

Andrea Abel and Christian M. Meyer

8 User Participation



Fig. 8.1: Direct, indirect, and complementary forms of political participation.

Andrea Abel, Free University of Bozen-Bolzano, Faculty of Education, Regensburger Allee 16/Viale Ratisbona 16, 39042 Brixen/Bressanone, Italy, e-mail: andrea.abel@unibz.it; EURAC Research, Institute for Applied Linguistics, Drususallee 1/Viale Druso 1, 39100 Bozen/Bolzano, Italy, e-mail: andrea.abel@eurac.edu

Christian M. Meyer, Technische Universität Darmstadt, Ubiquitous Knowledge Processing Lab, 64289 Darmstadt, Germany, e-mail: research@chmeyer.de

Democracy lives in the participation of citizens, who can contribute to the politics of a state, both directly by electing representatives and also indirectly by expressing their ideas, suggestions, or wishes. In addition to these ways of directly and indirectly shaping politics, people inform themselves about political happenings and exchange their views about them with one another. For dictionaries, a similar breadth of options exists for participation: users swap ideas with those providing the dictionaries and among themselves. They formulate requests, give feedback, or play an active role in creating dictionary entries.

The Internet is increasingly shaping our society and connecting us ever more strongly with one another. By the mid-1990s, the use of so-called social media technologies, such as blogs, wikis, or social networks, had changed the Internet from a repository of curated expert information into an interactive platform for exchanging user-generated content. In this changed environment, referred to as “Web 2.0”, Internet dictionaries increasingly involve dictionary users in lexicographic activities. The degree of user participation ranges from user-driven compilations of entire dictionaries and feedback about the quality of individual entries to dialogues between lexicographers and users or among users themselves.

These new forms of user participation on the Web have hardly been researched. In this chapter, we provide an overview of the different possibilities for involving dictionary users directly, indirectly, or in accessory ways when compiling a dictionary. Above all, the assessment of quality and clarification of legal issues are of paramount interest in order to be able to evaluate the potential of user-generated content. In addition to a systematic categorisation of user contributions, we discuss several practical examples of individual organisational forms and examine possible motivations for active involvement. Furthermore, we seek to prompt a critical discussion about the strengths and weaknesses of different forms of user participation.

8.1 Introduction

The participation of users in compiling a dictionary is hardly a new topic in lexicography, but dates back to the pre-electronic era. Already in the 19th century, the OXFORD ENGLISH DICTIONARY (OED) established reader programmes in which the public was asked to read books, collect examples of the customary ways in which a word was used, and then send them in. Initially, these examples were unsystematic but they were then compiled more and more deliberately by prescribing lists of words, literature, and particular thematic areas for the volunteers (cf. Thier 2014).

However, the development of the Internet and the World Wide Web permitted completely new ways for involving dictionary users so that the question of user participation has become an increasingly important factor in the planning, development, and

use of a dictionary. For one thing, the Web allows dictionary users to communicate with each other, which was a very laborious and time-consuming activity before – if at all possible. For another, it offers totally new possibilities for interaction, putting users in a position to compose dictionary entries independently and to improve them collaboratively. This user-driven creation of dictionary content represents a fundamental change in the lexicographic process (→ Chapter 3.4.2). Carr (1997: 214) describes this method as “bottom-up lexicography” since dictionaries are assembled “bottom up” from individual entries and user contributions into a whole. By contrast, in the traditional model of commercial or academic lexicography, dictionaries are systematically planned as a whole and then compiled “top down” by expert lexicographers.

The new forms of user participation mean that the boundary between dictionary users and dictionary editors is becoming increasingly blurred. In this context, Lew (2014: 9) proposed the portmanteau “*prosumer*” since a user is both actively involved in compiling the dictionary as a “*producer*” and interested in the compiled content as a “*consumer*”.

Research into user participation in Internet dictionaries is still a fairly new topic area. For example, Wiegand et al. (2010: 17) observed:

Allerdings sind die lexikographischen Prozesse, wie man sie bei der Entstehung von gemeinschaftlich erstellten Online-Wörterbüchern, wie dem Wiktionary, beobachten kann, mit der traditionellen Phaseneinteilung nicht mehr adäquat beschreibbar; ihre Abläufe sind bislang auch erst ansatzweise erforscht. [In any case, the lexicographic processes that we can observe in the creation of collaboratively compiled online dictionaries like Wiktionary can no longer be adequately described with the traditional division of phases; as yet, research into the way these processes work is only rudimentary.]

Storrer/Freese (1996) and Storrer (1998) were among the first to attempt a classification targeted at correcting errors, suggesting new headwords, obtaining expert contributions on certain topics, and collecting contributions by laypeople. Köhler Simonsen (2005) distinguished between *active user involvement*, by which he means feedback on the design and development of a dictionary with the help of surveys or test groups, and *lexicographic democracy*, which he describes as feedback on final articles (e.g. error corrections). While this definition is limited to giving feedback to the editorial staff, Fuertes-Olivera (2009) used the term *democratisation* in a different way, focussing on “*collective free multiple-language Internet dictionaries*” (ibid.: p. 101) such as WIKIPEDIA and WIKTIONARY, which are compiled entirely by users without editorial control. Storrer (2010) introduced a similar distinction between user contributions controlled by professional editors and those created by the users themselves in a collaborative effort. Further works by Lew (2011, 2014) suggested a more detailed classification of dictionaries that allow for direct user contributions, additionally introducing *collaborative-institutional dictionaries* and dictionaries making use of external *user-generated content*. Melchior (2012, 2014) used the term *semi-collaborative dictionary* for his analysis of the LEO dictionary portal as being supported by users rather than generated by users.

In a different strand of research, de Schryver/Prinsloo (2001) coined the term *fuzzy simultaneous feedback* to point out user feedback which is available during the development of a lexicographic product. For electronic dictionaries, de Schryver/Joffe (2004) focussed primarily on log file analyses, which are a way of exploring a user's behaviour and demands without additional effort by the users (→ Chapter 1.3).

Insight into the variety of forms of user participation is of great relevance. Dictionary entries and background material that are contributed voluntarily as well as feedback on new and existing content have the potential to speed up the production of a dictionary, enhance its quality, and enrich its content. Publishers can save money, and users can acquire knowledge about the structure of a dictionary and its use, thereby identifying more strongly with the product. Conversely, assessing user contributions often means more work when a dictionary is being compiled. A large number of poor quality or inappropriate contributions can also lead to disruptions in the lexicographic process or to wrong and inconsistent lexicographic products.

The aim of this chapter is to describe the different types of user participation and organise them systematically. We discuss several practical examples for each type of participation, examining in particular the motivation of users, legal questions, quality issues, and the processes for submitting user contributions.

At the highest level, we distinguish between three basic types of user participation, which we shall consider in more detail below:

1. Direct user participation (→ Section 8.2);
2. Indirect user participation (→ Section 8.3);
3. Accessory user participation (→ Section 8.4).

This categorisation and the corresponding descriptions are based on Abel/Meyer (2013, 2016) and Meyer/Abel (2017), taking as their starting point the reflections by Mann (2010). → Tab. 8.1 gives an overview of the three types of user participation and their characteristics. As we shall see, this does not prevent multiple forms of user participation from being used in parallel within a single dictionary project. Our categorisation is suitable for describing user participation in individual Internet dictionaries (e.g. OED ONLINE, DUDEN ONLINE, WIKTIONARY) and dictionary portals (cf. Storrer 2010;

Tab. 8.1: Overview of different types of user participation.

Direct user participation:	<ul style="list-style-type: none">– Contributions to open-collaborative dictionaries– Contributions to collaborative-institutional dictionaries– Contributions to semi-collaborative dictionaries
Indirect user participation:	<ul style="list-style-type: none">– Explicit feedback– Implicit feedback
Accessory user participation:	<ul style="list-style-type: none">– Exchanges between dictionary makers and dictionary users– Exchanges between dictionary users

Engelberg/Müller-Spitzer 2013; → Chapter 2), i.e. interfaces that permit access to a whole series of dictionaries (e.g. LEO, DICT.CC).

8.2 Direct user participation

Direct user participation denotes the compiling, changing, or deleting of dictionary entries by dictionary users. We distinguish between contributions to open-collaborative dictionaries, collaborative-institutional dictionaries, and semi-collaborative dictionaries. These categories will be described in more detail below.

User contributions to open-collaborative dictionaries are not subject to any editorial supervision by a fixed group of experts. Rather, the dictionary is based on entries composed and revised by a potentially unlimited number of volunteer users. All changes are directly visible in the dictionary and can, therefore, be immediately examined by other users and, if necessary, revised again. The collective knowledge of the participants – frequently referred to as “swarm intelligence” (cf. Krause/Krause 2011, Rosenberg 2015), as “collective intelligence” (cf. Malone et al. 2010), or as the “wisdom of crowds” (cf. Suro-wiecki 2005) – takes the place of expert knowledge. The basic assumption of this approach is that the different subjective perspectives and knowledge of the individuals involved is consolidated into a communal group dynamic and is thereby bound together into a larger whole. The open-collaborative process has especially gained popularity through the free online encyclopaedia WIKIPEDIA and its sister project WIKTIONARY. With versions in 168 languages and a total of 38.7 million dictionary entries, WIKTIONARY is currently the largest open-collaborative dictionary.¹

Malone et al. (2010) distinguish between economic factors (direct financial advantages, future earning potential, and practising personal competences), pleasure (enjoyment, altruism, sociability), and reputation (recognition by peers) as fundamental motivational factors behind contributing to open-collaborative resources. Other studies that deal with the possible driving forces behind contributions to online communities point to similar, and also wider, factors, such as acquiring and exchanging information, identifying with particular groups, or a sense of belonging (cf. Lampe et al. 2010, Rafaeli/Ariel 2008).

