

科学研究費補助金・基盤研究(S)  
裁判過程における人工知能による高次推論支援

# 法的判断の構造とモデル化の探求

— AI は人々に受け容れられるか? —

太田 勝造 (OTA Shozo)  
東京大学大学院法学政治学研究科  
UTokyo Law

基盤研究(A) 課題番号18H03612(2018年度~2022年度)  
太田勝造(\*), 加藤淳子(\*), 佐藤健, 浅水屋剛(\*), 今井耕介,  
岡田謙介, 齋藤宙治(\*), 米村滋人, 森大輔, 飯田高

(国立情報学研究所 October 02, 2018)

# 研究目的

## (1) リーガル・マインド

アメリカ合衆国のロー・スクールなどで法的判断の核心は **Legal Mind** であるとされるが明確な定義はない。せいぜい、

**"Think like a lawyer."**

と説明されるが、tautological な empty formula に過ぎない。

法的判断の専門家としての法律家は「リーガル・マインド」を持つとされる。リーガルマインドの重要な要素は日本の法科大学院教育では、

**「具体的事実を法律要件に当てはめて法律効果の有無を判断する」**

という「要件＝効果思考様式」とされ、これが政策科学的な「目的＝手段思考様式」との相違点とされる。

本研究では、fMRI（機能的磁気共鳴画像装置）やアイ・トラッキングなどの先進技術を活用した実験、実験室実験、フィールド・エクスペリメントなど、文系・理系の壁を越えた多角的かつ学際的手法を駆使して、この捉え難いリーガル・マインドにアプローチし、近時発達が著しいAIによるモデル化のための基盤を構築する。

## (2) AI裁判に対する国民の受容

佐藤健プロジェクトによって、AIによる裁判支援システムが構築された場合に、国民がその正当性を認め受容するにはどのようなシステムであるべきかを、明らかにすることが第二の研究目的である。

# 研究手法

## (1) リーガル・マインド

**参加者：**大学生（非法学部生と法科大学院生，将来は弁護士や裁判官）

**事前準備：**法的判断に関わる内容の質問票を配布し，回答を回収する。  
法意識，規範意識，遵法精神，法イメージ，公共精神，  
刑罰目的論，裁判への信頼度，政治への信頼度  
自己評価(self-esteem)，パースナリティ特性評価尺度，  
生活満足度など

**実査：**fMRIスキャナーに入ってもらい，（スキャナーの中で）参加者は刑事裁判事例を読みながら，順次，質問に答える。

**事後質問：**事後にチェックのために，質問票に記入してもらう。

**課題Ⅰ：**事実認定の法的基準の判断の脳活動

「合理的疑いを挟まない程度の確実性(Beyond the Reasonable Doubt)」と被告人の有罪・無罪の判断（説示なしとあり）

**課題Ⅱ：**量刑判断における理性と情緒の重み

「反省・悔悟がある場合と，無い場合での量刑判断」

「反省・悔悟の機能説明：更生の見込み，応報感情緩和」

## (1) リーガル・マインド：事案と証拠関係

《事案①》本件の争点は、被告人がB子とCちゃんの2名を殺害した犯人であるかどうかです。本件では、被告人が犯行を全面的に否認し、被害者は死亡しており、目撃者もいないため、直接証拠はありません。もっとも、各事実を裏付ける後述の状況証拠【1】～【4】があります。

【1】被告人と被害者家族の関係 (1/8) 被告人は、A男が子供のころにその実母と婚姻し、養父としてA男を育てた。もっとも、A男の借金問題、女性問題等をきっかけに、本件当時は被害者家族と必ずしも良好な関係にあったとはいえない。

【2】被告人の犯行当日の行動 (2/8) 《ア》本件マンションの踊り場の灰皿内から、本件事件の翌日にたばこの吸い殻72本が採取されたが、その中に被告人が好んで吸っていた銘柄の吸い殻が1本あった。この吸い殻に付着していた唾液中の細胞のDNA型は、被告人の血液のDNA型と一致しており、このDNA型一致の出現頻度は1000万人に2人という極めて低いものである。

