TECHNISCHE UNIVERSITÄT CHEMNITZ

E-Science-Tage 2021

BIRD: USING CONVERSATIONAL USER INTERFACES TO PROVIDE RELEVANT METADATA FOR INTERDISCIPLINARY RESEARCH DATASET PUBLISHING

André Langer, Lukas Schmolke, Martin Gaedke

Professorship for Distributed and Self-organizing Systems, Faculty of Computer Science, Chemnitz University of Technology, Germany

SITUATION

„**Research data** (RD) refers to the results of observations or experiments that validate research findings" [1]. Scientists are encouraged to **publish** research datasets (RDP) together with relevant meta information in a FAIR way so that others can find and reuse them also in an **interdisciplinary** context.

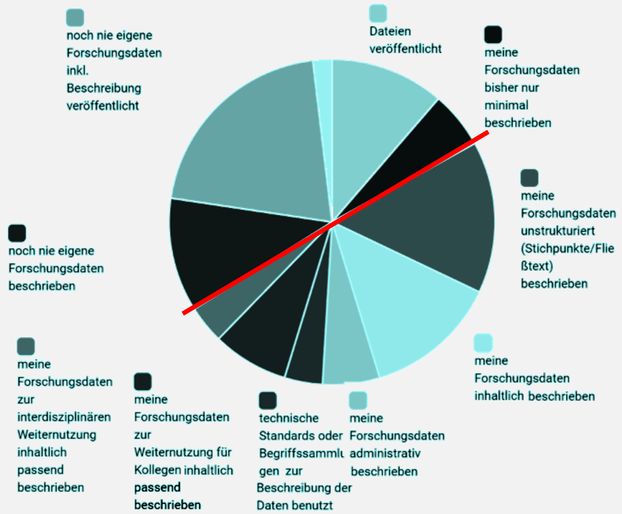


Figure 1: Preliminary survey among 24 early-career researchers on RDP [2]

PROBLEM DOMAIN

- (Early Career) Researchers are not necessarily aware of existing RDM (metadata) standards
- Motivation to complete extensive submission forms limited
- Descriptions often unstructured, general and not context-aware

OBJECTIVES

OBJ1: Chatbot dialog - based user interaction

OBJ2: Collection of descriptive research meta data

OBJ3: Adaptive conversation progress

OBJ4: Possibility for additions and changes

OBJ5: Structured, platform-independent result export

EXISTING APPROACHES

Multiple existing standards for describing research datasets

- DCAT-AP
- DataCite / OpenAIRE Guidelines
- Schema.org/Dataset

Input UIs to create such a metadata description still quite static

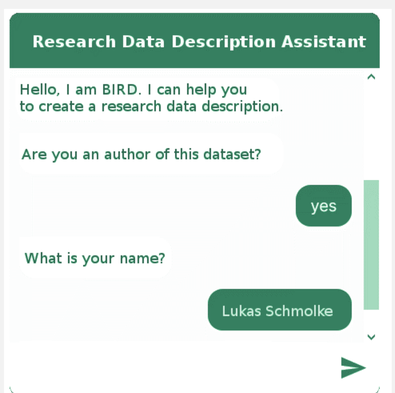
- Submission interfaces in RDM applications
- DataCite Metadata Generator
- Schema.org Markup Generator

BIRD: a Bot-based Interface for Research Descriptions

BIRD uses a **conversational user interface** to facilitate the creation of FAIR research data meta descriptions with a focus on **descriptive metadata** decoupled from a particular application.

It demonstrates different user interaction scenarios exemplarily based on the **Rasa framework**, the OpenAIRE/DataCite class and property definitions and provides a structured XML metadata download functionality.

The implementation of an intent-based solution was straight-forward for general metadata. NLP processing performance is OS-dependent and the framework documentation has pitfalls. Handling robustness and a posteriori corrections is challenging.



CONTRIBUTION + FUTURE WORK

To foster the interdisciplinary publication and discovery of research data, tools with a better **user interface experience** have to be developed that allow a natural and effective provision of **structured, relevant information** with limited prior knowledge..

The **BIRD prototype** shifts away from traditional forms and allows a dialog-based textural or even lingual metadata collection [3].

After deployment, it is going to be assessed in a real-world usage scenario. The main focus is then on incorporating taxonomical persistent semantic concept identifiers for common research characteristics and to improve the adaptive dialog behavior.

BIRD prototype: **https://www.pirol-data.de/bird**

Git Repository: **https://purl.org/net/vsr/bird** Public soon

REFERENCES

1. Elsevier (2020). Sharing Research Data (Accessed: 2021-02-17) https://www.elsevier.com/authors/author-resources/research-data
2. Tietz, M., Langer, A. & Gaedke, M. (2020). Umfrage-Ergebnisse zur Beschreibung von Forschungsdaten (Accessed: 2021-02-17) https://purl.org/net/vsr/storch/survey
3. Kim, S., Lee, J. & Gweon, G. (2019). Comparing Data from Chatbot and Web Surveys: Effects of Platform and Conversational Style on Survey Response Quality. 1-12.
4. Langer, A. (2018). PIROL: Cross-domain Research Data Publishing with Linked Data technologies. Proceedings of the Doctoral Consortium Papers Presented at the 31st CAiSE 2019. Rome, Italy. June 05 - 07, 2019; pp. 43-51

ACKNOWLEDGEMENT

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 416228727 – SFB 1410

CONTACT

Dipl.-Inf. André Langer
PIROL: Publishing Interdisciplinary Research over Linked Data [4]
Phone: +49 371 531 30 469
Mail: andre.langer@informatik.tu-chemnitz.de