The contents of the dictionary are not tied to a particular provider or publisher, and many such resources use free licences, also known as open content licences, such as the CREATIVE COMMONS series of licences through which – unlike in the classic copyright model – the distribution of content in unchanged form is generally possible for anyone as well as, to some extent, commercial use and modification – depending on the specific licence (cf. Kreutzer 2011).

¹ https://meta.wikimedia.org/wiki/Wiktionary/List_of_Wiktionaries [last access: March 22, 2024].

In addition to the licence under which the contents are published, the source or origin of the contributions is also a relevant issue. Uncovering and preventing the copying of dictionary information from other works protected by copyright pose great challenges for the providers of collaborative dictionaries. Plagiarism is hardly a new phenomenon in lexicography (cf. Hausmann 1989), where, on the whole, transcribing or copying from existing dictionaries seems to have been a common and long known practice, albeit one that has been judged differently in different contexts (cf. Landau 2001). However, this aspect has to be considered anew given the high number of participants in collaborative resources. The user communities around WIKIPEDIA and WIKTIONARY have defined comprehensive guidelines and workflows in relation to copyright issues.² Here, the attempt is made to provide as much information as possible through references to sources. Data of questionable origin are first put up for discussion or, possibly, removed from the dictionary.

At the same time, there were phases when large bodies of lexicographic data or entire dictionaries from sources that are freely licensed or whose copyright had already expired were integrated into WIKTIONARY. Hanks (2012) and Hanks/Franklin (2019) noted that numerous outdated meaning paraphrases were found in the English WIKTIONARY that could be traced back, in particular, to the adoption of sources that were copyright free. For example, some parts of the digitised version of WEBSTER'S REVISED UNABRIDGED DICTIONARY (WEBSTER) from 1913 were used in the English WIKTIONARY, with sometimes negative consequences for the quality of entries. Lew (2014) demonstrated this in relation to the English verb *handle* in WIKTIONARY, for which an uncommon interpretation in contemporary English was listed as the first meaning (i.e. "To use the hands"). This was one result of copying content from the old edition of WEBSTER, along with numerous uncommented archaic quotations from the bible or classical literature that were provided as lexicographic examples. In the meantime, the entry *handle* has been changed: The meaning is no longer in first place, but it is still there and still with a quotation from the bible.³

A large proportion of open-collaborative dictionaries is based on clearly prescribed lexicographic instructions that describe, at the macrostructural level, the choice of headwords to be included and, at the microstructural level, the construction of dictionary entries and the information classes to be included in them, such as pronunciation, meanings, or example sentences. The URBAN DICTIONARY, for example, focuses on English colloquial language and nonce words. It collects the associated information through an online form that permits the headword to be entered, along with an explanation of its meaning, an example of usage, and further freely selected tags (such as synonyms, misspellings, etc.). SPRACHNUDEL is a German equivalent that also covers slang and neologisms.

2 See <https://en.wiktionary.org/wiki/Wiktionary:Copyrights> [last access: July 28, 2023].

3 <https://en.wiktionary.org/wiki/handle#Verb> [last access: March 28, 2024].

Many of the collaborative dictionaries with fixed lexicographic instructions are translation dictionaries such as BAB.LA or GLOSBE, whose entries are often very simply structured. For these dictionaries, the input form only covers the headword in the source and target language, which makes participation possible without great effort. Because of the large variety of language pairs and their wide potential user community, these kinds of bilingual or multilingual dictionaries benefit particularly from direct user contributions (cf. Meyer/Gurevych 2012).

More comprehensive input forms are required for more complex entry structures in order to capture all of the classes of information, for example separate text fields for meaning paraphrases, usage examples, sources, a tabular input for synonyms and translations, or a selection field for the word class. For instance, the multilingual dictionary project KAMUSI allows explanations of meaning to be given in multiple languages. The input forms are adapted for exactly this case, and compiling or editing an entry is modelled as a multi-stage, interactive process. → Fig. 8.2 shows an extract from these input options. Benjamin (2014) discusses contributions to KAMUSI and also discusses the challenges of such a complex article structure, such as users frequently typing information in the wrong field.

Key Information * **Word Forms and Origins** **Translations and Concepts** **Related terms within the same language**

Concept

Term (lemma) * Term Language * Headword * (?) ☐

Part of Speech * Part Attribute

Definition

Definition of the term in the language to which it belongs (?) Definition Source URL

Definition Translations

Translated Definition Language

Fig. 8.2: Extract from the input forms in KAMUSI.

Beside these form-based input options, we can also find dictionaries whose entries are composed in a markup language. Here, the dictionary content contributed is not distributed across several predefined fields belonging to particular information classes but is rather written using a specific syntax that defines the formatting (e.g. bold face, italics, coloured text) and logical structure (e.g. headlines or the beginning and end of a meaning explanation) for individual information classes.

The XML markup language that is often employed for expert-built lexicographic products was already introduced in → Chapter 4. XML and the dictionary standards based on it, such as the Text Encoding Initiative (TEI)⁴ and the Lexical Markup Framework (LMF; Francopoulo 2013), make it possible to represent very complex lexicographic data structures. However, this expert markup requires a high degree of knowledge to master it. As a result, dictionary standards of this type are predominantly aimed at professional lexicographers and are hardly ever employed for collaboratively compiled dictionaries.

Instead, these dictionaries are often based on wiki technology, which provides fairly simple ways of writing and revising content. A wiki is an online platform through which texts can be collaboratively composed and edited by users themselves. The texts are written with the help of a special markup language, the so-called wiki markup, which provides both the established formatting options (e.g. bold face) and the definition of reusable text blocks (e.g. a table of inflected word forms) while also being easy to learn. Above all, the wiki markup language should reduce the inhibitions of users who are less comfortable with technology. The English-language RAP DICTIONARY is an example of a wiki-based dictionary of this kind; its entry for *Cheeser* is structured as follows in the Wiki markup language:⁵

```
===noun===
```

```
'''Cheeser'''
```

```
A person that trys to become closer to you using all ways for the
purpose of having your money.
```

```
''Becareful of the cheesers, the teasers''-- [[ Grand Pupa ]](Song:I
like It, Album: 2000 - 1995)
```

```
[[ Category:Terms ]]
```

```
[[ Category:Nouns ]]
```

⁴ <https://tei-c.org/> [last access: September 9, 2023].

⁵ <https://web.archive.org/web/20140429154439/http://www.rapdict.org/Cheeser> [last access: March 29, 2024].

The text set within three equals signs produces a heading. Italics are activated by two inverted commas and bold text by three inverted commas. Cross-references to other headwords can be marked with square brackets, as are classifications of the entry into the categories “terms” and “nouns”.

In contrast to form-based input, markup languages make it possible to define, organise, and position lexicographic information freely. Thus, wiki-based dictionaries are not limited to particular, pre-defined lexicographic instructions but instead allow participants to determine these and change them themselves, thereby becoming involved in the planning and conceptual development of the dictionary. Matuschek et al. (2018) compare OMEGA WIKI and WIKTIONARY, two open-collaborative dictionaries with more rigid vs more flexible microstructures and with prescribed vs variable lexicographic instructions. Here, it becomes clear that a more flexible approach, like that in WIKTIONARY, offers noticeably more organisational options for an entry, for example, through hierarchical division of the explanations of meanings for entries with many possible meanings. At the same time, inconsistencies between the various entries arise very easily in this kind of dictionary, which can, in turn, hinder efficient use of the dictionary.

Since direct contributions to collaborative dictionaries are not checked by experts, we find two types of quality issues: first spam and vandalism; and second descriptions that are vague, incorrect, old-fashioned, too general, and/or too complicated. By spam and vandalism, we mean nonsensical and crude content, such as clearly false information, swearwords, or derogatory comments in existing texts as well as deleting actually useful dictionary content without reason or without making at least a correction. As a result, there is a need for quality control mechanisms, particularly in large dictionary projects. In the German WIKTIONARY, for example, individual stages of a dictionary entry are marked as so-called flagged revisions once a certain quality standard has been reached.⁶ While the label is only aimed at the absence of spam and vandalism, there were discussions as to whether a second label should be assigned for reaching minimum quality requirements in relation to the second type of quality issues. However, defining these requirements is notably more difficult than for spam and vandalism, and the questions of quality and defining quality criteria are constant points of discussion even for expert-built dictionaries (cf. Kemmer 2010).⁷

Only active users who had worked on at least 200 entries were assigned the right to flag revisions in the German WIKTIONARY. This prevented the label from being misused. So-called “construction site” labels were a further measure to ensure quality. Anyone who noticed a quality issue in an entry but who could not or did not want to

⁶ https://en.wikipedia.org/wiki/Wikipedia:Flagged_revisions/fact_sheet [last access: March 22, 2024].

