



**Universidade Federal de Pernambuco - UFPE**  
**Centro de Informática - CIn**

**Proposta de TCC**

Gabriel Estevam Longuinhos (gel)

**Recife**  
**2021**

- (i) nome completo do aluno;**
- (ii) curso (CC, EC ou SI);**
- (iii) endereço eletrônico do aluno no CIn-UFPE;**
- (iv) título do trabalho;**
- (v) área;**
- (vi) nome do(a) orientador(a);**
- (vii) uma lista de possíveis avaliadores (de comum acordo com o orientador);**
- (viii) resumo (abstract) de no máximo 10 linhas**

i. Gabriel Estevam Longuinhos

ii. Sistemas de Informação

iii. [gel@cin.ufpe.br](mailto:gel@cin.ufpe.br)

iv. Catálogo de Metadados de Rastreabilidade em Data Lakes

v. Banco de Dados

vi. Bernadette Farias Loscio

vii. Robson Fidalgo

viii. A Lei Geral de Proteção de Dados Pessoais (LGPD) dispõe, sobre o tratamento de dados pessoais com o objetivo de proteger os direitos fundamentais de liberdade e privacidade da pessoa natural. Data lakes são soluções de Big Data que armazenam dados heterogêneos em seu formato mais bruto, dentre esses dados, podem haver muitos dados pessoais que recebem diversos tratamentos. Uma maneira de fornecer transparência sobre o tratamento realizado nos dados em data lakes é por meio da captura de metadados de rastreabilidade. Atualmente existem várias abordagens para coleta de metadados, no entanto, poucas estratégias exploram o uso desses metadados. Sendo assim, o objetivo deste trabalho é disponibilizar um catálogo que permita consultar e explorar os metadados coletados em ambientes de data lake.