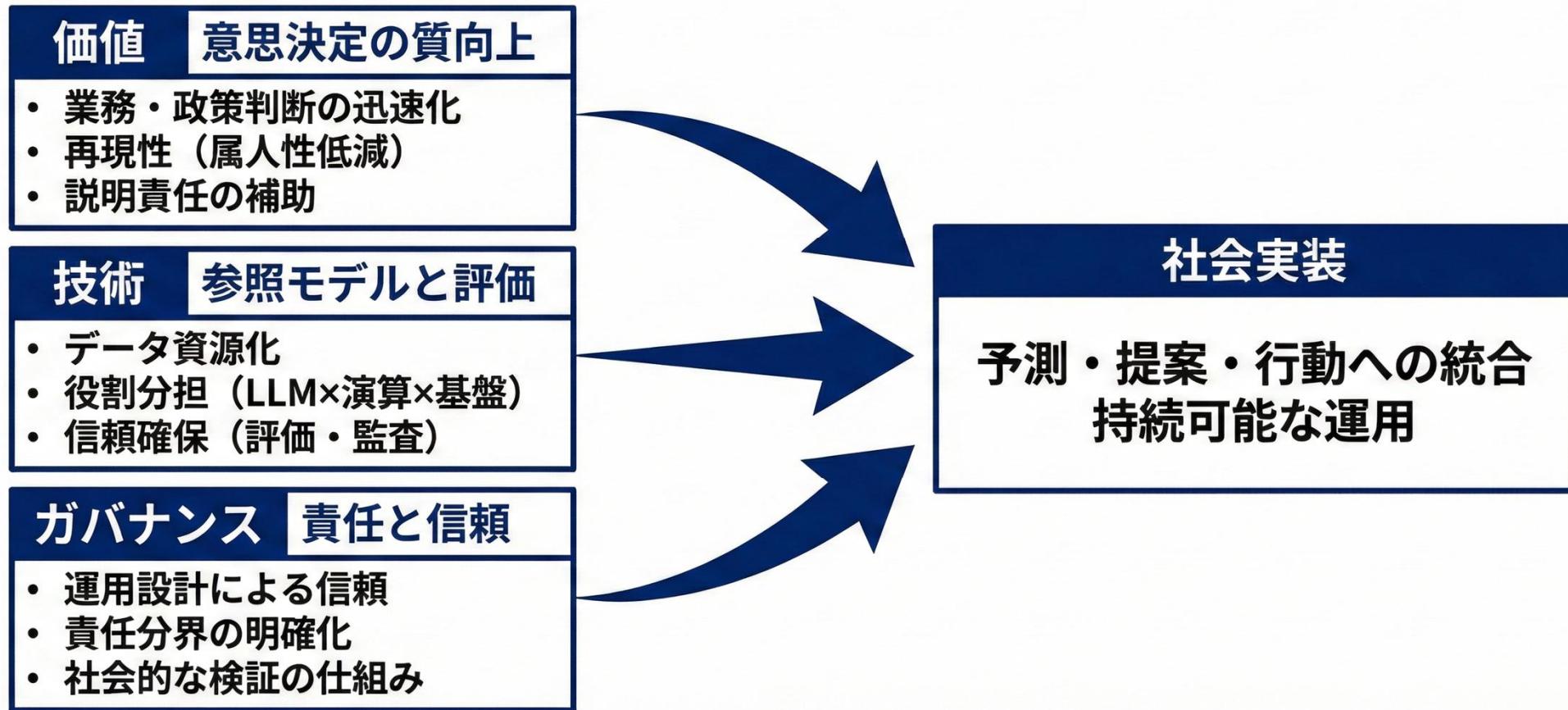


AIの観点からの論点提示： 価値・技術・ガバナンス

麗澤大学 工学部 教授／(一社)人工知能学会 理事
清田 陽司

AI観点からの論点提示

社会実装に向けた「価値・技術・ガバナンス」の統合的整理が必要



価値：ジオAIで“何が変わるか”（精度以外の価値軸）

「精度の追求」から「意思決定プロセスの質的転換」へ

・業務変革

- ・迅速化、再現性（属性の低減）

・政策支援

- ・説明責任の補助

・展開の方向

- ・判断の支援（提案・選択肢提示）まで視野に

・評価軸

- ・意思決定プロセスの質（運用負荷、監査可能性）

【事例1】Flood Hub: 洪水予測と早期警報（行動支援）



転換点：「当たる/外れる」の精度より、避難・配備のための「リードタイム確保」が価値

指標の変化：精度KPI → 行動KPI（予測期間[最大7日]）、対象カバレッジ、意思決定の再現性）

根拠：「行動のための時間」を提供する設計

引用元: <https://www.cities.google/flood-hub>

【事例2】ミャンマー地震: 衛星画像による発災直後の被害推定

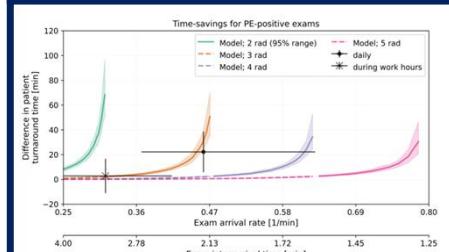


転換点：高精度な確定より、「どこから確認・派遣すべきか」の順位付けが重要

教訓：発災直後は「正確さ」より「速度」が最大の価値
前提：現地検証とセットで運用される初動判断ツール

引用元: https://satelliteimagergdemostg.z5.web.core.windows.net/damage-assessment/mandalay_earthquake_3_29_2025.html

【事例3】(参考) 医療画像におけるAIトリアージ（業務フロー改革）



転換点：診断の代替ではなく、緊急度に基づく「読影順序の並べ替え」でシステム全体の効率化

本質：AIが「答え」を出すのではなく、人間のための「優先順位」を設計する

成果：放射線科医のレポート作成までの時間(TAT)の短縮など、業務負荷軽減に寄与

引用元: Yee Lam Elim Thompson et al. Impact of AI-Triage on Radiologist Report Turnaround Time: Real-World Time-Savings and Insights from Model Predictions. 2025.
<https://arxiv.org/abs/2510.15237>

技術(1): 共通データ資源×共通評価が技術普及を加速

「モデルの進化」以上に「共通の評価を行えるデータ資源」が普及を左右する

- ・比較可能性の確保：同一データ・同一タスク・同一指標で「前進」が可視化
- ・参入障壁の低下：研究・実装が「再現可能」になり、参加者が増える
- ・ジオAIへの示唆：共通データ資源・評価手順の整備は協調領域として優先度が高い

事例	データ資源 (共有)	共通タスク／指標 (評価)	波及効果 (普及)																												