⁷ The extent of research still needed, in particular in relation to Internet dictionaries, was shown by the workshop “Was ist ein gutes (Internet-)Wörterbuch? – Alte und neue Fragen zur Qualität lexikographischer Produkte im ‘digitalen Zeitalter’” [What is a good (Internet) dictionary? – Old and new questions on the quality of lexicographic products in the ‘digital age’] at the XVI International EURALEX Congress in 2014 (cf. <http://internetlexikografie.de>).

correct the issue themselves was able to assign a pre-defined text block to draw attention to missing sources, necessary or useful additions, or inconsistent structure, among other things. Other contributors could then revise the entry or further discuss the quality issue. If an entry was judged to be unsuitable or irrelevant for the planned dictionary content, a separate label could be used to suggest the deletion of the entire entry (cf. the entry on the plural form *Erdoberflächen*⁸ which was flagged for some time due to there being no plural for this lemma). Meyer/Gurevych (2014) discuss quality control measures in collaborative online dictionaries in more detail.

Contributions to collaborative-institutional dictionaries constitute a second kind of direct user participation (cf. Lew 2011). These collaborative-institutional dictionaries used to be provided by established publishers; examples include the former MERRIAM-WEBSTER OPEN DICTIONARY and MACMILLAN OPEN DICTIONARY projects. The user contributions in these dictionaries mostly took the form of complete, submitted dictionary entries that were checked by the editors of the dictionary for vandalism, insults, or defamation. In contrast to semi-collaborative dictionaries and explicit feedback (see below), the contributions to collaborative-institutional dictionaries remained by and large unedited. At the same time, however, there is a close connection to these two types of user contributions, which we shall consider in more detail in what follows. One difference to open-collaborative resources is that users had no way of changing or deleting someone else's contributions.

While contributors to these dictionaries presumably had similar motivations to those for open-collaborative dictionaries, providers of collaborative-institutional dictionaries had two aims: first, to gather suggestions for preparing professionally and editorially compiled dictionaries; and, second, to advertise the publishers' own activities and products. The contributions could be collected without precise guidelines on the type and scope of the entries, as with the MACMILLAN OPEN DICTIONARY, or with a particular section of language in mind, as was intended for youth language in the former Duden SZENESPACHENWIKI. Since collaborative-institutional dictionaries were, for the most part, accompanied by dictionaries provided by professional editorial teams, they tended to contain entries that were not included in those expert-built resources. Hence, the resulting dictionaries were usually smaller than open-collaborative dictionaries. The MACMILLAN OPEN DICTIONARY, for example, contained about 11,700 entries in 2023, the year when it closed down⁹

Unlike open-collaborative resources, collaborative-institutional dictionaries mostly did not use free licences to publish their contributions. The rights of use remained either entirely with the contributing user or they were transferred entirely, or in part, to the dictionary.

⁸ <https://de.wiktionary.org/w/index.php?title=Erdoberfl%C3%A4chen&oldid=3753791> [last access: March 22, 2024].

⁹ https://web.archive.org/web/20230216132652/https://www.macmillandictionary.com/open-dictionary/index-chronological-order_page-1.htm [last access: March 28, 2023].

Given that well-known collaborative-institutional dictionaries such as the ones listed above have been closed down and we are not aware of any active projects of this kind, this indicates that this type of direct user participation was not successful. This can be due to the economic situation of the institution, the costs of running such a service, the quality of the submissions, or the amount of work that would be needed to make professional use of these contributions (i.e. considering them as explicit feedback for an expert-built dictionary, as discussed in the next section).

Contributions to semi-collaborative dictionaries constitute the third kind of direct user participation. They are carefully checked by professional lexicographers or other language experts before being integrated into the dictionary. The TECHDICTIONARY, for example, is based on contributions on topics related to computers and technology that are written by users and only included in the dictionary after being checked. LEO, a portal of 12 bilingual dictionaries, is a prominent example of semi-collaborative dictionaries. Its central components are translation entries contributed by users as well as lists of words, terminology, and glossaries donated to the portal. After being carefully checked, the contributions are generally added directly to the dictionary but are not substantially revised, as is the case with explicit feedback (→ Section 8.3). Nevertheless, the decision whether, and how, to include a contribution in the dictionary always remains the responsibility of the dictionary publishers, so that quality control and a consistent dictionary structure are made possible.

In semi-collaborative dictionaries, the rights of use are either transferred to the providers of the resource, which is usually the case with commercial providers (e.g. LEO), or are channelled into a dictionary with a free licence, as is the case, for example, in the semi-collaborative synonym dictionary OPENTHESAURUS.

While these kinds of resources often enjoy high numbers of visitors, Naber (2005) found that only a small proportion of the registered users were actually actively involved in the writing of entries in the case of the synonym dictionary OPENTHESAURUS and that most contributions represented newly suggested synonyms, even though alterations and deletions would also be possible. Similar findings were reported for WIKIPEDIA (cf. Javanmardi et al. 2009) and WIKTIONARY (cf. Meyer 2013). This kind of distribution, with extremely few very active users, on the one hand, and a very high number of users who only make a small contribution, on the other, can be found with virtually all types of user participation. This distribution can be described as a power law, which became very familiar in linguistics as Zipf's Law, for example, also in relation to the distribution of word frequency in corpora. Furthermore, it is well known that online communities have high numbers of lurkers, that is, members who only observe, without being active, for example, by making contributions or revisions (cf. Rafaeli/Ariel 2008 for a summary).

What is common to all three types of direct user participation is that user contributions are integrated directly into the dictionary. This mode of compiling dictionaries is referred to as *collaborative lexicography*. The dictionaries discussed benefit particularly from the diversity of the participants, which is, in principle, high. This applies both to the areas of knowledge covered and the forms of use of the linguistic units represented

by the participants. For language varieties (e.g. youth language, technical languages, dialects) and translation dictionaries, this provides clear additional value (cf. Meyer/Gurevych 2012).

Direct user participation as has been described in this section has only become possible with the advent of Internet dictionaries and the corresponding technology since user contributions are based on the new possibilities for interaction available on the Internet. By contrast, previous options for user participation almost exclusively involved forms of indirect participation, which we consider in more detail in the next section.

8.3 Indirect user participation

Indirect user participation denotes feedback from dictionary users on existing or missing lexicographic content, on the use of the dictionary, and on the dictionary project as a whole. Among its characteristic forms are suggestions, additions, corrections, requests and opinions, externally generated content, and dictionary usage data. What is common to all of these user participation forms is that the dictionary users have no possibility of directly changing the dictionary but only the possibility of effecting an indirect change through their feedback. In the rest of this section, we distinguish between explicit and implicit feedback as the two main forms of indirect user participation.

Explicit feedback refers to contributions that users express explicitly and that they intentionally make available to the dictionary providers, e.g. suggestions for new words, corrections of errors, or comments on the organisation or presentation of the entries. Such contributions may address both new and existing dictionary content.

Submitting explicit feedback is popular, and the motivations for engaging as an active user are similar to those for direct contributions. Above all, the motivational factors referred to by Malone et al. (2010) as pleasure and reputation play a large role. In an online survey on DUDEN ONLINE, Rautmann (2014) noted that just under half of the respondents were interested in feedback options for dictionary entries. For OED ONLINE, Thier (2014: 70) found:

Die Beiträge stammen bei weitem nicht nur von Akademikern, sondern von Menschen aus allen Teilen der Bevölkerung, die sich für ihre Sprache interessieren. [The contributions come very much not only from academics but also from people from all sections of the population who are interested in their language.]

When it was launched, for example, DUDEN ONLINE offered a button for “Suggested Words” through which users could propose new headwords to be included.¹⁰ However, additions and corrections can still be submitted by email. Rautmann’s (2014) analysis showed that more than half of the words suggested in this way met the inclu-

¹⁰ This function is no longer available.

sion criteria of the dictionary and were envisaged to be included in a new edition, e.g. *Burgerbude* (“burger stall”). Overall, the quality and usefulness of the explicit feedback on DUDEN ONLINE was perceived by Duden’s editors as predominantly positive (cf. Rautmann 2014).

We already mentioned in the introduction that submitting additional material has a long tradition at the OED, especially concerning examples of attestation. Since the mid-19th century, reader programmes have involved volunteers being encouraged to send in citations. The “Wordhunt” campaign between 2007 and 2008 and the “Science Fiction Citations” initiative (cf. Thier 2014) constitute more recent examples. The “Wordhunt” involved a BBC television programme in which viewers were encouraged to submit examples of words from a list that could be dated to an earlier point in time than given in the dictionary. By contrast, the “Science Fiction Citations” call is framed more openly; although it is no longer an official OED project, it still aims to receive submissions of examples of any concepts from science fiction literature.¹¹ The OED has continued with participatory campaigns in the recent past. As part of the “OED M-R antedatings” initiative launched in 2020, members of the public should find the earliest possible evidence for dictionary entries in the alphabetical range from M to R and submit their findings via an online form.¹²

In addition to new and supplementary information, user-driven assessments of quality are also included in the form of explicit feedback. For example, the DICT.CC Internet dictionary asks users to judge the accuracy of translation equivalents. Questionable equivalents and their word class are displayed on the screen and users are able to choose between “YES (100% correct)” und “NO/MAYBE” or skip to the next translation without making a decision. For example, → Fig. 8.3 shows the translation *loodering* – *heftige Prügelei*. In order to integrate only high-quality translations in the dictionary, the labels on the buttons have been chosen so that translations are only marked as correct if the user is certain about their decision (“YES (100% correct)”).

