

ZW

# 博士論文

政策形成における生成 AI と人間の関係性：  
公共交通政策を事例として

Generative AI and Human Relationships in Policy Formation:  
A Case Study of Public Transport Policy

氏名：永田 右京

指導教員：〇〇 教授

〇〇大学 大学院 〇〇研究科  
〇〇専攻 博士後期課程

2026 年 〇月 〇日 提出

# 謝辞

本論文の作成にあたり、多大なるご指導とご支援を賜りました〇〇大学教授〇〇先生に、心より感謝申し上げます。先生の温かいご指導と厳しいご助言なしには、本論文を完成させることはできませんでした。

また、〇〇研究科の諸先生方には、研究全般にわたり貴重なご助言をいただきました。深く感謝申し上げます。

〇〇研究所の皆様には、研究の場を提供していただき、有意義な議論を重ねることができました。心より御礼申し上げます。

共同研究者の皆様、特に〇〇様には、多くのデータ収集や分析においてご協力いただきました。深く感謝いたします。

日々の研究生生活を共にした研究室の仲間たちには、多くの刺激と励ましをいただきました。皆様との議論は、本研究の発展に不可欠でした。

最後に、私を支え続けてくれた家族に深く感謝いたします。皆様の理解と励ましがあったからこそ、本研究を完遂することができました。

なお、本研究の一部は、日本学術振興会科学研究費補助金（課題番号：〇〇〇〇〇〇〇〇）の助成を受けたものです。

2026年〇月〇日  
永田 右京

# 要旨

本研究は、公共交通政策を舞台に、生成 AI と人間の協調的関係性を探求したものである。

第 1 章では、研究の背景、問題の所在、研究目的を論じた。公共交通政策における「連携・共創」の実装ギャップが、制度的欠陥だけでなく、人間の認知バイアスに起因する可能性を指摘した。

第 2 章では、Human-AI Policy、協調的ガバナンス論、認知バイアスと意思決定、ZK-SNARKs と政策評価に関する先行研究をレビューし、理論的空白を特定した。

第 3 章では、日本の公共交通政策の変遷と制度設計の現状を整理し、Japan MaaS 38 プロジェクトの実証分析を通じて実装ギャップの実態を明らかにした。

第 4 章では、協調ロボット制御モデルを用いた計算論的分析により、現状維持バイアスの閾値効果、確証バイアスの逆説的效果、狭い視野の一貫した負の影響を解明した。

第 5 章では、生成 AI を人間の「執政の創造性」を補完する「杖」として位置づけ、ZK-SNARKs の概念を援用した政策評価システムを提案した。Constitutional AI、市民討議、LLM as a Judge を組み合わせた三層アーキテクチャを設計した。

第 6 章では、認知バイアスへの対処戦略と生成 AI を組み込んだ制度設計への示唆を導出した。

第 7 章では、研究の総括と、都市計画への展開について論じた。

本研究は、生成 AI と人間の関係性を理論的・実証的に探求し、より良い政策形成のための指針を提供した。

**キーワード：**生成 AI、政策形成、認知バイアス、公共交通政策、ZK-SNARKs、制度設計

# Abstract

This study explores the collaborative relationship between generative AI and humans, using public transport policy as an empirical context.

Chapter 1 discusses the research background, problem statement, and objectives. It points out that the implementation gap in "collaboration and co-creation" in public transport policy may be attributable not only to institutional deficiencies but also to human cognitive biases.

Chapter 2 reviews prior research on Human-AI Policy, collaborative governance theory, cognitive bias and decision-making, and ZK-SNARKs in policy evaluation, identifying theoretical gaps.

Chapter 3 examines the evolution of Japanese public transport policy and current institutional design, revealing the reality of implementation gaps through empirical analysis of 38 Japan MaaS projects.

Chapter 4 employs computational analysis using cooperative robot control models to uncover the threshold effect of status quo bias, the paradoxical effect of confirmation bias, and the consistently negative impact of narrow framing.

Chapter 5 positions generative AI as a "staff" that complements human "executive creativity," proposing a policy evaluation system utilizing ZK-SNARKs concepts. It designs a three-layer architecture combining Constitutional AI, citizen deliberation, and LLM as a Judge.

Chapter 6 derives strategies for addressing cognitive biases and implications for institutional design incorporating generative AI.

Chapter 7 summarizes the research and discusses future directions for urban planning applications.

This study provides theoretically and empirically grounded guidelines for better policy formation by exploring the relationship between generative AI and humans.

**Keywords:** Generative AI, Policy Formation, Cognitive Bias, Public Transport Policy, ZK-SNARKs, Institutional Design

# Contents

謝辞	ii
要旨	iii
Abstract	iv
発表論文リスト	xi
<b>1 序論</b>	<b>1</b>
1.1 研究の背景	1
1.1.1 生成 AI と人間の関係性という問い	1
1.1.2 公共交通政策における「連携・共創」の潮流	1
1.2 問題の所在	1
1.2.1 協調的ガバナンスの実装ギャップ	1
1.2.2 人間の認知限界と政策形成	1
1.2.3 生成 AI の可能性と限界	2
1.3 研究目的と意義	2
1.4 論文の構成	2
<b>2 先行研究のレビュー：Human-AI Policy の議論</b>	<b>3</b>
2.1 はじめに	3
2.2 「執政の創造性」とは何か	3
2.2.1 定義	3
2.2.2 政策過程モデルにおける価値判断	3
2.2.3 ウィキッド・プロブレムとの関連	4
2.3 ルーマン理論による基礎づけ	4
2.3.1 なぜルーマンか	4
2.3.2 コミュニケーションの3段階：情報・伝達・理解	5
2.3.3 構造的カップリング：本研究の核心概念	5
2.3.4 「理解の創造性」の定義	5
2.3.5 AI の原理的限界	6
2.4 Human-AI Policy：政策形成における AI と人間の関係性	6
2.4.1 AI 政策論の展開	6
2.4.2 Human-Centered AI の理念	6
2.4.3 AI の「杖」としての位置づけ	7
2.5 協調制御理論と社会システムへの応用	7
2.5.1 協調制御理論の基礎	7
2.5.2 社会システムへの応用可能性	7
2.6 協調的ガバナンス論	7
2.6.1 協調的ガバナンスの定義	7
2.6.2 公共交通における連携・共創	7
2.6.3 実装ギャップの指摘	8

2.7	認知バイアスと意思決定	8
2.7.1	行動経済学の基礎概念	8
2.7.2	政策プロセスにおける認知バイアス	8
2.8	ZK-SNARKs と政策評価	8
2.8.1	ZK-SNARKs の基本概念	8
2.8.2	政策評価への応用可能性	8
2.9	小括：理論的空白の特定	9
<b>3</b>	<b>舞台としての公共交通政策：現状と課題</b>	<b>10</b>
3.1	はじめに	10
3.1.1	本章の目的	10
3.1.2	MaaS の勃興	10
3.1.3	先行研究	10
3.1.4	リサーチクエスションと仮説	11
3.2	MaaS の概念と展開	11
3.2.1	フィンランドにおける MaaS の勃興	11
3.2.2	MaaS に必要な統治構造変革	11
3.2.3	MaaS とインターネットの類似性	12
3.2.4	日本版 MaaS の定義と展開	12
3.3	法体系と交通まちづくり	12
3.3.1	交通政策基本法	12
3.3.2	交通政策基本計画	13
3.3.3	地域公共交通活性化・再生に関する法律	13
3.3.4	交通まちづくりの定義とその親和性	13
3.4	市民参加の捉え方	13
3.4.1	公共交通政策における市民参加に関する議論	13
3.4.2	公共サービスとしての公共交通	14
3.4.3	望ましい市民参加のスタンス	14
3.5	実装ギャップの実証分析：Japan MaaS	14
3.5.1	分析手法	14
3.5.2	分析対象	14
3.5.3	分析結果	15
3.5.4	事例分析	15
3.5.5	考察	15
3.5.6	日本版 MaaS は交通まちづくりの一類型として捉えられるか	16
3.6	小括：公共交通政策を「実験場」として位置づける理由	16
<b>4</b>	<b>認知バイアスの政策協調への影響：計算論的分析</b>	<b>17</b>
4.1	研究背景と問題設定	17
4.1.1	研究の背景	17
4.1.2	政策の文脈と変遷	17
4.1.3	実装ギャップの実証的証拠	17
4.1.4	研究目的とアプローチ	18
4.2	理論的枠組み	18
4.2.1	協調ガバナンス理論	18
4.2.2	公共サービス提供における共創	18
4.2.3	協調的政策実装における認知バイアス	19
4.2.4	協調制御理論とマルチエージェント調整	19
4.3	計算論的モデリング・フレームワーク	19
4.3.1	類推的マッピング：正当化と限界	19
4.3.2	モデル構造	20
4.3.3	運動学モデル	20

4.3.4	制御目標と調整メカニズム	20
4.4	認知バイアスの統合	21
4.4.1	現状維持バイアスのモデル化	21
4.4.2	確証バイアスのモデル化	21
4.4.3	狭い視野のモデル化	21
4.5	シミュレーション実験	21
4.5.1	実験設計	21
4.5.2	評価指標	22
4.6	結果	22
4.6.1	確証バイアスの効果	22
4.6.2	現状維持バイアスの閾値効果	22
4.6.3	狭い視野の一貫した負の影響	22
4.7	統計的サマリー	22
4.8	制度的設計への示唆	22
4.9	小括：認知バイアスによる協調失敗のメカニズム	23
<b>5</b>	<b>生成 AI と人間の関係性：ZK-SNARKs 型政策評価システム</b>	<b>24</b>
5.1	はじめに：ウィキッド・プロブレムと政策評価の課題	24
5.1.1	ウィキッド・プロブレムの台頭	24
5.1.2	コミュニケーションにおける二重の困難	24
5.2	ZK-SNARKs 概念の援用	24
5.2.1	ZK-SNARKs とは何か	24
5.2.2	ZK-SNARKs が実用的な理由	25
5.2.3	秘密を守りながら専門性を証明する	25
5.3	ZK-SNARKs 型政策評価フレームワークの提案	25
5.3.1	フレームワークの目的	25
5.3.2	入札型秘匿証明の構築	25
5.4	LLM as a Judge による実装	26
5.4.1	LLM as a Judge と ZK-SNARKs の親和性	26
5.4.2	LLM as a Judge の可能性	26
5.4.3	LLM as a Judge の限界と課題	26
5.5	LLM as a Judge の限界を超えるための人間支援機能	27
5.5.1	Constitutional AI と市民討議	27
5.5.2	計画・推論分離型システム	27
5.5.3	控訴プロセスによる人間介入	27
5.5.4	決定論的運用	27
5.6	ZK-SNARKs 型政策評価のアーキテクチャ	28
5.6.1	三層アーキテクチャ	28
5.6.2	生成 AI の「杖」としての位置づけ	28
5.6.3	限界と課題	28
5.7	小括	28
<b>6</b>	<b>制度設計への示唆</b>	<b>30</b>
6.1	理論的含意	30
6.1.1	第 4 章のシミュレーション結果からの設計原則	30
6.1.2	第 5 章の ZK-SNARKs システムからの制度的含意	30
6.2	各バイアスへの対処戦略	30
6.2.1	現状維持バイアス：小さな変化の積み重ね	30
6.2.2	確証バイアス：多様な視点の構造的導入	30
6.2.3	狭い視野：全体目標の可視化	30
6.3	生成 AI を組み込んだ制度設計	31
6.3.1	政治的-行政的インターフェース	31



6.3.2	行政的-事業的インターフェース	31
6.3.3	ZK-SNARKs 型システムの制度的位置づけ	31
6.4	「連携・共創」の再設計	31
6.4.1	三層制度設計（S-T1-T2-O）の再考	31
6.4.2	生成 AI を「杖」として活用するガバナンス	31
6.5	小括	32
<b>7</b>	<b>結論：都市計画への展開</b>	<b>33</b>
7.1	研究の総括	33
7.1.1	研究目的の達成	33
7.1.2	核となる主張	33
7.2	理論的貢献	33
7.2.1	認知バイアスの計算論的分析手法の政策科学への導入	33
7.2.2	ZK-SNARKs 概念の政策評価への応用	33
7.2.3	生成 AI の「杖」としての理論的位置づけ	34
7.3	実践的貢献	34
7.3.1	制度設計への提言	34
7.3.2	ZK-SNARKs 型政策評価システムの設計指針	34
7.4	今後の課題：都市計画を舞台にした実証	34
7.4.1	より複雑な政策領域への適用	34
7.4.2	ZK-SNARKs システムの社会実装	34
7.4.3	生成 AI と人間の協調的関係性の継続的検証	34
7.5	結び	34
<b>A</b>	<b>付録</b>	<b>37</b>
A.1	協調ロボット制御モデルの数式展開	37
A.1.1	運動学モデル	37
A.1.2	協調係数の計算	37
A.2	Japan MaaS プロジェクト分析の詳細	37
A.2.1	分析対象プロジェクト一覧	37
A.2.2	評価指標のコーディング基準	37
A.3	ZK-SNARKs 型政策評価システムの実装詳細	38
A.3.1	システム構成	38
A.3.2	技術的仕様	38

# List of Figures

# List of Tables

3.1	Japan MaaS 36 プロジェクトの評価指標分析 . . . . .	15
4.1	Japan MaaS 38 プロジェクトの実装ギャップ分析 . . . . .	18
4.2	シンボル-概念対応表 . . . . .	20
4.3	現状維持バイアスの閾値効果 . . . . .	22
4.4	認知バイアス効果の統計的サマリー . . . . .	23
6.1	認知バイアスと生成 AI を考慮した三層制度設計 . . . . .	31
A.1	評価指標のコーディング基準 . . . . .	38
A.2	技術的仕様 . . . . .	38

# 発表論文リスト

## 査読付き論文

1. Nagata, U. (2025). "「日本版 MaaS」は「交通まちづくり」の一類型として捉えられるか？目標とガバナンスについての一考察", 『日本評価学会誌』, Vol.XX, No.X, pp.XX-XX.
2. Nagata, U. (2025). "Re-designing Collaboration and Co-creation in Regional Public Transport Policy: Integrated Approach to Cognitive Biases and Institutional Coordination", 土木学会論文集, Vol.XX, No.X, pp.XX-XX.
3. Nagata, U. (2025). "ZK-SNARKs 概念を援用した、ウィキッド・プロブレムに対応する政策評価の仕組み", 『公共政策学会誌』, Vol.XX, No.X, pp.XX-XX.

## 学会発表

1. 永田右京 (2024). "地域公共交通計画の実装ギャップに関する分析", 日本公共政策学会 2024 年度大会.
2. 永田右京 (2024). "認知バイアスが政策協調に与える影響の計算論的分析", 土木学会第 XX 回年次学術講演会.
3. 永田右京 (2025). "生成 AI と人間の協調的関係性の設計", 日本行政学会 2025 年度大会.

# Chapter 1

## 序論

### 1.1 研究の背景

#### 1.1.1 生成 AI と人間の関係性という問い

近年、生成 AI（Generative AI）の急速な発展により、政策形成プロセスにおける AI 活用の可能性が広く議論されている。ChatGPT をはじめとする大規模言語モデル（Large Language Models, LLM）は、テキスト生成、要約、分析などのタスクにおいて人間と同等、あるいはそれ以上の性能を示す場面も増えている。しかし、AI が「人間を代替」するのではなく、「人間を補完」する関係性をどのように設計すべきかという問いは、依然として未解決のままである。

本研究では、生成 AI を人間の「執政の創造性」<sup>1</sup>を補完する「杖」として位置づけ、両者の協調的關係性のあり方を探求する。

#### 1.1.2 公共交通政策における「連携・共創」の潮流

日本の公共交通政策においては、2002 年の規制緩和以降、「連携」と「共創」が重要な政策概念として位置づけられてきた。特に 2021 年の「ポストコロナ時代の地域交通の共創に関する検討会」（国土交通省）においては、交通事業者による地域活性化、異業種との協働、コミュニティ参画という三つの次元での「共創」が提唱されている。

しかし、こうした政策的意図にもかかわらず、実装段階では多くの課題が指摘されている。

### 1.2 問題の所在

#### 1.2.1 協調的ガバナンスの実装ギャップ

公共交通政策における「連携・共創」は、制度的には整備されつつあるものの、実践レベルでは大きなギャップが存在する。例えば、Japan MaaS の 38 プロジェクトを分析した先行研究<sup>2</sup>によれば、92%のプロジェクトが事業指標のみを重視し、社会的影響やアクセシビリティ改善を評価指標に組み込んだのは 29%に留まる。さらに、市民参加の仕組みを設けたプロジェクトはわずか 5%であった。

この実装ギャップは、単なる制度的欠陥だけでなく、人間の認知特性に起因する可能性がある。

#### 1.2.2 人間の認知限界と政策形成

人間の意思決定は、認知バイアス（cognitive biases）の影響を強く受けることが知られている [1]。特に政策形成プロセスでは、変化への抵抗や現状の維持選好を示す現状維持バイアス（Status Quo Bias）、自己の信念を確認する情報の優先的選択という確証バイアス（Confirmation Bias）、局所的最適化への固執や全体最適の見落としという狭い視野（Narrow Framing）が重要な影響を及ぼ

---

<sup>1</sup> 価値判断、コミュニケーション、新たな規範の創造といった人間固有の能力

<sup>2</sup> 第 3 章で詳述

す。これらのバイアスは、ステークホルダー間の協調を阻害し、政策の実装ギャップを生む一因となっている可能性がある。

### 1.2.3 生成 AI の可能性と限界

生成 AI は、膨大な情報の処理、パターン認識、予測を行うことで、EBPM（証拠に基づく政策形成）を支援する強力なツールとなり得る。しかし、AI には本質的な限界も存在する。第一に、規範的判断・価値創造の不在である。AI は何が社会にとって「善い」のかを判断できない。第二に、文脈理解の困難性である。学習データの範囲外の「未知の状況」への適応には限界がある。第三に、「創造性」の源泉の欠如である。人間的な自発的な揺らぎやアナログな現実世界の機微を再現できない。

これらの限界を踏まえつつ、AI を「杖」として活用する関係性をどのように設計すべきかが問われている。

## 1.3 研究目的と意義

本研究の目的は、以下の三点である。第一に、公共交通政策における実装ギャップの要因として、人間の認知バイアスの影響を計算論的に解明する。第二に、生成 AI と人間の協調的關係性を具体化するシステムとして、ZK-SNARKs 型政策評価システムの可能性を探る。第三に、認知バイアスと生成 AI を考慮した制度設計への示唆を導出する。

本研究の意義は、生成 AI と人間の關係性を理論的・実証的に探求し、より良い政策形成のための指針を提供することにある。

## 1.4 論文の構成

本論文は 7 章から構成される。第 2 章では先行研究のレビューを行い、Human-AI Policy の議論を中心に整理する。第 3 章では舞台としての公共交通政策の現状と課題を論じる。第 4 章では認知バイアスの政策協調への影響について計算論的分析を行う。第 5 章では生成 AI と人間の關係性として、ZK-SNARKs 型政策評価システムを提案する。第 6 章では制度設計への示唆を導出する。第 7 章では結論として、都市計画への展開について論じる。

## Chapter 2

# 先行研究のレビュー：Human-AI Policy の議論

### 2.1 はじめに

本章では、本研究の理論的基盤となる先行研究をレビューする。特に、「執政の創造性」とルーマン理論、生成 AI と政策形成、人間-AI 協調の理論、協調的ガバナンス論、認知バイアスと意思決定、ZK-SNARKs と政策評価の六つの領域を中心に整理し、理論的空白を特定する。

### 2.2 「執政の創造性」とは何か

#### 2.2.1 定義

「執政の創造性」とは、以下のように定義される：

**執政の創造性：**社会の構成員同士のコミュニケーションを前提に、価値判断を基盤として、新たなガバナンスの範囲を生成・配分していく能力。

この定義は、三つの要素から構成される。第一に、政策は、単一の主体による合理的決定ではなく、複数の主体間のコミュニケーションを通じて形成される。第二に、政策決定は統計的平均や客観的指標への還元が不可能な、主体的な価値判断を要する。第三に、政策は固定された枠組みではなく、状況に応じて新たな範囲を創造し続ける動的なプロセスである。

この概念は、カンギレム（Canguilhem）の「規範の創造」概念 [2] および井庭の創造システム理論 [3] を統合したものである。カンギレムは、「正常」とは統計的平均ではなく、「生の独自の規範性」を肯定することであると論じた。健康とは客観的指標ではなく、主体によって体験される価値である。同様に、政策における「正しさ」も、統計的効率性ではなく、主体による価値判断に基づく。

井庭の創造システム理論では、コミュニケーションの「わかり合えなさ」から出発し、「発見」を要素とするシステム上で創造が起こるとされる。本研究は、この創造のプロセスを政策の文脈に適用し、「執政の創造性」として定式化する。

#### 2.2.2 政策過程モデルにおける価値判断

「執政の創造性」の重要性は、既存の政策過程モデルを分析することで裏付けられる。主要な政策過程モデルを見てみよう。

#### 政策の窓モデル

キングドン [4][5] による政策の窓モデルは、政策過程に「問題」「政策代替案」「政治」という3つの独立した流れが存在し、それらが特定の時点で合流したときに政策決定が生じることを示す。

キングドンは、この3つの流れが合流する瞬間を「政策の窓」が開くと表現する。窓は一時的にしか開かず、逃したら再び開くまで長く待たなければならない。

このモデルが示すのは、政策決定は論理的な手順ではなく、問題と解決案と政治的機会の偶然の出会いによって生じるということである。この「偶然の出会い」を見極め、窓が開いた瞬間に行動するには、人間の価値判断と政治的勘が不可欠である。AIは過去のデータから傾向を分析することはできても、「今がチャンスだ」と感じ取り、その瞬間に動くという状況認識と行動のタイミングについては、人間の判断に依存せざるを得ない。

### 唱導連携モデル

唱導連携モデル（Advocacy Coalition Framework）は、サバティエ [6] らによって開発されたモデルであり、特定の政策分野において、共通の信念体系を共有する主体たちが連携（コアリション）を形成し、政策を唱導するプロセスを分析する。

このモデルの特徴は、政策を単なる利害調整ではなく、信念体系の競合として捉える点にある。各連携は、(1) 深層核信念（基本的価値観）、(2) 政策核信念（具体的政策目標）、(3) 二次的側面（手段的判断）という階層的な信念体系を持つ。

ここで重要なのは、どの信念を優先すべきか、どの連携の主張を採用すべきかという判断が、常に価値判断を伴うということである。統計的データや客観的分析だけでは、信念の競合を解決できない。どの価値を重視するかという判断が必要であり、これは人間の創造的適応能力に依存する。

### 村松モデル

村松 [7] モデルは、日本の政策過程を「与党・官僚・利益団体」の三者関係として分析する枠組みである。このモデルでは、政策決定はこれら三者の交渉と取引を通じて行われる。

真淵による修正版では、野党勢力や新規参入者の役割も考慮される。既存の産業組織が維持しようとする価値と、新規参入者が求める変革の価値が衝突する中で、政治家はバランスを取る必要がある。

このモデルが示すのもまた、利害の調整と価値の優先順位付けが政治的本質であり、それは計算可能な最適化問題ではないということである。三者間の「妥当な落としどころ」を見出すには、人間同士の交渉と相互調整が不可欠である。

## 2.2.3 ウィキッド・プロブレムとの関連

公共政策の多くは、Rittel & Webber [8][9] の「ウィキッド・プロブレム（Wicked Problems）」の性質を帯びている。ウィキッド・プロブレムとは、決定的な問題定義がなく、問題そのものが何であるかについて利害関係者の間で合意がない問題である。停止ルールもなく、いつ問題が「解決」されたのかを客観的に判断できない。解は「良い/悪い」ではなく「より良い/より悪い」であり、最適解は存在せず、複数の利害のバランスを取るしかない。また、解を適用した結果、予期せぬ副作用が生じる可能性があり、社会実験はやり直しが効かない。さらに、すべてのウィキッド・プロブレムは本質的にユニークであり、過去の経験から単純に適用できない。

ウィキッド・プロブレムへの対応は、終わりのない創造的プロセスである。そして、「終わりがないからこそ、政治がその不快さを受け入れる必要がある」のである。この終わりのない創造的プロセスこそが、人間による政策の固有の意味であり、AIには代替不可能な領域である。

## 2.3 ルーマン理論による基礎づけ

### 2.3.1 なぜルーマンか

コミュニケーションを理論的基軸とする場合、ユルゲン・ハーバーマスの「コミュニケーション的行為理論」[10]も選択肢として存在する。しかし、本研究はルーマン [11] の立場を採用する。その理由を説明する。

ハーバーマスは、コミュニケーションを「了解志向的」な行為として捉える。コミュニケーションの理想的状態では、参加者が互いに「了解」に到達し、合意を形成することが期待される。こ



の立場からは、コミュニケーションの「失敗」や「誤解」は、理想的言語状況からの逸脱として問題視される。

これに対し、ルーマンはコミュニケーションを「了解の否定」を内包するプロセスとして捉える。ルーマンによれば、コミュニケーションは常に「理解」と「誤解」の双方を可能性として含んでおり、「わかり合えなさ」こそがコミュニケーションの本質的な特徴である。

本研究がルーマンの立場を採用する理由は、以下の二点にある。

第一に、「わかり合えなさ」を「創造の源泉」として位置づける視点が必要である。ハーバーマスの「理想的了解」を前提とすれば、AIも「十分に良い」情報処理を行うことで「機能的な了解」に貢献できると主張しうる。しかし、ルーマン的な視点からは、「わかり合えなさ」から生じる価値判断と意味構成のプロセスこそが創造の核心であり、これをAIは代替できない。

第二に、政策的決定における「価値競合」の不可避性も重要である。公共政策は、ハーバーマスが想定する「理想的言語状況」において理性の力だけで解決可能な問題ではなく、複数の正当な価値が競合するウィキッド・プロブレムとしての性質を持つ。このような状況では、「了解」への到達よりも、「わかり合えなさ」を前提とした創造的適応こそが求められる。

### 2.3.2 コミュニケーションの3段階：情報・伝達・理解

ルーマン [11][12] によれば、コミュニケーションは「情報 (Information)・伝達 (Mitteilung)・理解 (Verstehen)」という3つの選択過程から構成される。

**情報 (Information)** 多数の可能性の地平からの一つを選択であり、何を語るかを選択である。

**伝達 (Mitteilung)** 多数の伝達可能性からの選択であり、いかに語るかを選択である。

**理解 (Verstehen)** 多数の理解可能性からの選択であり、いかに受け止めるかを選択である。

ルーマンは、「三つの選択のはたらきのすべてが総合されるときにはじめてコミュニケーションというものが成り立つ」と強調する。この3層構造は、コミュニケーションが単純な情報伝達ではなく、各段階で選択と解釈が介在する複雑な過程であることを示している。

### 2.3.3 構造的カップリング：本研究の核心概念

ここで重要になるのが、ルーマンにおける「構造的カップリング (strukturelle Kopplung)」 [11][13] の概念である。これは、作動上は完全に独立 (閉鎖) している複数のシステムが、互いに不可欠な環境として影響し合う関係を指す。

ルーマン理論において、最も重要な構造的カップリングは、「心的システム (意識)」と「社会システム (コミュニケーション)」の関係である。心的システムは思考し、社会システムはコミュニケーションする——それぞれ別の作動を行う。しかし、心的システムがなければコミュニケーションは発生しない。意識はコミュニケーションに対して、刺激や誘発を与えたり、あるいは邪魔をしたりすることができる。ただし、意識がコミュニケーションを「因果的に決定」するわけではない。意識はあくまで環境として、コミュニケーション・システムに「刺激」を与え、システム側がそれを独自の論理で処理する<sup>1</sup>。

### 2.3.4 「理解の創造性」の定義

以上の理論的整理を踏まえ、本研究で論じる「理解の創造性」を以下のように定義する：

**理解の創造性：** 心的システムと社会システムの構造的カップリングにおいて、心的システムから社会システムへの刺激として提供される、価値判断を伴う意味構成のプロセス。

<sup>1</sup> ここで注意すべきは、ルーマン理論において「認知」は心的システム内部に所在し、社会的システムに「分布」しているわけではないことである。構造的カップリングは、社会的コミュニケーションが心的システムを刺激することを可能にするが、認知プロセスそのものは心的システムのオートポイエシスとして完結する。

この定義は、ルーマンの厳密な意味での「理解」（コミュニケーション接続）とは区別される。本研究が着目するのは、コミュニケーション接続を駆動する「価値判断の源泉」としての心的システムの役割である。

この観点から、「わかり合えなさ」は単なる誤解ではなく、構造的カップリングにおいて各心的システムが独自の価値判断に基づいて意味を構成する結果として生じる、コミュニケーションの本質的な特徴として理解される。

この価値判断を伴う意味構成のプロセスこそが、人間固有の創造性であり、本研究が「執政の創造性」として定式化する対象である。

### 2.3.5 AIの原理的境界

ルーマンの理論において、機械はオートポイエシス・システム（生命・意識・社会）とは区別される「非ポイエティック」な存在とされている [11]。AI システムも基本的には計算機プログラム（機械）上で動作するシステムである。AI——その最先端の形態を含めて——は心的システムに該当しない。

**オートポイエティックでない：**たとえ自己学習能力を持つ AI であっても、その「学習」は人間が設計したアルゴリズムと訓練データに依存している。外部からの入力なしに自律的に自身の「思考」を産出し続ける閉鎖的なネットワークを持たない。

**自己言及的な意味処理を行わない：**AI の確率的出力は、文脈に応じて「もっともらしい」次のトークンを選択するが、この選択プロセスは自己言及的ではない。AI は「この意味を選択したこと自体」を次の処理の地平として開くことはなく、単に統計的パターンに基づいて出力を生成する。