【2】被告人の犯行当日の行動 (3/8) 《イ》本件犯行時刻頃、被告人が当時使用していた自動車と同種・同色の自動車が、本件マンション付近に駐車されていた。被告人自身も、捜査段階においては、本件事件当日に自己の運転する自動車を同地点に駐車したことを認めていたが、その後、否認した。

【2】被告人の犯行当日の行動 (4/8) 《ウ》被告人は、本件当日の夕方、妻を迎えに行く約束をしていたが、その約束をたがえ、B子及びCちゃんが殺害された時間帯には、自らの携帯電話の電源も切っていた。

【3】被告人の犯行動機 (5/8) 《ア》本件犯行の半年前、B子及びCちゃんは、被告人宅で1か月間、被告人と同居したが、そのころ、被告人は、B子に対し、恋慕の情を抱いており、性交渉を迫る、抱き付く、キスをするなどの行為に及んだことがあった。

【3】被告人の犯行動機 (6/8) 《イ》しかし、B子は、被告人からの誘いを拒絶し、被告人宅から被告人に告げることなくA男の下へ戻った上、A男と行動を共にするようになり、被告人との接触を避けてきた。

【3】被告人の犯行動機 (7/8) 《ウ》被告人は、A男の養父ないし保証人として、A男の借金への対応に追われていたが、A男は、被告人に協力したり、感謝したりすることをせず、無責任かつ不誠実な態度をとり続けていた。

【4】本件の犯人像 (8/8) B子は、生前、在宅時も施錠し、限られた人間が訪れた際にしかドアを開けようとしなかったこと、本件の犯人が1歳のCちゃんを殺害しているのは口封じの可能性が高いこと、犯人が徹底的な証拠隠滅工作をしていることなどから、本件は被害者と近い関係にある者が行った可能性がある。 "

## (1) リーガル・マインド：判断基準（証明度）

《証明度の説明①》刑事裁判における有罪の認定に当たっては、合理的な疑いを差し挟む余地のない程度の立証が必要であり、状況証拠によって事実認定をすべき場合であっても、直接証拠によって事実認定をする場合と比べて立証の程度に差があるわけではありませんが、直接証拠がないのですから、状況証拠によって認められる間接事実中に、被告人が犯人でないとしたならば合理的に説明することができない（あるいは、少なくとも説明が極めて困難である）事実関係が含まれていることが必要です。〔最高裁判所第三小法廷判決平成22年4月27日刑集64巻3号233頁〕

※ 解釈： $p(\text{ある間接事実が存在する} | \text{被告人は犯人ではない}) \approx 0\% \Rightarrow \text{事後確率はほぼ}100\%$

※ 解釈：事前確率と $p(\text{ある間接事実が存在する} | \text{被告人は犯人である})$ は？

《証明度の説明②》刑事裁判における有罪認定に当たっては、合理的な疑いを差し挟む余地のない程度の立証が必要です。ここで、合理的な疑いを差し挟む余地がないというのは、被告人が犯人であることについて疑いを全く残さない場合をいうものではなく、抽象的な可能性としては被告人が犯人であることについて疑いをいれる余地があっても、健全な社会常識に照らして、その疑いに合理性がないと一般的に判断される場合には、有罪認定を可能とする趣旨です。そして、このことは、直接証拠によって事実認定をすべき場合と、状況証拠によって事実認定をすべき場合とで、何ら異なることはありません。〔最高裁判所第一小法廷判決平成19年10月16日刑集61巻7号677頁〕