Is this translation correct and useful?	
English	loodering [Scot.] [Irish]
German	heftige Prügelei {f}
Class	noun
Comment	<p>• http://www.oed.com/view/Entry/... "loodering n. a severe beating; a hiding.</p> <p>1866 W. Gregor Dial. Banffshire (Philol. Soc.) 227 Lloutheran, the act of beating with severity; a severe beating. 1912 J. Campbell Judgment i. 9 She'll get over it. It's not the first time she's got a loodering. 1985 L. Shannon in Stories (Boston) No. 12 14 Whenever she walks through that door she'll get the biggest loodering she's ever had in her whole life.</p>
<div>YES (100% correct) NO / MAYBE ... Skip » » Guidelines</div>	

Fig. 8.3: Evaluation of quality in dict.cc.

11 <https://sfdictionary.com/about> [last access: March 28, 2024].

12 <https://pages.oup.com/ol/cus/1646166399178702002/oed-m-r-antedatings> [last access: March 22, 2024].

It is not unusual to find this kind of evaluation task in the field of (paid) crowdsourcing, a common strategy of companies to outsource certain tasks to volunteer participants on the Internet and thereby benefit from the “wisdom of many” or “crowd intelligence” (→ Section 8.2). Reviewing a newly developed product or online service is one example of this kind of evaluation task, often described as a Human Intelligence Task (HIT) since those asked are bringing their intuitions and intelligence to bear on solving the task, which would be impossible or difficult to complete with a machine. For example, businesses set out HITs in which participants have to indicate the best shop category for a particular product that is perhaps difficult to categorise. Designers of user interfaces can use HITs to test whether, for example, the colour is felt to be pleasant and whether users are able to find their way around quickly. Equally, product developers can survey a wide user group to assess the importance of particular product features. Crowdsourcing is also used in the field of computational linguistics research to generate training and evaluation data on, for example, whether an automatically created summary has been successful or not. In order to find participants for these kinds of tasks, the HITs are posted on crowdsourcing platforms like CROWDFLOWER or AMAZON MECHANICAL TURK and remunerated with small sums of money (e.g. USD 0.05; cf. Fort et al. 2011). From the perspective of dictionary research, we can consider not only quality evaluations but also user research questioning (→ Chapter 9) as crowdsourcing activities. As far as we are aware, though, crowdsourcing platforms have not yet been used for this kind of questioning.

However, the basic idea behind crowdsourcing is not limited to paid evaluation tasks. In the broadest sense, all forms of “crowd intelligence” can be understood as crowdsourcing, including the volunteer contributions in collaborative dictionaries (→ Section 8.2). A particular form of crowdsourcing is crowdfunding, a way of fundraising on the Internet, in which a project is intended to be financed by small payments from as many users as possible (cf. Howe 2008). In the field of dictionaries, crowdfunding could be used to finance new dictionaries or existing active dictionaries that are under development. Meyer/Gurevych (2014) discussed this form of user participation with the example of NITTY GRITS: a crowdfunding campaign run by the Southern Food and Beverage Museum was intended to raise the resources necessary to revise a dictionary of food and culinary terms in order to make it the definitive International Culinary Dictionary, but this was not achieved.¹³

We distinguish form-based feedback, where user submissions take place through an online form with fixed, pre-determined fields, and free-text feedback, where no further restrictions on the form of the feedback are provided, such as an email with an arbitrary text. To a certain extent, form-based feedback makes it possible to guide the type and volume of submissions received. For example, the LEO dictionaries provide

¹³ <https://www.indiegogo.com/projects/nitty-grits-the-international-culinary-dictionary/#/> [last access: October 24, 2023].

different forms for corrections and for suggesting new entries or translations, meaning that submissions from users are pre-sorted and relevant details can be requested in a targeted way. The forms used by LEO contain easily understood fields and can therefore be completed by contributors with little effort. It should also be noted that the suggestions are submitted in the forum area so that other users can also comment on or add to the suggestions. The OED ONLINE uses one single form, which is shown in part in → Fig. 8.4. Detailed information can be requested in a form of this kind (cf. Thier 2014), e.g. the bibliographic data of the sources indicated. Complex forms can, however, inhibit users. As a result, fewer, but possibly more accurate, user contributions can be expected when using complex rather than simple forms. For dictionary providers, this can be a way to steer the quality and volume of user feedback.

In addition, the OED ONLINE offers the option to submit free-text feedback by post or email. These contributions first have to be checked and categorised by the editors and so sometimes must represent considerable additional work. When it comes to error corrections, for example, they must check whether the problem listed can actually be found in the dictionary with the information provided. However, the free-text feedback evaluated by DUDEN ONLINE shows that the majority of submissions were useful for the editorial work on the dictionary (cf. Rautmann 2014).

As well as providing explicit feedback on particular dictionary entries, users can comment on the dictionary as a whole. This includes both content-related aspects (e.g. the choice of headwords) and layout or organisational aspects. Melchior (2012: 359–367) analysed these kinds of user submissions for the LEO German-Italian dictionary and characterised eight different types of users on this basis. Tensions arise when different types of users come into contact with one another, for example, users who wish for neologisms and nonce words to be included promptly and users who view the dictionary as a “moral compass”, demanding that vulgar expressions are removed.

Feedback on the structure and organisation of a dictionary can also be sought by lexicographers by publishing beta or advance versions (cf. Melchior 2014). This enables different layout versions to be tested at the same time or one after the other, without compromising access to the actual dictionary. This kind of beta version was, for example, made available for the DIGITALES WÖRTERBUCH DER DEUTSCHEN SPRACHE (DWDS) (cf. Klein/Geyken 2010).

The boundary between direct contributions to semi-collaborative dictionaries and indirect contributions in the form of explicit feedback is fluid. For example, the submission of a new translation to one of the LEO dictionaries can be integrated into the dictionary without extensive editorial work (as long as the translation is accurate). In this case, we are talking about a direct user contribution to a semi-collaborative dictionary. However, we count a citation submitted to supplement an entry in the OED ONLINE as explicit feedback since the submission is neither a complete dictionary entry nor will the submission be immediately integrated into the dictionary. Rather, the editorial team has to decide whether the citation is relevant and informative for the existing

Fill out the form below to submit your contribution to the OED

What are you submitting?*

Word or phrase*

Part of speech

Pronunciation (please use IPA, tell us what it rhymes with, or link to a recording)

Definition or sense number as defined in the OED (e.g. 2.a.)

Quotation evidence. Each quotation needs: full quotation text; information about where you found the quotation (e.g. bibliographical reference, web link)

Is there anything else you'd like to add?

*Mandatory field.

Submit

Fig. 8.4: Input form for submitting examples to the OED Online.

entry, whether it can be verified, and in which form it can be integrated into the entry (e.g. which context is required).

In contrast to explicit feedback, *implicit feedback* arises without any input from individuals; it is often, in fact, unintentional and without dictionary users being aware that they are providing feedback. This kind of user contribution includes records about dictionary usage and external contributions that are integrated into a dictionary without being compiled specially for this purpose.

Records of dictionary usage are employed in dictionary research as an instrument to understand the behaviour of users and thereby adapt the dictionary more effectively to their information needs. Log data (→ Chapter 1.3) often form the basis of these kinds of analysis. These logs automatically capture every access to a dictionary entry along with the access date and time and potentially the retention time, search terms, and navigation history. Ready-made software solutions are available to analyse log data, e.g. GOOGLE ANALYTICS or MATOMO. Such tools process the raw log files and report the most frequently visited pages, the average time spent by users, and frequent navigational patterns. At the same time, the data protection requirements in the countries concerned always have to be taken into consideration when recoding and analysing log data.

This kind of evaluation is known in the context of DUDEN ONLINE for example (cf. Rautmann 2014). In the process, the Duden editors receive access to the list of the most frequently read entries. In addition to optimising the dictionary towards the entries that are regularly consulted, log data can be used to improve access to dictionary contents. To achieve this, they are filtered to show unsuccessful searches so that the users' search strategies can be analysed more closely or potential gaps may be revealed. It has been shown, for example, that expressions of more than one word, such as *im Folgenden* ('in the following') or *des Weiteren* ('furthermore') are often entered into the DUDEN ONLINE search field. For reasons of space in print dictionaries, information on these constructions are primarily found in the examples section for the relevant lemma. However, high demand indicated by the log data analysis has prompted the editors to broaden their headword guidelines so that these frequently consulted multi-word expressions appear as separate dictionary entries in addition to the existing descriptions.

Some dictionaries provide returning users with a log-in screen in order to personalise their use of the dictionary, for example by being able to view a list of their own previous search requests. For these dictionaries, more extensive log data can be captured and user behaviour evaluated over a longer time period. Already in the early 2000s, profiles were generated for the ELEKTRONISCHES LERNERWÖRTERBUCH DEUTSCH–ITALIENISCH (ELDIT) based on user log-ins; these are characterised by the headwords and information classes that users consult (cf. Abel et al. 2003). A similar analysis was conducted for the BASE LEXICALE DU FRANÇAIS (BLF) in which the search and consultation behaviour of the participating users was analysed in addition to the headwords and multi-word expressions in their search requests. Among other things, this showed that users were mostly seeking information about meanings and grammatical gender, the latter being a typical problem for learners of French (Verlinde/Binon 2010).

