

A confiança e a Ecologia da Desinformação: análise do comportamento de busca informacional

Os dados desta tese são provenientes de pesquisa quantiquantitativa com 77 usuários nas entrevistas e questionário (24 participantes da Geração X, 26 da Geração Y e 27 da Geração Z) e 65 no rastreamento ocular (19, 21 e 25, respectivamente). Todos os dados são primários, coletados pela própria pesquisadora durante os testes; não há reuso de dados existentes ou de terceiros.

Os dados incluem métricas extraídas do rastreamento ocular: Duração Total de Fixação na Área de Interesse (Total Duration of Fixation in AOI), Duração Média de Fixação na Área de Interesse (Average Duration of Fixation in AOI), Número de Fixações na Área de Interesse (Number of Fixations in AOI) e Número de Visitas (Number of Visits), permitindo analisar a alocação atencional e o esforço cognitivo envolvidos na avaliação de confiança nas notícias. Foram incluídas também duas métricas de pupilometria, por seu potencial em captar variações na carga cognitiva e na ativação emocional durante o processo de avaliação da confiança, complementando os dados comportamentais de atenção visual com um indicador fisiológico implícito. Foram coletados ainda dados de verbalização, como a nota atribuída pelos participantes a cada um dos 10 estímulos apresentados no teste de rastreamento ocular, as entrevistas gravadas em formato de tabela e os resultados do questionário.

Quanto aos formatos, as métricas de rastreamento ocular, os resultados do questionário e as entrevistas gravadas estão organizados em planilhas Excel (.xlsx); os 10 estímulos utilizados nos testes estão em formato HTML e os mapas de calor em imagem (.png); o instrumento do questionário está em formato Word (.docx); e os dados brutos do rastreamento ocular estão nos arquivos nativos do software Tobii Studio, versão 3.4. Trata-se de formatos amplamente utilizados e de fácil abertura, e os dados tabulares podem ser exportados para .csv, formato aberto e interoperável, assegurando compartilhamento e acesso de longo prazo independentemente de software proprietário.

Importante destacar que a coleta de dados foi autorizada pelo Comitê de Ética da Universidade Estadual Paulista - Campus Marília e com registro na Plataforma Brasil.

Coleta e criação dos dados

Os dados foram gerados a partir de três instrumentos complementares, triangulados para analisar a confiança como fenômeno multidimensional: (i) rastreamento ocular (eye tracking), incluindo dados de pupilometria e mapas de calor, capturados durante a exposição a 10 estímulos noticiosos apresentados em formato HTML; (ii) entrevistas semiestruturadas, registradas em áudio e posteriormente transcritas; e (iii) questionário estruturado, com atribuição de notas de confiabilidade em escala Likert de 1 a 5. O desenho experimental seguiu procedimentos consolidados na literatura de eye tracking, com estruturação da coleta em dois intervalos (Intervalo 1, de comportamento visual espontâneo, e Intervalo 2, de julgamento ativo).

O teste de comportamento ocular foi projetado e executado por meio do software Tobii Studio, versão 3.4, em conjunto com o rastreador ocular Tobii TX300, com taxa de amostragem de até 300 Hz (binocular). A calibração foi realizada com o modelo de cinco pontos. Para a pupilometria, aplicou-se correção de linha de base (baseline correction), utilizando um ponto neutro composto por uma tela cinza (RGB 128, 128, 128) apresentada antes de cada um dos estímulos.

Organização, nomenclatura e estrutura de pastas

Os arquivos foram organizados numericamente e por instrumento. Os estímulos foram identificados de 1 a 10 no Intervalo 1 e de 11 a 20 no Intervalo 2, mantendo a correspondência temática entre as duas etapas. Os dados são armazenados em uma estrutura hierárquica de pastas separadas por tipo de instrumento (rastreamento ocular, entrevistas, questionário) e por participante, adotando-se convenção de nomenclatura padronizada que combina identificador do participante, geração (X, Y ou Z), instrumento e data de coleta, evitando o uso de dados pessoais identificáveis nos nomes dos arquivos.

