4611 Committee
James Bailey, Ling Chen, Sunil Gupta, Tongliang Liu, Jilles Vreeken, Geoff Webb, Lina Yao

Additional reviewers
Michael Winikoff, Noa Agmon

Conflicts of Interest
Lina Yao, strong, Managed by exclusion from all processes on this submission

Decision
Conference to be ranked as B

Justification
CoNLL was on the review list to ascertain whether it should retain its A ranking. Like many of
the NLP conferences, it had never actually been reviewed.

While it was in the Machine Leaning committee, it was in many ways more suitable for it to be
reviewed by the AI committee, along with a substantial number of other NLP conferences.
Consequently the ML committee deferred mostly to the AI committee, which recommended a rank of B.

In terms of impact CoNLL has an h5 of 43. This is high, but the 2 A* NLP conferences have h5 of
192 and 176. COLING and LREC, both newly ranked B have an h5 of 73 and 61 respectively. So in the
context of NLP it is not so outstanding. The citation centiles are at A in 4602, but this is
weaker than COLING which is above A. All the A* and A NLP conferences are well above the A bar in
4602. NLP gets high citations.

The PC is rather weak. The median h-index of the PC is 7, and only 39 of 164 PC members are above
the established researcher threshold of h-index=12. 58 PC members have h-index in 0-4 range and a
further 52 in range 5-9. Of the 39 established researchers in the PC only 1 published there in 3
years of the last 5, 1 in 2 years, and 9 in 1 year. This is a very weak level of publishing
engagement for the PC.

10 strong people were identified who published 1-7 papers each in the last 5 years. However only
4 of the 10 attend often. Author strength in the 4602 context is at the A level, but COLING (B)
is higher, and A and A* conferences are much higher. Of the area leaders identified 1 published
there in 4 years, 1 in 3 years, 4 in 2 years and 9 in 1 year. The area leader list is not
reproducible as the submitters removed from the list people publishing substantially in
application areas such as vision, physics, speech processing and data mining. This seems
unreasonable and calls into question the validity of the list as area leaders. The h-index of
people in the list ranges from 51 to 152.

CoNLL was compared in detail with A*, A and B conferences in NLP and it was agreed that it was a
clear step down from the A conferences on all factors. The author strength and citation
statistics looked more favorable in the AI context than in the Machine Learning context, so it
was not disadvantaged by considering it in that context.