However, analysing log data usually does not provide precise results: for one thing, access by automated computer programs and search engines cannot be filtered out well enough; for another, there is no exact record of reading time or whether an attempt to look something up was successful. For example, Verlinde/Binon (2010) observed that over 90% of the page visits were caused by automated search engines checking the website for new or updated content. However, these automated visits cannot always be distinguished clearly from a human visit to the page, which leads to so-called noise, that is, inaccuracies in measurement in the generated data. This noise can be reduced to a certain extent through automatic procedures to clean the data, but log data analysis is often criticised for being superficial and limited in its meaningfulness (cf. Müller-Spitzer/Möhrs 2008; Verlinde/Binon 2010). Newer works rely on data cleaning and statistical measures in order to analyse the relation of look-up frequency and corpus frequency (Müller-Spitzer et al. 2015; de Schryver et al. 2019).

At the same time, users' registered accounts supply dictionary providers with additional implicit feedback about the use of the dictionary. On MERRIAM-WEBSTER ONLINE users can, for example, mark individual dictionary entries as favourites; in this way the editors receive additional information on particularly popular entries so that they can cultivate and develop these. Similar options are available in DICTIONARY.COM and WORDNIK, where favourites can also be organised in user-defined lists. The title and composition of this kind of word list provide further information on users' needs and their behaviour when looking up words. In WORDNIK, for example, we can find a user-generated word list of about 3,000 academic terms, a list of 100 colour names and a list of about 600 words that a user has marked as "learned".¹⁴

Indirect user contributions are not limited exclusively to the dictionary itself but can also be drawn from external sources and displayed as part of dictionary entries. This form of implicit feedback involves external user-generated content. This external content includes messages or blog posts about a particular headword as well as illustrations, videos, and audio data that have been contributed by users to other online sites. For example, WORDNIK allows photographs from Flickr to be included in their dictionary entries (cf. McKean 2017: 473). As with direct user contributions, adhering to copyright for external user-generated content is also an important aspect of dictionary planning. For example, when incorporating Flickr images, WORDNIK indicates that the photographs are subject to a CREATIVE COMMONS licence. Avoiding inappropriate content is another important issue. Lew (2014), for example, discussed how inappropriate images were displayed in the retired GOOGLE DICTIONARY, which showed automatically retrieved illustrations from the Google image search in its dictionary entries until 2011. As external content is continuously changing and as the images were integrated into the display for a dictionary article in a fully automated way, it was

¹⁴ <https://www.wordnik.com/lists/academic-words-4>, <https://www.wordnik.com/lists/great-color-names>, <https://www.wordnik.com/lists/learned-words-1> [last access: March 29, 2024].

almost impossible to manually check whether they improve the lexicographical descriptions and respect copyright and social norms. This is why publishers declare limitations on liability when (external) user-generated content is used, e.g. Wikipedia.¹⁵ Another approach is using artificial intelligence methods to filter out unsuitable contributions. For example, Wang/McKeown (2010) employed language technology to detect vandalising changes in WIKIPEDIA. To do this, they modelled and automatically analysed different forms of vandalism with particular attention paid to syntactic features (e.g. syntactically incorrect sentences), lexical features (e.g. certain lexical elements, including Web slang like “LOL”, “haha”, etc., often accompanied by noticeable repetition of punctuation, such as “!!!!!!”, and comments on revisions), and semantic features, which is a particularly challenging task (e.g. words or word meanings that do not fit in the given context or are thematically unsuitable), as well as the editing history of individual authors.

8.4 Accessory user participation

Accessory user participation denotes exchanges between dictionary compilers and users or among dictionary users themselves. In this way, it describes a kind of integration that is located beyond the contents of the dictionary but focussing on the macrostructure (i.e. the selection and organisation of the lemmas) or microstructure (i.e., the organisation and format of individual dictionary entries).

If these are exchanges in which the dictionary compilers address the users and provide them with information, without a reaction being demanded or being possible, we can refer to *unidirectional communication*. Blogs represent a typical example of this kind of communication. For example, some dictionary publishers post blogs in which they report on interesting, surprising, or amusing topics about language use or language history. The MACMILLAN DICTIONARY blog, for example, used to contain a collection of dictionary-related resources that is now partially accessible on Macmillan Education’s website.¹⁶ In 2013, the publisher launched the rubric “Stories behind Words”,¹⁷ in which teachers, authors, linguists, and general language enthusiasts were asked to write about anecdotes or experiences relating to words. In this case, the publishers employed user contributions to address their audience in a unidirectional manner.

Blog contributions often contain hyperlinks to dictionary entries and are thereby intended to help advertise the publishers’ own products as well as to bind users and

¹⁵ https://foundation.wikimedia.org/wiki/Policy:Terms_of_Use#16._Limitation_on_Liability [last access: March 29, 2024].

¹⁶ <https://www.onestopenglish.com/adults/vocabulary/macmillan-dictionary-blog> [last access: October 24, 2023].

¹⁷ <https://www.onestopenglish.com/stories-behind-words/552993.article> [last access: October 24, 2023].

customers to their brand. Using newsletters or social networks like Facebook or microblogging services like X, formerly known as Twitter, to disseminate product information represents a similar approach. For example, the OED ONLINE – like other publishers – uses a whole spectrum of unidirectional communication options in order to reach users. Services including blogs, social media, and video platforms are driven by the marketing department but they make available content created by dictionary staff (Thier 2014).

Language games are another type of popular service offered by different publishers or institutions. In 2010, for example, the Dutch ALGEMEEN NEDERLANDS WOORDENBOEK (ANW) invited users to search for “the lost word” in their game “Het Verloren Woord”. Those interested received a series of cryptic descriptions at set intervals: for example, the phrase *niet vroeg* (‘not early’) led to the word *laat* (‘late’) and from this palindrome read backwards, the word *taal* (‘language’) had to be deduced (Schoonheim et al. 2012: 975). Here, participants were able to exchange ideas with other users and receive feedback from the organisers. However, in order to solve the task, it was necessary more than anything to use the dictionaries of the Instituut voor de Nederlandse Taal; this not only raised awareness of those dictionaries but also encouraged the use of the dictionaries in a playful way, thereby achieving an educational goal (Schoonheim et al. 2012).

In 2023, the Danish Dictionary DDO launched a quiz in collaboration with the magazine “DM Akademikerbladet”. Under the motto “Test dig selv: Fostår du ordbogens nye ord?” [“Test yourself: do you understand the dictionary’s new word?”] participants could playfully find out how well they understood Danish neologisms already included in the dictionary. Such initiatives not only contribute to the visibility of dictionaries, but also raise awareness of the fact that dictionaries adapt to changes in the language.¹⁸

In many cases, users also have the option to engage more actively in these forms of communication, for example, by commenting on an announcement, evaluating contributions or suggesting new topics, thereby helping publishers to orient their offer more effectively to the demand. If this kind of mutual exchange between dictionary makers and dictionary users takes place, we can talk of *bidirectional communication*. The boundaries between unidirectional and bidirectional communication are fluid in many ways since users may also respond to forms of unidirectional communication (e.g. in an email or phone call) and, likewise, there might be no response at all, even if bidirectional communication were technically possible.

Language advice services constitute a particular kind of bidirectional communication. Since the 1960s, the Duden editors have offered telephone help, providing further assistance on language-related questions to users who, for example, have been unable to find what they need in one of the publisher’s dictionaries. In keeping with the motto “There are no stupid questions! – Every question is answered”, users can

¹⁸ <https://www.akademikerbladet.dk/aktuelt/2023/november/test-dig-selv-forstaar-du-ordbogens-nye-ord> [last access: March 22, 2024].

direct language-related questions by email to an expert in the CANOONET language blog “Ask Dr. Bopp”, which moved to the LEO language blog in 2020.¹⁹ In addition to a personalised answer to the specific language question, recurring or interesting examples are often made available on the blog for a wider number of users. These resources offer useful insights into the information needs of users and, thus, can contribute to improving and adapting dictionary content. Furthermore, the expert answers in the particular case of Dr. Bopp often refer users to dictionary content or other content from the website so that these are indirectly promoted.

Accessory user contributions are not limited to communication between dictionary providers and users. Thanks to the technology of Web 2.0, the opportunities for users to communicate among themselves are also increasing. One popular option in this context is the forum in the LEO portal. If we take the German compound *Nutzerbindung* (literally: ‘user binding’) as an example, there was still no English translation given in the German-English LEO dictionary when this chapter was first written.²⁰ One user posted their query about a suitable equivalent in the forum, describing the meaning of the term in German as follows: “It means binding users to a website (e.g. with an interesting offer) and motivating them to return to the website”. The user wanted to know if the literal translation “user binding” could be used in English. In response, another user suggested “. . . to build a loyal customer base . . . to get repeat business (or customers)”. This example brings home to us that reciprocal user participation sometimes represents an important addition to the dictionary content itself, above all by allowing users to explore specific language questions in a very specific way.

Discussion pages and comments are a further form of mutual exchange among users. On WORDNIK, users can comment on dictionary entries. This function is meant to be used to react to entries, to ask questions, or simply to express one’s own opinion on words but it is also used to express views on content that is hardly related to the content of the entry at all. For example, comments on the headword *dictionary* range from preferences for particular Internet dictionaries to descriptions of terms like *lexicography*.²¹

Discussion pages in WIKTIONARY make it possible to discuss each individual dictionary entry on its own page. Unlike comments or forums, user contributions on these pages are not tied to a chronological structure but can be placed anywhere. As a result, different aspects can be discussed in parallel (cf. Ferschke et al. 2013). → Fig. 8.5 shows an extract from a discussion on the meaning description in WIKTIONARY for the headword *Kreuzung* (‘crossroads, junction, intersection’). A core question being discussed is whether a road, by definition, ends at an intersection or continues across it, which has implications when defining the term as a crossing of four or only two

¹⁹ <https://blog.leo.org/> [last access: October 24, 2023].

