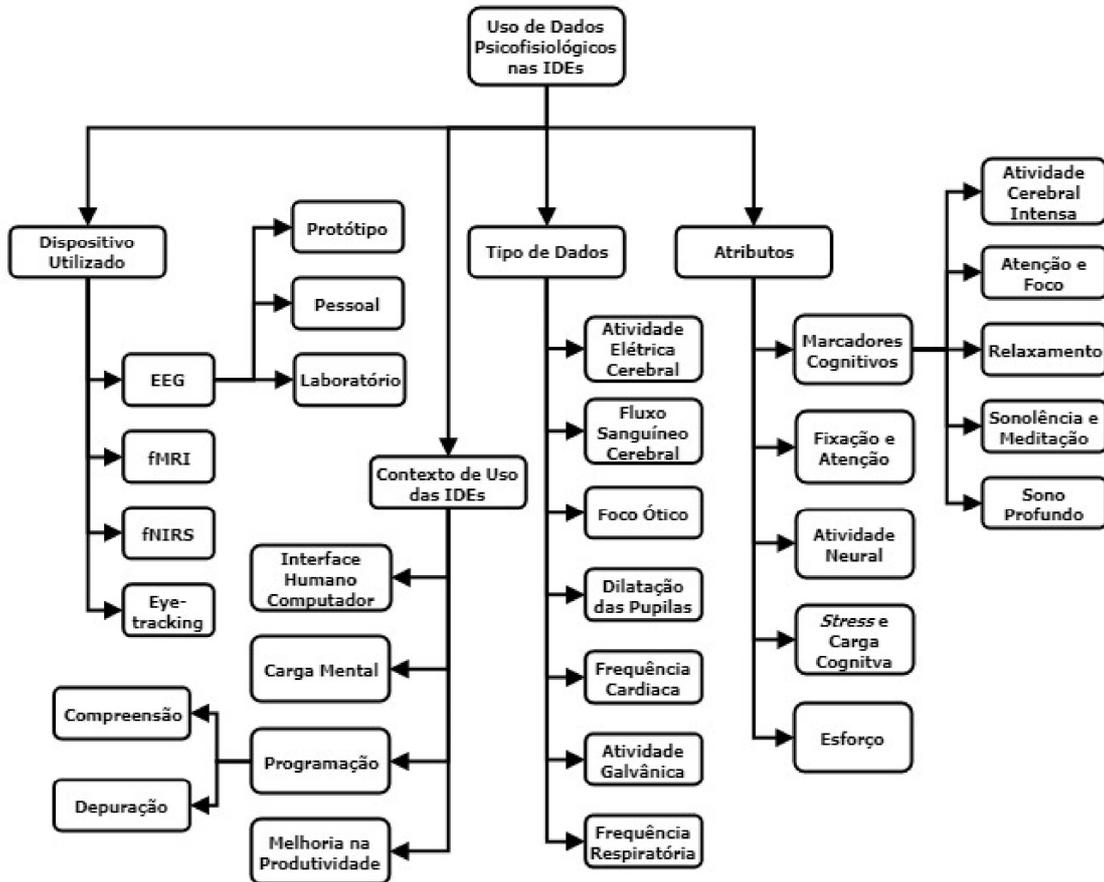
Nesta seção estão dispostos os resultados obtidos por meio da execução processo de filtragem descrito na Seção 5 segundo o delineamento discutido na Seção 4. Para apresentação dos dados quantitativos referentes a cada Questão de Pesquisa, foram utilizadas tabelas contendo a distribuição Estudos Primários, representados pelos seus identificadores enumerados no Tabela 12 da Seção 5, em suas respectivas categorias de respostas.

### 6.1. QP01: Taxonomia

No decorrer no desenvolvimento deste estudo, foi possível confeccionar uma taxonomia utilizada para a classificação de trabalhos que tenham utilizado dados psicofisiológicos e Ambientes Integrados de Desenvolvimento (IDEs), segundo ilustrado na Figura 3.

É possível observar que o segundo nível da taxonomia foi desenvolvido tendo como base os critérios de classificação dos estudos abordados no desenvolvimento deste

trabalho, sendo eles: (1) dispositivos que foram utilizados nos estudos para coletas de dados psicofisiológicos; (2) o contexto no qual os dados coletados foram utilizado no escopo das IDEs; (3) os tipos de dados que foram coletados pelos dispositivos utilizados e (4) o que os dados coletados são capazes de indicar.



**Figura 3. Taxonomia do uso de dados psicofisiológicos nas IDEs**

Ainda que a taxonomia possa trazer uma visão geral no âmbito deste trabalho e facilitar o processo de classificação e identificação de estudos, é válido ressaltar que tal ferramenta foi desenvolvida com o propósito de ser aplicada neste estudo, sendo assim, é necessária a sua validação e/ou adaptação quando utilizada no desenvolvimento de trabalhos similares.

## 6.2. QP02: Dispositivos

A Tabela 6 apresenta os resultados referentes à QP02. É possível observar que na maioria dos estudos envolvendo a coleta de dados psicofisiológicos (29,63%), são utilizados eletroencefalógrafos (EEG) como dispositivo de coleta de sinais. Tal fenômeno pode ser atribuído à redução do custo dos artefatos de eletroencefalografia, além do desenvolvimento de versões acessíveis pelo público não especializado, democratizando o seu uso em diversos contextos [Rostami et al. 2015]. Ademais, estudos sinalizam que a eletroencefalografia, por meio de ondas cerebrais provenientes da ativação neural, é capaz de traduzir

processos cognitivos durante o desenvolvimento de tarefas diversas, inclusive àquelas ligadas ao desenvolvimento de software, compreensão de código e estratégias de depuração [Rostami et al. 2015, Fritz et al. 2014, Fritz e Muller 2016a, Radevski et al. 2015].

**Tabela 6. Distribuição dos estudos por dispositivos utilizados (QP02)**

Dispositivo	Trabalhos	Percentual	Estudos Primários
Sem Dispositivos	3	11,11%	S15, S26, S27
<i>Eye-tracker</i>	5	18,52%	S02, S03, S12, S16, S20
Eletroencefalograma	8	29,63%	S01, S04, S11, S14, S18, S19, S23, S25
Imagem por Ressonância Magnética Funcional	5	18,52%	S05, S08, S17, S21, S22
Espectroscopia Funcional de Infravermelho Próximo	2	7,41%	S13, S24
Combinação de Dispositivos	4	14,81%	S06, S07, S09, S10

Em seguida, *Eye-trackers* e *Imagem por Ressonância Magnética Funcional* representaram, igualmente, 18,52% dos dispositivos utilizados nos Estudos Primários. Nota-se ainda o uso de *Espectroscopia Funcional de Infravermelho Próximo* (7,41%) como dispositivo de coleta.

Representaram 14,81% dos estudos analisados, àqueles que utilizaram mais de um dispositivo, usualmente EEGs combinados com *Eye-trackers* ou ainda, dispositivos complementares, tais como pulseiras de medição e faixas peitorais equipadas com sensores de frequência cardiorrespiratória, sendo estes responsáveis por complementar as medições em cenários onde o *Eye-tracker* e o EEGs não são capazes de indicar, de forma isolada, atributos como *stress* e atenção focada [Beelders e du Plessis 2016].

Dada a natureza de alguns dos Estudos Primários, assim como a metodologia adotada, não foi constatada a utilização de dispositivos para a coleta de dados. Tal gama de estudos representaram 11,11% do volume total.

### 6.3. QP03: Atributos

No que diz respeito aos atributos analisados, a Tabela 7 elucida as suas distribuições nos Estudos Primários. Observa-se a predominância (29,63%) de trabalhos que tiveram como objeto de estudo a atividade neural dos indivíduos. Tais trabalhos fizeram uso de dispositivos como EEGs, fNIRs e fMRI para avaliar quais regiões cerebrais eram ativadas durante a execução de seus experimentos, além de relacionar determinadas semânticas às atividades cerebrais observadas [Lane et al. 1997].