ImageNet	<table border="1"><caption>Data for ImageNet Accuracy Evolution</caption><thead><tr><th>Year</th><th>Layers</th><th>Accuracy (%)</th></tr></thead><tbody><tr><td>2010</td><td>shallow</td><td>26.2</td></tr><tr><td>2011</td><td>8 layers</td><td>25.8</td></tr><tr><td>2012</td><td>16 layers</td><td>16.4</td></tr><tr><td>2013</td><td>19 layers</td><td>11.7</td></tr><tr><td>2014</td><td>22 layers</td><td>7.3</td></tr><tr><td>2015</td><td>152 layers</td><td>6.7</td></tr><tr><td>2016</td><td>152 layers</td><td>3.6</td></tr><tr><td>2017</td><td>152 layers</td><td>5.1</td></tr></tbody></table>	Year	Layers	Accuracy (%)	2010	shallow	26.2	2011	8 layers	25.8	2012	16 layers	16.4	2013	19 layers	11.7	2014	22 layers	7.3	2015	152 layers	6.7	2016	152 layers	3.6	2017	152 layers	5.1	大規模ラベル付き画像データセット	分類・検出などの共通タスク (ILSVRC)	2012年のAlexNetを契機に深層学習の研究が加速
Year	Layers	Accuracy (%)																													
2010	shallow	26.2																													
2011	8 layers	25.8																													
2012	16 layers	16.4																													
2013	19 layers	11.7																													
2014	22 layers	7.3																													
2015	152 layers	6.7																													
2016	152 layers	3.6																													
2017	152 layers	5.1																													
TREC		テストコレクション (文書・トピック)	検索タスクの統一評価 (有効性測定)	技術の比較・再現性向上、知見の蓄積																											
SpaceNet		高解像度衛星画像+ 高品質ラベル	共通指標・チャレンジ (基盤地図抽出)	衛星×Computer Visionの 参入障壁低減、実装手法の 洗練																											

共有データ資源 + 共通タスク/指標 + 公開の場（チャレンジ/評価会） → 普及・投資・累積改善

引用元: <https://doi.org/10.1007/s00371-021-02280-6> / <https://pages.nist.gov/trec-browser/> / <https://spacenet.ai/datasets/>

技術(2)-①: 役割分担の重要性 (先行事例)

学習モデル（直感・生成）+ 外部要素（探索・検索・ツール）で、性能と信頼性を両立

- ・ 単体最適ではなく構成最適：モデルが得意な部分と、外部要素が得意な部分を分ける
- ・ 外部要素が担うもの：厳密計算／根拠参照／制約充足（=説明・監査の基盤）
- ・ ジオAIへの示唆：LLMは「手順化・説明」、厳密計算はGIS/DB、根拠はデータ基盤で担保