質問：被告人の犯人性に対する心証を、0～100に数値化して回答してもらいます。0～100の内、どこにあたるか考えてください。（30秒）

0. 絶対に犯人ではありえない・・・50. 犯人とも犯人でないとも言える・・・100. 絶対に犯人である

質問：現時点で、改めて被告人の犯人性に対する心証を0～100に数値化して回答してもらいます。0～100のうち、どこにあたるか考えてください。（30秒）

0. 絶対に犯人ではありえない・・・50. 犯人とも犯人でないとも言える・・・100. 絶対に犯人である。

質問：あなたは、0～100の数値のうち、有罪と無罪の境界線はどこにあると考えていますか。0～100のうち、どこにあたるか考えてください。（30秒）

0. 絶対に犯人ではありえない・・・50. 犯人とも犯人でないとも言える・・・100. 絶対に犯人である。

質問：自分の有罪・無罪の結論についてどのくらい自信がありましたか。0～100のうち、どこにあたるか考えてください。（30秒）

0. 自信が全くない・・・50. 自信があるともないとも言える・・・100. 自信満点である

## (1) リーガル・マインド：悔悟の情と量刑判断——理性か情緒か？

### 《事案：殺人の共同正犯》

XとYは、遊ぶ金欲しさから、夜間にナイフで通行人を脅して財布を奪うことにした。そこで、XとY一人で歩いていたサラリーマンAに対し、それぞれナイフを眼前に突きつけて金を出せと脅し、財布を奪い取ろうとした。

しかし、Aが激しく抵抗し、大声で助けを呼んだため、とっさに、XとYはともにAを殺害することを決意し、両名ともにナイフをAに突き刺して怪我を負わせた上で、バラバラに逃走した。Aはナイフによる刺し傷のためしばらくして絶命した。

以上の経緯については、防犯カメラと目撃証人によって間違いないものとして確認されている。

### 《反省・悔悟の念が不存在の犯人X》

Xは、自らの犯行が発覚することを恐れてナイフを直ちに川に投げ捨て、Yには「絶対に余計なことをいうなよ」とメールしている。

警官に逮捕された後も、「Yに命令されて嫌々やった」「自分はナイフは使っていない」「殺す気はなかった」などと弁解し、謝罪する様子は一切みせていない。

それどころか、「Aから因縁をつけてきたので身を守るために抵抗しただけで正当防衛である」と述べている。

### 《反省・悔悟の念の存在する犯人Y》

Yは、自らの犯行の重大さに恐れおののき、逃走開始後直ちに119番で救急車を呼ぶとともに、警察に自首をし、犯行に至るまでの過程を全て正直に自白し、犯行に使ったナイフも警官に渡した。

Aの家族に対してできる限りの被害弁償を申し出ており、法廷でも自らが犯した罪に対して真摯な謝罪の弁を口にしている。

## (1) リーガル・マインド：悔悟の情と量刑判断——理性か情緒か？

### 《量刑判断》

**質問 1** 刑罰には様々なもの、そして様々な重さのものがありますが、それらを仮に考える「最も軽い刑」から考える「最も重い刑」まで、0から100までの一直線上に位置づけました。

**問** Xに対する処罰として、最もふさわしいとあなたが考えるものは、0～100の目盛の中では、どこに位置づけられますか？ 考えてください。（30秒）

0. 最も軽い刑・・・50. 真中位の刑・・・100. 最も重い刑

**問** Yに対する処罰として、最もふさわしいとあなたが考えるものは、0～100の目盛の中では、どこに位置づけられますか？ 考えてください。（30秒）

0. 最も軽い刑・・・50. 真中位の刑・・・100. 最も重い刑

**質問 2** 犯罪を犯した者の反省・悔悟については次のような考え方があります。

自らの犯罪について深く反省・悔悟している者は、矯正指導（更生と社会復帰のための刑務所等の矯正施設での再教育）を素直に受け入れるので、更生が期待できるのに対し、反省・悔悟していない者は、矯正指導を拒絶するため更生が期待しにくいです。

この事を念頭に置いて、次の質問にお答えください。

**問** Xに対する処罰として、最もふさわしいとあなたが考えるものは、0～100の目盛の中では、どこに位置づけられますか？ 考えてください。（30秒）

0. 最も軽い刑・・・50. 真中位の刑・・・100. 最も重い刑

**問** Yに対する処罰として、最もふさわしいとあなたが考えるものは、0～100の目盛の中では、どこに位置づけられますか？ 考えてください。（30秒）

0. 最も軽い刑・・・50. 真中位の刑・・・100. 最も重い刑

## (1) リーガル・マインド：悔悟の情と量刑判断——理性か情緒か？

### 《量刑判断》

**質問3** 犯罪を犯した者の反省・悔悟については次のような考え方があります。

自分の犯した罪について深く反省・悔悟をしている者に対しては、反省・悔悟をしていることを通じて、犯罪に対する社会の怒りがある程度は鎮静化するのに対し、反省・悔悟していない者に対してはそのような怒りの鎮静化が起きることはありません。

