

特集 人工知能学会 25 周年特集「四半世紀を越えて」〔第 3 部〕学会の在り方を考える

学術分野の Web サービスと成果評価の現状

神 嶋 敏 弘 産業技術総合研究所ヒューマンテクノロジー研究部門

本稿では、学術分野における Web 上のサービスの現状と、研究成果の評価をめぐる事例と議論について報告する。

1. Web 上の学術サービス

1.1 文献の電子アーカイブ

論文などの出版物は学術活動において重要であるので、Web 上の学術文献の出版サイトを最初に紹介する。国際的な学会としては、コンピュータ科学全般の双壁である ACM と IEEE があり、ほかには人工知能の AAI, 言語処理の ACL, 応用数学の SIAM などがある。これらの学会には出版部門があり、その出版物は Web 上のアーカイブより入手できる。

ACM Digital Library <http://portal.acm.org/>

IEEE Explore <http://ieeexplore.ieee.org/>

IEEE Computer Society Digital Library

<http://www.computer.org/portal/web/csd1>

AAAI Digital Library

<http://www.aaai.org/Library/library.php>

ACL Anthology

<http://www.aclweb.org/anthology-new/>

SIAM <http://epubs.siam.org/>

ACM のサイトは、自学会の雑誌・論文誌・国際会議録のアーカイブに加え、IEEE などの関連学会や学術系の出版社の文献などの書誌情報をも管理している。この書誌情報は、アーカイブとしての役割だけではなく、ここに登録されていること自体が、論文誌や国際会議録のレベルの保証とも解釈されるようになってきている。IEEE は、IEEE の全 Society の文献を配布する Explore と、その一部である Computer Society のみの文献を入手できる Computer Society Digital Library とがある。これらのサイトからは会員資格の取得や購読料の支払いにより文献を入手できる。

国内の学会関連の文献も Web 上で配布されている。

J-Stage <http://www.jstage.jst.go.jp/browse/>

J-Stage Information and Media Technologies

<http://www.jstage.jst.go.jp/browse/imt/>

情報処理学会 <http://www.bookpark.ne.jp/ipsj/>

電子情報通信学会 <http://www.ieice.org/jpn/books/>

J-Stage は、科学技術振興機構が運営する、国内学会の電子出版サイトであり、人工知能学会論文誌も含めた多くの学会誌・論文誌が配布されている。ここには、コ

ンピュータ科学関係で英語で執筆された文献を集めて発信する *Information and Media Technologies* があり、人工知能学会論文誌の中で英語で執筆されたものは収録されている。国内のコンピュータ科学分野での大きな学会である情報処理学会や電子情報通信学会も独自の配布サイトを運営している。

学術文献は一般の出版社からも出版されている。

SpringerLink <http://www.springerlink.com/>

Elsevier <http://www.sciencedirect.com/>

合併・買収により Springer 社と Elsevier 社に多くの学術出版物が集約されてきている。Springer 社の SpringerLink は、多くの国際会議録などを含む *Lecture Notes in Artificial Intelligence/Computer Science* を有している。また、Kluwer Academic Publishers 社との合併により、機械学習の有力誌 *Machine Learning* など多くの論文誌をも出版している。Elsevier 社の Science Direct も、Academic Press, Morgan Kaufmann, Pergamon Press などの合併・吸収により多くの有力誌を有している。*Cell* や *Lancet* など各学術分野で非常に著名な雑誌や、人工知能関係の *AI Journal* などここから配布されている。

その他の動向も紹介しておこう。物理学などの分野では、論文を投稿し査読を受ける前に、プレプリントとして公開することが行われ、その公開には arXiv.org (<http://arxiv.org/>) が利用されている。プレプリント以外にも、テクニカルレポートや、文献の追加情報などの公開にも利用されるようになってきている。Project Euclid (<http://projecteuclid.org/>) は統計・応用数理の学会が、効率と利便性の向上を図るため共同で運用している文献アーカイブである。