Versionamento

O controle de versões foi realizado por meio de nomenclatura sistemática com sufixos de versão (por exemplo, v01, v02) e registro de data, preservando-se sempre os dados brutos originais em cópia inalterada, separados dos arquivos de trabalho e das versões processadas ou tratadas.

Garantia de qualidade e consistência

A qualidade e a consistência da coleta foram asseguradas por: calibração do equipamento de rastreamento ocular antes de cada sessão; padronização das condições de aplicação (mesmos estímulos, na mesma ordem, entre os participantes); captura padronizada dos dados fisiológicos e comportamentais; transcrição e conferência das entrevistas; e validação da entrada de dados do questionário. A triangulação entre os três instrumentos funcionou, ainda, como mecanismo de verificação cruzada, permitindo confrontar os achados fisiológicos, comportamentais e discursivos.

Documentos que acompanham os dados

Cada conjunto de dados será acompanhado de descrição que permita a compreensão e a reutilização por usuários secundários. No nível descritivo básico, serão registrados: autoria, título do conjunto de dados, data de criação, e as condições de acesso e uso.

No nível metodológico e analítico, incluirá a descrição do desenho experimental (triangulação entre rastreamento ocular, entrevista semiestruturada e questionário), as especificações técnicas dos instrumentos (software Tobii Studio versão 3.4, rastreador Tobii TX300 a 300 Hz, calibração de cinco pontos e correção de linha de base com tela neutra cinza RGB 128,128,128), a descrição dos 10 estímulos e sua numeração correspondente entre os Intervalos 1 e 2, a definição das variáveis coletadas (por exemplo, dilatação pupilar, fixações,

mapas de calor, notas de confiabilidade em escala Likert de 1 a 5), as unidades de medida, os pressupostos assumidos e os formatos e tipos de arquivo de cada instrumento.

Para descrição e localização dos dados, será adotado o padrão de metadados da UNESP, por sua ampla aceitação, interoperabilidade e adequação à descrição de recursos em repositórios de acesso aberto.

Consentimento para preservação e compartilhamento

Devido ao envolvimento de seres humanos, os protocolos da pesquisa foram submetidos à Plataforma Brasil para registro (CAAE: 73313823.7.0000.5406), com autorização do Comitê de Ética da Universidade Estadual Paulista (UNESP). Todos os participantes assinaram Termo de Consentimento Livre e Esclarecido (TCLE) antes da coleta, no qual foram informados sobre os objetivos do estudo, os procedimentos (rastreamento ocular, entrevista e questionário), o caráter voluntário da participação e o tratamento dos dados, incluindo a possibilidade de preservação e compartilhamento dos dados anonimizados para fins de reutilização científica.

Proteção da identidade dos participantes

A identidade dos participantes é protegida por anonimização: cada participante é identificado por ordem numérica (por exemplo, P1, P2...), sem uso de nome ou dados diretamente identificáveis nos arquivos ou em suas nomenclaturas. Nas transcrições das entrevistas, informações que possam permitir a identificação (nomes próprios, locais, referências pessoais) são removidas ou substituídas. A chave que vincula os códigos aos participantes, quando existente, é mantida separada dos dados de pesquisa e sob acesso restrito à pesquisadora.

Armazenamento e transferência seguros de dados sensíveis

Como parte da gestão dos dados, as informações resultantes da pesquisa e o Plano de Gestão de Dados foram disponibilizados no Repositório Institucional da UNESP após a defesa da tese. Os dados sensíveis, especialmente as gravações e transcrições de entrevista, são armazenados em ambiente com acesso controlado e cópias de segurança (backup), e sua transferência, quando necessária, ocorre por meios seguros. O tratamento dos dados observa a Lei Geral de Proteção de Dados Pessoais (LGPD — Lei nº 13.709/2018), aplicando-se os princípios de finalidade, necessidade e segurança.