Estudos que dissertaram em torno dos *Estados Mentais* representaram 25,93% do Estudos Primários. Nestes estudos optou-se primordialmente pelo uso de EEGs para a observação dos estados mentais, tendo em vista a íntima ligação destes à ondas cerebrais (Alfa, Beta, Gama, Delta e Theta), além de ser um tópico que se mostrou muito presente na literatura especializada [Szu et al. 2013].

A *Fixação/Atenção* foi o atributo estudado em 18,52% dos estudos presentes no *corpus* literário adotado neste mapeamento sistemático. Usualmente a fixação foi atribuída pelos estudos ao foco ocular em pontos de interesse, predominantemente mensurado por meio de *Eye-trackers*, vide dados apresentados na Tabela 6 entretanto, nos

**Tabela 7. Distribuição dos estudos por atributos analisados (QP03)**

Atributo	Trabalhos	Percentual	Estudos Primários
Sem Atributos	3	11,12%	S15, S26, S27
Fixação/Atenção	5	18,52%	S02, S03, S12, S16, S20
Estados Mentais	7	25,93%	S01, S04, S14, S18, S19, S23, S25
Atividade Neural	8	29,63%	S05, S08, S11, S13, S17, S21, S22, S24
Atributos Combinados	2	7,40%	S09, S10
Fixação/Atenção, Atividade Neural	2	7,40%	S06, S07

casos em que a atenção era objeto de estudos, fez-se o uso de sensores complementares de forma concomitante.

Em 7,40% do trabalhos, realizou-se a análise combinada dos atributos de *Fixação/Atenção* e *Atividade Neural*. Neste estudos observou-se que o enfoque dos autores permeava a compreensão de código-fonte, para tanto, foram utilizados *Eye-trackers* para a coleta de dados relativos ao posicionamento e/ou foco ocular e fNIRS, a fim de identificar a ativação neural de áreas cerebrais ligadas ao aprendizado e leitura [Fakhoury et al. 2018, Fakhoury 2018].

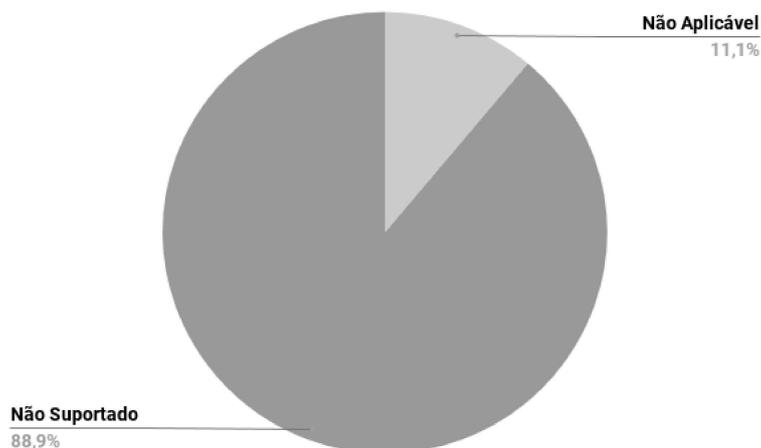
Estudos que analisaram mais de três atributos em seus experimentos correspondem a 7,40% do total. Nestes trabalhos foram utilizadas situações de maior especificidade se comparados aos demais, tais como a mensuração de *Fixação/Atenção*, *Estados Mentais* e atributos complementares (frequência cardiorrespiratória) durante a execução de tarefas cotidianas por parte de um desenvolvedor [Fritz et al. 2014], ou ainda, a combinação de *Estados Mentais* com atributos oriundos de sinais mioelétricos (piscar de olhos, tensão temporomandibular), a fim utilizá-los na comparação do nível de compreensão de código-fonte entre desenvolvedores novatos e experientes [Lee et al. 2016].

Ainda que 88,88% dos estudos tenham aplicado análises de atributos relacionados aos dados psicofisiológicos coletados, em 11,12% destes não houve a presença de análises em torno de tais atributos. Tal característica se deu pela natureza destes trabalhos, além do processo metodológico adotado em seu desenvolvimento.

#### 6.4. QP04: Integração

Esta Questão de Pesquisa visava obter como resposta quais os tipos de suporte que as IDEs utilizadas nos experimentos proveram para a integração dos dados psicofisiológicos coletados, possibilitando assim a sua análise no contexto de Engenharia de Software pelos pesquisadores. A Figura 4 ilustra como se deu a divisão das respostas para esta Questão de Pesquisa.

Observa-se que, do total de 27 estudos analisados, 11,1% destes fizeram algum tipo de uso de dados psicofisiológicos entretanto, sem a utilização de IDEs em seus experimentos, conforme discriminado anteriormente na Seção 4.3.2 e ilustrado por meio da Figura 4. Por outro lado, cerca de 88,9% dos estudos foram enquadrados na categoria *Não Suportado*, ou seja, realizaram experimentos envolvendo IDEs e também a análise de dados psicofisiológicos. Ainda sim, nenhum destes estudos especificou mecanismos



**Figura 4. Suporte à integração de dados psicofisiológicos nas IDEs (QP04)**

capazes de integrar os dados coletados e os Ambiente Integrados de Desenvolvimento (IDEs), sendo assim, as análises foram realizadas de forma externa e por meio de ferramentas especializadas.

### 6.5. QP05: Propósitos

Neste item estão dispostos os propósitos adotados pelos autores nos Estudos Primários, relativos ao uso de dados psicofisiológicos no contexto de Engenharia de Software. Mediante os dados elucidados pela Tabela 8, constatou-se que, em sua maioria (66,70%), os Estudos Primários tinham como enfoque analisar a *Compreensão de Código-fonte e as Estratégias de Depuração* utilizadas pelos desenvolvedores nos experimentos realizados.

**Tabela 8. Distribuição dos estudos por propósito adotado (QP05)**

Propósito	Trabalhos	Percentual	Estudos Primários
Compreensão de Código-fonte e Estratégias de Depuração	18	66,70%	S02, S03, S05, S06, S07, S08, S09, S12, S13, S14, S16, S17, S20, S21, S22, S25, S26, S27
Interfaces Cérebro Computador	2	7,41%	S01, S23
Carga Cognitiva	4	14,81%	S04, S11, S19, S24
Melhoria na Produtividade	2	7,41%	S10, S18
Fora do Contexto	1	3,70%	S15

Observou-se ainda que, em 14,81% dos Estudos Primários, o propósito focal do trabalho foi a avaliação da do desenvolvedor enquanto este executava um série de atividades relacionadas ao seu contexto. É válido ressaltar que, apesar da possibilidade de avaliar a Compreensão de Código-fonte e a estratégia de depuração de forma individual, as medições relacionadas a estes propósitos possuem uma maior validade quando atreladas à *Carga Cognitiva*, segundo levantamentos desenvolvidos por outros autores [Gonçales et al. 2019a].

Constituindo 7,41% dos Estudos Primários estavam àqueles cujo propósito principal era a busca pela *Melhoria na Produtividade* dos desenvolvedores mediante o uso de

dados psicofisiológicos e IDEs.

Trabalhos envolvendo o desenvolvimento de *Interfaces Cérebro Computador* corresponderam a também 7,41% dos Estudos Primários. Esta categoria engloba a construção de um protótipo de dispositivo para o controle de periféricos domésticos por meio de ondas cerebrais [Szu et al. 2013] e a idealização de um EEG portátil capaz de comunicar-se com *smartphones* [Alshbatat et al. 2014].

## 6.6. QP06: Contribuições

Esta seção tem por objetivo apresentar as principais contribuições identificadas em cada um dos Estudos Primários. Tais contribuições foram extraídas durante a leitura do texto completo (Seção 5) e visavam responder à esta Questão de Pesquisa. Os textos foram sintetizadas e dispostos por meio da listagem a seguir, a fim de possibilitar uma leitura estruturada, mesmo que a origem dos dados tenha sido de natureza qualitativa e com aspectos subjetivos.

