EDIE - Elexis DIctionary Evaluation Tool

Seung-Bin Yim[1], Lenka Bajčetić[1], Thierry Declerck[1,2], and John P. McCrae[3]

[1] Austrian Academy of Sciences
[2] DFKI GmbH, Multilinguality and Language Technology
[3] Data Science Institute, NUI Galway

Abstract. This paper describes the current status of the EDIE - Elexis DIctionary Evaluation tool, which is aiming at evaluating the availability and usability of linked lexical resources and dictionaries which are accessible using the Elexis infrastructure.

Keywords: Dictionary Evaluation · eLexicography · Linked Data.

1 Introduction

In order to assess the quality, availability and usability of the data contained in a complex lexicographic infrastructure such as Elexis [7],[4] it is necessary to define the most important aspects of such a lexicographic repository which require evaluation. In Elexis, the relevant aspects for this have been grouped into three categories:

1. the technical quality, which ensures that the resource maintains a valid encoding of the data schemes covered by the infrastructure
2. the operational quality, which verifies that the lexical resources are available and accessible as they are deployed on the Web
3. the scientific quality, which confirms that the results of each service are correct for the tasks they are accessed for

In order to provide the end users of the infrastructure more insight into the quality, availability and usability of the lexical resources made available by Elexis, we implement EDIE - the Elexis DIctionary Assessment tool[5]. This tool is designed to assist users in context-dependent qualitative assessment of the linguistic resources instead of providing an general absolute evaluation score that makes a verdict about whether a lexical resource is of a "good" or "bad" quality. This design decision is made due to the complexity to derive general evaluation measure that apply to all types of lexical resources, especially in a multilingual context, and also the fact that the quality of a resource cannot be evaluated without context. In general, we consider newer, bigger, richer dictionaries as better in

---

[4] See `https://elex.is/` for more details.
[5] The code is available here: https://github.com/elexis-eu/edie and the service will be deployed shortly on the Elexis platform

comparison to older, smaller and scarcer resources. However, it is quite conceivable that for a particular purpose a user might find the "older, smaller and scarcer" resource to better serve her/his purposes. In essence, we want to design a tool that provide users with useful information that would help them compare lexicographic resources in various relevant aspects, so they can make their own assessment according to their particular needs.

Dictionary evaluation has two major goals: (1) to assist users in their decision-making in acquiring the best lexical resources and dictionaries for their usage needs by presenting them with a well-founded analysis of their various (qualitative and quantitative) features, and (2) to assist lexicographers in optimizing the functionality of their dictionaries [6].

Our work mostly focuses on the first of the two goals, aiming to assist users of lexicographic resources in their choice for a particular need. However, evaluating the content of a dictionary in a multilingual and a multifaceted context is too complex of a problem to be handled automatically. Even asking an expert to provide their opinion on a lexicographic resource does not mean that their measure of quality and usability is suitable for a certain use-case. Our tool is thus delivering information on various aspects, so that the end-users can make up their mind.

2   Background

The work described in this paper is done in the context of the Elexis project. This project aims to expand and enrich the European lexicographic infrastructure by providing access to NLP tools and resources. The multilingual infrastructure is intended to be used by academics, students, researchers, programmers, dictionary creators, etc. Elexis offers tools which can help a user create, modify, and publish a dictionary, or link an existing dictionary to other lexical resources.

At the core of Elexis is the so-called dictionary matrix, a universal repository of linked senses, meaning descriptions, etymological data, collocations, phraseology, translation equivalents, examples of usage and all other types of lexical information found in all types of existing lexicographic resources, multilingual,monolingual, modern, historical etc. Data from this resource is available through a RESTful Web service, as part of the platform. EDIE is situated at this access interface. Figure 1 shows the overall architecture of Elexis and the place the dictionary matrix has in this infrastructure.

