

リスク配分システムとしての法

人間中心のAI社会原則検討会議 2018年6月1日
慶應義塾大学法学部 大屋雄裕

「じこはおこるさ」(きかんしゃトーマスオールスターズ)

起きてしまうトラブルと損害

①事後の観点：どのように社会的に配分するか

②事前の観点：事後の損害配分を前提にした注意義務への転化

損害を負うことになるであろう → 結果回避へのインセンティブ
一つのデフォルト値＝「泣き寝入り」(被害者負担)

配分制度を人為的に作らない場合に発生し得る状態 → 社会的反感
適切な配分制度の構築・社会的合意の形成

eg. 過失責任主義＝故意・過失による責任の基礎付け／無過失免責
ここで「過失」の内容は多分に客観化(結果回避義務・予見義務)

AI技術の進展による問題＝人間による注意力の限界

cf. 第1回参考資料2 (Preferred Networks 意見)

①統計的機械学習技術の応用 (C)

判断過程の(相対的)ブラックボックス化を惹起

開発者・利用者による予測可能性↓ = 予見義務の実効性低下

②ネットワーク化した情報システム (B)

開発者・利用者の指示を遵守?…… cf. 牛馬による交通

一定の自律性+高速化 → 阻止・操作可能性↓ = 結果回避義務の実効性低下
義務を配分された主体が現実に注意・結果回避を実現できるか?

過失責任主義の機能不全

議論の枠組

存在するリスク＝保護されるべき法益は何か?

eg. 生命、身体、同意のある財産、実験室の安全……

解決の可能性は何か?

被害者負担、損害分配ルールの形成、保険による分散、無過失責任、行為規制
解決手法の速度と強度は?

遅い：法規制 — 共同規制 — 自主規制 — 無規制 : 速い

強い：刑事罰を伴う強制—民事違法—訓示規定—ガイドライン：弱い

cf. 公表など弱い負のサンクション、褒賞など正のサンクションも

「悪人の視点」the bad man's point of view

(Oliver Wendell Holmes, Jr., "The Path of the Law" (1897))

制度の悪用を想定し、それに対して強靱な制度を構想する必要

環境規制……未然防止 prevention から事前配慮 precaution へ

しかし：ゼロ・リスクは存在しない、実験室は別の話、法益の違い

③汎用AI (A)

いまはまだないもの＝リスクは不確定・不確実 → 一定の対処が必要

被害者負担に陥る危険、「自称汎用AI」の危険