- **S01 e S18** Possibilitou a aquisição e o processamento de sinais de atividades neurais em tempo real, além da extração de *features* a fim de utilizar tais sinais como BCI (*Brain-Computer Interface*).
- **S02:** Trouxe uma visão de como a exibição das informações referentes ao código são interpretadas por programadores novatos e experientes.
- **S03:** O estudo indicou que, apesar de o tempo de fixação ser maior nos casos em que o código é apresentado em preto-e-branco, se comparado com o mesmo trecho com realce de sintaxe, essa diferença não é significativa no âmbito de compreensão do código.
- **S04:** Neste estudo vale salientar duas principais contribuições: (1) a metodologia para utilização de EEG na avaliação de carga cognitiva e (2) a evidência de que programadores mais experientes resolvem as tarefas de forma mais ágil, gerando uma menor carga cognitiva no mesmo espaço de tempo que os novatos.
- **S05:** Como principal contribuição deste estudo, os autores trouxeram um mapeamento da ativação cerebral durante o processo de depuração.
- **S06:** Os autores propuseram metodologias para relacionar a carga cognitiva do desenvolvedor com a baixa qualidade do código-fonte.
- **S07:** Este estudo foi importante por trazer uma revisão de como outros autores estão utilizando o *Eye-tracker* e o fNIRs como ferramentas para a mensuração da carga cognitiva/mental de desenvolvedores no quesito de compreensão de código-fonte.
- **S08:** Entre as contribuições deste estudo, duas são de grande valia para o desenvolvimento deste mapeamento: (1) elucidar que a linguagem de programação e a linguagem natural são interpretadas em regiões diferentes do cérebro e (2) com o aumento da *expertise* do programador, a linguagem de programação começa a ser interpretada como uma linguagem natural pelo seu cérebro.
- **S09:** O trabalho trouxe como contribuição a possibilidade de se utilizar sinais neurofisiológicos para a avaliação da dificuldade em atividades de programação, além de trazer a possibilidade do desenvolvimento de ferramentas para a melhoria no processo de desenvolvimento de software.
- **S10:** Os autores trouxeram uma revisão de grande valia em torno de dados psicofisiológicos e seus dispositivos sensores. Ainda apresentaram as técnicas utilizadas

para a análise dos dados utilizados e os algoritmos para predição de casos semelhantes.

- **S11:** Como resultado os autores apresentaram que, com o aumento da carga cognitiva, a produtividade degrada-se no decorrer do tempo.
- **S12:** Ainda que tenha sido realizado um experimento no âmbito do uso de linguagens de programação, o estudo apresentou como resultado os movimentos oculares dos participantes, sem traçar paralelos com a literatura a fim de encontrar padrões ou ainda explicações para os padrões identificados.
- **S13 e S14:** Os resultados são úteis no entendimento do que acontece no cérebro do programador, quando observados determinados aspectos (lógica, fluxo de código) durante o processo de compreensão de linguagens de programação. Por meio do entendimento dos padrões de ativação cerebral, é possível criar métricas de avaliação das habilidades de programadores novatos e experientes. Ainda é possível determinar fatores que tornam o ensino de programação mais eficaz em termos de aumento da habilidade cognitivo.
- **S15:** Foi desenvolvida uma IDE sensível ao contexto de trabalho.
- **S16 e S20:** Os autores expuseram o estado-da-arte em torno da utilização de *eye-tracking* na Engenharia de Software, além dos avanços da tecnologia no decorrer dos anos.
- **S17:** Desenvolveu diretrizes para estabelecer a fMRI como um instrumento padrão para observar a compreensão de código-fonte e outros processos relacionados.
- **S19:** O trabalho trouxe como contribuições a possibilidade de uso de EEGs de baixo custo como instrumentos para a medição da facilidade de uso de IDEs por parte dos desenvolvedores, permitindo o aprimoramento da ferramenta, possibilitando ainda uma melhora na sua usabilidade.
- **S21:** Como contribuição do estudos, os autores evidenciaram que existe um padrão de ativação associado à memória de trabalho e ao processamento de linguagem natural. Ainda como contribuição, o trabalho trouxe uma metodologia para utilização de fMRI no contexto de compreensão de linguagens de programação.
- **S22:** Os autores identificaram que o uso de sugestões semânticas reduzem a ativação cerebral durante o processo de compreensão do código-fonte, além de reproduzir os estudos de outros autores e validar os seus resultados.
- **S23:** O trabalho em questão trouxe a viabilidade do desenvolvimento de um EEG doméstico, tanto para o uso como equipamento de monitoramento, como para BCI, dada a possibilidade de análise de dados em tempo real. Além disso elucidou os estados mentais vinculado aos espectros das ondas cerebrais.
- **S24:** O presente estudo contribuiu com a evidência de que, durante o processo de ensino de programação orientada à objetos, IDEs visuais como o BlueJ, reduzem a carga cognitiva, possibilitando inferir um processo de aprendizado mais facilitado. Ainda os autores fomentam estudos em torno do uso de IDEs e fNIRS.
- **S25:** Ainda que o trabalho tenha avaliado o comportamento das ondas Theta e Alpha, os autores não trouxeram uma perspectiva psicofisiológica dos processos cognitivos envolvidos durante as atividades.
- **S26:** Propôs uma métrica para avaliar complexidade cognitiva baseada em grafo de chamadas de métodos.

- **S27:** Os autores mostraram que as metodologias de depuração utilizada pelos desenvolvedores variam conforme as funcionalidades das IDEs, opondo-se às formas de depuração tradicionais.

Em adição aos resultados apresentados na forma de textos sintéticos, foram dispostas as métricas de acordo com as categorias as quais as respectivas contribuições foram enquadradas, conforme apresentado na Tabela 9.

**Tabela 9. Distribuição dos estudos por tipo de Contribuições (QP06)**

Contribuição	Trabalhos	Percentual	Estudos Primários
Métrica	10	37,04%	S03, S04, S05, S10, S12, S13, S14, S21, S22, S26
Ferramenta	2	7,41%	S09, S15
Modelo	6	22,22%	S01, S16, S18, S19, S20, S23
Método	6	22,22%	S02, S06, S07, S08, S24, S25
Processo	3	11,11%	S11, S17, S27

Denota-se um maior número (37,04%) de estudos que contribuíram de forma significativa fornecendo algum tipo de *Métrica* acerca da utilização dos dados psicofisiológicos aplicados aos atributos elencados na QP03 e propósitos da QP05. Por conseguinte, observa-se que *Método* e *Modelo* apresentaram as mesmas frequências (22,22% cada), trazendo ambas contribuições que pudessem ser aplicados para a definição de um modelo de arquitetura ou método experimental. A categoria *Processo* trouxe contribuições (11,11%) para a definição de processos que pudessem ser adotados no que tange coleta, processamento e análise de dados psicofisiológicos. Por último, apenas 2 estudos (7,41%) apresentaram alguma ferramenta que possuísse aplicabilidade no contexto deste trabalho.

É válido ressaltar que nenhum dos estudos analisados foi enquadrado na categoria *Integração*, já que estes não apresentaram quaisquer mecanismos capazes de suportarem a integração dos dados coletados com os ambientes integrados de desenvolvimento ou outras ferramentas utilizadas.

### 6.7. QP07: Natureza

Nesta Questão de Pesquisa buscou-se trazer um panorama acerca da distribuição das naturezas de cada estudo selecionado, utilizando-se os critérios definidos na Seção 4.3.2. Os resultados obtidos pelo processo de condução do mapeamento sistemático estão organizados e dispostos na Tabela 10.

