

Technological Implications of an Open Biodiversity Knowledge Management System (OBKMS)

Viktor Senderov

PhD Candidate at Pensoft Publishers & Bulgarian Academy of Sciences (BAS)

Affiliation: Pensoft Publishers

Supervisor: Prof. Lyubomir Penev

14.10.2016

Main objectives of the project

PhD project plan published as part of an **open thesis**.

Objectives:

- Models and workflows for import, peer-review, and publishing of **structured biodiversity data**
- A general **semantic model** of a biodiversity publication
- A **Biodiversity Knowledge Graph**
- Competency questions and **use-cases**




Taxon treatment ▼

Heser stoevi Deltshev, sp. n.

- ZooBank urn:lsid:zoobank.org:act:E4D7D5A0-D649-4F5E-9360-D0488D73EEE8

Materials

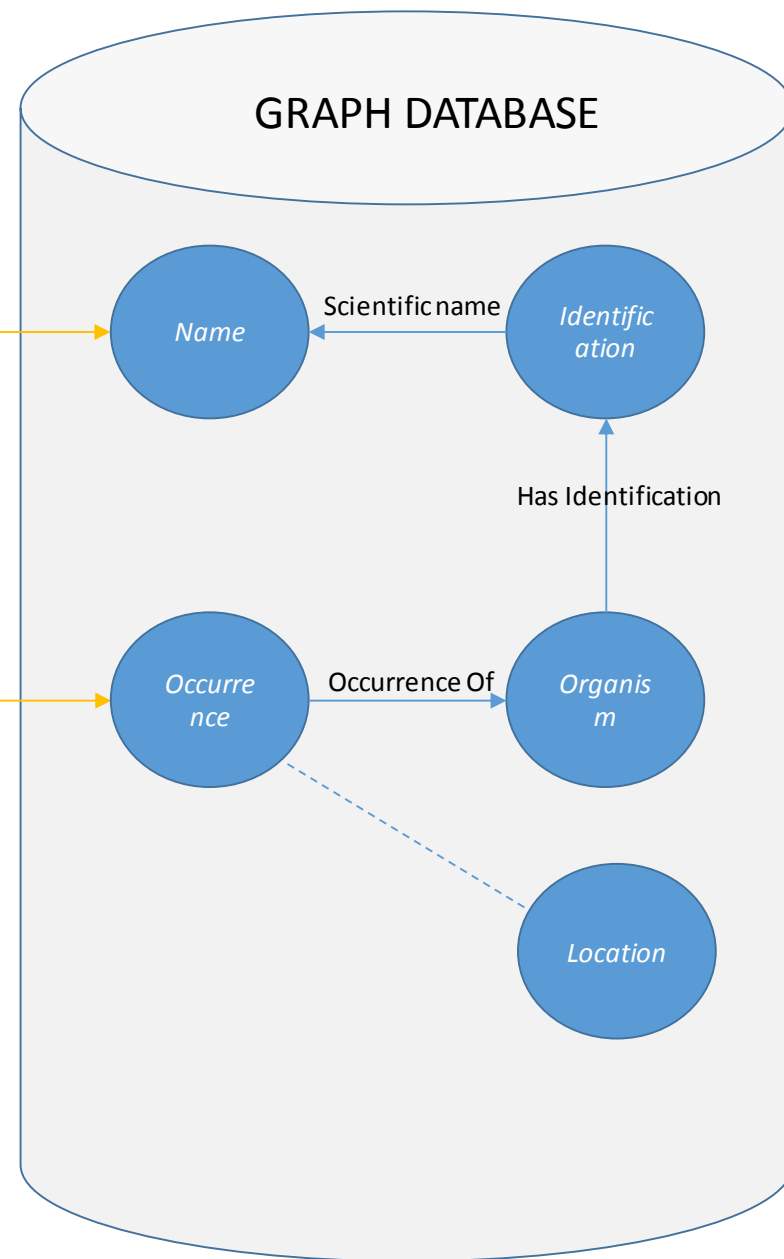
Download as CSV 

Holotype:

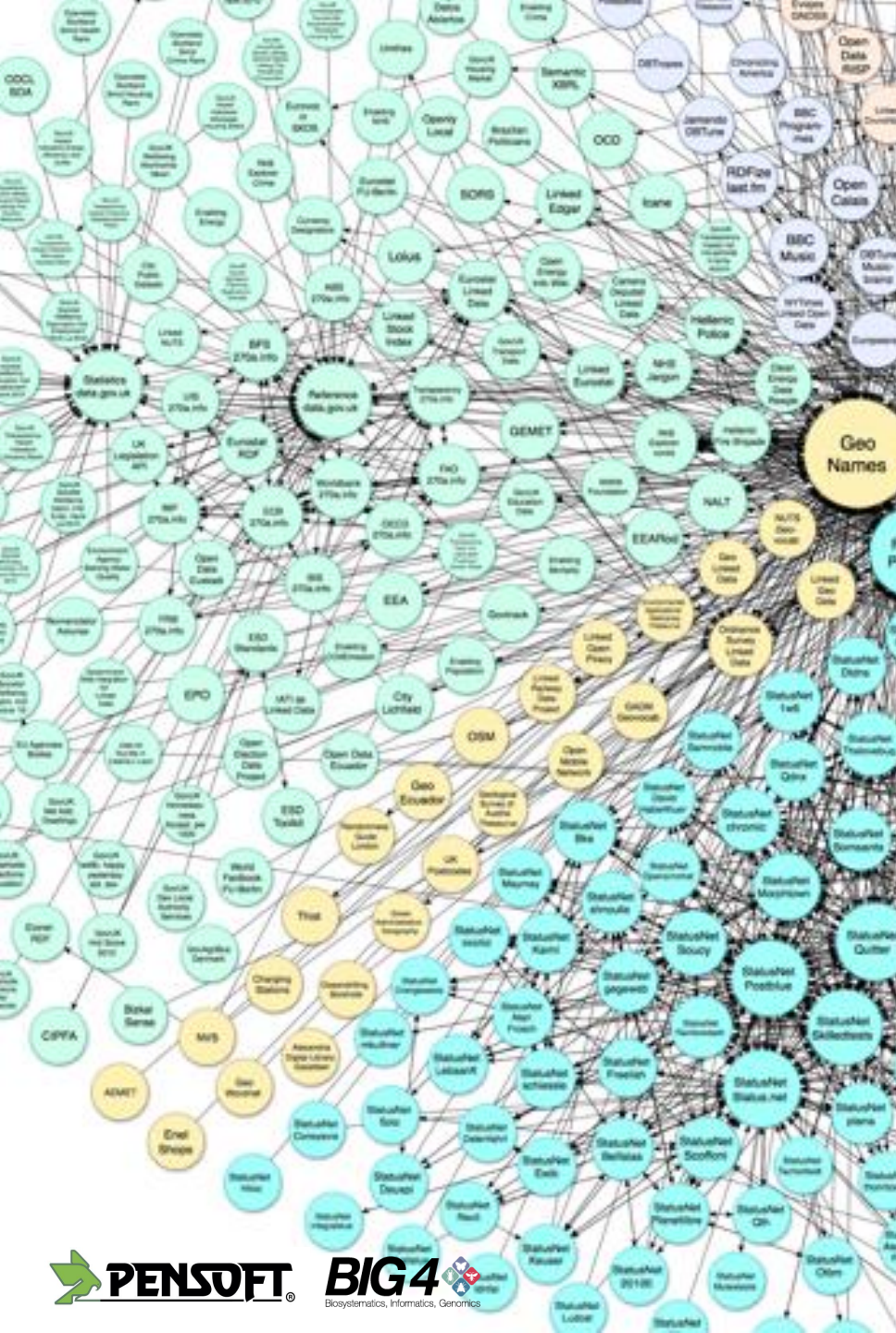
a. scientificName: *Heser stoevi*; family: *Gnaphosidae*; taxonRank: species; scientificNameAuthorship: Deltshev; continent: Asia; country: Turkmenistan; stateProvince: Lebap; county: Koytendag District; locality: v. Hojeypil, around the Dinosaurs tracks Site; verbatimElevation: 1150; decimalLatitude: 37 56.443; decimalLongitude: 66 37.597; samplingProtocol: hand collecting; eventDate: 05/25/2015; habitat: under stones; individualCount: 1; sex: male; lifeStage: adult; recordedBy: P. Stoev; identifiedBy: Christo Deltshev; dateIdentified: 2015; institutionID: NMNHS; collectionID: Archnida; basisOfRecord: PreservedSpecimen

Paratype:

a. scientificName: *Heser stoevi*; family: *Gnaphosidae*; taxonRank: species; scientificNameAuthorship: Deltshev; continent: Asia; country: Turkmenistan; stateProvince: Lebap; county: Koytendag District; locality: v. Hojeypil, around the Dinosaurs tracks Site; verbatimElevation: 1150; decimalLatitude: 37 56.443; decimalLongitude: 66 37.597; samplingProtocol: hand collecting; eventDate: 05/25/2015; habitat: under stones; individualCount: 1; sex: female; lifeStage: adult; recordedBy: P. Stoev; identifiedBy: Christo Deltshev; dateIdentified: 2015; institutionID: NMNHS; collectionID: Archnida; basisOfRecord: PreservedSpecimen



LINKED
OPEN
DATA



Methodology

- **Ontology engineering** to create the model
- **Semantic Web** technologies to create the knowledge graph
- **Data mining** to answer competency questions

Main results obtained so far

- Data import workflows (aggregators) for **structured biodiversity data** and novel article types
- Automated creation of data paper manuscript from EML metadata
- Decision on which **ontologies** to use and what to further develop
- A **preliminary semantic model** of a biodiversity publication and examples of nano-publications
- A **prototype** of the **Biodiversity Knowledge Graph**



