The role of crowdsourcing in lexicography

Jaka Čibej[1], Darja Fišer[1], Iztok Kosem[2,3]

[1] Department of Translation, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia
[2] Trojina, Institute for Applied Slovene Studies, Ljubljana, Slovenia
[3] Faculty of Social Sciences, University of Ljubljana, Ljubljana, Slovenia
E-mail: jaka.cibej@ff.uni-lj.si, darja.fiser@ff.uni-lj.si, iztok.kosem@trojina.si

Abstract

In the past decade, crowdsourcing has been used with great success in specialized lexicographic tasks, such as collecting candidate lexemes for dictionary updates or validating automatically identified synonyms. However, professional lexicography is only now starting to explore crowdsourcing as an integral part of the workflow, thereby opening a number of important questions that could have lasting consequences on the nature of lexicographic work, its management and financing, as well as the perception, use and life-cycle of the lexicographic product. In this paper, we address these questions through the perspective of a proposal for a new monolingual dictionary of Slovene, in which crowdsourcing will play an integral role at a number of stages of dictionary construction – from headword list creation to dealing with stylistic issues.

Keywords: crowdsourcing; microtask design; crowd motivation; quality control; legal and ethical aspects of crowdsourcing

1. Introduction

Crowdsourcing is a term first introduced in 2006 to signify a process that involves a group of people (also called a crowd) that contribute towards achieving a goal by distributing the overall workload among the individuals in the group (Howe, 2008). The crowd does not necessarily consist of experts in the relevant field. In fact, a number of crowdsourcing projects have shown that even groups of non-expert individuals are talented, creative and productive enough to solve complicated tasks that in the past were solely the domain of experts. Today, due to modern technology and the global spread of the internet, channelling the potential of the crowd is becoming increasingly simple, more affordable and effective.

Although crowdsourcing is discussed with increasing frequency in lexicography, it has not yet been tested in large-scale, diverse and comprehensive settings. As shown by Abel & Meyer (2013), user contributions to dictionaries are currently limited to collaborative lexicographic projects or dictionary correction after publication. At the same time, lexicographers are facing increasing time constraints and amounts of data. What is more, the increasing (semi-)automatization of lexicographic work is turning some stages of dictionary creation into routine processes, for which lexicographers are overqualified. This calls for the introduction of crowdsourcing and user contributions

in dictionary creation. If established, it could have lasting consequences on the nature of lexicographic work, its management and financing, as well as the perception, use and life-cycle of the lexicographic product.

In this paper, we propose to integrate crowdsourcing into the overall workflow of lexicographic projects. We also address a number of important questions that arise in the process, such as the importance of appropriate microtask design, crowd motivation, quality control as well as legal and ethical aspects of crowd payment, all through the perspective of a proposal for a new monolingual dictionary of Slovene, in which crowdsourcing will play an integral role in a number of stages of dictionary construction – from headword list creation to dealing with stylistic issues.

2. The crowd and lexicography

One of the earliest examples of obtaining active participation of the general public in dictionary production was the creation of the Oxford English Dictionary (OED) in the late 19th century, when the OED editorial board encouraged volunteers to send in their contributions containing words and examples of use (Lanxon, 2011).

In the last decade, crowdsourcing has already been used successfully in a number of linguistic projects. For example, when evaluating Puzzle Racer, an annotation game, Jurgens & Navigli (2014) find it to be equally effective compared to annotation by experts, with the costs being 73% lower. Using the CrowdFlower platform, Fossati et al. (2013) crowdsourced the annotation of FrameNet, a lexical database of English, and found the crowdsourcing method to be both faster and more accurate than conventional annotation methods. Using sloWCrowd, a custom developed open-source crowdsourcing tool for lexicographic tasks, Fišer et al. (2014) corrected errors in the automatically developed WordNet for Slovene, and found the annotators' average accuracy to be 80.12%, which is high for complex lexical semantic tasks. When annotating a silver standard corpus of Croatian, Klubička & Ljubešić (2014) find the accuracy of a single worker to be approximately 90%, and the accuracy of the majority answer of three workers to be approximately 97%.