**Tabela 10. Distribuição dos estudos por Natureza (QP07)**

Natureza	Trabalhos	Percentual	Estudos Primários
Experimento Controlado	15	55,56%	S02, S03, S04, S05, S06, S08, S09, S10, S11, S12, S13, S14, S21, S22, S25
Proposta de Solução	7	25,93%	S07, S15, S18, S23, S24, S26
Pesquisa Avaliação	1	3,70%	S19
Artigo de Opinião	4	14,81%	S16, S17, S20, S27

Observa-se que predominantemente os estudos selecionados (55,56%) são de natureza *Experimento Controlado*, ou seja, estudos em que os autores realizaram algum tipo de experimento de forma a validar uma técnica já existente ou ainda, corroborar algum conceito encontrado na literatura. Em seguida estão os estudos de natureza *Proposta de Solução*, representando 25,93% do montante total. Nesta natureza de estudos, estão àqueles em que o principal objetivo foi o de apresentar uma solução para um problema já existente ou uma versão melhorada de uma solução já apresentada, por meio de uma prova de conceito. Os *Artigos de Opinião* compuseram 14,81%, trazendo trabalhos que expressam a opinião dos autores acerca de determinado tópico, instigando a discussão e abrindo oportunidade para novos trabalhos. Por fim, 3,70% dos trabalhos foram de natureza *Pesquisa Avaliação*, onde buscou-se avaliar uma ferramenta, metodologia ou ideia apresentada por outros autores.

### 6.8. QP08: Veículos de Publicação

Esta Questão de Pesquisa objetivava quantificar os veículos nos quais os Estudos Primários foram publicados. Os dados referente às respostas adquiridas mediante o desenvolvimento deste estudo estão dispostos na Tabela 11.

**Tabela 11. Distribuição dos estudos por veículos de publicação (QP08)**

Veículo	Trabalhos	Percentual	Estudos Primários
Conferência	16	59,26%	S03, S04, S06, S07, S08, S09, S10, S11, S12, S13, S14, S15, S17, S21, S25, S26
Periódico	6	22,22%	S01, S02, S16, S20, S23, S24
Oficina	2	7,41%	S18, S19
Simpósio	3	11,11%	S05, S22, S27

Constata-se que em sua maioria (59,26%), os Estudos Primários foram publicados em conferências da área, indicando uma maior incipiência no que tange a temática dos trabalhos. Por conseguinte, estão os trabalhos que foram veiculados em periódicos (22,22%), caracterizando assuntos consolidados na comunidade científica. Por último são apresentados os trabalhos que foram publicados em eventos menores, tais como simpósios (11,11%) e oficinas (7,41%).

#### 6.8.1. Frequência de Publicação

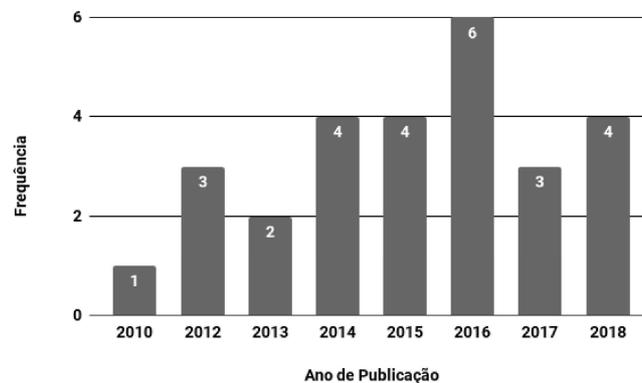
Em adição aos resultados apresentados acerca dos *Veículos de Publicação*, realizou-se também o levantamento da distribuição de publicação dos estudos no decorrer dos anos, conforme quantificado por meio da Figura 5.

É possível constatar que, no decorrer da última década – período avaliado no desenvolvimento deste mapeamento sistemático – houve um aumento no número de publicações referente ao uso de dados psicofisiológicos aplicados no contexto da Engenharia de Software.

## 7. Discussão e Desafios

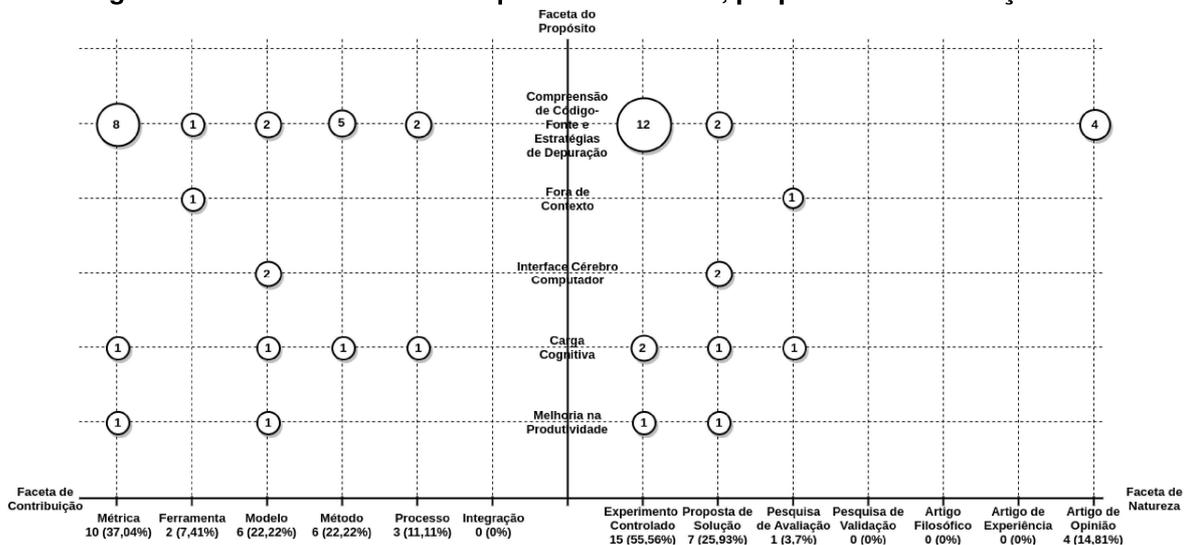
Por meio da Figura 6 é apresentado um gráfico de bolhas onde os estudos primários são organizados em três facetas: contribuição (eixo horizontal esquerdo), propósitos (eixo

**Figura 5. Distribuição dos estudos pelos anos de publicação**



vertical) e natureza (eixo horizontal direito). Tal gráfico traz uma visão combinada dos estudos primários objetivando identificar as relações entre as contribuições de cada estudo, as suas naturezas e os propósitos adotados. Já o mapa de calor ilustrado na Figura 7, elucida a densidade de estudos para cada par de respostas às questões de pesquisa apresentadas neste trabalho. Sendo assim, fazendo uso de tais recursos, foi possível identificar os seguintes aspectos de discussão:

**Figura 6. Densidade de estudos pela sua natureza, propósito e contribuições.**



*Carência de Suporte à Integração de Dados Psicofisiológicos:* os estudos que realizaram experimentos utilizando ambientes integrados de desenvolvimento e dados psicofisiológico mostraram que não existem mecanismos nativos de suporte a tais tipos de dados por parte das IDEs, tampouco por meio de extensões que adicionem tal funcionalidade. Neste aspecto, os dados eram coletados por ferramentas proprietárias que comunicavam-se com os dispositivos sensores, para posteriormente serem pré-processados e analisados. Tal carência de suporte à integração mostra-se oportuna para a utilização de ferramentas de *Big Data* que trabalhem com grandes volumes de dados e/ou processamento de fluxos de dados em tempo real. Ademais, a produção de uma tecnologia infraestrutura para a integração de dados psicofisiológicos aos ambientes integrados

de desenvolvimento por meio de funcionalidades extensíveis, mostra-se como uma necessidade presente em estudos experimentais aplicados à Engenharia de Software, tanto na indústria como sob o viés acadêmico.