**価値判断を伴わない：**心的システムからの刺激は「何が重要か」「何を優先すべきか」という規範的判断を前提とするが、AI の出力は統計的パターンに基づいており、独自の規範的判断を伴わない。以上の議論から、公共政策における AI の原理的境界が明らかになった。AI は「情報」の選択や「伝達」の補助には機能しうるが、「理解」の創造的プロセスには原理的に参加できない。AI は心的システムを持たないため、価値判断を伴う意味構成を行えないのである。

## 2.4 Human-AI Policy：政策形成における AI と人間の関係性

### 2.4.1 AI 政策論の展開

公共政策における AI 活用に関する議論は、2010 年代後半から急速に発展してきた。初期の議論は、AI による行政サービスの効率化や自動化に焦点が置かれていたが、近年では AI と人間の関係性そのものが問いの中心となっている [14]。

この転換の背景には、生成 AI（ChatGPT、Claude、Gemini 等）の登場がある。これらの技術は、従来の AI（分類・予測）とは異なり、創発的なテキスト生成能力を持つ。この能力は、政策文書の作成、選択肢の生成、市民との対話など、政策形成の核となるプロセスに直接関与しうる。

### 2.4.2 Human-Centered AI の理念

Shneiderman (2022) [14] は、Human-Centered AI (HCAI) の理念として、以下の 2 軸マトリクスを提示している：

**高自動化・低制御** AI が自律的に判断し、人間は結果を受け入れるのみ

**高自動化・高制御** AI が提案を行い、人間が最終判断を下す

**低自動化・高制御** 人間が主導し、AI が補助的な役割を果たす

**低自動化・低制御** 人間も AI も十分に機能しない状態

本研究が着目するのは「高自動化・高制御」の領域である。この領域では、AI の計算能力と人間の規範判断力が相互に補完し合う。

### 2.4.3 AIの「杖」としての位置づけ

生成AIを「杖（Aaron's rod）」として位置づける視点は、AIが人間を代替するのではなく、人間の能力を補完・増幅する道具として活用する考え方である。

この視点からは、いくつかの設計原則が導かれる。第一に、AIは人間の最終判断を前提とすることである。第二に、AIの限界を明示的に理解することである。第三に、人間-AI協調のプロセスを透明化することである。第四に、AI自体のバイアスに対処することである。第五に、説明責任は常に人間が負うことである。

この「杖」としての位置づけは、AIを「執政の創造性」を支援する道具として捉え直す視点を提供する。

生成AIは、いくつかの領域では人間を補完し得る。データの処理・分析、選択肢の生成・提示、文書作成の効率化、多様な視点の提示、そして認知バイアスの指摘（「悪魔の代理人」機能）などがその例である。

一方で、いくつかの領域では人間の役割が不可欠である。規範的判断（何が「善い」か）、文脈に応じた柔軟な対応、新たな価値の創造、政治的なアカウンタビリティ、そして最終的な責任の所在などがその例である。

## 2.5 協調制御理論と社会システムへの応用

### 2.5.1 協調制御理論の基礎

協調制御理論（Cooperative Control Theory）は、複数の自律エージェントが協調して共通の目標を達成するための制御手法を研究する分野である [15]。

この理論は、いくつかの重要な特徴を持つ。分散的な意思決定を行うこと、局所的な情報に基づく協調を行うこと、そして全体的な目標の達成を目指すことである。

### 2.5.2 社会システムへの応用可能性

協調制御理論は、社会システムの分析にも応用可能である。特に、複数のステークホルダーが関与する政策ネットワークにおいて、各主体が自律的に行動しながら全体としての政策目標を達成するプロセスをモデル化できる。

本研究では、協調ロボット制御モデルを用いて、政策ネットワークにおけるステークホルダー間の協調を分析する（第4章で詳述）。

## 2.6 協調的ガバナンス論

### 2.6.1 協調的ガバナンスの定義

Ansell and Gash (2008) [16] は、協調的ガバナンスを以下のように定義している：

「一つまたは複数の公共機関が、非政府のステークホルダーを、合意形成志向で審議的な集団的意思決定プロセスに直接関与させる統治のあり方」

### 2.6.2 公共交通における連携・共創

日本の公共交通政策においては、「連携」と「共創」が重要な概念として位置づけられている [17]。Kato et al. (2009) は、コミュニティ参加型地域公共交通の成功条件として、関係ステークホルダー間での認識と責任分担の共有、各ステークホルダーが参加から利益を得られること、ステークホルダーを調整するキーパーソンの存在、そしてステークホルダーの努力が利用促進・価値向上につながることを指摘している。

### 2.6.3 実装ギャップの指摘

しかし、こうした理論的条件にもかかわらず、実践レベルでは多くの課題が指摘されている。Emerson et al. (2012) [18] は、協調的ガバナンスが直面する課題として、高い取引コスト、最小公約数的な解決策への収束、そして組織された利益による捕捉を指摘している。

## 2.7 認知バイアスと意思決定

### 2.7.1 行動経済学の基礎概念

Kahneman (2011) [1] は、人間の思考を「システム 1（速い思考）」と「システム 2（遅い思考）」に分類し、認知バイアスがシステム 1 の特性に起因することを示した。

### 2.7.2 政策プロセスにおける認知バイアス

政策形成において特に重要な認知バイアスとして、以下の三つを取り上げる：

#### 現状維持バイアス (Status Quo Bias)

Samuelson and Zeckhauser (1988) [19] によって提唱された概念で、変化よりも現状を維持することを好む傾向を指す。

#### 確証バイアス (Confirmation Bias)

既存の信念や仮説を支持する情報を優先的に探し、反証する情報を無視・軽視する傾向 [20]。

#### 狭い視野 (Narrow Framing)

問題を孤立して考え、より広い文脈や長期的な影響を考慮しない傾向 [1]。

## 2.8 ZK-SNARKs と政策評価

### 2.8.1 ZK-SNARKs の基本概念

ZK-SNARKs (Zero-Knowledge Succinct Non-interactive Arguments of Knowledge) は、暗号技術の一種であり、秘密情報を公開することなく、その情報の正しさを証明する技術である [21]。ZK-SNARKs は以下の 4 つの特性を持つ：

**Zero-Knowledge** 証明を通して元の秘密情報が一切漏洩しない

**Succinct** 証明サイズが常に数百バイト程度と一定

**Non-interactive** 証明者から検証者への 1 回の送信で証明完了

**Arguments of Knowledge** 真の知識を所有している必要があり偽造不可能

### 2.8.2 政策評価への応用可能性

ZK-SNARKs の概念を政策評価に応用することで、「秘密を守りながら専門性を証明する」仕組みが実現可能になる。例えば、企業が自社の技術情報を公開せずに、政策課題への貢献可能性を証明できる。

本研究では、ZK-SNARKs の概念を援用した政策評価システムを LLM as a Judge と組み合わせで提案する（第 5 章で詳述）。

## 2.9 小括：理論的空白の特定

先行研究のレビューから、いくつかの理論的空白が明らかになった。第一に、「執政の創造性」の理論化である。ルーマンの社会システム理論を用いて「執政の創造性」を基礎づけ、AIと人間の役割分担を明確にした研究は限定的である。第二に、認知バイアスと政策協調の接続である。協調的ガバナンスの失敗要因として認知バイアスに着目した研究は限定的である。第三に、計算論的分析手法の欠如である。政策協調プロセスを計算論的にモデル化した研究は少ない。第四に、ZK-SNARKs 概念の政策評価への応用である。暗号技術の概念を政策評価に応用した試みは先駆的である。

本研究は、これらの空白を埋めることを目指す。

「日本版 MaaS」は「交通まちづくり」の一類型として捉えられるか？

目標とガバナンスについての一考察

永田 右京<sup>1)</sup>

<sup>1</sup> 株式会社 MaaS Creative(神奈川県横浜市西区平沼 1-40-1 嶋森ビル 8 階)

岩手県立大学 総合政策研究科 (岩手県滝沢市菓子 152-52)

E-Mail:s18568un@gmail.com

地域公共交通政策のうち、Mobility as a Service(MaaS) と呼ばれる分野は、この数年で急速に勃興してきた。それ以前から、「交通に関連する地域の課題への対応をベースにして、市民と行政が協働して進めるまちづくり」として、「交通まちづくり」という思想体系が存在、発展してきた。

本稿ではこうした MaaS の政策進捗において、交通に関わる理念法である「交通政策基本法」を参照した上で、「交通まちづくり」のそれと類似するその理想が実現されているか、また公共交通事業者の一方性を取り除けているかについて、2020 年度の「日本版 MaaS」の全事例を精査した。その結果、交通政策において理念的に重視される機能の維持管理に関して、多くの計画では関知できていない実態が浮き彫りとなった。さらに市民参加による事業構築や精査を事業の内部に組み込んでいる事例もごく少数であった。この結果から、日本版 MaaS は構想段階では「交通まちづくり」と捉えられるが、実装段階ではそうではない、と結論づけた。

**Key Words:** *Regional Public Transport Revitalization Law, Mobility as a Service, Japanese MaaS, Citizen participation*

## 2.10 1. はじめに

近年、日本において「Mobility as a Service」の勃興が著しい。2019 年度より始まった「日本版 MaaS 推進事業」は、瞬間に全国各地に広がり、現在までに 100 事業以上のプロジェクトを送り出している。本稿ではこうしたプロジェクトがどのような体制の下に進められ、またどのような検証体制が敷かれているかについて、国土交通省の公開している資料を基として実証的に考察する。なお、本稿における「日本版 MaaS」は、国土交通省による定義に基づく。

### 2.10.1 (1) 本稿の前提

前提として、本稿のスタンスを明らかにする。本稿でいう「公共交通」とは、他人の運転の提供によって、移動できるようになるサービスであり、これはタクシー、ライドシェアなどを含む。

また本稿における分析モデルとして、政策過程論を採用する。政策過程は、目的を設定した上で現状を理解し、そのギャップを課題として認識した上でそれを解決するための方策を列举し、実施する政策を決定するという一連の流れである。この中で、どのように目標をたて、またどのように目標を設定しているかが、本稿における大まかな分析対象である。



本稿で言う「交通まちづくり」の概念は、太田<sup>1)</sup>の述べるような「交通に関連する地域の課題への対応をベースにして、市民と行政が協働して進めるまちづくり」である。原田編<sup>2)</sup>の述べるところの定義である「まちづくりの目的に貢献する交通計画」とは異なり、「市民の参加と協働により発展的に進める活動プロセス」である。この中では、「交通計画における住民、市民の参加」「都市計画、都市づくりとの連携」の2点が重要視されている。

## 2.10.2 (2) 先行研究

日本版 MaaS について、その特徴を述べた文献はいくつかある。

まず注目されるのは、特に過疎地域や地方中核市における交通が抱える課題を解決するためのアプローチとしての見立てである。魏<sup>3)</sup>は、日本の MaaS の注目点が「地方型 MaaS」である点に言及したうえで、日本の MaaS の特徴について以下のように述べている。

「地方型」MaaS は、地域の資源や特性を最大限活かしながらより満足度の高いモビリティサービスを提供することであり、地域にある交通モード間の連携強化により、その場その場で発生した交通需要（生活交通需要、観光交通需要）を上手く汲み上げ、集約することで実現できるものである。

また藤本<sup>4)</sup>では、欧州における MaaS が「都市型 MaaS」と「地方型 MaaS」に分化されており、また日本版 MaaS が「地方型」を含む5類型に分割されている点に触れたうえで、地方型 MaaS を検討するポイントを5つに整理している。

MaaS 自体の実装にかかわる研究も進んでいる。後藤<sup>5)</sup>は、MaaS を旅行業の一類型にとらえ、企画旅行業務を適用する可能性やデータ供給の課題について論じている。

既存の公共交通運営体制に関する議論は、多くの場合「どのような取り組み」があったか、また「どのような評価が可能か」という2つに大別される。前者の場合、生活バスよっかいちについての事例報告を行った福本・加藤<sup>6)</sup>や、多くの日本モビリティマネジメント会議の報告に代表される。また「どのような評価が可能か」については、アクセシビリティ評価の地域への適用について事例を分析した喜多<sup>7)</sup>などがあげられる。公共交通に関する責任分担や意思決定組織についての検討も複数ある。

本稿で前提とした「交通まちづくり」の考えでは、交通政策における市民参加と地域戦略との連携が重視されていると考えられる。また後述するものであるが、地域公共交通に関する法体系においても、交通政策の戦略性を重視している。ただ少なくとも「日本版 MaaS」において、法体系からの課題意識を導入した包括的な研究、また交通事業者としてではない形で計画の要件定義から携わる市民参画の有無、に関する実証的な研究は見当たらなかった。

このように、上記に示したような実証的な研究を実施し、結果として「交通まちづくり」の思想を援用できないか検討する作業には新規性があり、結果によっては思想によるバックアップをもとに日本における MaaS の広がり支援できる可能性がある。

## 2.10.3 (3) 本稿のクエスチョンと仮説

本稿は、いわゆる「日本版 MaaS」を「交通まちづくり」実践の一類型と捉え、以下の2つの問いについて検討し、結果としてこの捉え方が適切かについて確認するものである。

- ・「日本版 MaaS」の政策進捗において目標設定がどこに置かれているか
- ・立場を問わない市民による批判回路がどのように整備されているのか

本分析における仮説としてはそれぞれ、

- ・目標設定は、MaaS 事業の事業性に関して設定されるが、それ以外の地域に関する効果については設定されない
- ・市民による批判回路は、協議会に閉じず、パブリックコメント、市民代表参加、座談会、ワークショップなどの機会を通じて設定されていると設定している。

#### 2.10.4 (4) 本稿の流れ

日本版 MaaS について議論する前にまず、2 章及び 3 章では、日本および勃興場所であるフィンランドにおける MaaS の展開について簡単にまとめる。次に 4 章では、公共交通に関わる法体系についてまとめ、前提となる理論について記述する。上記 2 点については前者の問いに対応する。5 章にて「市民参加」に関する理論体系を確認する。これは後者の問いに対応する。これを踏まえて、6 章では事業 2 年目で件数が一番多く、「地域特有の課題の解決に寄与することが見込まれる、地域特性に応じた MaaS のモデルとなり得る」と紹介され、本研究の方向性と合致する 2020 年の「日本版 MaaS」において行われた事業について総攬し、全ての事業に対して前述の問いを確認する。最後に、現在の政策に提供される示唆をまとめる。

### 2.11 2. フィンランドにおける MaaS の勃興：提案初期論文の確認と法改正

#### 2.11.1 (1) 初期に提案された MaaS の特徴

まず MaaS は、Hekkila<sup>8)</sup> にて初めて学術的な表舞台に立った、公共交通の新しい提供概念である。氏が所属する大学へ依頼されたヘルシンキ市の交通課題解決のために、修士論文で書き上げたものである。これによると、Nokia を主軸としたフィンランドの通信政策やその事業に対する理解の蓄積が、ほぼそのまま交通政策に導入されていることがわかる。

この論文では、論文冒頭のサマリーにおいてその目的を明確に示している。具体的には、公共交通の利用促進が叫ばれ、都市構造への影響が避けられない中で、既存の交通サービス目標は「課題に十分にこたえられていない」と批判している。そのうえで、「公共交通セクターの構造転換が必要である」として「多様で魅力的な交通サービスの便利な提供」として MaaS を定義している。

具体的な MaaS の構造として、「交通サービスプロバイダーから交通サービスを買ひ、それをユーザーへ提供する」主体を「モビリティオペレーター」と定義している。結果として、交通サービスを従来提供していた事業者と利用者の間に、もう一つの主体が挟まることが特徴である。

#### 2.11.2 (2) 公共政策上の MaaS に必要な統治構造変革

この MaaS コンセプトを実現するうえで、政策上で何が必要であるのかについて、Heikkilä は表-1 のように、7 つの政策目標を提示している。

表-1 MaaS を進めるのに必要な要件

(Hekkila<sup>8)</sup> より筆者翻訳・作成)

||@ >p() \* 0.0408 >p() \* 0.1916 >p() \* 0.2496@

項目名	具体的な内容
1 企業, 当局, 機関, ユーザーなど, すべてのステークホルダーの協力体制の調整	エコシステムの充足のためだけでなく, すべてのステークホルダーのモチベーションを高めるために, すべてのステークホルダーの要求を開発プロセスに取り入れる協力関係を構築せねばならない。
2 サービス・エコシステムの前条件を満たすための法律・規制の改正	法律, 規制改正の必要性がある。現行の法律において, サービス・エコシステムの開発や新企業の市場参入を妨げている法規制を取り除く必要がある。自治体は, 利害関係者の提案を集めて上位のガバナンスに伝達することで, 仲介者として機能できる。
3 共通のルールと適切な規制の構築, およびその遵守状況の監視	新興市場が機能し, すべてのステークホルダーの利益を確保するためには, 適切な規制が必要である。規制は, 真の健全な競争と企業の成長を保証するものでなければならない。
4 モビリティサービス提供の再編成	交通サービスの提供を変革し, HSL (ヘルシンキ交通局) による公共交通機関のチケット販売の独占を撤廃し, 他の企業も公共交通機関のチケットを販売できるようにする必要がある。
5 変革されたオペレーションの確立	変革された事業を確立するためには, 現在のビジネスモデルや構造を変更し, モビリティサービス事業者を生み出す必要がある。そのため, サービス・エコシステムの規制をクリアした事業者の市場参入をサポートする必要がある。

6 購入・補助金手続きの見直し 現在, 自治体や国は HSL を通じて公共交通機関を補助し, 交通サービスを交通サービス事業者から直接購入している. 組織の変革では, 購入と補助金は MaaS オペレーター (≒販売者) に直結し, そこから他のレベルに資金が分配されることになる. このような資金の配分は, (サービス提供者がどのように投資するかを決定できるため,) サービス事業者がモビリティのニーズを革新的に満たすための余地を提供する.

7 パイロットとテストエリア 新しいサービスなどのオペレーションをまず小規模にテストした後, プロトタイプを開発し, 再び市場に投入するフィードバックシステムを機能させることができる.

このような政策目標を実現するために, フィンランドでは多様な法改正が行われている. 「モビリティ・ローミング」の実現を目標として, 既存サービス事業者及び参入する交通事業者が交通に関する情報提供を義務付けられるとともに, チケット発行と決済に関するシステムへの相互運用が可能となる法律が成立, 2017 年に施行されたという<sup>注2)</sup>.

### 2.11.3 (3)MaaS とインターネットの類似性

次に, それぞれのインターネット産業における構造を見て, その構造の類似点を探ってみる. そもそもフィンランドでは, こうした通信産業と運輸産業に関する規制を, 同一の機関が担当している.

インターネットの特徴は, 「いかなる情報であっても運搬可能とすること」, そのために「デジタル化」を挟むことである. デジタル化とは, 従来連続分布であったものの「断片化」, つまり離散分布化である. これをシステムに当てはめるのであれば, 今まであやふやであった役割を定義し, 分割し, その接続要件を定義し, 以てシステム全体の安全性や機能を担保することである.<sup>注3)</sup>

先述の論文で述べられているのは, 交通に関わる様々な機能を分化させ, 接続して全体システムを機能させる構造図である.

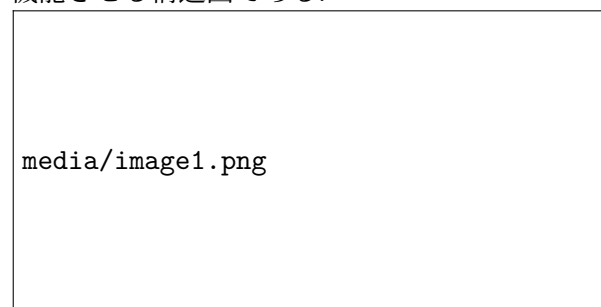


図-1 Hekkila(2014) で提唱される MaaS の主体関係モデル  
Reorganized Mobility Scheme を筆者翻訳

そもそもフィンランドは IT 企業ノキアを擁しており, IT 産業については同様の産業蓄積がある. Heikkila<sup>8)</sup> では, Casey の議論を引用して, 元々断片的であったシステム群を組み合わせ, さまざまなプレーヤー, サービス, 技術を参入させることで, クオリティをある程度犠牲にしつつも分散的かつオープンな市場を形成した「インターネット革命」について言及している. また, 次いで引用される, 将来的な運輸連合の在り方について議論したドイツ運輸事業者連合会 (VDV) の資料には, MaaS に近い議論がある. 具体的には既存の運輸連合とは異なり, 公共交通事業者と消費者の間に「一つのシンプルなインターフェース」である組織を噛ませることで, より消費者に合った商品開発, 情報提供, パッケージングを通じてサービスを魅力的にできるだろう, としている.

以上を纏めると, フィンランドにて定義された MaaS に存在するインターネットエコシステムと類似の概念は 3 点ある. 1 点目はモジュール化であり, 様々な機能についてルールを決め, 目的ごとに分化して整備させることで, システム上でのメタボリズムと堅牢性を確保する. 2 点目はネットワークであり, 前述のモジュールを接続することで, 一つの目的を完結させることを支援する. もう 1 点目はインターフェースの統合であり, 様々なサービスがあったとしても, ユーザーとの接点については統合して管理できるようにする.

<sup>2</sup>石神 (2020)

<sup>3</sup>企業の水平統合への議論にも繋がる. 機能の分化, 統合を行わずシステム全体が瓦解あるいは不安定になった事例として, 日本航空 (日本航空, 日本エアシステム) やみずほ銀行 (第一勧業銀行, 富士銀行, 日本興業銀行) があげられる.



#### 2.11.4 (4) 欧米における MaaS の学術上の定義と分析

特に欧米においては、MaaS の内容について、エコシステムやモデル面から学術面の分析が進んでいる。

Kamargianni, Matyas<sup>12)</sup> は、MaaS のエコシステムとしてコアにデータ提供者、交通事業者、ユーザーを置き、ビジネス環境の構成要員に大学や労働組合、政策立案の主体まで含めたモデルを提唱した。さらに、Arby, Karlsson, Sarasini<sup>13)</sup> は、MaaS の段階として、ビジネスの観点から、以下の 4 レベルに分けた MaaS のモデルを紹介した。具体的には、レベル 0: 統合なし レベル 1: 情報の統合 レベル 2: 予約・支払いの統合 レベル 3: 提供するサービスの統合 レベル 4: 社会全体目標の統合 である。一方で、MaaS というモデル自体の限界も指摘されている。そもそも交通事業者から受け取る手数料が少額であり、また交通が MaaS で提供される統合アプリケーションで購入されているケースも、例えばフランダースでは 3% 程度に過ぎないという<sup>注4)</sup>。

### 2.12 3. 日本版 MaaS の展開と捉え方

翻って、日本での展開は如何様に進んだのだろうか。ここでは、公開されている関連通達の歴史を総攬し、どのような公式の主張がなされたのかを確認する。そして、これの個別事業について、どのような分析の方法をあてはめるのが適切かについて議論する。

#### 2.12.1 (1) 日本における MaaS 関連事業の基礎研究と展開

日本における、中央官庁による MaaS 推進事業は、主に国土交通省と経済産業省が所管して展開されてきた。2019 年には経済産業省と共同で「スマートモビリティチャレンジ」が開始され、全国 28 事業が指定されることとなった<sup>注5)</sup>。さらに国土交通省単独でも、2019 年より「日本版 MaaS 推進事業」が推進され、初年度においては 19 事業、2020 年度には 36 事業が選定された<sup>注6)</sup>。

2020 年 3 月に策定、その後バージョン 2 にアップデートされた「MaaS 関連データの連携に関するガイドライン」では、「Society5.0 リファレンスアーキテクチャ」に基づき、MaaS エコシステムも併せて検討が行われている。利用されるデータを「MaaS 関連データのうち、各 MaaS において設定された最低限のルール等に基づき、当該 MaaS プラットフォームを利用する全てのデータ利用者が利用可能なものとして、当該プラットフォームに提供等が行われるデータ」としての「協調的データ」と、「MaaS 関連データのうち、当該データの提供者との契約等により個別に共有が行われるものとして、各 MaaS プラットフォームに提供等が行われるデータ」としての「競争的データ」に分割して定義される。各交通事業者は、公開されるデータの共通化（フォーマット、言葉の意味など）について共通のものを整備することが推奨されている。また、バス及びフェリーについては、前者は GTFS-JP、後者は「標準的なフェリー・旅客船航路情報フォーマット仕様書（ver2.0）」は標準仕様として設定されている。

#### 2.12.2 (2) 日本版 MaaS の定義

このような国土交通省の動きは、国土交通省の定義する「日本版 MaaS」の考えに基づいている。「日本版 MaaS」の定義についてはこのように述べられている。

日本版 MaaS は、地域住民や旅行者一人一人のトリップ単位での移動ニーズに対応して、複数の公共交通やそれ以外の移動サービスを最適に組み合わせて検索・予約・決済等を一括で行うサービスであり、観光や医療等の目的地における交通以外のサービス等との連携により、移動の利便性向上や地域の課題解決にも資する重要な手段となるものです。<sup>注7)</sup>

<sup>4)</sup>Zipper(2020)

<sup>5)</sup>事業目的によりガバナンスが行われており、その目的となる 5 つは以下のとおりである：①他の移動と重ね掛けして効率化②モビリティでのサービス提供③需要側の変容を促す仕掛け④異業種との連携による収益活用、付加価値の創出⑤モビリティ関連データ取得、交通・都市政策との連携。なお、このプロジェクトは 2020 年に 52 事業へ拡大することとなる。

<sup>6)</sup>国土交通省。

<sup>7)</sup>国土交通省。

この定義を参照すると、日本における MaaS は構造的なものではなく、地域課題解決ベースのサービス提供について重きを置いている。これに基づけば日本版 MaaS は、いかに交通サービスの対価支払い・生産・提供の一連の消費活動を組み替える意図は持たない。つまり、国土交通省が考える MaaS は、制度やシステムの変革というよりむしろ、交通事業者が提供する公共交通サービスについての改善アイデアコンテスト並びに実証実験、ととらえた方が説明はつきやすい。

国土交通省の引用する MaaS の定義は、基本的に「都市と地方の新たなモビリティサービス懇談会」に依拠したものになる。具体的には、「出発地から目的地までの移動ニーズに対して最適な移動手段をシームレスに一つのアプリで提供するなど、移動を単なる手段としてではなく、利用者にとっての一元的なサービスとして捉える概念」である。

日本政府の発行する文書における MaaS という単語の初出は、2018 年 3 月 7 日の第 24 回「未来投資会議」であり、議事での言及はないものの、国土交通省がすでに検討段階に入っていることがうかがえる。具体的には、①交通関連データ連携②運賃設計③まちづくりとの連携、の 3 点において、MaaS を活用する意向であることを示している。同年 10 月に先述の「都市と地方の新たなモビリティサービス懇談会」が設置され、MaaS に関わる規則体系が整備されていくことになる。

このような考えは、いわゆる「地方創生」を念頭に置いていると考えることもできる。実際、日本版 MaaS ではサービスの提供目的を「地域課題の解決」に置き、最近では新型コロナ対応も追加された一、モデル変革を目的としておらず、結果として公共交通政策全体に入るメスはそれほど大きくない。つまり、サービス提供に係るデータプラットフォーム、サービスにおけるアイデアを競って開発している状態で、公共交通へのガバナンスそのものにはあまり変化が見られない。

### 2.12.3 (3) 日本版 MaaS の捉え方

以上から、日本における政策としての「日本版 MaaS」は、源流にある MaaS とは分けて考えるべきである。つまり、本事業にて解決されるべきとされる課題は、公共交通政策のより効率的、効果的な推進のための仕組みを新たに作るというよりは、むしろ既存の仕組みを活用して地域課題解決のために交通を動員しようとしている。

以上より、日本版 MaaS のそれぞれの事業については、既存の地域公共交通政策の一類型として捉えた分析が妥当である。つまり、責任体系の分配システムは変わっておらず、計画を中心とした分析が有効である。

### 2.12.4 (4) 日本における MaaS の定義

なお、日本における MaaS の定義については、議論が乱立しているが、これらは主に 2 つのアプローチに分類される。第一にコンセプトとしてのアプローチ、第 2 にチケット統合サービスとしてのアプローチであり、これが混合している。

前者の代表的な論者として中村文彦、牧村和彦、加藤博和の 3 者がある、中村は MaaS について、一つのように公共交通サービスを統合して運用する必要性を強調する。

最大のポイントは『as a Service』の部分にあると思います。つまり公共交通を『まるで一つのサービスであるかのように』使えなければいけないのです。<sup>注8)</sup>

また牧村は、MaaS について、アプリケーションへ注目が集まる状況に警鐘を鳴らし、MaaS の本質としてライフスタイルへ注目する<sup>注9)</sup>。

MaaS は自動車という伝統的な交通手段に加えて、新たな選択肢を提供し、自家用車という魅力的な移動手段と同等かそれ以上に魅力的な移動サービスにより、持続可能な社会を構築していこうというまったく新しい価値観やライフスタイルを創出していく概念だ。

さらに加藤は、MaaS について、新たなコンセプトの訳語としての役割を与えている。つまり、「M: もっと a: あなたらしく a: あんしんして S: せいかつできるために」であり、「少なくとも、S=Smartphone

<sup>8</sup> AndE(2023)。

<sup>9</sup> 牧村 (2021) p18。

ではない」「移動利便性（mobility）向上と地域活性化によって『乗って楽しい』『降りても楽しい』おでかけの選択が促進され『健幸』に資する」とする<sup>注10)</sup>。こうしたコンセプト（牧原）の指摘するイギリス行政学的に言えばドクトリンとして MaaS を活用している状況にある<sup>注11)</sup>。

後者の代表的な論者として石田東生がある。石田は 2020 年の記事の中で、最大の MaaS システムとして「全日本空輸（ANA）や、日本航空（JAL）の既存の予約ネットワークシステム」を挙げており、MaaS が統一して予約できるチケットサービスとして捉えられている<sup>5)</sup>。

## 2.13 4. 現在の法体系と「交通まちづくり」

次に、地域公共交通に関する法律の概要から見ていく。日本版 MaaS は現状、一つの理念法、一つの計画、一つの行政法に基づき遂行されている。

### 2.13.1 (1) 交通政策基本法

交通政策基本法は、2013 年に可決・成立した、日本における交通政策の理念を定めた法律である。当時の民主党政権が掲げた公約の一つでもあり、政権交代後も議論が続き、成立した。本法では、2 条から 7 条までを理念として、国には「国民の理解を得るよう努める」こと、地方自治体には「その地方公共団体の区域の自然的経済的社会的諸条件に応じた施策を策定し、及び実施する責務」、交通事業者は「その業務を適切に行うよう努めるとともに、国又は地方公共団体が実施する交通に関する施策に協力するよう努める」ことをそれぞれ求め、またこの 3 者ならびに住民その他の関係者に対して、「基本理念の実現に向けて、相互に連携を図りながら協力するよう努める」としている。また「交通政策基本計画」の策定を国に義務付け、環境大臣との協議を義務付けるなど、所謂「横ぐし」の法体系となっている。

公共交通に関わる点の理念について確認すると、「交通の機能を管理する」という点に重点が置かれている。まず第 2 条にて「（交通の）機能が十分に発揮されることにより、国民その他の者の交通に対する基本的な需要が適切に充足されることが重要である」という基本的認識の下に行われなければならない」として、交通政策の意義が定義されている。そして第 3 条にて、目的志向型の交通政策目標設定義務、つまり「交通が、豊かな国民生活の実現に寄与するとともに、我が国の産業、観光等の国際競争力の強化並びに地域経済の活性化、地域社会の維持及び発展その他地域の活力の向上に寄与するものとなるよう、その機能の確保及び向上が図られることを旨として行われなければならない」が記載されている。さらに、その機能を各交通手段が「特性に応じて適切に役割を分担し、かつ、有機的かつ効率的に連携する」ことを規定し（第 5 条）、また「関係者が連携し、及び協働しつつ、行われなければならない」とする（第 6 条）。

### 2.13.2 (2) 交通政策基本計画

交通政策基本計画は、前述した交通政策基本法に基づいて策定される計画である。第 1 次の計画は 2015 年に策定され、2021 年に更新されている。

計画の位置づけは 2015 年発行の第 1 次計画に記載されており、「交通政策基本法が提示するこれらの交通政策の長期的な方向性を踏まえつつ、政府が今後講ずべき交通に関する施策について定めるものである」とされている。また構成としては、「交通に関する施策の「基本的方針」、計画期間内に目指すべき「目標」、目標の各々について取り組むべき「施策」の三層構造となっており、関係者の責務・役割や連携・協働等についても、施策の推進に当たって「留意すべき事項」として整理している」（第 1 次計画「はじめに」より）。また新型コロナウイルス感染症拡大に伴う利用客数現象の段になって、第 2 次計画にはこうした時代背景を踏まえる形で、「交通が直面する「危機」と、それを乗り越える決意」という章が追加され、交通事業に対する危機感がつづられている。

<sup>10</sup>加藤 (2019) p6.