電子出版を中心とする論文誌も現れている。

Journal of AI Research <http://www.jair.org/>

Journal of Machine Learning Research

<http://www.jmlr.org/>

IEICE ELEX <http://www.elex.ieice.org/>

国際的には *Journal of AI Research* や、*Journal of Machine Learning Research* はそれぞれ人工知能や機械学習の分野でトップジャーナルとして認められている。国内でも、人工知能学会論文誌は電子出版であり、電子情報通信学会も電子出版専門の論文誌 ELEX を出版している。国際会議でも、ACM/IEEE 主催ではない人工知能関連の IJCAI, 機械学習関連の ICML, NIPS, UAI などでは会議録が全文公開されるようになってきている。

1.2 文献の書誌情報サービス

前述のようにACM Digital Libraryは書誌情報サービスをも兼ねているが、書誌情報専門のサイトもある。

DBLP <http://www.informatik.uni-trier.de/~ley/db/>

CiNii <http://ci.nii.ac.jp/>

編集者が管理する書誌情報サービスとして著名なものにDBLPである。DBLPがDataBase systems and Logic Programmingであったように、初期の頃はデータベース関連の書誌情報を収集していたが、現在ではコンピュータ科学分野全般の書誌情報をまとめてDigital Bibliography & Library Projectとなっている。ACMのライブラリの広い範囲をカバーしており、また、ACMと同様に収録されていることが文献・会議のレベルの基準としても使われることもある。しかし、FAQによれば、特に審査などはなく、定められたフォーマットで書誌情報を送信するといった、一定水準以上の運営が行われていれば採録されるとのことである。国内の書誌情報サイトとしては、国立情報学研究所が運営するCiNiiがあり、書誌情報のアーカイブだけではなく、その情報を下記のようなほかの書誌情報サイトに配信するといったことも行っている。

上記のサイトは人手での管理だが、Webクローラーにより自動収集するタイプの書誌情報サービスもある。

CiteSeerX <http://citeseerx.ist.psu.edu/>

Google Scholar <http://scholar.google.co.jp/>
Microsoft Academic Search

<http://academic.research.microsoft.com/>

CiteSeerXは、自動収集型の書誌情報サービスの先駆けであったCiteSeerの後継サイトであり、引用文献の類似性に基づく類似文書検索などの特徴がある。Google Scholarは、現在、最も多くの書誌情報を集積していると思われる。Microsoft Academic Searchは、著者間の共著関係の可視化などの特徴がある。

以上のサイトは、サービス側が書誌情報をまとめて提供するタイプであった。それに対し、利用者が個々に自身の関心のある文献情報を登録し、それらを共有できるサイトも現れている。

CiteULike <http://www.citeulike.org/>

Mendeley <http://www.mendeley.com/>

CiteULikeはそうしたサイトの中で現在、最も有力と思われる。論文での引用に使えるようbibtexなどの形式で書誌情報を出力できたり、タグ付けなどにより文献を分類できたり、またそれらの情報をグループなどで共有できたりといった機能を提供している。Mendeleyは電子化された文献を保持しておくためのレポジトリを提供するサイトである。利用者は、文献本体のファイルは自分が置いたものしかアクセスできないが、書誌情報などは共有することができる。

1.3 その他の学術関連サービス

文献だけでなく、口頭発表・講演をビデオ撮影し公開するサイトも登場している。

videolectures.net <http://videolectures.net/>

TechTalks.TV <http://techtalks.tv/>

videolectures.netはそうしたサイトの代表で、ビデオだけでなく、ビデオと同期して講演資料も表示されるなど、学術発表用に向けた工夫がなされている。また、発表に対する評価付け、閲覧ランキング、閲覧者同士でメッセージを交換できる機能などを備えている。評価付けについては、私見ではあるが、著名な研究者でよく閲覧されるビデオについては機能しているが、それ以外では、自身の発表に五つ星を付けるといった感じになっており、それほど機能しているようには見えない。発表・講演の公開にあたっては、発表者の許諾を個別に得ており、同じ会議であっても発表者が許可しなかったものについては閲覧できない。TechTalks.TVは、videolectures.netほどの機能はなく講演ビデオが閲覧できるだけである。しかし、videolectures.netは本拠はヨーロッパにあるため、アメリカで行われる会議では、こちらで講演ビデオがアーカイブされることがあるようだ。