²⁰ <https://dict.leo.org/forum/viewUnsolvedquery.php?idThread=88976&lang=en> [last access: October 24, 2023].

²¹ <https://www.wordnik.com/words/dictionary> [last access: October 24, 2023].

roads. Among other things, the extract makes clear what an important role sources perceived as authorities, like DUDEN or the DWDS, play in users' argumentation but also how vehemently these discussions can be conducted, particular if, as in this specific example, an "edit war" is to be averted. In the specific example (→ Fig. 8.5), a registered user is annoyed about the reversion of a change he/she made to the entry *Kreuzung*: "Sorry for my strong choice of words, but I don't know what kind of 'experts' are reverting and reviewing here!" She/he emphasises that only "a place where 4 or more roads meet is called a crossroads. This is equivalent to saying: a place where 2 (or more) roads intersect OR a place where one road crosses a second road. I really don't know what meanings you are trying to 'palm off' on the readers here, but this is borderline behaviour. [. . .] I don't want to start an edit war here, which is why I won't change the reversion of my changes again and ask someone with expertise and understanding to take care of the matter." Another user replies: "Gladly. A road that leads to a crossroad doesn't end there but simply runs through it. So you end up with two or more roads meeting. Defined in the same way in the DWDS, in Duden Das große Wörterbuch der deutschen Sprache [. . .]. Where does it say otherwise? Apart from that, your tone is once again completely inappropriate. [. . .] Or am I not seeing the problem now? You see a difference between 'meet' and 'cross'?" And the dispute continues in this tone of voice.

In essence, accessory user contributions are affected by the same quality criteria that were discussed above for direct and indirect contributions. Removing inappropriate content manually is possible, especially in smaller projects, while larger initiatives make use of collaborative engagement or automatic systems like spam filters. In WORDNIK, for example, the option exists for every comment to send a feedback email to the editors. This is activated by clicking on the symbol of a downturned thumb at the end of each comment; the editors can then remove anything unsuitable if necessary. In open-collaborative resources like WIKTIONARY, however, the removal or correction of misplaced or false content rests in the hands of the contributors alone.

As long as a discussion about relevant lexicographic issues actually takes place (in contrast, for example, to vague comments or comments largely not relevant to the topic in environments like WORDNIK), comments and discussion pages like those in WIKTIONARY can also constitute a quality measure for the development of the relevant dictionaries. However, this applies not only to purely collaboratively compiled dictionaries like WIKTIONARY but also to the field of professional and, in part, commercial lexicography, which can gather qualitative and quantitative feedback and information in this way to develop their own dictionary outputs.

Ensuring quality is a major incentive for publishers to provide opportunities for communication or exchange associated with dictionaries, together with the opportunity to advertise their products and bind users to their brand, for which a wider variety of online channels are used. Educational initiatives around dictionaries can also play a role in this context, as we saw through the example of the ANW.

On the part of the users, the motivations for being involved can be as varied as the ways of making the contributions themselves. One reason undoubtedly lies in the desire

Diskussion:Kreuzung

Treffpunkt [Bearbeiten]

Entschuldigt meine heftige Wortwahl, aber ich weiß nicht, was hier für "Fachleute" revertieren und sichten!!

- ein Ort, wo sich 2 Straßen treffen, nennt man "Straßenknick" oder "Straßenecke", im einfachsten Fall einfach nur Straße, wenn eine gerade verlaufende Straße von der a-Straße zur b-Straße wird.
- ein Ort, wo sich 3 Straßen treffen, wird Straßengabel oder auch Abzweigung genannt
- ein Ort, wo sich 4 oder mehr Straßen treffen, wird Kreuzung genannt. Das ist sinngleich mit der Aussage: ein Ort, wo sich 2 (oder mehr) Straßen kreuzen oder ein Ort, wo eine Straße eine zweite Straße quert.

Ich weiß wirklich nicht, was ihr hier den Lesern für Bedeutungen "unterjubeln" wollt, aber das ist schon grenzwertiges Verhalten. Da ich bei einer massiven Revertierung ohne jede Rückfrage erheblichen Sachverstand erwarte bzw. unterstelle, kann ich hier nur VM annehmen.

Ich möchte jetzt hier keinen Editwar anzetteln, weshalb ich die stattgefundene Revertierung meiner Änderungen nicht erneut ändere und bitte, jemand mit Sachkunde und Verständnis möge sich der Angelegenheit annehmen. —JÄhh (Diskussion) 12:03, 7. Mär. 2011 (MEZ)

Gerne. Eine Straße, die in eine Kreuzung führt, endet nicht dort, sondern läuft, salopp gesagt, einfach durch. So kommt man auf zwei oder mehr sich treffende Straßen. Definitorisch ebenso erfasst im DWDS [↗](#), in *Duden Das große Wörterbuch der deutschen Sprache in 10 Bänden. 3., völlig neu bearbeitete und erweiterte Auflage. Mannheim, Leipzig, Wien, Zürich: Dudenverlag 1999*, unter englisch "crossroads" im *Oxford Dictionary of English* und bei *Merriam-Webster* [↗](#). Wo steht es anders? Davon abgesehen ist dein Ton wieder einmal völlig unangemessen. —Pill (Kontakt) 12:50, 7. Mär. 2011 (MEZ)

Oder sehe ich jetzt das Problem nicht? Du siehst einen Unterschied zwischen "treffen" und "kreuzen"? —Pill (Kontakt) 14:44, 7. Mär. 2011 (MEZ)

Pill, ich kann es nicht fassen: Du definiert in Deinem Einleitungssatz eine Kreuzung damit, das Straßen, die in eine Kreuzung führen, dort nicht enden? Ist das jetzt hier pillepalle?

Wenn sich also 2 Straßen im Winkel von 45 Grad treffen, dann ist das für die ganze Welt eine Spitzkehre oder auch eine sehr scharfe Kurve, aber für Euch oder Dich ist das eine Kreuzung - ja? Wenn sich 3 Straßen jeweils im Winkel von 45 Grad treffen, dann ist das für Euch hier eine Kreuzung?

Wenn auf eine Straße (eine durchgehende Hauptstraße, zur besseren Beurteilung) eine (zur Verdeutlichung: kleine) Seitenstraße im Winkel von 90 Grad trifft, dann ist das bei Euch eine Kreuzung?

Weiter führst Du als Beleg an (Definitorisch ebenso erfasst im [<http://www.dwds.de/?qu=Kreuzung&view=1> [↗](#) DWDS), wenn sich dort 2 Straßen treffen, übersiehst oder ignorierst aber, das unter der Quelle zu finden ist: "sich zwei Straßen kreuzen". Jetzt wirst Du mir sicherlich belegen, daß zwischen "sich treffen" und "sich kreuzen" überhaupt kein Unterschied besteht - dann bin ich zufrieden und gebe Ruhe - würde aber kollidieren mit dem Eintrag [kreuzen](#).

Du argumentierst: "Eine Straße, die in eine Kreuzung führt, endet nicht dort, sondern läuft, salopp gesagt, einfach durch." Beleg - natürlich Fehlanzeige! Aber ich bin gerne bereit, das mal durchzukauen: Wenn nach Deiner Lesart eine Straße an einer Kreuzung nicht aufhört, sondern durchläuft, wie nennst Du denn dann bitte einen Verkehrs(koten)punkt, von dem aus sich jeweils vom Mittelpunkt entfernend 5 Straßen wegführen? Zur Besseren Kenntlichmachung nennen wir die Straßen mal A, B, C, D und E, und zwischen ihnen jeweils ein Winkel von 72 Grad.

Fig. 8.5: Extract from the discussion contribution on the headword *Kreuzung* in WIKTIONARY.

to fill gaps in information quickly, for which forms of bidirectional communication appear to be particularly well suited (for more on possible motivations → Section 8.2).

8.5 Discussion

A classification of different forms of user participation like the one presented above serves, first, as an instrument to describe existing dictionaries and as a basis for further research into user participation in lexicographic contexts. Second, it is helpful when planning new resources and platforms or when revising existing ones.

A thorough discussion of user participation has also shown that having recourse to the potential power of collective intelligence is in no way a particularly new phenomenon in the field of lexicography nor one that has scarcely been used before. Above all, explicit feedback has been encouraged by dictionary compilers from their early days, for example by Duden's editors or by Oxford University Press, in the form of postal submissions. However, what is new is social interaction via social media and associated technologies, which have paved the way for user participation to become a mass phenomenon in its current scale and format. In particular, the forms of direct user participation were not – or only barely – possible before the emergence of the Internet.

All forms of user participation exhibit specific strengths and weaknesses, which have to be recognised and balanced for a dictionary to be planned effectively. The potential of collaboratively compiled dictionaries lies in the fact that, in theory, there are an unlimited number of participants – instead of single individuals or teams of a clearly defined size – with varying expertise who can devote themselves to these dictionaries for an unlimited time and in very particular ways. Not only the compilation of these dictionaries can be essentially unrestricted and free of charge but also access to them.

