

Data audit for technical evaluation of **Spider (Arachnida, Araneae) fauna of the lowland part of the Balkhash-Alakol basin (SE Kazakhstan): an integrated assessment**

Resource links: <https://www.gbif.org/dataset/92cc4b3a-97f2-4aa3-b260-883ba061c1a9>
<https://www.gbif.org/dataset/791d91fe-ce27-4044-8a3d-a2a8b203f979>

the Darwin Core archives **dwca-almaty_literature-v1.3.zip** and **dwca-almaty_collection-v1.2.zip** were downloaded on 2025-10-12 from:

http://gbif.ru:8080/ipt/archive.do?r=almaty_literature
http://gbif.ru:8080/ipt/archive.do?r=almaty_collection

Maxim Shashkov (m.shashkov@pensoft.net)

About this evaluation

Pensoft does a technical evaluation of the dataset (or datasets) referred to in the data paper. If the dataset passes or has only minor problems, the data paper manuscript is referred to reviewers. If the dataset has major problems, a review of the paper is postponed until the dataset has been corrected.

To see what features of a dataset are checked in a technical evaluation, please go to <https://zookeys.pensoft.net/about#Data Quality Checklist Recommendations>

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.

Recommendation: the datasets were improved but still contain a row of issues that should be addressed before the datapaper is accepted for review.

Part 1 - literature

1. (partly repeated) The ***bibliographicCitation*** field is intended to provide a reference for the resource itself rather than the source of initial details:

<https://dwc.tdwg.org/terms/#dcterms:bibliographicCitation>

For literature sources, it is more appropriate to use ***associatedReferences***:

<https://dwc.tdwg.org/terms/#dwc:associatedReferences>

If you duplicated the ***bibliographicCitation*** field as ***associatedReferences***, please remove the first one.

The example you provided in the response to the initial technical evaluation report is regarding the event, not a particular occurrence.

2. There is no "references" field in the Darwin Core standard. You can use the ***associatedReferences*** for specifying URLs as well.
3. (partly repeated) Please explain more clearly in the ***informationWithheld*** field, that for more details data user could refer to the source

4. Please verify the field lists in the **Data resources** section of the manuscript thoroughly. All fields from the datasets should be listed in the manuscript, and vice versa - all fields mentioned in the manuscript must appear in the datasets.
5. The **eventDate** and **verbatimEventDate** fields are inconsistent:
1993-06-30/08-12 1993-6-30/1995-8-12
1993-06-30/08-12 1993-6-30/1995-8-12
1993-06-13/06-14 1993-6-13/1996-6-14
 and others.
Changes in the year were not taken into account.
6. If the month does not change, it is not necessary to specify it twice:
1919-06-01/06-29
1920-06-01/06-10
1993-05-28/05-15
1993-06-13/06-14
7. Is there any difference between the values in the **habitat** field:
 building
 buildings
8. Please round geographic coordinates up to 5 decimal places: 43.881100000000004
9. Please provide full scientific names with authorship and year:
 Agalena taurica
 Agroeca thorelli
 Dictyna latens
 Euophrys thorelli
 Larinioides folium
 and others
If you want to keep initial names according to the literature, you can use the
verbatimIdentification field: <https://dwc.tdwg.org/terms/#dwc:verbatimIdentification>
10. If you use both **scientificName** and **acceptedNameUsage** fields, the **taxonomicStatus**
field would be useful as well: <https://dwc.tdwg.org/terms/#dwc:taxonomicStatus>
11. Please unify the spelling of the scientificName field:
 Pardosa turcestanica
 Pardosa turkestanica (Roewer, 1951)
12. (partly repeated) Please unify spelling of the taxonRemarks field:
 29 sp.n. holotype
 4 sp.n.: holotype

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

part2

13. There are multiple inconsistencies between geographic coordinates and localities:

- | | | |
|---|-------------------|---|
| 4 | 46.22432,75.01124 | |
| 2 | 46.22432,75.01124 | Balkhash Lake, S coast |
| 1 | 45.9475,74.97284 | 20 km from Karaoi village |
| 1 | 45.9475,74.97284 | NW part of Saryesik-Atyrau desert, 20 km NE from Karaoi village |
| 2 | 43.97606,77.41633 | near Shengeldy village |
| 6 | 43.97606,77.41633 | Balkhash Lake |
| 2 | 44.91361,76.51472 | |
| 1 | 44.91361,76.51472 | 22 km NE from Bakanas village |
| 1 | 43.2452,76.9328 | |
| 1 | 43.2452,76.9328 | Shevchenko str., 89b |
| 2 | 43.685,76.589 | |
| 3 | 43.685,76.589 | Sorbulak Lake |
| 5 | 44.11592,78.70265 | |
| 3 | 44.11592,78.70265 | Altyn-Emel National Park, Shygan cordon |
| 1 | 43.62823,76.74617 | |
| 1 | 43.62823,76.74617 | Sorbulak Lake |
| 1 | 43.542,79.436 | |
| 1 | 43.542,79.436 | Chundzha village |
| 1 | 43.2142,76.9226 | |
| 5 | 43.2142,76.9226 | Al-Farabi ave. |
| 1 | 43.846,79.419 | |
| 1 | 43.846,79.419 | Chundzha, wooden lake |
| 4 | 44.03047,79.5535 | |
| 2 | 44.03047,79.5535 | 2.5 km E from Aidarly village |
| 1 | 43.87487,78.19593 | |
| 1 | 43.87487,78.19593 | Altyn-Emel National Park, Zhantogai |
| 1 | 43.9782,77.4147 | |
| 1 | 43.9782,77.4147 | near Shengeldy village |
| 1 | 43.9782,77.4147 | Kapshagay reservoir N coast, near Shengeldy village |

7	43.95361,77.04639	
10	43.95361,77.04639	Ili river bank, abandoned station of the Mining Institute
24	43.239,76.866	
4	43.239,76.866	Sairan reservoir
1	43.239,76.866	Sairan reservoir, W coast
1	44.11639,78.70111	
1	44.11639,78.70111	Altyn-Emel National Park
10	43.224,76.921	
9	43.224,76.921	KazNU campus
15	43.219,76.923	
3	43.219,76.923	KazNU campus
1	43.91278,75.50111	Tamgaly
1	43.91278,75.50111	Tamgaly village
7	43.9516,77.0603	Ili river bank
11	43.9516,77.0603	Ili river bank, abandoned station of the Mining Institute

14. There are inconsistencies between the **fieldNumber** and **eventDate** fields: No. of records | fieldNumber | eventDate

1 | WP37 | 2022-10-14

1 | WP37 | 2023-04-05

1 | WP247 | 2023-09-28

1 | WP247 | 2023-09-29

3 | WP262 | 2024-05-08

1 | WP262 | 2024-05-09

1 | WP266 | 2024-05-27

2 | WP266 | 2024-05-18

1 | 43 | 2021

1 | 43 | 2020-07-27

1 | 48 | 2020-07-07

1 | 48 | 2020-07-27

1 | 609 | 2021-08-15

1 | 609 | 2021-08-19

1 | 627 | 2021-12-15

1 | 627 | 2021-08-19

15. It can be advisable to add the **collectionCode** field, as your collection was registered in the GRSciColl registry:

<https://scientific-collections.gbif.org/institution/7e82dc97-c81e-4361-845f-4338170452b2/collections>