Since there are numerous possible use-cases, as well as different types of end users, we needed to create a generic dictionary assessment tool which would work best under these ambiguous circumstances. Since we cannot make any definitive assumptions regarding the goal of the end users and their priorities regarding dictionary quality, we have decided to create a tool which would leave the final evaluation to the end users, while providing them with enough information to make their own estimate. The tool is described in the next section.
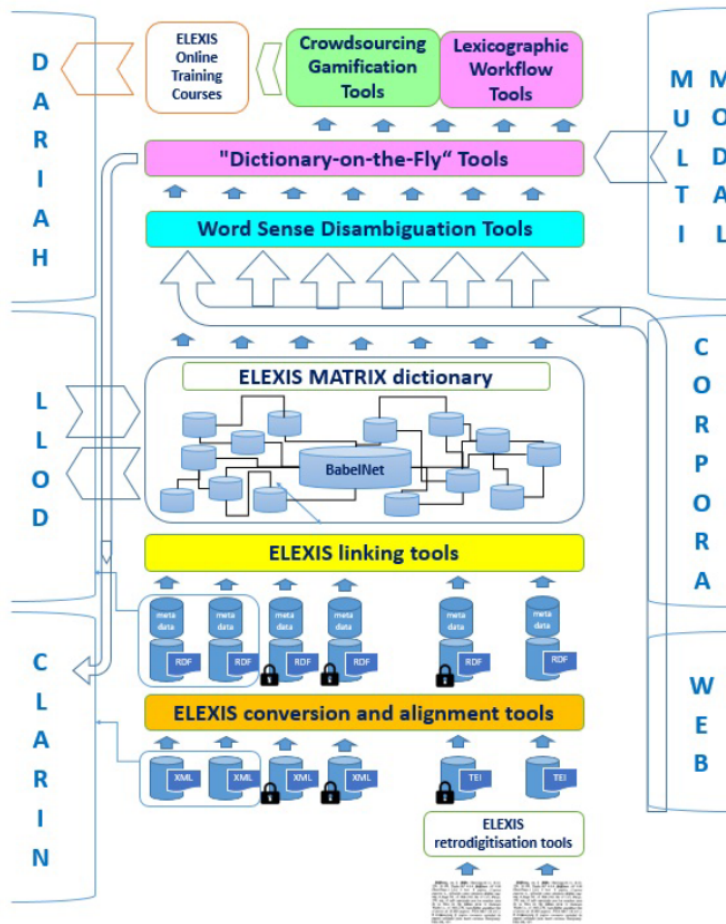
Fig. 1. Overall Architecture of Elexis

3   EDIE – the Elexis DIctionary Assessment tool

The general structure of EDIE is shown in Fig. 2.

The EDIE tool includes the main evaluator controlling the evaluation logic, an API client to retrieve the required data for evaluation and helper functions. The three Evaluator packages contain specific evaluators for different aspects of the resources. The model package contains the two domain classes, that are the main focus of EDIE, the EntryModel and the Metadata, representing the two relevant aspects of a resource: content and metadata respectively. The content of a lexical resource is represented using the EntryModel class. This class has all the fields an entry could have, e.g. lemma, senses, examples, part of speech, etc. Iterating through the entries of a lexical resource, EDIE creates a statistical
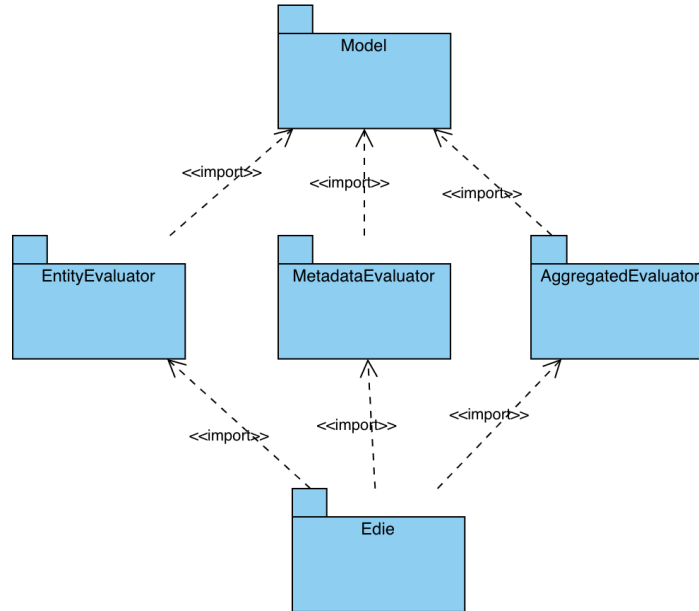
Fig. 2. EDIE package diagram

overview of a 'typical' entry, defining the average structure and type of information which can be found in such a dictionary. These statistics can provide useful insight into the content of a specific lexical resource included in the Elexis lexicographic infrastructure.

Besides the Entry class which is used to model the content of a dictionary, EDIE also relies on the Metadata class, which encapsulates all the metadata information which can be found in the Elexis infrastructure for a particular dictionary. The metadata class has all fields defined by Dublin Core and those used by the whole Elexis infrastructure.

