

## **Gerenciamento ético participativo e reflexivo na ciência de dados e uma abordagem a questões de opacidade em algoritmos.**

### **RESUMO**

O primeiro artigo mostra que as etapas do trabalho na ciência dos dados são amplamente distribuídas, com grupos de culturas diferentes e, até mesmo para um único conjunto de dados, centenas de pessoas podem estar envolvidas em decidir quais experimentos fazer e como calibrar os instrumentos; como os dados devem ser extraídos, formatados e visualizados; e quais insights mostrar. Assim, são grandes os desafios para avaliar a qualidade e os impactos éticos das declarações dos dados disseminados. Além disso, o que significa ‘participação’ e quais as responsabilidades dos provedores e usuários de dados (pesquisadores que reutilizam os dados) e quais as implicações, de se usar ferramentas de mineração e interpretação de dados, desenvolvidas fora do contexto das preocupações éticas com as situações de uso dos dados. O autor defende a supervisão ética através do gerenciamento participativo e reflexivo das práticas de dados, como o governo do Reino Unido fez na questão de experimentos com animais, baseado nos princípios dos 3Rs, adequados porque sua aplicação varia de acordo com as situações, dependendo dos objetivos, métodos e ferramentas da pesquisa (critérios que não são satisfeitos com a adoção de diretrizes éticas gerais). Assim, é necessário um modelo que entrelaça decisões técnicas com considerações éticas, através: (i) do treinamento contínuo em ética aplicado à prática científica; (ii) da interação entre inspetores, instituições que hospedam projetos e pesquisadores através de encontros regulares para avaliação ética participativa em que os indivíduos trocam idéias sobre potenciais implicações do seu trabalho. As estruturas reguladoras devem promover a conscientização de como os diferentes estágios da ciência de dados se relacionam e destacar como cada perspectiva individual é inevitavelmente parcial e precisa de coordenação com as outras. Isso mantém um diálogo sobre as melhores práticas favorecendo a qualidade e confiabilidade dos resultados da pesquisa e enfatiza como intervenções individuais fazem diferença nos resultados, podendo ter implicações éticas severas, como na escolha do tratamento de pacientes, que depende da maneira que seus dados são gerenciados por algoritmos e visualizações, que processam informações médicas, inclusive determinando quem pode acessar e processar esses dados. Logo, é importante engajar o colaborador para que documente e avalie alegações feitas como resultado de processamento e interpretação de dados em cada estágio. Eles precisam assumir essa responsabilidade e debater escolhas passadas, avaliar a sustentabilidade de soluções atuais e analisar futuras implicações.

O objetivo do segundo trabalho foi o de analisar os reais motivos por trás da opacidade de algoritmos de classificação de aprendizagem de máquina, principalmente aspectos de discriminação. A opacidade foi dividida em três tipos: sigilo corporativo - seria vantagem competitiva e/ou para ocultar regulamentações, manipulação de consumidores ou padrões de discriminação?; analfabetismo técnico - o público não tem condições de avaliar esses mecanismos (resultados da análise de crédito, etc.) que afetam suas oportunidades de vida; e a opacidade intrínseca da complexidade dos algoritmos de máquina - para determinados problemas uma instância simples desses algoritmos não é precisa na sua classificação. O artigo aborda que os algoritmos de aprendizagem de máquina não são totalmente objetivos, isentos de inadequações ou injustiças na classificação, dado o grau de julgamento humano envolvido no processo, o que inclui recursos de definição, pré-classificação dos dados de treinamento e ajuste de parâmetros. Por exemplo, o caso em que o classificador do Google Fotos identificou fotos de pessoas negras como "Gorilas". Algumas soluções para os problemas sociais com essa "caixa preta" são auditorias nas empresas, projetos open source, educação computacional do público em geral e denúncias de exclusões e formas de discriminação vivenciadas.

## REFERÊNCIAS

BURRELL, Jenna. **How the machine 'thinks': Understanding opacity in machine learning algorithms**. In: Sage Journals, 2016. Disponível em:

<<http://journals.sagepub.com/doi/10.1177/2053951715622512>> Acesso em: mai, 2018.

LEONELLI, Sabina. **Locating ethics in data science: responsibility and accountability in global and distributed knowledge production systems**. In: The Royal Society, 2016. Disponível em:

<<http://rsta.royalsocietypublishing.org/content/374/2083/20160122>> Acesso em: mai, 2018.