ACL 2019 Program Co-Chairs Report

1. ACL 2019 Organizing committee

General Chair

Lluís Màrquez - Amazon

Program Co-Chairs

Anna Korhonen - University of Cambridge

David Traum - University of South California

Local Organisation Co-Chairs

Alessandro Lenci - University of Pisa

Bernardo Magnini - Bruno Kessler Foundation

Simonetta Montemagni - Institute for Computational Linguistics of CNR (CNR-ILC)

Workshop Co-Chairs

Barbara Plank - IT University of Copenhagen

Sebastian Riedel - University College London and Facebook AI Research

Student Research Workshop Co-Chairs

Fernando Alva-Manchego - University of Sheffield

Eunsol Choi - University of Washington

Daniel Khashabi - University of Pennsylvania

SRW Faculty Advisors

Hannaneh Hajishirzi - University of Washington

Aurelie Herbelot - University of Trento

Scott Wen-tau Yih - Facebook AI Research

Yue Zhang - Westlake University

Tutorial Co-Chairs

Preslav Nakov - Qatar Computing Research Institute, HBKU

Alexis Palmer - University of North Texas

Demo Co-Chairs

Enrique Alfonseca - Google

Marta R. Costa-jussà - Technical University of Catalonia

Publication Co-Chairs

Douwe Kiela - Facebook

Ivan Vulić - University of Cambridge

Shay Cohen - University of Edinburgh (Advisory)

Kevin Gimpel - Toyota Technological Institute at Chicago (Advisory)

Conference Handbook Co-Chairs

Elena Cabrio - Université Côte d'Azur, Inria, CNRS, I3S

Rachele Sprugnoli - Università Cattolica del Sacro Cuore

Conference App Chair

Andrea Cimino - Institute for Computational Linguistics of CNR (CNR-ILC)

Local Arrangement Committee

Sara Goggi - Institute for Computational Linguistics of CNR (CNR-ILC) (Co-Chair)

Maria Cristina Schiavone - MCI Group (Co-Chair)

Sacha Bourdeaud'Hui - Mamoka (Web Manager)

Local Sponsorship Co-Chairs

Roberto Basili - University of Rome, Tor Vergata

Giovanni Semeraro - University of Bari

Publicity Co-Chairs

Felice Dell'Orletta - Institute for Computational Linguistics of CNR (CNR-ILC)

Lucia Passaro - University of Pisa

Sara Tonelli - Bruno Kessler Foundation

Mentorship Co-Chairs

Rada Mihalcea - University of Michigan

Robert Frederking - Carnegie Mellon University (CMU)

Aakanksha Naik - Carnegie Mellon University (CMU)

Student Volunteer Coordinators

Dominique Brunato - Institute for Computational Linguistics of CNR (CNR-ILC)

Marco Senaldi - University of Pisa, Scuola Normale Superiore (Pisa)

Giulia Venturi - Institute for Computational Linguistics of CNR (CNR-ILC)

2. Senior Program committee

Program Co-Chairs

Anna Korhonen- University of Cambridge

David Traum - University of South California

Assistants

Simon Baker, Qianchu (Flora) Liu, Olga Majewska, Edoardo Ponti, Ehsan Shareghi

Senior Area Chairs (SACs) and Area Chairs (ACs)

Dialogue and Interactive Systems

SACs: Kallirroi Georgila, Ryuichiro Higashinaka

ACs: Michel Galley, Zhou Yu, Milica Gasic, Rebecca J. Passonneau, Gabriel Skantze, Matthew Marge, Helen Hastie, Kazunori Komatani, Yun-Nung Chen, Pascale Fung

Discourse and Pragmatics

SACs: Annie Louis, Andrew Kehler

ACs: Benamara Farah, Giuseppe Carenini, Michael Strube, Bonnie Webber, Smaranda Muresan, Manfred Stede

Document Analysis

SACs: Bracha Shapira, Eugene Agichtein

ACs: Michael Bendersky, Dilek Hakkani-Tur, Anton Leuski, Andrei Popescu-Belis, Peng Zhang, Xiang Ren, Sujian Li

Generation
SACs: Cecile Paris, Kees van Deemter
ACs: Leo Wanner, Anya Belz, Leila Kosseim, Advaith Siddarthan, Nadjet Bouayad, Kathy McCoy, Stephanie M. Lukin, Matthew Stone, Nina Dethlefs, John Kelleher, Paul Piwek, Yoav Goldberg

Information Extraction and Text Mining
SACs: Alessandro Moschitti, Heng Ji, Mausam, Hannaneh Hajishirzi
ACs: Isabelle Augenstein, Nazli Goharian, Ruihong Huang, Kevin Cohen, Siddharth Patwardhan, Sumithra Velupillai, Yunyao Li, Gerard de Melo, Mark Stevenson, Avi Sil, Aurélie Névéol, Kenneth Church, Alan Ritter, Hoifung Poon, Nut Limsopatham

Linguistic Theories, Cognitive Modeling and Psycholinguistics
SACs: Frank Keller, Aline Villavicencio
ACs: Afra Alishahi, Yevgeni Berzak, Shuly Wintner, Vera Demberg, Emily Prud'hommeaux

Machine Learning
SACs: Chris Dyer, Ariadna Quattoni
ACs: Ashish Vaswani, Kai-Wei Chang, Fei Sha, Barbara Plank, William Wang, Tim Rocktäschel, Le Sun, Jason Naradowsky, Alice Oh, Amir Globerson, Pontus Stenetorp, Andreas Vlachos

Machine Translation
SACs: Trevor Cohn, Yang Liu
ACs: Dekai Wu, Kevin Duh, Jörg Tiedemann, Deyi Xiong, Taro Watanabe, Philipp Koehn, Marine Carpuat, Arianna Bisazza, Alexander Fraser, Zhaopeng Tu, Qun Liu, Yvette Graham, Daniel Cer, Minh-Thang Luong

