

人工知能 (AI) サービスのミスは、ブランドの信頼にどのように影響しますか？

台湾大学 ○王 淑美 台湾海洋大学 黄 昱凱 趙 時樑

シンガポール社会科学大学 謝 承憲

武蔵野大学 張 巧韵 中国科技大学(台湾) 李 福斯

1. まえがき

人工知能 (Artificial Intelligence, AI) は、現代における技術革新と経済発展の主要な推進力の一つである。AI の概念は 20 世紀中頃にまで遡ることができ、その核心的な目標は人間の知能を模倣し、自動的な意思決定、学習、そして問題解決能力を実現することである。近年、計算能力の飛躍的な向上、大規模データの普及、アルゴリズムの急速な進化に伴い、AI 技術は理論研究の段階を超えて、実用的応用が急速に拡大する発展期に入っている。現在、AI の応用範囲は音声認識、自然言語処理、画像解析、自動運転、医療診断など多岐にわたり、人間の生活様式や経済構造に深甚な影響を及ぼしている。さらに、AI の発展は伝統的産業のデジタル化転換を促進するだけでなく、スマートシティ、スマート交通、スマート製造といった新たな応用シナリオを生み出している。

しかしながら、サービス産業における AI の応用にはいくつかの課題も伴っている。第一に、AI の運用はデータに強く依存しているため、データのプライバシーとセキュリティの問題が重要な論点となる。第二に、AI システムがサービス過程において故障や偏差を起こした場合、顧客の不満や信頼危機を招く可能性がある。特にサービスの失敗に直面した際には、顧客のブランドに対する信頼が深刻に損なわれるおそれがある。したがって、企業は AI に起因する問題に対応するために、効果的なサービス回復戦略を設計する必要がある。例えば、人間のカスタマーサービスを組み合わせたハイブリッド型の回復策や、即時フィードバック機構を活用して迅速にサービス欠陥を修復する方法が挙げられる。これらの課題は、AI を応用する際に同時に堅牢な回復メカニズムを設計することの重要性を浮き彫りにしている。

ホテル産業はサービス業における重要な分野であり、AI 技術の応用において先駆的な役

割を果たしてきた。その活用は、スマート予約・自動チェックイン・チェックアウト、音声アシスタント、さらには施設のインテリジェント管理にまで及び、業務効率と顧客体験の向上に寄与している。しかし、AI への過度な依存はサービスの人間味を損なう可能性があり、誤った予約情報や機器の不具合は顧客の不満やブランド信頼の低下を招き得る。したがって、ホテル業界は迅速なカスタマーサポートや補償制度など、効果的な回復戦略を整備し、効率性とサービス品質のバランスを図る必要がある。本研究は、AI サービスの失敗がブランド信頼に与える影響と、その際における補救戦略の調整効果を検討することを目的とする。

2. 研究方法

カタストロフィー理論 (Catastrophe Theory, CT) は、フランスの数学者ルネ・トム (René Thom, 1923 - 2002) によって 1970 年代に発展させられた数学的枠組みである。従来の微分方程式研究が注目してきた焦点とは異なり、カタストロフィー理論は特に方程式構造の安定性、動力学的安定性 (パラメータの変化に伴う系の定性的性質の変化)、および臨界点集合といった特性に焦点を当てる。そして、方程式の解に内在する定性的現象が、方程式中のパラメータの変化によってどのように不連続的な挙動を示すかを記述することを目的としている。カタストロフィー理論が主に探究する課題は、あるシステムが複数の定常状態 (steady state) を持つ場合に、システムパラメータの変化に応じて、どのように一つの定常状態から別の定常状態へ移行するのか、その過程を解明することである。

ここでは、カタストロフィー理論の数学的構造を簡潔に説明する。あるシステム S を想定し、 n 個の状態変数 (x_1, x_2, \dots, x_n) によってその時点での状況を表す。一方、 m 個の制御変数 (c_1, c_2, \dots, c_m) によってシステムの状態

How does AI service failure impact brand trust?