All this suggests that crowdsourcing could also be used in lexicography to great effect – not as a final or main phase of dictionary creation, but as a method to filter and process data before its implementation in actual dictionary creation by lexicographers. However, in order to ensure the effectiveness of the crowdsourcing method, several factors must be taken into consideration: crowd motivation, microtask design, quality control, choice of crowdsourcing platforms, and legal or financial issues. An overview of these aspects is provided in the following sections.

2.1 Crowd motivation

Motivated contributors are crucial for the success of any crowdsourcing project, even more so with languages of limited diffusion, which cannot rely on a large pool of crowdsourcers. According to Lew (2013), the motivation provided by the project initiator can be psychological, social or economic.

Psychological motivation is based on the fact that many internet users find participating in crowdsourcing projects or contributing user-generated content psychologically satisfying or personally fulfilling, either as an act of altruism, a way of expressing their identity or simply because they find it entertaining. This motivational aspect was the basis for the development of games with a purpose (GWAP) – applications that enable individuals to solve tasks while playing a game. Examples include *Phrase Detectives,* an online game for anaphora resolution (Chamberlain et al., 2008); *Verbosity*, a game for collecting common-sense facts (von Ahn et al., 2006); *Puzzle Racer* and *KaBoom!*, both annotation games (Jurgens & Navigli, 2014); and *JeuxDeMots*, a game aimed at building a large-scale lexical network for French (Joubert & Lafourcade, 2012).

With social motivation, individuals are driven by their urge to interact with others who share similar interests. Such a group is willing to contribute to a project that will benefit their community, perhaps by resulting in a useful product or by providing a chance for the individuals to improve their skills or to express their enthusiasm for a particular topic. A subcategory of social motivation is educational motivation (e.g. students solving tasks either as part of their academic obligations or as an extra-credit activity). Another aspect of social motivation involves the recognition a contributor receives for their work and effort in a community; for instance, an esteemed title (e.g. Wikipedia Editor) or credit on a hall-of-fame list. Successful projects involving social motivation include a number of well-known collaborative projects, such as *Wiktionary* and *Urban Dictionary* or its Slovene counterpart *Razvezani jezik*[1].

When crowdsourcing is used for large-scale or commercial projects where a substantial input or long-term involvement is expected from crowdsourcers, researchers typically resort to economic motivation by offering *micropayments*, i.e. small remuneration paid to the contributor for every successfully completed task (cf. Rumshisky, 2011; Akkaya et al., 2010; Fossati et al., 2013). Other types of economic motivation include prizes and vouchers (cf. El-Haj et al., 2014; Fišer et al., 2014). If using economic motivation, it is important to bear in mind the ethical aspects of recruiting and paying the crowdsourcers relative to the difficulty level and time spent on the task, cost of living in their country of residence, easy access to the earnings, etc. (cf. Sabou et al., 2014).

---

[1] http://razvezanijezik.org

2.2 Microtask design

Since microtasks are often undertaken by non-experts, they need to be simple to process both mentally and logistically. They should not be too time-consuming nor should they require a high degree of expertise or too much introductory training. As pointed out by Rumshisky (2011) and Biemann & Nygaard (2010), crowdsourcing tasks should be kept simple (with clear, short instructions) and designed to enable maximum effectiveness by splitting complex annotation into simpler steps. The importance of well-designed microtasks is also pointed out by Kosem et al. (2013), who showed that complex, multi-dimensional questions, or those that require subjective evaluations, do not yield satisfactory results.

2.3 Quality control

There are a number of ways to control the quality and consistency of crowdsourcing results. The first method is the *gold standard*, a dataset which contains a number of microtasks that have been pre-annotated (already answered correctly) by experts. These tasks are offered to crowdsourcers at various points during their work in order to test their reliability. If an individual fails to pass a threshold, his or her answers are deemed unreliable and are excluded from the final results (Rumshisky, 2011).

Another way of controlling quality is to observe *inter-annotator agreement*. This is achieved by offering different crowdsourcers the same task, thus obtaining multiple answers for each task. The final decision is achieved by taking into consideration the *majority vote*, i.e. the answer chosen by the most annotators. Based on the distribution of the multiple answers, a *confidence score* per microtask or per crowdsourcer can be computed (Oyama et al., 2013). However, it is important to consider that an optimal balance must be achieved between multiple annotations for the same task and new annotations, as multiple annotation is costly.

