

Many thanks for these valuable corrections. All possible changes in the data and the manuscript are made and accompanied by replies **painted by blue color**. Also, below we cite the disclaimer from the head of the each technical evaluation report:

- > Please note that Pensoft does not check the details of the content of a dataset, for example,
 - > whether the correct author is given for a scientific name or whether the correct latitude /
 - > longitude is given for a locality.
-

Part 1 - literature

1. According to the Darwin Core definition the **bibliographicCitation** is "A bibliographic reference for the resource.", so it should be unique for a particular occurrence or event within the dataset. While, for example, "Huber B.A. 2011. Revision and cladistic analysis of Pholcus and closely related taxa (Araneae, Pholcidae) // Bonner Zoologische Monographien. Nr.58. P.1–509" mentioned for 26 different occurrences.

This is not true. DwC doesn't say **bibliographicCitation** values must be unique for a particular occurrence or event within the dataset. We have full confidence in the technical evaluator's high level of expertise and we respect the opinion expressed in the recommendation; however, we are not obligated to follow it.

Furthermore, if you provide a reference to some literature source in this field, you must be sure that almost all information from this source is represented by a particular occurrence or event in the dataset, and vice versa, there are no more occurrences or events described in this source besides the considered record. In other words, the literature source that you provide as a bibliographic citation and the related event or occurrence must be (near) equivalent. I am not sure that most of the cited literature sources in the dataset contain details regarding only one occurrence.

This is not true. DwC doesn't say all occurrences should be extracted. Scientific ethics ask to cite source even if just a single number is taken of this. We have full confidence in the technical evaluator's high level of expertise and we respect the opinion expressed in the recommendation; however, we are not obligated to follow it.

If you can prove that some bibliographic references contain description only of the particular occurrence or event, you can keep them in this field; otherwise, please move them to the **associatedReferences** field. Briefly, any bibliographic citation in the dataset must unambiguously relates to a particular record, just as you can find exactly one article you need by bibliographic citation.

We respect the technical evaluator's opinion. However, we respect original literature authors' authorship not less. All these reputable researches, books, articles and journal, provided these data at firsts, deserve to be cited properly. The associatedReferences field is not proper enough for this because can contain any literature sources concerning the record not specifying these recommended to cite. Our implemented decision to provide "the preferred way to cite the resource itself" (<https://dwc.tdwg.org/terms/#dcterms:bibliographicCitation>) has no any conflict with this DwC term. Yes, the specification has two opposite words in the same sentence ("to cite the resource" and "to cite this record") what can confuse interpretation of the term and it is unclear why technical evaluator assert his way as an only correct. The technical

evaluator' suggestion does not contradict DwC specification. But our decision doesn't contradict DwC specification as well. We were fully prepared to discuss this matter and were deeply concerned that all of our arguments on this point presented at earlier stages were ignored.

2. The **informationWithheld** field in the dataset means that some details from the dataset were not provided in the published version of the dataset. It implies that withheld information had been digitized, however. Please provide what information and for what reasons were withheld rather than reference to sources, from which data users must extract information themselves.

All this additional data concern exact occurrences so can be located in occurrenceRemarks field. We have transferred it there.

3. It is not a good idea to remove useful details from the dataset. Please get the **acceptedNameUsage** field back, and add the **taxonomicStatus** field that clarify relation the **scientificName** field with the first one.

Previous note from technical elevation report:

If you want to keep initial names according to the literature, you can use the verbatimIdentification field: <https://dwc.tdwg.org/terms/#dwc:verbatimIdentification>

We did exactly as you said at the previous stage of the technical evaluation. Now initial names according to the literature are provided in verbatimIdentification field meanwhile scientificName field contains the most relevant scientific names according the World Spider Catalog (2025) so **acceptedNameUsage** and **taxonomicStatus** fields are not required anymore.

4. Point regarding the **references** field was my mistake. Please get it back to the dataset. It is got back.

5. The **occurrenceRemarks** field contains details regarding collection code and catalog number. Please place these details to the correspondent fields.

All this non-obligatory information is removed. We kindly remind that it is author decision how detailed should be openly published data (while these follows Darwin Core specification and common sense). And, as it is written at the head of the report,

>Please note that Pensoft does not check the details of the content of a dataset, for example,
>whether the correct author is given for a scientific name or whether the correct latitude /
>longitude is given for a locality.