Multidisciplinary (also for AC COI)
SACs: Patrick Pantel, Julia Hockenmaier
ACs: Yoav Artzi, Bowen Zhou, Grzegorz Chrupała, Dong Nguyen, Simone Paolo Ponzetto, Sara Rosenthal

Multilinguality
SACs: Joakim Nivre, Timothy Baldwin
ACs: Anders Søgaard, Jonathan May, Christian Hardmeier

Phonology, Morphology and Word Segmentation
SACs: Graham Neubig, Hinrich Schütze
ACs: Ryan Cotterell, Manaal Faruqui, Hai Zhao, Kemal Oflazer, Miikka Silfverberg

Question Answering
SACs: Sanda Harabagiu, Zornitsa Kozareva
ACs: Kang Liu, Yansong Feng, Shafiq Joty, Eric Nyberg, Preslav Nakov, Giovanni Da San Martino, Jennifer Chu-Carroll, Idan Szpektor

Resources and Evaluation
SACs: Sara Tonelli, Ron Artstein
ACs: Gina-Anne Levow, Thierry Declerck, Nancy Ide, Kenji Sagae, Udo Kruschwitz, Beata Megyesi, Roberto Navigli, Owen Rambow

Sentence-level Semantics
SACs: Mona Diab, Ivan Titov
ACs: Wei Xu, Siva Reddy, Steven Bethard, Eduardo Blanco, Wenpeng Yin, Liang Huang, Edward Grefenstette, Michael Roth, Mehrnoosh Sadrzadeh, Anette Frank

Sentiment Analysis and Argument Mining
SACs: Marie-Francine Moens, Bing Liu
ACs: Saif Mohammad, Els Lefever, Liang-Chih Yu, Yulan He, Oren Tsur, Claire Cardie, Yue Zhang, Swapna Somasundaran, Jinho D. Choi

Social Media
SACs: Cristian Danescu-Niculescu-Mizil, Jacob Eisenstein
ACs: Kalina Bontcheva, Nigel Collier, Dirk Hovy, David Jurgens, Tim Finin, Diyi Yang, Wei Gao, Wei Wei
Summarization
SACs: Mirella Lapata, Chin-Yew Lin
ACs: Wenjie Li, Xiaojun Wan, Jackie Chi Kit Cheung, Shashi Narayan, Xiaodan Zhu, Fei Liu

Tagging, Chunking, Syntax and Parsing:
SACs: Phil Blunsom, Noah A. Smith
ACs: Roi Reichart, Marek Rei, Daisuke Kawahara, Emily Pitler, Omri Abend, Weiwei Sun

Textual Inference and Other Areas of Semantics:
SACs: Sabine Schulte im Walde, Raffaella Bernardi
ACs: Omer Levy, Angeliki Lazaridou, Jonathan Berant, Vivek Srikumar, Dimitri Kartsaklis, Christopher Potts, Roy Schwartz

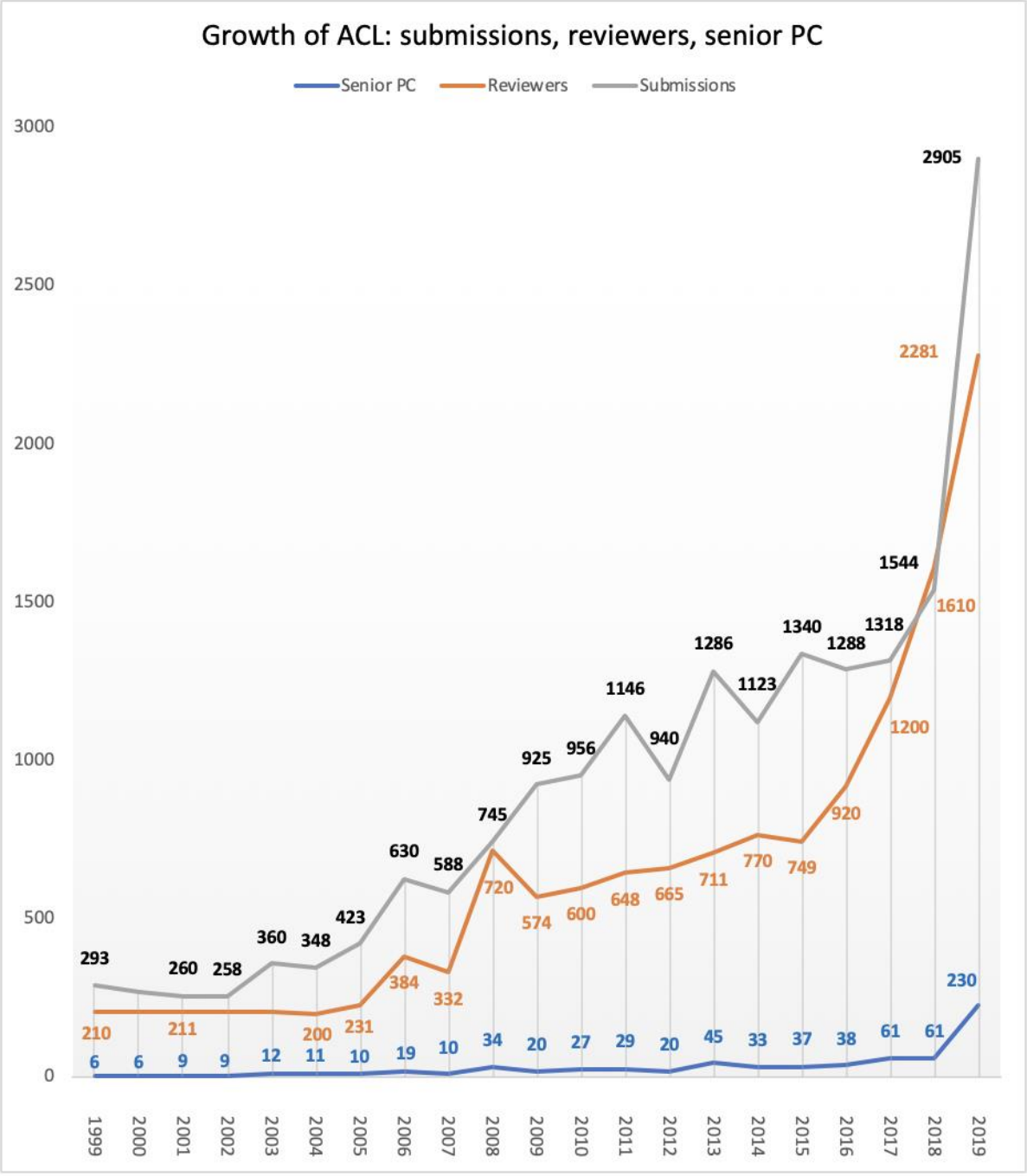
Vision, Robotics, Multimodal, Grounding and Speech:
SACs: Louis-Philippe Morency, Michael Johnston
ACs: Catharine Oertel, Matthias Scheutz, Sakriani Sakti, Elia Bruni, Manny Rayner, Douwe Kiela, Yonatan Bisk, Yale Song