Retenção e acesso

Os dados anonimizados são de acesso aberto no Repositório Institucional da UNESP, enquanto os que possam identificar participantes permanecem sob acesso restrito, conforme o TCLE e as normas éticas aplicáveis.

Titularidade dos dados

A titularidade dos dados coletados e criados nesta pesquisa (registros de rastreamento ocular, pupilometria, mapas de calor, transcrições de entrevista e respostas ao questionário) observa

as políticas institucionais aplicáveis da Universidade Estadual Paulista (UNESP) e a legislação vigente, no contexto de sua produção no âmbito do Programa de Pós-Graduação em Ciência da Informação (PPGCI).

Licenciamento para reutilização

Os dados anonimizados passíveis de compartilhamento são licenciados sob Creative Commons Atribuição 4.0 Internacional (CC BY 4.0), que permite reutilização, redistribuição e adaptação, inclusive para fins comerciais, desde que atribuído o devido crédito à autoria. Os dados sensíveis, que possam identificar participantes, não são licenciados para reutilização aberta e permanecem sob acesso restrito, conforme as condições éticas e o TCLE.

Restrições relativas a dados de terceiros

Os estímulos utilizados na pesquisa foram 10 matérias jornalísticas reais, extraídas de portais de notícias brasileiros e em circulação nos dias que antecederam a coleta. Os direitos autorais desses conteúdos pertencem aos respectivos veículos de comunicação. As notícias foram utilizadas exclusivamente para fins de pesquisa científica, ao amparo das limitações e exceções ao direito autoral previstas na legislação brasileira (uso para fins de estudo, crítica e pesquisa). Por essa razão, o conteúdo integral dos estímulos não é redistribuído sob a licença CC BY 4.0, que se aplica apenas aos dados de autoria própria da pesquisadora. Os estímulos são disponibilizados nos apêndices da tese com indicação da fonte, do veículo e da data de publicação original, preservando os créditos aos detentores dos direitos.

Restrições ou adiamento do compartilhamento

Não há previsão de embargo por motivos de patenteamento, uma vez que a pesquisa, de natureza teórica e empírica no campo da Ciência da Informação, não gera invenção passível de proteção patentária. As únicas restrições ao compartilhamento decorrem da proteção aos participantes (dados sensíveis sob acesso restrito), e não de interesses comerciais ou de propriedade intelectual.

Capacidade de armazenamento

O volume de dados gerado é compatível com os recursos de armazenamento disponíveis, não havendo necessidade de contratação de serviços adicionais pagos. A pesquisadora utilizou a licença do Google Workspace disponibilizada pela UNESP aos seus estudantes, que oferece capacidade de armazenamento suficiente para o conjunto de dados da pesquisa.

Rotina e locais de backup

Foram mantidas três cópias dos dados, em locais distintos: (1) no Google Drive, por meio da conta institucional disponibilizada pela UNESP aos estudantes; (2) em cópia local no computador pessoal da pesquisadora; e (3) em backup no iCloud. A manutenção de múltiplas cópias em serviços e dispositivos diferentes reduz o risco de perda em caso de falha de qualquer um deles. Após a defesa, a preservação de longo prazo é assegurada pelo depósito no Repositório Institucional da UNESP.

Responsabilidade pelo backup e recuperação

A pesquisadora, Suellen Elise Timm Barros, foi a responsável pela realização e verificação dos backups durante todo o período da pesquisa. A preservação e a integridade dos dados depositados após a defesa ficam a cargo da equipe responsável pelo Repositório Institucional da UNESP.

Recuperação em caso de incidente

Em caso de perda, corrupção ou indisponibilidade de dados, a recuperação é feita a partir de qualquer uma das três cópias mantidas em locais distintos. A redundância entre serviços de nuvem e armazenamento local garante que a falha em um único ponto não resulte em perda definitiva dos dados.