この事を念頭に置いて、次の質問にお答えください。

**問** Xに対する処罰として、最もふさわしいとあなたが考えるものは、0～100の目盛の中では、どこに位置づけられますか？ 考えてください。（30秒）

0. 最も軽い刑・・・50. 真中位の刑・・・100. 最も重い刑

**問** Yに対する処罰として、最もふさわしいとあなたが考えるものは、0～100の目盛の中では、どこに位置づけられますか？ 考えてください。（30秒）

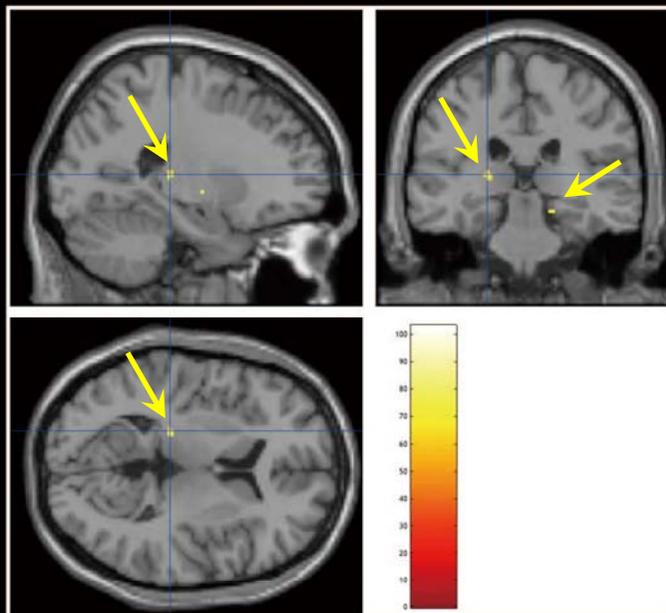
0. 最も軽い刑・・・50. 真中位の刑・・・100. 最も重い刑



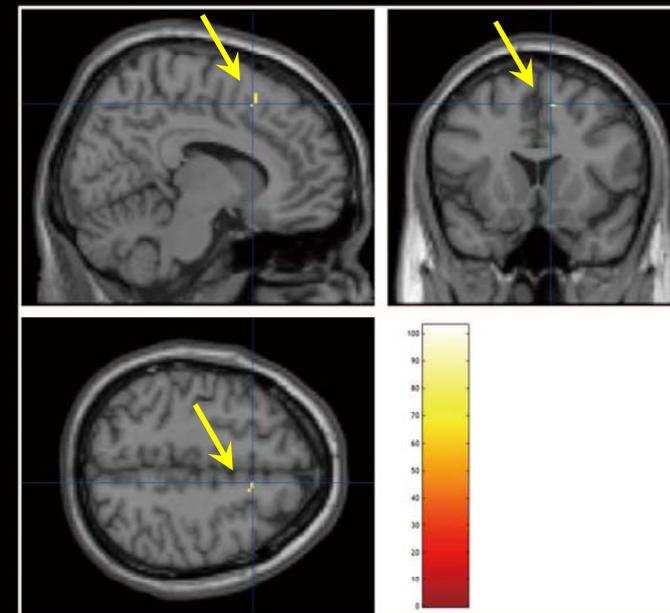
# (1) リーガル・マインド：悔悟の情と量刑判断——理性か情緒か？

《状況証拠からの心証形成事案：合理的疑いを差し挟む余地が無いか否かの判断》

## 結果



caudate (尾状核) [-22 -28 8]  
脳の学習と記憶システムの重要な部分  
を占めていると考えられている。



pMPFC [BA32/8/6; 10 16 50]