Word-level Semantics
SACs: Eneko Agirre, Diana McCarthy
ACs: Mohammad Taher Pilehvar, Ekaterina Shutova, Ivan Vulić, Laura Rimell, Paul Cook, Chris Biemann, Marianna Apidianaki, Diarmuid Ó Séaghdha, Jose Camacho-Collados, Aitor Soroa

Applications
SACs: Joel Tetreault, Karin Verspoor
ACs: Ekaterina Kochmar, Vinodkumar Prabhakaran, Razvan Bunescu, Sarvnaz Karimi, Filip Ginter, Vincent Ng, Beth Ann Hockey, Jens Edlund, Maria Liakata

3. Main Innovations

Given the rise of AI, Natural Language Processing has become increasingly popular and almost all recent conferences have reported a record breaking number of submissions. Yet, never in the history of ACL have we seen such a dramatic growth: within just a single year, we have gone from 1544 submissions to 2905! This is illustrated in the following graph that shows the growth of ACL over the past 20 years in terms of the number of submissions, reviewers and (Senior) Area Chairs.



Review of such a large number of submissions requires a large, well-organised Program Committee. Extending the ACL 2018 practice, we created a structure similar to the conferences that have a Senior Program Committee alongside the Program Committee. For the Senior PC, we recruited a relatively large number of Senior Area Chairs (46, 2-4 to head each area) and Area Chairs (184, 3-15 per area). We also differentiated between their roles so that SACs assign papers to ACs and reviewers and make recommendations for their area, while ACs each manage a smaller set of papers within the area, lead discussions with reviewers, write meta-reviews and make initial recommendations. This structure also helps to compensate for the problem that our rapidly growing field is suffering from: the lack of experienced

reviewers. As ACs focus on a smaller number of papers, they can pay more attention to the review process. As for reviewers, we recruited many of them this year: 2281 (ACL 2018 had 1610).

With such a huge number of submissions, every step of conference organization (from the initial checking of submissions to decision making) takes longer than before. Knowing the timeline would be extremely tight, we looked into ways of improving efficiency. We wanted to improve efficiency in ways that would optimise the experience for authors and PC members. In particular, we reduced the number of deadlines requiring a short turn-around of 3 days (or less). Such deadlines at best are stressful for all, but often work poorly, given the diversity of work and life situations in the community (i.e. the great variation in times / days when people are actually available for conference-related work).

We implemented the following changes:

- We dropped the paper bidding phase. This phase can take several days of time, and given the large number of submissions, reviewers find it increasingly time consuming, moreover, not all reviewers complete it. However, the time considerations aside, we were also worried about the impact of reviewers choosing their favourite papers for review, as opposed to choosing to review papers that they are qualified to review (for an interesting blog post on the topic, see https://naacl2018.wordpress.com/2018/01/28/a-review-of-reviewer-assignment-methods/). Our plan was to rely on the Toronto Paper Matching System (TPMS) in allocating papers to reviewers. Unfortunately, this system didn't prove as useful as we had hoped for (it requires more extensive reviewer profiles for optimal performance than what we had available) and the work had to rely largely on the manual effort. Our fantastic SACs did an outstanding job here, but this is clearly a task that needs better automated support.

- Like NAACL 2019, we didn't have an author response phase this year. Originally introduced as an improvement to the review process, author response has proven time-consuming (taking not only authors but also reviewers and chairs time) and not hugely impactful on a larger scale. For example, the following paper (that appeared in NAACL 2019) summarises relevant data from ACL 2018:

  Does My Rebuttal Matter? Insights from a Major NLP Conference
  Yang Gao, Steffen Eger, Ilia Kuznetsov, Iryna Gurevych and Yusuke Miyao

  So, instead of author response, we decided to invest in promoting discussion within the PC, and on ensuring that discussions, papers and reviews have the full attention of ACs.

- Finally, in contrast with the elaborate review forms of some recent conferences, we adopted much simpler, streamlined review form, adapted from EMNLP 2018 (many thanks to Julia Hockenmaier, David Chiang and Junichi Tsujii!). While encouraging thorough review, this form is less laborious for reviewers and more focused on highlighting the key points for decision making and feedback to authors.