国際会議の開催情報は各種のメーリングリストを通じて長らく配布されてきたが、その状況はWikiCFP(<http://www.wikicfp.com/>)の登場により大きく変わった。WikiCFPは、国際会議の開催場所、開催日時、締切といった情報を利用者が登録・閲覧できるWikiサイトである。利用者はアカウントをつくと、自分が関心のある会議をまとめてリストをつくることできる。そのリストに含まれる会議の情報は、各種の予定管理サイト・ソフトで利用できるiCalendar形式として参照できたり、見やすく可視化された一覧としても閲覧することもできる。また、会議にはタグ情報も付与でき関心のある分野の会議情報を知ることができたり、利用者の被登録数に基づいて多数の人が関心をもつ会議がわかったりするといった機能も備える。

会議への投稿・参加だけでなく、会議の開催・運営をサポートするサービスもある。

EasyChair <http://www.easychair.org/>

Microsoft's Academic Conference Management Service

<http://cmt.research.microsoft.com/cmt/>

これらのサイトは、投稿論文の登録・投稿、査読者への論文割当て、査読結果の回収、査読に関する討議・決定、査読結果の著者への通知といった、会議運営に必要な作業を効率的に管理できるようにするものである。これにより国際会議の運営にかかる作業・コストの負担は大幅に軽減されるようになった。しかし、会議の開催コストの低下は、国際会議の粗製濫造を招き、あまりまともに運営されていないような会議も生じさせるという残念な結果の原因ともなっているようにも思える。

Scholarpedia (<http://www.scholarpedia.org/>)

は、誰でもが執筆できる百科事典である Wikipedia とは異なり、学術的背景をもった各分野の専門家が記名で記事を書くことで、高い品質の記事を集めることを目標としている。Web を利用する特徴としては、読者・利用者が review をすることができ、著者に意見を伝えることができる。ただし、利用者の判断で記事を編集できる Wikipedia とは異なり、その意見を採用するかどうかの最終判断は著者に委ねられている。該当分野の専門家に執筆を依頼しても引き受けてくれるとは限らないため、記事の総数は Wikipedia よりどうしても少なくなるようだ。

2. Web と学術成果の評価

学術成果の評価について Web 上で話題になったことを取り上げた後、成果評価の現状についての問題提起をいくつか紹介する。

数学の重要な未解決問題について、Web 上で話題となった事例を二つ紹介する。一つ目は、位相幾何学に関するポアンカレ予想についてである。G. Perelman によって予想が正しいことが証明されたが、その最初の論文は学術誌ではなく、1.1 節で紹介した arXiv.org で公開された*1。その後、関連分野の専門家の査読を、通常とは異なる手続きで受け、その証明が正しいことが認められた。二つ目は、計算量複雑度に関する $P \neq NP$ 予想についてである。V. Deolalikar によりテクニカルレポートが Web で最初に公開された。こうした著名な未解決問題に対する証明と称する文書は非常に多数出回っており、また、その多くは精査に耐えないものであるため、ほとんどは注目されることがない。しかし、計算複雑度の理論で NP 完全の概念を示しチューリング賞受賞者である S. A. Cook が、検討する価値のある論文であるとコメントをしたため、注目されることとなった。このときは、専門のサイト <http://rjlipton.wordpress.com/2010/08/15/the-p%E2%89%A0np-proof-is-one-week-old/> がつくられ、著名な研究者が議論に参加することとなった。最終的に、この証明には欠陥が見つかり否定されることとなったが、その解決に至るまでの過程は Web 上で進行した。

こうした状況を受け、査読者を割り当てるのではなく、公開した文書に関心のあるほかの研究者が検証を行う open peer review など考えられている [Wing 11]。実際に試験的に実施されたこともあるようだが、結果は良いとも、悪いともいえないものだったとのことである。すでに出版された研究成果を集合的に評価するサイトとして、PubZone (<http://www.pubzone.org/>)