Using the model, we can easily see which elements of the metadata are present and which are missing, and this provides an assessment of the metadata completeness. Current implementation only supports evaluation of completeness, but more metadata evaluators can be easily be added by simply implementing the MetadataMetric base class. The provided metadata is also used when a context-specific evaluation should be run. In such case, language, genre and type metadata of dictionaries is used to filter out non-relevant dictionaries, then the rest of the information about a particular dictionary to other Elexis dictionaries

of the same category, so that the output of our assessment would provide the user information within context.

For example, if a dictionary is categorized as a terminological dictionary of French, we can compare its properties to other terminological dictionaries of French, and the output of the system would be something like: "this dictionary is bigger and more recent than the average dictionary in its category". This way, we make sure that the comparisons we make are useful and reasonable. If the end user wants to make sure that they are using the biggest and newest resource, they can use the system output as proof that their resource is the most recent and biggest available in the Elexis lexicographic network.

Table 1. shows different available metrics of EDIE. We have mostly focused on the entry metrics since the user can quickly inspect the available metadata, but cannot quickly glance through hundreds or thousands of entries. The proposed metrics are selected because they provide a quick insight into the dictionary structure, sense granularity, and the type of information we can encounter. In addition, it is also useful to see which entry formats are supported by the dictionary, and to what extent. These metrics can be calculated for all available lexical resources of ELEXIS dicationary matrics, or for specific genres or types or lexical resources.

Metadata Metrics	Entry Metrics	Aggregated Metrics
SizeOfDictionary	formsPerEntry	DictionarySize:
Recency	sensesPerEntry	min
Digital Availability	definitionPerSense	max
Licence	definitionPerEntry	mean
	DefinitionLengthPerEntryBaCharacter	median
	DefinitionLengthPerEntryByToken	
	DefinitionLengthPerSenseByCharacter	
	DefinitionLengthPerSenseByToken	
	FormatsPerEntry	
	JsonSupportedEntries	
	TeiSupportedEntries	
	OntolexSupportedEntries	
	JsonCoverage	
	TeiCoverage	
	OntolexCoverage	

Table 1. Different metrics of EDIE Dictionary evaluation

Figure 3 shows a sample visualization output for an aggregated metric, namely dictionary size. Using a visual output such as this one is a simple way to represent a particular dictionary in reference to other ones, and make a comparison with respect to a certain aspect which is relevant to the user. We intend to allow the user to choose the category within which the comparison is made, and also the relevant aspect.
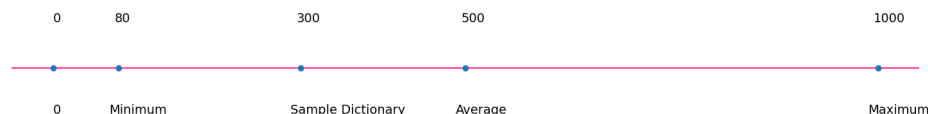
0	80	300	500	1000

0	Minimum	Sample Dictionary	Average	Maximum

Fig. 3. Visualization output for number of entries aggregated comparison

4   Related Work

Evaluation of dictionaries and linguistic resources relies on the accuracy and thoroughness of the metadata which accompanies them. Without relevant information regarding the resource, the user cannot create a verdict about the quality or the usability of a particular resource for their purpose. The assessment of metadata provided with a lexicographic resource is also called *metalexicography* [6].

One example of metadata schema used to evaluate and connect language resources is given by the META-SHARE ontology, which is described in [1].[6] While the META-SHARE ontology is a very important resource for our work, we are not aware of any initiative using it for (automatic) usability assessment of lexical resources. But we are aware of work pursued within the COST Action "NexusLinguarum"[7] and dealing with data profiling and data quality analysis in the Linguistic Linked Open Data (LLOD)[8], using for this the ABSTAT tool ([5] ; [4])[9] As many of the lexical data included in the dictionary matrix ([3])[10] of Elexis are also in RDF, we plan to join forces with the NexusLinguarum initiative, under consideration of the META-SHARE ontology.

Another initiative related to our work is "LingHub" ([2])[11], which is combining metadata from different schemes, like LRE-MAP, META-SHARE, CLARIN and more. This integration is resulting in an RDF-based set of metadata that are greatly improving the discovery of language resources. But LingHub is not

---

[6] The latest version of the META-SHARE ontology is available at `http://www.meta-share.org/ontologies/meta-share/meta-share-ontology.owl/documentation/index-en.html`.

[7] `nexuslinguarum.eu/`.

[8] See `https://linguistic\-lod.org/` for more details on the LLOD cloud.

