How reliable is Google Scholar for Bibliometrics?

A summary of the literature


- Compared metrics from Israeli authors from HighlyCited.com and Israeli Nobel Prize Winners in 2004/5 (not in the latter)
- Excluded authors difficult to disambiguate or had no pubs in date range
- Analysis from 1996-2006
- found significant differences when compared to Web of Science and Scopus
- whether metrics are higher depends on the discipline - maths/computing are higher, physics lower
- Need for data cleansing in GS – errors in Year & Inexact matching of authors (“J Smith” finds SJ Smith)


* data incomplete or work (unsurprising as the data is not validated by humans)

- F PASSWORD = FORGOTTEN PASSWORD
- M Profile = My Profile
- Errors in author names e.g.Julie M Still > Julie M
- Duplicates – from A&I Databases (Eric etc.) can cause errors in citation counts
- Some specific errors have been resolved (i.e. Google read the article!)


- Academics at University of Udine
- Top 20 computer science journals + top 20 in the ‘theory and methods” category
- h-indexes about 3x higher
• Best correlation – citation based metrics (cites, cites/yr, etc.) + g

• Moderate – paper based metrics (e.g. cites/paper), h

• Metrics are higher because of the discipline's use of high-quality, very competitive conferences as a mode of primary publication

• Correlation between ranking of academics so highly ranked academics in Google Scholar are generally also highly ranked in Web of Science even if the numerical value of the metrics differs


• Highly downloaded articles from Elsevier social sciences journals

• Articles have more citations in GS – increased coverage of books, conferences & preprints & some duplication

• Monograph publication, both sole authors and chapter in edited books


• Random sample of academic neurosurgeons

• Significant correlation with academic rank

• H-index tends to be higher in GS, with small differences at low academic rank & increasing differences at higher rankings


• Research outputs from 3 UK Business schools

• Broad coverage picks up fringes of subject

• Higher metrics in GS

• Monograph coverage

• But lack of transparency as to what is actually indexed and what may be missing