The (borderline or difficult) cases with insufficient consensus among crowdsourcers may then be manually annotated by an expert. This process is called *refereeing*. If the microtasks were designed properly and the annotation process successful, the expert is only required to evaluate a small number of ambiguous examples, while the bulk of the work is still crowdsourced. If, on the other hand, the annotators disagree in a significant number of cases, it might indicate that the microtasks were not designed efficiently, were not assigned to the appropriate target group, or that the annotation guidelines need to be further refined to provide clearer instructions (Fossati et al., 2013).

The last approach to quality control is observing *intra-annotator agreement*, which measures the consistency of a single crowdsourcer in answering the same microtasks at various points of their engagement (Gut & Bayerl, 2004). This allows for the exclusion

of unreliable annotators who are either 'spam workers', not knowledgeable enough or not confident enough to provide consistent answers. This process, however, is also costly. The more common the iteration of previous questions, the smaller the number of new annotations that will ultimately be available. Also, iteration should not be noticed by crowdsourcers as this may affect their motivation.

2.4 Legal and financial issues

When using crowdsourcing for lexical resource development, a number of legal and financial issues arise. Although these depend heavily on local legislation and project funding, we provide a general overview of the key issues that need to be taken into consideration. Although they are not central to the content and quality of lexicographic projects, they often act as a significant barrier to lexicography embracing crowdsourcing since most lexicographic teams, especially in academic settings, are unfamiliar with the legislation restrictions in this area and rarely get sufficient support from legal experts in the field.

Dataset availability – If the datasets used in crowdsourcing are to be made available to the public, a suitable license needs to be selected in accordance with local legislation on copyright and personal data protection.

Disclaimer – Before contributing to the project by solving tasks, crowdsourcers should agree to a disclaimer that informs them on how the results of their work will be used.

Crowdsourcer acknowledgement – Because crowdsourcers typically contribute a sizeable amount of work to the project, it needs to be determined if and how they should be credited on the final product in accordance with local copyright legislation.

Recruitment restrictions – Local legislation may impose restrictions on crowdsourcer recruitment. This is especially true in the case of under-aged workers.

Payment restrictions – Another matter to consider is potential payment restrictions, e.g. how local tax legislation treats micropayments or prizes for participating in crowdsourcing projects.

2.5 Crowdsourcing platforms

In this section, we provide an overview of the platforms that have either already been used for crowdsourcing in linguistics or show potential in lexicography. Both commercial and open-source crowdsourcing platforms exist.

The most widely known and used crowdsourcing platform is **Amazon Mechanical Turk**[2] (cf. Rumshisky, 2011; Rumshisky et al., 2012; Biemann & Nygaard, 2010; Snow et al., 2008). Campaign management, quality control measures and payment support are already integrated in the administrator's interface, and a substantial crowdsourcing community has already been recruited, at least for the bigger languages. Similar examples are **CrowdFlower**[3] and **Clickworker**[4] which offer a number of applications, ranging from data categorisation to sentiment analysis. Microtasks can be uploaded using CML, CSS or Javascript. Crowdsourcers can be filtered according to age, expertise or geographic location.

Among open-source platforms, the most notable is **Crowdcrafting**[5], which is based on **PyBossa**[6], a Python-based open-source framework for creating crowdsourcing projects that can be installed locally and is available under the Creative Commons License BY-SA 4.0. Another open-source tool is **sloWCrowd**[7] (Tavčar et al., 2012), which is PHP/MySQL-based and was originally developed for correcting mistakes in automatically generated semantic lexicons (such as Wordnet), but has been upgraded to allow for project-specific task specifications.

3. Crowdsourcing workflow for lexicography

In this section, we provide an overview of proposals to utilise crowdsourcing methods in the various stages of corpus-based dictionary construction projects. We propose a modular approach that can be adapted to the specific nature of the project at hand and the budget available. Not all stages need to be followed. Their order can be changed and some can be done in parallel, but it is important to at least consider the recommended phases and address the issues raised in each of them, as crowdsourcing is a complex, time-consuming and potentially costly procedure that cannot yield useful results without careful planning and task design.

