



## ACM SIGIR 2018 参加報告

### SIGIR とは

SIGIR<sup>☆1</sup> は情報検索分野のトップカンファレンスです。国際学会 ACM の情報検索分科会 (ACM SIGIR<sup>☆2</sup>) が主催し、1978 年からは毎年開かれています。近年はアジア・オセアニア、南北アメリカ、ヨーロッパを順に回っていて、前回は日本、今回は米国、次回以降はフランス、中国、カナダで行われる予定です。

筆者は 2016 年から続けて参加し、周辺分野の動きに敏感でオープンなコミュニティという印象を持ちました。特に最近ではニューラルネットの急速な広がりや中国の勢いを感じています。以下、SIGIR の最近の動向と次回に向けた動きを紹介します。

### 周辺の国際会議とのかかわり

SIGIR の最近の研究対象は周辺分野の多くの国際会議と共有されています。SIGIR とほかの国際会議の論文の参照関係 (2015 年から 2017 年) を見ると<sup>☆3</sup>、情報検索の CIKM<sup>☆4</sup> と ECIR<sup>☆5</sup>、WWW<sup>☆6</sup>、データマイニングの WSDM<sup>☆7</sup>、KDD<sup>☆8</sup> とは相互に参照し合っています。

☆1 ACM SIGIR Conference on Research and Development in Information Retrieval  
 ☆2 Special Interest Group on Information Retrieval. 本稿では区別のため、分科会は 'ACM SIGIR' と表記。  
 ☆3 文献 1) p.36, 2 SIGIR, p.37, 図 1, 図 2.  
 ☆4 ACM International Conference on Information and Knowledge Management  
 ☆5 European Conference on Information Retrieval  
 ☆6 ACM International World Wide Web Conference  
 ☆7 ACM International Conference on Web Search and Web Data Mining  
 ☆8 ACM SIGKDD Conference on Knowledge Discovery and Data Mining

表-1 SIGIR 2018, SIGIR 2017, SIGIR 2016 の概要

会議 (回)	SIGIR2018(41)	SIGIR2017(40)	SIGIR2016(39)
会期	July 8-12	Aug 7-11	July 17-21
開催地	Ann Arbor Michigan, USA	Tokyo, Japan	Pisa, Italy
参加者数	740	911	566
フルペーパー	86/409 (21%)	78	62
ショートペーパー	98/327 (30%)	121	104
デモ	18/36 (50%)	17	21
チュートリアル	11/13 (85%)	8	12
ワークショップ	10/15 (67%)	8	7
SIRIP(Industry Track)	9/16 (56%)	12	12

す。ほかに SIGIR から参照しているのは、機械学習の NeurIPS (NIPS)<sup>☆9</sup> や ICML<sup>☆10</sup>、自然言語処理の ACL<sup>☆11</sup> や EMNLP<sup>☆12</sup>、ヒューマンコンピュータインタラクションの CHI<sup>☆13</sup> などで、SIGIR を参照しているのは情報検索の ICTIR<sup>☆14</sup>、推薦の RecSys<sup>☆15</sup> などです。

### SIGIR 2018

SIGIR 2018 は米国ミシガン州アナーバーのミシガン大学で、2018 年 7 月 8 日から 12 日まで 5 日間開催されました。表-1<sup>☆16</sup> に過去 3 回の概要を示します。今回も例年通り、チュートリアルとワークショップが各 1 日、本会議が 3 日でした。

第 40 回の節目に参加者 911 名 (前年比 1.6 倍) と記録を大幅に更新した<sup>☆17</sup> 昨年に続き、今回も SIGIR 史上 2 番目の約 740 名が参加しました。日本はアカデミアから 13 名以上が参加し、産業界からもヤフー、リクルート住まいカンパニー、リクルートジョブズ、メルカリ、NTT コムウェア、本田技術研究所 などから 17 名以上が参加しました<sup>☆18</sup>。

今回のスポンサーの大半は米国と中国の企業で、分野は情報通信が多く、アジアの企業が増えています<sup>☆19</sup>。新しいスポンサーは中国の車のライドシェアサービスの DiDi、音楽ストリーミングサービスの Spotify、米国の小売業、e コマースの企業です。

### 投稿・採択状況

今回のフルペーパーの投稿数は昨年を上回る 409 本で、うち 86 本が採択され、採択率は 21.0% でした (表-1)。

2016 年以降の国別の投稿・採択状況では、中国の躍進が目立ちます。2016 年は全論文の投稿数、フルペーパー採択数とも米国が 1 位、中国が 2 位でしたが、全論文投稿数では 2017 年、フルペーパー採択数では 2018 年に、中国が逆転して首位に立っています<sup>☆20</sup>。

