

第 3 回 GLOCOM 研究交流会 開催レポート

テーマ：『優れた知の生成に関する希望と困難』

報告者：渡辺智暁（国際大学 GLOCOM 教授・研究部長・主幹研究員）

コメンテーター：八田真行（国際大学 GLOCOM 客員研究員／駿河台大学経済経営学部 准教授）

司会：山口真一（国際大学 GLOCOM 講師・主任研究員）

日時：2019年12月16日（月）15:00～17:00

場所：国際大学グローバル・コミュニケーション・センター



概要

GLOCOM 研究交流会は、研究員と参加者が、カジュアルに研究について意見交換・交流をすることを目的とし開催している会です。第3回となる本会では、国際大学 GLOCOM 教授/研究部長/主幹研究員の渡辺智暁より、「優れた知の生成に関する希望と困難：データ駆動型 AI&ビッグデータ、エビデンスベースの科学的知、Wikipedia の「集合知」、ネットの自由な言論」というテーマで 45 分間の話題提供を行い、コメンテーターに八田真行氏（駿河台大学経済経営学部准教授）をお迎えしてディスカッションを行った後、1 時間のフリーディスカッションを行いました。講演では、Wikipedia が専門家が執筆する百科事典と遜色が無い品質であることや現在の科学の手法に弱点があること、AI やビッグデータによる新しい科学がどのような方向へ向かっていくのかについて紹介され、ディスカッションも多様な質問が参加者からなされました。

◆講演「優れた知の生成に関する希望と困難 : データ駆動型 AI&ビッグデータ、エビデンスベースの科学的知、Wikipedia の「集合知」、ネットの自由な言論」



Wikipedia とネット言論

最初に、Wikipedia とネット言論について考えるが、まずは「Wikipedia は案外正しい」というところから話を始めたい。Wikipedia の執筆者の層は厚くページの分量も多く、英語版 Wikipedia の品質検査に関する論文は何本もあるが、専門家が執筆する百科事典と遜色ないというものが多い。

Wikipedia はネット言説とは異なり、自浄作用がある。投稿される内容は相対的に見ると、問題の量は少ない。Wikipedia にわざわざ投稿しようとする人は、アクティブであればあるほど、Wikipedia の方針・目的に概ね共感している人が多く、入り口のところでフィルタがかかっていると言える。Wikipedia がネットと言論とだいぶ異なるのは、それが「集約するメディアである」ということだ。一つのテーマに対して一つのページに一貫性を持たせてまとめるので、合意形成のプロセスが大変ということもある。

Wikipedia を編集する過程で争いが起きやすいが、Wikipedia にはこうした争いを解決するための方針がいくつかある。そのうちの一つとして、情報源に関する方針がある。「世間的に信頼できる情報源になっているものの情報を出典付きで記載する」というのが今の Wikipedia のルールだ。究極的には、信頼されている情報源に掲載されている情報であれば、それが真実か否かは問わないというのが Wikipedia の方針ということもできる。この他に、「Wikipedia は紹介や解説をする場であって主張する場ではない」というルールもある。こうしたルールがあっても完璧には程遠く、ぶれやバイアスは少なからずかかるにしても、実態としてはそれなりにはうまくいっている。

ただし、これを集合知の典型と呼ぶかは疑問だ。ネットの言論でも真似出来ないところがある。また、人種や性など、その時代の主流派の説は偏見であっても記述するのが Wikipedia だ。病気の原因やメカニズムが誤ったものであっても、主流派であれば大きく取り上げている。それを正す力は Wikipedia には内在されておらず、外部の文献に任されている。信頼できる情報源よりも、参加者の判断を下に置いている。真実の追求はそもそも放棄されているということも言える。