があり、DBLP などの書誌情報サイトと連携しており、文献に対して評価やコメントを付与できる。上記の CiteULike のような文献管理サイトも同様に論文へのタグ付け機能などを備える。人工知能学会全国大会でも 2009 年より開始しているが、Twitter による情報発信は、多くの国際会議で実施されるようになってきている。国際会議でも、各種のソーシャルネットワークサービスに議論の場を設けたり、機械学習関連 ICML でも議論用サイト (<http://mldiscuss.appspot.com/>) を併設するなどの試みが行われている。

これらの Web 上での評価について、著者が考える問題点をいくつかあげておきたい。まず、上記のほとんどのオンライン上での評価・議論は、CiteULike や videlectures.net で多少機能してはいるが、ほかのサイトではほとんど利用されていないように思える。その理由としては、評価のために論文を読み、理解するにはかなりの労力を要するため、論文の評価付けは容易にはできない。議論についても、複雑な内容を伝えたり、テンポの良い議論には向いておらず、細部にわたる内容を論じるには適さないだろう。また、人気に基づく評価は、多様な論文が評価される機会を減らす問題もあると考えている。国際会議や編集委員会に投稿された論文は一通り評価される機会を得られるが、単に Web 上で配布されただけではそうした機会を得られない。評価される機会を得るには、Deolalikar の場合での Cook による示唆のように、何らかの要因で注目されなければならないだろう。こうした論文の質のフィルタリングは、著名な国際会議や論文誌に採録されていたり、すでに著名な研究者であるといった従来のプロセス以上に機能しているものは、現在のところないと考える。加えて、人気による評価は注目されたものがより注目され、それ以外への注目を失わせる効果がある。多数の候補の中から関心のあるものを選び出す推薦システムには、人気のあるものを上位に選び出す人気バイアスがあり、人気を加速的に強め、ほかを抑制する働きをもつ。このバイアスを補正し、多様性を広げさせるには、関心が薄いかもしれない非常に人気のないものを薦めるという手段を取る必要がある [Levy 10]。気に入らない音楽を聞くことと比べ、関心の薄い論文の読解に時間を費やすコストはずっと大きく、こうした手段をとるのは難しいだろう。加えて、良い評価を見たあとでは、その対象について、人間は良い評価をしてしまうバイアスも知られている [Cosley 03]。最後に、研究分野では、たった一人の研究者によって見いだされることで、長い間注目されなかった研究が大きく発展する事例が多々あることは、読者には周知のことであろう。このことから、少なくとも人気によるフィルタリングだけでは、研究分野では不十分であると考えられる。

一方で、現在の評価システムにも問題が指摘されている。まず、1.1 節で紹介したオンライン論文誌 *Journal of Machine Learning Research (JMLR)* の設立時に、

*1 <http://arxiv.org/abs/math.DG/0211159>, <http://arxiv.org/abs/math.DG/0303109>, <http://arxiv.org/abs/math.DG/0307245>

機械学習の有力誌 *Machine Learning* から多数の有力研究者が編集委員会を辞職した事例 (<http://www.sigir.org/forum/F2001/sigirFall01Letters.html>) を紹介する。辞職した研究者の声明では、採録から掲載までの期間の長さや、閲覧がオープンアクセスでないことなどを問題として指摘した。それに対し、*Machine Learning* 誌側も、掲載までは採録原稿をプレプリントとして公開していることや、購読料の適正さなどの反論を示した。この多数の有力研究者の移動に伴い、オープンアクセスの *JMLR* 誌の影響が拡大し、オープンアクセスが意識されるようになった。コンピュータ科学の分野では、論文誌より、国際会議のほうが影響力をもつようになっている [Chen 10]。この文献では、国際会議の採択率と論文のその後の影響についても調査しており、20%前後までは、採択率の降下は引用数の増加と関連があるが、15%より採択率を下げると注目され得る論文を不採録にする可能性が増えるとのことである。データベースの国際会議 *ICDE2010* の基調講演 [Naughton 10] では、採択率が低くなりすぎることを指摘している。データベースのトップ会議 *SIGMOD2010* では、350本の投稿のうち、3人が *accept* 以上だったのは1件のみ、平均で *accept* を超えたものも4件に留まり、評価が厳しくなっていることを指摘している。こうした傾向は論文の寄与よりも、いかに欠点がないかで評価してしまう問題を生じさせている。アルゴリズム系の研究者 O. Goldreich による意見 (<http://www.wisdom.weizmann.ac.il/~oded/on-stocfocs.html>) でも同様の指摘がされている。国際会議への採録が人事と直結しているため、会議に通ったかどうか話題となり、肝心の内容が軽視される傾向を指摘している。会議が競争の場となることで、コンセプトよりも、いかに技術的に解くのが難しいかが競われるようになっているとの傾向も問題であると指摘している。