日本からのフルペーパーは、前回少なくとも 3 本が採択されましたが、今回は採択なしでした。本会議のテクニカルセッションでの日本からの発表は、ジャーナル TOIS<sup>☆21</sup> の論文の口頭発表のみでした<sup>☆22</sup>。

☆9 Conference on Neural Information Processing Systems  
 ☆10 International Conference on Machine Learning  
 ☆11 Annual Meeting of the Association for Computational Linguistics  
 ☆12 Conference on Empirical Methods in Natural Language Processing  
 ☆13 ACM CHI Conference on Human Factors in Computing Systems  
 ☆14 ACM International Conference on the Theory of Information Retrieval  
 ☆15 ACM Conference on Recommender Systems  
 ☆16 文献 1) p.37, 表-1 より引用。フルペーパー以降の行は SIGIR2018 については採択数/投稿数 (採択率)、ほかは採択数を示す。  
 ☆17 2016 年までの最高記録は 2006 年の 719 名。(文献 2) p.198)  
 ☆18 文献 1) pp.36-37, 3.1 開催概要。  
 ☆19 文献 1) pp.37-38, 3.1, 表 2。  
 ☆20 文献 1) pp.37-38, 3.2, 図 5, 図 6。  
 ☆21 ACM Transactions on Information Systems  
 ☆22 文献 1) pp.37-38, 3.2. SIGIR2017 から TOIS の論文の著者は CHIIR, SIGIR, ICTIR のいずれかの国際会議で口頭発表を行えることになった。

2016年以降のフルペーパーの著者の所属はアカデミアが増えています。SIGIR2018はアカデミアが81.6%、企業が15.6%です(図-1<sup>☆23</sup>)。フルペーパーの企業別著者数(図-2<sup>☆24</sup>)を見ると、2016年はヤフーとマイクロソフト、2017年以降はマイクロソフトが突出しています。また、2018年はSpotifyなどのスポンサー企業がフルペーパーでも存在感を見せています。

今回は例年と異なり、6つのトラック<sup>☆25</sup>別に投稿、採択が行われました。投稿数が多かったのは「コンテンツの推薦、分析、分類」(124件、投稿論文の30.0%)、採択数が多かったのは「分野に特化した応用」と「ヒューマンファクターとインタフェース」(各20件、採択論文の各23.2%)です<sup>☆26</sup>。ただし、投稿論文に推薦関連が多く、採択論文にユーザに関するものが目立つ傾向は前回も見られました<sup>☆27</sup>。

### テクニカルセッションの変化

2016年以降のテクニカルセッションの構成と論文の分布を見ると<sup>☆28</sup>、2016年は推薦、評価、効率関連の論文が多く、新しいセッションでは音声と対話、セマンティックWebのエンティティとナレッジグラフができ、ユーザ寄りセッションもありました。

2017年はフィルタリングと推薦、検索モデルとランキング、検索のインタラクションの論文が多く、前年に続き、セマンティックWeb関連や対話関連のセッションもありました。

2018年はセッション単位で見るとランキング学習の論文数が突出していますが、ユーザの行動の分析に関するもの、推薦に関するものは各々複数のセッションがあります。対話やセマンティックWeb関連のセッションは定着しています。新しいセッションとしては情報検索の新しい応用(New IR Applications)、ソーシャルグッド<sup>☆29</sup>(Social

Good)などがありました。情報検索の新しい応用の論文(計4本)は以下の2つに大別され、すべてニューラルネットを用いるものでした。

(a) 複数のモダリティ(テキスト、画像、音声など)にまたがる情報の統合的な検索に関する研究(クエリテキストに適合する動画の区間の検索、料理レシピのテキストと対応する画像の双方向の検索)

(b) 専門家のノウハウとして蓄積された分野特有の知識を利用したマッチングに関する研究(ファッションアイテムのコーディネート、求人の人と仕事のマッチング)

### ニューラルネットに関する動向

SIGIR 2016の基調講演でChristopher Manning氏(スタンフォード大学)は今後2,3年は深層学習がSIGIRを支配すると予測しました(文献3) p.72)。

SIGIR 2017の最大の話はニューラルネットで、ニューラル情報検索のワークショップ(Neu-IR)<sup>☆30</sup>のオープニングではSIGIR 2017の約4分の1の論文がニューラル情報検索に関するもので、「ニューラルネットの年」であると報告されました。

そこで、SIGIR 2018のテクニカルセッションの論文を対象にニューラルネットに関する論文かどうか調べたところ(図-3)<sup>☆31</sup>、フルペーパーの40/87本(46.0%)、TOISの論文も含めると45/102本(44.1%)が該当しました。該当する論文についてセッションでの分布を見ると、フルペーパーの論文はセッションの73.1%(19/26)、TOISの論文も含めると同80.8%(21/26)に分布しています。