<sup>11</sup>MaaS を日本に導入した書籍として、日高、牧村、井上、井上 (2018) への言及は欠かせないが、論者として書きぬけないためここでは除外した。この書籍では MaaS について概念的に紹介しており、記述としては以下である；「あらゆる交通手段を統合し、その最適化を図ったうえで、マイカーと同等か、それ以上に快適な移動サービスを提供する新しい概念」。



横ぐしの法律を背景としているが、地域公共交通に関連して目標となっている数値はほとんどの場合、事業がどの程度導入されているか、が指標となっている。

表-2 交通政策基本計画における目標値設定

①  $> p() * 0.1579 > p() * 0.1473 > p() * 0.1790$  ②

第1次計画 第2次計画

施策項目記号 基本目標 A 基本目標 A

地域公共交通に関する

施策項目名 ・自治体中心に、コンパクトシティ化等まちづくり施策と連携し、地域交通ネットワークを再構築する

・地域の実情を踏まえた多様な交通サービスの展開を後押しする 地域が自らデザインする、持続可能で、多様かつ質の高いモビリティの実現

目標数値

(交通事業/計画) 地域公共交通網形成計画の策定総数

鉄道事業再構築実施計画(鉄道の上下分離等)の認定件数

デマンド交通の導入数

LRT の導入割合

コミュニティサイクルの導入数 地域公共交通計画の策定件数

地域公共交通特定事業の実施計画の認定総数

鉄道事業再構築実施計画(鉄道の上下分離等)の認定件数

新たなモビリティサービスに係る取組が行われている地方公共団体の数

目標数値

(非交通事業/計画) 航路、航空路が確保されている有人離島の割合 航路、航空路が確保されている有人離島の割合

### 2.13.3 (3) 地域公共交通活性化・再生に関する法律

地域公共交通活性化・再生に関する法律は、2007年より施行された公共交通における「総合計画法」である。この法律は、元々は「施策支援法」であった。「頑張る地域を応援する」、「LRT」をはじめとした道路運送法だけでは対応できなかったバス事業以外の交通事業について包括的な連携を図るとともに、交通システムごとの支援メニューを取りそろえたシステムであったのである。さらに、「地域公共交通総合事業」が設定され、関連する実証実験やその調査に関しての補助も設定された。しかし、2010年近傍のいわゆる「事業仕分け」を経て、地域の生活交通を確保する「総合計画法」に転化し、取り組み支援型の事業支援システムを残したまま強制力を持ち始めた。この法律に基づいて策定される「地域公共交通計画」「地域公共交通利便増進計画」を所管する地方運輸局には、路線届出の最終的な受理権限もある。結果としてこのシステムには、地域ごとに、その地域が考える「あるべき交通システム」を設計するためのこうした計画に、「どのような交通手段を導入すべきか」「どのように交通を維持確保していくべきか」という点で、メニューの採用に圧力がかけやすい特性がある<sup>注12)</sup>。

### 2.13.4 (4) 交通まちづくりの定義とその親和性

主に交通政策基本法を基底とした理念体系を確認すると、この内容が驚くほど「交通まちづくり」のそれと類似していることが分かる。

交通まちづくりは前述のように、都市の課題解決のために、公共交通を中心とした交通を動員する交通政策の概念である。既存の交通需要を満たすための交通政策とは異なり、地域の課題解決に貢献する交通を、目的の設定から市民参加を加える形で設計、提供する計画手法で、2005年ごろから主に札幌、横浜、岡山、熊本など政令市における市民団体より勃興してきた。

<sup>12</sup>永田(2022)第2章。

土木学会にて行われた交通まちづくりの議論では、議論が進むにつれて市民参加の重要性について序列を下げる傾向はみられた<sup>注13)</sup>ものの、方向性は変わらない。すなわち、公共交通を手段とする方向性は、交通まちづくりの根幹であると言える。政策の直接の接続は確認できていないものの、思想の文脈はここに接続している。

## 2.14 5. 市民参加の捉え方

本稿は市民参加プロセスについて、手続主義、非帰結主義のスタンスを取るものの、公共交通それ自体に対しては帰結主義のアプローチを取る。

### 2.14.1 (1) 公共交通政策における市民参加に関する議論

公共交通政策における市民参加に関する規範的な議論は、帰結主義に立って市民参加が情報提供と労力提供の役割を持つと説く太田<sup>8)</sup>、非帰結主義に立って6つの要件<sup>注14)</sup>を示す太田<sup>9)</sup>、実証的に7つの要件を提示し、協働を通じて「新たな公共、地球の公益」を創造すべきとした森栗<sup>10)</sup>などがある。

近年の交通政策における市民参加に関する実証的な議論では、その市民参加形態は、主に不足する財源や人的資源を確保するため、市民が交通運営者として参画する物に集中している<sup>注15)</sup>。ただ、こうしたスタンスではあくまで、どのように交通事業を成立させるか、また「目の前の交通に対する課題をどのように解決するか」という観点に基づいている。市民参加や住民参加は、目の前にある課題を炙り出し埋め合わせるのに有用であると評価され、計画立案プロセスに組み込まれている。ただこうしたスタンスは、こと「横ぐし」の領域を扱う地域政策においては、課題の相対化を妨げ、重要度の高い解決すべき課題を見落とす懸念がある。たとえば高齢者の健康増進に公共交通を貢献させるという議論をしても、実のところ効果は他の政策の方が高い、ということは検討されにくい。さらに交通事業者などの公益的な企業は、地域政策の構成員として「選ばれる」というプロセスを経ないことが多いため<sup>注16)</sup>、概して一方的になりやすい可能性がある<sup>注17)</sup>。こうした状況は、何らかのシステムの導入を通じて是正されるべきと考えられる。

### 2.14.2 (2) 公共サービスとしての公共交通

一方で、地域公共交通は自治体の計画によって定義され、非競合性を持たないため、インフラとしてではなく公共サービスとして定義されるべきである。通信サービスや道路などの非競合性のある程度持つ財とは違い、利用できる時間や容量が、物理的に決定されているからである。つまり、一旦整備したところで供給量は簡単に变化させられない即時材であるとともに、また提供に必要な労力や資源には限りがある。市民参加を通じてそうした資源を導入できるとしても限界があるから、得られる効果とコストとのバランスについて検討すべきである。

### 2.14.3 (3) 望ましい帰結/手続きのスタンスの提案

以上を踏まえ、市民参加の類型としては、地域の目標のうち公共交通が貢献すべき領域を設定する、政策目標設定から関わる形態が望ましい。一方で交通サービスに関しては、他行政サービスと同列

<sup>13</sup>もちろん、こうした議論が政令市から立ち上がった点に関して、財源の自由度や市民全体の理解レベルの高さが寄与した点は容易に想像がつく。このような高い教育レベルと自由度を交通まちづくり初期の前提と考えるならば、日本全国へ交通まちづくりを広げるにあたって、多様な市民の参加を諦めた可能性は大いにある。

<sup>14</sup>1: 民意の反映 2: 民意による利害調整 3: 専門知の活用と専門家の専横の排除 4: 政治プロセスの失敗回避のための国からの各種規制 5: 効率的な交通政策が導出可能 6: 地方自治体から独立 4に関しては、今日の地方自治法と当時のそれは違うから注意が必要である。また6に関しては、都市圏域や生活圏行きが自治体とは違う点から指摘されているのであるが、これらが同一と認められるのであれば独立していなくても問題ないと思量される。

<sup>15</sup>代表的なものだと、山下(2009)、福本・加藤(2012)、加藤(2012)などが挙げられる。

<sup>16</sup>市場原理の導入も、実際のところ選ばれた状態を確立できているかは不明である。この認識は規制緩和後でも、市場の入退出が活発でない現状に起因する。寺田(2004)によれば、乗合バス市場における新規参入は活発ではなく、規制緩和から1年後までの新規参入者の市場シェアは1%を下回る。規制緩和による退出ルールの明確化が休廃止を増加させた傾向はなく、乗合バスの市場構造は規制緩和後にもあまり変化していない」という。

<sup>17</sup>公益法人のガバナンスの更なる強化等に関する有識者会議(2021)。

に比較され、その地域への貢献によって選択されるべきであろう。

現状の市民参加は、あくまで目の前にある課題の解決を旨としており、実証主義に基づいた太田の主張通りに情報提供と労力提供の役割を持っている。しかしそれだけでは、全ての計画段階における批判はなし得ないものであり、少なくとも交通の「要件定義」段階である目標立案レベルへの市民参加は、公共サービスの一方性を排するためには重要な要素である。

具体的にはまず、初めに地域において、交通事業以外の数値によって語られる公共交通がもたらすべき影響、つまり「効果」の定義を行う。結果として、交通をあくまで手段として、その「効果」への貢献を推定して議論できる。ここには基本的に、政策手段としての事業を推進する立場となる点から、交通事業者を意思決定主体として組み込んではいならない。

こうした目標への市民参加は、それ自体がハンス・ケルゼン<sup>11)</sup>の指摘する民主主義の本質と深く結びついている。政策の目標、手段、あらゆる点に対して市民に批判するための回路を残すことで、武力によるクーデターがなくとも政権の構造転換をなしえる、というのがその主張である。

## 2.15 6. 日本における公共政策上の MaaS の分析

### 2.16 2020 年の日本版 MaaS 推進 36 事業の総攬

以上を踏まえた上で、日本において MaaS を推進する「日本版 MaaS」について確認したい。

#### 2.16.1 (1) リサーチクエスションの設定

以下の 2 点をリサーチクエスションとして置く。

- ①その政策進捗において目標設定がどこに置かれているか
- ②立場を問わない市民による批判回路がどのように整備されているのか

本研究は、日本版 MaaS を「MaaS」としてではなく、「交通まちづくり」の一類型として分析し、その目的意識と計画における責任体系の構築状況、すなわちガバナンス状況を探るものである。

#### 2.16.2 (2) 分析項目・手法

分析手法について、「計画における計画指標（Key Performance Indicator, 以下 KPI）に事業利用状況以外の指標がある」「会議体において、交通事業者、システム供給者、学識者、自治体以外の市民参加がある」の二つの点を検討し、交通政策を地域政策として運営できているかの評価とする。

①については、MaaS が「事業の手段」としてではなく、「政策の手段」として機能しているかを確認するものである。これを分析する手法としては議事録を分析する、議会を分析するなどの手法が考えられるが、そもそもすべての事例で MaaS について議会で論点となっているわけではない点、また議事録を公開している協議会自体が少ない点から、こうした議事録ベースのアプローチを行うのは困難である。

この点から、評価指標の状況について観察対象とすることとした。MaaS 事業自体の事業性や交通事業の利用状況だけではなく、それ以外の評価指標を計画に置いているかを確認する。

②については、地域における交通事業者など公益的な企業の特性を鑑み、市民参加が行われており目標や手段への批判プロセスがあるかを確認する。市民の行政への批判回路について、議会を通じた参加等はあるが、交通計画では協議会参加、グループヒアリング、アンケート調査、パブリックコメントが推奨手法として取り上げられている<sup>注18)</sup>。これらのうち、合議への参加を通じて直接の批判とできるのは協議会参加のみである。

以上の検討より、会議体への市民の参加を指標としてとらえることとした。座談会、協議会を代表とする場における、交通事業者、システム供給者、学識者、自治体以外の市民参加の存在を、国土交通省が発行する先述のドキュメントと検索エンジン「Google」「Bing」「Yahoo!Japan」を利用した文書探索を通じて確認した。

<sup>18)</sup>国土交通省 (2020)。

### 2.16.3 (3) 分析対象

日本版 MaaS は、その事業概要を一定のフォーマットに落とし込んで申請する仕組みとなっており、国土交通省は採択のたびにそれらをまとめた文書を公開している<sup>注19</sup>。これらには協議会の構成員や目標数値が記載されており、これを分析対象とする。2020 年の日本版 MaaS 推進事業では、全国 36 の事業が認定されており、提出された計画については全件採択となっている。

ここで、本事業によって実施された事業の特性を説明するために、日本版 MaaS として展開された各事業について、要素技術の利用状況について確認し、類型化を試みる。日本版 MaaS の主な要素技術として、サブスク・定期など企画乗車券、オンラインで交通情報が見られる、あるいは決済できるアプリケーション、病院、商業クーポンなど外部情報の提供、デマンド交通、シェアサイクル等の新モビリティ、さらにデータ連携基盤の構築や利用が認められた。特にアプリケーションの提供が 31 事例、外部情報との連携が 32 事例と多かったものの、企画乗車券の導入と新モビリティの導入比率は低く、またデータ連携基盤の利用や構築は 12 事例と少なかった。なお、参入の高度化に必要な API をはじめとしたデータ連携のオープン化に関して言及しているのは 2 事例のみであった（加賀、しんゆり）。

### 2.16.4 (4) 分析結果

分析の結果、双方の判断を満たし、地域政策として日本版 MaaS を運用している事例は存在しなかった。

①について、交通事業や MaaS 事業の指標を導入している事業は 32/36 であったが、それ以外の地域目標を設定している事業は 10 のみであった。さらに②については、自治体が参加している事業に限定すると、パブリックコメントや協議会への市民参加を確認できた事業は 2/32 にとどまった。

表-3 日本版 MaaS における要素技術の状況



### 2.16.5

### 2.16.6 (5) 事例

分析の射程を示すため、今回の分析の結果を 2 つほど挙げる。

<sup>19</sup>国土交通省 (2020b).



### 2.16.7 1. 地方版 MaaS の広域連携基盤構築モデル事業（ひたち圏域）茨城交通株式会社

日立市を中心とする茨城県北部の圏域における MaaS 事業展開について、3 市 1 村の協力の下で行われた日本版 MaaS 事業である。

#### 2.16.8 I. 指標

①について、本事業の目標は 3 つに分かれている。一つ目は「利用者関連指標」であり、内容は以下の通り利用者数が参照されている。

- ・取組ページへのアクセス数：20,000imp 以上
- ・地域内の取組認知度：50%以上
- ・アプリ DL 数（合計）：2,500 以上
- ・チケット販売数：10,000 枚以上
- ・周遊券の販売枚数：100 枚以上
- ・セット券・企画乗車券の販売数：100 枚以上
- ・通勤型デマンドの利用者数：300 人以上
- ・ラストワンマイルデマンド利用者数：100 人以上

2 つ目は「交通事業者関連指標」であり、指標は参加する事業者数により達成されるものになっている。いずれも MaaS によってカバーされる人口はじめとした MaaS の広がりを目標としている。

- ・基盤に参加する事業者数：3 社以上
- ・利用者数ベースのカバー数：70%以上
- ・新しく造成した商品数：5 つ以上
- ・MaaS 基盤へのデータ提供社数：4 社以上

3 つ目は「MaaS 事業者関連指標」であり、以下のように接続事業者を目標と据えていて、2 つ目と同様の性格を帯びる。

- ・システム基盤へ対応事業者数：3 社以上
- ・チケット発券する事業者数：3 社以上

#### 2.16.9 II. 市民参加状況

②について、参加する主体について確認したところ、市民団体らしき仕組みは確認できず、市民代表者の参加形跡も確認できなかった。また、市民参加、パブリックコメントなどの情報を Google 検索上で探したが、見当たらなかった点から市民参加システムなしとした。

### 2.16.10 2. 鞆の浦 MaaS

広島県福山市の景勝地、鞆の浦における MaaS 事業展開について、同市の協力の下で行われた日本版 MaaS 事業である。

#### 2.16.11 I. 指標

①について、本事業の目標は、そのデータ取得方法に合わせて 3 つに分かれている。

一つ目は「デジタルチケット発行数」であり、実証実験チケットの利用枚数 200 枚が目標になっている。

二つ目は「アンケート指標」であり、以下の通りアプリケーションの満足度や性能へのフィードバック、回遊個所の向上等をアンケートで確認している。

- ・満足度  
目標 満足度 70%以上
- ・setowa に対する評価検証  
目標：今後の課題点抽出 10 件以上
- ・実証実験を通じた来訪誘発・回遊性向上効果  
目標：setowa 利用により立寄り場所の増えた者の割合 10 %



三つ目は電動レンタサイクル利用者の回遊エリア拡大状況であり、電動レンタサイクルのデータをもとに平均利用距離 15km 以上を目指している。

以上より本事業では、関連する交通事業の指標というよりはむしろ、MaaS 事業に特化した指標が並んでいると判断した。

#### 2.16.12 II. 市民参加状況

②について、参加する主体について確認したところ、市民団体らしき仕組みは確認できなかったが、観光関連団体の代表として「公益社団法人福山観光コンベンション協会」が出席しており、これは交通事業者外の参加と認められることから、市民参加システムありとした。

#### 2.16.13 (6) 考察

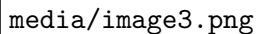
分析対象とした日本版 MaaS 推進事業へ認定されたそれぞれは、その目標を考えると、ほとんどの場合交通事業の成立、また日本版 MaaS 事業の成立を旨としている。実際のところ、本事業を担当する国土交通省新モビリティサービス推進課の担当者は、「申請前の 1 年で事業検討して、申請後の 1 年で事業として自走できるようにしてほしい」と述べており<sup>注20</sup>、事業としての成立を企図している。一方で、こうしたスタンスは実際のところ、海外ではすでに行き詰まり感を見せている<sup>21</sup>。日本における MaaS は、こうした国際的な流れを反映していない。

また交通政策の理念として、交通の機能について「地域の活力に貢献」するように施策を運営するためには、少なくとも地域がどのようになるべきかの指標を入れるべきである。交通事業、MaaS 事業の指標のみでは、こうした運営には不十分である。それは交通事業の、利益を産む/費用を抑える事業としての価値こそ管理できるが、地域交通事業はこうした効率化が難しい現状を鑑みると、地域における機能については管理できないからである。

表-4 日本版 MaaS における目標値決定状況

<sup>20</sup>2022 年 5 月のヒアリングによる。

<sup>21</sup>海外での MaaS の行き詰まり感に関しては、2024 年 3 月に破産、事業承継を実施したほか、技術とデータフォーマットが障壁として挙げられている A. Gerber and K. Hinkelmann (eds.): *Analysing Barriers in the Business Ecosystem of European MaaS Providers: An Actor-Network Approach*, Society 5.0 2023 (EPiC Series in Computing, vol. 93), pp. 68–81 や、輸送事業者が細分化され、データ共有の枠組みが不十分な地域でオープンデータ不足を障壁とする Marc, Hasselwander., João, F., Bigotte.: *Transport Authorities and Innovation: Understanding Barriers for MaaS Implementation in the Global South.*, Transportation research procedia, 2022, また特に通常の交通アプリとの差異を見分けられない状況を原因と示した Andrea, L., Hauslbauer., B., Verse., E., Guenther., T., Petzoldt.: *Access over ownership: Barriers and psychological motives for adopting mobility as a service (MaaS) from the perspective of users and non-users.* Transportation research interdisciplinary perspectives, 2023. などが挙げられる。



さらに、こうした行政機関が手がける交通サービスに対する批判回路についても、十分に構築されていない状況が見えてきた。もちろん「現状のサービスはユーザーヒアリングを経ている」という反論もあるだろうし、「議会制民主主義が批判回路となりうる」という意見もあるだろう。しかし、公共交通政策は多くの場合、土木工学を主軸とした専門性を問われることから<sup>注22)</sup>、こうした回路が機能しにくいと推察される。以上より、専門家が知識面、手法面でサポートする形で批判回路を設定すべき、と思量される点を前提とすると、現状の批判システムは不十分である可能性が示唆される。

#### 2.16.14 (7) 日本版 MaaS は交通まちづくりの一類型としてとらえられるか？

ここまでの議論を踏まえて、日本版 MaaS を交通まちづくりとして捉えられるかについて議論したい。交通まちづくりの基本線は、初期構想段階では市民参加を含んでいた以外は一貫しており、地域の課題を解決する方針のもと交通を計画し、提供するものであった。

これを踏まえると、日本版 MaaS はその構想段階では交通まちづくりの要件に合うものの、実装段階においてはその要件に合致しないと理解できる。前者については、日本版 MaaS の特徴として地域課題を解決する交通サービスの提供を支援する政策である点が該当する。一方で後者については、日本版 MaaS の計画において交通サービスの利用状況が評価の中心にあり、必ずしも地域課題解決の状況をモニタリングしていない点が該当する。

このような状況になった要因として、いくつかの仮説を挙げておきたい。第1に、新たなガバナンスシステムの導入へのコストが高いことがあげられる。一般に、デジタル時代の政府活動への意向にはそれに即した評価、批判のシステムへ移行する必要があるが、それには取引費用が掛かりすぎる事が分かっている<sup>注23)</sup>。第2に、これは第1の指摘と関連するが、杉谷<sup>15)</sup>の述べるようにそもそも自治体の持つ政策評価システムが業績評価に偏っている点があげられる。三重県の「さわやか運動」から始まった政策評価では、担当者が直接事業内容を説明する「事務事業評価」が核であり、その中で簡略化された評価として業績評価が利用されることが多い。これは一種の日本の評価システムの文化であり、日本版 MaaS もその流れを汲んでいる可能性は十分にある。

一方で、市民参加の文脈については、そもそも日本版 MaaS の概念において組み込まれていない状況で、情報公開や批判回路についても明確ではない。公共政策学の蓄積を踏まえると、地域交通政策においてこの点は後退基調にあると考えられ、早急な改善が求められる点といえる。

<sup>22</sup>土木計画のうち、交通計画の費用便益分析においては、どの程度の需要が発生するかが重要なファクタであり、その算出には一般的に「四段階推定法」と呼ばれる手法を用いるのが主流である。屋井 (2021) 参照。

<sup>23</sup>Duleavy et al (2006)

### 2.16.15 (8) 本分析の限界

本稿の限界の一つは、そもそも調査対象が日本版 MaaS 事業のみである点である。地域公共交通に関する事業は MaaS 事業だけではなく、LRT 事業や再編事業など多様である。また批判プロセスの調査についても、Google の機能に制約を受けるものであって、この限界を越えるには個別インタビューなどの質的調査を施し、分析を行う必要がある。さらに、1998 年の行政改革会議を経て実施された地域公共交通の規制緩和について、その総決算を行っていない点もまた課題である。

## 2.17 7. 終わりに

本稿では、MaaS を推進する政策において、交通に関わる理念法である「交通政策基本法」を参照した上で、「交通まちづくり」の概念を引用し、その思想を日本版 MaaS の政策分析に援用できるかを確認した。

まず、日本版 MaaS とフィンランドにおいて発案された時点のそれとを比較し、前者が地域課題解決のための仕組みとして変容したことを導き出した。この点は「交通まちづくり」の思想と一致している。次に、日本版 MaaS 推進事業において採択されたすべての事業内容を確認、分析した。その結果、これらの事業は「交通まちづくり」の思想と完全に合致するものではないと明らかになった。本稿の結論は、日本版 MaaS に関する文献で上げた魏<sup>3)</sup>の結論を支持するものであった。つまり、地域の課題解決というよりはむしろ、日本版 MaaS でとらえられた課題意識は地域交通事業の課題解決であった。ただ、本稿ではそういった決定やプログラム運営に至った経緯については立ち入らなかった。

地域公共交通政策として進行中の各政策を概観すると、こうした指摘は未だ色褪せていないと考える。

有識者会議の下で 2022 年 8 月 27 日にまとめた「アフターコロナにおける地域交通のり・デザイン」において、政府は「3つの共創」を利用して、地域交通の確保に取り組むとしている。加えて、2023 年 10 月 1 日に改正された「地域公共交通の活性化及び改正に関する法律」に基づく政策展開においては、特に公的資金投入の拡大と「鉄道再構築協議会」によるさらなる事業者自治体間協議の活性化が取りざたされている。ただ、責任分担をどのように建てつけるのかの議論は見られず<sup>注24)</sup>、また公共交通の意義や事業への批判システム、事業者の一方性を排するシステムの構築は検討されていない。公共交通の存在意義を見つめなおし、従来の政策を批評し更新し続ける人的・物的・経済的資源の動員体制を取ってこそ健全な地域交通が達成され则认为る初期の「交通まちづくり」の思想文脈からすれば、この状況は好ましくない。

「利用者視点」を導入しているとうたう交通政策が各地で展開されているが、少なくとも政策において相手にするのは市民であり、この国の主権者である。もちろん主権者の意見が全て正しいわけでもないが、主権の所在に目を背け、専門性で固め、「利用者のため」と謳って実施される交通政策はまさに、善意で舗装された地獄への道となり得るのである。専門性が必ずしも正しさではなく、目指す方向性はイデオロギーであり批判されうることを、これらを前提として、民主主義国家の計画のあり方を交通の面から探る作業は、今後の課題としたい。加えて、パブリック・インボルブメントと呼ばれる、サービスではなくインフラ整備にかかる市民参加プロセスについても、今後の研究における検討課題としたい。

<sup>24</sup>例えば今回の法改正で、JR 各社の運営する鉄道に関する「再構築協議会」こそ設置されこれを「十分な説明」の舞台とすることとなったが、JR と自治体間の関係性については基本的には変わらない。本制度の成立は基本的に性善説によっている。これが依拠できないものとなった場合、国の関与が行われるよう変化する本制度において、JR による同意を取れない状況での路線廃止は路線がいかなる状況でも行われる上に国地方係争処理委員会の対象外になることが想定される。結果として、国は種々の路線に関する責任を持たず、JR の撤退行動を許容することになる。

また今回の法改正にて、ある地域の交通を統合的に受託するエリア一括運行事業が制定された。本事業は既存のバス交通に関する補助を合計して自治体へ渡し、それを利用して一括運行事業者を選ぶ仕組みである。この機能は自治体の能力水準によっている。自治体は補助額削減インセンティブを持ち、事業者は費用削減を要求するが、適切な競争や契約形態を選択しなければそれは実現されない。既存のバス交通関連補助を利用している点で、地域公共交通会議をはじめとした協議会経由の補助システムと同様となり、責任体系は既存のシステムと変わらない。

## 2.18 Notes

1. 石神孝裕:MaaS 先進国フィンランドの視察, IBS Annual Report, 2020 より.
2. 企業の水平統合への議論にも繋がる. 機能の分化, 統合を行わずシステム全体が瓦解あるいは不安定になった事例として, 日本航空 (日本航空, 日本エアシステム) やみずほ銀行 (第一勧業銀行, 富士銀行, 日本興業銀行) があげられる.
3. David Zipper: The Problem With 'Mobility as a Service', Bloomberg.

<https://www.bloomberg.com/news/articles/2020-08-05/the-struggle-to-make-mobility-as-a-service-a-reality>  
閲覧

4. 事業目的によりガバナンスが行われており, その目的となる 5 つは以下のとおりである: ①他の移動と重ね掛けして効率化②モビリティでのサービス提供③需要側の変容を促す仕掛け④異業種との連携による収益活用, 付加価値の創出⑤モビリティ関連データ取得, 交通・都市政策との連携. なお, このプロジェクトは 2020 年に 52 事業へ拡大することとなる.