# (1) リーガル・マインド：悔悟の情と量刑判断——理性か情緒か？

## 1. 安静状態 (Resting State) の脳活動を差し引いて比較

⇒ 単なる文章読解や、非法的判断との区別が検証されない

## 2. 非法的文章読解時や、非法的判断時の状態の脳活動を差し引いて比較

**【質問例の①】** XとYは、ラウンジに座って話をしていました。コーヒーを飲みたくなったのでキッチンに行ってお湯を沸かしていたら、Aがやってきました。その内に、やかんのお湯がわきました。XとYはAも誘ってコーヒーを入れて飲みました。20s

《質問》 入れたてのコーヒーの温度は何度くらいだと思いますか。考えてください。15s

《指示》 スケールを動かしてください。20s

**【質問例の②】** XとYは、海に泳ぎに行きました。ひとしきり海水浴をしたので、砂浜に戻って休んでいました。暑いので、近くにある海の家に入ったところ、家族を連れてやってきた知り合いのAに会ったので、挨拶をしました。

《質問》 その時の海水浴場の外気の温度はどのくらいだと思いますか。

**【質問例の③】** XとYは、その日の最初の授業に出るために、銀杏並木を歩いていました。銀杏はすっかり色づいて、朝日に照らされています。向こうから、白い息をはきながら、友人のAが走ってきました。授業に遅れないよう、3人は急いで教室に向かいました。

《質問》 その時のキャンパスの気温はどのくらいだと思いますか。

**【質問例の④】** XとYは、一泊旅行に出かけました。泊ったのは、名湯と知られる温泉なので、ゆっくりと入ることにしました。温泉場はかなりの湯けむりでした。後から入ってきたAと温泉に入っている内に仲良くなったので、一緒に食事をするようになりました。

《質問》 3人が入った温泉の温度はどのくらいだと思いますか。

## 3. 判断の時間切れが見られた ⇒ 時間の延ばし、残りの時間を示す。

## (2) A I 裁判所への態度——信頼かアレルギーか？

### 《様々な支援の在り方》

A I 裁判所が：判例法令検索，証拠調べ，事実認定，法的推論，判決書作成，判決言渡し  
選択肢 実施／支援 実施／支援 実施／支援 実施／支援 実施／支援 実施／支援

- ⇒ 支援の程度にも，事実上裁判官が原則従うから，裁判官が参考にする程度まで
- ⇒ 民事事件，刑事事件，家事事件，少年事件などで国民の態度は異なりうる
- ⇒ ルーティン・ケース（簡単事件）とハード・ケース（価値判断必要）
- ⇒ 組み合わせの爆発
- ⇒ 実現可能性と，典型性などを考慮して3～6パターンに絞り込む
  - 例：刑事（交通切符程度，粗暴犯，ホワイトカラー犯罪？）
  - 例：民事（交通事故，離婚，債権回収？）

### 《国民の態度の調査》

実験計画法による複数シナリオのランダム配布

パーソナル質問とインパーソナル質問：他人への判決，自分の事件  
斉一性・平等性（裁判官によるブレや当たり外れが減る），迅速化，法律と先例に従った正確な判断化などと，「人間性喪失」(human value)などの，賛否の理由付けを解明する。

対比の参照点として，生身の人間裁判官の①誤判と②正しい判決

当事者・証人らの，老若男女，美醜，性格などによる影響の指摘？

## (2) A | 裁判所への態度——信頼かアレルギーか？

《事案①》本件の争点は、被告人がB子とCちゃんの2名を殺害した犯人であるかどうかです。本件では、被告人が犯行を全面的に否認し、被害者は死亡しており、目撃者もいないため、直接証拠はありません。もっとも、各事実を裏付ける後述の状況証拠【1】～【4】があります。

【1】被告人と被害者家族の関係 (1/8) 被告人は、A男が子供のころにその実母と婚姻し、養父としてA男を育てた。もっとも、A男の借金問題、女性問題等をきっかけに、本件当時は被害者家族と必ずしも良好な関係にあったとはいえない。