Before deciding on a crowdsourcing campaign, an estimate of the required investment should be made with respect to time, money and personnel, as the campaign should not take up more time and financial and/or human resources than conventional annotation methods. However, if crowdsourcing is integrated into dictionary construction from the very beginning, different crowdsourcing tasks at all dictionary construction levels can be designed according to the same principles and use the same pre- and post-processing chains and crowdsourcing platform, making the effort of setting up a viable crowdsourcing environment all the more worthwhile.

---

[2] https://www.mturk.com

[3] http://www.crowdflower.com/

[4] http://www.clickworker.com/en/

[5] http://crowdcrafting.org/

[6] http://pybossa.com/

[7] http://nl.ijs.si/slowcrowd/about.php?project=slowcrowdmain
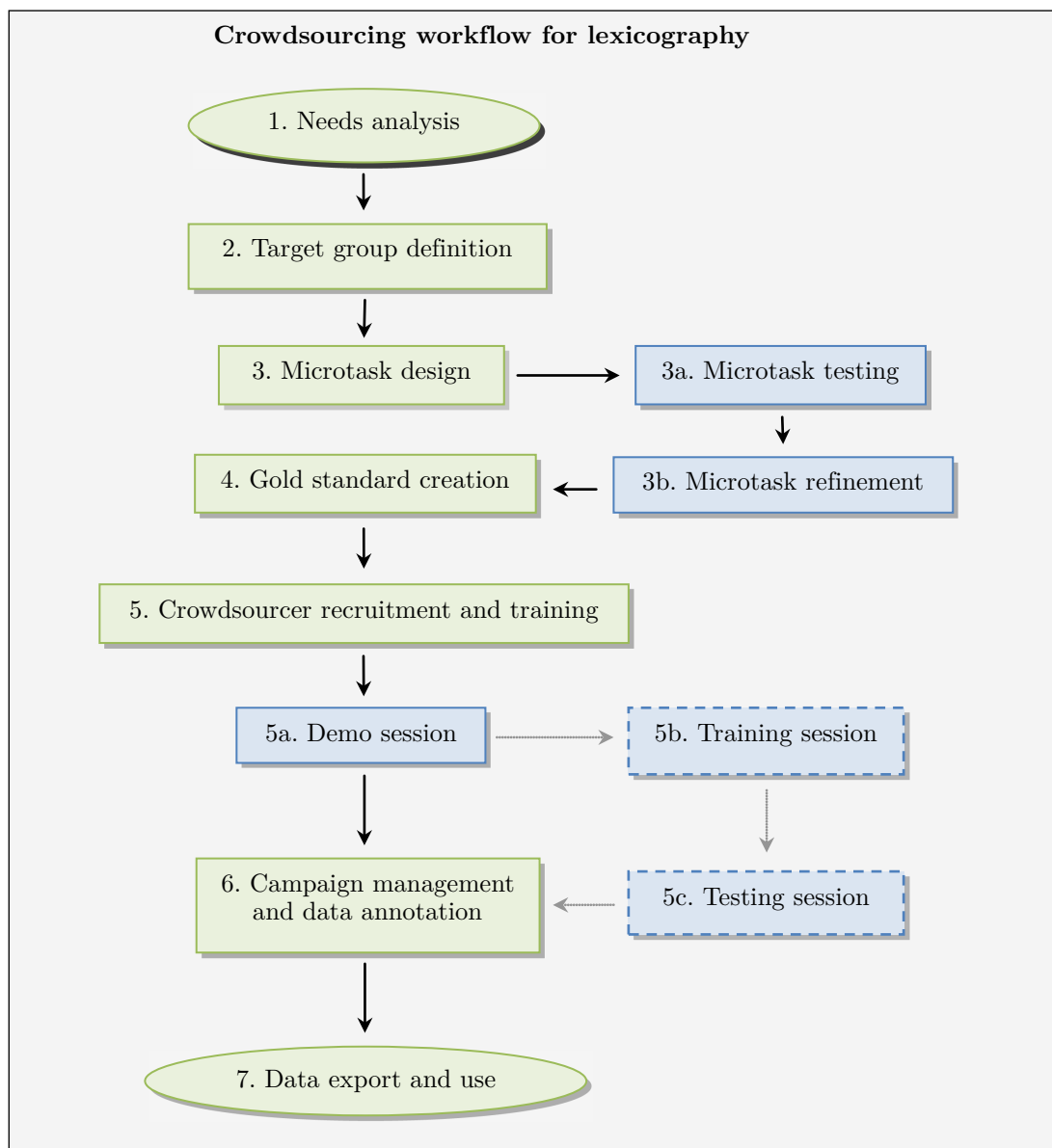
Figure 1: Crowdsourcing workflow for lexicography. Green-coloured boxes represent main phases and blue-coloured ones subphases. Dashed boxes and arrows represent optional phases which can be omitted in small-scale, low-budget campaigns