First and foremost, added value arises for dictionary content through direct user participation and explicit feedback. In open-collaborative dictionaries, users and providers are one and the same to some extent, and all content is compiled and revised in a participatory manner. Particularly in the case of contributions to collaborative-institutional dictionaries and semi-collaborative dictionaries or in the form of explicit feedback, this added value can extend from closing individual gaps in lemmas via supplementing important examples of usage to whole dictionary entries and the supply of larger bodies of material. This not only means that the coverage of a dictionary can be extended and content gaps closed but also that the lexicographic work can be undertaken more quickly and at lower cost. Lexicographers and language experts can save time and money when research tasks or the draft formulation of whole entries can be given to users. Dictionary providers and users benefit equally if content is available more quickly and in a more up-to-date form.

Furthermore, the strengths of collaborative lexicography lie in the diversity of the user group, which facilitates a wide-ranging description of different speech varieties and language pairs. This includes numerous dialect and regional expressions and phrases (e.g. *bostitchen*: Swiss *tackern* ‘to staple’²²), slang terms and Internet jargon (e.g. *Karen*: a pejorative term used to refer to a middle-aged and middle-class white woman who puts herself first, is rude, insensitive, pushy, and whiny;²³ *ROFL*: rolling on the floor laughing²⁴) and technical language/jargon (e.g. *ageotype*: a category of ageing biomarkers;²⁵ *shewee*: a portable female urinary device²⁶). Among the languages and translation equivalents, we can find languages with only a few speakers and endangered languages (*siissisoq*: nose horn in Greenlandic) as well as language combinations that are scarcely of any commercial interest (e.g. Greenlandic–Italian; cf.; Matuschek et al. 2018; Meyer 2013; Meyer/Gurevych 2012; Rundell 2012).

However, it is not only newly contributed descriptions that bring added value but also the reporting or correcting of errors, which can raise the quality of a dictionary considerably. On the one hand, this kind of collaborative checking of quality can be used to perfect information that has been professionally compiled; on the other, it can fulfil its own purpose in selecting the best user entries. Here, it is the large number of users that is, first and foremost, an advantage since inappropriate user contributions (e.g. inappropriate comments and discussion contributions but also external user-generated content) can hardly be checked by single individuals or a few moderators. The example of WORDNIK shows some of the possible ways of employing users to monitor comments. While many forms of user participation express the opinion or understanding of an individual user, there are multiple efforts to consolidate the different perspectives of a larger group of speakers, for example by collaboratively formulating a dictionary article in Wiktionary or by jointly evaluating the usefulness of a particular translation equivalent in DICT.CC (→ Fig. 8.3).

However, dictionary users also benefit directly from the different forms of user participation. The use of open licences in collaborative dictionaries makes lexicographic content accessible to a large body of users. Furthermore, accessory forms of user participation increase the popularity of dictionaries while direct and indirect forms of participation provide the opportunity to actively shape the dictionary as a resource and to have a stake in the final product. In addition, binding users closely into the lexicographic process can serve an educational purpose and help to develop

22 <https://de.wiktionary.org/wiki/bostitchen> [last access: February 22, 2024]. The verb derives from *Bostitch*, the name of a company producing staplers.

23 <https://www.urbandictionary.com/define.php?term=Karen&page=2> [last access: February 22, 2024].

24 <https://en.wiktionary.org/wiki/ROFL> [last access: February 22, 2024].

25 <https://en.wiktionary.org/wiki/ageotype> [last access: February 22, 2024].

26 <https://www.urbandictionary.com/define.php?term=shewee> [last access: February 22, 2024]. The noun derives from *Shewee*, the company producing the devices.

important competences in using dictionaries. Complementary services and products achieve this in particular, for example, in playful ways or through engaging blog posts that prompt users to consult the dictionary. Direct user participation also has a contribution to make in this respect since checks have to be undertaken to see whether information is already contained in the dictionary and to see how language descriptions can be most effectively formulated. In turn, the exchange of views among dictionary users and language advice services represent added value for users if language questions are discussed that are not answered in the dictionary or at least not for a specific, given context.

Overall, user contributions lead to a negotiation of content according to the principles of supply and demand, from which both users and providers can, in theory, benefit. Implicit feedback reveals what is actually looked up by users. Explicit feedback and user comments provide information about the wishes and expectations of users in relation to the dictionary. On the one hand, direct user contributions reflect the usage of language on the part of the users (i.e., the user's "supply" of content). On the other hand, newly created content may be oriented towards demand in cases where users come across a gap while consulting the dictionary and then research and add the relevant material.

In contrast to this, the forms of participation also entail numerous challenges and weaknesses, which is an argument against planning a particular participatory resource or which demand further lexicographic, technological, or educational solutions. For example, the potential to save time and money described above is in no way clear cut. For collaborative-institutional and semi-collaborative forms of participation as well as with explicit feedback, the editorial checking of user contributions leads, in the first instance, to an increase in work for dictionary providers. The extent to which the usefulness and quality of contributions exceeds the time invested in checking them undoubtedly varies between individual dictionary projects and the different modes of participation. While explicit feedback from providers has been predominantly judged as positive (cf. Rautmann 2014; Thier 2014), it has also been shown that implicit feedback from log data only has a limited significance, no matter how much effort and expense is devoted to organising the analysis of the results. Dealing with plagiarism is also a particular problem. It is a challenging task to identify what appear to be high-quality user contributions as direct, unacknowledged use of data from other secondary sources, a task that can bring with it greater effort than would be involved in compiling a whole new lexicographic description based on primary sources.

Questions concerning the quality of user contributions in comparison with resources maintained purely by editors require particular reflection. User-generated dictionaries contain information about extremely varied language varieties and specialist vocabularies or about rarely occurring lemmas, while, to draw on examples from WIKTIONARY, common German words like *Fehlalarm* ('false alarm') or *Einzugsgebiet* ('catchment area') are missing. Frequent interpretations of lemmas are sometimes

also not captured, such as the interpretation of the German lemma *Favorit* as “preferred object”²⁷.

In addition, if poor quality, inappropriate, or false descriptions are posted in a dictionary, this carries with it the danger of users’ language questions no longer being able to be answered reliably and the reference work thereby becoming unusable. The survey described in → Chapter 9.3.1 on the importance of criteria for Internet dictionaries demonstrates that users assign the highest priority to the reliability of the information available in dictionaries. Thus, it remains to be determined whether, and in what ways, user contributions really provide added value, something which has only been the subject of rudimentary research thus far. The studies by Fuertes-Olivera (2009), Hanks (2012), Meyer/Gurevych (2012), Rundell (2012), and Lew (2014) exhibit qualitative shortcomings in collaboratively compiled sources that can be traced back to mistaken, non-specific, old-fashioned, and partly obsolete descriptions. Whether user contributions bring anything new at all in qualitative or quantitative terms is of central importance for evaluating their potential. Meyer/Gurevych (2014) demonstrated that edited sources (e.g. DWDS or DUDEN ONLINE) were often listed in descriptions in the German WIKTIONARY that had been contributed collaboratively. This, and also a look at the traditional microstructure of dictionary entries, point to a comparatively conservative lexicographic approach in this collaborative space while collaborative innovation tends instead to be found in wide-ranging collections of material and new ways of integrating existing material. Overall, user-generated dictionaries seem to have considerable gaps and shortcomings in quality, which limits their usefulness. As a result, the expertise of professional lexicographers is indispensable if Internet dictionaries of high quality are to be created.

8.6 Summary and outlook

In this chapter, we addressed a relevant topic area in Internet lexicography, namely user participation, which should not be underestimated in its relevance. Indeed, it constitutes an important basis for enriching the quality and quantity of dictionary resources and in some cases is even the sole source for their entire construction. Using specific examples, we discussed three basic types of user participation in a systematic overview.

Forms of direct user participation encompass communal efforts in the construction and development of open-collaborative, collaborative-institutional, and semi-collaborative dictionaries, albeit with different degrees of editorial control and input

²⁷ For example, Wiktionary captures only the traditional meanings of the lemma in German, i.e. a) living being that is favoured by someone, b) participants in a competition with the best chance of winning. <https://de.wiktionary.org/wiki/Favorit> [last access: February 22, 2024].

options but with the common characteristic of directly integrating user contributions into the relevant dictionary. To a large degree, this type of user participation has only become possible with the advent of social media technologies.

We have to distinguish forms of indirect participation from direct participation. These are based on the principle of feedback or a mediated influence on the content and form of dictionaries. These include, on the one hand, form-based and free-text feedback, which dictionary users make available knowingly and of their own volition, and, on the other hand, implicit feedback gathered through lexicographically motivated analyses of log data or the integration of external user-generated content not intended *a priori* for lexicographic purposes.

Finally, the concept of accessory user participation covers different forms of exchange between dictionary compilers and users beyond the actual dictionary contents on the macro- and microstructural levels, which can proceed in a unidirectional or bidirectional manner.

It has also become clear, among other things, that different forms of user participation do not rule out one another but rather can be applied in parallel or in combination within the same dictionary or dictionary portal. This can be seen through the example of the LEO portal, which facilitates all three types of user participation presented in this chapter: it makes strong use of translation contributions and donated word lists, which are generally included directly in the relevant dictionaries, after they have been checked by editors, as is customary for semi-collaborative contexts. In addition, there are opportunities for users to contribute indirectly to the dictionary via corrections or suggestions for headwords. Feedback on provisional versions, for example, with different layouts can also be gathered. Finally, LEO provides forums through which users can exchange views with one another and the “Ask Dr. Bopp” language blog, which constitute forms of accessory user participation.

