



Questões de consciência sobre a coleta de dados de usuários por bibliotecas digitais

Elaine Parra Affonso^a

Transcrição da apresentação

Essa apresentação faz parte das atividades do 8º CEDIT (Ciclo de Estudos Dados Informação e Tecnologia) da UNESP de Tupã. O objetivo é explicitar sobre a insciência do usuário na fase de coleta de dados, para tanto relata se os resultados alcançados no artigo intitulado “Privacy awareness issues in user data collection by digital libraries”, publicado na IFLA journal.

Meu nome é Elaine Parra Affonso, professora da Fatec/Presidente Prudente, autora deste artigo que foi desenvolvido juntamente com o Professor Ricardo Gonçalves Sant’Ana da UNESP de Tupã.

Com o aumento exponencial dos dispositivos tecnológicos ampliam-se também as atividades que realizam coleta de dados. Além da coleta de dados realizada nos ambientes digitais de modo explícito, existem dados que circulam silenciosamente nas redes de computadores, o que resulta na insciência sobre o processo de coleta de dados.

^a Doutora em Ciência da Informação. Professora na Faculdade de Tecnologia de São Paulo – FATEC. E-mail: elainepff@gmail.com. ORCID: <https://orcid.org/0000-0002-3953-462X>. Currículo: <http://lattes.cnpq.br/8697314376216832>.

No cenário das bibliotecas digitais, a coleta de dados pode ocorrer no momento de interação do usuário ao realizar uma busca no site, ao preencher um cadastro para solicitar informações, incluindo os dados de tráfego nas redes de computadores.

O objetivo deste trabalho foi investigar os aspectos de privacidade na fase de coleta de dados pelas Bibliotecas Nacionais Digitais da América do Sul.

A metodologia deste trabalho baseou – se em:

- Identificação de Bibliotecas Nacionais Digitais de países da América do Sul.
- Pesquisa exploratória nos sites das bibliotecas digitais para identificar as seguintes questões: disponibilização de políticas de privacidade; protocolo de comunicação utilizado; identificação de dados coletados com a consciência do usuário; identificação de possíveis dados coletados implicitamente durante a interação do usuário com o ambiente digital, por meio da ferramenta WireShark.

Os pacotes trafegados na rede de computadores durante a interação com o site da Biblioteca Digital do Brasil foram coletados e foi selecionado um pacote para análise e exemplificação.

Ressalta-se que a coleta realizada durante a interação do usuário com o ambiente da biblioteca digital se deu pelo equipamento do autor e utilizando rede doméstica.

A coleta de dados desta pesquisa foi realizada no mês de julho de 2017.

Resultados: Foram identificados nove países da América do Sul que disponibilizam Bibliotecas Digitais Nacionais.

A Biblioteca Nacional Digital do Brasil e da Colômbia apresentam orientações sobre questões que envolvem privacidade.

Na exploração dos sites foi observado que a disponibilidade de políticas de privacidade para promover a consciência sobre o processo de coleta de dados, é limitada.

A maioria das bibliotecas utilizam o protocolo HTTP, configurando assim problemas com a segurança dos dados e consequentemente ameaças à privacidade e proteção de dados pessoais.

Três bibliotecas (Argentina, Brasil e Chile) solicitam algum cadastro para que seja possível reservar documentos, Isso ampliar acesso aos dados sobre os usuários e possíveis implicações na privacidade.

Os dados coletados que são explícitos para os usuários são aqueles solicitados no momento de cadastro, tais como: nome, e-mail, login, password, termo de busca, profissão e comentários.

Ao analisar um pacote de dados de tráfego durante a interação do usuário com ao ambiente nota – se a presença de outros dados, tais como: data e hora de solicitação, endereço IP, dados de localização, informações de navegador e sistema operacional, cookies e endereço MAC. Esses dados isolados podem não caracteriza ameaças à privacidade, no entanto, quando combinado com outros dados podem implica em ameaças à privacidade do indivíduo.

Esses dados não são perceptíveis para o usuário no momento da interação com o ambiente digital, o que aumenta a abstração para o usuário sobre o processo, ao contrário do detentor que pode ampliar o conhecimento sobre esse processo, cenário que confirma a assimetria informacional entre detentores de dados e usuário. Desta forma a percepção da coleta e em apenas em relação aos dados que são relatados pelo próprio usuário.

Destacamos os problemas de privacidade na fase de coleta, analisando dados que são coletados explicitamente e implicitamente.

As camadas de abstração presentes nas redes de

computadores ao organizar e simplificar o seu contexto complexo, encapsulam detalhes de comunicação, o que implica em problemas relacionados à privacidade.

As ameaças à privacidade podem ser ampliadas pela insciência do usuário sobre quando, como e onde a coleta de dados ocorre.

Outros aspectos também pode ser preocupantes, como a possível correlação dos dados com outros bancos de dados, formando perfis e aumentando o conhecimento dos detentores de dados em relação ao usuário.

Assim bibliotecas digitais precisam disponibilizar políticas de privacidade a fim de orientar os usuários em relação a coleta de seus dados.

Essas políticas não devem específicas apenas dados que os usuários fornecem voluntariamente mas também dados que são abstraídos das camadas de redes de computadores.

Essa é a referência do artigo citado na apresentação, e segue o contato dos autores.

Vídeo da Apresentação

Título: Questões de consciência sobre a coleta de dados de usuários por bibliotecas digitais



Privacy awareness issues in user data collection by digital libraries

Publicado: **IFLA journal**, v. 44, n. 3, p. 170-182, 2018.



Elaine Parra Affonso
Fatec Presidente Prudente



Ricardo César Gonçalves Sant'Ana
Unesp/Tupã

Disponível em:

http://dadosabertos.info/enhanced_publications/8ccedit/video.php?id=4

Slides da Apresentação

Título: Questões de consciência sobre a coleta de dados de usuários por bibliotecas digitais



Questões de consciência sobre a coleta de dados de usuário por bibliotecas digitais

Disponível em:

http://dadosabertos.info/enhanced_publications/8cedit/presentation.php?id=4