Riscos à segurança e sua gestão

Os principais riscos referem-se ao acesso não autorizado a dados sensíveis (gravações e transcrições de entrevista, que podem identificar participantes) e à eventual perda ou vazamento desses dados. Esses riscos são mitigados pela anonimização dos dados, pela restrição de acesso e pela manutenção das informações em contas protegidas por autenticação. Os dados sensíveis não são divulgados publicamente e permanecem separados dos dados anonimizados destinados ao compartilhamento.

Controle de acesso

O acesso aos dados durante a pesquisa foi restrito à pesquisadora, em contas protegidas por senha e autenticação, sem compartilhamento público de links ou pastas. A chave que vincula os códigos de identificação aos participantes, quando existente, é mantida separada dos demais dados e sob acesso exclusivo da pesquisadora.

Acesso seguro por colaboradores

Como a coleta e o tratamento dos dados foram conduzidos pela própria pesquisadora, não houve compartilhamento amplo de dados sensíveis com terceiros. Quando necessário o acesso da orientação ou da banca, este ocorreu por meio de dados já anonimizados ou de resultados agregados. Eventuais compartilhamentos são feitos por meios seguros (compartilhamento restrito e autenticado no Google Drive institucional), evitando canais não protegidos como e-mail aberto ou dispositivos removíveis.

Transferência segura dos dados coletados

Os dados coletados no ambiente de aplicação (rastreamento ocular e entrevistas) foram transferidos, o mais brevemente possível após cada sessão, para os ambientes de armazenamento seguros descritos, reduzindo o tempo em que permaneceram apenas no equipamento de coleta. O tratamento observa a Lei Geral de Proteção de Dados Pessoais (LGPD — Lei nº 13.709/2018), aplicando-se os princípios de segurança, finalidade e necessidade.

Dados a reter ou destruir por obrigações legais e éticas

Os dados que possam identificar os participantes (gravações e transcrições de entrevista) são retidos apenas pelo período necessário e sob acesso restrito, em conformidade com o Termo de Consentimento Livre e Esclarecido, as normas do Comitê de Ética (CAAE: 73313823.7.0000.5406) e a LGPD (Lei nº 13.709/2018). Findo o período de guarda estabelecido junto ao Comitê de Ética, os dados identificáveis podem ser eliminados de forma segura, caso não haja justificativa para sua manutenção.

Crítérios para decidir o que preservar

São preservados prioritariamente os dados com maior valor de reutilização e que sustentam os achados da tese: os dados anonimizados de rastreamento ocular e pupilometria, os mapas de calor, as respostas do questionário e os dados anonimizados/agregados das entrevistas. Dados brutos redundantes, intermediários ou de baixa relevância analítica podem não ser preservados a longo prazo. A decisão considera o valor científico, a viabilidade de armazenamento e o esforço adicional necessário para preparar cada conjunto para compartilhamento.

Usos previstos para os dados

Os dados têm potencial de reutilização para: validação e replicação dos achados desta tese; realização de novos estudos comparativos sobre confiança informacional, heurísticas geracionais e desinformação; desenvolvimento metodológico em pesquisas que combinem rastreamento ocular, pupilometria e métodos qualitativos; e uso para fins didáticos e de formação em Ciência da Informação.

Prazo de retenção e preservação

Os dados anonimizados passíveis de compartilhamento são preservados de forma duradoura no Repositório Institucional da UNESP, com identificador persistente, garantindo acesso continuado à comunidade científica. Os dados sensíveis são mantidos sob acesso restrito conforme as normas éticas aplicáveis e, findo o período de guarda estabelecido, são eliminados de forma segura.

Local de preservação

Os dados de valor duradouro são preservados no Repositório Institucional da UNESP, que oferece armazenamento gerenciado, identificador persistente e curadoria de longo prazo, para além da vigência da pesquisa.

Custos

O depósito no Repositório Institucional da UNESP não implica custos para a pesquisadora, por se tratar de infraestrutura institucional de acesso aberto.

