ACT: An Annotation Platform for Citation Typing at Scale

David Pride
KMI, The Open University. UK.
orcid.org/0000-0002-7162-7252
david.pride@open.ac.uk

Jozef Harag
KMi, The Open University. U.K.
orcid.org/0000-0002-5445-0141
jozef.harag@open.ac.uk

Petr Knoth
KMi, The Open University U.K.
orcid.org/0000-0003-1161-7359
petr.knoth@open.ac.uk

ABSTRACT

In this paper we introduce the Academic Citation Typing (ACT) Platform, a highly scalable online tool that takes as its input any full text research paper and which then enables rapid annotation and classification of in-text citations according to purpose and influence. In contrast to previous work, we employfi rst authors as annotators. Our evaluation shows that these authors are able to quickly classify the citations within their own papers. Over 200 authors have thus far annotated their papers using the ACT platform. This approach has already enabled the collection of the largest dataset of citations annotated according to their purposes and influence on the citing paper. Furthermore, this process is ongoing and the dataset will continue to expand following this initial phase.

CCS CONCEPTS

• **Information systems → Data mining**; **Digital libraries and archives**;

KEYWORDS

Citation Typing, Citation classification, Data Mining, Open Access, Scholarly Data, Research Evaluation

1 INTRODUCTION

Prior studies that have introduced and built on the notion of classifying citations according to purpose have limitations. The annotated datasets used in, or produced by, these studies are relatively small. Building high-performance machine learning models to automatically and accurately identify citation type requires significantly larger training datasets.

We therefore propose a new method of annotating citations. Our methodology differs from previous works in employing authors as annotators as opposed to independent annotators. We have developed a new platform we call ACT which makes it possible to recruit large numbers of annotators and enables them to classify citations rapidly and authoritatively. ACT displays the full text of the research paper alongside a point-and-click style classification interface and automatically highlights the in-text citation marker for each citation as a visual prompt for the annotator, hence making the annotation process far less laborious.

2 RELATED WORK

In 2006, Teufel & Siddharthan [5] introduced the largest dataset of 2, 829 citations, annotated according to 12 types. Further recent work in the domain has been that of Shoton & Peroni[4] who introduced CiTo, the citation ontology which consists of 107fi ne grained reasons for citation. In 2016, Jurgens & Jurafsky [1] introduced a dataset of 1, 969 citations and simplified the 12 typesfi rst suggested by [5] into seven types. Most recently, Valenzuela et al. [6] introduced a method for identifying a cited paper as either *incidental* or *influential* to the citing paper and also released a dataset of 465 citing-cited paper 'pairs'.

Our work builds primarily on the studies of [1], [5] and [6] as we choose to collect annotations according to both purpose and influence. To annotate citations according purpose, we keep our classification schema compatible with those of [1] and [5], but add an additional layer to the compare/contrast category; show similarities, show differences or show disagreement (Table 1).

Citation purpose	Description
Background	The paper you are citing provides relevant information or is part of the body of literature in this domain.
Uses	Your paper uses the methodology or tools created by the paper you are citing.
Compare / contrast - similarities - differences - disagreement	Your paper expresses similarities or differences to, or disagrees with, the paper you are citing.
Motivation	Your paper is directly motivated by the paper you are citing.
Extension	Your paper extends the data, methods etc. of the paper you are citing.
Future	The paper you are citing is potential avenue for future work.

3 METHODOLOGY

For thefi rst phase, we collated a multidisciplinary dataset of 4,274 full text papers from which we successfully extractedfi rst author names, emails and approximately 93k citing sentences using Grobid [3]. The dataset was then uploaded to the ACT platform which automatically generates a unique URL token for each paper. Using an email delivery service we then invited an initial sample of authors to annotate their own paper. Each author was sent a personal link to the ACT platform with their paper displayed (Figure 1).

4 RESULTS

The ACT platform was, prior to data collection, tested for user experience and reliability with 6 internal evaluators. An email service was then used to send customised invitations, each with a unique URL link to the specific ACT survey for that author. This was send to an initial sample of 4,274 authors. One limitation of this approach