4. Timeline

- Sept 2018 — Call for Nominations for reviewers, SACs, ACs and invited speakers
- Dec 2018 - Jan 2019 — [PC chairs] Recruitment of SACs and ACs
- Jan - Feb 2019 — [SACs] Recruitment of reviewers and additional ACs
- March. 4, 2019: — Paper submission deadline (both long and short)
- March 5-7, 2019: — [PC chairs and assistants]] check papers for format, assign them to areas
- March 8-11: — [SAC] check papers in each area, report COIs and other problems, recruit more reviewers (and ACs) if needed
- Mar. 12-15: — [SAC] allocate papers to reviewers and to ACs, and ask them to check their assignments
- March 14-16: — [AC and reviewers] check their assignments and report COIs and other issues.
- March 17- April 9: — Review period
- April 10-23: — [AC] chase late reviewers, lead discussion
- April 24-25: — [AC] finalise meta-reviews and recommendations
- April 26-May 1: — [SAC] rank papers and make initial decisions and recommendations
- May 2- May 12: — [PC chairs] make final decisions
- May 13: — [PC chairs] notification of acceptance
- June 3: — Camera ready due
- July 28, 2019 — Tutorials
- July 29-31, 2019 — Conference
- August 1-2, 2019 — Workshops and Co-located conferences

All deadlines were 11:59PM UTC-12:00 ("anywhere on Earth").

5. Submissions

5.1 Overview of statistics

The conference received as many as 2905 submissions by the submission deadline - a 75% increase over ACL 2018 and is an all-time record for ACL-related conferences! 119 of the submissions were withdrawn by authors and 92 desk-rejected due to issues such as dual submissions, plagiarism or submissions not conforming to the submission guidelines. The resulting 2694 valid submissions, including 1609 long and 1085 short papers, were sent to review.

5.2. Detailed statistics

The following table shows, for each area, the number of submissions (long, short and total) that underwent review.

	Area	Long	Short	Total
1.	Information Extraction, Text Mining	156	93	249
2.	Machine Learning	148	73	221
3.	Machine Translation	102	105	207
4.	Dialogue and Interactive Systems	125	57	182
5.	Generation	97	58	155
6.	Question Answering	99	55	154
7.	Sentiment Analysis, Argument Mining	91	60	151
8.	Word-level Semantics	78	59	137
9.	Applications	65	72	137
10.	Resources and Evaluation	70	60	130
11.	Multidisciplinary, AC COI	70	44	114
12.	Sentence-level Semantics	70	42	112
13.	Tagging, Chunking, Syntax, Parsing	50	49	99
14.	Social Media	51	42	93
15.	Summarization	48	35	83
16.	Document Analysis	48	33	81
17.	Vision, Robotics Multimodal Grounding, Speech	56	23	79
18.	Multilinguality	43	32	75
19.	Textual Inference, Other Areas of Semantics	44	30	74
20.	Linguistic Theories, Cognitive, Psycholinguistics	39	21	60
21.	Discourse and Pragmatics	33	24	57
22.	Phonology, Morphology, Word Segmentation	26	18	44
	Total	**1609**	**1085**	**2694**

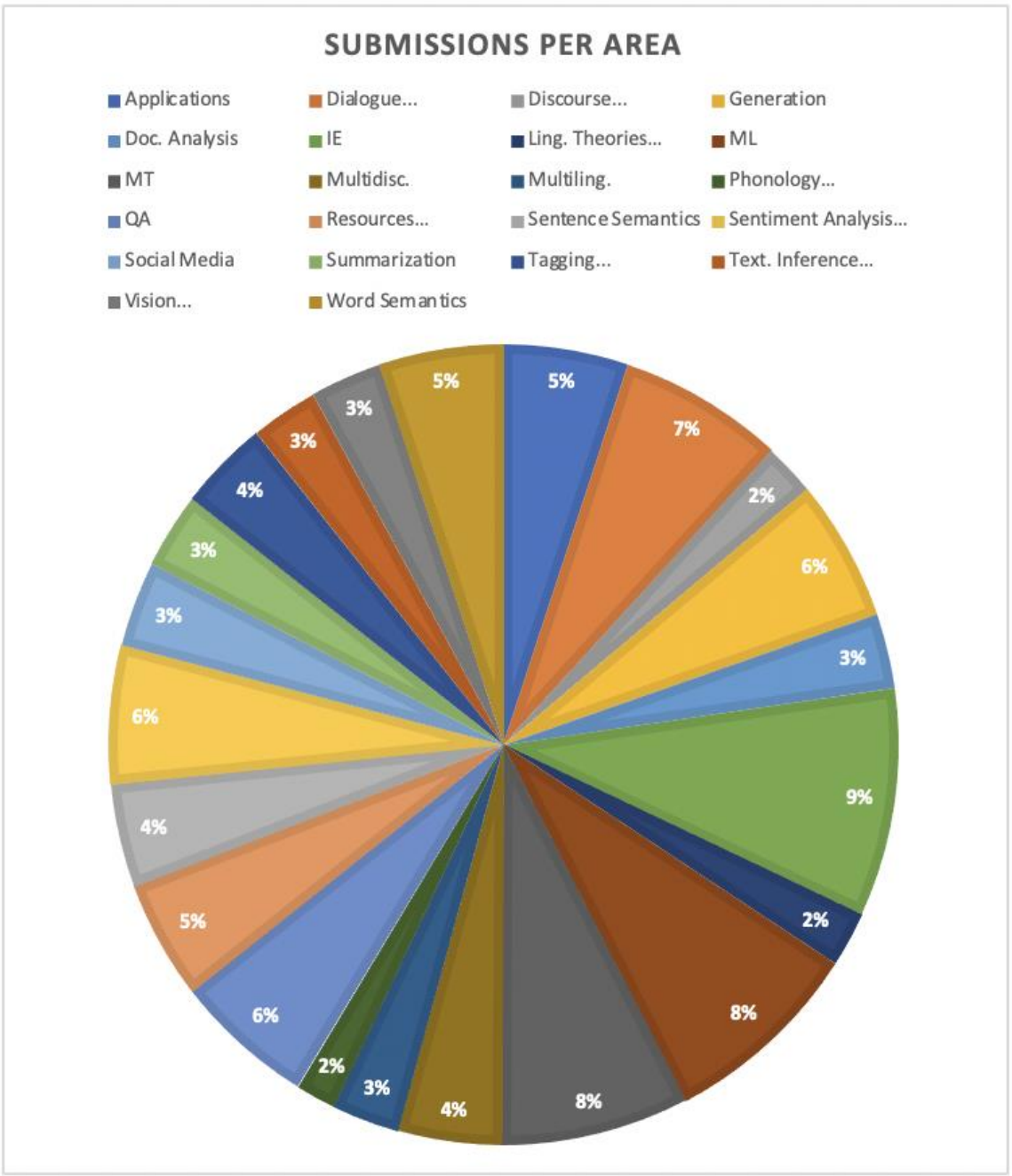
Our 3 largest areas in terms of submissions are the same as in ACL 2018:

- Information Extraction and Text Mining (9.2% of all valid submissions vs. 11.5% in ACL 2018 – note that the percentages are not fully comparable because this year's conference features an additional area, Applications)
- Machine Learning (8.2% vs. 7.4% in ACL 2018)
- Machine Translation (7.7% vs. 8.3% in ACL 2018)

Also Dialogue and Interactive systems are among the top 5 areas in both conferences. However, Document Analysis, which was the 4th largest area last year, ranks only the 16th this year, while Generation (which ranked the 14th last year with 59 submissions) is ranked now the 5th with 156 submissions (the increase

in submissions is much larger here than our overall growth rate!). Another surprise is Linguistic Theories, Cognitive Modeling and Psycholinguistics, which clearly grew in popularity: 24 submissions last year, 60 this year.

Submissions remain still relatively evenly distributed across the different areas (see the below pie chart) in comparison with e.g. in ACL 2017 where IE was clearly dominating (23.4% of submissions).



6. Review Process

6.1 A call for nominations:

In October 2018, we issued a Call for Nominations for potential reviewers, ACs, SACs and invited speakers. We did this in order to find the best candidates from our research community and help us ensure broad coverage of expertise in all areas of NLP. People could nominate candidates in any of the following four categories. Self-nominations were welcome and encouraged for the first three of them:

- **Reviewers:**

  As a minimum requirement, nominated reviewers must have a good publication record in NLP/CL. Apart from that, the most important factors are thoroughness and reliability.

- **Area Chairs:**

  Nominated ACs must have a PhD in an area related to NLP/CL, a strong publication record in leading NLP/CL venues, and an extensive experience in reviewing for such venues.

- **Senior Area Chairs:**

  Nominated SACs must be experienced NLP/CL researchers with an impressive research and publication record in leading NLP/CL venues. Previous Area Chairing experience from *CL and EMNLP conferences is highly desirable.

- **Invited Speakers:**

  It is expected that the Invited Speaker is not only an exceptional researcher but also a community influencer and someone who can deliver an inspiring talk. Nominations from outside the mainstream NLP/CL community are also welcome!

By November 2018, we received 862 valid responses proposing candidates for

- 751 Reviewers
- 114 Area Chairs
- 28 Senior Area Chairs
- 11 Invited Speakers

6.2 Recruiting (senior) area chairs

During December 2018 – January 2019, we recruited the Senior Program Committee, i.e. SACs and ACs. We had decided on our 22 thematic areas (the same areas as in ACL 2018 plus one additional one: Applications) and wanted each area to be headed by 2-4 SACs who would be in charge of the overall review process within the area. We also wanted each area to be assisted by ACs who would look after a subset of the papers (not more than 15 on average). This meant that we needed a fairly large committee. We used the nominations as a starting point for recruitment, but also carried out additional recruitment to obtain a wider and more diverse pool of qualified candidates.

We paid attention to gender and geographical balance. Out of the 46 individuals who finally accepted our invitation to act as SAC, 27 (58.7%) were male and 19 (41.3%) female. Most of the areas (13 out of 22) were chaired by a female and a male SAC. 7 were chaired by two males and 3 by two females. From the final set of 184 ACs, 123 (66.8%) were male and 61 (33%) were female. In some of the areas (in particular Machine Learning, Machine Translation, Phonology, Morphology and Word Segmentation, Textual Inference and Multilinguality) it proved challenging to obtain a good gender balance. The geographic

balance of Senior PC includes the following numbers: Europe - 87 (38%), North America - 99 (43%), Asia and Pacific - 43 (18%), none from Africa and South America.

6.3 Recruiting reviewers

In February 2019, SACs were provided with the list of reviewer nominations along with a list of reviewers from previous conferences. They used these as a starting point for recruiting reviewers for their area. In the end, we had a large pool of 2281 potential reviewers, 59-319 per area.

6.4 Checking submissions for violations, areas and COIs

Immediately after the submission deadline on 3 March, 2019, the paper submissions were checked by the PC chairs and assistants. Plagiarism detection software was ran on the entire submission archive, which identified a few cases of plagiarism. Assistants also opened all the submission files and checked that they conformed to the submission requirements. Reported issues were checked by PC chairs who desk-rejected the papers with serious violations of the submission guidelines.

The remaining papers were checked for areas. When submitting their paper, authors had selected the primary area for their paper, along with a number of keywords. The areas and keywords were based on the 22 subject areas of the conference. Papers that were deemed a better fit with another area where moved to that area. Papers that had COIs with one or both of the SACs with were also moved to another area. In total, 592 papers (20% of the submissions) moved areas during this process.

The main problem we faced during this phase was lack of time: the timeline was way too tight considering the number of submissions. For example, the PC chairs couldn't complete the checking of areas for papers but got only 40% of it done. They had to ask SACs to complete this task and had to also encourage them to report any violations of the submission guidelines that were missed during the pre-checks.

