



# AI法廷の模擬裁判

{ mock\_trial\_in\_AI\_court }

## モノガタリと思考実験

東京大学法学部3年

岡本隼一



# 自己紹介 なにをやっているひとですか？

○ 東京大学法学部 3年生

○ 専門（にしたいこと） 分析哲学、法哲学

→ 数理的、分析的な思考を通じて法や政治体制の根本原理を考える学問

有名なところだと、「トロリー（トロッコ問題）」「カルネアデスの筏」「功利主義」「リバタリアニズム」など、聞いたことのある言葉かも

→ 現実にどんな政治が行われているか（実証）ではなく、どのような政治原理を採択すべきか（基礎）を扱う学問

○ 趣味 SFとTRPG

（後述）

# SFの魅力

- 本格的にSFを読み始めたのは、翻訳版『三体』から
- 偏愛作家はグレッグ・イーガンとジョン・バーレイ、小川一水、伊藤計劃
- 最近SFマガジンにエッセイが掲載されました。
- SFの魅力

Science Fiction（科学小説）であると同時に、Speculative Fiction（思弁小説）

→ 現実とは異なる世界設定が人類の社会と暮らしにどんな変化をもたらすのだろうか？

Cf. 『三体』：宇宙人との侵略に際し、全体主義に向かう社会

『闇の左手』：性別が変化する人類の暮らす惑星社会



三体/劉慈欣/2018

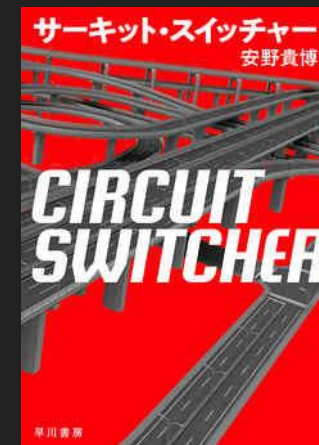
# モノガタリの持つ力

○ Speculative Fiction は、異なる世界の人類を「モノガタリ」の形で表現する

思考実験：例えば「トロリー問題」だけで政治原理や未来社会を考えることができるのだろうか？

5人を生かすか、1人を生かすか。数理的な実験を超えた、個々のモノガタリこそが真に役立つのではないか。

個人の物語や複数の視点を想像していくべきではないだろうか。



サーキット・スイッチャー  
/安野貴博/2022

# 文学作品の重層性

○ 文学作品とはポリフォニーである（ミハイル・バフチン）

登場人物の善人・悪人を議論するべきではない

作者の思想や人格の「表現」としてとらえるべきでもない

登場人物ひとりひとりがあたかも独立した人格として、個々の思想を主張することで、登場人物相互の対話が生まれ、多次的・多元的な表現ができるのが文学作品の魅力

トロリー問題の事例

自動運転車の事故で亡くなった1人の被害者遺族は、5人の生存者は、システムの開発者はなにを考えるのだろうか？

# AI法廷を劇にした理由

- 『AI法廷の模擬裁判』は、AI裁判官が実用化された近未来の法廷をイメージする裁判劇形式で行った
- プレゼンテーション手段として劇を採用したのは2つの理由から

多視点から未来を想像してもらうため

→ たとえば証人はどんな心情だったのか。被告人は「機械に裁かれるとき」どういう思いだったのか。弁護士と検事は裁判長に対してなにを訴えようとしたのか。

未来を想像し、よりよい社会を選択するためには、モノガタリ的センスが不可欠

未来のあり方を自分事としてとらえてもらうため

→ 登場人物に感情移入し、じっさいにその世界に暮らす自分を想像してもらうことで、より問題意識を受け取りやすくなる。

# まとめ：モノガタリ的な思考実験

- 技術や情報を学ぶことは非常に重要。現状の技術で何ができて、何ができないかを知ることは、誤った情報に惑わされないためにかかせない。
- しかし、未来についてより深く考えて正しい選択をくだすためには、情報と同時に、物語の力が必要である。
- 科学の成果を、非専門家にとどけ、自分事として考えてもらうサイエンスコミュニケーションとしてもSF的アイデアは役に立つ。