Wikipedia は集合知ではない

ウィキペディアがやっていることは外部後を集約するのみであることから、本来の意味と比較すると、集合知として限定的といえる。そこからネット言論一般を考えると、取捨選択の指針が Wikipedia のように共有されているわけではないので、集約をネット言論に当てはめるのは難しい。Wikipedia は一次資料を援用して自分で独自の解釈を加えて論評することは認められていないが、ネット言論でやらないのは非生産的だ。言論のためのプラットフォームのアナロジーで考えると、参加者による自治はソフトウェアという技術的なものとルールという制度的なものの両方があり、それらのおかげでうまくいっていることがあることから、学ぶことは多いかもしれない。

Wikipedia の強みと現行の科学の弱点

学術的な百科事典と Wikipedia を比較すると、正確さにおいては実は Wikipedia の方が優れているという調査結果が出ることもある。「専門家が書けば誤りは含まれない」ということがないのは周知の通りだ。Wikipedia は百科事典と比肩する品質に英語版ではなっているがそれはなぜなのかというと、Wikipedia を荒らしたい人よりも貢献したい人のが多い、ルールが決まっている、問題のある投稿を発見・対処するための仕組みがかなりサイトに埋め込まれているということが挙げられる。Wikipedia の仕組みはボランティアで作られており、ウェブサイトで攻撃の報告を受けるとツールの開発もされるため、普通の Wiki に比べてそうした部分がすごく充実していると言える。その結果、荒らしたい側と対処側のパワーバランスが対処側に有利になるように保たれていることが大きい。

学術論文の世界で、対応する部分を考えてと査読がある。査読の弱点としては、捏造、隠ぺいに弱いことの他に、意外なパターンの発見や強い効果は評価されやすく、「あると思ったパターンが見つからなかった」報告は評価が低い「出版バイアス」と呼ばれるものがある。強い効果のかなりの部分は、他の目的の研究をしている際に、偶然見つかることがあるが、本来の目的とはずれていることもあり、小規模データを元に報告され、大規模データでは再現できないことも多い。査読制度が機能してもこれらの弱点が出るのは問題だ。また、心理学の世界で、心理学上の知見がどれだけ再現できるのかを再検証する研究を実施したところ、40%しか再現できなかった。他にも、がんに関する生物学上の知見を再検証すると 10%しか再現できなかったという報告もある。Wikipedia よりもはるかにひどい惨状とも言える。

科学の自浄作用に期待

では科学は絶望的かということ、こういうものを自ら見つけて、それに対する対策を考えて、少し洗練された形でまた別のプロセスをデザインしていくということが科学の中では起きているので、そこまで絶望的ではないかと考えている。隠ぺいを防止するために臨床試験する前に登録にしておき、透明性を挙げるようにしている。自浄作用はあるが、現状はそこまで優れたものとは言い切れない。これらのことから比較してみると、「Wikipedia は査読を受けていないから信用できない」という理論は、査読を過剰評価しすぎているといえるかもしれない。ただ、学術研究というものは、そもそも信頼できる情報源が報じていない領域について探索し、パターンを見つけていくような営みであり、Wikipedia より遥かに合意形成が困難な知的活動に取り組んでいるといえる。ただし、制度としてみると査読があれば万全というわけではないだろう。

これまでの科学とデータ駆動型の知

科学と AI やビッグデータなどのデータ駆動型の知を比較する。データ駆動型の知についてよく言われるのは「データの上でのパターンはあるものの理論がない」ということだ。例えば、「おむつと缶ビールを近い売り場に置くと売れ行きが良い」ということが発見されたときは、当初は何の理論もなかったが、のちに「仕事帰りにおむつを買ってくることを頼まれた人が来店ついでにビールを買う」という背景があることがわかっている。膨大なデータ解析で得られる知は人間にも理解ができないし、細部の判断のアルゴリズムの部品は不可解・非合理的な判定方法を使っていて、そもそも信頼できる知ではないのではないかということが言われている。

