**Análise do artigo: On the Societal Impact of Open Foundation Models (Kapoor and Bommasani et al. 2024)**

Os modelos de Machine Learning (ML) distribuídos abertamente apresentam aspectos positivos e negativos. Na perspectiva positiva, considera-se que os modelos distribuídos abertamente trazem inovações para ciência, aumenta a tomada de decisão e a partir desses modelos, outros modelos podem ser customizados para diversas aplicações. Esses modelos apresentam grande quantidade de dados no treinamento, tornando assim o modelo mais robusto e com menor probabilidade de erros do que os modelos específicos. O modelo aberto não precisa ser acompanhado de código, dados, ou computação, bem como, o modelo pode ser utilizado em *hardware* local, o que elimina o compartilhamento de dados para o desenvolvedor. Os modelos base são marcados por acesso mais amplo e maior customização e a usabilidade influencia fortemente no processo de inovação.

No entanto, a distribuição aberta dos modelos de ML também apresentam riscos e pontos negativos à sociedade. Eles podem ocorrer a partir de ataques cibernéticos, golpes de clonagem de voz, fraudes e relacionados à biossegurança, entre outros. Os autores apontam que a maior parte de pesquisas anteriores não analisam os riscos a que os modelos de ML podem ser submetidos. Existem vários riscos relacionados ao uso indevido e quanto aos que desejam competir com os desenvolvedores do modelo base. Para tanto, restrições de uso para quem deseja assim competir são formalizadas a partir de licença do desenvolvedor, bem como, restrição na idade dos usuários e à exposição de conteúdos sensíveis. De tal modo, os riscos irão evoluir à medida que as capacidades dos modelos e as defesas sociais também evoluírem.

Portanto, a partir dos modelos abertos, é possível uma compreensão mais profunda de como os modelos estão operando. Logo, os desenvolvedores que os utilizam têm opções mais específicas a partir do modelo base, tendo em vista aplicações posteriores. É importante ressaltar que o direcionamento de recomendações para os desenvolvedores da IA, para pesquisadores de riscos da IA, para os reguladores de concorrência e os gestores, com foco na compreensão dos benefícios e na mitigação dos riscos, é essencial para uma fundamentação empírica das pesquisas sobre a distribuição dos modelos. Em tese, a partir da utilização dos modelos abertos, uma melhor investigação sobre seus impactos sociais podem ser fomentados. Por outro lado, os modelos abertos também podem permitir que atores mal-intencionados utilizem os modelos para gerar desinformação, fraudes ou mesmo ataques cibernéticos.