**Original Manuscript ID:** Access-2022-20656

## Original Article Title: “The state of big data reference architectures: a systematic literature review”

**To:** IEEE Access Editor

**Re:** Response to reviewers

Dear Editor,

Thank you for allowing a resubmission of our manuscript, with an opportunity to address the reviewers’ comments.

We are uploading (a) our point-by-point response to the comments (below) (response to reviewers), (b) an updated manuscript with yellow highlighting indicating changes (*Supplementary Material for Review*), and (c) a clean updated manuscript without highlights (*Main Manuscript).*

Best regards,

Pouya Ataei et al.

**Reviewer#1, Concern # 1:** The presented work is timely, and the authors made great efforts in this paper. However, a few big data architectures can be surveyed and added to mention a few:  
https://ieeexplore.ieee.org/document/8417407  
https://ieeexplore.ieee.org/abstract/document/9393907

**Author response:** While we appreciate the feedback given, the mentioned studies by the reviewer do not really fit in the scope of our work. This is due to the fact that these studies are revolving around actual concrete architectures in a very specific domain and do not elaborate much of architectural constructs, while our study is focused on reference architectures and high-level architectural constructs. These concrete architectures also seem to be driven by traditional Hadoop based ecosystems without any attention to data engineering cross-cutting concerns. This is well covered in our review by other reference architectures.

In addition, evaluation of these studies does not seem to be very relevant to industrial practices. It is worth mentioning that these studies and dozen others have not been overlooked in the process of our research and we’ve been well-aware of them. We’ve been researching on big data reference architectures for the past 4 years; thus, we have evaluated an analyzed many recent big data architectures. Given all, we do not think that these studies can strengthen our work or add value to it.

**Author action:** Despite the rationale above, we still did another search and looked at available big data architecture and reference architectures in both academia and industry. We could not find any new study that we have not covered in our systematic literature review.



**Reviewer#1, Concern # 2:** Also, the citations and reference list should be revised, for example reference [19], [24], [25], [4] are off.

**Author response:** This was a good catch, and we realized some BibTex references that we downloaded from Google Scholar did not have the best qualities.

**Author action:** We fixed the mentioned references; we reviewed the list of references once more for any errors.



**Reviewer#2, Concern # 1:** Could the authors define acronyms of RA & BD under Abstract for better clarity? Thanks!

**Author response:** While some journals are against defining acronyms in abstract, we respect reviewer’s opinion on this matter.

**Author action:** RA and BD acronym has been added to the abstract.



***Note:*** *References suggested by reviewers should only be added if it is relevant to the article and makes it more complete. Excessive cases of recommending non-relevant articles should be reported to ieeeaccesseic@ieee.org*