Spreadsheet Quality Assurance: A Literature Review

Pak-Lok POON, Man Fai LAU, Yuen Tak YU, Sau-Fun TANG

Frontiers of Computer Science, DOI: 10.1007/s11704-023-2384-6

Problem & Ideas

- Problem of spreadsheet quality assurance (QA):
  - There does not exist a holistic & systematic review of various spreadsheet QA techniques over the entire spreadsheet life cycle & from different disciplines/perspectives
- Ideas:
  - A software engineering-based spreadsheet life-cycle framework is firstly developed
  - A comprehensive investigation of various spreadsheet QA techniques at different spreadsheet life-cycle stages is performed

Technical-oriented IT		Business-oriented IS		Dual-focused (IT & IS)		Management Science (MS) /operational research (OR)	
Target journals	No. of papers	Target journals	No. of papers	Target journals	No. of papers	Target journals	No. of papers
ASE	4	CAIS	3	DSS	6	JORS	5
CSUR	1	EJIS	0	IPM	0	OMEGA	3
TOCHI	1	IM	2	IWC	1		
TOIS	2	IO	2				
TOSEM	2	ISJ	0				
CACM	2	ISR	0				
COMP	0	ITP	1				
SW	2	JAIS	2				
TR	1	JIT	0				
TSC	0	JMIS	4				
TSE	6	JOEUC	16				
IST	2	JSIS	0				
IJHCS	4	MISQ	0				
JFP	2	MISQ-E	0				
JCL	6						
JSS	6						
Total count:	41	**Total count:**	30	**Total count:**	7	**Total count:**	8

Target conferences	No. of papers
OOPSLA (or *PACMPL* from 2017 onwards)	2
ESEC/FSE	3
ISSTA	2
ICASE	2
VL/HCC	18
ICSE	9
HICSS	0
ICIS	0
Total count:	36

Left: Relevant journal papers related to spreadsheet QA; *Right:* Relevant conference papers related to spreadsheet QA.

Main Contributions

- An extensive literature review on spreadsheet QA over a 35.5-year period for target journals and a 10.5-year period for target conferences
- A holistic and comprehensive review of spreadsheet QA issues over the entire spreadsheet life cycle so that the quality of spreadsheets and the data generated from them can be improved
- Identification of existing research gaps in the spreadsheet QA area, shedding light on future research directions

	Values of *P*	Values of *T*
Spreadsheet life cycle	6	13
Problem and scope identification	1	2
Specification, modeling, and design	26	37
Implementation	13	17
Testing and debugging	81	107
Usage and maintenance	6	7

The number of “primary” papers (*P*) and the number of both “primary” & “secondary” papers (*T*) for the spreadsheet life cycle & its major stages.