6.5 Assigning papers to reviewers and to ACs

SACs assigned papers to reviewers and to ACs using the Toronto Paper Matching System (TPMS) and/or manually. Unfortunately, because many reviewer profiles were incomplete or missing, TPMS was not as useful as we had hoped for and this phase proved more time consuming than intended.

After the assignments were ready, they were still checked by reviewers and ACs. Once we had dealt with the requests for re-assignment, the review period started.

6.6 Reviewing

Each paper was reviewed by at least three reviewers. Reviewers were given around three weeks time to complete the reviewing. Despite the usual reminders, over 1000 reviews were still missing at the deadline (out of over 8000 total). Fortunately, we had some ability to compensate given our lack of author response to allow reviewers to complete their reviews (and recruit replacement reviewers), but this would have been a bigger problem in trying to implement author response fairly.

6.7 Decision making

Area Chairs:

After the reviews were in, ACs took over. They were given two weeks for leading discussions on papers and for writing meta-reviews and making initial decisions and recommendations.They were asked to recommend for each paper whether it should be

- Best Paper candidate
- Accept
- Lean to accept
- Can't Decide if it should be accepted or not
- Lean to Reject
- Reject

The ACs were advised to based their recommendations on the reviews, the reviewer discussion, and their own expert judgement. They were also encouraged to read and review the papers themselves to make the final judgement where needed.

Senior Area Chairs:

After ACs had finished with their task, SACs were given a week to rank papers and make recommendations to PC chairs. They were advised to base their decisions on meta-reviews and, for any paper that ACs had recommended as potential 'accept' they were also asked to look at reviews and discussion. Based on all these materials and their own judgement, they were asked to recommend, for each paper, whether it should be

1. Definitely Accepted (no more than 20% of submissions)
2. Could be Accepted (no more than 20%)
3. Rejected (at least 60%)

and were asked to edit the meta-reviews accordingly.

PC chairs:

The PC chairs reviewed the recommendations from SACs, and considered all the potential 'accept' recommendations carefully, looking at all the available materials. They also looked for any inconsistencies between recommendations from reviewers, ACs and SACs, and re-considered papers where such inconsistencies were detected. They determined the acceptance thresholds for each area depending on the overall quality of the area.

6.8 Statistics on acceptances

Especially in light of the record number of submissions, the selection process was very competitive. 660 papers were finally accepted to appear in the conference, resulting in the overall acceptance rate of 22.7%. This is a little lower than the acceptance rate for ACL 2018 (24.9%) or ACL 2017 (23.3%) – yet remarkably similar when we consider the 75% increase in submissions from ACL 2018. We accepted 447 long paper and 213 short paper submissions. As in previous years, the acceptance rate is clearly higher for long papers (25.7% vs. 18.2% for short papers), showing once again that short papers are harder to get accepted than long ones:

Conference		Submissions	Accepts	Accept rate (%)
ACL 2019	**All**	**2905**	**660**	**22.7**
	Long	1737	447	25.7
	Short	1168	213	18.2
ACL 2018	**All**	**1544**	**384**	**24.9**
	Long	1018	258	25.3
	Short	526	126	24.0
ACL 2017	**All**	**1297**	**302**	**23.3**
	Long	737	195	26.5
	Short	560	107	19.1

The most challenging areas in terms of acceptance are Document Analysis (18.5%) and Sentence-level Semantics (19.8%), along with Information Extraction and Text Mining (20.6%), Word-level Semantics (20.7%) and Phonology, Morphology and Word Segmentation (20.9%). In contrast, the area with the highest acceptance rate is Multidisciplinary and Area Chair COI (31.5%). This area handled the papers Senior Area Chairs had conflicts of interest with in their own areas. Other relatively high acceptance areas include Vision, Robotics, Multimodal Grounding and Speech (30.0%), Dialogue and Interactive Systems (28.4%) and Resources and Evaluation (28.1%).

	Area	All submissions	Accepts	Accept rate (%)
1.	Applications	136	32	23.5
2.	Dialogue and Interactive Systems	183	52	28.4
3.	Discourse and Pragmatics	55	15	27.3
4.	Document Analysis	81	15	18.5
5.	Generation	153	40	26.1
16.	Information Extraction and Text Mining	247	51	20.6
7.	Linguistic Theories, Cognitive Modeling and Psycholinguistics	60	14	23.3
8.	Machine Learning	223	56	25.1
8.	Machine Translation	205	46	22.4
10.	Multidisciplinary and Area Chair COI	112	35	31.3
11.	Multilinguality	75	21	28.0
12.	Phonology Morphology and Word Segmentation	43	9	20.9
13.	Question Answering	155	39	25.2
14.	Resources and Evaluation	128	36	28.1
15.	Sentence-level semantics	111	22	19.8
15.	Sentiment Analysis and Argument Mining	150	33	22.0
17.	Social Media	93	23	24.7
18.	Summarization	81	21	25.9
19.	Tagging Chunking Syntax and Parsing	99	27	27.3
20.	Textual Inference and Other Areas of Semantics	74	21	28.0
21.	Vision Robotics Multimodal Grounding and Speech	80	24	30.0
22.	Word-level Semantics	135	28	20.7
	Desk reject or withdrawn	225		
	Total	**2905**	**660**	**22.7**

There are also interesting differences between countries/regions (as defined by the START system). We looked at these taking into account the countries/regions of the corresponding authors of papers only (which

is clearly a simplification). We have 64 countries/regions represented among the corresponding authors. The 51 with more than one submission are shown in the table. The distribution of submissions across countries/regions is too skewed for fair comparison of acceptance rates. However, if we consider the top 15 countries in terms of the number of submissions (each with more than 30 submissions) only, the ones with the highest acceptance rates are Singapore(34.8%), Israel (34.1%), the UK (29.7%), Hong Kong (29.4%) the US (28.8%), and Germany (28.7%).