Shu-Mei WANG, Yu-Kai HUANG, Cheng-Hsien HSIEN, Connie CHANG and Fu-Ssu LEE

が決定される。制御変数の組み合わせによって、複数の安定状態が対応し、定点アトラクタ（例：P1・P3）や有限周期アトラクタ（例：P2）が形成され得る。システム S が動的システムである場合、これらの安定状態は安定平衡点に相当する。カタストロフィー理論が注目するのは、状態変数が制御変数の変動により不連続かつ突発的に変化する現象である。

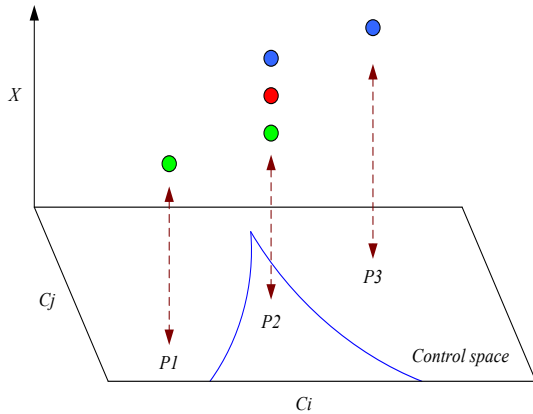


Fig.1 カタストロフィー・モデルの特質

カタストロフィー理論では、システム S にポテンシャル関数 $f(x, c)$ が存在すると仮定する。このとき、システム S の力学系は、ポテンシャル関数の勾配ベクトル場（gradient vector field） ∇f によって表すことができる。

$$\text{grad } f(x, c) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

上記の力学系は、次のように書き換えることができる：

$$\frac{dX}{dt} = -\text{grad } f(x, c) = -\frac{\partial f(x, c)}{\partial x_i}, \quad i = 1, 2, \dots, n$$

システムが長期的な変動の下で、その状態変数 X がある一点に収束し、かつその位置に維持される場合、 X はシステム S の平衡状態と呼ばれ、 e_i で表される：

$$e_i : \frac{dX}{dt} = -\text{grad } f = -\frac{\partial f}{\partial x_i} = 0, \quad i = 1, 2, \dots, n$$

平衡状態には安定（attractor）と不安定（repellor）があり、前者はその平衡状態の近傍にある他の状態が軌道に沿って徐々にこの平衡点へ収束することを意味する。逆に、近傍の状態が平衡点から離れていく場合は、不安定な平衡点と呼ばれる。カタストロフィー理論が特に関心を持つのは、システムに複数の安定平衡状態が存在する場合に、制御変数の変化によ

ってこれらの安定平衡点がどのように不連続な動的变化を示すかという点である。

システムの安定性は、ヘッセ判別式によって判定することができる：

$$\text{attractor} : \det\left(\frac{\partial^2 e_i}{\partial x_j^2}\right) \geq 0$$

$$\text{repellor} : \det\left(\frac{\partial^2 e_i}{\partial x_j^2}\right) < 0$$

ヘッセ判別式が 0 以上の場合、システムには安定平衡点のみが存在し、それはアトラクタとして記述される。ヘッセ判別式が 0 未満の場合には、安定平衡点に加えて不安定平衡点、すなわちリペラも現れる。アトラクタは系が平衡状態に収束する特性を示すのに対し、リペラは平衡状態が存在しても極めて不安定であり、わずかに逸脱すれば系はその平衡から大きく離れてしまうことを意味する。したがって、システム S に安定平衡点が存在する限り、カタストロフィー理論を用いてその挙動を分析することが可能である。

カusp・カタストロフィー・モデル（Cusp Catastrophe Model, CCM）は構造が単純でありながら多くの非線形挙動を説明できるため、最も頻繁に利用されるカタストロフィーモデルである。本モデルは、2 つの制御変数 (u, v) によって 1 つの状態変数 (x) を記述し、そのパラメータ空間は「制御空間」と呼ばれる。CCM のポテンシャル関数は次のように表される。

$$F(u, v, x) = -\frac{1}{4}x^4 + \frac{1}{2}ux^2 + vx$$

制御パラメータの組み合わせによって、ポテンシャル関数は異なる構造を持つ。その安定解はポテンシャル関数を微分することで求められ、安定曲面 MF として表される。：

$$\frac{\partial F}{\partial x} = -x^3 + ux + v = 0$$

$$M_F : \{(u, v, x) \mid -x^3 + ux + v = 0\}$$

對上式求算赫斯判別式・並令赫斯判別式為零可以得到奇異點集K(singularity set)：

$$3x^2 + u = 0$$

上記の二式から状態変数を消去することで、分岐集合 B を得ることができる：

$$4u^3 = 27v^2$$

分岐集合 B はカルダン判別式 (Δ) とも呼ばれ、方程式が複数の安定解を持つ場合の制御変数の集合を指す。前式から分かるように、分岐集合 B は CCM の制御変数によって構成される。カタストロフィー理論において、この 2

