

# 「ビッグデータ」と「民主主義」

**戦**前、米大統領選挙の事前予測でLiterary Digest社の大量データよりもGallup社の統計的標本調査のほうの信頼性が高いことが確認され、「1人1票の民主主義」を擬する「科学的世論調査」が確立されてきた。ここで「科学的」とは統計理論により精度が計算できることである。

しかし、2016年の英国EU離脱国民投票や米大統領選挙の事前予測調査の失敗は、世論調査の信頼性を大きく損なわせた。面接調査から電話調査、WEB調査などへの表面上の技術革新が、実はコスト削減が主目的で、統計調査の質の低下へ確実につながっている。その後、先述の国民投票や選挙でCambridge Analytica社が人々のFacebook関連情報などを利用して世論操作ともいべき「選挙活動」を行ったことが浮かび上がってきた。ロシアとの関係にも疑惑がもたれ、両社に捜査の手が及び、米国議会で喚問も行われた。ビッグデータが民主主義を壊そうとしているかの様相である。

各人のウェブの閲覧や購買の履歴情報から、その人の政治傾向、宗教、性的嗜好等まで特定され得るビッグデータの時代となっている。携帯電話番号は地域特定ができず選挙では携帯電話で標本調査による予測はできないとされているが、GPS情報と絡めると、各個人の居住地は勿論、勤務場所、勤務体系、通勤中の寄り道場所すら推

察するのは容易となる。匿名加工されたデータの各々は個人情報保護法に違反していなくとも、それらを融合して個人情報を特定することは法律で禁じられている。ただし、法律で禁じていても現実に問題が起り広く報道されれば、調査関係者一般の仕事に大きな支障が生じるであろう。

ソーシャル・ビッグデータの利用における個人情報保護に関しては、情報技術の急速な進展に法律が追いつかないうえに、国によって考え方や運営の仕方がかなり異なるなど、課題が山積している。このような状況ではあるが、実証的データに基づく学術の発展や政策立案のためにも、まずは個人情報の収集や公開において遵守すべき国内外の法律や倫理は確認しておくことが肝要である。筆者の所属するセンターでは昨年3月に、専門弁護士による特別講義（要旨は日本分類学会学会誌『データ分析の理論と応用』に掲載予定）とW. Jagodzinski氏によるEUのGDPR（一般データ保護規則）に関する特別講演（本誌に解説）を公開シンポジウムとして行ったが、これらが参考になろう。

経済振興のために情報技術が期待されているが、世論調査の世界では民主主義と個人の権利を守る活用の仕方が問われている。さて、民主主義を守る「科学的な世論調査」の再建のために、ビッグデータを如何に活用すべきであろうか。

吉野諒三

情報・システム研究機構 データサイエンス共同利用基盤施設・社会データ構造化センター センター長  
(統計数理研究所 教授 兼任)