【2】被告人の犯行当日の行動 (2/8) 《ア》本件マンションの踊り場の灰皿内から、本件事件の翌日にたばこの吸い殻72本が採取されたが、その中に被告人が好んで吸っていた銘柄の吸い殻が1本あった。この吸い殻に付着していた唾液中の細胞のDNA型は、被告人の血液のDNA型と一致しており、このDNA型一致の出現頻度は1000万人に2人という極めて低いものである。

【2】被告人の犯行当日の行動 (3/8) 《イ》本件犯行時刻頃、被告人が当時使用していた自動車と同種・同色の自動車が、本件マンション付近に駐車されていた。被告人自身も、捜査段階においては、本件事件当日に自己の運転する自動車を同地点に駐車したことを認めていたが、その後、否認した。

【2】被告人の犯行当日の行動 (4/8) 《ウ》被告人は、本件当日の夕方、妻を迎えに行く約束をしていたが、その約束をたがえ、B子及びCちゃんが殺害された時間帯には、自らの携帯電話の電源も切っていた。

【3】被告人の犯行動機 (5/8) 《ア》本件犯行の半年前、B子及びCちゃんは、被告人宅で1か月間、被告人と同居したが、そのころ、被告人は、B子に対し、恋慕の情を抱いており、性交渉を迫る、抱き付く、キスをするなどの行為に及んだことがあった。

【3】被告人の犯行動機 (6/8) 《イ》しかし、B子は、被告人からの誘いを拒絶し、被告人宅から被告人に告げることなくA男の下へ戻った上、A男と行動を共にするようになり、被告人との接触を避けてきた。

【3】被告人の犯行動機 (7/8) 《ウ》被告人は、A男の養父ないし保証人として、A男の借金への対応に追われていたが、A男は、被告人に協力したり、感謝したりすることをせず、無責任かつ不誠実な態度をとり続けていた。

【4】本件の犯人像 (8/8) B子は、生前、在宅時も施錠し、限られた人間が訪れた際にしかドアを開けようとしなかったこと、本件の犯人が1歳のCちゃんを殺害しているのは口封じの可能性が高いこと、犯人が徹底的な証拠隠滅工作をしていることなどから、本件は被害者と近い関係にある者が行った可能性がある。”

## (1) リーガル・マインド：悔悟の情と量刑判断——理性か情緒か？

### 《事案：殺人の共同正犯》

XとYは、遊ぶ金欲しさから、夜間にナイフで通行人を脅して財布を奪うことにした。そこで、XとY一人で歩いていたサラリーマンAに対し、それぞれナイフを眼前に突きつけて金を出せと脅し、財布を奪い取ろうとした。

しかし、Aが激しく抵抗し、大声で助けを呼んだため、とっさに、XとYはともにAを殺害することを決意し、両名ともにナイフをAに突き刺して怪我を負わせた上で、バラバラに逃走した。Aはナイフによる刺し傷のためしばらくして絶命した。

以上の経緯については、防犯カメラと目撃証人によって間違いないものとして確認されている。

### 《反省・悔悟の念が不存在の犯人X》

Xは、自らの犯行が発覚することを恐れてナイフを直ちに川に投げ捨て、Yには「絶対に余計なことをいうなよ」とメールしている。

警官に逮捕された後も、「Yに命令されて嫌々やった」「自分はナイフは使っていない」「殺す気はなかった」などと弁解し、謝罪する様子は一切みせていない。

それどころか、「Aから因縁をつけてきたので身を守るために抵抗しただけで正当防衛である」と述べている。

### 《反省・悔悟の念の存在する犯人Y》

Yは、自らの犯行の重大さに恐れおののき、逃走開始後直ちに119番で救急車を呼ぶとともに、警察に自首をし、犯行に至るまでの過程を全て正直に自白し、犯行に使ったナイフも警官に渡した。