Phase 1: Needs analysis – The first step of each crowdsourcing campaign requires a thorough needs analysis. Apart from the goal and expectations of the campaign (i.e. what can be expected in terms of volume and usability of the obtained results), it is also necessary to determine the type, amount, availability and format of the data required.

Phase 2: Target group definition – Once the needs have been analysed, it is necessary to determine the required crowdsourcer profile to ensure results of a suitable quality. The problem at hand may be suitable for the general public without any specialized linguistic or lexicographic knowledge or may require a certain degree of expertise and can only be solved effectively by e.g. language students or even expert lexicographers.

Phase 3: Microtask design, testing and refinement – The most important and

difficult part of crowdsourcing is microtask design. As already mentioned, microtasks should be one-dimensional questions with short, clear instructions, suited to the knowledge prerequisites of the target crowdsourcer profile. In addition, solving microtasks should be carried out through a user-friendly interface. No tasks should be included that do not benefit from this method and are likely to provide unreliable results. The designed microtasks need to be tested in a pilot study so that any identified incongruences and inconsistencies can be removed and any unclear, confusing or too complex microtasks refined.

Phase 4: Gold standard creation – A certain number of microtasks needs to be annotated by experts to create a gold standard that is later used to ensure the accuracy of crowdsourcing results, i.e. to filter out unreliable crowdsourcers or answers. The dataset should be as representative of the entire set of microtasks as possible, especially in terms of difficulty and complexity (e.g. it should not include only simple, transparent examples nor should it contain too many borderline examples to make it impossible for the annotators to achieve a sufficient degree of accuracy).

Phase 5: Crowdsourcer recruitment and training – Crowdsourcers need to be recruited and trained. Usually, a *demo session* (e.g. a presentation or a video) is held to introduce the crowdsourcers to the annotation process. The demo session is then followed by a *training session*, which either consists of a live annotation session supervised by an expert who offers advice and additional information to the crowdsourcers should they require it (e.g. with ambiguous borderline examples) or an online annotation session where automated feedback is provided with each answer. The next step is the *testing session*, which is used to determine whether the crowdsourcer has achieved a sufficient degree of accuracy to be recruited. In low-budget scenarios, the training and testing sessions are often skipped.

Phase 6: Data annotation and campaign management – In this step, the recruited crowdsourcers solve the microtasks provided by the initiator. The initiator needs to monitor the campaign and decide whether any additional fine-tuning is necessary, e.g. if the set of microtasks needs to be expanded, if the crowdsourcers are motivated enough to provide a consistent flow of answers, if the results meet the expectations of the project, etc.

Phase 7: Data export and use – The final phase involves exporting the crowdsourced data into an appropriate format for further use in the project (e.g. algorithm training or inclusion in a dictionary). The crowdsourcing platform should allow the data to be exported at any point of the crowdsourcing campaign for preliminary analyses.

This crowdsourcing workflow will play an integral role in the creation of a new monolingual dictionary of Slovene, the plans for which are presented in the following section.