	All submissions			Long submissions			Short submissions		
Country or Region	Sub.	Acc.	Rate (%)	Sub.	Acc.	Rate (%)	Sub.	Acc.	Rate (%)
Australia	46	11	23.9	22	4	18.2	24	7	29.2
Austria	5	0	0.0	2	0	0.0	3	0	0.0
Belgium	8	1	12.5	3	1	33.3	5	0	0.0
Brazil	11	0	0.0	6	0	0.0	5	0	0.0
Canada	74	16	21.6	44	12	27.3	30	4	13.3
Chile	2	0	0.0	2	0	0.0	0	0	N/A
China	817	155	19.0	567	118	20.8	250	37	14.8
Czech Republic	12	2	16.7	5	0	0.0	7	2	28.6
Denmark	25	4	16.0	11	1	9.1	14	3	21.4
Egypt	2	0	0.0	1	0	0.0	1	0	0.0
Estonia	2	0	0.0	2	0	0.0	0	0	N/A
Finland	6	0	0.0	2	0	0.0	4	0	0.0
France	60	11	18.3	32	4	12.5	28	7	25.0
Germany	136	39	28.7	73	26	35.6	63	13	20.6
Greece	7	4	57.1	1	1	100.0	6	3	50.0
Hong Kong	34	10	29.4	26	9	34.6	8	1	12.5
Hungary	7	1	14.3	3	1	33.3	4	0	0.0
India	107	18	16.8	54	16	29.6	53	2	3.8
Iran	3	0	0.0	2	0	0.0	1	0	0.0
Ireland	10	1	10.0	4	1	25.0	6	0	0.0
Israel	41	14	34.1	30	11	36.7	11	3	27.3
Italy	50	6	12.0	25	3	12.0	25	3	12.0
Japan	125	23	18.4	58	13	22.4	67	10	14.9
Luxembourg	2	0	0.0	2	0	0.0	0	0	N/A
Macau	5	1	20.0	3	1	33.3	2	0	0.0
Malta	2	0	0.0	0	0	N/A	2	0	0.0
Mexico	2	0	0.0	0	0	N/A	2	0	0.0
Netherlands	36	9	25.0	22	8	36.4	14	1	7.1
Norway	6	2	33.3	4	1	25.0	2	1	50.0
Pakistan	2	0	0.0	1	0	0.0	1	0	0.0
Peru	2	0	0.0	1	0	0.0	1	0	0.0
Poland	7	1	14.3	5	1	20.0	2	0	0.0
Portugal	8	3	37.5	4	2	50.0	4	1	25.0
Qatar	4	0	0.0	2	0	0.0	2	0	0.0
Republic of Korea	72	7	9.7	36	4	11.1	36	3	8.3
Romania	2	1	50.0	2	1	50.0	0	0	N/A
Russian Federation	14	4	28.6	7	2	28.6	7	2	28.6
Singapore	46	16	34.8	39	13	33.3	7	3	42.9
Slovakia	2	0	0.0	1	0	0.0	1	0	0.0
South Africa	2	1	50.0	1	0	0.0	1	1	100
Spain	29	6	20.7	12	1	8.3	17	5	29.4
Sri Lanka	5	0	0.0	1	0	0.0	4	0	0.0
Sweden	9	0	0.0	4	0	0.0	5	0	0.0
Switzerland	23	4	17.4	12	2	16.7	11	2	18.2
Taiwan	46	6	13.0	18	3	16.7	28	3	10.7
Thailand	2	0	0.0	1	0	0.0	1	0	0.0
Turkey	7	0	0.0	3	0	0.0	4	0	0.0
United Arab Emirates	4	2	50.0	1	1	100.0	3	1	33.3
United Kingdom	138	41	29.7	84	30	35.7	54	11	20.4
United States	820	236	28.8	485	154	31.8	335	82	24.5
Others	18	2		12	0		6	3	
TOTAL	**2905**	**660**	22.7	**1737**	**447**	25.7	**1168**	**213**	18.2

7. Presentations

Out of the 660 accepted papers, 245 were selected for oral presentation and the remaining 415 to be presented as posters. In addition, 21 TACL papers are presented at the conference (16 oral and 5 poster).

We have 48 oral sessions in total, with 6 running in parallel. There we two types of oral session. One featured 5 long papers and the other one 3 long papers and 3 short. Long papers were given a 20 minute slot (15 min for the talk, 4 min for questions, 1 minute for setting up) and the short papers were given a 13 min slot (10 min for the talk, 2 for questions and 1 for setting up).

We have 8 poster sessions in total where posters are grouped thematically according to areas, and larger areas are split between different sessions.

Sessions chairs were found by opening a google document for SACs and ACs and letting them sign up and choose the sessions they wanted to chair.

There were several requests for remote presentations. Remote presentations were granted only when no authors capable of presenting the conference could attend the conference (e.g. for health or visa reasons). The presentations are to be recorded in advance and the presenting author would answer questions via teleconferencing.

8. Best papers

During and following the review process, 79 papers were nominated as potential best paper candidates by reviewers, ACs and/or SACs. The PC chairs evaluated all of these papers and, after careful consideration, created a short-list of 28 best paper candidates. They clustered these papers thematically into 6 broad themes and recruited 2-3 experts to evaluate and rank the 4-5 papers under each theme. The whole committee consisted of 14 experts, including 6 women and 8 men representing North America, Europe and Asia and Australia. Each subcommittee had access to the camera-ready version of the paper along with the review materials. They were asked to evaluate them and to come up with a consensus selection of the top paper (or 2, if needed). They were asked to justify their choice for the best paper and to briefly describe its strengths. The following list shows the thematic clusters, the areas of candidate papers, and the experts in each area:

1. Semantics

Massimo Poesio

Aline Villavicencio

2. MT and Multilinguality

Matt Post

Lucia Specia

3. Phonology to syntax

Yuji Matsumoto

Yusuke Miyao

Mark Johnson

4. Dialogue

Dilek-Hakkani-Tur

Koichiro Yoshino

5. Generation, Summarization, QA

Xiaodan Zhu

Ed Hovy

6. Beyond the sentence

Lun-Wei Ku

Raquel Fernandez

Kathy McKeown

A small committee consisting of the two PC chairs and the general chair of the conference considered the recommendations from the sub-committees and took the final decision on the awards. The awards included the overall best long and short paper award along with a number of outstanding paper awards.