### 受賞論文

ベストペーパー<sup>☆32</sup>の受賞はCanamaresら<sup>4)</sup>です。推薦システムでは、popularity(人気)を推薦に利用すべきか、むしろバイアスとして排除すべきか、これまで答えが出ていませんでしたが、利用するのが有効な場合とそうでない場合の条件などを理論的なアプローチで明らかにしました。さらに、実験用データでもすでに知っているものに評価が集まり、偏りがあったことに注目し、偏りのない評価

☆23 文献1) p.38, 図7より引用。

☆24 文献1) p.39, 図8より引用。

☆25 (a) 検索とランキング, (b) 基礎と将来の方向性, (c) 分野に特化した応用, (d) コンテンツの推薦, 分析, 分類, (e) 人工知能, セマンティクス, 対話, (f) ヒューマンファクターとインタフェース。

☆26 <https://dl.acm.org/citation.cfm?id=3209978>. p.iii.

☆27 文献1) p.39, 3.3, 文献2) p.199, 図-4を参照。

☆28 文献1) p.39 3.4, 図10, p.40, 図11, 図12。

☆29 社会に対して良いインパクトを与えようとする活動やサービス、製品などのこと。

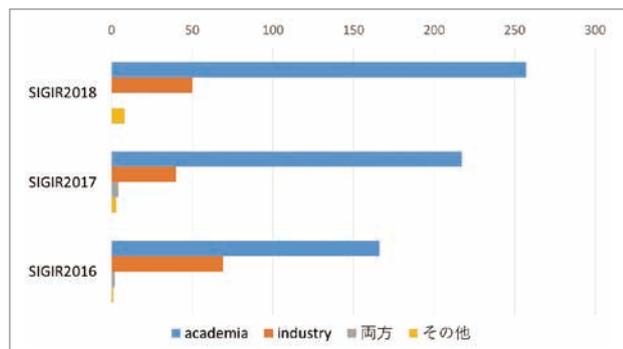


図-1 フルペーパーの著者の所属グループ別著者数(2016～2018年)

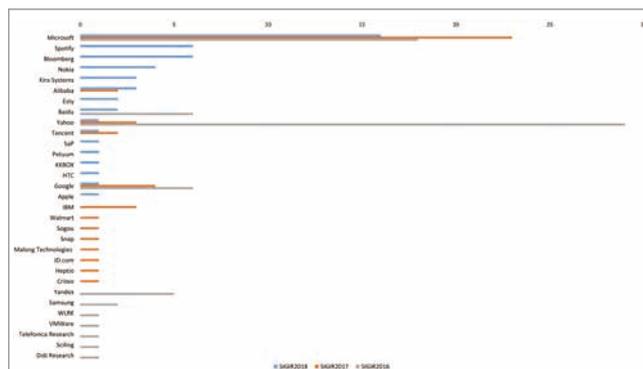


図-2 フルペーパーの企業の著者の所属 (2016～2018年)

用データを新たに作成して検証しました。

### 基調講演

シカゴ大学の研究助教<sup>☆33</sup>の Ghani 氏は、かつての Obama 大統領の選挙キャンペーンのデータサイエンスチームの中心メンバです。講演ではソーシャルグッドと公共政策を取り上げ、機械学習やデータサイエンスを政策課題や社会問題の解決に活用する例として、サービスとそれを必要としている人のマッチングや、特別な支援の必要な生徒の発見などの取り組みを紹介しました。

Gerald Salton 賞<sup>☆34</sup>を受賞したタンペレ大学教授の Jarvelin 氏は情報探索や情報検索の評価に関する研究で著名な人物です。講演ではこれまでの研究を振り返り、情報科学の研究は、人間を仕事や生活の文脈で情報の環境を認識し、それらとインタラクションするアクターとして研究するものと考えたこと、情報検索の研究でも人にかかわる部分が重要であることを強調していました。

### チュートリアルとワークショップ

今回はニューラルネットの対話型 AI への応用、テキストからの知識抽出と推論、周辺分野の技術(確率的トピックモデル、深層学習、ナレッジグラフ、敵対的生成ネットワーク)の情報検索への応用など、11 のチュートリアルがありました。

ワークショップは計 10 で、検索への対話的アプローチ、ナレッジグラフの応用、e コマースは前回から続いています<sup>☆35</sup>。新しいワークショップでは特定分野の専門家による複雑な検索のほか、ITS 関連、驚きの定量化、理由を説明可能な推薦と検索などがありました。e コマースのワークショップでは、商品タイトルから商品のカテゴリを予測する、機械学習のデータチャレンジ (Rakuten Data Challenge)<sup>☆36</sup>の結果が報告されました。