4. Crowdsourcing for the new Slovene dictionary

Slovar sodobnega slovenskega jezika (SSSJ) is a new monolingual dictionary of Slovene planned by the Centre for Language Resources and Technologies of the University of Ljubljana (CJVT UL)[8]. The goal of the proposed project is to construct a comprehensive corpus-based dictionary of Slovene that will reflect contemporary language use and will be built in accordance with modern lexicographic trends and the increasingly digital and online nature of lexicographic products. The project envisions the creation of an open-source database that will ultimately serve not only as the basis for a new monolingual dictionary of Slovene, but will also enable the development and improvement of both existing and new language technologies for Slovene, as well as the creation of a number of specialised Slovene dictionaries for different user profiles (e.g. linguists, students, learners of Slovene as a foreign language).

The initial proposal by Krek et al. (2013), based on which the plans for SSSJ and related resources are currently being made, envisioned that the SSSJ database would be completed in five years. Judging by experience from similar projects, such as the *Algemeen Nederlands Woordenboek* (Tiberius & Schoonheim, 2014) and the *Great Dictionary of Polish* (Żmigrodzki, 2014), this is a rather short period to create a database of any language from scratch, which is why the proposal includes an important innovation in lexicography: initial automatic extraction of corpus data. This method has already been tested on Slovene by Kosem et al. (2013) and is currently being used for the purposes of the *Estonian Collocation Dictionary* (Kallas et al., forthcoming). However, automatically extracted data requires a great deal of post-processing, including many routine and trivial tasks for lexicographers; this has led to the decision to make crowdsourcing an integral part of the SSSJ database creation, based on numerous good practice examples from abroad (Klubička & Ljubešić, 2014; Jurgens & Navigli, 2014; Fossati et al., 2013; inter alia) and the successful implementation of crowdsourcing in other Slovene projects (Kosem et al., 2013; Fišer et al, 2014).

4.1 SSSJ crowdsourcing scenarios

An example of a crowdsourcing task is distributing automatically extracted examples into different senses and subsenses. During the analysis, a lexicographer first makes a rough draft of sense division with one or more short glosses or an indicator for each sense, and then distributes the (automatically extracted) examples, collocates and grammatical relations, deleting any irrelevant or incorrect information in the process. To a large extent, the distribution of information can be carried out by crowdsourcers with a microtask in which they are asked to assign the extracted corpus examples to

---

[8] http://www.cjvt.si/projekti/sssj/

the relevant (sub)sense. In addition to the available senses and subsenses, crowdsourcers may also categorise examples as *None of the above senses*, when the example cannot be attributed to any (sub)sense offered; or as *Unclear example*, when the provided context is insufficient for the crowdsourcer to select one of the (sub)senses. The final decisions are then achieved through a majority vote or, if the majority vote is not unanimous or sufficiently clear (according to a predetermined threshold), through refereeing by a lexicographer.

While the only task for crowdsourcers is the distribution of examples, the results have many other uses. For instance, crowdsourcers indirectly distribute collocates attested in the examples as well as the grammatical relations under which the collocates are provided. Moreover, the examples marked as unclear are candidates for removal from the database or at least for omission from the dictionary entry. If a significant number of examples for a particular collocate is marked as unclear, the collocate itself will also need to be inspected. While a similar approach can be used for the examples categorised as *None of the above senses*, those examples carry two other potentially valuable pieces of information as they can alert the lexicographer to an overly coarse sense division or even to an overlooked (sub)sense.

Crowdsourcing can also be implemented in a number of other aspects of dictionary compilation and language resources (both new and existing); the improvement or development of which is an integral part of a dictionary project, in our case SSSJ. We provide a number of preliminary suggestions in the following paragraphs, but many more can and will be explored within the framework of the SSSJ project, depending on the budget available.

Lexicon – Microtasks concerning the creation of the SSSJ lexicon could involve determining the standard declension paradigm of headwords, the relation between words in terms of word-formation, the categorisation of marked (e.g. non-standard) word forms, and the pronunciation of the headword and its declined forms. In addition, crowdsourcing could be used to expand the lexicon of word forms for further use in the development of language technologies for Slovene.

Grammar – In terms of grammar, solving microtasks could help determine the relationship between certain interchangeable suffixes (e.g. the plural of *študent* 'student', which can be either *študenti* or *studentje*) or word forms (e.g. the demonstrative pronouns *oni* and *tisti*).

