

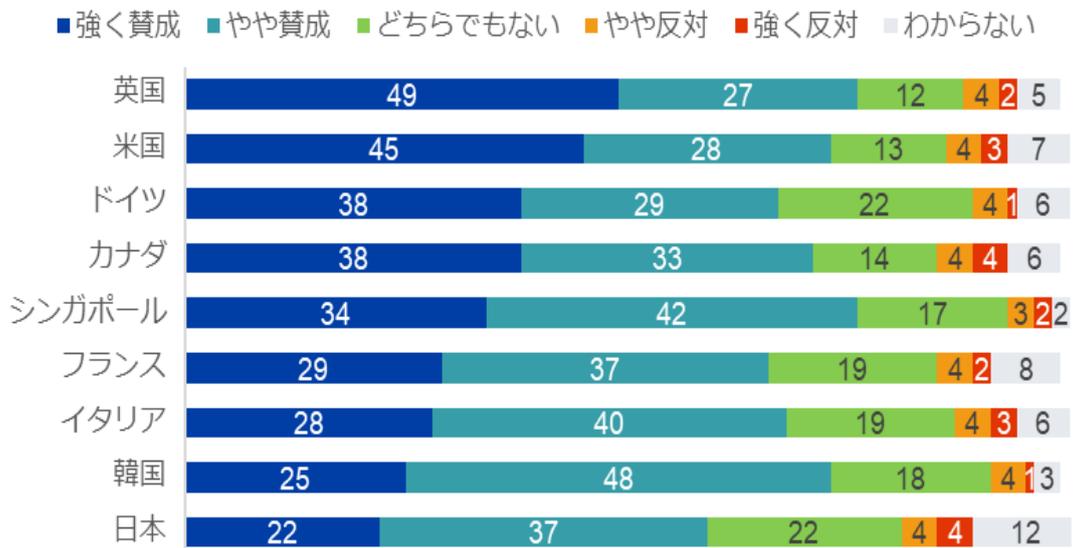
表 フロンティアAI企業の安全性確保に向けた取り組み分野

1	責任ある能力拡張
2	モデルの評価と攻撃への対策評価
3	モデルに関する報告と情報共有
4	セキュリティ管理
5	脆弱(ぜいじゃく)性に対する報告体制
6	AIが生成した作品に対する識別子
7	AIによるリスクに関する研究の優先化
8	モデルの誤用の防止と監視
9	インプットしたデータの管理と監査

(出所) 英国政府

図 9カ国の国民への意識調査結果

質問：強力なAIは、独立した専門家によって安全性を検査されるべきか



(注) 合計が100にならない場合がある。

(出所) 英国政府