**Figura 7. Mapa de calor apresentando a densidade de estudos para cada par de respostas às Questões de Pesquisa**

EEG	0	0	0	0	0	0	7	0	1	0	0	0	0	8	0	2	2	3	1	0	3	4	0	1	2	4	0	2	
Eye	0	0	0	0	0	0	0	5	0	0	0	0	0	5	0	0	5	0	0	0	0	0	3	2	0	3	2	0	0
fMRI	0	0	0	0	0	0	0	0	5	0	0	0	0	5	0	0	5	0	0	0	0	0	4	1	0	0	3	2	0
DevComb	0	0	0	0	0	0	0	0	0	2	2	0	0	4	0	0	3	0	1	0	0	1	3	0	0	0	4	0	0
fNIRS	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	1	1	0	0	0	1	1	0	0	1	1	0	0
NoDev	0	0	0	0	0	0	0	0	0	0	0	3	0	3	0	2	0	0	1	0	1	2	0	1	0	0	2	1	0
MentSta	7	0	0	0	0	0	0	0	0	0	0	0	0	7	0	2	2	2	1	0	3	3	0	1	2	3	0	2	
Fix/Att	0	5	0	0	0	0	0	0	0	0	0	0	0	5	0	0	5	0	0	0	0	3	2	0	3	2	0	0	
Act	1	0	5	0	2	0	0	0	0	0	0	0	0	8	0	0	6	2	0	0	1	6	1	0	1	5	2	0	
Fix/Att,Act	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	1	1	0	0	0	2	0	0
AttbComb	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	0	0	2	0	0	0	0	2	0	0
NoAttb	0	0	0	0	0	3	0	0	0	0	0	0	0	0	3	0	2	0	0	1	2	0	1	0	0	2	1	0	
NoSupp	8	5	5	4	2	0	7	5	8	2	2	0	0	0	0	2	16	4	2	0	5	15	3	1	6	14	2	2	
NoApp	0	0	0	0	0	3	0	0	0	0	0	3	0	0	0	2	0	0	1	0	2	0	1	0	0	2	1	0	
BCI	2	0	0	0	0	0	2	0	0	0	0	0	0	2	0	0	0	0	0	0	2	0	0	0	2	0	0	0	0
Compr,Dbg	2	5	5	3	1	2	2	5	6	2	1	2	2	16	2	0	0	0	0	0	2	12	4	0	3	12	3	0	
CogLoad	3	0	0	0	1	0	2	0	2	0	0	0	0	4	0	0	0	0	0	0	1	2	0	1	1	2	0	1	
ImprPrd	1	0	0	1	0	0	1	0	0	0	1	0	0	2	0	0	0	0	0	0	1	1	0	0	0	1	0	1	
OutCtx	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	
SolProp	3	0	0	1	1	2	3	0	1	1	0	2	5	2	2	2	1	1	1	1	0	0	0	0	3	3	0	1	
CtrlExp	4	3	4	3	1	0	3	3	6	1	2	0	15	0	0	12	2	1	0	0	0	0	0	0	1	12	2	0	
OpiPpr	0	2	1	0	0	1	0	2	1	0	0	1	3	1	0	4	0	0	0	0	0	0	0	0	2	1	1	0	
EvalRsc	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	
Jnl	2	3	0	0	1	0	2	3	1	0	0	0	6	0	2	3	1	0	0	0	3	1	2	0	0	0	0	0	0
Conf	4	2	3	4	1	2	3	2	5	2	2	2	14	2	0	12	2	1	1	3	3	12	1	0	0	0	0	0	0
Symp	0	0	2	0	0	1	0	0	2	0	0	1	2	1	0	3	0	0	0	0	2	1	0	0	0	0	0	0	0
Work	2	0	0	0	0	0	2	0	0	0	0	0	2	0	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0
EEG																													
Eye																													
fMRI																													
DevComb																													
fNIRS																													
NoDev																													
MentSta																													
Fix/Att																													
Act																													
Fix/Att,Act																													
AttbComb																													
NoAttb																													
NoSupp																													
NoApp																													
BCI																													
Compr,Dbg																													
CogLoad																													
ImprPrd																													
OutCtx																													
SolProp																													
CtrlExp																													
OpiPpr																													
EvalRsc																													
Jnl																													
Conf																													
Symp																													
Work																													

**EEG:** eletroencefalograma – **Eye:** eye-tracker – **fMRI:** Imagem por Ressonância Magnética Funcional – **DevComb:** Combinação de Dispositivos – **fNIRS:** Espectroscopia Funcional de Infravermelho Próximo – **NoDev:** Sem Dispositivos – **MentSta:** Estados Mentais – **Fix/Att:** Fixação/Atenção – **Act:** Atividade Neural – **Fix/Att,Act:** Fixação/Atenção, Atividade Neural – **AttComb:** Combinação de Atributos – **NoAttb:** Sem Atributos – **NoSupp:** Não Suportado – **NoApp:** Não Aplicável – **BCI:** Interfaces Cérebro Computador – **Compr,Dbg:** Compreensão de Código-fonte e Estratégias de Depuração – **CogLoad:** Carga Cognitiva – **ImprPrd:** Melhoria na Produtividade – **OutCtx:** Fora do Contexto – **SolProp:** Proposta de Solução – **CtrlExp:** Experimento Controlado – **OpiPpr:** Artigo de Opinião – **EvalRsc:** Pesquisa de Avaliação – **Jnl:** Periódico – **Conf:** Conferência – **Symp:** Simpósio – **Work:** Oficina

*Experimentos de Compreensão de Código-fonte e Estratégias de Depuração Aplicadas à Engenharia de Software:* os autores dos estudos revisados neste trabalho adotaram como propósito principal, em sua maioria, o uso dos dados psicofisiológicos para avaliação da compreensão de código-fonte por parte dos desenvolvedores, além das suas estratégias adotadas para o processo de depuração. Esta preferência elucida-se pelo

domínio dos estudos analisados, os quais usualmente buscavam aplicabilidade no contexto de Engenharia de Software. Analisando-se a Figura 7, fica evidente que este tipo de propósito foi adotado nos estudos que fizeram uso de experimentos controlados como metodologia principal.

*Análises de Dados por Meio de Ferramentas Externas:* apesar de os estudos terem realizado análises de dados para obtenção dos resultados e desenvolvimento das suas discussões e conclusões, observou-se a necessidade de análises externas ao ambiente controlado do experimento, mediante o uso de ferramentas como MATLAB, ou ainda, soluções proprietárias desenvolvidas pelos fornecedores dos dispositivos sensores utilizados nos experimentos. Esta característica ficou evidente nos experimentos em que foram utilizados ambientes integrados de desenvolvimento (vide item *NoSupp* da Figura 7). Nestes experimentos, além dos dados coletados pelos dispositivos, demandou-se que as métricas obtidas nas IDEs fossem correlacionadas aos dados psicofisiológicos de forma independente. No aspecto apresentado, encontra-se uma oportunidade para o desenvolvimento de extensões às IDEs que possibilitem a utilização de técnicas de *Data Analytics* de forma nativa, permitindo assim a extração de *insights* em torno dos dados psicofisiológicos de forma integrada ao contexto de uso das IDEs.

*Crescimento do Interesse por Dados Psicofisiológicos Aplicados à Engenharia de Software:* é possível observar uma maior presença de estudos publicados em conferências vinculadas à computação aplicada. Tal característica sugere a incipiência deste tópico de pesquisa, dado o interesse pelo desenvolvimento de trabalhos do gênero e o aumento de publicações nos últimos anos (Figura 5 da Seção 6). Não obstante, as integrações multidisciplinares entre Neurociências e Engenharia de Software abrem margem para a produção de estudos diversificados e inovadores no que tange a aplicabilidade da psicofisiologia aliada ao processo de desenvolvimento de software. Este fenômeno torna-se observável por meio da pluralidade de contribuições (Item 6.6) oriundas dos artigos analisados, assim como pelos trabalhos relacionados utilizados (Item 3), os quais serviram de embasamento teórico no decorrer da produção deste mapeamento sistemático.