9. Invited speakers

The conference featured two invited talks, one of which was nominated in the public call:

Simultaneous Translation: Recent Advances and Remaining Challenges
Liang Huang

Principal Scientist and Head of Institute of Deep Learning USA (IDL-US) at Baidu Research and Assistant Professor (on leave) at Oregon State University.

Loquentes Machinae: Technical Approaches, Applications and Ethical Issues of Conversational Systems

Pascale Fung

Professor in the Department of Electronic & Computer Engineering and the Department of Computer Science & Engineering at the Hong Kong University of Science & Technology(HKUST).

10. Issues and recommendations

The timeline is too tight considering the current size of the conference

Even with our time saving measures the conference schedule was still too tight, not only for us PC chairs but also for (S)ACs and reviewers. Interestingly, although ACL has grown significantly over the past 20

years, the schedule remains almost the same as it was back in 1999. In particular, the time between the submission deadline and the notification of acceptance is exactly the same (2 months) as it was in 1999, although the number of submissions has increased tenfold and the size and the complexity of the PC even more. It may be time to adopt the practice of related conferences (e.g., IJCAI, NeurIPS, SIGIR) and extend the schedule to allow for 3-4 months for this process. This could be critical for maintaining the quality of reviewing as the conference grows further.

The conference schedule is also impacted by the schedules for other conferences (e.g. this year NAACL and EMNLP-IJCNLP) and the ACL Guidelines and desire for preprints. We made a concerted effort with other conferences to avoid overlap in the review period (which otherwise shortened the available time for each conference). Overlapping review periods will either result in unhappy authors (when they cannot submit to all conferences) or chaos for PC chairs who struggle to manage multiple submissions and large numbers of withdrawn papers. Also reviewers may be less likely to review for multiple conferences at the same time. Even with no overlap this year, there is still a desire from authors for a longer period in between conferences, so that they can revise and resubmit papers based on the feedback from previous conferences, and rejection from one conference does not leave enough time for people to submit before the anonymity period of the next conference begins. In general, not just the dates of the conferences but also the number and scheduling of deadlines should be given community-wide attention going forward.

Problems with the current double blind policy

The ACL policies on submissions and preprints established in 2018 is an attempt to strike a compromise between the fears of bias towards prestigious authors and institutions, if authors identities are present, and authors' desires to get their work into the community early in pre-prints. While the policy of allowing pre-prints until a "blackout period" one month before the start of the conference is conceptually simple, the policy is very difficult to implement fairly in practice. There were not a huge number of controversial cases, but each one required careful consideration and sometimes more detailed investigation and consultation as to what constitutes acceptable vs unacceptable revealing of identities, including not just preprint servers like arxiv, but also issues like git servers, leaderboards, and other conference and workshop submissions.

We will follow with more detailed recommendations, but perhaps a better way to satisfy both early dissemination of results and anonymity is to have separate tracks - a main one that is more strict and clear about author anonymity, and one that allows authors to reveal identities - to be used in those cases where material from the paper is already in the public domain and anonymization is not really feasible.

Incomplete Reviewer Profiles

Probably the biggest need is for consistent reviewer profiles that can provide essential information about reviewers for the purposes of assigning papers appropriately and avoiding conflict of interest. Many SACs had great difficulty assessing reviewer competence for unknown reviewers, and TPMS profiling was not a complete solution. Many SACs call for bringing back reviewer bids for papers, which would make the assignment process easier, but as mentioned above, not necessarily better, and includes other problems.

We need better handling of Conflicts of Interest

There is a need for far better automatic identification and handling of COIs than that currently offered by the START system. A lot of time gets wasted when papers get wrongly allocated to chairs and reviewers,

and then re-allocated after COIs are detected and reported by these individuals. Also, when individuals end up with papers they have COIs with we are relying on their integrity and honesty when we expect them to report COIs. This opens up the possibility for potential abuse.

There were some recommendations for automatically importing profile and conflict of interest information from 2018 PC chairs of ACL (and related) conferences:
https://docs.google.com/document/d/1rXVO2zWDRT3cDPXnG4v7mFNYFzGl_EKFcDXf_zmt-bQ/edit?usp=sharing

Such suggestions should be developed further and implemented for next year, well in time for ACL 2020.

Problems with START

The START system has been an essential feature in managing the large number of submissions, and for most purposes it worked very well. However, there were some challenging aspects to configure to suit our process. Rich Gerber at SoftConf has been extremely responsive to bug reports, requests for help, and the addition of new features (particularly to support Area Chairs as meta-reviewers who did not see author identity, or the full set of papers in the area). Some necessary functions are not currently supported by START and require computation on spreadsheets, or ad-hoc representations. There is also still a fairly high learning curve for some functions, and non-parallelism in parts of the interface (pages which have almost the same information or links to other pages, but not all). Perhaps the most serious issue was that the notification emails stopped being sent out about ⅔ of the way through, so that despite the PC chairs to notify everyone at the same time, some were delayed by several hours until the problem could be diagnosed and addressed.

ACL 2019 Program Co-Chairs
Anna Korhonen, University of Cambridge
David Traum, University of Southern California