

2019年8月8日

消費者委員会本会議ヒアリング

## プロファイリングと消費者保護

山本龍彦（慶應義塾大学）

5

### I. はじめに——2つの事例

#### Case A

ある統計調査により、女性はうつ状態にあるときに化粧品の購入傾向が強くなるということが明らかとなった<sup>1</sup>。また、女性がうつ状態にあるかどうかを予測するアルゴリズムが構築された<sup>2</sup>。そこで、化粧品会社 Y は、スマートフォン・ユーザーである X の行動記録をネットワーク上で網羅的に収集し、X の性別・年齢・職業及びその精神状態等をプロファイリングした。この結果、X が化粧品を購入しうる年齢層に含まれる女性で、いままさにうつ状態にあることが予測されたため、X に対して化粧品のネット広告を配信した。

15

#### Case B

メタボリック症候群を気にし始めた 30 代男性である X は、ネット上でダイエット関連の商品をチェックしたり、見学のためトレーニング・ジムに立ち寄ったりしていた。ダイエットサプリ等を扱う健康食品会社 Y は、X の閲覧記録や行動履歴（位置情報）等を収集しており、これらの情報を用いたプロファイリングから、X がメタボを気にする 30 代男性であることを把握していた。そこで Y は、検索エンジン等のサービスも行うインターネット広告事業者 A に、同社のポータルサイト上で、一般的なニュースとともに、メタボの健康への危険性を報じるニュースを、X に対してのみ数日間にわたって集中的に配信するよう依頼した。

Y は、X がこのニュースを閲覧し、メタボへの不安を増大させていることを他の行動履歴等から確認したうえで、X に対してダイエット関連商品のネット広告を配信した。

○Ryan Calo の議論（*Digital Market Manipulation*, 82 GEO. WASH. L. REV. 995(2014)）  
→プロファイリングに基づく選択的な情報フィード（選択環境の操作）によって、事業者は  
「カモを待つのではなく、カモを自ら作り出すことができる」（“prime vulnerability”）

<sup>1</sup> See e.g., Lucia Moses, *Marketers Should Take Note of When Women Feel Least Attractive, What Messages to Convey and When to send them*, AdWeek(Oct. 2, 2013)..

<sup>2</sup> Munmun De Choudhury, Scott Counts & Eric Horvitz, *Predicting Postpartum Changes in Emotion and Behavior via Social Media*, Microsoft Research(2013).

moment”の発見と創出可能性)。

※Mattha Busby, *Revealed: how bookies use AI to keep gamblers hooked*, The Guardian, 30 April 2014.

この記事は、イギリスのギャンブル業界が、AIを用いてギャンブル依存症の者を予測・発見し、彼らを「はまらせておく」ための広告を送信している実態を批判的に検討している。また、こうしたAIの利用  
5 に対し、規制の必要性を説くイギリス議会議員の見解も紹介している。

※「オンラインゲームのアルゴリズムは、常習性を作り出し、これを維持するためにデザインされている。オンラインショッピングのウェブサイト GILT は、顧客を引っかけておく (hook) ために、いわゆる『マイクロ・クリフハンガー』を使っていると言われている」(Gerhard Wagner & Horst Eidenmüller<sup>3</sup>, *Down  
10 by Algorithms? Siphoning Rents, Exploiting Biases, and Shaping Preferences: Regulating the Dark Side of Personalized Transactions*, 86 U. CHI. L. REV. 581(2019); ADAM L. ALTER, *IRRESISTIBLE: THE RISE OF ADDICTIVE TECHNOLOGY AND THE BUSINESS OF KEEPING US HOOKED*(2017).

※オハイオ州立大学の実験<sup>4</sup>：個人の趣味嗜好を予測（プロファイリング）して送られる個別化広告が、個人の「自意識 (self-perception)」自体に重大な影響を与える（感情操作）。この実験によれば、被験者とな  
15 った大学生は、自分のオンライン上の行動の結果として送られてくる個別化広告を——それが本当に自分の性向とマッチしているかどうかにかかわらず——「自己の反映 (reflection of the self)」として認識する傾向があるという。個別化広告が示す特性を、「もともと自分もっていた特性」であると認識してしまう  
20 というのである。実験によれば、例えば、被験者のために「個別化」された広告として、環境保護的なメッセージを含む広告を受け取った被験者は、自分自身を環境保護に熱心な人間であったと評価づけし、その後、環境にやさしい商品を購入する傾向、環境保護に向けた慈善活動に募金する傾向が高まったと報告されている。

