Accuracy of affiliation information in Microsoft Academic Implications for institutional level research evaluation

Bijan Ranjbar-Sahraei⁽¹⁾, Nees Jan van Eck⁽¹⁾ and Rutger de Jong⁽²⁾

(1) Centre for Science and Technology Studies, Leiden University
(2) Science Library, Leiden University

{b.ranjbarsahraei,ecknjpvan}@cwts.leidenuniv.nl; r.m.de.jong@library.leidenuniv.nl

Summary

We have studied the accuracy of affiliation information in Microsoft Academic (MA) platform. To conduct this study, we have considered the full set of publications assigned to Leiden University (LU) as provided by two different data sources: MA and Web of Science (WoS). The results of this study suggest that a considerable number of publications in MA have missing or wrong affiliation information.

Data

- 1. All publications of LU available in MA database 130K
- 2. All publications of LU available in WoS database 110K

Matching Method

- 1. Matching publications based on title and year of publication
- 2. Matching based on identical meta date: volume number, first page number, last name of first author, and publication year

Results (part 1)

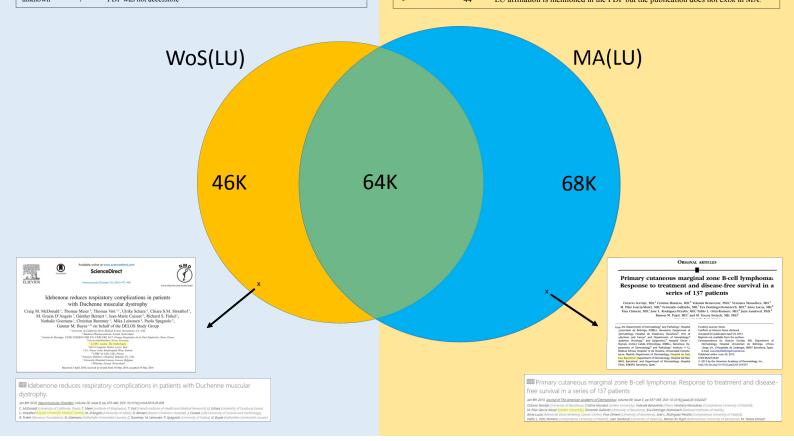
Manual validation of a sample of 100 WoS(LU) publications without a match in MA(LU).

MA affiliation	No. of	Observation
information	pub.	
incorrect	29	LU affiliation is not mentioned in the PDF.
correct	41	LU affiliation is mentioned in the PDF and the publication does not exist in WoS.
	23	LU affiliation is mentioned in the PDF and the publication exists in WoS(LU), but
		the match is not detected.
unknown	7	PDF was not accessible

Results (part 2)

Manual validation of a sample of 100 MA(LU) publications without a match in WoS(LU).

			Observation
ı	information	pub.	
	incorrect	40	LU affiliation is mentioned in the PDF and the publication exists in MA, but LU
l			affiliation is not provided by MA.
	correct	16	LU affiliation is mentioned in the PDF and the publication exists in MA(LU) with
l			correct affiliation, but the match is not detected.
		44	LLI affiliation is mentioned in the PDF but the publication does not exist in MA



Conclusion

We manually checked 200 publications, out of which 69 publications had missing or wrong affiliation information in MA. This is an indication that a considerable number of publications in MA has unreliable affiliation information. Therefore, institutional level research evaluation based on MA data should be performed only when proper attention is paid to this limitation.