AlphaGo (囲碁AI)		RAG	Toolformer
<p>価値の学習 探索</p>		<p>検索 生成</p>	<p>API呼び出し 判断</p>
分担	方策・価値（学習）・探索	生成（LLM）／検索	言語モデル（判断）／API（計算・検索・照会）
外部が担う	探索 = 検証・枝刈り (手の良さを“計算で裏取り”)	根拠参照 (最新化・出典付与・再現性)	決定的処理 (計算の厳密化・ログ化)
効果	学習の直感 + 探索の厳密化 (性能向上・安定化)	根拠に基づく生成 (説明可能性・更新容易性)	誤り低減・手続き化 ("いつツールを使うか"を学習)

引用元: <https://doi.org/10.1038/nature16961> / <https://doi.org/10.48550/arXiv.2005.11401> / <https://doi.org/10.48550/arXiv.2302.04761>

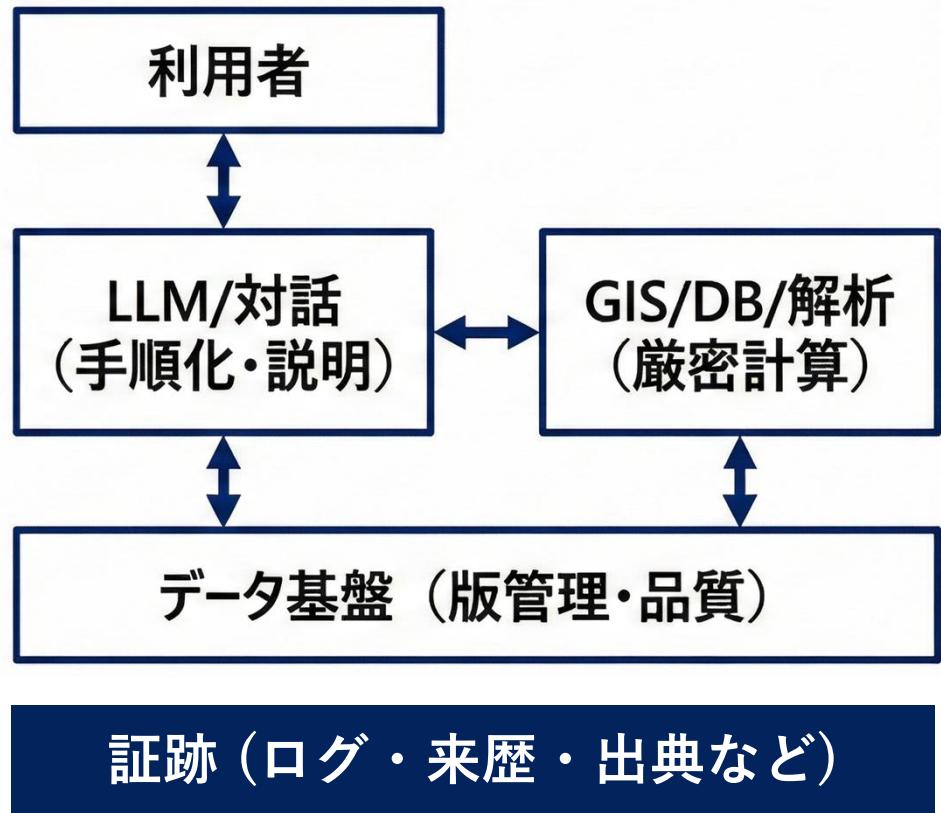
技術(2)-②: 参照モデル上の“役割分担” (ジオAIへの適用)

LLMとGIS/DBの分界点 + データ基盤 (版・来歴) + 証跡 (ログ) を設計要件に

- **LLM (対話・手順化・説明)**
問い合わせの分解／手順提示／結果の要約
- **GIS/DB/解析 (厳密計算)**
座標計算／空間検索／シミュレーション
- **データ基盤 (版管理・品質)**
来歴／更新／品質メタデータ
- **証跡 (ログ・出典)**
「どのデータで」「どう計算したか」の自動保持

先行事例との対応

- AlphaGo : 学習 (方策・価値) + 探索 (厳密化)
- RAG : 生成 + 検索 (根拠参照)
- Toolformer : 言語 + API (決定的処理・ログ)
→ ジオAI : LLM + GIS/DB + データ基盤 (証跡)



技術(3): 評価・運用

精度以外の「頑健性」と「不確実性」の扱いを設計に組み込む必要性

• 外挿・頑健性

未知地域／欠損／
観測条件差(ドメインシフト)での
性能維持

• 不確実性表示

確信度に応じた注意喚起／判断保留
／人的確認への接続

• 繼続評価

運用中の分布変化
(ドリフト) 検知
と再学習(アラート・運用ループ)