Standard – Microtasks concerning standard Slovene might include checking lists of individual paradigms and their potential corrections, as well as adding information on pronunciation and syntax.

Stylistics – Microtasks in stylistics could contribute towards developing the taxonomy of stylistic qualifiers and determining (or confirming) the stylistic qualifiers for dictionary headwords (or at least those that are deemed problematic).

User feedback – Crowdsourcing could also contribute towards the development of a user-friendly interface for the dictionary. By solving microtasks, potential dictionary users could decide between various options in terms of design, transparency, usefulness, etc., and choose the one they find suits the best.

5. Conclusion and future work

Crowdsourcing has great potential in lexicography, as evidenced by a number of linguistic projects that have already successfully used crowdsourcing as an effective method for data processing. To ensure the successful implementation of crowdsourcing in the lexicographic workflow, many aspects need to be considered: from microtask design, data preparation, crowd profiling and motivation to legal and financial issues.

The SSSJ project aims to be one of the first dictionary projects to give crowdsourcing a prominent role in the development of a database for a new monolingual dictionary of Slovene. The experience from the project so far has already shown that the need for crowdsourcing input extends beyond the dictionary database to any related existing or future language resource, such as a lexicon or a user interface. In addition, the crowd could be used to establish a permanent user feedback channel through crowdsourcing.

It is noteworthy that the results obtained from lexicographic crowdsourcing tasks can also be used for other purposes, e.g. for the improvement of language tools used by lexicographers. For example, corpus examples identified as unclear could form a training corpus for the improvement of a tool for extracting good dictionary examples. Similarly, identifying incorrect examples of collocates under a particular grammatical relation can help fine-tune scripts for extracting grammatical relations and their collocates from the corpus.

Crowdsourcing may well become a common tool in the next generation of lexicographic projects, making it much less time- and resource-consuming to keep up with the constant changes in language use as well as the increased demand for linguistic data-processing. We can therefore envisage the emergence of in-house crowdsourcing teams focused solely on providing support to lexicographers, linguists and researchers with language-related crowdsourcing tasks.

6. Acknowledgement

7. References

Abel, A. & Meyer, C. (2013). The dynamics outside the paper: user contributions to online dictionaries. In *Proceedings of eLex 2013*, pp. 179–194.

Akkaya, C., Conrad, C., Wiebe, J. & Mihalcea, R. (2010). Amazon Mechanical Turk for Subjectivity Word Sense Disambiguation. In *Proceedings of NAACL-HLT 2010 Workshop on Creating Speech and Language Data With Amazon's Mechanical Turk*.

Biemann, C. & Nygaard, V. (2010). Crowdsourcing WordNet. In *Proceedings of the 5th Global WordNet Conference*. Mumbai, India.

Chamberlain, J., Poesio, M. & Kruschwitz, U. (2008). Phrase Detectives: A Web-based collaborative annotation game. In *Proceedings of iSemantics*. Graz, Austria.

El-Haj, M., Kruschwitz, U. & Fox, C. (2014). Creating Language Resources for Under-resourced Languages: Methodologies, and experiments with Arabic. In Language Resources and Evaluation 2014. Springer.

Fišer, D., Tavčar, A. & Erjavec, T. (2014). sloWCrowd: A crowdsourcing tool for lexicographic tasks. In *Proceedings of LREC 2014*, pp. 4371–4375.

Fossati, M., C. Giuliano & Tonelli, S. (2013). Outsourcing FrameNet to the Crowd. In Proceedings of the *51st Annual Meeting of the Association for Computational Linguistics*. Sofia, Bulgaria: Association for Computational Linguistics, pp. 742–747.

Gut, U. & Bayerl, P. S. (2004). Measuring the Reliability of Manual Annotations of Speech Corpora. In *Proceedings of Speech Prosody 2004, Nara*, pp. 565–568.

Howe, J. (2008). *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. New York: Crown Publishing Group.

Joubert, A. & Lafourcade, M. (2012). A new dynamic approach for lexical networks evaluation. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, 23-25 May 2012.

Jurgens, D. & Navigli, R. (2014). It's All Fun and Games until Someone Annotates: Video Games with a Purpose for Linguistic Annotation. *Transactions of the Association for Computational Linguistics*, 2. Association for Computational Linguistics, pp. 449–463.