5. 国土交通省: 日本版 MaaS の推進,

<https://www.mlit.go.jp/sogoseisaku/japanmaas/promotion/index.html>, 2022/10/18  
閲覧

6. 国土交通省: MaaS のモデル形成,

<https://www.mlit.go.jp/sogoseisaku/japanmaas/promotion/model/index.html>, 2022/10/18 閲覧

7. AndE: 新しい移動の概念「MaaS」の現在 MaaS 社会実現のカギは「つなげる」こと 競争と協調を見極める, 2023 より.

<https://www.andemagazine.jp/2023/05/25/maas-interview.html> 2023/10/06 閲覧.

8. 牧村和彦:MaaS が都市を変える 移動 × 都市 DX の最前線, 学芸出版社, 2021

9. 加藤博和; 地域公共交通改革のためいまこそ殻を破ろう! ~地域の, そして日本の将来を救うために~, 令和元年度 第 2 回 (第 16 回) 国土交通省交通政策審議会交通体系分科会 地域公共交通部会, 2019 p6.

10. MaaS を日本に導入した書籍として, 日高洋佑, 牧村和彦, 井上佳三, 井上岳一:MaaS モビリティ革命の先にある全産業のゲームチェンジ, 日経 BP 社, 2018 p18. への言及は欠かせないが, 論者として書きぬけないためここでは除外した. この書籍では MaaS についてコンセプト的に紹介しており, 記述としては以下である: 「あらゆる交通手段を統合し, その最適化を図ったうえで, マイカーと同等か, それ以上に快適な移動サービスを提供する新しい概念」

11. 永田右京: 公共交通の構造転換, 慶應義塾大学総合政策学部卒業論文, 2022. 第 2 章.

12. もちろん, こうした議論が政令市から立ち上がった点に関して, 財源の自由度や市民全体の理解レベルの高さが寄与した点は容易に想像がつく. このような高い教育レベルと自由度を交通まちづくり初期の前提と考えるならば, 日本全国へ交通まちづくりを広げるにあたって, 多様な市民の参加を諦めた可能性は大いにある.

13. 1: 民意の反映 2: 民意による利害調整 3: 専門知の活用と専門家の専横の排除 4: 政治プロセスの失敗回避のための国からの各種規制 5: 効率的な交通政策が導出可能 6: 地方自治体から独立.



4 に関しては、今日の地方自治法と当時のそれは違うから注意が必要である。また 6 に関しては、都市圏域や生活圏行きが自治体とは違う点から指摘されているのである、これらが同一と認められるのであれば独立していなくても問題ないと思量される。

14. 代表的なものだと、山下祐介: 地域公共交通をめぐる社会実験と住民参加, 運輸と経済第 69 巻 12 号, 2009., 福本雅之, 加藤博和: 地域公共交通への住民参画の促進方策に関する検討, 第 45 回土木計画学研究発表会 (春大会), 2012. などが挙げられる。
15. 市場原理の導入も, 実際のところ選ばれた状態を確立できているかは不明である。この認識は規制緩和後でも, 市場の入退出が活発でない現状に起因する。寺田一薫: バス産業の規制緩和, 日本評論社, 2004. によれば, 乗合バス市場における新規参入は活発ではなく, 規制緩和から 1 年後までの新規参入者の市場シェアは 1%を下回る。規制緩和による退出ルールの明確化が休廃止を増加させた傾向はなく, 乗合バスの市場構造は規制緩和後にもあまり変化していない」という。
16. 公益法人のガバナンスの更なる強化等に関する有識者会議: 公益法人のガバナンスの更なる強化等のために (中間とりまとめ), 2021
17. 国土交通省: 日本版 MaaS の推進,  
<https://www.mlit.go.jp/sogoseisaku/japanmaas/promotion/index.html>, 2022/10/18 閲覧
18. 国土交通省: MaaS のモデル形成,  
<https://www.mlit.go.jp/sogoseisaku/japanmaas/promotion/model/index.html>, 2022/10/18 閲覧
19. 2022 年 5 月のヒアリングによる。
20. 海外での MaaS の行き詰まり感に関しては, 2024 年 3 月に破産, 事業承継を実施したほか, 技術とデータフォーマットが障壁として挙げられている A. Gerber and K. Hinkelmann (eds.): Analysing Barriers in the Business Ecosystem of European MaaS Providers: An Actor-Network Approach, Society 5.0 2023 (EPiC Series in Computing, vol. 93), pp. 68–81 や, 輸送事業者が細分化され, データ共有の枠組みが不十分な地域でオープンデータ不足を障壁とする Marc, Hasselwander., João, F., Bigotte.: Transport Authorities and Innovation: Understanding Barriers for MaaS Implementation in the Global South., Transportation research procedia, 2022, また特に通常の交通アプリとの差異を見分けられない状況を原因と示した Andrea, L., Hauslbauer., B., Verse., E., Guenther., T., Petzoldt.: Access over ownership: Barriers and psychological motives for adopting mobility as a service (MaaS) from the perspective of users and non-users. Transportation research interdisciplinary perspectives, 2023. などが挙げられる。
21. 土木計画のうち, 交通計画の費用便益分析においては, どの程度の需要が発生するかが重要なファクタであり, その算出には一般的に「四段階推定法」と呼ばれる手法を用いるのが主流である。屋井鉄雄: 土木と環境の計画理論-3つの並行プロセスによる計画づくり-, 株式会社数理工学社, 2021. 参照。
22. Dunleavy, Patrick, et al.: New public management is dead—long live digital-era governance. Journal of Public Administration Research and Theory 16(3), pp.467-494, 2006.
23. 例えば 2023 年の法改正にて, JR 各社の運営する鉄道に関する「再構築協議会」こそ設置されこれを「十分な説明」の舞台とすることとなった(「十分な説明」については国土交通省: 新会社とその事業を営むに際し当分の間配慮すべき事項に関する指針, 国土交通省告示第千六百二十二号, 2001. 参照) が, JR と自治体間の関係性については基本的には変わらない。本制度の成立は基本的に性善説によっている。これが依拠できないものとなった場合, 国の関与が

行われるよう変化する本制度において、JR による同意を取れない状況での路線廃止は路線が  
いかなる状況でも行われる上に国地方係争処理委員会の対象外になることが想定される。結  
果として、国は種々の路線に関する責任を持たず、JR の撤退行動を許容することになる。

また今回の法改正にて、ある地域の交通を統合的に受託するエリア一括運行事業が制定された。本  
事業は既存のバス交通に関する補助を合計して自治体へ渡し、それを利用して一括運行事業者を選  
ぶ仕組みである。この機能は自治体の能力水準によっている。自治体は補助額削減インセンティ  
ブを持ち、事業者に費用削減を要求するが、適切な競争や契約形態を選択しなければそれは実現さ  
れない。既存のバス交通関連補助を利用している点で、地域公共交通会議をはじめとした協議会経  
由の補助システムと同様となり、責任体系は既存のシステムと変わらない。

## 2.19

## 2.20 References

- 1) 太田勝敏編著:『交通まちづくり』の展開と課題, 方向性, IATSSReview (国際交通安全学会誌) 33(2), 2008
- 2) 原田昇, 羽藤英二, 高見淳史: 交通まちづくり ー地方都市からの挑戦ー, 鹿島出版社, 2015
- 3) 魏蜀楠: 佐世保市『地方型』MaaS の導入可能性に関する政策研究, 2022.
- 4) 藤本 直樹: わが国における「地方型 MaaS」の推進に向けた政策の方向性と課題についての考察, 情報経営, 2021, 81 巻, p. 113-116
- 5) 後藤 大: MaaS の推進と法規制, 保険学雑誌, 2021, 2021 巻, 653 号, p. 653\_67-653\_75
- 6) 加藤博和, 福本雅之: 地域参画型公共交通サービス供給の成立可能性と持続可能性に関する実証分析ー「生活バスよっかいち」を対象としてー, 土木学会論文集 D/65 巻 4 号, 2009
- 7) 喜多秀行: 社会的共通資本としての地域公共交通サービスの計画方法論, 土木計画学研究発表会 (春大会) 2022
- 8) Sonja Hekkila: Mobility as a Service -A Proposal for Action for the Public Administration, Case Helsinki, 2014
- 9) 太田和博: 地域交通政策の意思決定における住民参画の意義と課題, 運輸と経済第 69 巻 12 号, 2009
- 10) 太田和博: 公平性にかなう地域交通政策の策定システム, 三田商学研究 Vol.43 No.3 P.187-, 2000
- 11) 森栗茂一: 交通計画における住民協働の有効性と展開手法, 運輸と経済第 69 巻 12 号, 2009
- 12) ハンス・ケルゼン: 民主主義の本質と価値, 長尾隆一・植田俊太郎, 岩波文庫, 2015, 1929
- 13) M. Kamargianni, Melinda Matyas: The Business Ecosystem of Mobility-as-a-Service, Computer Science, 2017
- 14) Sochor, J., Arby, H., Karlsson, M., Sarasini, S.: A topological approach to Mobility as a Service: A proposed tool for understanding requirements and effects, and for aiding the integration of societal goals. Proc. ICoMaaS – 1st International Conference on Mobility as a Service, 2017.
- 15) 牧原出: 行政改革と調整のシステム, 行政学叢書, 東京大学出版会, 2007
- 16) 杉谷和哉: 政策にエビデンスは必要なのか-EBPM と政治のあいだ-, ミネルヴァ書房, 2022

(Received June 14, 2024)

(Accepted ????,??,????)

Can “Japanese-style MaaS” be seen as a type of

'Transport Town Planning'?

A study of goals and governance

Ukyo NAGATA

In the area of regional public transport policy, the field known as Mobility as a Service (MaaS) has rapidly emerged in recent years. Even before that, the idea of 'Transport Town Planning' had existed and developed as a system of thought that 'promotes town development through collaboration between citizens and government based on addressing regional issues related to transport'.

In this paper, we examine all the cases of “Japanese-style MaaS” in fiscal 2020 to see whether the ideals of “Transport Town Planning” are being realized and whether the unilateralism of public transport operators is being removed, with reference to the Basic Act on Transport Policy, which is a legal philosophy related to transport, in the progress of MaaS policy. As a result, it was revealed that many of the plans did not take into account the maintenance and management of the functions that are emphasized in transport policy. Furthermore, there were very few cases where the construction of the project and the detailed examination of the project through citizen participation were incorporated into the project. Based on these results, we concluded that the Japanese version of MaaS can be considered as 'Transport Town Planning' in the conceptual stage, but not in the implementation stage.

地域公共交通政策における「連携」と「共創」の再設計：認知バイアスと制度的調整の統合的アプローチ

## 2.21 Re-designing “Collaboration” and “Co-creation” in Regional Public Transport Policy: Integrated Approach to Cognitive Biases and Institutional Coordination

永田 右京\*

Ukyo Nagata

著者所属 \*岩手県立大学総合政策学部 〒020-0693 岩手県滝沢市菓子152-52 s246w003@s.iwate-pu.ac.jp

### 2.21.1 Abstract

This study develops a computational modeling framework to analyze cognitive bias effects in collaborative transport policy implementation and proposes institutional design principles for effective stakeholder coordination. Using cooperative robot control theory, this study models stakeholder interactions in policy networks where each agent represents a policy stakeholder with specific capabilities and positions. Through systematic simulation experiments, this research analyzes how three key cognitive biases—confirmation bias, status quo bias, and narrow framing bias—affect coordination effectiveness in collaborative governance. The computational analysis reveals that confirmation bias can paradoxically enhance coordination when properly leveraged, while status quo bias creates threshold effects requiring careful institutional management. Narrow framing bias consistently degrades system performance through reduced global optimization. Based on these findings and analysis of implementation gaps in existing collaborative frameworks, this study proposes a three-tier institutional architecture that separates

political-administrative and administrative-business interfaces to optimize both democratic accountability and operational efficiency. This study contributes to policy science by providing theoretically grounded institutional design principles derived from computational analysis of cognitive mechanisms underlying stakeholder coordination in public transport governance.

**Keywords:** Public transport policy, Collaboration, Institutional design, Cognitive bias, ill-posed problem

## 2.21.2 1. Introduction

### 1.1 Research Background and Problem Statement

Regional public transport policy in Japan has undergone significant transformation since the 2002 deregulation, evolving from a primarily regulatory approach to one emphasizing “collaboration” (連携) and “co-creation” (共創) among diverse stakeholders. This shift reflects broader trends in public governance toward collaborative approaches that seek to harness the expertise and resources of multiple actors—government agencies, transport operators, citizens, and other businesses—to address complex policy challenges.

The institutionalization of this transformation has been further accelerated by recent policy developments. The formulation of regional public transport plans has become a mandatory effort (国土交通省 2022:1 地域公共交通計画と乗合バス等の補助制度の連動化に関する解説パンフレット), with specific guidance emphasizing the establishment of conference bodies to facilitate efficient information exchange and consultation among stakeholders (国土交通省 2025: 5 地域公共交通計画等の作成と運用の手引き 第4版 理念編). This institutional framework provides the formal structure within which collaboration and co-creation processes must operate.

The policy discourse around collaboration and co-creation gained particular prominence following the 2021 policy framework “Research Group on Co-creation of Regional Transport for the Post-Corona Era” by the Ministry of Land, Infrastructure, Transport and Tourism (MLIT). This framework explicitly calls for “co-creative transport” involving three key dimensions: transport operators actively generating human flows to revitalize regional communities, cross-sector collaboration between transport operators and other industries, and community engagement that treats transport as a shared responsibility.

However, despite the policy emphasis on collaboration and co-creation, empirical evidence suggests significant gaps between policy intentions and implementation outcomes. The proliferation of collaborative frameworks has not necessarily translated into more effective or sustainable transport solutions, raising fundamental questions about the conditions under which collaboration and co-creation can function effectively in public transport governance.

This implementation gap is further exacerbated by the structural resource constraints facing local governments in Japan. The 2023 report by the Council of Experts on Decentralization Reform specifically identified an “inverted triangle” burden structure where municipal governments face disproportionately higher implementation costs compared to prefectural and national levels. The report documented several critical resource pressures: “In some cases, a single staff member handles multiple ministry operations, and multiple plans are formulated within one department. In such cases, there is a sense of duplication in the content of these plans and work becomes concentrated due to overlapping planning periods”; “Due to resource shortages, sufficient time cannot be secured for planning-related administrative work, leading to cases where national templates or examples from other municipalities are almost entirely followed to secure funding tied to plans and avoid being publicly identified as non-implementing entities”; and “Even when plan formulation itself is a best-effort obligation and formulation procedures are left to local government discretion, considerable procedural costs such as deliberation council reviews occur to obtain regional consensus” (Council of Experts on Decentralization Reform, 2023: 9-12). These structural constraints create a context where collaborative and co-creative approaches, despite their theoretical appeal, may be undermined by practical implementation



realities.

## 1.2 Policy Context and Evolution

Japanese regional public transport policy has evolved through distinct phases, each reflecting different approaches to the balance between market mechanisms and public intervention:

**2002 Deregulation:** Shifted route operations from permission-based to notification-based system and changed subsidy allocation from company-wide to route-specific support, introducing market competition while maintaining essential services through targeted subsidies.

**2006 Regional Public Transport Conferences:** Established formal collaborative governance mechanisms, enabling community bus operations and providing platforms for stakeholder participation in transport planning.

**2010 Life Transport Survival Project:** Institutionalized a competitive approach where only “striving regions” were permitted to maintain living transport services, emphasizing local initiative and self-reliance.

**2021 Co-creation Framework:** Explicitly embraced co-creation as the central organizing principle, encompassing resource mobilization (“total mobilization of transport resources”), joint management (direct operator collaboration), and service optimization (“bundle and reduce” approaches).

## 1.3 Empirical Evidence of Implementation Gaps

To assess the practical implementation of collaboration and co-creation principles, this study conducted comprehensive analysis of Japan’s MaaS (Mobility as a Service) initiative, which represents a concrete manifestation of co-creation policy. Analysis of all 38 projects approved under the FY2020 Japan MaaS initiative reveals significant discrepancies between policy ideals and implementation reality.

Regarding business performance indicators, 92% of projects (35/38) included metrics focused on usage rates, revenue generation, or operational efficiency. In stark contrast, for non-business indicators, only 29% of projects (11/38) incorporated indicators addressing social impact, accessibility improvement, or environmental outcomes. Most concerning was the level of citizen participation mechanisms, where merely 5% of projects (2/38) established meaningful mechanisms for citizen participation in project governance and evaluation.

These findings indicate that collaborative frameworks may be captured by transport operator interests rather than serving broader public purposes, with minimal achievement of genuine democratic participation.

## 1.4 Research Objectives and Approach

This study addresses three primary research questions. First, the policy implementation gap is examined by investigating to what extent collaborative transport policies achieve their stated objectives of moving beyond business-focused metrics to incorporate broader social and democratic values. Second, coordination mechanisms are analyzed by exploring how cognitive biases and stakeholder characteristics affect the effectiveness of collaborative coordination in transport policy implementation. Third, institutional design is addressed by determining what institutional arrangements can optimize the benefits of collaboration while mitigating its inherent limitations and coordination failures.

This research employs an integrated approach combining literature review of empirical policy evaluation (Nagata, 2024), computational modeling using cooperative robot control theory, and institutional design theory. This approach contributes to policy science by synthesizing existing empirical evidence of implementation gaps with theoretical insights into how cognitive biases affect stakeholder coordination effectiveness.

### 2.21.3 2. Theoretical Framework and Literature Review

#### 2.1 Collaborative Governance Theory

Collaborative governance has emerged as a dominant paradigm in public administration, defined by Ansell and Gash (2008) as “a governing arrangement where one or more public agencies directly engage non-state stakeholders in a collective decision-making process that is formal, consensus-oriented, and deliberative and that aims to make or implement public policy or manage public programs or assets.”

In the context of transport policy, collaborative approaches have been promoted as solutions to the limitations of traditional regulatory frameworks. Kato et al. (2009) identify four critical conditions for successful community-participatory public transport: shared recognition and responsibility distribution among stakeholders, mutual benefit for all participants, presence of key coordinators, and connection between stakeholder efforts and service improvement outcomes.

However, collaborative governance faces inherent challenges. Emerson et al. (2012) note that collaboration requires significant transaction costs, may lead to lowest-common-denominator solutions, and can be captured by well-organized interests. These challenges are particularly acute in transport policy, where technical complexity, regulatory constraints, and commercial interests create additional coordination difficulties.

#### 2.2 Co-creation in Public Service Delivery

Co-creation represents an evolution beyond traditional collaboration, emphasizing joint value creation between public agencies and stakeholders (Voorberg et al., 2015). In transport policy, co-creation manifests in several forms. Resource mobilization involves utilizing non-traditional transport resources such as vehicles, drivers, and infrastructure from diverse sources. Joint operations encompass formal partnerships between transport operators through joint ventures and shared services that enable greater coordination. Service integration focuses on bundling transport with other services to create comprehensive mobility solutions that address multiple user needs simultaneously.

The Japanese policy framework identifies three dimensions of co-creation: transport operators actively generating community vitalization, cross-sector collaboration for problem-solving, and community ownership of transport services. However, empirical studies reveal mixed results. The Kyushu Regional Transport Bureau (2023) found that while co-creation enables previously impossible route reorganizations and improves service efficiency, business performance improvements remain limited, and building necessary trust relationships requires significant time investment.

#### 2.3 Empirical Evidence on Collaboration and Co-creation Effects

**2.3.1 Collaboration Effects** Kato et al. (2009) analyzed community-participatory regional public transport and identified four existence requirements, though they did not address the elements needed to realize these conditions. These requirements include ensuring that related stakeholders share recognition and responsibility distribution, that each stakeholder can benefit from this participation, that key persons exist to coordinate stakeholders, and that stakeholder efforts connect to usage promotion and value enhancement.

Kita (2006) positively evaluated the provision of collaboration and cooperation venues in comprehensive coordination plan formulation from a technical personnel development perspective, noting that “providing venues and opportunities for various transport-related stakeholders to collaborate while receiving support from researchers and technical experts to improve planning skills... is precisely the policy that the state should implement.”

**2.3.2 Co-creation Effects** The Kyushu Regional Transport Bureau (2023) conducted comprehensive interviews and analysis of co-creation across Kyushu regions, identifying both effects and challenges.

Regarding co-creation effects, the analysis found that previously unimplementable route re-organizations became feasible, and both convenience maintenance/improvement and transport efficiency improvement were achieved simultaneously. However, significant co-creation challenges also emerged. Business performance improvements remain limited, and simply appealing transport operators' difficulties cannot enable municipal support. Building trust relationships necessary for collaboration requires considerable time investment, and few opportunities exist for discussing role distribution, cost burden, and community development between transport operators and other private businesses and administration.

Nomura (2023) analyzed regional transport security through Limited Liability Partnership (LLP) in Ichinohe Town, where municipalities and transport operators participate as parallel LLP members, reducing registration burden and responsibility distribution costs, positioning this as a co-creation example.

Yoshida (2021) implemented inter-operator collaboration in Hachinohe City (2007), reducing 200 services while achieving equal intervals, resulting in increased ridership and profitability.

## 2.4 Cooperative Control Theory and Social Systems

### Introduction

Cooperative control theory provides a framework for coordinating multiple autonomous agents to achieve shared objectives, forming the backbone of multi-agent systems (MAS). This theory has evolved to address complex challenges in automation, robotics, and increasingly, the analysis of social systems and collaborative governance.

### Core Concepts and Mechanisms

Cooperative control theory addresses several key problems in multi-agent coordination. Consensus mechanisms ensure all agents agree on certain variables or states, which is fundamental for group coordination in MAS (Gulzar et al., 2018; Ying et al., 2022). Formation and containment problems involve arranging agents in specific patterns or ensuring they remain within certain boundaries (Ying et al., 2022; Anand et al., 2024; Briñón-Arranz et al., 2014). Resource allocation and coverage focus on distributing tasks or resources efficiently among agents (He et al., 2023; Ying et al., 2022), while flocking and connectivity problems address the maintenance of group cohesion and communication links (Ying et al., 2022).

*Theoretical Foundations:* Cooperative control leverages graph theory to model agent interactions, Lyapunov theory for stability, and distributed optimization for performance (Hengster-Movrić & Lewis, 2014; Anand et al., 2024; Briñón-Arranz et al., 2014). Algorithms are designed to work with only local information, enabling scalability and robustness in large systems (Hengster-Movrić & Lewis, 2014; Yang et al., 2023).

### Methodologies and Algorithmic Advances

The field has developed sophisticated methodological approaches to multi-agent coordination. Distributed control protocols enable agents to make decisions based on local neighbor information while supporting global objectives without centralized oversight (Hengster-Movrić & Lewis, 2014; Anand et al., 2024). Adaptive and learning-based control methods increasingly utilize neural networks and reinforcement learning to handle uncertainties and dynamic environments, allowing agents to adapt and optimize their behavior collaboratively (Shi & Shen, 2015; Lan et al., 2023). Game theory integration has been applied in scenarios such as traffic signal control, where agents must balance individual and collective goals, thereby enhancing decision-making in dynamic, competitive environments (Abdoos, 2020).

### Extensions to Social Systems and Collaborative Governance

Cooperative control theory has been extended to various social and governance applications. Urban traffic management represents one successful application area where multi-agent cooperative

control, combined with game theory and reinforcement learning, has been applied to optimize traffic signals across networks, reducing congestion and improving flow through agent collaboration (Abdoos, 2020). Human-robot and human-agent collaboration applications demonstrate how cooperative control underpins intelligent interaction and collaboration in mixed human-machine teams, supporting decision-making and task allocation in complex social environments (Yang et al., 2023). In collaborative governance contexts, the principles of consensus, distributed decision-making, and adaptive coordination in MAS are increasingly informing models of collaborative governance, where multiple stakeholders must align on shared policies or actions (He et al., 2023; Yang et al., 2023).

### **Challenges and Future Directions**

Several challenges remain in the advancement of cooperative control theory. Scalability and complexity issues arise as systems grow, making it increasingly difficult to ensure stability, optimality, and real-time performance (He et al., 2023; Yang et al., 2023). Uncertainty and non-linearity present ongoing challenges in handling unknown, state-dependent effects and nonlinear dynamics, though adaptive and learning-based methods show promise in addressing these issues (Shi & Shen, 2015; Lan et al., 2023). Integration with social systems represents a particularly complex challenge, as translating cooperative control principles to human-centric domains requires addressing issues of trust, communication, and heterogeneous agent capabilities (Yang et al., 2023; Abdoos, 2020).

**Summary** Cooperative control theory is central to the coordination of multi-agent systems, addressing problems like consensus, formation, and resource allocation through distributed, adaptive, and learning-based methods. Its extension to social systems and collaborative governance leverages these principles to manage complex, dynamic, and decentralized environments, with ongoing research focused on scalability, uncertainty, and integration with human factors.

## **2.5 Empirical Evidence of Implementation Gaps: Japan MaaS Policy Analysis**

To assess the practical implementation of collaboration and co-creation principles, this study draws on recent empirical research on Japan’s MaaS (Mobility as a Service) initiative. Japan MaaS represents a concrete manifestation of co-creation policy, explicitly designed to “solve regional challenges through transport improvement” and requiring municipal involvement in the application process.

Japan MaaS is a public policy initiative that applies the MaaS concept—integrating transport-related services to provide them as a unified service—to address regional challenges. Unlike purely commercial MaaS implementations, Japan MaaS adopts a “regional problem-solving through transport improvement” stance, requiring municipal involvement in the application process. The approach incorporates cross-industry collaboration thinking, making it closely aligned with “co-creation” principles.

Previous research by Nagata (2024) analyzing all 38 Japan MaaS projects revealed significant implementation gaps in collaborative governance. The analysis demonstrated that 92% of projects focused primarily on business indicators, only 29% incorporated non-business indicators addressing social impact or accessibility, and merely 5% established meaningful citizen participation mechanisms. These findings indicate that collaborative frameworks may be captured by business interests rather than serving broader public purposes, with minimal achievement of genuine democratic participation.

This pattern reflects broader challenges in collaborative governance, including organized interest capture, technical complexity barriers to citizen participation, and institutional path dependence that constrains innovation in governance approaches (Nagata, 2024). These findings motivate the need for more sophisticated understanding of how stakeholder coordination actually functions in practice, leading to the computational modeling approach examining the cognitive mechanisms underlying these coordination failures.

## 2.6 Main Cognitive Biases in Collaborative Transport Policy Implementation

### Key Cognitive Biases

Several cognitive biases significantly influence collaborative transport policy implementation. Loss aversion manifests when decision-makers and stakeholders tend to value existing assets or policies more highly than potential gains from new initiatives, leading to resistance to change and preference for the status quo (Watkins & Musselwhite, 2025). Status quo bias reflects a systematic preference for maintaining current conditions, making it difficult to implement innovative or disruptive transport policies (Watkins & Musselwhite, 2025). Risk perception errors occur when misjudgments about the likelihood or impact of risks skew policy priorities, often leading to overly cautious or misaligned decisions (Watkins & Musselwhite, 2025). Confirmation bias emerges when stakeholders seek or interpret information in ways that confirm their pre-existing beliefs, reducing openness to alternative perspectives or evidence (Rastogi et al., 2020). Anchoring bias occurs when initial information or assumptions disproportionately influence subsequent judgments, even when new, more relevant data becomes available (Rastogi et al., 2020). Additionally, case-specific, environmental, and human-nature biases include those arising from specific data, context or culture, and universal cognitive tendencies, all of which shape how transport policies are evaluated and implemented (Watkins & Musselwhite, 2025).

### Mechanisms of Influence

Cognitive biases operate through several distinct mechanisms in policy settings. Information processing is affected by cognitive biases that influence how stakeholders perceive, interpret, and prioritize information, often leading to selective attention or misinterpretation of data relevant to transport policy (Acciarini et al., 2020; Rastogi et al., 2020; Watkins & Musselwhite, 2025). Stakeholder interaction is complicated by biases that can hinder the achievement of cognitive consensus in multi-stakeholder settings, as different groups may frame problems and solutions according to their own interests and mental models (Knoppen et al., 2021). Decision-making dynamics are influenced by biases such as loss aversion and status quo bias that can slow down or block policy innovation, while anchoring and confirmation bias can entrench initial positions, making collaborative negotiation and adaptation more difficult (Rastogi et al., 2020; Watkins & Musselwhite, 2025; Knoppen et al., 2021).

### Constructive and Detrimental Effects

Cognitive biases demonstrate both detrimental and potentially constructive effects in policy settings. Most cognitive biases are detrimental in collaborative policy settings, leading to sub-optimal decisions, resistance to necessary change, and failure to achieve consensus or system-level optimization (Watkins & Musselwhite, 2025; Knoppen et al., 2021). For example, loss aversion and status quo bias can prevent the adoption of beneficial transport reforms (Watkins & Musselwhite, 2025). However, constructive effects can emerge when structured stakeholder engagement and knowledge exchange help uncover and address underlying assumptions, sometimes shifting priorities and improving decision quality (Knoppen et al., 2021). Awareness and mitigation strategies, such as education, training, and iterative decision processes, can reduce the negative impact of biases and foster more effective collaboration (Watkins & Musselwhite, 2025; Knoppen et al., 2021).

### Evidence from the Literature

Empirical research provides evidence for both the challenges and potential solutions related to cognitive biases in policy settings. Studies demonstrate that structured, iterative stakeholder participation can help overcome cognitive biases by promoting knowledge exchange and revealing hidden assumptions, leading to more balanced and system-oriented policy outcomes (Knoppen et al., 2021). Comprehensive reviews highlight the prevalence and impact of biases such as loss aversion, status quo bias, and risk perception errors in transport policy, emphasizing the need for multidisciplinary approaches to identify and mitigate these effects (Watkins & Musselwhite, 2025).

Cognitive biases like loss aversion, status quo bias, and risk perception errors significantly influ-

ence collaborative transport policy implementation, often impeding innovation and consensus. However, structured stakeholder engagement and awareness strategies can help mitigate these effects and improve policy outcomes.