※表情解析技術の発展

25 ○今後の消費者モデルは、「媒介される消費者 (mediated consumer)」  
→「他の誰かによってデザインされた技術を媒介に市場と相対する消費者」：特徴として、  
①継続的・網羅的監視、②詳細な特性把握、③選択環境の調律、④時間的優位性（最も脆弱な瞬間にアプローチできる）

30

---

<sup>3</sup> それぞれフンボルト大学教授（民法）、オックスフォード大学教授（消費者法）。

<sup>4</sup> Rebecca Walker Reczek et al., *Targeted Ads Don't Just Make You More Likely to Buy: They Can Change How You Think About Yourself*, HARVARD BUSINESS REVIEW(April 4, 2016) at <https://hbr.org/2016/04/targeted-ads-dont-just-make-you-more-likely-to-buy-they-can-change-how-you-think-about-yourself>

## II. いくつかのアプローチ

### 1. プライバシー権(不法行為法)による対応(入口規制)

5 ※「一般の人が……当該私人の私生活であると誤認しても不合理でない程度に真実らしく受け取られるものであれば、それはなおプライバシーの侵害としてとらえることができる」(宴のあと事件判決。東京地判昭和39年9月28日下民集15巻9号2317頁)

### 2. 個人情報保護法制による対応(入口規制)

(省略)

10

### 3. 消費者保護制度による対応

(1) 防衛的アルゴリズム(「自己抑制的選好アルゴリズム (self-restraint preference algorithms)」の使用促進(推奨)<sup>5</sup>

15

#### (2) セレンディピティ・アーキテクチャの選択可能性

→「消費者は、パーソナライズ化された経験と、パーソナライズ化されない経験の間で選択することができるべきである。それは、すなわち、過去の自らの選択から構築された『管理されたアーキテクチャ』と、平均的市民を反映しているか、あるいは他の選択原理・順位に従った『セレンディピティ・アーキテクチャ』の間での選択である」(Wagner & Eidenmüller, at 605)。

20

→「メタ選択」(meta-choice)の可能性が消費者に留保された仕組み

→問題は法で強制するか? (「個別化アルゴリズムによって消費者を標的化するプラットフォームその他のインターネット企業は、消費者が精選された選択 (curated choice) からオプトアウトすることを可能にするよう法によって要求されうる」。Wagner & Eidenmüller, at 606)。「軽度の規制 (light-touch regulation)」であり、許容される。At 609

25

#### (3) 取消権の新設

→「操作的なネット広告」というカテゴリ?

30

→「高度な情報技術を用いて、特定消費者のセンシティブな心理状況をプロファイリングしたうえで、当該消費者の意思形成に直接的な影響を及ぼすことを目的に、当該消費者が精神的に脆弱な状態にあるときを狙って、当該消費者に個別的に送られる広告」(山本龍彦「ビッグデータ社会における『自己決定』の変容」NBL1089号(2017年))。

---

<sup>5</sup> Michael S. Gal, *Algorithmic Challenges to Autonomous Choice*, 25 MICH TECH. L. REV. 59(2018).

→「囚われの聴衆」？（消費者契約法 4 条との類似性：操作的・攻撃的なネット広告、ターゲティング広告による消費者被害に対して消費者契約法の取消権で対応できないか？）

See Wagner & Eidenmüller：「消費者の選択が、ビッグデータおよび AI 技術によって影響を受けた場合、消費者に対して取消権を与えられることは正当であるように見える」（Id. at 597）。「消費者は、脆弱な瞬間（a moment of vulnerability）にアルゴリズムの影響下で行った取引については、これを取り消すことが可能であるべきである」。At 609）

※GDPRガイドライン<sup>6</sup>は、GDPR 22 条（完全自動化決定の原則禁止）について、「多くの典型的なケースにおいて、ターゲティング広告は、個人に重大な効果を与えるものではない」とし、原則として 22 条の適用対象外であるとするが、「事案の具体的性質によって、そのように考えられることはあり得る」とする。同ガイドラインは、ターゲティング広告が 22 条の適用を受けるかどうかを判断する要素として、①プロフィール過程の侵害性、②個人の期待や願望、③広告が送られる方法、④標的とされたデータ主体の脆弱性（particular vulnerabilities）を挙げている。そして、「一般には、個人に対してわずかな影響しか与えない処理が、社会における特定集団に重大な効果を実際上与えることがあり得る」とし、「例えば、財政的な困難に陥っている者で、オンラインのギャンブルの広告を規則的に見せられた者は、こうした申し出にサインし、さらなる負債を負う可能性がある」と述べている。22 条の運用状況については慎重な調査が必要であるが、脆弱な集団を狙った操作的なターゲティング広告については同条の適用を受け得るという趣旨のようにも解される。

