

信用スコアの課題と今後

慶應義塾大学法科大学院教授

山本龍彦
やまもと たつひこ



信用スコア活用の可能性

AIがさまざまなデータから個人の社会的信用力を予測・数値化する信用スコアリング。日本でも、ソフトバンクとみずほ銀行が共同設立したJ.Score、Yahoo!スコア、LINE Scoreなどが同サービスを開始し、一般化する兆しを見せている。信用スコアは、個人向け融資の審査等に役立てられるだけでなく、連携先による広告配信の最適化やマッチング、高スコア保持者への優遇特典の提供などに利用されるという。今後、AIの予測精度が向上し、スコアに対する信頼が高まれば、「スコア先進国」中国のように、さまざまな社会領域で利用されていくことが予想される。

このような信用スコア。もちろんメリットも大だ。ビッグデータ解析により、「すべてのデータは信用データとなった」といわれるように、信用力の算出に使える情報は増えた。その恩恵にあずかり、個人のあらゆる行動記録を算出に利用することで、伝統的な信用情

報を持たない若者や少数派への融資を可能にするなど、「ファイナンシャル・インクルージョン」を実現させる。また、データ・ポイントの拡充・多様化によって、より正確な信用力の予測が可能となるうえ、こうして測られた信用力に見合うサービスや特典を享受できることで、個人の努力を正しく反映した公正な社会システムが構築されることにもなる。さらに、信用スコアがデジタルプラットフォーム上の正確な信頼指標となることで、より安全なオンライン取引やシェアリングエコノミーが実現するかもしれない。

ブラックボックス問題とスコアリングプロセスの閉鎖性による課題

他方で、現状のサービスには課題も多い。例えば、スコアの算出にどのような情報が使われたのか、どの情報にどれぐらいの比重がかけられたのかが明らかではない。いわゆる「ブラックボックス問題」である[L.I.N.E Score]のユーザー向けHPでは、「[L.I.N.E] 関連サ

ービス上の利用傾向」やサービス利用前になされる質問への回答データが算出に利用されるといふ。通信の秘密に該当するメッセージ内容等は利用しないとあるが、それ以上の具体的記述はない。

もちろん、フィードバックにより算出用データやアルゴリズムは常に変化し得ること、理解可能なかたちで完全に透明化すると、ユーザーはそれを踏まえて行動を調整するようになるため、予測精度が落ちること、アルゴリズムは知的財産として保護され得ることなどから、具体的な開示には限界がある。しかし、閉鎖的なスコアリングのプロセスは、さらに以下のような問題を生む。

第1に、本人のIDを使った行動がすべて本人の行動である保証はないため、他人の行動記録が本人のスコアに影響を与える可能性がある。第2に、どのような行動がスコアに影響を与えるかが不透明だと、不安から特定の行動を差し控えるようになる。また、一般的な道

図表1 EU、米国における信用スコアへの法的対応

EU General Data Protection Regulation, GDPR	
• 人間が介在せずに最終的な決定を自動処理のみにより行われたい権利(完全自動意思決定の原則禁止) ※AIが自動算出したスコアのみで融資を拒否することは原則許されない	
• 例外的に許容される場合でも、①人間の介在を得る権利、②自らの見解を表明する権利、③AI等の自動処理のみによる決定を争う権利を保障する必要がある。	
• 事業者は完全自動意思決定の存在、当該決定のロジックについて有意義な情報(算出用データの源泉や重み付け)等を告知しなければならない。	
米国 Equal Credit Opportunity Act, ECOA	
• 与信判断に際して、人種・宗教・出身国・性別・妊娠状況・年齢等に基づき差別することを禁止。これらの属性に基づく「異なる取り扱い(disparate treatment)」を禁止。	
• 一見中立的だが、結果的に少数派に「過度に不利な影響(disproportionate adverse effect)」を与える実践も原則禁止。	