以上のように現状の評価制度への問題提起と同時に、いくつかの改革案も示されている。自然言語処理の研究者である H. Daumé は、自身のブログで (<http://nlpers.blogspot.com/2009/05/how-to-reduce-reviewing-overhead.html>) のような提案を行っている。最初は二人で査読し、一定以上のもののみを三人目が査読する *tiered review*, *chair* に大きなスクリーニングの権限を与える、短い論文を推奨する、査読した数に応じて被査読可能数を割り当てるなどのアイデアを

示している。自然科学全般での有力誌 *Nature* では、投稿が多数あるため、編集者が査読の前にスクリーニングを実際に実施している (<http://www.natureasia.com/japan/nature/authors/guideline.php>)。査読者によって、査読基準が厳しいものから、甘いものまで幅があり、公平でない場合がある。この変動を吸収すべくデータマイニングの国際会議 *KDD* では、スコアの補正 [Flach 09] を実施している。コンセプトのみをまとめた短い論文と、詳細を述べた長い論文とを同時に投稿し、査読を短い論文で行い、必要に応じて詳細を見る、本会議の前にワークショップをスクリーニングに用いるといった対応も、個人的には、可能なのではないかと考えている。

以上、学術分野を取り巻く Web 上のサービスを紹介し、現状の評価制度について論じた。成果の評価は、研究の発展にとって非常に重要な役割を果たすため、今後とも活発な議論が必要となるであろう。

◇ 参 考 文 献 ◇

- [Chen 10] Chen, J. and Konstan, J. A.: Conference paper selectivity and impact, *Commun. ACM*, Vol. 53, No. 6, pp. 79-83 (2010)
- [Cosley 03] Cosley, D., Lam, S. K., Albert, I., Konstan, J. A. and Riedl, J.: Is seeing believing? How recommender interfaces affect users' opinions, *Proc. SIGCHI Conf. on Human Factors in Computing Systems*, pp. 585-592 (2003)
- [Flach 09] Flach, P. A., Spiegler, S., Golenia, B., Price, S., Guiver, J., Herbrich, R., Graepel, T. and Zaki, M. J.: Novel tools to streamline the conference review process: Experiences from *SIGKDD09, SIGKDD Explorations*, Vol. 11, No. 2, pp. 63-67 (2009)
- [Levy 10] Levy, M. and Bosteels, K.: Music recommendation and the long tail, *WOMRAD 2010: Recsys 2010 Workshop on Music Recommendation and Discovery* (2010)
- [Naughton 10] Naughton, J. F.: DBMS: Lessons from the first 50 years, speculations for the next 50, *ICDE2010 Keynote* (2010), <http://www.cs.utexas.edu/~7edsb/Inspiration/naughtonicde.pdf>
- [Wing 11] Wing, J. M. and Chi, E. H.: Reviewing peer review, *Commun. ACM*, Vol. 54, No. 7, pp. 10-11 (2011)

—— 著 者 紹 介 ——



神 嘉 敏 弘 (正会員)

1992年京都大学工学部情報工学科卒業。1994年同大学院工学研究科情報工学専攻修士課程修了。同年、電子技術総合研究所入所。2001年博士(情報学)。同年電子技術総合研究所は産業技術総合研究所へ再編。機械学習とその応用の研究に従事。AAAI, ACM, 電子情報通信学会各会員。