[9] See the "Intermediate Activity Report Working Group 1 'Linked data-based language resources'" of the NexusLinguarum COST Action at `https://nexuslinguarum.eu/wp-content/uploads/2021/11/D1.3_IntermediateActivityReport.pdf`.

[10] Available at `https://github.com/elexis\-eu/dictionary-matrix`.

[11] See also `https://linghub.org/`.

dealing directly with the data itself, and the quality issues dealt with by the developers of LingHub are primmarily concerning the encoding of the metadata.

5   Future Work

Current implementation of EDIE does not have any graphical user interface for interactive exploration of the lexicographic resources. Such an user interface in combination with different statistics and comparative visualizations based on different criteria selected by users (dictionary types, genres, languages, etc.) would help the users to assess different dictionaries in a more user-friendly manner.

6   Conclusion

We have introduced the current state of development of EDIE, an evaluation tool for dictionaries, that allows users to assess different aspects of dictionaries based on their metadata and entries. Furthermore, aggregated metrics over dictionaries of interests/contexts let users compare different dictionaries for their specific use cases.

References

1. Gavrilidou, M., Labropoulou, P., Desipri, E., Piperidis, S., Papageorgiou, H., Monachini, M., Frontini, F., Declerck, T., Francopoulo, G., Arranz, V., Mapelli, V.: The METASHARE metadata schema for the description of language resources. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12). pp. 1090–1097. European Language Resources Association (ELRA), Istanbul, Turkey (May 2012), `http://www.lrec-conf.org\ /proceedings\/lrec2012\/pdf\/998\_Paper.pdf`
2. McCrae, J.P., Cimiano, P.: Linghub: a linked data based portal supporting the discovery of language resources. In: Filipowska, A., Verborgh, R., Polleres, A. (eds.) Joint Proceedings of the Posters and Demos Track of 11th International Conference on Semantic Systems - SEMANTiCS 2015 and 1st Workshop on Data Science: Methods, Technology and Applications (DSci15) 11th International Conference on Semantic Systems - SEMANTiCS 2015, Vienna, Austria, September 15-17, 2015. CEUR Workshop Proceedings, vol. 1481, pp. 88–91. CEUR-WS.org (2015), `http: //ceur-ws.org/Vol-1481/paper27.pdf`
3. McCrae, J.P., Tiberius, C., Khan, A.F., Kernerman, I., Declerck, T., Krek, S., Monachini, M., Ahmadi, S.: The ELEXIS interface for interoperable lexical resources. In: Proceedings of the sixth biennial conference on electronic lexicography (eLex). pp. 642–659. Sintra, Portugal (10 2019), `https://elex.link/elex2019/wp\-content/ uploads/2019/09/eLex\_2019\_37.pdf`
4. Principe, R.A.A., Spahiu, B., Palmonari, M., Rula, A., Paoli, F.D., Maurino, A.: ABSTAT 1.0: Compute, manage and share semantic profiles of RDF knowledge graphs. In: Gangemi, A., Gentile, A.L., Nuzzolese, A.G., Rudolph, S., Maleshkova, M., Paulheim, H., Pan, J.Z., Alam, M. (eds.) The Semantic Web: ESWC 2018 Satellite Events - ESWC 2018 Satellite Events, Heraklion, Crete, Greece, June 3-7,

2018, Revised Selected Papers. Lecture Notes in Computer Science, vol. 11155, pp. 170–175. Springer (2018). https://doi.org/10.1007/978-3-319-98192-5_32, `https://doi.org/10.1007/978-3-319-98192-5\_32`

5. Spahiu, B., Maurino, A., Palmonari, M.: Towards improving the quality of knowledge graphs with data-driven ontology patterns and SHACL. In: Skjæveland, M.G., Hu, Y., Hammar, K., Svátek, V., Lawrynowicz, A. (eds.) Proceedings of the 9th Workshop on Ontology Design and Patterns (WOP 2018) co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 9th, 2018. CEUR Workshop Proceedings, vol. 2195, pp. 52–66. CEUR-WS.org (2018), `http://ceur-ws.org/Vol-2195/research\_paper\_2.pdf`

6. Swanepoel, P.H.: Towards a framework for the description and evaluation of dictionary evaluation criteria. Lexikos **18**, 207–231 (2008)

7. Woldrich, A., Goli, T., Kosem, I., Matuška, O., Wissik, T.: ELEXIS: Technical and social infrastructure for lexicography (Jul 2021). https://doi.org/10.5281/zenodo.4607957, `https://doi.org/10.5281/zenodo.4607957`, published in K Lexical News (28), pp. 45-52