### SIRIP (Industry days)

SIRIP<sup>☆37</sup>は企業での情報検索の応用に関するもので、

2016 年と 2017 年は本会議中の 1 日にテクニカルセッションと並行して行われましたが、今回は 2 日間開催されました。

アマゾンではアレクサショッピングの商品に関する質問応答の研究を紹介しました。ユーザから意見やアドバイスなどを求められる主観的な質問に取り組んでいて、課題は複数の観点からの回答や評価などです。マイクロソフトはインド市場向けのチャットボット、ルー (Ruuh) の開発を紹介しました。若者の心をつかむには、シェアされた画像や動画について意見や感情を表すことも求められるそうです。リンクトインからは人材の検索・推薦を行うシステム (LinkedIn Recruiter) の開発の話がありました。

### 国内のコミュニティの動向

日本での SIGIR 2017 の成功を受けて、2017 年 12 月、ACM SIGIR の初めての支部である東京支部<sup>☆38</sup>が誕生しました。SIGIR 関連の 세미나、論文読み会、メンタリングワークショップなどのイベントや会員向けの情報発信を通じて、国内外の研究コミュニティの交流を促進し、情報検索研究での日本のプレゼンスが向上することを目指しています。2018 年 10 月には論文の読み会<sup>☆39</sup>を開催し、2019 年 1 月にはメンタリングワークショップ POWIR 2018 Winter<sup>☆40</sup>を開催する予定です。

### SIGIR 2019 に向けて

次回、SIGIR 2019 は第 42 回で、2019 年 7 月 21 日から 25 日までパリで ARIA<sup>☆41</sup>の後援を受けて開催されます。公式ページ<sup>☆42</sup>には論文等の募集や次回のトラックの変更<sup>☆43</sup>など、すでに一部の情報が掲載されており、今後の更新情報は公式 Twitter<sup>☆44</sup>でも発信されます。

今後の SIGIR での日本のプレゼンス向上を願っています。

### 参考文献

- 1) 野本昌子, 富樫 陸, 鈴木翔吾, 椎野弘章: SIGIR2018 参加報告, 人工知能学会研究報告, SIG-AM-20-06, pp.36-43 (2018).
- 2) 酒井哲也: 会議レポート: ACM SIGIR 2017 開催報告, 情報処理, Vol.59, No.2, pp.198-199 (Feb. 2018).
- 3) Manning, C.: Natural Language Inference, Reading Comprehension and Deep Learning, <https://nlp.stanford.edu/manning/talks/SIGIR2016-Deep-Learning-NLI.pdf> (2016).
- 4) Canameres, R. and Castells, P.: Should I Follow the Crowd?: A Probabilistic Analysis of the Effectiveness of Popularity in Recommender Systems, Proceedings of the 41st International ACM SIGIR Conference on Research Development in Information Retrieval, pp.415-424 (2018).

(野本昌子/ヤファー (株))

<sup>☆33</sup> Research Assistant Professor

<sup>☆34</sup> Gerald Salton 賞は情報検索の研究に継続的に顕著な貢献を行ってきた人に対して 3 年ごとに贈られる賞。受賞理由は以下を参照。<http://sigir.org/awards/gerard-salton-awards/>

<sup>☆35</sup> 文献 1) p.42 表 5.

<sup>☆36</sup> <https://sigir-ecom.github.io/data-task.html>

<sup>☆37</sup> SIGIR Symposium on IR in Practice. 従来の Industry Track.

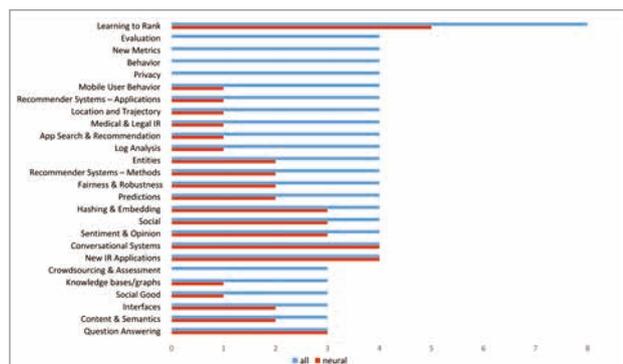


図-3 セッション別の論文数およびニューラルネット関連論文数 (2018 年)

<sup>☆38</sup> Tokyo ACM SIGIR Chapter, <http://sigir.jp/>

<sup>☆39</sup> IR Reading 2018 秋, [http://sigir.jp/post/irreading\\_2018fall/](http://sigir.jp/post/irreading_2018fall/)

<sup>☆40</sup> <https://easychair.org/conferences/?conf=powir2019winter>

<sup>☆41</sup> French Association for Information Retrieval and Applications

<sup>☆42</sup> <http://sigir.org/sigir2019/>

<sup>☆43</sup> 追加は Evaluation (評価), 変更は Future Directions.

<sup>☆44</sup> <https://twitter.com/sigir2019>