## 8. Considerações Finais e Trabalhos Futuros

O desenvolvimento deste trabalho proporcionou aos autores o vislumbre teórico do atual estado da arte acerca da utilização de dados psicofisiológicos aplicados à Engenharia de Software, desde a sua obtenção, processamento e extração de *insights* sobre os processos físicos e psicológicos envolvidos na produção de artefatos de software por parte do desenvolvedor. Foi possível observar que, em sua maioria, os autores optaram pela utilização de eletroencefalografia (29,63%), *eye-tracking* (18,52%) e imagem por ressonância magnética funcional (18,52%) em seus experimentos, sendo este fato corroborado pela maturidade do uso de equipamentos presente na literatura, assim como a redução dos custos e facilidade na configuração, operação e extração dos dados oriundos das medições.

Constatou-se um expressivo número de estudos que analisavam os princípios da teoria da carga cognitiva que, assim como o uso de eletroencefalografia e *eye-tracking*, mostrou-se bem consolidado na literatura especializada, reforçando o embasamento teórico dos autores, em especial nos experimentos que tinham como propósitos entender como o desenvolvedor compreende o código-fonte e quais as estratégias que este utiliza

durante o processo de depuração (66,7%), assim como naqueles que buscavam avaliar os níveis de carga cognitiva apresentado pelos desenvolvedores durante a execução das suas tarefas (14,81%).

Um ponto que pôde ser observado durante a leitura dos artigos na etapa de filtragem e seleção, foi a falta de consenso de autores em torno de possíveis processos cognitivos relacionados às atividades do desenvolvedor, excetuando-se a carga cognitiva, os demais conceitos, quando aplicados no contexto da Engenharia de Software, apresentaram teor ambíguo em relação, como é o exemplo da atividade neural, em que alguns estudos conceituaram a mesma como sendo a intensidade de ativação elétrica e/ou hemodinâmica no cérebro, enquanto outros estudos apresentaram este conceito relacionando apenas com áreas pontuais do cérebro e atribuindo uma semântica a tal.

No aspecto da análise de dados, os autores fizeram uso majoritariamente de ferramentas externas para o processamento, extração das *features* e obtenção de *insights* acerca dos conjuntos de dados, especialmente nos estudos de natureza experimental em que os autores aplicaram a manipulação de ambientes integrados de desenvolvimento como parte das fontes de dados. Nestes estudos não foram apresentadas quaisquer formas de integração dos dados oriundos das IDEs com as ferramentas de análises, demandando a utilização de suítes de software genéricas de terceiros para a manipulação, processamento e análise, ou ainda, de ferramentas fornecidas pelos desenvolvedores dos dispositivos sensores utilizados (88,9%). Tal fator mostrou-se como uma oportunidade para a aplicação de conhecimentos multidisciplinares, tais como a utilização de ferramentas de *Big Data* para a manipulação do volume de dados e ainda, processamento em tempo real ou *near real-time*, extração de *insights* por meio de algoritmos de aprendizagem de máquina (em inglês, Machine Learning), confecção de dispositivos vestíveis e de Internet das Coisas a fim de expandir a coleta de dados para fora de um ambiente específico, possibilitando compreender as características psicofisiológicas apresentadas pelo desenvolvedor em seu real ambiente de trabalho, entre outras aplicações.

Por meio das oportunidades observadas no decorrer da produção deste mapeamento sistemático, pretende-se desenvolver, como trabalhos futuros, mecanismos que possibilitem a coleta e processamentos de dados psicofisiológicos oriundos de dispositivos sensores, assim como a sua integração a ambientes integrados de desenvolvimento, de forma a possibilitar a geração de metadados que possam ser relacionados aos eventos gerados dentro das IDEs. Em posse de tais dados, ainda será possível aplicar técnicas de aprendizagem de máquina para a obtenção de padrões que possam sugerir potenciais falhas na produção de software de acordo com o padrão psicofisiológico de cada desenvolvedor, possibilitando a criação de IDEs inteligentes e adaptáveis ao usuário, além de recursos táticos para a priorização de *refactoring* em unidades de software.

## Referências

- Alshbatat, A. I. N., Vial, P. J., Premaratne, P., e Tran, L. C. (2014). EEG-based Brain-computer Interface for Automating Home Appliances. *Journal of Computers*, 9(9):2159–2166.
- Andreassi, J. L. (2007). *Psychophysiology: Human behavior & physiological response*, 5th ed. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US.

- Astromskis, S., Bavota, G., Janes, A., Russo, B., e Di Penta, M. (2017). Patterns of developers behaviour: A 1000-hour industrial study. *Journal of Systems and Software*, 132:85–97.
- Bednarik, R. (2012). Expertise-dependent visual attention strategies develop over time during debugging with multiple code representations. *International Journal of Human Computer Studies*, 70(2):143–155.
- Beelders, T. e du Plessis, J.-P. (2016). The Influence of Syntax Highlighting on Scanning and Reading Behaviour for Source Code. pages 1–10. Association for Computing Machinery (ACM).
- Bischoff, V., Farias, K., Gonçalves, L. J., e Victória Barbosa, J. L. (2019). Integration of feature models: A systematic mapping study. *Information and Software Technology*, 105(October 2016):209–225.
- Crk, I. e Kluthe, T. (2014). Toward using alpha and theta brain waves to quantify programmer expertise. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, pages 5373–5376.
- Duraes, J., Madeira, H., Castelhana, J., Duarte, C., e Branco, M. C. (2016). WAP: Understanding the Brain at Software Debugging. *Proceedings - International Symposium on Software Reliability Engineering, ISSRE*, pages 87–92.
- Fakhoury, S. (2018). Moving towards objective measures of program comprehension. *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pages 936–939.
- Fakhoury, S., Ma, Y., Arnaoudova, V., e Adesope, O. (2018). The effect of poor source code lexicon and readability on developers' cognitive load. *Proceedings of the 26th Conference on Program Comprehension*, pages 286–296.
- Fishburn, F. A., Norr, M. E., Medvedev, A. V., e Vaidya, C. J. (2014). Sensitivity of fNIRS to cognitive state and load. *Frontiers in Human Neuroscience*, 8(1 FEB):1–11.
- Floyd, B., Santander, T., e Weimer, W. (2017). Decoding the Representation of Code in the Brain: An fMRI Study of Code Review and Expertise. *Proceedings - 2017 IEEE/ACM 39th International Conference on Software Engineering, ICSE 2017*, pages 175–186.
- Fonseca, J. (2002). *Metodologia da pesquisa científica.*, volume 2. INTA.
- Fritz, T., Begel, A., Müller, S. C., Yigit-Elliott, S., e Züger, M. (2014). Using psychophysiological measures to assess task difficulty in software development. *Proceedings of the 36th International Conference on Software Engineering*, pages 402–413.
- Fritz, T. e Muller, S. C. (2016a). Leveraging Biometric Data to Boost Software Developer Productivity. *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering Leveraging*, pages 66–77.
- Fritz, T. e Muller, S. C. (2016b). Leveraging Biometric Data to Boost Software Developer Productivity. pages 66–77.