978-1-7281-1547-4/19/$31.00 ©2019 IEEE
DOI 10.1109/JCDL.2019.00055

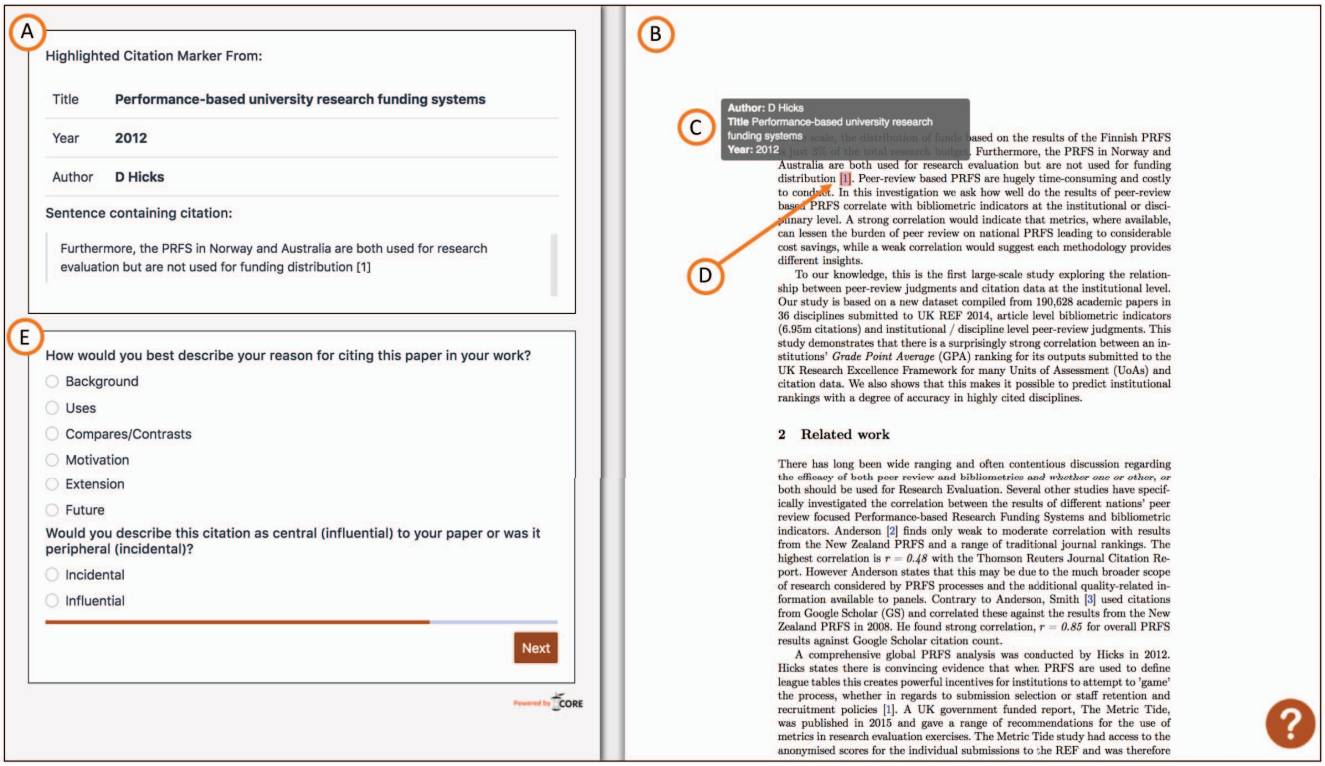

Figure 1: The ACT Platform in use. A. Details of current citation for annotation. B. Author's PDF manuscript. C. Hover-over pop-out for current citation details. D. In-text highlighted citation marker. E. Annotation section. (A live demo of the ACT Platform is available at: https://youtu.be/qQz_gB0Yjx4)

is that an academic may change institution and therefore email address may no longer be current. Alternately, the formatting of the email address may not be exactly correct. Overall 2,794 emails were delivered and 1,254 of these were opened. 224 authors then visited the ACT platform via the link provided and 212 of these completed the annotation process. Each author annotated an average of 24.5 citations and our total dataset thus far contains 5215 annotated citations. This is already the largest collection of citations annotated according to both type and influence, considerably larger than that of [2], [5] and [6]. Moreover, we can continue to build on this initial dataset as our process can be repeated with any set of full text research papers.

Additionally, the platform also records the time taken by each author to complete the annotation process, which was an average of nine minutes, around 22s per citation. Our observation, based on feedback from the evaluators and authors who completed the annotation process is thatfi rst authors, in almost all cases, remember their reasons for citing a particular paper without prompting and can therefore complete the process quickly and with confidence.

5 CONCLUSION

We have developed a fully scalable method for crowdsourcing large numbers of citations, annotated according to purpose and influence, rapidly and accurately. While our classification scheme is compatible with [5] and [1], we add similarities, differences and disagreement sub-classes into the compare / contrast class, as these are important for applications in Scientometrics. Early testing has shown that with the ACT platform we have the capacity to create a dataset of annotated citations on a previously unseen scale. These data can then be used in the future to train models for identifying citation purpose and influence with higher accuracy than previously possible.

REFERENCES

[1] David Jurgens, Srijan Kumar, Raine Hoover, Dan McFarland, and Dan Jurafsky. 2016. Citation classification for behavioral analysis of a scientificfield. *arXiv preprint arXiv:1609.00435* (2016).

[2] David Jurgens, Srijan Kumar, Raine Hoover, Dan McFarland, and Dan Jurafsky. 2018. Measuring the Evolution of a Scientific Field through Citation Frames. *Transactions of the Association for Computational Linguistics* 6 (2018), 391–406.

[3] Patrice Lopez. 2009. GROBID: Combining automatic bibliographic data recognition and term extraction for scholarship publications. In *International Conference on Theory and Practice of Digital Libraries*. Springer, 473–474.

[4] David Shotton. 2010. CiTO, the citation typing ontology. In *Journal of biomedical semantics*, Vol. 1. BioMed Central, S6.

[5] Simone Teufel, Advaith Siddharthan, and Dan Tidhar. 2006. Automatic classification of citation function. In *Proceedings of the 2006 conference on empirical methods in natural language processing*. Association for Computational Linguistics, 103–110.

[6] Marco Valenzuela, Vu Ha, and Oren Etzioni. 2015. Identifying Meaningful Citations. In *AAAI Workshops*. http://www.aaai.org/ocs/index.php/WS/AAAIW15/paper/view/10185