しかし現状、医学や薬学、経済学、エビデンスベーストの政策の分野で重要視されている「ランダム化比較試験」はたくさん問題がある。例えば、試験にかかるコストが高額で長期間なので再現が極めて困難だということだ。また、因果関係についてはわかるが、その背景となる理論が必ずしもあるわけではない。しかし、現状ではある種のゴールドスタンダードとして機能している。ある特定の症状に効く薬を開発している人からすると、試験の際には他の病気を持っている人を排除する必要があるが、そうすればするほど、実際の患者は排除される側になっている可能性が高い。患者が必ずしも特定の症状のみを罹患しているわけではないからだ。この実験方法では、本当に一部の患者にしか効くことが確かめられていないという批判もある。その他にも、大型資金が必要となることもあり、統計的に見れば製薬会社に不都合な結果が隠蔽されているらしいという事も言われている。

そう考えてみると、昨今の深層学習で生まれてくるようなアルゴリズムも、そこまで科学と比べて見劣りしないのではないかと考えるが、深層学習に携わっている人の論文を読むと、「自分たちは既存の科学と同等だからこれでいい」という態度でいるわけではなく、「データが偏っているからもう少し違うデータを使って検証する必要がある」、「人間が理解で

きない部分があるから、透明性を高める必要がある」という従来型の科学に従事していた人が指摘するようなことを自分たちでもしている。彼らは平凡な科学の一部になろうとしているし、そうなっていくと予想する。一部の過激な科学不要論はまた極端なものなのだろうと考える。

◆コメント



Eternal September 問題と、コミュニティ維持の条件

何かを生み出すポジティブなコミュニティを維持すること、成長することは難しい。「終わらない9月」と訳される Eternal September 問題と言われるものがある。掲示板のようなサービスの Usenet は当初大学関係者ののみが使っていた。新学期である9月になると新入生がこれに参加し、古株となるユーザーからネチケットや文化を学ぶ。

しかし、1993年9月に AOL が Usenet へのアクセスを一般に提供を開始したところ、参加者が桁違いに増えてさばききれなくなってしまった。この問題について、Reddit を対象に研究している人がいる (Kiene, Monroy-Hernandez & Mako Hill(2016))。彼らによれば、コミュニティ維持の条件として、①活発でよくコーディネートされた管理者グループ、②コミュニティ・モデレーションを円滑にする、共有されたコミュニティ感覚、③規範の侵害を緩和する技術的システムの3つが挙げられる。また別の研究 (Fiesler et al.(2017)) では、①モデレータに強力な権限、②明確なガイドラインと規範、③投票評価システムを条件としてあげている。また、行動規範となるガイドラインである「Code of Conduct」の是非がオープンソ

ース開発プロジェクトなどで10年ほど議論されている。賛成派はコミュニティの維持になるといい、反対派は自由が損なわれるという。個人的な経験則では無いとフレームワークが発生するので、あった方が良く考えている。

また、「Toxicな人問題」もある。Toxicな人とは「面倒な人」という意味で、これは有害だ。彼らの特徴として、ろくにコードも書かずに口だけ達者であら捜しが得意、なぜか一日に何通もメールを書くほど膨大なエネルギーがあることが挙げられる。そんなに多くは無いもののこうした人がいるとプロジェクトが大変なことになる。

足切りの有効性

「駄目な人は切った方が良く」ということが言える。「だめな言説」は機械的に切ることができないが、Wikipediaには「要出典」など明確なNGがあり、これは編集の過程で切れる。オープンソースプロジェクトにおいては、ダメなコードは動かないので、これが足切りとして作用する。コミュニティを不特定多数に開くとToxicな人やそれに引きずられる人が多数出てくるので、面倒なことになるが、足切りをすれば大体の問題はなくなるのではないかと考える。科学研究も形式的なチェックで足切りをすればいいのではないかと考えている。いくつかのポイントに関して形式的なチェックをするだけで品質が上がるのではないかと考えている。