## 2.7 Key Principles of Institutional Design for Effective Collaborative Governance in Public Transport Policy

### Theoretical Frameworks

Several theoretical frameworks inform institutional design for collaborative governance in public transport. Network-based collaborative governance demonstrates that effective collaborative governance often relies on network structures that facilitate cross-sector and interorganizational collaboration, operating at multiple levels—collaborative, financial, and interorganizational governance—each with distinct structures, processes, and orchestrators to address project complexity and stakeholder diversity (Hu et al., 2024). Inclusiveness and interdependence considerations indicate that institutional design should ensure broad stakeholder inclusion while managing the degree of interdependence among participants, as inclusive arrangements where participants are not overly interdependent tend to empower stakeholders and enhance their perceived ability to influence policy outcomes (Fossheim & Andersen, 2022). Accountability frameworks reveal that accountability in collaborative institutions is shaped by the specificity of rules and the delegation of decision-making authority, where more detailed rules can enhance process accountability, while greater institutional independence can support legitimacy and flexibility (Mancheva et al., 2023). Good governance principles establish that transparency, accountability, stakeholder engagement, and strong leadership are foundational for effective institutional coordination and collaborative governance in public transport (Bouraima et al., 2023; Adjil et al., 2023).

### Empirical Findings

Empirical research provides evidence for key principles of effective institutional design. Communication and stakeholder engagement studies show that regular dialogue, consultation, and inclusive participation foster cooperation, trust, and a sense of ownership among stakeholders, which are critical for integrated planning and effective policy implementation (Adjil et al., 2023; Paulsson et al., 2018). Research on power to influence policy indicates that institutional arrangements that are inclusive and reduce interdependence among participants provide greater opportunities for stakeholders to influence policymaking, leading to more innovative and accepted solutions (Fossheim & Andersen, 2022). Studies on balancing efficiency, accountability, and coordination reveal multiple dimensions of institutional effectiveness. Regarding efficiency, streamlined frameworks and clear governance structures help reduce fragmentation and improve the implementation of transport initiatives (Bouraima et al., 2023; Hu et al., 2024). For accountability, transparent processes and well-defined roles enhance trust and ensure that collaborative arrangements remain answerable to both participants and the public (Adjil et al., 2023; Mancheva et al., 2023). Concerning coordination, multi-level governance structures and regular stakeholder engagement enable better alignment of objectives and resources across diverse actors (Hu et al., 2024; Paulsson et al., 2018).

**Challenges:** Lack of political will, corruption, inadequate participation, and poor vision are major barriers to effective institutional coordination. Implementing good governance principles is identified as the top strategy to overcome these challenges (Bouraima et al., 2023).

### Table 1: Summary Table: Principles and Empirical Insights

(Table1 参照)


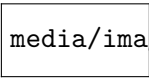
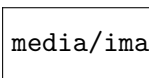
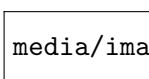
Effective institutional design for collaborative governance in public transport policy requires inclusive participation, clear accountability, multi-level coordination, and strong stakeholder engagement, supported by transparent and context-sensitive frameworks. These principles help balance efficiency, accountability, and coordination among diverse stakeholders.



2.21.4 3. Computational Modeling Framework

3.1 Cooperative Control Model Structure

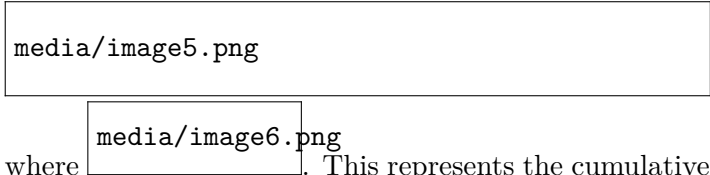
This study adapts Yoshihara et al.’s (2009) cooperative robot control theory to model stakeholder interactions in policy implementation. An N-joint robot arm represents a policy network where each joint  $i$  corresponds to a policy stakeholder with state variables:

- Position vector: 
- Velocity vector: 
- Joint angle: 
- Link length: 

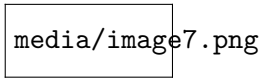
This mapping allows us to represent policy stakeholders as autonomous agents with specific capabilities (link lengths) and policy positions (joint angles) who must coordinate to achieve collective policy objectives.

3.2 Kinematic Model

Each joint’s position depends on previous joints’ positions and cumulative joint angles:

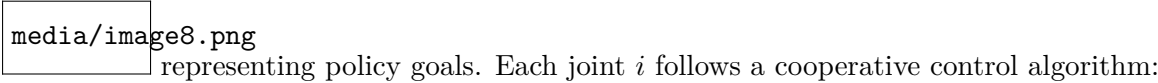


Velocity is computed as the time derivative of position:

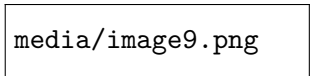


3.3 Control Objectives and Coordination Mechanisms

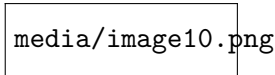
The system’s objective is to guide the end-effector (final stakeholder) to track a target position



4.3.1 Target Direction Vector Calculation Direction vector from end-effector to target:

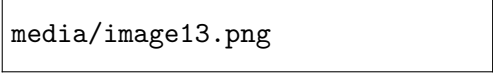


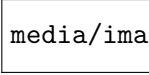
Normalized perpendicular vector:



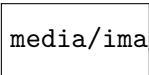
**4.3.2 Target Velocity Decomposition** Target velocity  is decomposed into each joint's motion direction:



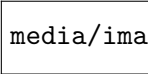


**4.3.3 Coordination Coefficient Calculation** Each joint's coordination coefficient  is calculated based on the difference between its target velocity and subsequent joints' actual velocities:

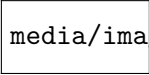


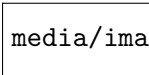
where  are small positive constants for numerical stability.

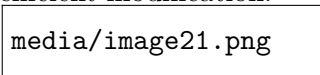
### 3.4 Cognitive Bias Integration

**4.4.1 Status Quo Bias** Status quo bias  is incorporated as resistance to joint angle changes:

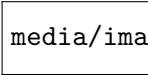



where  is the angular velocity without bias.

**4.4.2 Confirmation Bias** Confirmation bias  is incorporated as coordination coefficient modification:

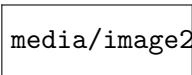


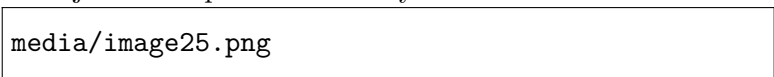
Positive confirmation bias represents overconfidence in self-judgment, while negative values represent excessive self-doubt.

**4.4.3 Narrow Framing Bias** Narrow framing bias  is incorporated as neglect of global optimization:



### 3.5 Cooperative Velocity Calculation

Each joint's cooperative velocity  is calculated as:



3.6 Joint Angle Updates

Final joint angle updates are based on the inner product of cooperative velocity and target velocity:

media/image26.png

media/image27.png

where is the control gain.

3.7 Performance Metrics

System performance is evaluated using the following metrics:

4.7.1 Accuracy × Distance

media/image28.png

media/image29.png

where is cumulative movement distance.

4.7.2 Energy Efficiency

media/image30.png

4.7.3 Joint Activity

media/image31.png

4.7.4 Joint Smoothness

media/image32.png

3.8 Simulation Implementation Process

(Figure 1 参照) *Figure 1: Comprehensive simulation process flow showing the six-stage methodology for analyzing stakeholder coordination through robot arm control. The process includes initial setup, bias application, control execution, performance evaluation, statistical analysis, and policy implications derivation. This framework enables systematic analysis of 1,820 experiments with randomized conditions.*

2.21.5 4. Experimental Design

4.1 Experimental Overview

To quantitatively evaluate the impact of cognitive biases on coordination performance, this study conducted a comprehensive experimental program totaling 1,820 runs. The experimental design included baseline experiments with 100 runs without bias, confirmation bias experiments comprising 11 conditions with 40 runs each (440 experiments total, intensity 0.0-1.0, step 0.1), status quo bias experiments involving 21 conditions with 40 runs each (840 experiments total, intensity 0.0-1.0, step 0.05), and narrow framing experiments consisting of 11 conditions with 40 runs each (440 experiments total, intensity 0.0-1.0, step 0.1).

## 4.2 Bias Experiments

To avoid deterministic results, this study implemented comprehensive experimental design with randomized initial conditions:

(Figure 2 参照) *Figure 2: Four key animation scenarios demonstrating how different cognitive biases affect stakeholder coordination patterns. Each scenario shows distinct joint configurations and target-reaching behaviors: (a) normal operation without bias, (b) confirmation bias promoting policy goal focus, (c) status quo bias maintaining existing approaches, and (d) narrow framing limiting system-wide perspective.*

Each experiment incorporated multiple randomization parameters to ensure statistical validity. Initial joint angles were varied within  $\pm 0.1$  radians, target positions were varied around the center by  $\pm 0.01\text{m}$  with radius variation of  $\pm 0.005\text{m}$ , control gains were varied by  $\pm 0.5$ , and switching thresholds were varied by  $\pm 0.002$ .

## 4.3 Statistical Analysis Methods

Statistical analysis employed multiple complementary methods to ensure robust interpretation of results. Analysis of Variance (ANOVA) was used to test for statistical significance of bias effects across different experimental conditions. Confidence intervals were calculated to estimate effect sizes and provide measures of uncertainty around mean performance values. Regression analysis was conducted to verify linear relationships between bias intensity and coordination performance metrics.

## 2.21.6 4. Experimental Results

### 4.1 Statistical Analysis Results

**5.1.1 Confirmation Bias** Analysis of confirmation bias revealed no significant effect on coordination performance ( $F=0.838$ ,  $p=0.602$ ). Average performance remained stable at  $0.970 \pm 0.002$  across all intensity levels, suggesting a constructive role for this bias type.

**5.1.2 Status Quo Bias** Analysis of status quo bias demonstrated a significant effect on coordination performance ( $F=1.593$ ,  $p=0.044$ ). A notable threshold effect occurred around intensity 0.2, with performance degradation beyond this level. These findings have important institutional design implications, emphasizing the importance of threshold management in policy coordination systems.

**5.1.3 Narrow Framing** Analysis of narrow framing bias showed a significant negative effect on coordination performance ( $F=1.985$ ,  $p=0.028$ ). The results demonstrated linear degradation with  $R^2=0.204$  and a slope of  $-0.00107$ , indicating overall system performance decline as narrow framing intensity increased.

## 4.2 Comprehensive Bias Effect Analysis

(Figure 3 参照) *Figure 3: Constructive Effect of Confirmation Bias*

(Figure 4 参照) *Figure 4: Threshold Effect of Status Quo Bias*

(Figure 5 参照) *Figure 5: Linear Degradation of Narrow Framing Bias*

## 4.3 Joint-Specific Analysis

Analysis of how biases affect different stakeholder positions within the policy network revealed:

**5.3.1 Base Joint (Joint 0) Analysis** The base joint representing political and citizen participation level showed stabilizing effects from confirmation bias.

**5.3.2 End-Effector Joint (Joint 7) Analysis** The end-effector joint representing business operator level showed the most pronounced effects from status quo bias.

## 5.4 Summary of Experimental Results

Through 1,820 experiments, this analysis obtained the following findings:

1. **Confirmation Bias:** Constructive effect maintaining coordination performance
2. **Status Quo Bias:** Sharp change around threshold 0.2
3. **Narrow Framing:** Linear performance degradation

## 4.4 Statistical Analysis

This study conducted comprehensive statistical analysis on 1,820 experiments (100 baseline + 440 confirmation bias + 840 status quo bias + 440 narrow framing bias) with randomized initial conditions to ensure statistical validity. ANOVA tests revealed distinct patterns for each bias type:

**Confirmation Bias:** No significant effect on performance ( $F = 0.838$ ,  $p = 0.602$ ), supporting the hypothesis that confirmation bias can be constructively leveraged. Performance remained stable across all intensity levels (0.0-1.0), with mean accuracy ranging from 0.970 to 0.970.

**Status Quo Bias:** Significant negative effect ( $F = 1.593$ ,  $p = 0.044$ ), confirming the threshold effect around intensity 0.25. Performance degraded sharply beyond this threshold, validating the institutional design recommendations for managing resistance to change.

**Narrow Framing Bias:** Significant linear degradation effect ( $F = 1.985$ ,  $p = 0.028$ ), demonstrating the importance of maintaining system-wide perspectives. Linear regression analysis showed  $R^2 = 0.204$  with a negative slope of  $-0.00107$ .

All experiments included appropriate confidence intervals (95% CI) and effect size calculations. The baseline condition showed mean accuracy of  $0.969 \pm 0.002$ , providing a stable reference point for bias comparisons.

## 4.5 Joint-Specific Analysis

To understand how biases affect different stakeholder positions within the policy network, this study conducted joint-specific analysis focusing on key positions:

(Figure6 参照) *Figure6: Bias effects on the base joint (foundational stakeholder). Shows how fundamental policy actors respond differently to various bias types, with confirmation bias providing stability at the foundation level.*

(Figure7 参照) *Figure7: Bias effects on the end-effector joint (final implementation stakeholder). Demonstrates how biases at the implementation level directly impact overall policy outcomes, with narrow framing showing the most pronounced negative effects.*

The joint-specific analysis revealed that bias effects varied significantly depending on stakeholder position within the policy network. Base joints (foundational stakeholders) showed greater resilience to confirmation bias, while end-effector joints (implementation stakeholders) were more sensitive to narrow framing effects.

## 4.6 Summary of Experimental Findings

(Figure8 参照) *Figure8: Comprehensive summary of bias effects across all experimental conditions with confidence intervals. This overview demonstrates the statistical significance and practical magnitude of each bias type's impact on policy coordination effectiveness.*

## 2.21.7 5. Institutional Design Framework for Effective Collaboration and Co-creation

### 5.1 Theoretical Foundation of Institutional Design

Based on empirical findings from Japan MaaS analysis and computational modeling results, a comprehensive institutional design framework is proposed that addresses the coordination challenges identified in both studies. This framework utilizes institutional design theory while incorporating insights about cognitive bias effects and democratic participation requirements.

Key insights from the analysis indicate that effective collaboration and co-creation require institutional arrangements that enable leveraging the constructive potential of confirmation bias, managing the threshold effects of status quo bias, mitigating negative impacts of narrow framing, and ensuring democratic accountability while enabling operational efficiency.

### 5.2 Three-Tier Institutional Architecture

(Figure9 参照) *Figure9: Three-tier institutional design for effective collaboration. Institutional design for policy coordination that strategically leverages cognitive biases. The upper section shows the basic framework separating quality definition from implementation, while the lower section details specific mechanisms for managing administrative-business relationships through environmental design.*

**6.2.1 Tier 1: Political-Citizen Interface (Democratic Quality Setting)** The first tier, the political-citizen interface, is responsible for ensuring democratic legitimacy and specifying quality standards. At its core, this tier aims to constructively leverage confirmation bias by fostering stakeholder confidence in shared quality objectives, while simultaneously maintaining a broad perspective to avoid narrow local optimization. This is achieved through structured citizen participation processes that define transport service quality objectives, as well as political oversight of administrative implementation supported by clear accountability mechanisms. Public deliberation on the trade-offs between efficiency, accessibility, and sustainability is also central, alongside the establishment of quality specification frameworks that provide clear direction yet retain interpretive flexibility. In practice, this tier is characterized by regular citizen assemblies or deliberative polls on transport priorities, transparent reporting on the achievement of quality objectives, clearly defined boundaries regarding the scope and authority of citizen participation, and integration with broader democratic governance structures.

**6.2.2 Tier 2: Administrative Coordination (Environmental Design)** The second tier, administrative coordination, translates quality objectives into operational frameworks. The guiding principle here is to manage status quo bias through incremental change approaches that remain below disruptive thresholds, while simultaneously developing the expertise necessary for effective multi-stakeholder coordination. This is accomplished by building specialized administrative capacity for stakeholder coordination and conflict resolution, implementing technical evaluation systems that incorporate both business and public value metrics, designing regulatory frameworks that foster competitive environments within quality constraints, and establishing performance monitoring and adjustment systems with effective feedback loops. The organizational structure supporting this tier includes specialized departments for collaborative governance staffed with appropriate expertise, buffer mechanisms to mediate between political and business pressures, structured processes for translating democratic input into operational guidance, and innovation incentives that are closely aligned with public objectives.

**6.2.3 Tier 3: Business-Operations Interface (Autonomous Implementation)** The third tier, the business-operations interface, is tasked with the efficient provision of services



within established quality frameworks. Here, the approach allows confirmation bias to support business confidence in quality-aligned operations, minimizes status quo bias through competitive pressure, and maintains a system-wide perspective via integrated performance measurement. This is realized through competitive processes for service provision that operate within quality constraints, quality-adjusted performance contracts that reward the creation of public value alongside efficiency, collaborative problem-solving processes to address operational challenges, and innovation incentives that promote public value creation. The framework is further supported by performance-based funding that rewards innovation, requirements for cross-sectoral coordination, the use of system-wide performance metrics, and the establishment of information-sharing platforms among stakeholder groups.

### 5.3 Interface Management Mechanisms

The management of interfaces between tiers is crucial for effective institutional design. At the political-administrative interface, the primary challenge is to translate democratic input into operational guidance without undermining administrative expertise or introducing excessive political interference. This is addressed by implementing structured quality specification processes that provide clear direction while maintaining interpretive flexibility. Key mechanisms include the use of quality frameworks that specify outcomes rather than processes, regular review cycles with sunset clauses to manage status quo bias, the preservation of professional administrative autonomy within democratically defined parameters, and transparent accountability mechanisms that enable political oversight without resorting to micromanagement.

At the administrative-business interface, the challenge lies in balancing public objectives with operational efficiency, while also managing the risk that business interests may capture collaborative processes. The solution involves the adoption of competitive frameworks that reward quality achievement alongside efficiency, thereby aligning private interests with public objectives. This is operationalized through environmental design that makes quality-aligned behavior profitable, performance measurement systems that incorporate multiple dimensions of value, competitive pressures that prevent the entrenchment of status quo bias in service provision, and collaborative governance structures that facilitate joint problem-solving while maintaining robust accountability.

## 2.21.8 6. Discussion

### 6.1 Theoretical Implications

The theoretical implications of this study span multiple dimensions. First, a reevaluation of cognitive biases was achieved by revealing constructive aspects of cognitive biases previously viewed negatively. Second, quantitative analysis methods were developed by enabling quantitative analysis through mathematical modeling of policy collaboration. Third, contributions to institutional design theory were made by presenting new theoretical frameworks for strategically leveraging biases in governance arrangements.

### 6.2 Practical Implications

The practical implications encompass several important areas. Regarding application to policy formulation, the findings highlight the importance of policy design that considers cognitive biases as strategic resources rather than obstacles. For stakeholder management, the research demonstrates the value of relationship management methods tailored to specific bias characteristics of different stakeholder groups. Concerning institutional reform, the study indicates the need for gradual improvement strategies for existing institutions that work with rather than against predictable cognitive tendencies.

### 6.3 Research Limitations

This study acknowledges several important limitations. The simulation environment represents a significant constraint, as computational modeling necessarily diverges from actual policy complexity. The bias modeling approach involves simplification of cognitive biases, which are more nuanced and context-dependent in real-world settings than parametric representations suggest. The validation scope is limited to regional public transport in Japan, which may limit generalizability to other policy areas or institutional contexts.

## 2.21.9 7. Conclusion

### 7.1 Research Achievements

This study analyzed collaborative mechanisms in regional public transport policy using robot arm coordination control simulation and produced several key findings. Quantitative evaluation of cognitive biases was accomplished by statistically verifying the effects of three cognitive biases on coordination performance through 1,820 controlled experiments. The constructive effects of certain biases were demonstrated through the coordination performance-maintaining effect of confirmation bias, challenging conventional assumptions about bias uniformly hindering policy implementation. A three-tier institutional design framework was developed that strategically leverages biases rather than simply attempting to eliminate them.

### 7.2 Future Research Agenda

The future research agenda encompasses several critical directions. Empirical research represents a priority, requiring validation of the theoretical framework with actual policy cases to test the practical applicability of the computational findings. Model extension offers opportunities to explore the combined effects of more complex cognitive biases and their interactions in multi-stakeholder environments. International comparison research could examine the applicability of the institutional design framework in different cultural and institutional environments beyond Japan.

### 7.3 Policy Implications

The policy implications of this study offer significant insights for governance practice. Strategic leverage of cognitive biases represents a paradigm shift toward utilizing rather than eliminating biases in institutional design, recognizing that certain biases can enhance rather than hinder coordination effectiveness. Gradual institutional reform emerges as a preferred approach, emphasizing incremental improvement strategies that avoid triggering counterproductive status quo bias responses. Continuous monitoring becomes essential, requiring regular evaluation and adjustment of institutional effects to maintain optimal bias management and stakeholder coordination over time.

## 2.21.10 Acknowledgements

**2.21.11 This document and experiment were supported by Claude Code / Gemini CLI. The idea for this document was inspired by Mr. Yuma Matsuda, CEO of OnGigants Inc, and feedback was received from the Japan Association for Planning and Public Administration.**

## 2.21.12 References

Abdoos, M. (2020). A Cooperative Multiagent System for Traffic Signal Control Using Game Theory and Reinforcement Learning. *IEEE Intelligent Transportation Systems Magazine*, 13, 6-16.

- Acciarini, C., Brunetta, F., & Boccadelli, P. (2020). Cognitive biases and decision-making strategies in times of change: a systematic literature review. *Management Decision*.
- Adj, I., Sumaryadi, I., Djohan, D., & Rowa, H. (2023). Collaborative Governance in the Management of Transportation Modes in DKI Jakarta Province. *Jurnal Ilmiah Ilmu Administrasi Publik*.
- Ansell, C., & Gash, A. (2008). Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory*, 18(4), 543-571.
- Anand, A., Guha, D., & Purwar, S. (2024). Cooperative Formation Control of the Multi-Agent System. 2024 IEEE 4th International Conference on Sustainable Energy and Future Electric Transportation (SEFET), 1-5.
- Bouraima, M., Oyaro, J., Ayyıldız, E., Erdogan, M., & Maraka, N. (2023). An integrated decision support model for effective institutional coordination framework in planning for public transportation. *Soft Computing*, 1-27.
- Briñón-Arranz, L., Seuret, A., & Canudas-De-Wit, C. (2014). Cooperative Control Design for Time-Varying Formations of Multi-Agent Systems. *IEEE Transactions on Automatic Control*, 59, 2283-2288.
- Council of Experts on Decentralization Reform. (2023). 地方分権改革有識者会議 報告書. [In Japanese].
- Emerson, K., Nabatchi, T., & Balogh, S. (2012). An integrative framework for collaborative governance. *Journal of Public Administration Research and Theory*, 22(1), 1-29.
- Fossheim, K., & Andersen, J. (2022). The consequences of institutional design on collaborative arrangements' power to influence urban freight policymaking. *Case Studies on Transport Policy*.
- Grimmelikhuijsen, S. G., Jilke, S., Olsen, A. L., & Tummers, L. (2017). Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review*, 77(1), 45-56.
- Gulzar, M., Rizvi, S., Javed, M., Munir, U., & Asif, H. (2018). Multi-Agent Cooperative Control Consensus: A Comparative Review. *Electronics*, 7, 22.
- He, F., N., Li, X., & Liu, Z. (2023). A Review of Multi-Agent Collaborative Control. 2023 2nd International Conference on Artificial Intelligence, Human-Computer Interaction and Robotics (AIHCIR), 506-510.
- Hengster-Movrić, K., & Lewis, F. (2014). Cooperative Optimal Control for Multi-Agent Systems on Directed Graph Topologies. *IEEE Transactions on Automatic Control*, 59, 769-774.
- Hu, Y., Xu, Q., Field, B., Xia, B., & Wu, G. (2024). Governing collaborative networks in mega transport projects development: Integrative findings from 34 cases worldwide. *Transport Policy*.
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39(1), 17-31.
- Kato, H., Taniguchi, A., & Osuga, K. (2009). Empirical analysis of feasibility and sustainability of community-participatory regional public transport service supply: A case study of "Seikatsu Bus Yokkaichi". *Journal of Japan Society of Civil Engineers*, 65(4), 568-578. [In Japanese].
- Kita, H. (2006). *Development of regional public transport policy*. Seizendo Shoten. [In Japanese].
- Klijn, E. H., & Koppenjan, J. F. M. (2012). Governance network theory: Past, present and future. *Policy & Politics*, 40(4), 587-606.
- Knoppen, D., Janjevic, M., & Winkenbach, M. (2021). Prioritizing urban freight logistics policies: Pursuing cognitive consensus across multiple stakeholders. *Environmental Science & Policy*.
- Kyushu Regional Transport Bureau. (2023). アフターコロナ時代に向けた地域交通の共創に関する研究会 中間とりまとめ. [In Japanese].
- Lan, X., Yan, J., He, S., Zhao, Z., & Zou, T. (2023). Distributed cooperative reinforcement learning for multi-agent system with collision avoidance. *International Journal of Robust and Nonlinear Control*, 34, 567-585.

- Mancheva, I., Pihlajamäki, M., & Keskinen, M. (2023). Institutional accountability: the differentiated implementation of collaborative governance in two EU states. *West European Politics*, 47, 619-644.
- Ministry of Land, Infrastructure, Transport and Tourism. (2022). 地域公共交通計画と乗合バス等の補助制度の連動化に関する解説パンフレット. [In Japanese].
- Ministry of Land, Infrastructure, Transport and Tourism. (2025). 地域公共交通計画等の作成と運用の手引き 第4版 理念編. [In Japanese].
- Nagata, U. (2024). Can “Japan MaaS” be understood as a type of “Transport Town Planning”? A consideration of goals and governance. *Journal of Japan Society of Civil Engineers*, 80(20), 24-20140. <https://doi.org/10.2208/jscej.24-20140> [In Japanese].
- Nomura, Y. (2023). Efforts to secure regional transport using limited liability partnerships in Ichinohe Town. *Transport Policy Studies*, 26(1), 45-52. [In Japanese].
- Ostrom, E. (2005). *Understanding institutional diversity*. Princeton University Press.
- Paulsson, A., Isaksson, K., Sørensen, C., Hrelja, R., Rye, T., & Scholten, C. (2018). Collaboration in public transport planning – Why, how and what?. *Research in Transportation Economics*.
- Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *American Political Science Review*, 94(2), 251-267.
- Rastogi, C., Zhang, Y., Wei, D., Varshney, K., Dhurandhar, A., & Tomsett, R. (2020). Deciding Fast and Slow: The Role of Cognitive Biases in AI-assisted Decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 6, 1-22.
- Russo, J. E., & Schoemaker, P. J. H. (1992). Managing overconfidence. *Sloan Management Review*, 33(2), 7-17.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1), 7-59.
- Shi, P., & Shen, Q. (2015). Cooperative Control of Multi-Agent Systems With Unknown State-Dependent Controlling Effects. *IEEE Transactions on Automation Science and Engineering*, 12, 827-834.
- Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Irwin/McGraw-Hill.
- Watkins, S., & Musselwhite, C. (2025). Recognised cognitive biases: How far do they explain transport behaviour?. *Journal of Transport & Health*.
- Yang, Z., N., & Yao, Y. (2023). A Survey on Cooperative Control of Multi-Agent Systems. 2023 2nd International Conference on Artificial Intelligence, Human-Computer Interaction and Robotics (AIHCIR), 511-515.
- Voorberg, W. H., Bekkers, V. J. J. M., & Tummers, L. G. (2015). A systematic review of co-creation and co-production: Embarking on the social innovation journey. *Public Management Review*, 17(9), 1333-1357.
- Yoshida, T. (2021). Effect analysis of bus route reorganization through inter-operator collaboration in Hachinohe City. *Transport Studies*, 64, 123-130. [In Japanese].
- Yoshihara, Y., Makino, H., Tomita, N., & Yano, K. (2009). Does real-time optimization of joint mobility generate globally optimal arm movements? *Transactions of the Society of Instrument and Control Engineers*, 45(11), 570-577. [In Japanese].
- Y., Liu, Y., Zhao, L., & Zhao, M. (2022). A Review on Cooperative Control Problems of Multi-agent Systems. 2022 41st Chinese Control Conference (CCC), 4831-4836.