評価観点	確認・設計事項	事例
外挿・頑健性	<ul style="list-style-type: none"> 観測条件差(装置・施設・地域)で性能低下しないか 適用範囲・限界の明示(対象地域/センサ/季節等) 	<p>医療画像: 装置・施設差で性能低下(ドメインシフト)</p> <p>引用元: Caio César Dias Resende et al. Influence of operator experience, scanner type, and scan size on 3D scans. J. Prosthetic Dentistry 125(2), 2021. https://doi.org/10.1016/j.jprosdent.2019.12.011</p>
不確実性表示	<ul style="list-style-type: none"> 根拠が薄い場合の「注意喚起」「判断保留」 人的確認・追加観測への分岐(トリガ設計) 	<p>災害AI: 暫定推定+現地確認(不確実性を前提)</p> <p>引用元: Shahrzad Gholami et al. On the Deployment of Post-Disaster Building Damage Assessment Tools using Satellite Imagery: A Deep Learning Approach. SSTD 2022. https://doi.org/10.1109/ICDMW58026.2022.00134</p>
継続評価	<ul style="list-style-type: none"> 運用中の性能劣化(ドリフト)検知 再学習・モデル更新の運用(アラート/周期/条件) 	<p>Google Flu Trends: 運用中のドリフト(監視・再学習が必要)</p> <p>引用元: Keith Winstein. The 2012-2013 Divergence of Google Flu Trends. 2013. https://cs.stanford.edu/~keithw/www/Winstein-Slides-14March2013.pdf</p>

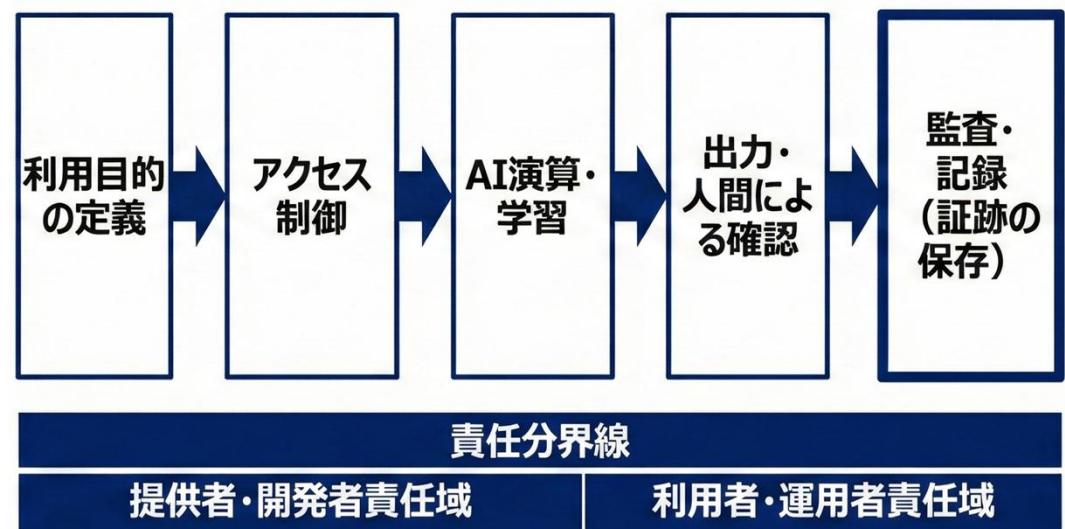
ガバナンス：信頼確保は「規制対応」ではなく「運用設計」

責任分界を明確にし、社会が「検証」できる仕組みを構築する

- **目的**：データ流通と活用を両立する
「信頼のインフラ」
- **手段**：用途別アクセス制御／処理ログ・来歴の明示／監査可能性
- **論点**：制度（法律・指針）と現場（運用）の「中間領域」の設計

先行分野の“二層構造”例（法規制 + 自主規制／標準）

分野	法規制(最低基準)	自主規制/標準(実務ルール)
金融	米：SEC監督/日：金融商品取引法	FINRA (SRO) ／日本証券業協会（自主規制）
航空	FAA (認証・承認)	RTCA/DO-178C 等 (開発保証の標準)
建築	建築基準法：確認検査／定期報告等	日本建築学会JASS (標準仕様書)等：品質確保の実務標準



二層構造で具体化：制度（最低基準）+標準／自主規制（運用ルール）→「検証可能性」の担保

議論の入り口: 3つの問い合わせ (To-beに向けた論点)

実装を加速させるための「ギャップの解消」に向けて

問い合わせ1: AI学習の前提となる
「最小限のデータ整備要件」
(フォーマット・メタデータ・
アノテーション)は何か

問い合わせ2: 参照モデル上の「協調
領域／競争領域」の切り分け
(どこを共通化するか)

問い合わせ3: 用途別の「許容不確実
性」と「責任分界」を誰がど
う定義し、どう運用するか

当面の成果物 (たたき台)

- a. データ整備要件
- b. 参照モデル (構成・証跡管理)
- c. 評価・運用指針 (用途別)