つの制御変数には特別な名称があり、 u は「分裂因子 (splitting factor)」、 v は「正則因子 (normal factor)」と呼ばれる。

3. モデル設計と分析

本研究はカタストロフィー理論 (Catastrophe Theory) を用いて、人工知能 (AI) サービスの失敗がいかんしてブランド信頼の非線形的崩壊を引き起こすのか、そして補償戦略がその過程においてどのような調整的役割を果たすのかを検討する。モデル設計においては、AI サービスの特性に即して各変数を明確に定義する。

状態変数 (x): ブランド信頼。ブランド信頼を最終的な結果変数とし、通常時における安定性と、特定条件下での急激な低下または回復といった非線形的特徴を表す。

分裂変数 (u): AI サービス失敗の知覚された深刻度。この変数は金銭的損失といった実際の結果だけでなく、AI 特有の不可解性 (ブラックボックス性) や不可制御性を包含する。知覚された深刻度が高まると、ブランド信頼は二重安定領域に入りやすくなり、高水準と低水準の間で突発的な崩壊や跳躍が生じる。

偏差変数 (v): AI 補償の透明性と説明性。従来の補償と異なり、この変数は補償策が AI の意思決定過程をどの程度透明化・説明可能化 (XAI) できるかに焦点を当てる。高い透明性を備えた補償策は、ブランド信頼の全体的水準を高めるだけでなく、低水準からの迅速な回復を促進する。

図2はカusp・カタストロフィー・モデル (Cusp Catastrophe Model) の応用を示しており、人工知能 (AI) サービスの失敗がブランド信頼の非線形的な「崩壊」および「回復」現象をどのように引き起こすかを説明している。本モデルは、1つの状態変数 x と2つの制御変数 u および v によって構成され、システムの挙動を表す三次元曲面を構築している。

本カタストロフィーモデルの特性に基づく、分析の核心は、AI サービス失敗後におけるブランド信頼 x の非線形的変化メカニズムを明らかにする点にある。すなわち、AI サービス失敗の知覚された深刻度 u が臨界値に達すると、信頼システムは二重安定領域に入り、そこでは信頼がヒステリシス

(hysteresis) と突発的ジャンプ (sudden jump) の両特性を示す。具体的には、深刻な AI 失敗は信頼の急激な崩壊を引き起こし、一度信頼が崩壊すると、企業は過剰な「AI 補償の透明性と説明性」 v を提供しなければ、信頼を低水準から高水準へと跳躍的に回復させることができない。このことは、信頼の維持よりもその回復がはるかに困難であることを示している。

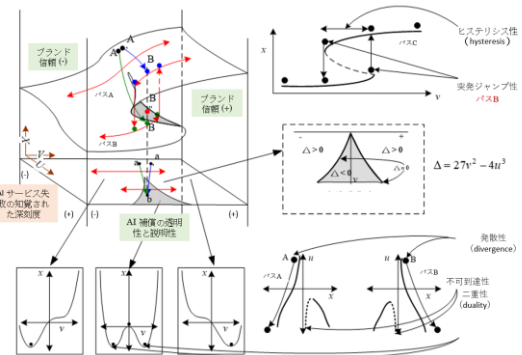


Fig.2 本研究において構築された理論モデル

4. 総合討論

本稿で構築した理論モデルは、AI 失敗の知覚された深刻度がブランド信頼に非線形的な変化をもたらす可能性を明らかにし、さらに補償戦略の透明性および説明性が、信頼の維持水準や崩壊・回復の臨界点に影響を及ぼすことを示している。本研究は探索的な質的分析であり、モデルの特性を参照しながら、AI サービス失敗がブランド信頼に与える影響と、その過程における補償戦略の調整効果を検討するものである。

人工知能 (AI) サービスにおける過失がブランド信頼に与える影響、およびその修復戦略の調節作用を研究するにあたり、「破局モデル (Catastrophe Model)」は、非線形な動態関係や突発的な行動変化を捉えるための革新的な分析フレームワークを提供します。このモデルは、システムの多重平衡状態の検討、特にブランド信頼という複雑な心理的・行動的メカニズムを研究する際に適しており、従来の線形モデルでは明らかにできない深層的な現象を解明するのに役立ちます。破局モデルは、AI サービス過失がブランド信頼に及ぼす影響の程度を定量化し、修復戦略が調節変数としてこのプロセスにどのように影響するかを捉えるのに役立ちます。具体的には、結果変数であるブランド信頼の変化は、AI サービス過失の直接的な影響を受けるだけでなく、修復戦略の介入