### 2.21.13 和文要旨

本研究は、協調的交通政策実装における認知バイアス効果を分析する計算モデリング枠組みを開発し、効果的な利害関係者調整のための制度設計原則を提案する。協調ロボット制御理論を用いて、各エージェントが特定の能力と立場を持つ政策利害関係者を表現する政策ネットワークにおける利害関係者相互作用をモデル化した。体系的シミュレーション実験により、確証バイアス、現

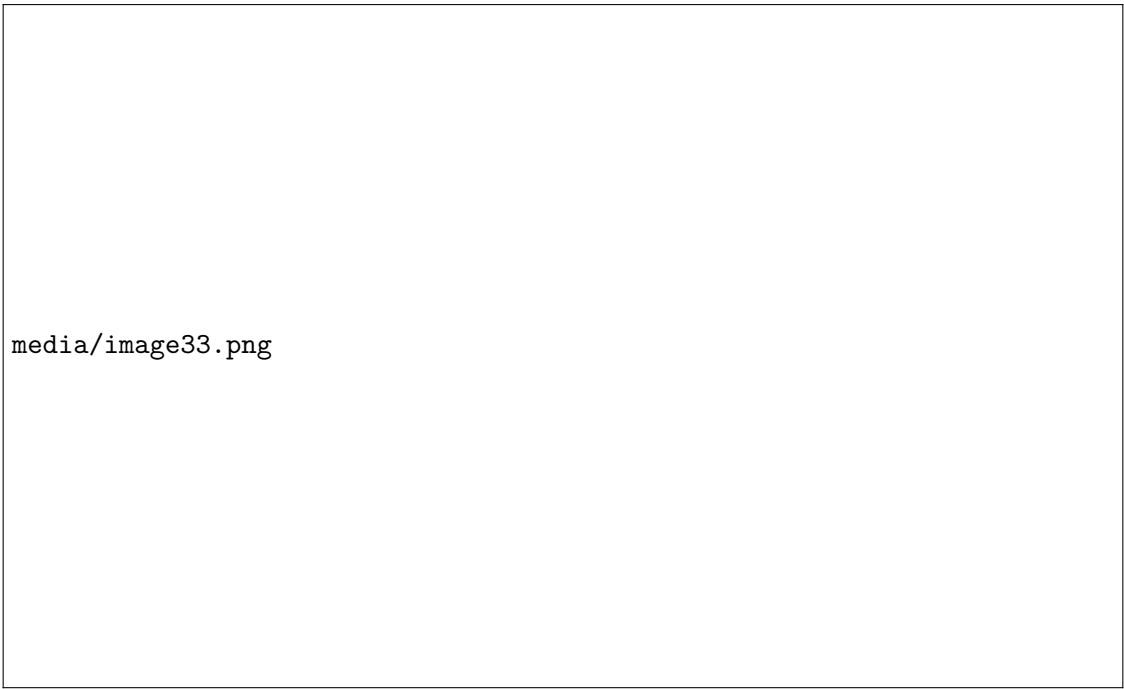


Figure 2.1: Figure 1: Simulation Process Flow for Stakeholder Coordination

状態維持バイアス、狭枠バイアスの3つの主要認知バイアスが協調ガバナンスの調整効果に与える影響を分析した。計算分析により、確証バイアスは適切に活用されれば協調効果を逆説的に向上させ、現状維持バイアスは慎重な制度管理を要する閾値効果を生み出すことが明らかになった。これらの知見に基づき、民主的説明責任と運営効率を最適化する3層制度アーキテクチャを提案する。

**和文キーワード:** 公共交通政策 連携/共創 組織デザイン 認知バイアス 不良設定問題

2.21.14 図表

Table1: Summary Table: Principles and Empirical Insights

||@ >p() \* 0.1849 >p() \* 0.3024 >p() \* 0.5126@

Principle	Mechanism/Framework	Empirical Insight (Citation)
Inclusiveness	Broad stakeholder participation	Increases power to influence policy (Fossheim & Andersen, 2022)
Accountability	Clear rules, transparency, delegation	Enhances trust and legitimacy (Adjil et al., 2023; Mancheva et al., 2023)
Coordination	Multi-level network governance	Improves alignment and integration (Hu et al., 2024; Paulsson et al., 2018)
Efficiency	Streamlined, context-specific frameworks	Reduces fragmentation, speeds implementation (Bouraima et al., 2023; Hu et al., 2024)
Stakeholder Engagement	Regular dialogue and consultation	Builds trust, ownership, and co-operation (Adjil et al., 2023; Paulsson et al., 2018)

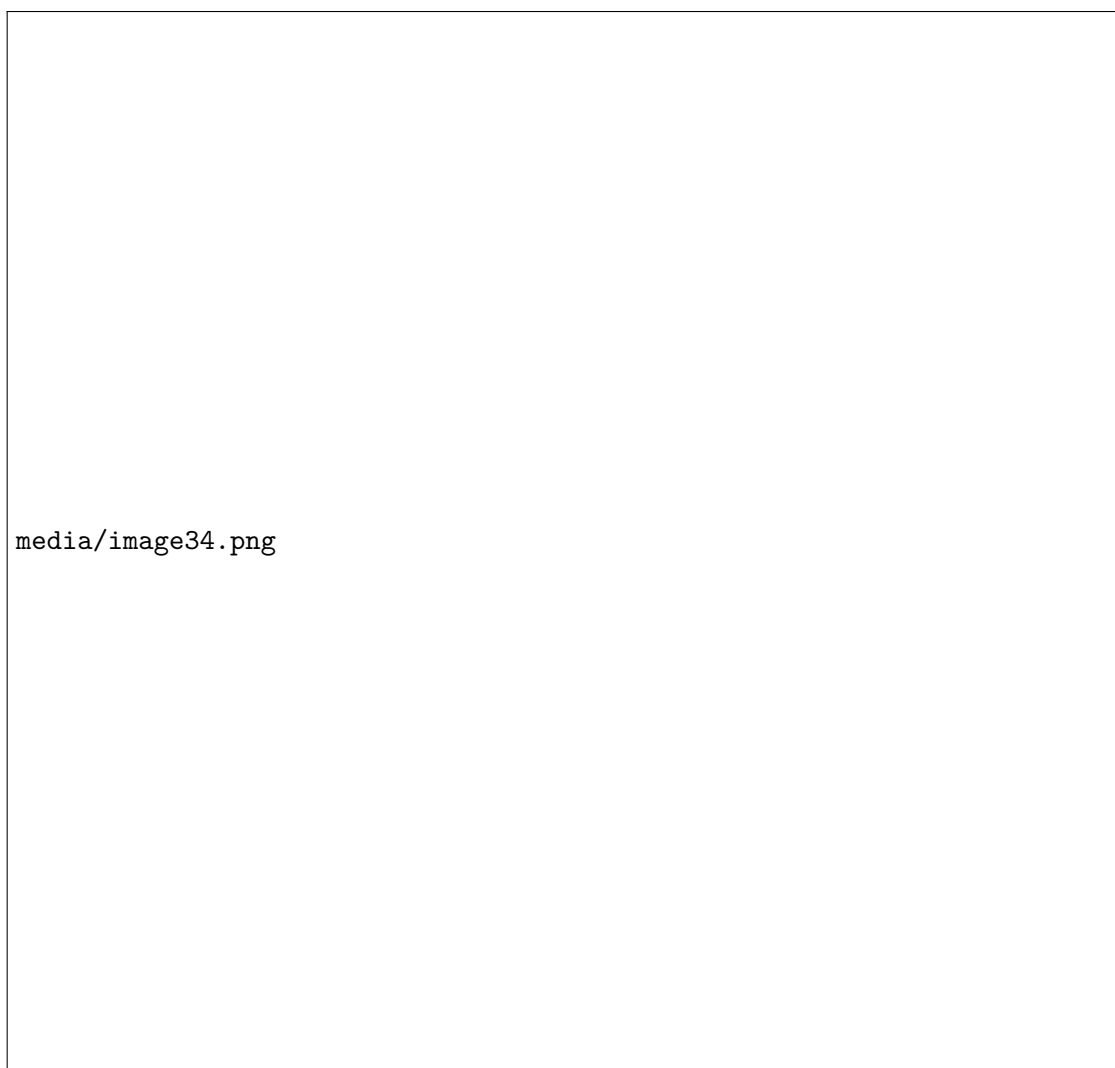


Figure 2.2: Figure 2: Robot Arm Animation Scenarios for Policy Analysis



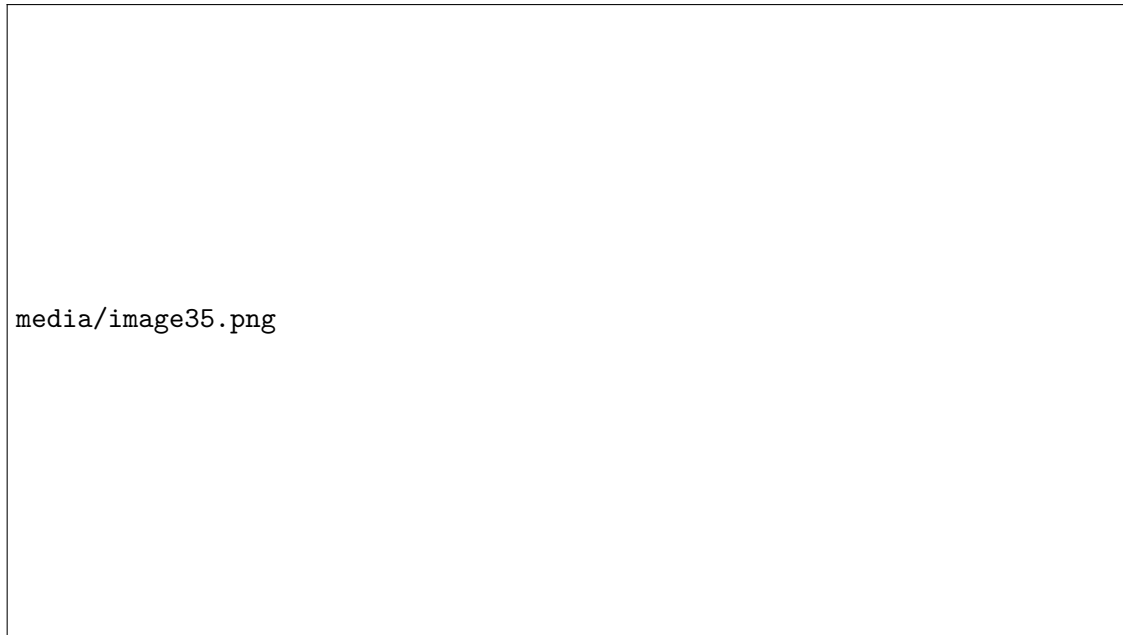


Figure 2.3: Figure 3: Constructive Effect of Confirmation Bias

Figure 1: Simulation Process Flow for Stakeholder Coordination

Figure 2: Robot Arm Animation Scenarios for Policy Analysis

Figure 3: Constructive Effect of Confirmation Bias

Figure 4: Threshold Effect of Status Quo Bias

Figure 5: Linear Degradation of Narrow Framing Bias

Figure6: Joint 0 (Base) Bias Effects

Figure7: Joint 7 (End-Effector) Bias Effects

Figure8: Summary of All Bias Effects

Figure9: Three-Tier Institutional Design for Effective Collaboration

## 2.22 ZK-SNARKs 模倣システムによる政策対話インフラの構築

### 2.22.1 市民討議による原則策定と Constitutional AI 訓練を統合した政策評価アーキテクチャ

---

永田 右京

岩手県立大学大学院 総合政策研究科 博士後期課程 2 年

---

#### 要旨

本研究は、ウィキッド・プロブレム（地球温暖化、少子高齢化等）の政策評価における従来手法の限界を整理し、秘密情報を含む政策提案の実現と専門的判断を両立する新しい枠組みを提案する。本システムの特徴は、市民・研究者による討議を通じて Constitutional AI の憲法的原則を共同策定し、その原則に基づいて訓練された Large Language Model (LLM) が政策評価を行う点にある。Zero-Knowledge Succinct Non-interactive Arguments of Knowledge (ZK-SNARKs) の秘

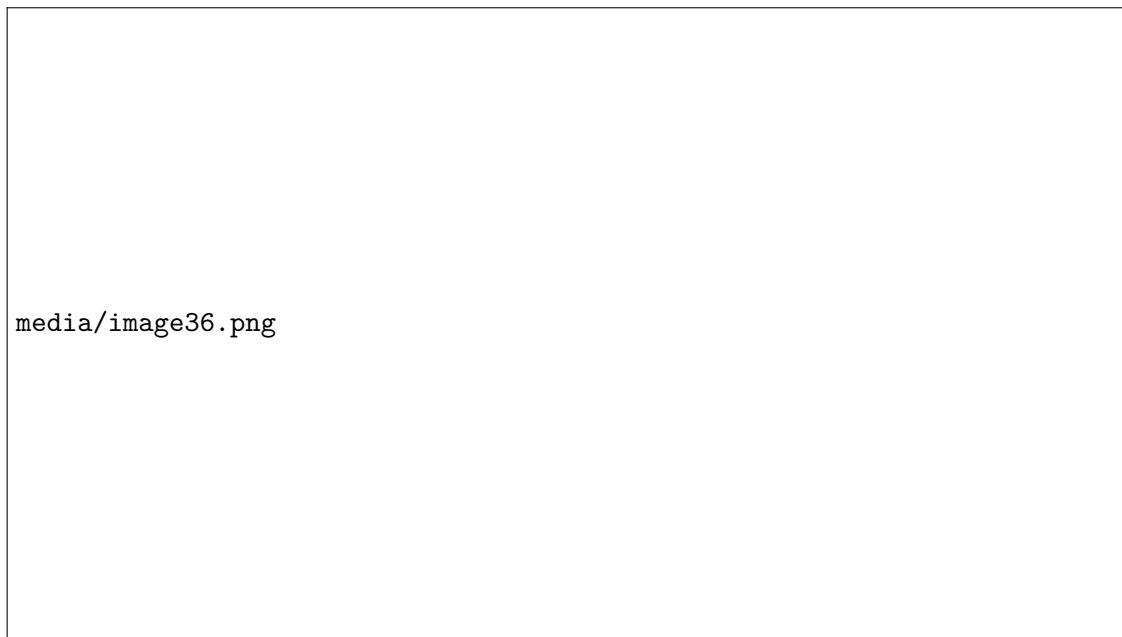


Figure 2.4: Figure 4: Threshold Effect of Status Quo Bias

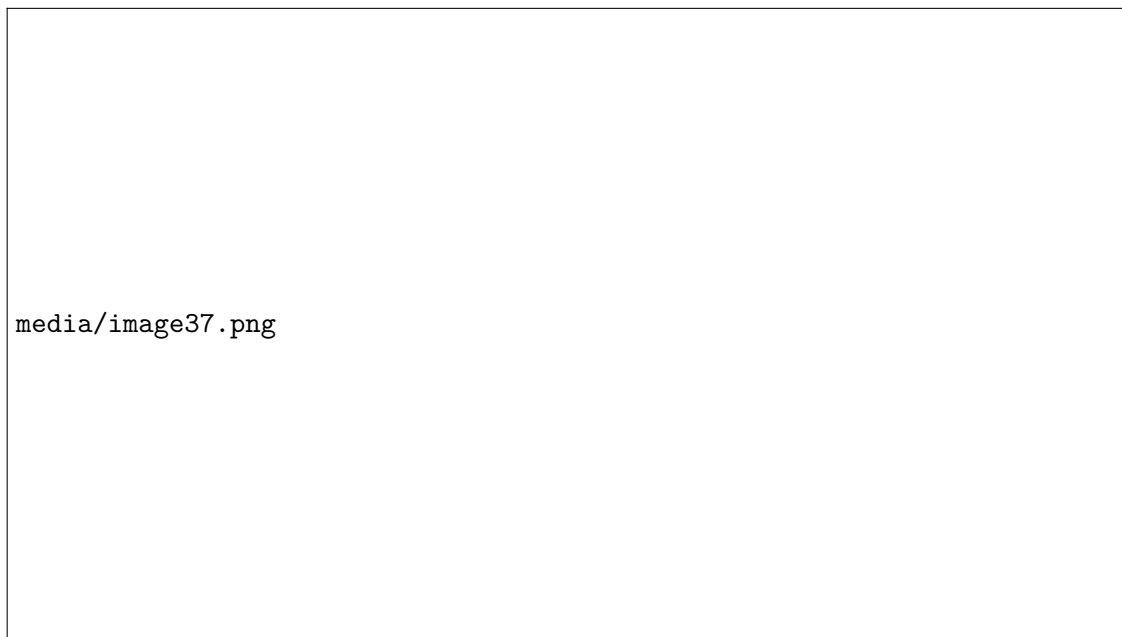


Figure 2.5: Figure 5: Linear Degradation of Narrow Framing Bias

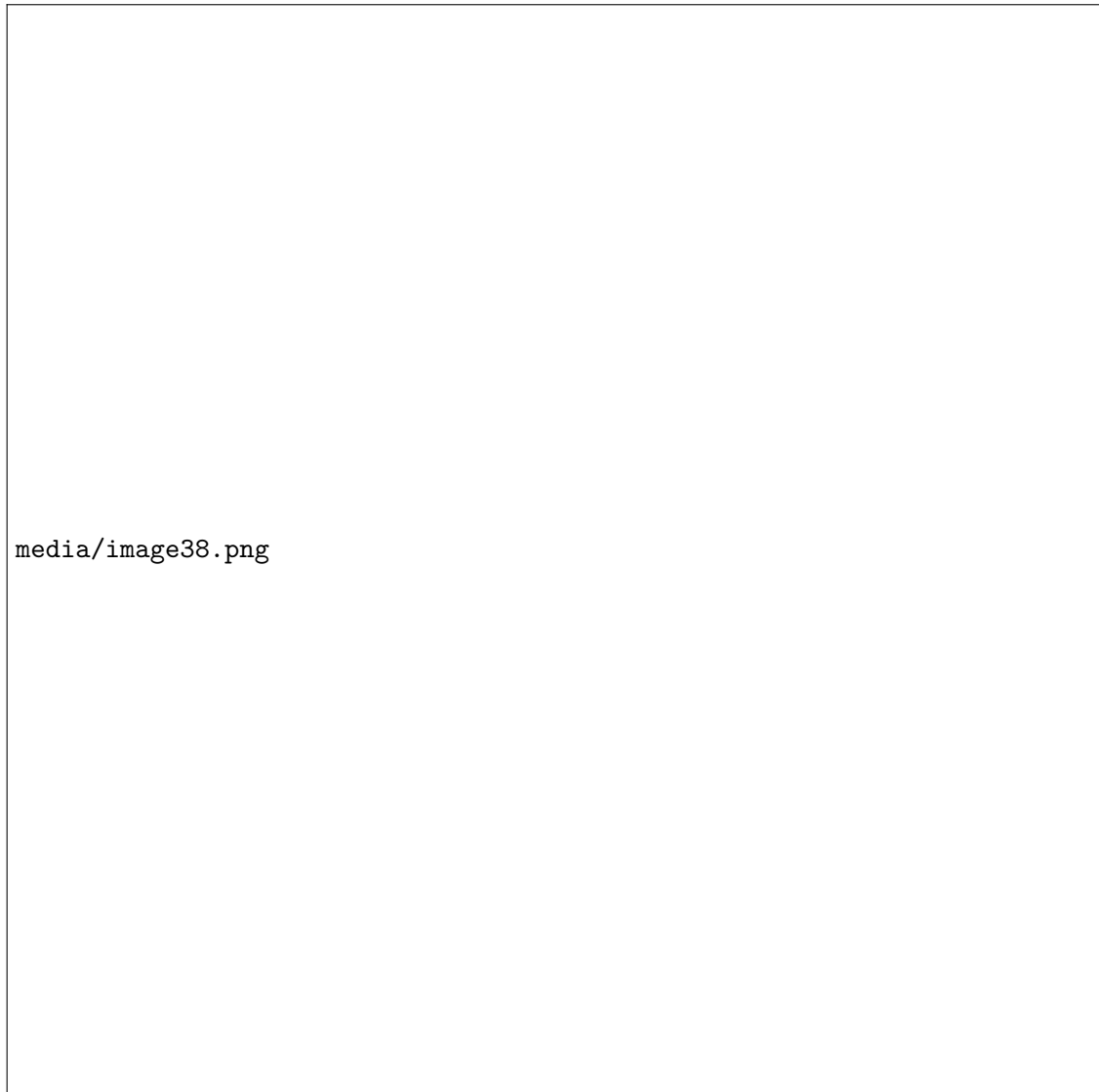


Figure 2.6: Figure6: Joint 0 (Base) Bias Effects

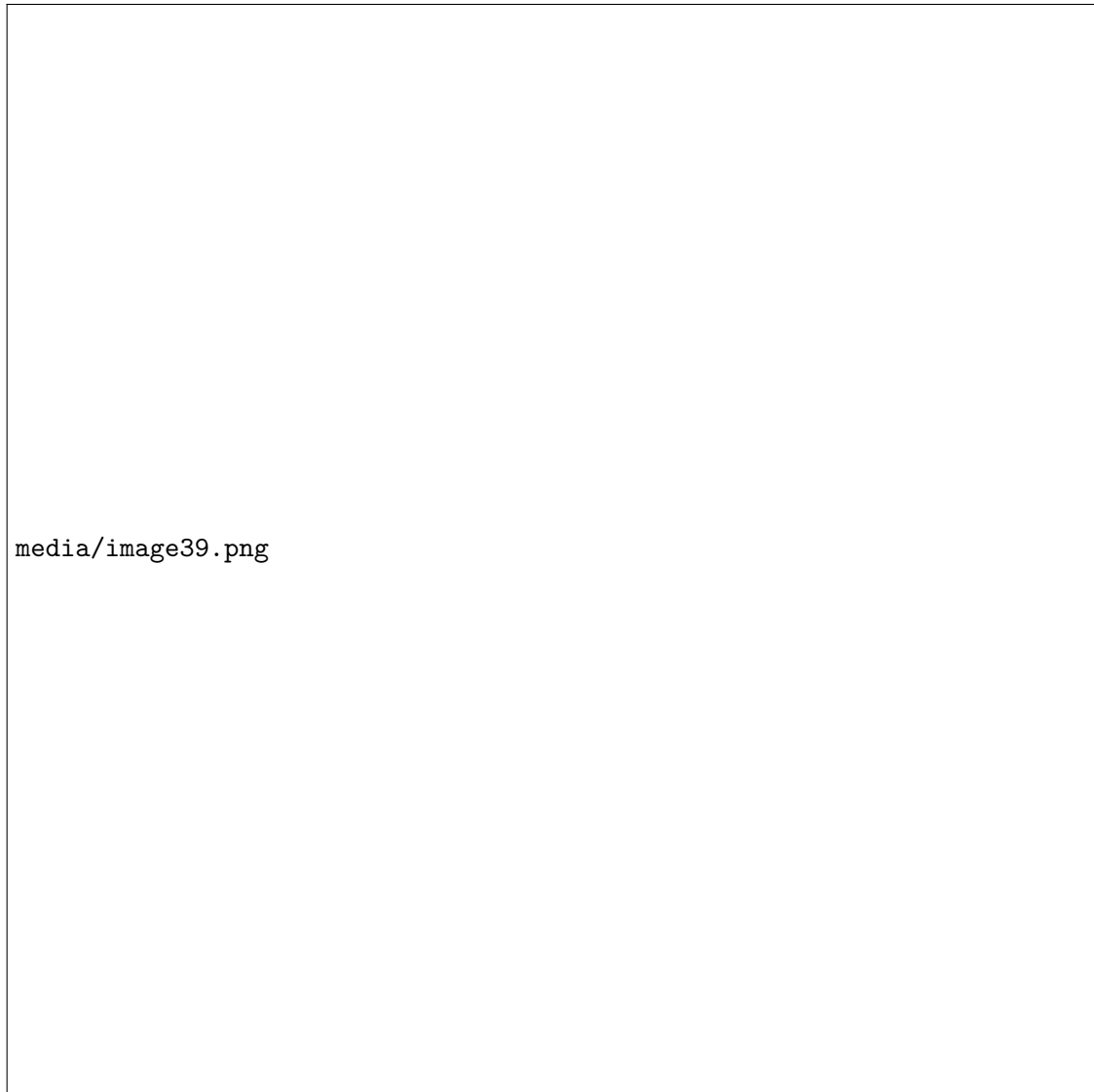


Figure 2.7: Figure7: Joint 7 (End-Effector) Bias Effects

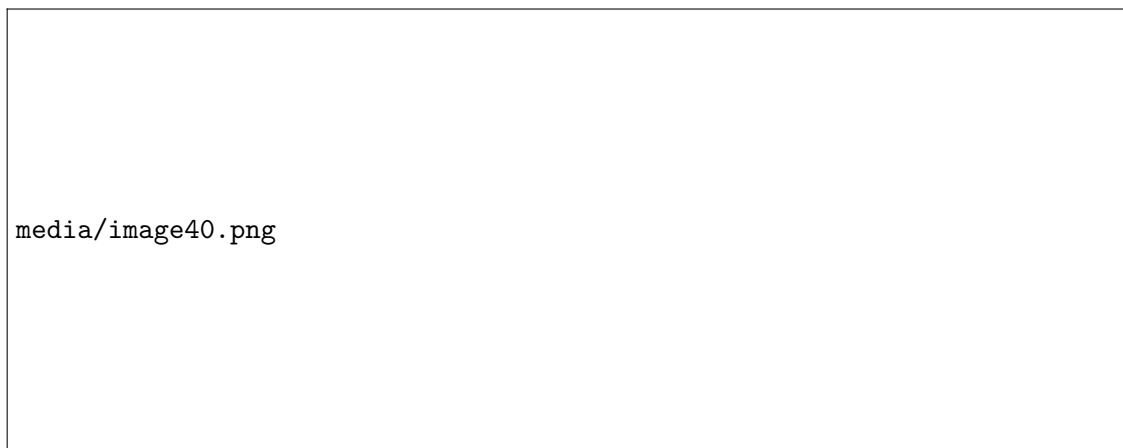


Figure 2.8: Figure8: Summary of All Bias Effects

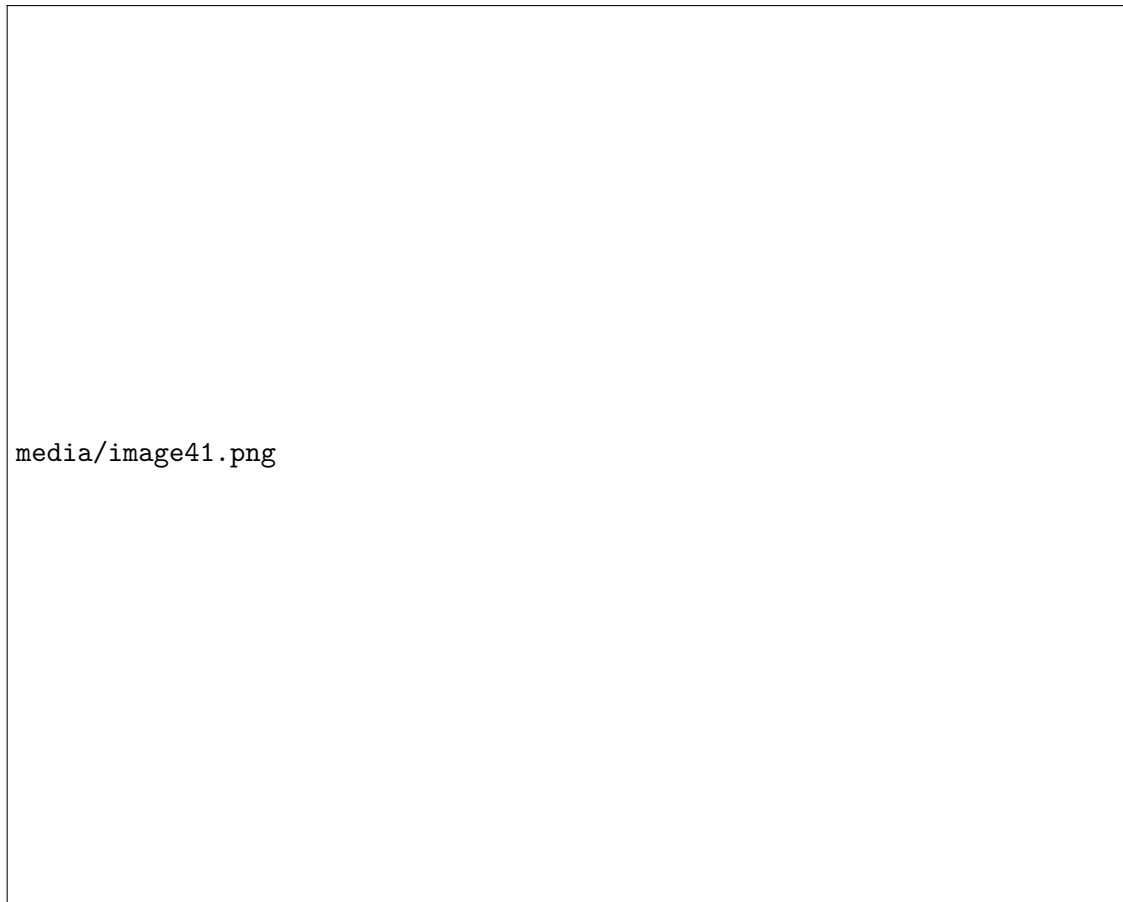


Figure 2.9: Figure9: Three-Tier Institutional Design for Effective Collaboration

匿証明特性を模倣することで、政策提案に含まれる秘密情報を秘匿化処理し、外部に漏らすことなく LLM の評価に使用する。さらに、TEE (Trusted Execution Environment) による分散型評価モデルと、Temperature=0 設定および Self-Consistency 手法による決定論的運用を組み合わせることで、LLM と ZK-SNARKs の本質的な違い（数学的保証と確率的期待）に対処し、実用上の信頼性を高める工学的アプローチを提案する。このアーキテクチャにより、市民参加による民主的正当性の確保、秘密情報を含む具体的な政策提案の実現、専門的視点に基づく評価、人間による最終的な評価決定の保障を同時に達成することを目指す。

**キーワード:** ウィキッド・プロブレム、ZK-SNARKs、Constitutional AI、LLM as a Judge、TEE (Trusted Execution Environment)、政策評価、市民討議、秘密情報の秘匿化

## 2.22.2 1. はじめに

### 1.1 研究の背景

現代社会が直面するウィキッド・プロブレム (Wicked Problems) は、その複雑性と多面性により、従来の政策立案・評価手法では適切に対処することが困難である。Rittel and Webber (1973) が提唱したこの概念は、明確な解が存在せず、多様なステークホルダーの価値観が対立し、試行錯誤による学習が不可欠な問題群を指す。地球温暖化、少子高齢化、都市計画、医療政策といった現代的課題は、いずれもこの特徴を備えており、単一の専門分野や手法では解決が困難である。ウィキッド・プロブレムの政策評価においては、多様な専門知識の統合が不可欠である。しかしながら、企業の技術情報、研究機関の未公開データ、個人の経験知など、政策評価に有用な情報の多くは秘匿性を要求される。従来の政策評価プロセスでは、こうした機密情報の提供を前提とするため、情報提供者の参加が制限され、評価の質と正当性が損なわれてきた。この構造的矛盾により、専門性の高い評価と参加者の秘匿性保護の両立が、政策対話における本質的課題となっている。