6. Please remove excessive quotation:
"Almaty Area, ""Shazyn riverside"""
"Долина Б. Алматинки, ""Приютская колония"""
"Алма-Атинский у., долина Б. Алматинки, ""Приютская колония"""
"Б. Алматинка, ""Приютская колония"" and
others

All this quotation is removed. We can't find others examples of such quotation aside of examples listed.

7. Is there some sence in the order of the identifiers?

- 79 Logunov D.V. | Rakov S.Yu.
30 Rakov S.Yu. | Logunov D.V.

The order of collectors makes sense of course as order of authors in articles has. However, we can't find examples you listed. Only "Logunov D.V. | Rakov S.Yu." is presented in the dataset.

Part 2 - collection

8. The **verbatimIdentification** provides mostly other values than in previous version.
What caused this inconsistency? Make sure it was not resulted from improper table sorting.

Sorry, but the second dataset doesn't have the **verbatimIdentification** field, and never had. Collection data have been accompanied by the most relevant scientific names always. You can easily verify it looking at the versions 'history'. If you mean the first (literature) dataset, it is checked and no mistakes were found in this field.

9. (partly repeated) Please unify spelling of the **taxonRemarks** field:
5 sp.n.: paratypes
2 sp.n.: paratypes
1 syntype
2 syntypes

Sorry, but the second dataset doesn't have these values in the **taxonRemarks** field. If you mean the first (literature) dataset, we don't see any difference between:

- 5 sp.n.: paratypes
2 sp.n.: paratypes

Also, DarwinCore doesn't say the values of this field should be unified, like it is obligatory for the fields limited by set of 'dictionary' values. So, "syntype" and "syntypes" delimits single and multiple specimens and are valid values both.

10. The **eventRemarks** field contains details that could be formalized as **vitality**. All this non-obligatory information is removed. We kindly remind that it is author decision how detailed should be openly published data (while these follows Darwin Core specification and common sense).

11. There are inconsistencies between the ***eventRemarks*** and ***verbatimEventDate*** fields.

03-05.06.2024	05.09.2024 - found dead.
05.06.2024	04.11.2024 - found dead.
05.06.2024	Found dead 02.10.2024 and 05.02.2025. and others

Indeed, sometimes collector's labels can contain mistakes. However, we suppose over-interpretation is bad idea and better not to replace exact label data written on the tubes by figured out values. For this reason, all this non-obligatory information is removed. We kindly remind that it is author decision how detailed should be openly published data (while these follows Darwin Core specification and common sense). And, as it is written at the head of the report,

>Please note that Pensoft does not check the details of the content of a dataset, for example, >whether the correct author is given for a scientific name or whether the correct latitude / >longitude is given for a locality.

12. There are possible inconsistencies between the coordinates and ***locationRemarks***

2	46.22432,75.01124
4	46.22432,75.01124 KATO: 193600000
1	45.9475,74.97284
1	45.9475,74.97284 KATO: 193657100
6	43.97606,77.41633
2	43.97606,77.41633 KATO: 191037100
1	43.2452,76.9328
1	43.2452,76.9328 KATO: 750000000
2	43.685,76.589
3	43.685,76.589 KATO: 196800000
1	43.2451,76.9326
10	43.2451,76.9326 KATO: 750000000
1	44.11592,78.70265
7	44.11592,78.70265 KATO: 334035100
1	43.542,79.436
1	43.542,79.436 KATO: 196630100
5	43.2142,76.9226
1	43.2142,76.9226 KATO: 750000000
7	43.2664,76.8178

26 43.2664,76.8178 KATO: 750000000
1 43.9782,77.4147
2 43.9782,77.4147 KATO: 191037100
4 43.95361,77.04639
13 43.95361,77.04639 KATO: 191000000
5 43.239,76.866
24 43.239,76.866 KATO: 750000000
5 43.224,76.921
10 43.224,76.921 KATO: 750000000
4 43.224,76.921 N part
15 43.219,76.923 KATO: 750000000
3 43.219,76.923 S part
1 43.2349,76.9203
2 43.2349,76.9203 KATO: 750000000
1 43.2259,76.9224
1 43.2259,76.9224 KATO: 750000000

All this non-obligatory information about location is removed, geographical coordinates are remain the same. We kindly remind that it is author decision how detailed should be openly published data (while these follows Darwin Core specification and common sense).