This paper makes a valuable contribution to the methods needed to deal with the 'big data' of variable quality that citizen science observations are rapidly accumulating. The paper and novel method it describes examine aspects of the validity of the accompanying 'metadata' for online images (including crowd-sourced identification at various taxonomic levels). It checks whether each image uses a standard (selected) taxonomy. It enables automated contrasts of the stated geographical location of images with the existing documented range of the taxon. It automatically assesses image quality as another potential filter on images of a taxon.

This method will enable more rapid filtering of large databases such as iNaturalist for taxa of interest, retrieving taxa of interest and removing the less reliable records. It will also enable rapid detection of range changes, such as through introductions. It helps researchers and others make the most of the efforts of numerous citizen science image providers and assessors.

The pipeline can be used to check the consistency of taxonomic nomenclature between image metadata (such as a crowd-sourced identification) and the preferred taxonomic list for a set of images (in this case spiders on iNaturalist). The taxonomic accuracy of metadata for images is compared with an 'official' database, in this case the World Spider Catalogue. However, for many taxa there are disputed taxonomies at several levels including species, genus and family. Some additional comments on how to include capacity for expansion of the pipeline to include alternative taxonomic lists would be helpful.

Geo-location metadata of images can be checked for major errors and were quite precise for spiders in this case study. The method permits potentially erroneous species records to be flagged for expert attention if they are in a surprising location. The pipeline can help filter images automatically by computer-assessed image quality, complementing scores given by crowd-sourced identification, and thus helping researchers select the most reliable.

The author sees data-volume as the biggest barrier to crowd-sourced data. Perhaps this is true for taxa that can be identified from images, but such taxa are a tiny fraction of all recorded taxa. For the Araneae, used in this paper as a case-study, there are only a few percent of species, even in a fauna as small and relatively well-known as that of Britain, for which I would accept photographic evidence of a species' identity. Most spider species, even as adults, require microscopic identification of a dead specimen. The value will be higher for genus and family level, or for screening for those taxa where photographs can permit reliable identification of some individuals. The method is transferrable and for other major taxonomic groups there may be fewer, or more, limitations.

The methods are not easy to follow for a non-programmer, and the figure captions (such as Figure 1) need to be simpler and more self-explanatory. I am taking the programming methods on trust since I do not have the specialist knowledge to assess them, but other reviews can assess validity. My focus is on the scientific applications, assuming the methods to be sound.

Based on the novel analysis and case-study of spiders, the author makes helpful suggestions on limitations, potential improvements and applications of the method on databases more generally. There are potential applications of this tool, so long as one remains aware of the taxonomic limitations. I think it could more rapidly alert people to the spread of harmful invasive species such as the false widow *Steatoda nobilis*, by flagging images for expert validation - perhaps including those from newspaper archives. It could detect other range-expansions of conservation or ecological interest. From sets of images before and after a disturbance it could perhaps examine shifts in the balance of taxa - for example declines in orb or scaffold web building spider families when vegetation structure is simplified. Other applications will doubtless be discovered. There are complimentary proposals for those designing citizen science studies, such as what data are most valuable or most essential to keep records consistent and accessible.

A few minor issues in presentation: I suggest the Abstract be simpler and more like the end of the Discussion. Some of the punctuation could be improved, most importantly in the Abstract, and there are a few typos (eg large should be larger in the geo-tagging methods section; "was" should be were in Figure 3 caption). I suggest clarification in the Abstract: "genus level and the highest image quality according to the BRISQUE scores" should presumably be: genus level and had the highest image quality according to the BRISQUE scores. Perhaps one instance of "observations taxonomic" should be observations' taxonomic in the discussion. For me, the PDF preview and download had a formatting problem which cut off parts of some figure captions. Please describe what a 'pipeline' is for the novice!