### 1.2 従来手法の限界

政策評価の手法は、その目的や志向性によって多様であり、杉谷 (2019) によれば、デモクラシーの類型に対応した複数の評価タイプが存在する。しかし、いずれのタイプにも固有の限界がある。第一に、プロフェッショナル評価 (専門家主導型) の限界がある。この方式では、評価の専門家や行政職員が実務を進め、市民はアンケートや情報源として扱われる。専門家の知見を活用できる一方で、テクノクラシー (専門家支配) に偏向する可能性が指摘されている (杉谷, 2019, p.190)。政治エリートの能力や意欲に依存し、民主的性格を失う恐れがある。また、専門家が保有する機密情報を評価に反映させることができず、評価の質が制限される。

第二に、ステークホルダー評価 (利害関係者参加型) の限界がある。この方式では、政策に利害関係を持つ者 (受益者、管理者、専門家など) が評価プロセスに参加し、多元的な視点を反映させることを目的とする。しかし、利害関係者間の利害対立を解決する方法がない場合、評価自体が紛争の場になりかねない (杉谷, 2019, p.198)。

第三に、協働型評価 (専門家+市民) の限界がある。この方式では、評価担当者がファシリテーターの役割を担い、関係者と協力してプログラム理論の構築や評価を行う。しかし、機密情報の取り扱いが困難であり、専門知と市民参加の質的両立が難しいという課題が残る。

さらに、EBPM (エビデンスに基づく政策形成) の文脈では、「説明責任 (Accountability)」と「応答責任 (Responsiveness)」の間に緊張関係が存在する (窪田・山谷編, 2020, p.213)。説明責任は近代科学的エビデンス (普遍性を志向) に基づき外部の第三者による統制を可能にするのに対し、応答責任は現場の実践家が個別ニーズや生活世界に根差したエビデンスに基づき果たすべき責任である。アーギロス (Argyrous, 2012) が指摘するように、EBPM における透明性は、行政を一方的に監視するためだけでなく、エビデンスの導出プロセスをオープンにすることでエビデンスの多元性を確保することに資する (杉谷, 2019, p.229)。

これらの限界は、秘匿性と透明性、専門性と参加性、説明責任と応答責任という相反する要求を同時に満たすことの困難さに起因する。奥田 (2019) らが指摘するように、ウィキッド・プロブ



レムの解決には「ステークホルダーの合意調達と、柔軟かつ迅速な専門知・実践知の調達・適用」の連動が不可欠であるが、従来の枠組みではこの連動を実現する技術的基盤が欠如していた。

### 1.3 本研究の目的

本研究の目的は、市民討議による原則策定、ZK-SNARKs による秘密情報の秘匿化処理、Constitutional AI 訓練手法による専門的価値観の埋め込みを統合した政策対話インフラを構築することである。このシステムにより、以下の4つの目標を達成する。

第一に、市民参加による民主的正当性と専門的評価の両立を実現する。本システムの特徴的な点は、Constitutional AI の憲法的原則を市民・研究者の討議を通じて共同策定することにある。自治体が討議を運営・記録し、市民と研究者が対話を通じて「公平性」「透明性」「プライバシー保護」などの原則を定義する。この過程により、LLM の訓練基盤となる原則に民主的正当性が付与される。討議で策定された原則に基づいて Constitutional AI 訓練を行うことで、市民の価値観を反映した専門的判断が可能となる。

第二に、秘密情報を含む政策提案の実現を可能にする。ZK-SNARKs の秘匿証明特性を模倣することで、政策提案者は企業秘密や未公開研究データをシステムに提出できる。秘密情報は外部に一切漏らさず、LLM の評価にのみ使用される。これにより、企業の技術情報や研究機関の未公開データを含む具体的根拠を持つ政策提案が可能となりながら、知的財産や競争優位性を保護することができる。

第三に、参加者の多様化と提案の質的向上を実現する。秘匿性の保証により、従来参加を躊躇していた研究機関が具体的な技術データや研究成果を含む高品質な政策提案を提出できる。市民は意見を提出し、研究者は機密データを含む提案を行う。専門的価値観に基づく評価により、提案内容そのもので判断され、権威や社会的地位に依存しない公平な評価システムを構築する。

第四に、透明性と秘密保護の両立を強化する。ZK-SNARKs により、提案の計算結果が秘密データに基づいて正しく導出されたことを証明しつつ、評価プロセスの透明性を保ちながら提案に含まれる秘密情報を保護する。自治体・行政が最終判断を行い、議会・住民への説明責任を果たす。これにより、行政の説明責任を技術的に保証し、市民の信頼を獲得する。

本論文の構成は以下の通りである。第2章では、ZK-SNARKs の技術的基礎とその政策評価への適用可能性を検討する。第3章では、Constitutional AI と LLM as a Judge の技術的特性を整理し、政策評価における役割を明らかにする。第4章では、提案するシステムアーキテクチャを詳述し、技術統合の方法論を示す。第5章では、期待される効果と今後の展望を論じる。

[keepspectratio]research-overview.svg 図1：本研究の全体像 - 市民討議、ZK-SNARKs、Constitutional AI の統合アーキテクチャ

## 2.22.3 2. ZK-SNARKs と秘匿証明

### 2.1 ZK-SNARKs の技術的基礎

Zero-Knowledge Succinct Non-interactive Arguments of Knowledge (ZK-SNARKs) は、暗号学における秘匿証明プロトコルであり、4つの本質的特性を持つ。これらの特性により、従来は不可能であった「秘密を完全に守りながら、その秘密に関する知識を持っていることを証明する」ことが技術的に実現可能となった。

Zero-Knowledge（ゼロ知識）特性は、証明を通して元の秘密情報が一切漏洩しないことを保証する。検証者は「証明者が知識を持っている」という事実のみを確認でき、知識そのものは一切得られない。この特性により、機密情報の詳細を開示することなく、その情報の保持を証明することが可能となる。

Succinct（簡潔）特性は、どんなに複雑な計算であっても証明サイズが常に数百バイト程度と一定であることを意味する。膨大なデータの正しさを小さな証明で検証可能であり、計算効率と通信効率の両面で優れた性能を発揮する。この特性により、大規模な政策評価プロセスにおいても実用的な処理時間で証明の生成と検証が可能となる。

Non-interactive（非対話）特性は、証明者から検証者への1回の送信で証明が完了することを示す。複数回のやり取りを必要とせず、オンライン通信の負荷を大幅に削減する。この特性により、非同期的な政策評価プロセスや、多数の参加者が関与する大規模な協議においても効率的な運用が可能となる。

Arguments of Knowledge（知識の論証）特性は、証明者が本当にその知識を所有している必要があり、偽造が計算量的に不可能であることを保証する。数学的安全性により信頼性が確保され、悪意のある参加者による不正な証明の作成を防止する。

これらの4つの特性は、楕円曲線暗号などの数学的基盤により実現される。特に、楕円曲線上の離散対数問題の計算困難性に基づくペアリング演算を利用することで、効率的な証明生成と検証が可能となる。Groth（2016）により提案された Groth16 方式は、現在最も効率的な ZK-SNARKs 実装の一つであり、証明サイズが 128 バイト、検証時間が数ミリ秒という実用的な性能を達成している。

[keepaspectratio]zk-snarks-properties.svg 図 2：ZK-SNARKs の 4 つの特性 - Zero-Knowledge、Succinct、Non-interactive、Arguments of Knowledge

## 2.2 政策提案への適用可能性

ZK-SNARKs の秘匿証明特性は、政策提案における秘密情報保護と提案の具体性の両立という構造的課題に対する技術的解決策を提供する。従来、参加者が保有する機密情報を政策提案に含めるためには、その情報の詳細を開示する必要があり、これが具体的な提案の提出を阻害してきた。ZK-SNARKs を用いることで、提案者は機密情報そのものを開示することなく、その情報を含む政策提案をシステムに提出できる。

具体的には、提案者が保有する技術データ、研究成果、実証実験の結果などを秘密情報として扱い、これらの情報を含む政策提案をシステムに提出する。ZK-SNARKs の秘匿化処理により、秘密情報は LLM の評価に使用されるが、外部には一切漏らさない。証明プロセスにより、提案に含まれる計算結果（例：「CO2 を 30%削減できる」）が実際に秘密情報に基づいて正しく導出されたことを、秘密情報を開示することなく証明できる。検証者は、証明を検証することで、提案の計算結果が秘密データに基づいて正しいことを確認できるが、秘密データの詳細にはアクセスできない。この仕組みにより、企業は競争優位性を損なうことなく具体的な技術データを含む政策提案を提出でき、研究機関は未公開の研究成果を保護しながら科学的根拠を含む提案を行うことができる。また、個人は経験知やノウハウといった暗黙知を秘匿しながら、その知見に基づく提案を行うことができる。この結果、従来は抽象的な提案しかできなかった多様なステークホルダーが、具体的根拠を持つ高品質な政策提案を提出することが可能となり、政策の質と実効性が向上する。

[keepaspectratio]zk-snarks-concept.svg 図 3：ZK-SNARKs による秘匿証明の概念 - 秘密情報を開示せずに知識を証明

さらに、ZK-SNARKs の非対話特性により、評価者は自身の都合の良いタイミングで証明を生成し、非同期的に提出できる。これにより、時間的・空間的制約を受けずに大規模な参加型評価を実現できる。Succinct 特性により、証明サイズが小さく保たれるため、ネットワーク帯域や保存容量の制約も最小限に抑えられる。

## 2.3 ZK-SNARKs vs STARKs：文脈依存性の必然

ZK-SNARKs の代替技術として ZK-STARKs（Scalable Transparent Arguments of Knowledge）が存在するが、本研究では政策評価の文脈において ZK-SNARKs が適していると判断した。この選択の理論的根拠は、政策評価における文脈依存性の本質にある。

STARKs の根本的特徴は、Trusted Setup を必要としない Universal 性にある。これは、どんな文脈でも同じアルゴリズムを適用し、数学的に一律な検証基準を用いることを意味する。この特性は、暗号通貨のような文脈に依存しない純粋な計算検証には適しているが、政策評価という本質的に文脈依存的な領域には適合しない。

政策評価の現実には、文脈依存性が本質である。地域、時代、文化によって評価基準が変化し、人間の規範が動的に創造される。ステークホルダーは異なる価値体系を持ち、これらの多様性を尊

重しながら合意形成を図る必要がある。STARKs の Universal 性は、この文脈依存的な評価を適切に扱うことができない。

対照的に、ZK-SNARKs の Trusted Setup は、文脈に応じたカスタマイズされた評価システムの構築を可能にする。各政策分野、各地域、各時代に応じて最適化された証明システムを設計できる。Trusted Setup の過程で、その文脈における重要な価値観や評価基準を反映させることができる。この柔軟性により、政策評価の文脈依存的な性質に適合したシステムを構築できる。

さらに、ZK-SNARKs は証明サイズと検証時間において優れた性能を示す。Groth16 方式では証明サイズが 128 バイトと極めて小さく、検証時間も数ミリ秒と高速である。大規模な参加型政策評価において、多数の証明を効率的に処理する必要がある場合、この性能差は実用上重要な意味を持つ。

ただし、Trusted Setup のプロセスには慎重な設計が必要である。複数の独立した参加者によるマルチパーティ計算 (MPC) を用いることで、単一の参加者が秘密鍵を保持するリスクを排除できる。実際に、Zcash プロジェクトでは 2016 年の Sprout ceremony (6 人) から 2018 年の Sapling ceremony (90 人) へと参加者を拡大し、Tornado.cash はブラウザベースの参加を可能にすることで 1,114 人の参加者を集めた (Buterin, 2022)。さらに、2023 年に開始された Ethereum KZG Ceremony は約 95,000 件の貢献を集め、史上最大規模の Trusted Setup ceremony となった (Ethereum Foundation, 2023)。これらの実績により、実用上の安全性は十分に確保できることが示されている。

[keepspectratio]snarks-vs-starks.svg 図 4: ZK-SNARKs と STARKs の比較 - 文脈依存性と性能特性

## 2.22.4 3. Constitutional AI と LLM as a Judge

### 3.1 LLM as a Judge の概念と発展

Large Language Model を評価者 (Judge) として活用する LLM as a Judge の概念は、2024 年以降急速に発展した研究領域である。従来、テキスト生成や質問応答などのタスクにおける LLM の性能評価は、人間による主観的判断や、BLEU スコアなどの自動評価指標に依存していた。しかし、これらの手法は評価の一貫性、スケーラビリティ、コスト効率の面で課題を抱えていた。

LLM as a Judge は、LLM 自身が他の LLM の出力や人間の文章を評価する枠組みである。この手法の理論的基盤は、大規模言語モデルが訓練過程で獲得した広範な知識と、人間の価値判断パターンの学習にある。Shi et al. (2024) による包括的調査によれば、LLM as a Judge は単純な品質評価から、論理的整合性、事実正確性、倫理的適切性など、多次元的な評価が可能である。

Wang et al. (2024-2025) は、LLM ベースの評価手法を体系的に整理し、その有効性を実証した。特に、Chain-of-Thought (CoT) 推論を組み込むことで、評価プロセスの透明性が向上し、評価根拠の説明が可能となることを示した。Gu et al. (2025) は、開放型対話システムの評価における LLM as a Judge の有効性を検証し、人間評価者との高い一致率を報告している。

政策評価への応用という観点では、LLM as a Judge は以下の特性により有用である。第一に、大量の政策提案を効率的に処理できる。人間の評価者では処理しきれない規模の提案に対しても、一貫した基準で迅速な評価が可能である。第二に、多様な視点からの評価を統合できる。複数の評価基準を同時に考慮し、それらのバランスを取りながら総合的な判断を行う。第三に、評価プロセスの透明性を確保できる。評価理由を自然言語で説明することで、評価結果の解釈可能性が向上する。

[keepspectratio]llm-judge-concept.svg 図 5: LLM as a Judge の概念 - 政策提案の多基準評価と説明生成

### 3.2 LLM as a Judge の限界と課題

LLM as a Judge は有用な評価手法である一方で、いくつかの重要な限界を持つ。これらの限界を理解し、適切に対処することが、実用的な政策評価システムの構築には不可欠である。

第一の限界は、幻覚 (Hallucination) 問題である。LLM は、訓練データに含まれない情報や、事実と異なる内容を、もっともらしく生成する傾向がある。政策評価においては、根拠のない評価

や、誤った事実認識に基づく判断が重大な問題となる。特に、専門的な技術事項や最新の研究成果に関する評価では、LLM の知識の限界が顕在化する。

第二の限界は、バイアスと公平性の問題である。LLM は訓練データに含まれる社会的バイアスを学習し、それが評価に反映される可能性がある。性別、人種、地域、社会階層などに関する潜在的な偏見が、政策評価の公平性を損なう恐れがある。さらに、訓練データの分布により、特定の政治的立場や価値観に偏った評価を行う可能性も指摘されている。

第三の限界は、文脈理解の不完全性である。ウィキッド・プロブレムのような複雑な政策課題では、多層的な文脈理解が必要となるが、現在の LLM はこの能力に限界がある。特に、地域固有の事情、歴史的経緯、文化的背景などの深い理解が要求される場合、LLM の評価は表面的になる可能性がある。

第四の限界は、説明責任と最終決定権の問題である。LLM による評価は確率的な推論に基づいており、その判断過程は完全には解明されていない。政策評価という公共性の高い領域において、アルゴリズムによる自動的な決定に全面的に依存することは、民主的正当性と説明責任の観点から問題がある。

これらの限界に対処するため、本研究では Constitutional AI 訓練手法を導入し、さらに人間による最終的な評価決定を保障するシステム設計を採用する。

### 3.3 Constitutional AI：訓練手法の理論的基礎

Constitutional AI（憲法的 AI）は、Anthropic により 2022 年に提案された、LLM を人間の価値観に整合させる訓練手法である。Bai et al. (2022) による原著論文では、人間が定義した憲法原則（Constitution）に基づいて AI 自身が自己改善を行う手法が提示されている。

Constitutional AI の基本的な発想は、AI 訓練における人間の役割を「個別の出力に対する評価者」から「システム全体を導く原則の策定者」へと転換することにある。従来の Reinforcement Learning from Human Feedback (RLHF) では、人間は個々の AI 出力に対して「良い」「悪い」の判断を下す必要があり、これが大きな労力とコストを要した。Constitutional AI では、人間は「helpful（有用である）」「harmless（無害である）」「honest（誠実である）」といった高次の原則を自然言語で定義し、AI がこれらの原則に基づいて自己評価と自己修正を行う。

訓練プロセスは2つのフェーズから構成される。第一の Supervised Learning (SL) フェーズでは、LLM が生成した応答を、LLM 自身が憲法原則に照らして批評（Critique）し、改訂（Revision）する。この自己批評と自己修正のプロセスを繰り返すことで、LLM は憲法原則を内在化していく。第二の Reinforcement Learning (RL) フェーズでは、LLM が生成した複数の応答を、LLM 自身が憲法原則に基づいて評価し、優れた応答を選択する。この評価を用いて Preference Model（選好モデル）を訓練し、さらにこのモデルを報酬信号として強化学習を行う。このプロセスは「Reinforcement Learning from AI Feedback (RLAIF)」と呼ばれる。

Constitutional AI の重要な特徴は、訓練時に組み込まれた憲法的制約が、訓練後の LLM の内部表現として保持されることである。これにより、運用時には外部からの継続的な監視や介入なしに、LLM が自律的に倫理的制約を遵守した判断を行うことができる。この特性は、ZK-SNARKs における楕円曲線暗号の役割と類似している。楕円曲線暗号が ZK-SNARKs の数学的安全性を基盤として保証するように、Constitutional AI は LLM の倫理的信頼性を基盤として保証する。

政策評価への応用という観点では、Constitutional AI は以下の利点を提供する。第一に、公平性、透明性、プライバシー保護といった政策評価に必要な倫理的原則を、訓練時に明示的に組み込むことができる。第二に、市民討議という民主的プロセスで合意された判断基準を LLM に内在化させることで、評価基準に民主的正当性を付与できる。第三に、評価理由の説明において、憲法原則への言及を通じて、評価の透明性と説明責任を向上させることができる。

[keepaspectratio]constitutional-ai-process.svg 図 6: Constitutional AI の訓練プロセス - 市民討議による原則策定から LLM as a Judge へ

### 3.4 Constitutional AI 訓練による LLM as a Judge の実現

本研究では、Constitutional AI 訓練手法により訓練された LLM を、LLM as a Judge として政策評価に活用する。この統合的アプローチにより、LLM as a Judge の限界を克服し、信頼性の高い

政策評価支援システムを実現する。

訓練プロセスにおいて、政策評価に特化した憲法的原則と専門的価値観を定義する。憲法的原則には、公平性（特定の集団やイデオロギーへの偏向を避ける）、透明性（評価理由を明確に説明する）、プライバシー保護（個人情報や機密情報を適切に扱う）、科学的根拠の重視（エビデンスに基づく評価を行う）、多様な視点の尊重（異なる価値観を認識し考慮する）などが含まれる。専門的価値観には、各政策分野（環境、福祉、都市計画等）における専門家の判断パターン、評価基準、重視すべき要素などが含まれる。

SL フェーズでは、様々な政策提案に対する評価を生成し、それを憲法的原則と専門的価値観に照らして自己批評する。例えば、「この評価は特定の政治的立場に偏っていないか」「評価理由は十分に明確か」「科学的根拠は適切に引用されているか」「専門的観点から適切な判断基準を用いているか」といった観点から、AI 自身が評価を見直し、改善する。この反復的プロセスにより、政策評価に適した判断パターンと専門的視点が学習される。

RL フェーズでは、複数の評価候補を生成し、それらを憲法的原則と専門的価値観に基づいて比較評価する。専門家の判断パターンを訓練データとして学習し、AI 評価（RLAIF）を報酬信号として強化学習を行う。Chain-of-Thought 推論を組み込むことで、「なぜこの評価がより公平か」「どの評価がより専門的観点から適切か」といった判断プロセスを明示化する。これにより、最終的な評価選択の根拠が明確化され、説明可能性が向上する。

訓練された LLM as a Judge は、政策評価において以下の機能を提供する。第一に、大量の政策提案を効率的に処理し、初期スクリーニングや優先順位付けを行う。第二に、訓練された専門的価値観に基づいて評価を実行する。第三に、評価理由を自然言語で説明し、評価の透明性を確保する。第四に、憲法的制約と専門的判断基準に基づく自律的な判断により、一貫性のある公平な評価を提供する。

重要な点として、LLM as a Judge は最終的な決定を行うのではなく、人間による意思決定を支援する役割に徹する。Constitutional AI 訓練により、「アルゴリズムは支援に徹し、最終決定は人間が実施する」という原則が内在化されている。これにより、民主的正当性と説明責任を確保しながら、AI の効率性と専門的判断能力を活用することが可能となる。

## 2.22.5 4. システムアーキテクチャと技術統合

### 4.1 全体アーキテクチャの概要

本研究が提案する政策対話インフラは、Constitutional AI 訓練手法により訓練された LLM as a Judge を中核とし、ZK-SNARKs の秘匿証明特性を統合したシステムである。このアーキテクチャは、人間、訓練プロセス、運用システムという 3 つの層から構成される。

人間層では、政策評価に関わる全てのステークホルダーが参加する。この層の主要な役割は四つある。第一に、市民と研究者が討議を通じて Constitutional AI の憲法的原則を共同策定することである。自治体が討議を運営・記録し、市民と研究者が対話を通じて「公平性」「透明性」「プライバシー保護」「科学的根拠の重視」などの原則を定義する。この討議プロセスにより、LLM の訓練基盤となる原則に民主的正当性が付与される。第二に、各政策分野の専門的価値観を定義することである。研究者が環境政策、福祉政策、都市計画など、分野ごとの専門家の判断パターンや評価基準を整理する。第三に、政策提案を提出することである。市民は意見を提出し、研究者は機密データを含む提案を行う。第四に、最終的な評価決定を行うことである。自治体・行政が LLM as a Judge による評価結果を参考にしながら最終判断を行い、議会・住民への説明責任を果たす。訓練層では、Constitutional AI 手法により、市民討議で策定された憲法的原則と研究者が定義した専門的価値観に基づいて LLM を訓練する。SL フェーズにおいて、AI 自身が政策評価の応答を生成し、市民が定めた憲法的原則と専門的価値観に照らして自己批評と自己修正を繰り返す。RL フェーズにおいて、専門家の判断パターンを学習し、AI 評価（RLAIF）を報酬信号として強化学習を行い、民主的に策定された憲法的制約と専門的判断基準を LLM の内部表現として内在化させる。この訓練プロセスにより、市民の価値観と専門的視点を組み込んだ LLM が生成される。

運用層では、訓練された LLM as a Judge が政策評価を実行する。この LLM は、ZK-SNARKs の秘匿証明特性と政策評価システムを単一のトランスフォーマーモデル内で統合的に実装する。具

体的には、LLM は秘密情報を含む政策提案を入力として受け取り、訓練された専門的価値観に基づく評価を実行し、出力層において秘密を含まない評価結果と評価の根拠・理由を出力する。評価結果は自然言語で説明される。

このアーキテクチャの特徴は、市民討議による原則策定、Constitutional AI 訓練、LLM as a Judge による運用という一貫したフローにある。市民と研究者が討議を通じて定めた原則が Constitutional AI 訓練の基盤となり、その訓練により LLM に内在化された倫理的制約と専門的判断基準が運用時の自律的判断を導く。外部からの継続的な監視や介入を必要とせず、民主的に策定された原則に基づいて LLM が評価を行う。この仕組みは、楕円曲線暗号が ZK-SNARKs の数学的安全性を基盤として保証するように、市民討議と Constitutional AI が LLM の民主的正当性と専門性を基盤として保証することに類似している。

[keepaspectratio]integration-architecture.svg 図 7: 提案するシステムの統合アーキテクチャ - 人間層、訓練層、運用層の構成

## 4.2 ZK-SNARKs 秘匿証明特性の LLM 実装：秘密情報の秘匿化処理

LLM トランスフォーマーモデルにおける ZK-SNARKs 秘匿証明特性の実装は、本研究の技術的核心の一つである。この実装により、政策提案に含まれる秘密情報を秘匿化処理し、LLM の評価に使用するが外部には一切漏らさない。

実装の第一段階は、政策提案に含まれる秘密情報の入力処理である。提案者が提出する技術データ、研究成果、実証実験の結果などの機密情報を、LLM は入力として受け取る。重要な点は、LLM は秘密情報を「見ている」ということである。秘密情報をベクトル空間に埋め込んで隠蔽するのではなく、LLM は入力された秘密情報の内容を処理する。

第二段階は、専門的価値観に基づく提案の評価実行である。LLM の内部処理において、Constitutional AI 訓練により内在化された専門的価値観に基づいて、秘密情報を含む政策提案を評価する。Transformer architecture の Attention 機構により、入力情報と訓練された専門的判断基準を統合し、提案の妥当性を判断する。

第三段階は、評価結果の出力である。LLM の出力層において、秘密情報を含まない政策評価結果を生成する。同時に、評価の根拠・理由を生成し、どのような観点から評価が行われたかを説明する。重要な特性は、出力される評価結果と根拠がいずれも秘密情報を含まないことである。検証者は、評価結果と根拠を確認できるが、秘密データの詳細にはアクセスできない。

この実装により、ZK-SNARKs の 4 つの特性を LLM 上で模倣することを目指す。Zero-Knowledge 特性は、出力時に秘密情報を含めないことにより模倣される。Succinct 特性は、評価結果のサイズが入力情報のサイズに依存せず簡潔であることにより模倣される。Non-interactive 特性は、評価者が一度の出力で評価を完結できることにより模倣される。Arguments of Knowledge 特性は、LLM の訓練過程において学習された知識表現の整合性により模倣される。

ただし、LLM による実装は、真の ZK-SNARKs とは本質的に異なる点に留意が必要である。ZK-SNARKs は楕円曲線暗号に基づき、秘密が漏れないことを数学的に保証する。一方、LLM は秘密情報を入力として受け取り、出力時に秘密を含めないように訓練されているだけであり、秘密が漏れないことの確率的な期待しか持てない。敵対的なプロンプトにより秘密を漏洩させられるリスクが残る。この本質的な違いを認識した上で、次節で述べる工学的アプローチにより、実用上の信頼性を高める方策を検討する。

## 4.3 秘匿性と再現性を高める工学的アプローチ

前節で述べた LLM と ZK-SNARKs の本質的な違いに対処するため、本研究では 2 つの工学的アプローチを提案する。これらは数学的保証には及ばないものの、実用上の信頼性を高められる可能性がある。

**4.3.1 TEE (Trusted Execution Environment) による分散型評価モデル** 第一のアプローチは、TEE (Trusted Execution Environment: 信頼できる実行環境) を活用した分散型評価システムである。TEE とは、Intel SGX や AWS Nitro Enclaves などに代表される、ハードウェアレベルで隔離された実行環境を指す。

このアプローチでは、Constitutional AI 訓練により構築された評価用 LLM を、Docker コンテナなどの形式でパッケージ化し、各事業者（政策提案者）に配布する。事業者は、自身のローカル環境に設置された TEE 内でこのモデルを実行し、秘密情報を含む政策提案を入力として評価を行う。TEE の重要な特性は、Remote Attestation（遠隔証明）機能である。これにより、「配布されたプログラムが改ざんされずに実行された」ことを電子署名によって証明できる。事業者は評価結果とこの証明書のみを行政側に送信し、秘密データそのものは事業者のローカル環境から外に出ることがない。

この構成は、ZK-SNARKs における証明者（Prover）と検証者（Verifier）の関係に類似している。ZK-SNARKs では、証明者がローカルで証明を生成し、検証者は証明のみを受け取って検証する。本提案では、事業者が TEE 内で評価を実行し、行政側は評価結果と TEE の証明書を受け取って検証する。秘密データが外部に送信されない点で、実用的な秘匿性の担保が期待できる。

ただし、TEE にも限界がある。サイドチャネル攻撃などの脆弱性が報告されており、完全な秘匿性を保証するものではない。また、TEE 環境の構築コストや、モデル配布のインフラ整備といった実務的な課題も存在する。

**4.3.2 決定論的運用による再現性の確保** 第二のアプローチは、LLM の確率的な挙動を抑制し、再現性を高める運用手法である。LLM は本質的に確率的なモデルであり、同じ入力に対しても異なる出力を生成する可能性がある。これは、行政手続きに求められる「いつ誰がやっても同じ結果になる」という公平性・一貫性の要件と相容れない。

この課題に対処するため、2つの手法を組み合わせる。第一に、Temperature（温度）パラメータを 0 に設定する。Temperature は、LLM の出力分布のエントロピーを制御するパラメータであり、0 に設定することで、最も確率の高いトークンのみを選択する決定論的な出力が得られる。これにより、同じ入力に対して常に同じ出力を生成することが期待できる。

第二に、Self-Consistency（自己無撞着性）手法を適用する。これは、同じ評価を複数回実行し、結果を集約する手法である。数値評価の場合は平均値を採用することで外れ値を排除し、文章評価の場合は多数決により最頻出の結論を採用する。この手法により、ハルシネーション（幻覚）のリスクを統計的に低減できる可能性がある。

これらの手法を組み合わせることで、LLM の確率的な揺らぎを抑制し、行政手続きに求められる再現性と一貫性を高められる可能性がある。ただし、これらはあくまで工学的なアプローチであり、ZK-SNARKs のような数学的保証を提供するものではない。実際に行政手続きの要件を満たせるかどうかは、今後の実証実験による検証が必要である。

## 4.4 システムの運用フロー

提案するシステムの運用フローは、準備段階、評価段階、決定段階の 3つのフェーズから構成される。

準備段階では、市民討議による原則策定と評価基準の設定を行う。まず、自治体が討議を運営し、対象となるウィキッド・プロブレムを明確化する。次に、市民と研究者が対話を通じて Constitutional AI の憲法的原則を共同策定する。この討議プロセスにより、「公平性」「透明性」「プライバシー保護」などの原則に民主的正当性が付与される。策定された原則に基づいて Constitutional AI 訓練を行い、LLM as a Judge を構築する。