20

### Ⅲ. その他の問題

#### 1. スコアリング

##### (1) 中国の状況

アリババ傘下の信用情報機関である「芝麻信用」が、決済アプリの購入履歴や SNS データ、資産保有情報などから、個人の信用力を 950 点満点で採点しており、この信用力スコアが官民を跨いだ様々な領域で使われている。このスコアが高い人は、低金利でローンを組めたり、賃貸物件の契約で敷金が不要になったり、様々な便益を享受できる（スマート化）。

他方、スコアが低い人は、あらゆる場面で差別的な扱いを受けることになる（+下層で固定される。「バーチャル・スラム」化の懸念）。

※1 2017 年末時点で 4150 万人。アリペイのユーザー数は 2018 年 9 月時点で約 7 億人。

※2 Tencent の騰訊征信（テンセントクレジット）も一定のシェア

政府は、2014 年以降、「社会信用システム構築計画綱要」（2014 年 6 月国务院により発

---

<sup>6</sup> Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/678, Adopted on October 2017.

- 表)に基づき、統一的な社会信用システムの構築(統一的な賞罰メカニズム)の実現を目指している(2018年1月、中央銀行の個人情報データベースの他に、「百行征信(信聯)」という個人情報情報機関が設立。芝麻信用を含む8社と中国インターネット金融協会が共同出資。民間の信用スコアはこの「百行征信」にデータ提供)。一部の地方政府では、賞罰的スコア制度を運用(江蘇省蘇州市、福建省福州市など:ボランティア、献血、労働模範への選出で加点→バス代が安くなるなど)。

## (2)ポジティブな側面

### ① 正当な評価

- 10 →これまで人間が行う評価よりも「正しい」評価につながる。  
→高度情報ネットワークシステムのおかげで、サイバー空間と現実空間を跨いで、シームレスに個人の行動を把握できるようになった。さらに、AIの進歩によって、データ分析の能力が飛躍的に上がる。今後はAI導入のコストもますます下がる。  
→このような情報環境の変化によって、個人を評価する側は、個人のありのままの姿を正確
- 15 に評価できるようになる(「点」ではなく、「線」で、しかも立体的に評価できる)。  
→このようなことから、プロファイリングや格付けは、お互いにとって利点があることがある。

### ② 包摂可能性

→これまで融資を受けられなかった者、金融サービスを享受できなかった者が、信用スコアを担保に包摂される。

### 20 ③ 取引の安全

→プラットフォームビジネスでは相手の顔が見えない。

### ④ 社会の安全(規範の内面化)

## (3)ネガティブな側面

- 25 ①過去の差別の再生産(学習データに含まれるバイアス→AIによるバイアス承継)  
②超監視社会の出現(AI社会=More Data社会。自律的行動に対する制約)  
③バーチャル・スラム(スコアリングのプロセスが「ブラックボックス化」しうるため、再挑戦の機会が失われ、評価の低い者たちが固定化し、社会的に排除される可能性がある。  
Gated communities without gates, or Gated communities with “invisible” gates)
- 30 →ロー・スコアな者、ノー・スコアの者たちにとっては生きづらい?

## (4)消費者保護との関係

- スコアによって取引から排除されるリスク(ロースコアな者はプラットフォーム取引から排除される)
- 35 →スコアの適正さが問題となる。

## 2. 第一種価格差別(first-degree price discrimination)

→「足元」を見られる(市場価格で買えなくなる)。

※ネットフェリックスは、第一種価格差別を適用することで、自らの利益を12.18%上げ、それにより消費者余剰の合計は7.75%落ちたとされる。

- 5 →「第一種価格差別は、企業に利益を与え、消費者を害する」(Wagner & Eidenmüller, at 588)。

## 3. 消費者の福祉(倫理問題?)

- 10
- ・選択機会の限定(ハイキングのガイドブックの購入履歴が、エベレスト登頂につながる例)
  - ・子どもの場合(アルゴリズムが関与していない「世界」を知らない。選択がそれしかない、と感じてしまうリスク)