また、ジョージメイソン大学の法学者であるジェイソン・ブレナンはエピストクラシーについて話している。エピストクラシーは「知識がある人による支配」という意味だ。「三権分立の三権とは何か」など簡単なクイズをして正解しないと投票権が得られないようにした方が良くのではないかと考える。そうした時に、ハードルをどこに設定するのが難しい。ピーター・ティールは「インテリジェント・デザイン」という言葉を使い「頭がいい人がルールを設定すれば良い」と主張しているが、これは否定できないように思う。

◆ディスカッション

山口： Wikipediaで、信頼できる情報源からのものは記載するというものがあつたが、どういう基準で「信頼できる情報源」と判定しているのか。

渡辺： 学術論文や学術書など専門家が品質のチェックをしているようなメディアが中心で、その他には一部の報道メディアが挙げられる。その他の情報源は格下だ。企業のプレスリリース等は場合による。

山口： ネットの言論のプラットフォームにWikipediaのような自治機能を入れたほうがよりよい議論空間になるのではないかとされているが、その一方でいつまでたってもそうならないのは一部の人間しか求めていないのではないかと思うがどうか。

渡辺： そう思う。Wikipediaは始めた人たちも当初はこんなに大きくなると思っていなかった。実験してみる価値ぐらいはあるだろうと思われる。どれくらいドミナントなメディアになるかはわからない。自治が効いているメディアにどれだけ注目が集まるかはわからない

い

渡辺： Reddit はあまり良くない言論の場として言われることが多いが、コミュニティが上手く維持できる仕組みさえあればまっとうなネット言論が生まれると思うのは甘く、ニーズはもっと偏ったものにあると思っておいた方が良いのか。

八田： ネット言論と政策決定を切り離せばいいと思う。ネット言論はガス抜きで、政策決定は粛々とエリートが進めればいい。圧倒的多数の人はよくわかっていない。オープンソースは本当にダメな人が入ってこられなかったからうまくいった。社会にそれを拡大すると、ダメな人が入ってくる。全くランダムに投票するとなると、Yes が 50%、No が 50% となるはずだが、少しでも分かっている人が含まれることで 51% と 49% となるようになり、昔はそうになっていた。しかし、最近は偏っている人が増えてきて変な選択がなされるようになった。最近のイギリスがいい例だ。自由投票のようなものは何らかのシステマ的な仕掛けが入っていないと意見が分かれるようになりうまくいかないというのが意見だ。

——ジャーナリズムも資格制度を導入しようという議論がある。日本では言論の自由があることから難しいと思われるが、頭のいい人のみがやっても支持されるのか。

渡辺： 報道機関の免許制を導入するということを考えた時に、トランプ支持者のことを考えると、「主流派メディアは虚偽や誤報に満ちているから、それに汚染されていないオルタナティブなメディアを作る」という人たちが、免許を持っていなくても依然として活躍し、トランプを支持するという仕組みがある。免許制を導入したところで、その壁が解消するとは思えないし、やるなら選挙までいかないといけないと思う。

八田： 結局選挙権、人権を改正するしかないのではないか。意思決定における legitimacy (正当性) が重要だと思う。民主主義に legitimacy があるか怪しいと思っている。裁判をやり、有罪が確定し刑に服するのは、裁判官や陪審員が判断したという歴史的な legitimacy があるからだが、普通の選挙や意思決定が行われるときには、それが legitimacy がある意思決定なのかが怪しい。独立した個人が合理的な意思決定ができるというのが legitimacy の根拠だったが、現在においては怪しいと感じている。敗戦後、これまでは民主主義が良いものだと思われてきたが、そこに本当に根拠があるのか怪しいというのが最近噴出している議論だ。

——2016 年の米大統領選挙の際に、Facebook などのプラットフォームなどは積極的に正しい情報の流通に貢献しろということになった。この流れの中で登場したのが Facebook News だと思う。一般の人の「いいね！」や共有に任せていると偏った情報ばかりが流通してしまうので、正しい情報を流通するためには、ある程度権威のあるメディアが作ったニュースの中からプラットフォームがキュレーションするような方向に向かっているように思えるがどう思うか。