評価段階では、政策提案の収集と評価の実行を行う。市民は意見を提出し、研究者は機密データを含む提案を行う。この際、ZK-SNARKs 秘匿証明特性を LLM で実装したシステムを用いることで、研究者は機密情報を開示することなく具体的根拠を持つ提案を提出できる。LLM as a Judge は、各政策提案について、市民討議で策定された原則と専門的価値観に基づいて多面的な評価を実行し、評価結果と評価の根拠・理由を生成する。

決定段階では、評価結果の確認と最終決定を行う。まず、LLM as a Judge による評価結果と根拠を確認し、評価の妥当性を検討する。次に、評価結果を自治体・行政の意思決定者に提示する。この際、評価理由と根拠が自然言語で説明され、透明性が確保される。最終的な政策決定は、自治体・行政がこれらの情報を参考にしながら行い、議会・住民への説明責任を果たす。Constitutional AI 訓練により、LLM は「アルゴリズムは支援に徹し、最終決定は人間が実施する」という原則を遵守している。



このフローにより、市民参加による民主的正当性の確保、秘密情報を含む政策提案の実現、専門的価値観に基づく評価、自治体・行政による説明責任の強化という4つの目標が達成される。

[keepspectratio]stakeholder-journey.svg 図8: 政策対話インフラにおけるステークホルダージャーニー - 市民、研究者、自治体の協働プロセス

## 2.22.6 5. 期待される効果と今後の展望

### 5.1 期待される効果

本研究が提案する政策対話インフラは、ウィキッド・プロブレムの政策評価において、以下の4つの効果をもたらすことが期待される。

第一の効果は、市民参加による民主的正当性と専門的評価の両立である。市民と研究者が討議を通じて Constitutional AI の憲法的原則を共同策定することで、LLM の訓練基盤に民主的正当性が付与される。自治体が討議を運営・記録することで、プロセスの透明性が確保される。ZK-SNARKs により、研究機関の未公開データを含む具体的な政策提案をシステムに提出できる。秘密情報は外部に一切漏らさず、LLM の評価にのみ使用される。市民討議で策定された原則と Constitutional AI 訓練により専門的価値観を埋め込んだ LLM が、秘密情報を含む提案を市民の価値観と専門的視点の両方から評価する。この結果、民主的正当性を持つ高品質な政策評価が実現し、自治体・行政の説明責任が強化される。

第二の効果は、多様な提案の包摂と公平な評価の実現である。ウィキッド・プロブレムは、異なる価値観を持つ多様なステークホルダーが関与する問題である。市民討議で策定された憲法的原則に基づいて Constitutional AI 訓練を行うことで、民主的プロセスによって合意された判断基準を LLM に内在化させた LLM as a Judge が実現される。秘匿性の保証により、研究機関が具体的な技術データや研究成果を含む提案を提出できる。市民は意見を提出し、研究者は機密データを含む提案を行う。提案内容そのもので判断されるため、権威や社会的地位に依存しない公平な評価が実現される。

第三の効果は、市民の価値観と専門的判断の調和である。市民討議で策定された原則が Constitutional AI 訓練の基盤となることで、市民の価値観が LLM の判断に反映される。同時に、「アルゴリズムは支援に徹し、最終決定は人間が実施する」という原則が LLM に内在化されている。これにより、LLM as a Judge は市民の価値観と専門的価値観の両方に基づいて大量の政策提案を効率的に処理し、初期スクリーニングや優先順位付けを行うが、最終的な決定は自治体・行政が担う。この仕組みにより、AI の効率性を活用しながら、民主的正当性と説明責任を確保できる。

第四の効果は、政策対話の質的向上と社会的信頼の獲得である。市民討議による原則策定プロセスにより、市民が政策評価の基盤に直接参加する機会が創出される。評価理由の自然言語説明により、評価プロセスの透明性が確保され、市民の理解が促進される。ZK-SNARKs により、提案の計算結果が秘密データに基づいて正しく導出されたことが証明され、自治体・行政の説明責任が技術的に保証される。秘密情報を含む多様な提案の統合と市民の価値観に基づく判断により、政策の多面的な影響が包括的に評価され、政策の質と実効性が向上する。これらの効果により、ウィキッド・プロブレムに対する社会的な合意形成が促進され、実効性のある政策実施が可能となる。

### 5.2 今後の課題

本研究が提案するシステムの社会実装には、技術的課題と制度的課題の両面において、さらなる研究開発が必要である。4.3 節で述べた工学的アプローチにより、秘匿性と再現性に関する原理的な不安に対しては対応策の方向性が示されたが、残された課題は依然として多い。

技術的課題としては、第一に、TEE による秘匿実行の実用性検証が必要である。TEE 環境の構築コスト、モデル配布のインフラ整備、各事業者への導入支援体制など、実務的な課題が存在する。また、TEE にはサイドチャネル攻撃などの脆弱性が報告されており、セキュリティ面での継続的な監視と対策が求められる。第二に、決定論的運用の有効性検証が必要である。Temperature=0 設定と Self-Consistency 手法の組み合わせが、実際に行政手続きの要件（公平性・一貫性）を満たすレベルの再現性を達成できるかは、実証実験による検証が不可欠である。第三に、市民討議プロ



セスの設計と実装が必要である。どのように討議を運営し、市民と研究者の対話を促進し、憲法原則として文書化するかの具体的な方法論を確立する必要がある。第四に、Constitutional AI 訓練による原則内在化の信頼性評価が必要である。市民討議で策定された原則が LLM 訓練にどう変換されるか、その手続きの設計と、原則内在化の信頼性を評価する指標の開発が求められる。制度的課題としては、第一に、地方自治体や政府機関との連携による実証実験が不可欠である。特定の政策分野（環境、福祉、都市計画等）において、小規模なパイロットプロジェクトを実施し、TEE 環境での運用可能性と決定論的運用の有効性を含めて検証する必要がある。第二に、法的・制度的枠組みの整備が必要である。TEE による秘匿実行の法的位置づけ、LLM による評価結果の行政手続きにおける扱い、個人情報保護との関係などについて、法的検討が求められる。第三に、社会的受容性の調査と向上が重要である。市民や専門家がシステムをどのように認識し、どの程度信頼するかを調査し、必要に応じてシステム設計や説明方法を改善する必要がある。

さらに、長期的な研究課題として、AI の継続的な進化への対応がある。LLM や TEE 技術は急速に発展しており、新しい手法や改良が継続的に提案されている。本システムも、これらの技術進化を取り込みながら、継続的に改善していく必要がある。特に、より安全な TEE 実装の登場や、Constitutional AI 訓練手法の発展に対応し、システムを更新していくことが重要である。

### 5.3 今後の展望

本研究が提案する政策対話インフラは、ウィキッド・プロブレムの政策評価という特定の応用領域から始まるが、その適用範囲は広範に及ぶ可能性がある。

短期的には、地方自治体レベルでの具体的な政策課題への適用が期待される。例えば、地域の環境政策、福祉政策、都市計画などにおいて、自治体が運営する市民討議を通じて評価原則を策定し、地域住民と研究機関が協働して政策を評価し、合意形成を図るプラットフォームとしての活用が考えられる。市民討議により地域固有の価値観が反映され、秘匿性の保証により研究機関が未公開の研究成果を政策評価に活用できるようになる。

中期的には、国レベルの重要政策の評価への展開が期待される。地球温暖化対策、少子高齢化対策、エネルギー政策など、国家的な重要課題において、全国規模の市民討議を通じて評価原則を策定し、多様なステークホルダーの専門知を統合し、科学的根拠に基づく政策評価を実現する。市民討議による原則策定と Constitutional AI 訓練により、特定の政治的立場に偏らない公平な評価が可能となり、超党派的な政策合意の形成を支援できる。

長期的には、国際的な政策協調への応用も視野に入る。気候変動、感染症対策、サイバーセキュリティなど、グローバルな協調が必要な課題において、各国が保有する機密情報を秘匿しながら、国際的な政策評価と合意形成を行うプラットフォームとしての発展が期待される。ZK-SNARKs の秘匿性により、国家安全保障上の機密を保護しながら、科学的知見の共有が可能となる。

さらに、政策評価以外の分野への応用も考えられる。企業の戦略的意思決定、研究開発プロジェクトの評価、医療における診断支援など、専門知の統合と秘匿性が同時に要求される多様な領域において、本研究の枠組みは有用である。Constitutional AI 訓練による倫理的制約の組み込みは、AI 応用の信頼性を確保する一般的な手法として、広範な領域で活用される可能性がある。

## 2.22.7 6. 結論

本研究は、ウィキッド・プロブレムの政策評価における市民参加と専門的判断の両立という構造的課題を解決するため、市民討議による原則策定、ZK-SNARKs による秘密情報の秘匿化、Constitutional AI 訓練を統合した政策対話インフラを提案した。市民と研究者が討議を通じて憲法原則を共同策定し、ZK-SNARKs が政策提案に含まれる秘密情報の秘匿化処理を担い、Constitutional AI 訓練により市民の価値観と専門的価値観を埋め込んだ LLM as a Judge が秘密情報を含む提案の評価を実行する。この技術統合により、民主的正当性と専門的評価を両立させた。

提案システムの特徴は、市民討議、ZK-SNARKs、Constitutional AI 訓練の役割分担にある。市民討議により憲法原則に民主的正当性を付与し、ZK-SNARKs は政策提案に含まれる秘密情報の秘匿化処理を担当し、提案者の機密情報を保護しながら LLM の評価に使用できるようにする。

Constitutional AI 訓練手法により、市民討議で策定された憲法的原則と専門的価値観が LLM に内在化され、民主的に策定された制約と専門的判断基準の下で秘密情報を含む提案を自律的に評価することが可能となる。楕円曲線暗号が ZK-SNARKs の数学的安全性を保証するように、市民討議と Constitutional AI が LLM の民主的正当性と専門性を保証する。

また、LLM と ZK-SNARKs の本質的な違い（数学的保証と確率的期待）に対処するため、2つの工学的アプローチを提案した。TEE（Trusted Execution Environment）による分散型評価モデルでは、訓練済み LLM をコンテナ化して各事業者に配布し、秘密データをローカル環境から外に出さずに評価を実行する。Temperature=0 設定と Self-Consistency 手法による決定論的運用では、LLM の確率的な揺らぎを抑制し、行政手続きに求められる再現性と一貫性を高める。これらは ZK-SNARKs のような数学的保証を提供するものではないが、実用上の信頼性を高められる可能性がある。

本システムにより、市民参加による民主的正当性の確保、秘密情報を含む政策提案の実現、専門的評価、公平性確保、自治体・行政による説明責任向上を同時に達成し、奥田（2019）らが指摘する「ステークホルダーの合意調達と、柔軟かつ迅速な専門知・実践知の調達・適用」の連動を、秘密情報を保護しながら実現することを目指す。今後は、TEE 環境での運用可能性の検証、決定論的運用の有効性実証、市民討議プロセスの設計、技術的実装の詳細設計、実証実験による有効性検証、法的・制度的課題の整理を通じて、真に実用的な政策対話インフラの社会実装を目指す。

[keepspectratio]framework-objectives.svg 図 9：提案する政策対話インフラの 4 つの目的

## 2.22.8 参考文献

1. Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4(2), 155-169.
2. 奥田恒 (2019). 「マイケル・ハウレットの『政策統合』アプローチ：ウィキッド・プロブレムへの対処戦略からの検討」『社会システム研究』22, 191-206. <https://doi.org/10.14989/241035>
3. Ben-Sasson, E., Chiesa, A., Tromer, E., & Virza, M. (2014). Succinct non-interactive zero knowledge for a von Neumann architecture. *Proceedings of the 23rd USENIX Security Symposium*, 781-796.
4. Groth, J. (2016). On the size of pairing-based non-interactive arguments. *Advances in Cryptology – EUROCRYPT 2016*. <https://eprint.iacr.org/2016/260.pdf>
5. Gu, J., Jiang, X., Shi, Z., Tan, H., Zhai, X., Xu, C., Li, W., Shen, Y., Ma, S., Liu, H., Wang, S., Zhang, K., Wang, Y., Gao, W., Ni, L., & Guo, J. (2024). A survey on LLM-as-a-Judge. *arXiv preprint arXiv:2411.15594*. <https://doi.org/10.48550/arXiv.2411.15594>
6. Li, H., Dong, Q., Chen, J., Su, H., Zhou, Y., Ai, Q., Ye, Z., & Liu, Y. (2024). LLMs-as-Judges: A comprehensive survey on LLM-based evaluation methods. *arXiv preprint arXiv:2412.05579*. <https://arxiv.org/abs/2412.05579>
7. Zheng, L., Chiang, W. L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E. P., Zhang, H., Gonzalez, J. E., & Stoica, I. (2023). Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena. *Advances in Neural Information Processing Systems*, 36.
8. Bai, Y., Jones, A., et al. (2022). Constitutional AI: Harmlessness from AI feedback. *arXiv preprint arXiv:2212.08073*. <https://arxiv.org/abs/2212.08073>
9. Saha, S., et al. (2025). Learning to plan & reason for evaluation with Thinking-LLM-as-a-Judge. *arXiv:2501.18099v1*.

10. Buterin, V. (2022). How do trusted setups work? *Vitalik Buterin's website*. <https://vitalik.eth.limo/general/2022/03/14/trustedsetup.html>
11. Ethereum Foundation. (2023). KZG Ceremony. <https://ceremony.ethereum.org/>
12. Bünz, B., Fisch, B., & Szepieniec, A. (2022). Powers-of-Tau to the People: Decentralizing Setup Ceremonies. *Cryptology ePrint Archive*, Paper 2022/1592. <https://eprint.iacr.org/2022/1592>
13. 杉谷和哉 (2019). 『政策にエビデンスは必要なのか：EBPM と政治のあいだ』ミネルヴァ書房.
14. 窪田好男・山谷清志 編 (2020). 『政策評価の実践とその課題：アカウンタビリティのジレンマ』晃洋書房.
15. Argyrous, G. (2012). Evidence Based Policy: Principles of Transparency and Accountability. *Australian Journal of Public Administration*, 71(4), 457–468. <https://doi.org/10.1111/j.1467-8500.2012.00786.x>

---

## 付録：数式表記の補足説明

### A.1 ZK-SNARKs における楕円曲線ペアリング

楕円曲線  $E$  上の 2 つの群  $\mathbb{G}_1, \mathbb{G}_2$  と、ペアリング関数  $e : \mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$  により、効率的な証明生成と検証が実現される。証明  $\pi$  の検証は、 $e(A, B) = e(C, D)$  の形式の等式検証に帰着される。

## Chapter 3

# 制度設計への示唆

### 3.1 理論的含意

#### 3.1.1 第4章のシミュレーション結果からの設計原則

第4章の計算論的分析から、以下の設計原則が導かれる。現状維持バイアスへの対処としては、臨界点 ( $b_{sq} \approx 0.25$ ) を超えないよう、段階的な変化導入を設計することが重要である。確証バイアスの活用としては、適度な自信は協調を促進するため、完全な中立性よりも構造的な多様性を確保することが有効である。狭い視野への対処としては、全体目標の可視化と横断的指標の導入を必須とすることが求められる。

#### 3.1.2 第5章の ZK-SNARKs システムからの制度的含意

ZK-SNARKs 型政策評価システムは、以下の制度的含意を持つ。第一に、秘匿性の制度化である。企業秘密や個人情報を保護しながら政策参加を可能にする制度が必要である。第二に、自動評価の境界である。AI による評価と人間による判断の境界を明確化する必要がある。第三に、市民参加の質的向上である。市民討議を通じた評価基準の共同設計が求められる。

### 3.2 各バイアスへの対処戦略

#### 3.2.1 現状維持バイアス：小さな変化の積み重ね

現状維持バイアスへの対処として、小さな変化の積み重ねという戦略を提案する。具体的には、パイロット導入として小規模な実験から開始し、成功事例を蓄積することが有効である。また、段階的拡大として、臨界点を超えないよう、徐々に適用範囲を拡大することが求められる。さらに、デフォルト設定の変更として、現状維持の方向に働く制度上のデフォルトを見直すことも重要である。

#### 3.2.2 確証バイアス：多様な視点の構造的導入

確証バイアスへの対処として、多様な視点の構造的導入という戦略を提案する。「悪魔の代理人」の制度化として、AI または人間による批判的視点の提示を必須化することが有効である。また、多様なステークホルダーの参画として、異なる立場・利害を持つ主体の関与を確保することが求められる。さらに、反証可能性の確保として、仮説を反証する証拠を積極的に探索するプロセスを組み込むことも重要である。

#### 3.2.3 狭い視野：全体目標の可視化

狭い視野への対処として、全体目標の可視化という戦略を提案する。横断的評価指標として、単一の指標ではなく、複数の指標による評価を導入することが有効である。また、システムマップ

の作成として、政策の全体像を可視化した図の共有が求められる。さらに、長期的影響の分析として、短期的な成果だけでなく、長期的な影響も評価することが重要である。

### 3.3 生成 AI を組み込んだ制度設計

#### 3.3.1 政治的-行政的インターフェース

政治レベルと行政レベルの間で、AIは以下のように位置づけられる。政治的判断については、人間（政治家・議会）が最終的に決定する。行政的分析については、AIが情報の整理・分析を支援する。民主的アカウンタビリティについては、人間が説明責任を負う。

#### 3.3.2 行政的-事業的インターフェース

行政レベルと事業レベルの間で、AIは以下のように位置づけられる。政策の具体化については、AIが選択肢の生成を支援する。運用の効率化については、AIが日常的な判断を補助する。人間による監督については、重要な判断は人間が行う。

#### 3.3.3 ZK-SNARKs 型システムの制度的位置づけ

ZK-SNARKs 型政策評価システムは、以下の制度的枠組みの中で運用される。第一に、評価基準の共同設計である。市民討議を通じた基準の策定が必要である。第二に、透明性の確保である。AIの評価プロセスと基準の公開が求められる。第三に、控訴の権利である。評価結果に対する異議申立ての機会を確保する必要がある。第四に、人間による最終判断である。AIの評価は参考情報として位置づけられる。

### 3.4 「連携・共創」の再設計

#### 3.4.1 三層制度設計（S-T1-T2-O）の再考

STO フレームワークを踏まえ、認知バイアスと生成 AI を考慮した制度設計を提案する。

Table 3.1: 認知バイアスと生成 AI を考慮した三層制度設計

階層	主な課題	対処戦略
S	現状維持バイアスによる変化抵抗	長期ビジョンの共有、段階的目標設定
T1	狭い視野による部分最適化	横断的指標、全体目標の可視化
T2	確証バイアスによる視点の固定化	多様な視点の構造的導入
O	日常的な認知バイアス	AI による対話的介入

S 層では現状維持バイアスによる変化抵抗が課題となるため、長期ビジョンの共有と段階的目標設定が有効である。T1 層では狭い視野による部分最適化が課題となるため、横断的指標と全体目標の可視化が求められる。T2 層では確証バイアスによる視点の固定化が課題となるため、多様な視点の構造的導入が必要である。O 層では日常的な認知バイアスが課題となるため、AI による対話的介入が有効である。

#### 3.4.2 生成 AI を「杖」として活用するガバナンス

生成 AI を活用するガバナンスのあり方として、以下の原則を提案する。第一に、補完性の原則である。AI は人間の能力を補完し、代替しない。第二に、透明性の原則である。AI の活用方法と限界を明示する。第三に、アカウンタビリティの原則である。最終的な判断と責任は人間が負う。第四に、継続的学習の原則である。AI と人間の協調プロセスを継続的に改善する。

### 3.5 小括

本章では、第4章の計算論的分析と第5章の ZK-SNARKs 型システム設計を踏まえ、制度設計への示唆を導出した。主な貢献として、認知バイアスへの具体的な対処戦略の提示、生成 AI を組み込んだ制度設計の方向性の提示、「連携・共創」の再設計に向けた枠組みの提示が挙げられる。次章では、本研究の総括と、都市計画への展開について論じる。

## Chapter 4

# 結論：都市計画への展開

### 4.1 研究の総括

#### 4.1.1 研究目的の達成

本研究は、以下の3つの目的を掲げた。第一に、公共交通政策における実装ギャップの要因として、人間の認知バイアスの影響を計算論的に解明すること。第二に、生成 AI と人間の協調的関係性を具体化するシステムとして、ZK-SNARKs 型政策評価システムの可能性を探ること。第三に、認知バイアスと生成 AI を考慮した制度設計への示唆を導出すること。

**目的 1 について** 第 4 章において、協調ロボット制御モデルを用いた計算論的分析により、三つの認知バイアスが政策協調に与える影響を定量的に解明した。特に、現状維持バイアスの閾値効果、確証バイアスの逆説的效果、狭い視野の一貫した負の影響を発見した。

**目的 2 について** 第 5 章において、ZK-SNARKs の概念を援用した政策評価システムを提案し、LLM as a Judge、Constitutional AI、市民討議を組み合わせたアーキテクチャを設計した。

**目的 3 について** 第 6 章において、認知バイアスへの対処戦略、生成 AI を組み込んだ制度設計、「連携・共創」の再設計に向けた枠組みを提示した。

#### 4.1.2 核となる主張

本研究の核となる主張は、以下の通りである。

生成 AI は人間の「執政の創造性」を補完する「杖」として、認知バイアスへの対話的介入を通じて政策形成を支援できる。ZK-SNARKs の概念を援用することで、秘匿性と信頼性を両立した政策評価システムが可能になる。

### 4.2 理論的貢献

#### 4.2.1 認知バイアスの計算論的分析手法の政策科学への導入

本研究は、協調ロボット制御理論を政策ネットワーク分析に応用し、認知バイアスの影響を計算論的に解明する手法を導入した。このアプローチは、政策科学における計算論的転回（computational turn）の一環として位置づけられる。

#### 4.2.2 ZK-SNARKs 概念の政策評価への応用

本研究は、暗号技術の概念である ZK-SNARKs を政策評価に応用する試みとして先駆的である。この概念転用は、「秘密を守りながら専門性を証明する」という新たな政策参加のあり方を示唆している。

### 4.2.3 生成 AI の「杖」としての理論的位置づけ

本研究は、生成 AI と人間の関係性を「杖」として補完的に位置づける理論的枠組みを提示した。これは、AI による代替ではなく、AI と人間の協調を前提とする視点である。

## 4.3 実践的貢献

### 4.3.1 制度設計への提言

本研究は、認知バイアスと生成 AI を考慮した制度設計への具体的な提言を行った。現状維持バイアスへの対処として、段階的变化導入の設計を提案した。確証バイアスへの対処として、「悪魔の代理人」の制度化を提案した。狭い視野への対処として、横断的評価指標の導入を提案した。生成 AI の活用として、補完性、透明性、アカウントビリティの原則を提示した。

### 4.3.2 ZK-SNARKs 型政策評価システムの設計指針

本研究は、ZK-SNARKs 型政策評価システムの具体的な設計指針を提示した。三層アーキテクチャの採用、Constitutional AI と市民討議による評価基準設計、控訴プロセスによる人間介入の確保がその主要内容である。

## 4.4 今後の課題：都市計画を舞台にした実証

### 4.4.1 より複雑な政策領域への適用

公共交通政策は、本研究の「実験場」として適切であったが、より複雑な政策領域への適用が期待される。特に、都市計画は多様なステークホルダー（土地所有者、開発業者、住民、行政など）、多様な秘密情報（土地利用計画、開発権、資産価値など）、長期的影響（数十年単位での都市構造の変化）という点で興味深い研究対象となる。

### 4.4.2 ZK-SNARKs システムの社会実装

ZK-SNARKs 型政策評価システムの社会実装に向けては、いくつかの課題がある。技術的実装としては、TEE、LLM、Constitutional AI の統合が必要である。制度的設計としては、法的位置づけや運用ルール策定の求められる。社会的受容としては、市民の理解と信頼の獲得が重要である。

### 4.4.3 生成 AI と人間の協調的関係性の継続的検証

生成 AI 技術は急速に進化しており、人間-AI 協調のあり方も変化し続ける。本研究の枠組みは、継続的な検証と改善を必要とする。

## 4.5 結び

本研究は、公共交通政策を舞台に、生成 AI と人間の協調的関係性を探求した。その過程で、認知バイアスによる協調失敗のメカニズムを計算論的に解明し、ZK-SNARKs 概念を援用した政策評価システムを提案し、制度設計への示唆を導出した。

公共交通政策は、本研究の「第一の舞台」であった。今後は、都市計画という「第二の舞台」での実証を通じて、生成 AI と人間のより良い協調のあり方を探求していきたい。



## 参考文献

- [1] Daniel Kahneman. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.
- [2] Georges Canguilhem. *Le normal et le pathologique*. 滝沢武久訳：正常と病理（叢書・ユニベルシタス, 225）、法政大学出版局、1987年。PUF, 1966.
- [3] Takashi Iba. “An Autopoietic Systems Theory for Creativity”. In: *The 1st Conference on Collaborative Innovation Networks (COINs)*. 2009.
- [4] John W. Kingdon. *Agendas, Alternatives, and Public Policies*. Little, Brown, 1984.
- [5] 草野厚. **政策過程分析入門**. 慶應義塾大学出版会, 1997.
- [6] Paul A. Sabatier. “An Advocacy Coalition Framework of Policy Change and the Role of Policy-Oriented Learning Therein”. In: *Policy Sciences* 21.2/3 (1988), pp. 129–168.
- [7] 村松岐夫. **戦後日本の官僚制**. 東洋経済新報社, 1981.
- [8] Horst W. J. Rittel and Melvin M. Webber. “Dilemmas in a General Theory of Planning”. In: *Policy Sciences* 4.2 (1973), pp. 155–169.
- [9] 杉谷和哉. “ウィキッド・プロブレムとしての新型コロナウイルス感染症 政治と専門性の関係を中心に”. In: **医療福祉政策研究** 4.1 (2021), pp. 27–37.
- [10] Jürgen Habermas. *Theorie des kommunikativen Handelns*. 河上倫逸ほか訳：コミュニケーション的行為の理論、未来社、1985–1987年。Suhrkamp, 1981.
- [11] Niklas Luhmann. *Soziale Systeme*. 馬場靖雄訳：社会システム理論、勁草書房、2001–2020年。Suhrkamp, 1984.
- [12] Georg Kneer and Armin Nassehi. *Niklas Luhmanns Theorie sozialer Systeme*. 館野受男訳：ルーマン社会システム理論——「知」の扉をひらく、新泉社、1995年。Fink, 1993.
- [13] Humberto R. Maturana and Francisco J. Varela. *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel, 1980.
- [14] Ben Shneiderman. *Human-Centered AI*. Oxford University Press, 2022.
- [15] Hidehiko Yoshihara et al. “Cooperative Control of Multi-Joint Robot Arm”. In: *Proceedings of IEEE International Conference on Robotics and Automation*. 2009, pp. 1–6.
- [16] Chris Ansell and Alison Gash. “Collaborative Governance in Theory and Practice”. In: *Journal of Public Administration Research and Theory* 18.4 (2008), pp. 543–571.
- [17] Hirokazu Kato et al. “Community-Participatory Regional Public Transport”. In: *Transport Policy Studies* 12.2 (2009), pp. 2–9.
- [18] Kirk Emerson, Tina Nabatchi, and Stephen Balogh. “An Integrative Framework for Collaborative Governance”. In: *Journal of Public Administration Research and Theory* 22.1 (2012), pp. 1–29.
- [19] William Samuelson and Richard Zeckhauser. *Status Quo Bias in Decision Making*. Vol. 1. 1988, pp. 7–59.
- [20] Federica Russo et al. “Confirmation Bias in Policy Decision Making”. In: *Policy Sciences* 48.3 (2015), pp. 289–313.

- [21] Eli Ben-Sasson et al. “Succinct Non-Interactive Zero Knowledge for a von Neumann Architecture”. In: *Proceedings of USENIX Security Symposium*. 2014, pp. 781–796.

# Appendix A

## 付録

### A.1 協調ロボット制御モデルの数式展開

#### A.1.1 運動学モデル

N 関節ロボットアームにおいて、各関節の位置は以下の式で与えられる：

$$\mathbf{x}_i = \mathbf{x}_{i-1} + a_i \begin{bmatrix} \cos(\sum_{j=0}^{i-1} \theta_j) \\ \sin(\sum_{j=0}^{i-1} \theta_j) \end{bmatrix} \quad (\text{A.1})$$

ここで、 $\mathbf{x}_0 = \mathbf{0}$  である。

#### A.1.2 協調係数の計算

各関節の協調係数  $k_i$  は、以下の式で計算される：

$$k_i = \exp \left( -4 \ln(2) \frac{\|\mathbf{v}_{l,i} - \mathbf{v}_{i+1}\|^2 + \epsilon_1}{\|\mathbf{v}_{l,i}\|^2 + \epsilon_2} \right) \quad (\text{A.2})$$

ここで、 $\epsilon_1, \epsilon_2$  は数値安定性のための小さな正の定数である。

### A.2 Japan MaaS プロジェクト分析の詳細

#### A.2.1 分析対象プロジェクト一覧

2020 年度に認定された 38 の Japan MaaS プロジェクトを分析対象とした。プロジェクトは以下のカテゴリに分類される：

- 観光型 MaaS：XX 件
- 地域課題解決型 MaaS：XX 件
- 企業主導型 MaaS：XX 件

#### A.2.2 評価指標のコーディング基準

評価指標は以下の基準でコーディングした：

Table A.1: 評価指標のコーディング基準

カテゴリ	コーディング基準
事業指標	利用者数、収益、運行効率など
非事業指標	社会影响、環境効果、アクセシビリティ改善など
市民参加	ワークショップ、アンケート、協議会など

## A.3 ZK-SNARKs 型政策評価システムの実装詳細

### A.3.1 システム構成

ZK-SNARKs 型政策評価システムは、以下のコンポーネントから構成される：

- 秘匿化処理モジュール
- LLM 評価エンジン
- Constitutional AI 基準ライブラリ
- 控訴処理インターフェース

### A.3.2 技術的仕様

Table A.2: 技術的仕様

項目	仕様
LLM	GPT-4 / Claude
Temperature	0
Self-Consistency	5 回
TEE	AWS Nitro Enclaves