My expectations on the possible impact of the action on my future career

- Establish a **good scientific reputation** in the international biodiversity and bioinformatics communities thanks to **networking**, publications, and implementation of the OBKMS
- **Acquire new knowledge** in computer science and biology
leading to
Research-oriented position at the **intersection of artificial intelligence and biology**

Next steps

- Implementation of OBKMS
- Linking of OBKMS to other domains, in particular **genomics** through secondment to ERA7
- Planned publications:
 - “Semantic model of a biodiversity publication”
 - “Ontology population of OBKMS with Linked Open Data”
 - “Implementation, functionality, and architecture of the Open Biodiversity Knowledge Management System”

Journal articles published during action so far

1. Senderov V, Penev L (2016) The Open Biodiversity Knowledge Management System in Scholarly Publishing. Research Ideas and Outcomes 2: e7757.doi: [10.3897/rio.2.e7757](https://doi.org/10.3897/rio.2.e7757)
2. Faulwetter S, Pafilis E, Fanini L, Bailly N, Agosti D, Arvanitidis C, Boicenco L, Catapano T, Claus S, Dekeyzer S, Georgiev T, Legaki A, Mavraki D, Oulas A, Papastefanou G, Penev L, Sautter G, Schigel D, Senderov V, Teaca A, Tsompanou M (2016) EMODnet Workshop on mechanisms and guidelines to mobilise historical data into biogeographic databases. Research Ideas and Outcomes 2: e10445.doi: [10.3897/rio.2.e10445](https://doi.org/10.3897/rio.2.e10445)
3. Cardoso P, Stoev P, Georgiev T, Senderov V, Penev L (2016) Species Conservation Profiles compliant with the IUCN Red List of Threatened Species. Biodiversity Data Journal 4: e10356. doi: [10.3897/BDJ.4.e10356](https://doi.org/10.3897/BDJ.4.e10356)
4. Senderov V, Georgiev T, Penev L (2016) Online direct import of specimen records into manuscripts and automatic creation of data papers from biological databases . Research Ideas and Outcomes 2: e10617. doi: [10.3897/rio.2.e10617](https://doi.org/10.3897/rio.2.e10617)

Presentations and events

Dates	Venue	Attendance (y/n)	Presentation (y/n)
June 2015	EU BON, Cambridge, UK	Yes	No
June 2015	EMODNET, Crete, Greece	Yes	No
July 2015	BEXIS, Jena, Sofia	Yes	No
August 2015	BOLD, Guelph, Canada	No	Yes (junior author)
September 2015	BIG4, Copenhagen, Denmark	Yes	Yes
November 2015	PlutoF Workshop, Sofia, BG	Yes	No
Oct – Feb 2015	BAS, Sofia, Bulgaria	Yes (seminar)	Yes
March 2016	EU BON, Sofia, Bulgaria	Yes	Yes
March 2016	Plazi Workshop, Sofia, BG	Yes	No
April 2016	Ontotext, Sofia, Bulgaria	Yes (technical meeting)	Yes
June 2016	BIG4, Podiji, Czechia	Yes	Yes (3)
July 2016	iDigBio + Pensoft Webinar	Yes (webinar)	Yes
July 2016	Island Biology Conference, ESP	No	Yes (2, junior author)
August 2016	IUCN, Hawaii, USA	No	Yes (junior author)

Planned presentations

December 2016: Biodiversity Informatics Standards (TDWG) 2016 Annual Conference:

- Talk: “The Open Biodiversity Knowledge Management System: A Semantic Suite Running on top of the Biodiversity Knowledge Graph”
- Computer demo: “Demonstrating the prototype of the Open Biodiversity Knowledge Management System”
- Talk: “Nanopublications for biodiversity: concept, formats and implementation”
- Talk: “Next generation publishing for biodiversity using Pensoft’s ARPHA writing tool and publishing system”
- Computer demo: “Biodiversity Informatics Standards (TDWG)”
- Computer demo: “Online Import of Occurrence Records into Manuscripts from Taxonomic Databases using Pensoft’s ARPHA* Writing Tool”
- Computer demo: “Creation of Data Paper Manuscripts from Ecological Metadata Language (EML)”

International collaboration



Other activities

- Outreach:
 - 8 posts in own blog: <http://openbkms.blogspot.com>
 - 5 posts in the Pensoft blog: <http://blog.pensoft.net>
 - Social media activity (BIG4, own, and Pensoft accounts)
 - Webinar presentation with iDigBio
- Courses and completed exams:
 - Introduction to taxonomy (Prof. Boyko Georgiev, Dr. Pavel Nikolov)
 - MATLAB course (BAS)
 - Language course
 - BIG4 course in the Czech Republic

Thank you!

- PhD Financed through the EU Marie-Sklodovska-Curie Program, Grant Agreement Nr. 642241
- Advisor: Prof. Lyubomir Penev
- Consultants:
 - Dr. Kiril Simov
 - Dr. Eamonn O'Tuama
- BIG4 partners
- All international collaborators:
 - Plazi
 - ASU
 - Illionis