未来を想像し、選択するための文学的センスを養わねばならない



# AI法廷の模擬裁判

{ mock\_trial\_in\_AI\_court }

AIと自由民主主義の未来-

東京大学法学部3年 中島胡太郎



# 企画趣旨（はじめに）

## ➤ 岡本の立場

データやプログラムに基づいた機械的な正確さのほう  
が正統性として高く評価できる

→AIによる裁判を支持する

## ➤ 中島の立場

自由民主主義における「参加」の意義が司法の出力に  
与える正統化作用は極めて重要とみるべき

→AIによる裁判を支持しない

# 三権分立と自由民主主義

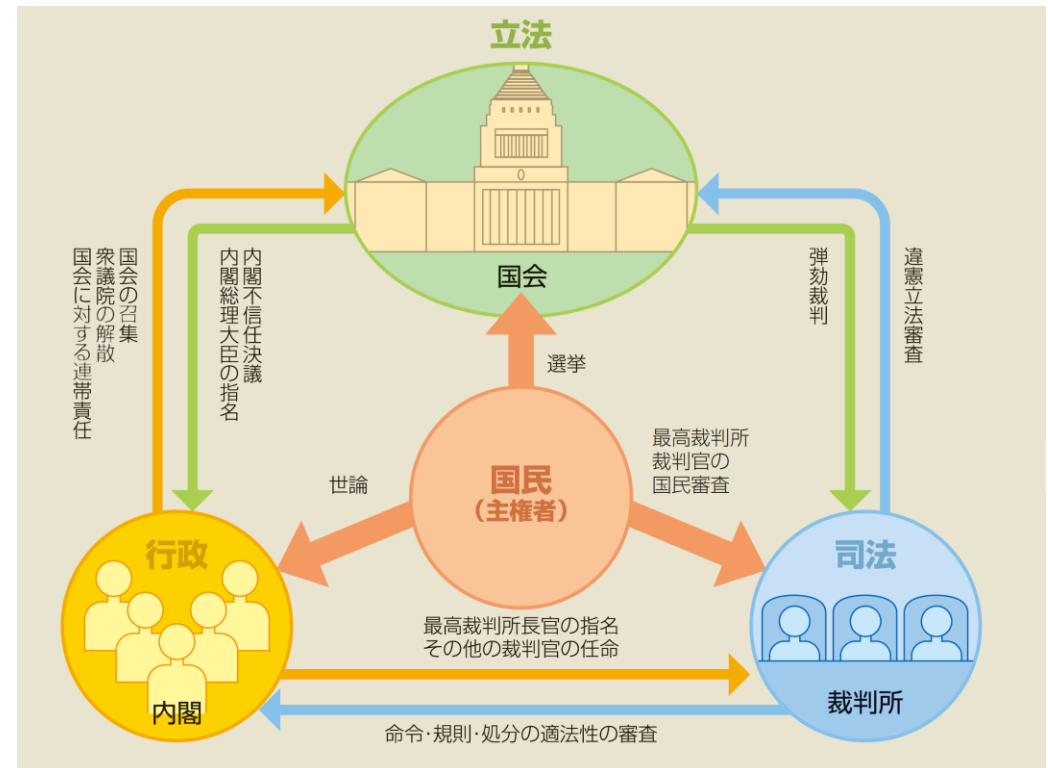
## ・ 制度的内容

お互いを監視し、拮抗させることで特定の権力の肥大化や暴走を抑える制度

## ・ 目的

国民の自由と権利を守る

⇒ 「自由主義的」な制度である





# 政治参加と自由民主主義

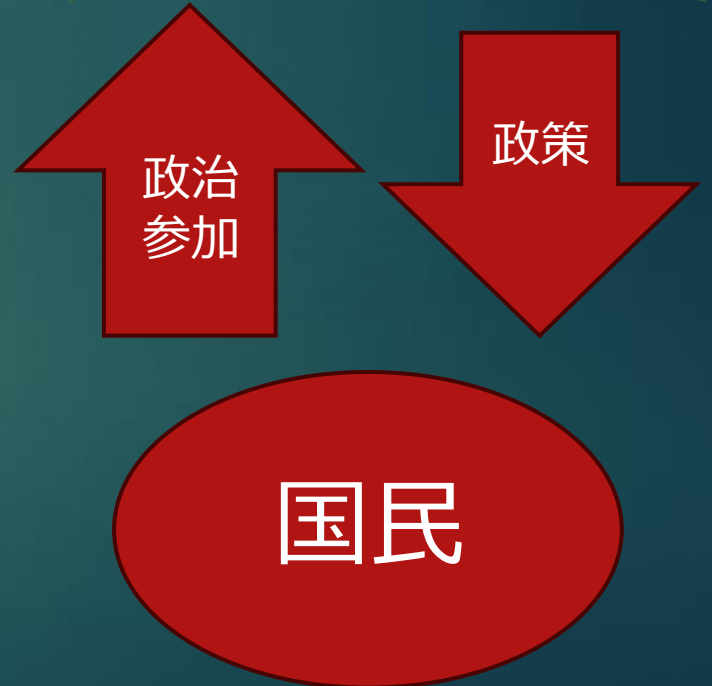


## ➤ 「参加」の意義

代表を選挙するという意味にとどまらず、我々がそのプロセスに参加したこと自体が重要性を持つ

→参加の事実が政治的決定に正統性を与える

私見：民主主義は我々の参加があってこそ成り立つ、人間的営みではないか？



# 「政治的疎外」と政治

## ➤ 市民の意識の問題

市民が「政治が我々の方向を向いている」という意識を失った結果としての欧州

：ポピュリズムや排外主義の台頭

→政治システムから人間を排除するAIの導入の結末は？



# AIの導入における課題

## ▶ 使用するAIという課題

ジェンダーや人種上の問題

：差別的な出力をする可能性がある（実例有）

→ 少数派を保護するという理念に相反するおそれ

（ただ、人間にも当てはまるという反論は成り立つ）

# 自由民主主義はAIを受容できるか？（企画趣旨）

## ➤ 個人的見解

多様な市民を抱える自由で民主的な社会を支える根幹と価値は制度に「人間が介在している」ことである

→AIによる制度は、それを突き崩す危険性を有する

## ➤ 想定される反論

司法は「民主」より「自由」を守る制度である

→むしろAIという、「多数派」から離れた主体がふさわしい



# 企画アンケートから考える

## ▶ 判決への支持

観客の「民意」と判決の一致・AI裁判官の受容

→少なくとも本裁判のAI裁判官への支持はあると考えられる