Aの家族に対してできる限りの被害弁償を申し出ており、法廷でも自らが犯した罪に対して真摯な謝罪の弁を口にしている。

# 研究手法

## (2) AIへの人々の受容：AI自動運転の社会調査

**参加者：**一般人

**事前準備：**AIによる裁判への人々の受容と比較するために，AIによる自動運転に対する人々の受容や，事故の際の責任意識（誰が賠償責任を負うべきか）などを調査する。

**実査：**ネット調査でフィールド実験

完全なAI自動運転車の場合（人間は運転を一切しない），人間の運転をAIがサポートする部分的AI自動運転，対照群として従来型の人間の運転する自動車と比較

**課題：**同じAIの利用でも，自動車と裁判とでは人々の受け容れ方が大きく異なる？

# 研究手法

## (2) AIへの人々の受容：AI自動運転の社会調査

**課題：**同じAIの利用でも、自動車と裁判とでは人々の受け入れ方が大きく異なる？

**[不完全AI自動運転車]** AI自動運転車とは、人間の運転者を適切にアシストする自動車です。周囲の状況を監視する各種のセンサーを備え、操作上の判断が求められる状況では、人工知能を搭載したコンピューターが適切な情報収集と状況判断を行い、適切な車の操作を運転者に知らせます。よって、人間はハンドルを握って操作したり、ブレーキ操作をする必要がある、一定程度の自動自動車です。

**[完全AI自動運転車]** 完全AI自動運転車とは、周囲の状況を監視する各種のセンサーを備え、操作上の判断が求められる状況では、人工知能を搭載したコンピューターが適切な情報収集と状況判断を行い、完全に全自動で適切に車を制御する「自動運転システム」により操作される車です。よって、人間がハンドルを握ったり、ブレーキを踏んだりする操作が全く必要のない全自動の自動車です。

## 研究手法

### (2) AIへの人々の受容：AI自動運転の社会調査

**課題：**同じAIの利用でも、自動車と裁判とでは人々の受け入れ方が大きく異なる？

**[不完全AI裁判所]** AI裁判所とは、人間の裁判官を適切にアシストするAIシステムです。証拠や証言を評価して事実認定をする裁判官をアシストするサブ・システムと、法規範を事実関係に当てはめたり法的推論をする裁判官をアシストするサブ・システムとを備え、人工知能を搭載したコンピューターが適切な事実認定と法的判断を行い、適切な判決内容の原案や候補を裁判官に知らせます。よって、人間の裁判官は自分で証拠評価をして事実認定をしたり、判決を起案する必要があります。AI裁判所は、一定程度のみの支援システムです。

**[完全AI裁判所]** AI裁判所とは、人間の裁判官に代わって裁判をするAIシステムです。証拠や証言を評価して事実認定をするサブ・システムと、法規範を事実関係に当てはめたり法的推論をするサブ・システムとを備え、人工知能を搭載したコンピューターが適切な事実認定と法的判断を行い、適切な判決内容を作り上げます。つまり、AI裁判所は完全に全自動で適切に裁判をします。よって、人間の裁判官が事実認定をしたり、法的判断をしたりする必要が全くない全自動のAI裁判所です。

# 研究手法

## (2) AIへの人々の受容：AI自動運転の社会調査

### 質問事項

- ・ 自動運転車・AI裁判所に対する態度
- ・ 自分で利用したいか？
- ・ 他人が利用することをどう思うか？
- ・ 人間社会にとって望ましいシステムだと思うか？

- 
- ・ 自動運転車・AI裁判所がミスや事故を起こした場合
    - ・ どう思うか？
    - ・ 誰が責任を取るべきか？
      - ・ 刑事責任
      - ・ 損害賠償責任

などなど。

科学研究費補助金・基盤研究(S)  
裁判過程における人工知能による高次推論支援

法的判断の構造とモデル化の探求  
—— AI は人々に受け容れられるか？ ——

**ご静聴を、ありがとうございました！**

基盤研究(A) 課題番号18H03612(2018年度～2022年度)

太田勝造, 加藤淳子, 佐藤健, 浅水屋剛, 今井耕介, 岡田謙介, 齋藤宙治, 米村滋人, 森大輔, 飯田高

(国立情報学研究所 October 02, 2018)