図表2 信用スコアの取り扱いに関する日本の原則・考え方

人間中心のAI社会原則(内閣府、2019年3月)
• プライバシー確保の原則 (第3原則: AIを利用したプロファイリングについて、単なる個人情報扱以上の慎重さが求められる場合がある)
• 公平性、説明責任及び透明性の原則 (第6原則: 不当な差別が生じないよう、公平性及び透明性のある意思決定とその結果に対する説明責任が適切に確保されると共に、技術に対する信頼性が担保される必要がある)
※年内に公表が予定される「AI活用ガイドライン」(総務省)も、尊厳・自律の原則や公平性の原則のなかで、AIを利用したプロファイリングを行う場合における不利益への配慮や、人間の判断の介在について検討を求めている
情報信託機能の認定スキームの在り方に関する検討会 とりまとめ(案) (総務省・経済産業省、2019年8月)
• 情報銀行において「信用スコア」を取り扱う場合の留意点を記載
• (例)情報銀行は、個人情報信用スコアの算定に利用されること及びそれによるリスクについて明示的に説明する。
• (例)算定につき、遺伝情報や差別に繋がる過去の情報を用いない。スコアに用いたデータ及びスコアの算出方法についてアカウンタビリティをもつ。人間の関与について検討する。

※信用スコアの取り扱いに関連する規律として、現状でも、「金融分野における個人情報保護に関するガイドライン」があるが、ビッグデータやAIを用いたスコアリングや与信判断を前提にしたものではなく、見直しが必要

「原則」が、信用スコアとも関係する諸原則を打ち出し、総務省が情報銀行における信用スコアの取り扱いに関する考え方を提示している。とはいえ、ルール化はそこにとどまる。「原則」や「考え方」は立派だが、信用スコアを具体的に規律するルールがないとなれば、データ連携を深めるEUや米国からの信頼も失うことになる。何より、「リクナビ問題」がその危険性を示唆したように、スコアというデジタル上の虚像が一人歩きし、個人の人生の機会が不当に奪われることになりかねない。他国の法制度を参照しながら、「原則」や「考え方」を踏まえた具体的なルールの策定を急ぐべきである。

徳規範や事業者が推奨する行動に自らを適合させるべきとの同調圧力が生じ、行動の自由や社会の多様性・活力が失われ得る。第3に、これまで存在してきた社会の差別構造が、アルゴリズムを設計する際のデータセットの偏りや、性別等のセンシティブ属性の使用を通じてスコアに反映し、差別構造が固定化ないし悪化する可能性がある。センシティブ属性そのものを使わなかったとしても、それと密接に関連する情報(代理変数)が使われることで、結果的に差別的な影響(disparate impact)が生じることもある。第4に、スコアの利用範囲が拡大すると、低スコア保持者が社会の至る所で冷遇される

うえ、スコアの算定基準が不透明なためにスコアアップの方法もわからず、その「身分」が社会の下層で固定化してしまう可能性がある。スコアに応じた新たな身分制が生まれ、最下層の者が仮想上身を寄せ合うパーチャル・スラムが形成され得るといふわけだ。

具体的なルール策定が急がれる日本

このように見ると、信用スコアには法的に対応しなければならぬ重要課題も少なくない。世界的に、すでにこうした課題に 대응する法的規律が用意され始めている。例えば、EUのGDPR(一般データ保護規則)は、AIの予測評価のみで重要な決定を行うことを禁

止している(完全自動意思決定の原則禁止)。本人の明示的同意などにより例外も許容されているが、その場合にも図表1にあるような権利を保障する必要がある。その背景には、予測精度をある程度犠牲にしてでも、公正さと透明性を確保し、スコアに対する本人関与を認めるべきとのEUの価値判断がある。

米国でも、ECOA(信用機会均等法)が与信判断における差別を禁止しているうえ、近年では、信用報告の利用に際して適正手続きを求めるFCRA(公正信用報告法)の射程を拡大すべきとの議論も見られる。

図表2のように、日本でも、今年3月に公表された内閣府の「人間中心のAI社会原則」が、信用スコアとも関係する