## ▶ 今後への示唆（？）

AI裁判官への一般的な支持は現状ないのではないか

—Q5,6の結果：人間の役割への高評価



# AI法廷の模擬裁判

{ mock\_trial\_in\_AI\_court }

ご清聴ありがとうございました



## シンポジウム (AI 模擬裁判) レジューメ

東京大学法学部 3 年 中島胡太郎

### ➤ はじめに (企画趣旨)

AI は司法の役割を担うことができるか？

- AI は法的な議論を行うことができるのか？ (AI 側の問題)
- 市民は AI による司法を受け入れることができるのか？ (市民・社会側の問題)
  - ・ 本発表では後者の議論を行う

AI による裁判のイメージ

— 今までにない迅速さと正確さを兼ね備えた技術によって正義が実現されるという希望

— 人間が機械の奴隷になる、というディストピア

企画趣旨：このような未来への展望をきわめてわかりやすい「模擬裁判」の形で実演し、来場者一人一人にその意義を訴えること

岡本の立場：人間社会の自律性という手続き的正統性よりも、データやプログラムに基づいた機械的な正確さのほうが正統性として高く評価できる

→ AI による裁判を支持する立場

中島の立場：自由民主主義における「参加」の意義を考えたとき、それが司法の出力に与える正統化作用は極めて重要とみるべき

→ AI による裁判を支持しない立場

### ➤ 三権分立と自由民主主義<sup>1</sup>

お互いを監視することで相互に暴走を抑えるための制度

：権力を拮抗させることで特定の権力の強大化を防ぐことになる

なぜ「権力を分散・拮抗」させる必要があるのか？

⇒ **国民の権利と自由を守る**ため

ゆえに、権力分立は、「**自由主義的**」な政治制度である、と言われる

もちろん、民主政の原理も権力分立と親和的である

：民主政の究極の目的は人権の保障であるから

現代の三権分立の特徴：理論上は違憲審査により司法が立法・行政を統制しうる

→ 司法が立法に与える影響は強いとみることができる

---

<sup>1</sup> 芦部信喜『憲法 第七版』、岩波書店、2019年。



## ➤ 政治参加と自由民主主義

現代議会制における「自由」とは「我々による統治」としての「政治的自由」である<sup>2</sup>

→この自由を制度的には唯一行使できるのが選挙である

(たとえ虚構に近くても)「国民は議会を通じて意思を表明できる」という事実(=政治の主体は我々である、という意識)が議会、ひいてはそこから選出される内閣の決定に政治的な正統性を与える