Kallas, J., Kilgarriff, A. Koppel, K., Kudritski, E., Langemets, M., Michelfeit, J., Tuulik, M. & Viks, Ü. (2015). Automatic generation of the Estonian Collocation Dictionary database. In Kosem, I., Jakubiček, M., Kallas, J., Krek, S. (eds.) *Electronic lexicography in the 21st century: linking lexical data in the digital age. Proceedings of the eLex 2015 conference, 11-13 August 2015, Herstmonceux Castle, United Kingdom.* Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd., pp. 1-20.

Klubička, F. & Ljubešić, N. (2014). Using crowdsourcing in building a morphosyntactically annotated and lemmatized silver standard corpus of Croatian. *Language technologies: Proceedings of the 17th International Multiconference Information Society IS2014*. Ljubljana.

Kosem, I., Gantar, P. & Krek, S. (2013). Automation of lexicographic work: an opportunity for both lexicographers and crowd-sourcing. In *Proceedings of eLex 2013*, pp. 33–48.

Krek, S., Kosem, I., Gantar, P. (2013). *Predlog za izdelavo Slovarja sodobnega slovenskega jezika* (A Proposal for a new Dictionary of Contemporary Slovene). Version 1.1. Accessed at: http://trojina.org/slovar-predlog/datoteke/Predlog_SSSJ_v1.1.pdf.

Lanxon, N. (2011). *How the Oxford English Dictionary started out like Wikipedia.* http://www.wired.co.uk/news/archive/2011-01/13/the-oxford-english-wiktionary (Access: 25. 10. 2014)

Lew, R. (2013). User-generated content (UGC) in online English dictionaries. *OPAL - Online publizierte Arbeiten zur Linguistik.*

Oyama, S., Baba, Y., Sakurai, Y. & Kashima, H. (2013). Accurate Integration of Crowdsourced Labels Using Workers' Self-Reported Confidence Scores. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pp. 2554–2560.

Rumshisky, A. (2011). Crowdsourcing Word Sense Definition. In *Proceedings of the Fifth Law Workshop (LAW V). Portland: Association for Computational Linguistics*, pp. 74–81.

Rumshisky, A., Botchan, N., Kushkuley, S. & Pustejovsky, J. (2012). Word Sense Inventories by Non-Experts. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC'12.* Istanbul, Turkey.

Sabou, M., Bontcheva, K., Derczynski, L. & Scharl, A. (2014). Corpus Annotation through Crowdsourcing: Towards Best Practice Guidelines. In *Proceedings of LREC 2014*, pp. 859–866.

Snow, R., O'Connor, B., Jurafsky, D. & Y Ng, A. (2008). Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing.* Association for Computational Linguistics, pp. 254–263.

Tavčar, A., Fišer, D. & Erjavec, T. (2012). sloWCrowd: orodje za popravljanje wordneta z izkoriščanjem moči množic. In *Proceedings of the Eighth Language Technologies Conference. Ljubljana: Jožef Stefan Institute*, pp. 197–202.

Tiberius, C. & Schoonheim, T. (2014). The Algemeen Nederlands Woordenboek (ANW) and its Lexicographical Process. In V. Hildenbrandt (ed.): *Der lexikografische Prozess bei Internetwörterbüchern. 4. Arbeitsbericht des wissenschaftlichen Netzwerks „Internetlexikografie".* Mannheim: Institut für Deutsche Sprache. (OPAL – Online publizierte Arbeiten zur Linguistik X/2014). Preprint accessed at: http://www.elexicography.eu/wp-content/uploads/2014/05/TiberiusSchoonheim_The-ANW-and-its-Lexicographical-Process_Preprint.pdf

von Ahn, L., Kedia, M. & Blum, M. (2006). Verbosity: a game for collecting common-sense facts. Zbornik konference *SIGCHI conference on Human Factors in computing systems.* ACM, pp. 75–78.

Żmigrodzki, P. (2014): Polish Academy of Sciences Great Dictionary of Polish [Wielki słownik języka polskiego PAN]. *Slovenščina 2.0*, 2 (2), pp. 37–52.