- Fucci, D., Girardi, D., Novielli, N., Quaranta, L., e Lanubile, F. (2019). A Replication Study on Code Comprehension and Expertise using Lightweight Biometric Sensors. *Proceedings of the 27th International Conference on Program Comprehension*.
- Funke, G., Knott, B., Mancuso, V. F., Strang, A., Estepp, J., Menke, L., Brown, R., Dukes, A., e Miller, B. (2013). Evaluation of subjective and EEG-based measures of mental workload. *Communications in Computer and Information Science*, 373(PART I):412–416.
- Gonçales, L., Silva, B., e Fessler, J. (2019a). Measuring the Cognitive Load of Software Developers : A Systematic Mapping Study. *Proceedings of the 27th International Conference on Program Comprehension*, (March).
- Gonçales, L. J., Farias, K., Oliveira, T. C. D., e Scholl, M. (2019b). Comparison of software design models: An extended systematic mapping study. *ACM Computing Surveys (CSUR)*, 52(3):48.
- Guarnera, D. T., Bryant, C. A., Maletic, J. I., e Sharif, B. (2018). iTrace : Eye Tracking Infrastructure for Development Environments. pages 2015–2017.
- Gui, Q., Ruiz-Blondet, M. V., Laszlo, S., e Jin, Z. (2019). A Survey on Brain Biometrics. *ACM Computing Surveys*, 51(6):1–38.
- Haag, A., Goronzy, S., Schaich, P., e Williams, J. (2004). Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System. In André, E., Dybkjær, L., Minker, W., e Heisterkamp, P., editors, *Affective Dialogue Systems*, pages 36–48, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Hejmady, P. e Narayanan, N. H. (2012). Visual attention switching patterns of programmers debugging with an IDE. *ETRA '12 Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 197–200.
- Ikutani, Y. e Uwano, H. (2014). Brain activity measurement during program comprehension with NIRS. *2014 IEEE/ACIS 15th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, SNPD 2014 - Proceedings*, pages 1–6.
- Lane, R. D., Fink, G. R., Chau, P. M.-L., e Dolan, R. J. (1997). Neural activation during selective attention to subjective emotional responses. *Neuroreport*, 8(18):3969–3972.
- Lee, S., Matteson, A., Hooshyar, D., Kim, S., Jung, J., Nam, G., e Lim, H. (2016). Comparing Programming Language Comprehension between Novice and Expert Programmers Using EEG Analysis. *Proceedings - 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering, BIBE 2016*, pages 350–355.
- Li, M. e Lu, B. (2009). Emotion classification based on gamma-band EEG. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1223–1226.
- Li, S. Z. e Jain, A. K., editors (2015). *Encyclopedia of Biometrics, Second Edition*. Springer {US}.
- Lin, Y. T., Wu, C. C., Hou, T. Y., Lin, Y. C., Yang, F. Y., e Chang, C. H. (2016). Tracking Students' Cognitive Processes during Program Debugging-An Eye-Movement Approach. *IEEE Transactions on Education*, 59(3):175–186.

- Logozzo, F., Barnett, M., Manuel, F., e Cousot, P. (2012). A Semantic Integrated Development Environment. *Proceedings of the 3rd annual conference on Systems, programming, and applications: software for humanity*, pages 15–16.
- Obaidallah, U., Al Haek, M., e Cheng, P. C.-H. (2018). A Survey on the Usage of Eye-Tracking in Computer Programming. *ACM Computing Surveys*, 51(1):1–58.
- Peitek, N. (2018). A neuro-cognitive perspective of program comprehension. *2018 ACM/IEEE 40th International Conference on Software Engineering: Companion Proceedings*, pages 496–499.
- Petersen, K., Vakkalanka, S., e Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64:1–18.
- Radevski, S., Hata, H., e Matsumoto, K. (2015). Real-time monitoring of neural state in assessing and improving software developers' productivity. *Proceedings - 8th International Workshop on Cooperative and Human Aspects of Software Engineering, CHASE 2015*, pages 93–96.
- Rostami, S., Shenfield, A., Sigurnjak, S., e Fakorede, O. (2015). Evaluation of Mental Workload and Familiarity in Human Computer Interaction with Integrated Development Environments using Single-Channel EEG. *Psychology of Programming Interest Group 2015 - 26th Annual Workshop.*, 1(July):7–22.
- Sharafi, Z., Soh, Z., e Guéhéneuc, Y.-g. G. (2015). A systematic literature review on the usage of eye-tracking in software engineering. *Information and Software Technology*, 67:79–107.
- Siegmund, J., Kästner, C., Apel, S., Parnin, C., Bethmann, A., Leich, T., Saake, G., e Brechmann, A. (2014). Understanding understanding source code with functional magnetic resonance imaging. *Proceedings of the 36th International Conference on Software Engineering*, pages 378–389.
- Siegmund, J., Peitek, N., Parnin, C., Apel, S., Hofmeister, J., Kästner, C., Begel, A., Bethmann, A., e Brechmann, A. (2017). Measuring neural efficiency of program comprehension. *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*, pages 140–150.
- Snipes, W., Murphy-Hill, E., Fritz, T., Vakilian, M., Damevski, K., Nair, A. R., e Shepherd, D. (2015). Chapter 5 - A Practical Guide to Analyzing IDE Usage Data. pages 85–138. Morgan Kaufmann, Boston.
- Solovey, E. T., Afergan, D., Peck, E. M., Hincks, S. W., e Jacob, R. J. K. (2015). Designing Implicit Interfaces for Physiological Computing: Guidelines and Lessons Learned Using fNIRS. *ACM Trans. Comput.-Hum. Interact.*, 21:35:1–35:27.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12(2):257–285.
- Sweller, J., Ayres, P., e Kalyuga, S. (2011). *Cognitive Load Theory*. Springer, London, 1st edition.

- Szu, H., Hsu, C., Moon, G., Yamakawa, T., Tran, B. Q., Jung, T. P., e Landa, J. (2013). Smartphone Household Wireless Electroencephalogram Hat. *Applied Computational Intelligence and Soft Computing*, 2013:1–8.
- Uysal, M. P. (2016). Evaluation of learning environments for object- oriented programming : measuring cognitive load with a novel measurement technique. *Interactive Learning Environments* 24, 4820(May 2015).
- Veltman, J. e Gaillard, A. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 41(5):656–669.
- Wieringa, R., Maiden, N., Mead, N., e Rolland, C. (2006). Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. *Requirements engineering*, 11(1):102–107.
- Yeh, M. K., Gopstein, D., Yan, Y., e Zhuang, Y. (2017). Detecting and comparing brain activity in short program comprehension using EEG. *Proceedings - Frontiers in Education Conference, FIE*, 2017-Octob(1444827):1–5.
- Yin, M., Li, B., e Tao, C. (2010). Using cognitive easiness metric for program comprehension. *2nd International Conference on Software Engineering and Data Mining, SEDM 2010*, pages 134–139.
- Zayour, I. e Hajjdiab, H. (2013). How much integrated development environments (IDEs) improve productivity? *Journal of Software*, 8(10):2425–2431.
- Zayour, I. e Hamdar, A. (2016). A qualitative study on debugging under an enterprise IDE. *Information and Software Technology*, 70:130–139.

## A. Seleção de Estudos Primários

- S01** Alshbatat, A. I. N., Vial, P. J., Premaratne, P., e Tran, L. C. (2014). EEG-based Brain-computer Interface for Automating Home Appliances. *Journal of Computers*, 9(9):2159–2166
- S02** Bednarik, R. (2012). Expertise-dependent visual attention strategies develop over time during debugging with multiple code representations. *International Journal of Human Computer Studies*, 70(2):143–155
- S03** Beelders, T. e du Plessis, J.-P. (2016). The Influence of Syntax Highlighting on Scanning and Reading Behaviour for Source Code. pages 1–10. Association for Computing Machinery (ACM)
- S04** Crk, I. e Kluthe, T. (2014). Toward using alpha and theta brain waves to quantify programmer expertise. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, pages 5373–5376
- S05** Duraes, J., Madeira, H., Castelhana, J., Duarte, C., e Branco, M. C. (2016). WAP: Understanding the Brain at Software Debugging. *Proceedings - International Symposium on Software Reliability Engineering, ISSRE*, pages 87–92
- S06** Fakhoury, S., Ma, Y., Arnaoudova, V., e Adesope, O. (2018). The effect of poor source code lexicon and readability on developers' cognitive load. *Proceedings of the 26th Conference on Program Comprehension*, pages 286–296
- S07** Fakhoury, S. (2018). Moving towards objective measures of program comprehension. *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pages 936–939
- S08** Floyd, B., Santander, T., e Weimer, W. (2017). Decoding the Representation of Code in the Brain: An fMRI Study of Code Review and Expertise. *Proceedings - 2017 IEEE/ACM 39th International Conference on Software Engineering, ICSE 2017*, pages 175–186
- S09** Fritz, T., Begel, A., Müller, S. C., Yigit-Elliott, S., e Züger, M. (2014). Using psycho-physiological measures to assess task difficulty in software development. *Proceedings of the 36th International Conference on Software Engineering*, pages 402–413
- S10** Fritz, T. e Muller, S. C. (2016a). Leveraging Biometric Data to Boost Software Developer Productivity. *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering Leveraging*, pages 66–77
- S11** Funke, G., Knott, B., Mancuso, V. F., Strang, A., Estep, J., Menke, L., Brown, R., Dukes, A., e Miller, B. (2013). Evaluation of subjective and EEG-based measures of mental workload. *Communications in Computer and Information Science*, 373(PART I):412–416
- S12** Hejmady, P. e Narayanan, N. H. (2012). Visual attention switching patterns of programmers debugging with an IDE. *ETRA '12 Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 197–200
- S13** Ikutani, Y. e Uwano, H. (2014). Brain activity measurement during program comprehension with NIRS. *2014 IEEE/ACIS 15th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, SNPD 2014 - Proceedings*, pages 1–6