⇒民主主義は極めて人間的な営みではないか？

## ➤ 政治的疎外と自由民主主義

逆に、政治に参加している・政治が自らの方を向いているという意識が薄れてくるとどうなるか？

AI法廷の文脈で言うと、「AIが司法、ひいては三権を運営するようになった結果は？」

:「政治的疎外」という議論である程度その未来を推測することができる

主権者であるはずの市民が政治をコントロールできなくなったという事実と、政府がそれに対し有効な対策を講じていない(政治不信)という感覚

—欧米におけるポピュリズム台頭の共通要因として指摘されている<sup>3</sup>

・「政治的疎外」の結果(例)

参加しているという感覚を失った市民の行動

—移民への批判(政治不信の原因を移民などに押し付ける)

—カリスマ的リーダーへの期待<sup>4</sup>

その具体例としての反EU・反移民(欧州)やトランプの当選(アメリカ)

→このような批判の受け皿として欧米で「**ポピュリズム**」が台頭することとなった

AIによる三権の運営は、市民から「主権者として政治を動かしている」という感覚を奪うのではないか？

---

<sup>2</sup> ケルゼン, ハンス『民主主義の本質と価値 他一篇』、長尾龍一・植田俊太郎訳、岩波文庫、2015年。

<sup>3</sup> 谷口将紀「ポピュリズムを招く新しい「政治的疎外」の時代」NIRA オピニオンペーパーNo.40、2018年。

<sup>4</sup> 谷口将紀・水島治郎編著『ポピュリズムの本質 「政治的疎外」を克服できるか』中央公論新社、2018年。

## ➤ 導入における課題

### ・ AI による出力における課題

AI が人種・ジェンダー的に極めて不適切な出力を行ったという事例は多い<sup>5</sup>

→少数者保護・法の番人としての「司法」にとって不適格ではないか  
ただし、同様の批判は人間にも当てはまる可能性も十分にある

例：政治家による女性蔑視的発言

一方、人間の場合は責任の主体が明確であるということも指摘できる

自動運転の時に騒がれた「責任の主体」という議論の再燃・大規模化の可能性がある

## ➤ 自由民主主義は AI を受容できるか？

### ・ 個人的見解

現代（日本）の自由民主主義社会には多様な思想を抱く市民が多く存在している

But 政策（政治システムからの出力）は全ての市民を満足させることはできない  
重要なことは、（間接的、虚構的であっても）「私たち」が参加し、決定した結果が政策である、ということ（司法においては「裁判員制度」）

→我々が政治的決定に「介在」しているからこそ、その出力をある程度受け入れられる

つまり、市民が政治的決定を正統なものと考えられるようになる

：政治学的に言えば、「民主的正統性」

「AI による司法」は、その正統性を破壊してしまう（=自由民主主義の根底を揺るがす）ことになるのではないか？

### ・ 想定される反論

司法はむしろ「民主的」な「多数者の専制」から市民の権利と自由を守る砦である

→むしろ多数派から距離を取った AI こそ、少数の「エリート」として適切ではないか？

この見解に対する個人的な反論

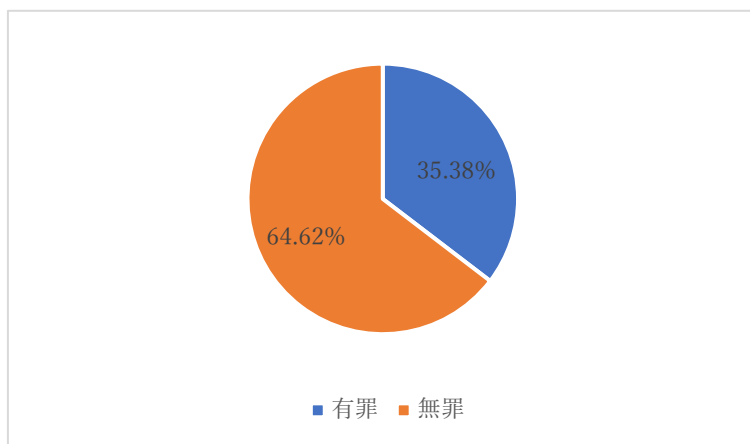
その AI が開発の過程で多数派のバイアスを受ける可能性がある（当然これは人間の裁判官にも当てはまるという譲歩をする必要があるが）

