language
science
press

LangSci Series Editors Meeting 2023

LangSci
November 17, 2023
2023-11-17, HU Berlin

Funding

- First funding round 2018–2020
  - **105** institutions * **1000** EUR = 100k€
- Second funding round 2021–2023
  - **116** institutions * **1000** EUR = 116k€
- Target third founding round 2024–2026
  - **100** institutions * **1300** EUR = 130k€

Countries

- Lion's share
  - **Germany**
- well represented
  - **Netherlands**, **Switzerland**, **Belgium**
- OK
  - **Sweden**, **Norway**, **Finland**, **Austria**
- can do better
  - **UK**, **US**, **Australia**, **France**
- low representation
  - **Spain**, **Eastern Europe**
- no representation
  - **Asia**, **Africa**, **Middle East**, **Latin America**, **Denmark**, **Ireland**, **New Zealand**, **Italy**, **Portugal**

Series evolution

- currently 35 series
- two more series accepted
  - *Current Issues in Celtic Linguistics*
  - *Frame-Based Approaches to Semantics*
- rejected
  - series on translation practices (not linguistic enough)
- in preparation
  - Papuan linguistics
- decommissioned
  - *Morphological Investigations*
  - *Topics at the Grammar-Discourse Interface*
  - *Studies in Diversity Linguistics*
- dormant
  - *Computational Models of Language Evolution*
  - *Classics in Linguistics*

List of forthcoming books 2024 (33)

Childs ALGAD
Jenks ALGAD
Green CAL
Essegbey CAL
Serrelli CAM
PutnamPolinsky CIB
Payne COGL
Terhart COGL
Dyck ELA
Gibson EOTMS
Winckel EOTMS
Herkel HPLS
Stockigt HPLS
Andreassen LV
Wagner LV
Bech OGL
Favaro ORL
Rosemeyer ORL
Howe ORL
Zimmermann OSL
Giouli PMWE
Miestamo RCG
Bartens SCL
Enke SIDL
Ahn SILP
Fuchs SILP
Brunner TBLS
Neacșu TBLS
Vaissiere TBLS
Tallman TPD
Bracks TPD
Nikulin TPD
Kuznetsova TPD

List of forthcoming books 2025 (14+)

Auer AHL
Haig CAM
Lamberti CAM
Stoynova CAM
Schroeder CAM
Soesman CIB
Niinaga COGL
Zahrer COGL
Rohleder COGL
Krifka EOTMS
Kuelpmann EOTMS
Sinnemäki RCG
Faller RCG
Richter TBLS

Priorisation of submissions

- In the next funding period 2024–2026, we will produce 33 books a year
- We will have 31 productive series, with 2 more accepted series
- ~ 1 book/year per series
- How to resolve competition for scarce slots?
  - proposal: series with less books in a given year goes first

Book policies

- We have been opposed to festschrifts from the beginning, but somewhat lenient
- We have suffered, and we have learned from our mistakes
- Festschrifts are by and large a nightmare
  - Stefan and Felix can attest to that
- We will not publish anything which resembles a festschrift from now on. Period.

Quality of submissions

- In general, the technical quality of submissions has significantly improved
- People use our templates, and series provide guidance to authors about what is expected from them
  - CAL has workshops at their annual meetings and some kind of LaTeX response team.
  - please share the experiences of your series

Quality of submissions

- One submission with very shoddy papers
- We cannot assure publication with the resources we have at our disposal.
- Cavalier disregard of any and all guidelines, which we cannot compensate.
- Volume editors will have to rework the papers, but they do not have the resources either
- Book stalled and unlikely to be published any time soon, or even at all.

- Given the increased output and workload, we will have less time for “rescue missions” for manuscripts in distress.

Open Text Collections

- New project joining the *Berlin-Brandenburg Academy of Sciences*, the *Endangered Language Documentation Programme*, and *Language Science Press*
- Get interlinearized texts in a structured format, treating them as research data
- Generate HTML pages, PDFs, but also NLP-compatible formats like JSON or CSV
- Print-on-demand books to be published via LangSci
- All costs covered until 2026, marginal costs after that

Room for open questions

⟩

In the context of a new series proposal, the advisory board raised the question of Open Science.

⟩ https://github.com/langsci/guidelines/issues/2

Open Science

- LangSci is Open Access, we should probably Open Science as well
- We already have
  - open source software
  - source code of the books available
  - bibliographies on Glottolog
  - examples in IMTVault
  - editable graphics
- we kind of have
  - a data/ folder on the GitHub repositories where tabular data can be stored
  - this is currently not well supported, nor used a lot

Issues in the field of Open Science

1. Replicability.
   - ⟩ Researchers should be able to redo the analysis (or parts thereof) and verify the results
   - ⟩ Full chain of scientific argumentation
2. Reuse
   - ⟩ Data once collected should be available for further/other/different research down the road

⟩ Discuss whether replicability or reuse is our main focus.

FAIR Principles

The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the FAIRification process.

- In the domain of reuse, the FAIR principles are a common yardstick
  - Findable
  - Accessible
  - Interoperable
  - Reusable

Findable

F1. (Meta)data are assigned a globally unique and persistent identifier

F2. Data are described with rich metadata (defined by R1 below)

F3. Metadata clearly and explicitly include the identifier of the data they describe

F4. (Meta)data are registered or indexed in a searchable resource

Accessible

Once the user finds the required data, they need to know how the data can be accessed, possibly including authentication and authorisation.

A1. (Meta)data are retrievable by their identifier using a standardised communications protocol

- A1.1 The protocol is open, free, and universally implementable
- A1.2 The protocol allows for an authentication and authorisation procedure, where necessary

A2. Metadata are accessible, even when the data are no longer available

Interoperable

The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.

I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (Meta)data use vocabularies that follow FAIR principles

I3. (Meta)data include qualified references to other (meta)data

Reusable

The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
  - R1.1. (Meta)data are released with a clear and accessible data usage license
  - R1.2. (Meta)data are associated with detailed provenance
  - R1.3. (Meta)data meet domain-relevant community standards

Additional considerations

- third party copyrights
- personality rights
  - audio
  - video
  - children

Take-home messages?

- What does this mean for your subfield?
- What existing strucures can we use?
  - Zenodo
  - OSF
  - Trolling
  - DELAMAN
  - ...
- What should be recommend to authors?
- What should we recommend to future series?
- What should we recommend to existing series?