

# Twitter アカウントを凍結されて 気付いたジレンマ

～守るべき表現の自由と不正対策との狭間で～



楠 正憲 | 国際大学 Glocom

筆者が2007年4月から12年以上も利用しているTwitterアカウントが2019年6月1日夕刻、突然凍結された。特にメールなどによる連絡はなく、外出中にTwitterアプリのフォロー数、フォロワー数がゼロとなってタイムラインが更新されず投稿できない状況が続いた。Yahoo! JAPANの提供するTwitter検索で自分のアカウントが凍結されていると騒ぎになっていることを知って慌ててPCのブラウザからTwitterにアクセスしたところ、アカウントが凍結されている旨について表示されて、異議申し立てを行うWebフォームに誘導された。ここで特にアカウントを凍結される心当たりがない旨を記入して送信したところ、自動応答で異議申し立てを受け付けた旨についてメールがきた。翌6月2日午前、筆者のアカウントが凍結されたことを心配していた友人から、アカウントが復活しているのではないかとFacebookに書き込みがあり、確かにタイムラインが再び更新されるようになり、0となっていたフォロワー数も数時間で回復した。Twitterから英文で短いメールがあり、スパムアカウントの判定アルゴリズムが誤動作して、誤ってアカウントを凍結してしまった旨について連絡があった。手短なメールはスパム判定アルゴリズムの誤動作で、利用者からの通報に基づく凍結ではない旨だけ簡潔に記載されていた。機械的な英文メールのみで日本語による説明はなかった。

この時期Twitterによるアカウント凍結が相次いだ。6月初頭には天安門事件が発生した日に合わせて中国関連をはじめとして数多くのアカウントが凍結された。

6月18日には新潟県・山形県で震度6強の地震が起きた前後に、これまで地震の情報を発信していたアカウントが凍結された。凍結はすぐに解除されたものの、その後も俗にシャドウバンと呼ばれる検索に引っかかりにくい状態が続いたようだ。

Twitterは民間の無償サービスで、アカウント開設時に同意した規約で、違反する行為があった場合にはアカウントが凍結される旨について同意していることは認識していた。しかしながら実際に自分のアカウントが凍結されてみて、どの書き込みが規約に違反したのか、どういった理由かについて何ら説明なく、アカウントの凍結は一方的に行われて、何ら説明なく唐突に解除されるものであることを実感できた。アカウントを凍結されてもログイン自体はできるので、Twitterアカウントと連携させているサービスを使い続けることはできる。しかしながら過去のツイートやダイレクトメッセージを見ることはできなくなり、これまでTwitterで連絡を取っていた人たちとの連絡手段は途絶してしまう。筆者が12年かけてフォローしてきた人々や、筆者をフォローしてくれてきた人々のリストにはアクセスできなくなってしまった。電子メールの場合、メーラーを使っていればメールアカウントが剥奪されたとしても、過去のやりとりや連絡先は手元に残すことができるのと対照的だ。Webメールなどクラウド上のサービスに依存している場合には、Twitterと同様にアカウントを凍結されることで、すべてを根こそぎ奪われることも考えられる。

データポータビリティを求める主として欧州当局の要

望に応じて、Twitter は保有しているデータのダウンロード提供を行っているが、データさえ手元に置いていれば解決する問題ではない。ダウンロードしたデータを別サイトで再び公開したところで、これまでの書き込みを参照していたリンクはすべて切れてしまうし、新たに発信した情報をこれまでのフォロワーに読んでもらえるとは限らないからだ。これまでデータに対する利用者の権利としては、プライバシーやデータの保護、データを手元に置いて好きな場所にアップロードできるポータビリティまでだったが、アカウント凍結によって侵害される権利は伝統的なプライバシー権だけでなく、過去から現在にかけての自分の蓄積や周囲との関係性、サイバー空間における人格そのものであって、それはデータポータビリティだけでは解決しないのである。

Web サービスを運営していく上で、スパムアカウントに対処しなければならぬ必要性は理解できる。昨今は有害な情報の書き込みやフェイクニュース、児童誘引などの対策も求められており、規約違反を監視して適切に凍結することが責任ある事業者として求められていることは承知している。とはいえ Twitter 社自身が理由を説明できないかたちで勝手にアカウントを凍結し、十分な説明なく凍結解除されたことは想定外だった。また自分のアカウントが凍結された理由を探った中で、Twitter の規約に対する考え方は時代を追って変遷しているが、過去に遡って NG ワードが含まれる投稿を見つけて、組織的に通報するなどの方法によって人為的にアカウントの凍結を行う手口や、米国 DMCA (デジタルミレニアム著作権法) を悪用して書き込みを削除させる方法があることも知った。私の友人でも非常に多くのフォロワーを抱えて有益なツイートを行っているにもかかわらず、アカウントを凍結されたまま解除されない方を幾人も知っている。Twitter 上でさまざまな表現や意見表明が行われている中でアカウントの凍結にかかわるルールが不透明で、その手続きを悪用されかねないことは、嫌がらせだけでなく言論統制にも容易に悪用できてしまう。

問題ある情報発信に対しては日本でもプロバイダ責

任制限法によって IP アドレスなどの発信者情報の開示や削除依頼の仕組みが用意されているが、プラットフォーム事業者による恣意的なアカウント・書き込みに対する削除や、虚偽の通報に基づく不当な取り扱いからユーザを守る仕組みは整備されていない。欧州の GDPR (EU 一般データ保護規則) では、個人データの自動処理によって不利益ある取り扱いをされた場合に、自然人による対応を求める権利を定めたが、日本や米国では同様の権利は今のところ認められていない。

世界中で AI 利用の倫理について議論されており、透明性をはじめとした原則が唱えられているが、不正対策の自動化でジレンマとなるのは、広く仕組みを公開してしまうと、悪意ある利用者が自動処理を回避する手がかりを与えてしまうことだ。とはいえ自動処理によって利用者に不利益ある措置が行われた場合に、異議申し立てが適切に処理されて、当事者間で同意に至らない場合には、司法なりの枠組みを通じて紛争解決するための仕組みは必要となるのではないか。

これまでサイバー空間において犯罪や違法行為、著作権侵害に対する対処として、データやアカウントの削除のための仕組みは整備されてきたが、サイバー空間における表現や対人関係の比重が高まる中で、表現の自由や通信の秘密といった憲法上の権利を保護するためには、悪意ある利用者を追い出す仕組みだけでなく、不当な通報や自動処理によって不利益を被った場合に、適切に異議申し立てできる仕組みを整備する必要があるのではないか。存在感のあるプラットフォーム事業者の多くが米中などの外国企業である実情も踏まえ、日本にいる利用者の権利を実質的に担保するためには、国際的な制度整備へ向けた多国間の連携や、日本における立法も含めて措置が必要と考えられる。

(2019年6月30日受付)

■楠 正憲 (正会員) masanork@gmail.com

マイクロソフト、ヤフーなどを経て 2017 年から Japan Digital Design CTO。内閣官房政府 CIO 補佐官としてマイナンバー制度を支える情報システム等の構築に従事。