渡辺： 一部のメディアに限る必要は無いと思う。Wikipedia は学位を執筆者に問うてない

が、ある程度の別のクォリティの要請をしている。「一定以上の発行部数がある」や、「報道関係の団体メンバーであり、特定の規範を守っていることを証明できる」という人は軽い審査で参加でき、そうでなくても、多くの人が「この人の言っていることは信頼できる」と言っていれば彼らも参加できる仕組みがあっても良いのではないか。Facebook がホワイトリスト作りをすると相当大変なことになり、名前が売れていて無難なところだけリスト入りということになりそうで、それはあまりにももったいない。

八田： ファクトチェックについては懐疑的だ。ファンタジーが見たい人は正しい情報を見たいわけではない。ファクトチェックをしようがしまいが、今でも正しい情報を見ようと思えば見えるが、彼らは見たくない。ファクトチェックをして正しい情報を定義づけて押し付けても、考えは変わらないだろう。ファクトチェックでなんとかなるというのは楽観的なのではないか。その代わりとして、ファンタジーが見たい人はファンタジーを見れば良いと思う。むしろやるべきなのはプロトコルで分割することだ。フェイスブックが問題なのは、ユーザーが見るコンテンツをフェイスブックが決めているということで、ユーザーが見たいものをプラットフォームが見せるべきだ。Facebook は独自のプロトコルを作成して、オープンな形で提供する。ユーザーが Facebook クライアントソフトでそれぞれフィルタを適用すればいい。ハブに力が集中しているので、エンド側に決める力を取り戻すべきで、具体的にはオープンなプロトコルでアクセスできるように強制すべきだ。そうすると、フェイスブック謹製のフィルタを使うことでファクトチェック済の言論が多く出るだろう。

——言論はそもそも思想や見解の発表なので違っていることが前提であり、オープンソースソフトウェアや Wikipedia と同じように議論はできないと感じた。2016 年の米大統領選でフェイクニュースが問題になった。日本において影響力が大きいのはテレビなので、テレビのコメンテーターの質をなんとかしないと駄目だろうと考える。

渡辺： 「ネットプロパガンダ」という本では、2007 年のリーマンショックが所得格差や社会の分裂を起こしたのが、言論の世界にも出てきたと紹介しており、根本的な原因は社会的な分断の方にあるという研究を出している。メディアの力がそこまで軽くはないとも思っているが、日本でも格差の問題は無視できないものでもある。そこまで考えると、隔離をするよりは格差をなくす政策を含めてやったほうが良くなっていくし、もっと根深いトラブルを避けられると思う。

八田： フリンジ(ひどい右翼系の言論)をメインストリームのメディアが紹介することで、マイナーなものが市民権を得たという話がある。ニュースの見せ方やインターネットとの親和性という点で残念ながらプロパガンダメディアの方がうまい。最近はストリーミングメディアの方がミームやお笑いのテクニクを取り入れておもしろいニュース番組を作っている。今の旧来のメディアは影響力を失っている。新興メディアが、ニュースの見せ方の点でここ 5~10 年いろいろなイノベーションを起こした。それをいち早く取り入れたのがキュレーションメディアだ。そういった点で旧来のテレビや新聞が追いかけておらず、「い

い記事を書けばなんとかなる」といっているのは幻想だと思う。

渡辺： Wikipedia は事実が何かを集約していくところなのでやりやすい。心理的な次元が入るところは格段に難しく、wiki でできるかはだいぶ疑問。事実関係を明らかにするだけでも意義はあるだろうと思う。論争になった時に「信頼されているメディアでこういう調査結果が出ている」というのは、ネットに出ているなら参照されやすいので、そういうところに信頼できる情報が沢山あるほどネットの言論も自浄作用が働きやすいところもあるだろうと思う。そこにはまだやれるところはあるだろう。

執筆：永井公成（国際大学 GLOCOM リサーチアシスタント）