---

<sup>5</sup> 横山美和「AI と人種・ジェンダー問題についてアメリカから学べること」産学官連携ジャーナル、2021 年 <[https://www.jst.go.jp/tt/journal/journal\\_contents/2021/12/2112-02\\_article.html#refer4](https://www.jst.go.jp/tt/journal/journal_contents/2021/12/2112-02_article.html#refer4)>（2023 年 8 月 28 日参照）。

➤ 自由民主主義はAIを受容できるか？—企画アンケートから考える

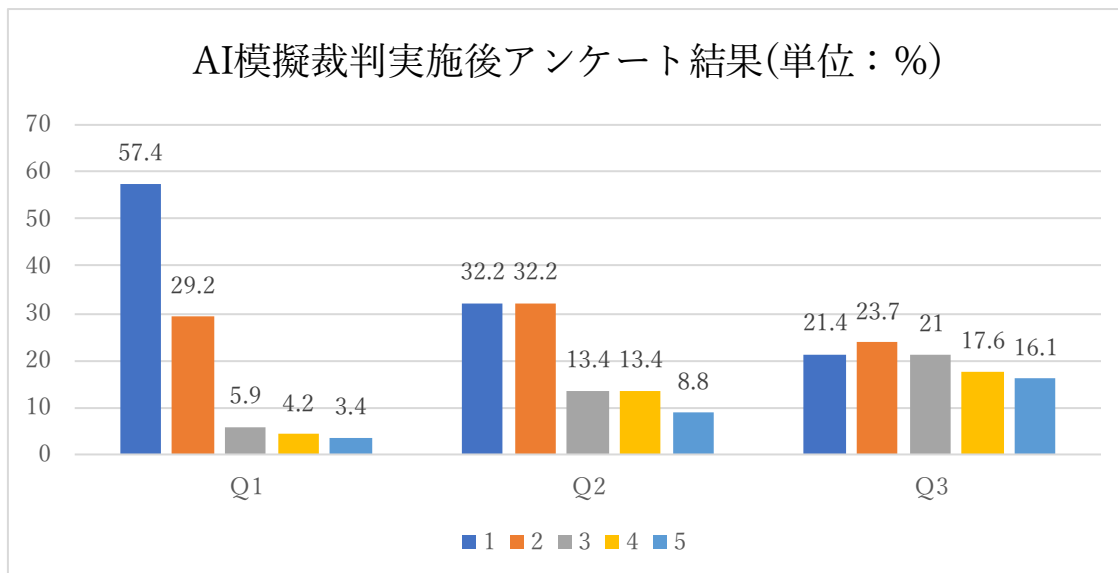
判決言い渡し前投票：被告人は有罪ですか、無罪ですか。あなたの判断を教えてください。