方法と効果にも依存します。カタストロフィー理論モデルの適用により、以下の状況を特定することが可能となります。

(1) 安定状態と破局点:破局モデルを通じて、AIサービス過失の発生によってブランド信頼の安定状態がどのように変化するか、また、いかなる条件下で劇的な信頼の崩壊(破局)が発生するかを検出できます。例えば、修復戦略の強度や効率が不十分な場合、ブランド信頼がある臨界点において突然低下する可能性があります。

(2) 非対称性と不可逆性:ブランド信頼の修復プロセスは通常、非対称性を伴います。すなわち、信頼の構築には長期的な努力が必要である一方で、一度の過失によって瞬間的に崩壊する可能性があります。破局モデルは、この非対称な特性を効果的に捉え、信頼回復における課題と重要な要因を特定するのに役立ちます。

(3) 調節効果の非線形表現:調節変数としての修復戦略の効果は、異なる状況に応じて変化する可能性があります。例えば、軽微な過失の下では、適切な修復措置が顧客の不満を効果的に緩和する可能性があります。重大な過失の場合、修復が講じられても、信頼の回復は過失の深刻度や顧客の感情的な反応に制限される可能性があります。破局モデルは、これらの非線形な影響を可視化し、企業が修復戦略を最適に実行するための条件を理解するのに貢献します。

カタストロフィー・モデルは、研究者に包括的かつ柔軟な分析ツールを提供し、AIサービスの失敗とブランド信頼の間に存在する非線形的関係を明らかにするだけでなく、救済戦略の調整効果を定量化することも可能にする。本モデルを応用することで、ブランド信頼の動態的变化に内在するメカニズムをより深く解明し、企業が効果的な救済戦略を策定するための科学的根拠を提供することができる。カタストロフィー・モデルの利点は以下の通りである。

(1) 非線形関係の把握:従来の線形回帰分析と比較して、カタストロフィー・モデルはAIサービス失敗がブランド信頼に及ぼす影響過程における非線形的ダイナミクスを明示し、さらに異なる影響要因の作用強度や臨界値を具体的に定量化することができる。

(2) 突発的変化の強調:カタストロフィー・モデルは、特定の条件下におけるブランド信頼の急激な変化を効果的に分析することができる。これは、サービス救済の成否やそれがブランド信頼に与える影響メカニズムを探究する上で極めて重要である。

(3) 多変数の統合:カタストロフィー・モデルは複数の重要変数とその相互作用を同時に考慮することを可能にし、AIサービス失敗(統制変数)、救済戦略(調整変数)、ブランド信頼(結果変数)の間に存在する複雑な関係を分析するのに特に適している。

5. まとめ

本研究は、カスプ・カタストロフィー・モデルを用いて、AIサービス失敗とブランド信頼との間に存在する非線形的関係を明らかにした。結果として、顧客がAI失敗の深刻度を臨界値として知覚した場合、ブランド信頼は突発的に崩壊し、その回復には通常を超える補償努力が必要であることが示された。さらに、補償戦略の透明性と説明性は信頼水準を高めるだけでなく、臨界点を調整し、信頼崩壊のリスクを低減させる役割を果たすことが確認された。経営的示唆として、企業は補償情報の透明化および説明可能性を優先的に強化し、段階的な補償メカニズムを構築するとともに、AI失敗の予防に注力すべきである。なぜなら、信頼の回復は維持よりもはるかに困難であるからである。本研究は、AI時代におけるブランドマネジメントが、技術的信頼性と補償戦略の双方に注目し、顧客関係の安定を確保する必要性を強調している。

参考文献

- 1) Jeong, E., & Lee, D. (2025). AI-human collaboration in services: an integrative framework to uncover the key success factors. *Service Business*, 19(3), 20.
- 2) Lv Xingyang, Yang Yufan, Qin Dazhi, Cao Xingping, Xu Hong (2022), "Artificial intelligence service recovery: The role of empathic response in hospitality customers' continuous usage intention," *Computers in Human Behavior*, Vol.126, 106993.
- 3) Sethi, V., & King, R. C. (1998). An application of the cusp catastrophe model to user information satisfaction. *Information & Management*, 34(1), 41-53.
- 4) Song, M., Xing, X., Duan, Y., & Mou, J. (2023). I can feel AI failure: the impact of service failure type and failure assessment on customer recovery expectation. *Industrial Management & Data Systems*, 123(12), 2949-2975.