Esforço de preparação

A preparação dos dados para compartilhamento e preservação (anonimização, organização, documentação (README) e conversão para formatos estáveis) foi prevista e realizada pela própria pesquisadora como parte das atividades da tese.

Como os usuários encontrarão os dados

Os dados são divulgados por meio do Repositório Institucional da UNESP, indexado e de acesso aberto, com metadados descritivos que facilitam sua localização por mecanismos de busca e agregadores. O conjunto de dados fica vinculado à tese, ampliando sua visibilidade.

Com quem e sob quais condições

Os dados anonimizados passíveis de compartilhamento são disponibilizados publicamente à comunidade científica sob licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0). Os dados sensíveis, que possam identificar participantes, não são compartilhados abertamente e permanecem sob acesso restrito, conforme o TCLE e as normas éticas.

Quando os dados serão disponibilizados

Os dados foram disponibilizados após a defesa da tese.

Identificador persistente

O conjunto de dados recebe identificador persistente atribuído pelo Repositório Institucional da UNESP, assegurando citabilidade e localização estável ao longo do tempo.

Reconhecimento da reutilização

Espera-se que a reutilização dos dados seja reconhecida por meio de citação formal ao conjunto de dados, com atribuição de crédito à autoria, conforme os termos da licença CC BY 4.0 e a forma de citação indicada na documentação (README) que acompanha os dados.

Restrições previstas e medidas para minimizá-las

A principal restrição decorre da natureza sensível de parte dos dados, ou seja, das gravações e transcrições de entrevista, que podem identificar os participantes. Para superar ou minimizar essa restrição sem comprometer a proteção dos participantes, adota-se a anonimização dos dados, disponibilizando-se abertamente as versões anonimizadas e agregadas, enquanto os dados identificáveis permanecem sob acesso restrito. Outra restrição refere-se aos estímulos, que são matérias jornalísticas de terceiros protegidas por direito autoral; nesse caso, o conteúdo é disponibilizado apenas com indicação de fonte e créditos, sem redistribuição sob licença aberta.

Necessidade de uso exclusivo

Foi necessário o uso exclusivo dos dados durante o período de desenvolvimento e defesa da tese, a fim de assegurar a originalidade e a integridade da análise até a conclusão do trabalho. Concluída a defesa, não há interesse em prolongar o uso exclusivo.

Necessidade de acordo de compartilhamento

Para os dados anonimizados de acesso aberto, não é exigido acordo de compartilhamento, uma vez que a licença CC BY 4.0 já define as condições de reutilização. Para eventual acesso aos dados sensíveis, quando cabível e eticamente justificável, o compartilhamento ficaria condicionado a um acordo específico e à observância das normas éticas, garantindo a proteção da confidencialidade dos participantes.

Responsável pela implementação e revisão do PGD

A pesquisadora Suellen Elise Timm Barros é a responsável pela implementação, revisão e atualização do Plano de Gestão de Dados ao longo da pesquisa, com a supervisão da orientadora, Silvana Aparecida Borsetti Gregorio Vidotti.

Responsabilidade por cada atividade de gestão de dados

Por se tratar de pesquisa individual de doutorado, a pesquisadora concentra a responsabilidade pelas atividades de gestão de dados: coleta e criação dos dados (rastreamento ocular, entrevistas e questionário), produção de documentação e metadados, controle de qualidade, armazenamento e backup, anonimização, arquivamento e compartilhamento. A orientadora acompanha e valida as decisões metodológicas e de gestão relacionadas ao projeto.

Expertise especializada

A pesquisa exigiu conhecimento em rastreamento ocular e pupilometria, competências desenvolvidas pela pesquisadora durante a formação, não sendo necessária contratação adicional de pessoal especializado.

Hardware e software

Foram utilizados o rastreador ocular Tobii TX300 e o software Tobii Studio 3.4, disponíveis na infraestrutura institucional, além do Google Workspace da UNESP para armazenamento. Não houve necessidade de recursos excepcionais além dos já providos pela instituição.