Q1：今回の模擬裁判のAIの判決について（1：妥当である～5：妥当でない）

Q2：AIが裁判官を務めたことについて（1：違和感を覚えなかった～5：覚えた）

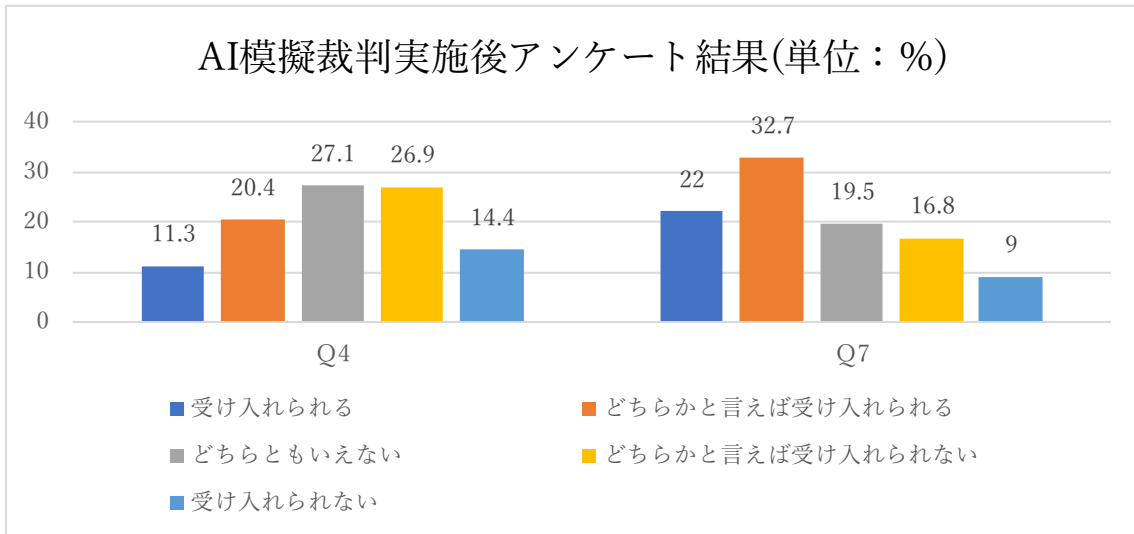
Q3：もしあなたがこの事件の被告人であったならば、「AIに裁かれること」に納得できるか（1：納得できる～5：できない）



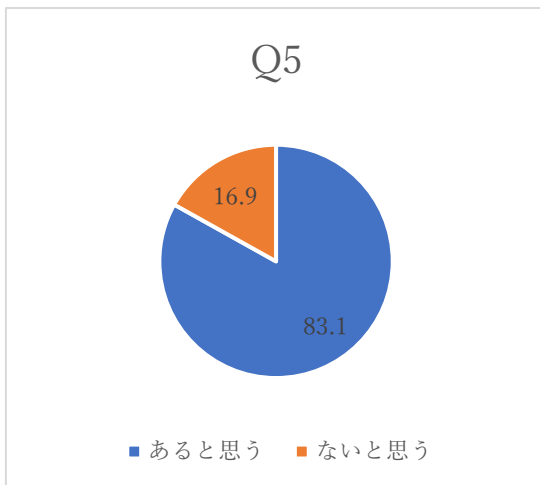


Q4：AI が膨大に学習した先例に基づいて、人の直感とは異なる判決を下した場合、あなたはそれを受け入れることはできるか

Q7：「AI が人を裁く時代」を、あなたは受け入れることはできるか



Q5：裁判において、機械に代替できない人の役割はあると思うか（単位：%）



Q6：AI による司法参加はどこまで許容されるか	%
AI による自動判決	7.3
AI による判決（裁判官の確認必須）	32.9
AI の判決予測による人間の支援	32.9
AI による人間の先例整理の提示	24.7
AI の参入は望ましくない	2.1

・検討と示唆

Q1 と Q2 を一体として考えれば、少なくとも当判決についてみれば、AI による判決は支持を得たとみることができる

： 妥当な判決を出した AI への違和感はない、とみることができる

≠ AI による裁判への支持がある

： Q3 から Q6 では、むしろ人間の裁判官への支持が見て取れる

→ AI のみの司法には支持が集まらないのではないか？

－ 「人間の役割」 への高評価・ AI による自動判決への不支持

⇒ 未だ、人間による司法への支持は存在している

➤ 最後に

本発表が、AI に司法を任せるとはできるのか、さらに広げて AI は自由民主主義社会にどのような影響を与えるのか、そして一市民としてどう行動すべきかを考える機会になれば幸いである。

参考資料（取材）

「東大五月祭『AI 裁判官』による模擬裁判で再認識 『人間裁判官』であるメリットと課題」(DG LAB Haus)

[〈https://media.dglab.com/2023/06/05-aijudge-01/〉](https://media.dglab.com/2023/06/05-aijudge-01/)

「ChatGPT 裁判官、殺人事件で「無罪」判決 東大生の模擬裁判、近未来を体感」(弁護士ドットコム)

[〈https://www.bengo4.com/c\\_18/n\\_16023/〉](https://www.bengo4.com/c_18/n_16023/)

「ChatGPT が裁判官？ AI に裁かれる未来 受け入れますか」(NHK)

[〈https://www3.nhk.or.jp/news/special/sci\\_cul/2023/05/special/ai-judgement/〉](https://www3.nhk.or.jp/news/special/sci_cul/2023/05/special/ai-judgement/)