- S14** Lee, S., Matteson, A., Hooshyar, D., Kim, S., Jung, J., Nam, G., e Lim, H. (2016). Comparing Programming Language Comprehension between Novice and Expert Programmers Using EEG Analysis. *Proceedings - 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering, BIBE 2016*, pages 350–355
- S15** Logozzo, F., Barnett, M., Manuel, F., e Cousot, P. (2012). A Semantic Integrated Development Environment. *Proceedings of the 3rd annual conference on Systems, programming, and applications: software for humanity*, pages 15–16
- S16** Obaidallah, U., Al Haek, M., e Cheng, P. C.-H. (2018). A Survey on the Usage of Eye-Tracking in Computer Programming. *ACM Computing Surveys*, 51(1):1–58
- S17** Peitek, N. (2018). A neuro-cognitive perspective of program comprehension. *2018 ACM/IEEE 40th International Conference on Software Engineering: Companion Proceedings*, pages 496–499
- S18** Radevski, S., Hata, H., e Matsumoto, K. (2015). Real-time monitoring of neural state in assessing and improving software developers' productivity. *Proceedings - 8th International Workshop on Cooperative and Human Aspects of Software Engineering, CHASE 2015*, pages 93–96
- S19** Rostami, S., Shenfield, A., Sigurnjak, S., e Fakorede, O. (2015). Evaluation of Mental Workload and Familiarity in Human Computer Interaction with Integrated Development Environments using Single-Channel EEG. *Psychology of Programming Interest Group 2015 - 26th Annual Workshop.*, 1(July):7–22
- S20** Sharafi, Z., Soh, Z., e Guéhéneuc, Y.-g. G. (2015). A systematic literature review on the usage of eye-tracking in software engineering. *Information and Software Technology*, 67:79–107
- S21** Siegmund, J., Kästner, C., Apel, S., Parnin, C., Bethmann, A., Leich, T., Saake, G., e Brechmann, A. (2014). Understanding understanding source code with functional magnetic resonance imaging. *Proceedings of the 36th International Conference on Software Engineering*, pages 378–389
- S22** Siegmund, J., Peitek, N., Parnin, C., Apel, S., Hofmeister, J., Kästner, C., Beggel, A., Bethmann, A., e Brechmann, A. (2017). Measuring neural efficiency of program comprehension. *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*, pages 140–150
- S23** Szu, H., Hsu, C., Moon, G., Yamakawa, T., Tran, B. Q., Jung, T. P., e Landa, J. (2013). Smartphone Household Wireless Electroencephalogram Hat. *Applied Computational Intelligence and Soft Computing*, 2013:1–8
- S24** Uysal, M. P. (2016). Evaluation of learning environments for object-oriented programming : measuring cognitive load with a novel measurement technique. *Interactive Learning Environments* 24, 4820(May 2015)
- S25** Yeh, M. K., Gopstein, D., Yan, Y., e Zhuang, Y. (2017). Detecting and comparing brain activity in short program comprehension using EEG. *Proceedings - Frontiers in Education Conference, FIE, 2017-Octob(1444827)*:1–5
- S26** Yin, M., Li, B., e Tao, C. (2010). Using cognitive easiness metric for program comprehension. *2nd International Conference on Software Engineering and Data Mining, SEDM 2010*, pages 134–139
- S27** Zayour, I. e Hamdar, A. (2016). A qualitative study on debugging under an enterprise IDE. *Information and Software Technology*, 70:130–139

## B. Trabalhos Relacionados

- TR01** Fritz, T., Begel, A., Müller, S. C., Yigit-Elliott, S., e Züger, M. (2014). Using psycho-physiological measures to assess task difficulty in software development. *Proceedings of the 36th International Conference on Software Engineering*, pages 402–413
- TR02** Fritz, T. e Muller, S. C. (2016a). Leveraging Biometric Data to Boost Software Developer Productivity. *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering Leveraging*, pages 66–77
- TR03** Fucci, D., Girardi, D., Novielli, N., Quaranta, L., e Lanubile, F. (2019). A Replication Study on Code Comprehension and Expertise using Lightweight Biometric Sensors. *Proceedings of the 27th International Conference on Program Comprehension*
- TR04** Gonçalves, L., Silva, B., e Fessler, J. (2019a). Measuring the Cognitive Load of Software Developers : A Systematic Mapping Study. *Proceedings of the 27th International Conference on Program Comprehension*, (March)
- TR05** Gui, Q., Ruiz-Blondet, M. V., Laszlo, S., e Jin, Z. (2019). A Survey on Brain Biometrics. *ACM Computing Surveys*, 51(6):1–38
- TR06** Radevski, S., Hata, H., e Matsumoto, K. (2015). Real-time monitoring of neural state in assessing and improving software developers' productivity. *Proceedings - 8th International Workshop on Cooperative and Human Aspects of Software Engineering, CHASE 2015*, pages 93–96
- TR07** Rostami, S., Shenfield, A., Sigurnjak, S., e Fakorede, O. (2015). Evaluation of Mental Workload and Familiarity in Human Computer Interaction with Integrated Development Environments using Single-Channel EEG. *Psychology of Programming Interest Group 2015 - 26th Annual Workshop.*, 1(July):7–22

### C. Strings de Busca Individuais

Base de Dados	String de Busca
ACM Digital Library	acmdlTitle: ( +(eeg fmri electroencephalography neural brain cognitive psychometric bioinformatic "functional magnetic resonance imaging") +(software developer programmer professional) +(ide code editor "integrated development environmentssoftware engineering")) OR recordAbstract: ( +(eeg fmri electroencephalography neural brain cognitive psychometric bioinformatic "functional magnetic resonance imaging") +(software developer programmer professional) +(ide code editor "integrated development environmentssoftware engineering"))
IEEE Xplore	(software OR developer OR programmer OR professional OR "software engineering") AND (eeg OR fmri OR electroencephalography OR neural OR brain OR cognitive OR psychometric OR bioinformatic OR "functional magnetic resonance imaging") AND (ide OR code OR editor OR "integrated development environment")
Science Direct	({"software engineering"} OR developer OR programmer OR code) AND (IDE ) AND (eeg OR Fmri OR cognitive OR bci)
Scopus	ALL ( ( "software engineering"OR developer OR programmer OR code ) AND ( ide OR "integrated development environment") AND ( eeg OR fmri OR cognitive OR bci ) ) AND PUBYEAR >2008 AND ( LIMIT-TO ( SUBJAREA , "COMP") OR LIMIT-TO ( SUBJAREA , "ENGI") ) AND ( LIMIT-TO ( LANGUAGE , "English") )
Springer Link	("integrated development environment"OR IDE) AND ("software engineering"OR "software development"OR programmer OR developer) AND (bci OR eeg OR fmri OR cognitive)
DBLP	IDE—"integrated development environment"
Semantic Scholar	(IDE OR "integrated development environment")