In a final discussion we illustrated the relevance of the present classification of types of user participation and explored their strengths and weaknesses. As the online publication of dictionary content increases, the issue of how to structure resources for user participation is increasingly gaining in significance for dictionary providers. This chapter should serve as an orientation for asking these questions and, at the same time, serve as a basis for further discussion on user contributions. In particular, the quality of contributions from the various characteristic forms of user participation has not been extensively researched to date.

Bibliography

Further reading

- Abel, Andrea/Klosa, Annette (2014) (ed.): *Der Nutzerbeitrag im Wörterbuchprozess. 3. Arbeitsbericht des wissenschaftlichen Netzwerks "Internetlexikografie"*. Mannheim: Institut für Deutsche Sprache. *The volume addresses user participation in Internet dictionaries in a targeted way and discusses different forms of participation for individual dictionary services or products.*
- Carr, Michael (1997): Internet Dictionaries and Lexicography. In: *International Journal of Lexicography* 10:3, 209–230. *The article is relevant to the field of electronic lexicography from a historical perspective, especially with regard to the term "bottom-up lexicography".*
- Melchior, Luca (2012): Halbkollaborativität und Online-Lexikographie. Ansätze und Überlegungen zu Wörterbuchredaktion und Wörterbuchforschung am Beispiel LEO Deutsch–Italienisch. In: *Lexicographica* 28, 337–372. *This article describes different types of users and the ways they are involved in the LEO dictionary portal.*
- Meyer, Christian M. (2013): Wiktionary: The Metalexicographic and Natural Language Processing Perspective. Dissertation, Darmstadt: Technische Universität Darmstadt. *Chapters 2–4 explore in detail user contributions, the organisation, and the content of the collaboratively compiled WIKTIONARY.*
- Rundell, Michael (2016): Dictionaries and crowdsourcing, wikis and user-generated content. In: Hanks, Patrick/de Schryver, Gilles-Maurice (eds.): *International Handbook of Modern Lexis and Lexicography*. Berlin/Heidelberg: Springer, 1–16. *This article critically discusses various forms of crowdsourcing in the context of dictionary projects.*

Literature

Academic literature

- Abel, Andrea, et al. (2003): Evaluation of the Web-based Learners Dictionary ELDIT. In: Lassner, David/McNaught, Carmel (eds.): *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, Chesapeake, VA, USA*, 1210–1217.
- Abel, Andrea/Meyer, Christian M. (2013): The dynamics outside the paper: user contributions to online dictionaries. In: Kosem, Iztok, et al. (eds.): *Proceedings of the 3rd eLex conference 'Electronic lexicography in the 21st century: thinking outside the paper'*. Ljubljana/Tallinn, 179–194.
- Abel, Andrea/Meyer, Christian M. (2016): Nutzerbeteiligung. In: Klosa, Annette/Müller-Spitzer, Carolin (eds.): *Internetlexikografie. Ein Kompendium*. Berlin/Boston: De Gruyter, 249–290.
- Benjamin, Martin (2014): Collaboration in the Production of a Massively Multilingual Lexicon. In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC)*. Reykjavik, Iceland, 211–215.
- Carr, Michael (1997): Internet Dictionaries and Lexicography. In: *International Journal of Lexicography* 10:3, 209–230.
- de Schryver, Gilles-Maurice/Joffe, David (2004): On how electronic dictionaries are really used. In: *Proceedings of the 11th EURALEX International Congress*. Lorient, France, 187–196.
- de Schryver, Gilles-Maurice/Prinsloo, Danie J. (2001): Fuzzy SF: Towards the ultimate customised dictionary. In: *Studies in Lexicography* 11:1, 97–111.

- de Schryver, Gilles-Maurice/Wolfer, Sascha/Lew, Robert (2019): The relationship between dictionary look-up frequency and corpus frequency revisited: a log-file analysis of a decade of user interaction with a Swahili-English dictionary. In: *GEMA Online Journal of Language Studies* 19:4, 1–27.
- Engelberg, Stefan/Müller-Spitzer, Carolin (2013): Dictionary Portals. In: Gouws, Rufus H., et al. (eds.): *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin/Boston: De Gruyter, 1023–1035.
- Ferschke, Oliver/Daxenberger, Johannes/Gurevych, Iryna (2013): A Survey of NLP Methods and Resources for Analyzing the Collaborative Writing Process in Wikipedia. In: Gurevych, Iryna/Kim, Jungi (eds.): *The People's Web Meets NLP: Collaboratively Constructed Language Resources*. Berlin/Heidelberg: Springer, 121–160.
- Fort, Karën/Adda, Gilles/Cohen, K. Bretonnel (2011): Amazon Mechanical Turk: Gold Mine or Coal Mine? In: *Journal of Computational Linguistics* 37:2, 413–420.
- Francopoulou, Gil (2009) (ed.): *LMF: Lexical Markup Framework*. London: Wiley-ISTE.
- Fuertes-Olivera, Pedro A. (2009): The Function Theory of Lexicography and Electronic Dictionaries: Wiktionary as a Prototype of Collective Free Multiple-Language Internet Dictionary. In: Bergenholtz, Henning/Nielsen, Sandro/Tarp, Sven (eds.): *Lexicography at a Crossroads: Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 99–134.
- Hanks, Patrick (2012): Word Meaning and Word Use: Corpus evidence and electronic lexicography. In: Granger, Sylviane/Paquot, Magali (eds.): *Electronic Lexicography*. Oxford: Oxford University Press, 57–82.
- Hanks, Patrick/Franklin, Emma (2019): Do Online Resources Give Satisfactory Answers to Questions About Meaning and Phraseology? In: Corpas Pastor, Gloria/Mitkov, Ruslan (eds.): *Computational and Corpus-Based Phraseology*. Cham: Springer International Publishing, 159–172.
- Hausmann, Franz Josef (1989): Dictionary Criminality. In: Hausmann, Franz J., et al. (eds.): *Wörterbücher: Ein internationales Handbuch zur Lexikographie*, Berlin/New York: De Gruyter, 97–101.
- Howe, Jeff (2008): *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. New York: Three Rivers Press.
- Javanmardi, Sara et al. (2009): User contribution and trust in Wikipedia. In: *Proceedings of the 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, 1–6.
- Kemmer, Katharina (2010): *Onlinewörterbücher in der Wörterbuchkritik. Ein Evaluationsraster mit 39 Beurteilungskriterien*. Mannheim 2010.
- Klein, Wolfgang/Geyken, Alexander (2010): Das Digitale Wörterbuch der Deutschen Sprache (DWDS). In: *Lexicographica* 26, 79–96.
- Köhler Simonsen, Henrik (2005): User Involvement in Corporate LSP Intranet Lexicography. In: Gottlieb, Henrik/Mogensen, Jens Erik/Zettersten, Arne (eds.): *Symposium on Lexicography XI: Proceedings of the Eleventh International Symposium on Lexicography*. Tübingen: Niemeyer, 489–510.
- Krause, Jens/Krause, Stefan (2011): Kollektives Verhalten und Schwarmintelligenz. In: Otto, Klaus-Stephan/Speck, Thomas (eds.): *Darwin meets Business: Evolutionäre und bionische Lösungen für die Wirtschaft*. Wiesbaden: Springer, 127–134.
- Kreutzer, Till (2011): *Open Content Lizenzen. Ein Leitfaden für die Praxis*, Bonn: UNESCO.
- Lampe, Cliff et al. (2010): Motivations to Participate in Online Communities. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, 1927–1936.
- Landau, Sidney I. (2001): *Dictionaries: The art and craft of lexicography*. Cambridge: Cambridge University Press.
- Lew, Robert (2011): Online dictionaries of English. In: Fuertes-Olivera, Pedro Antonio/Bergenholtz, Henning (eds.): *E-Lexicography: The Internet, Digital Initiatives and Lexicography*. London/New York: Continuum, 230–250.

- Lew, Robert (2014): User-generated content (UGC) in English online dictionaries. In: Abel, Andrea/Klosa, Annette (eds.): *Der Nutzerbeitrag im Wörterbuchprozess. 3. Arbeitsbericht des wissenschaftlichen Netzwerks "Internetlexikografie"*. Mannheim, 7–25.
- Malone, Thomas W./Laubacher, Robert/Dellarocas, Chrysanthos (2010): Harnessing Crowds: Mapping the Genome of Collective Intelligence. In: *Social Science Research Network Electronic Paper Collection*. Rochester, NY. <http://ssrn.com/abstract=1381502> [last access: May 2, 2024].
- Mann, Michael (2010): Internet-Wörterbücher am Ende der "Nullerjahre": Der Stand der Dinge. Eine vergleichende Untersuchung beliebter Angebote hinsichtlich formaler Kriterien. In: *Lexicographica* 26, 19–46.
- Matuschek, Michael/Meyer, Christian M./Gurevych, Iryna (2018): Multilingual Knowledge in Aligned Wiktionary and OmegaWiki for Translation Applications. In: Rehm, Georg, et al. (eds.): *Language technologies for a multilingual Europe. TC 3 III*. Berlin: Language Science Press, 139–180.
- McKean, Erin (2017): Wordnik. In: Fuertes-Olivera, Pedro A. (ed.): *The Routledge Handbook of Lexicography*. London: Routledge, 473–484.
- Melchior, Luca (2012): Halbkollaborativität und Online-Lexikographie. Ansätze und Überlegungen zu Wörterbuchredaktion und Wörterbuchforschung am Beispiel LEO Deutsch–Italienisch. In: *Lexicographica* 28, 337–372.
- Melchior, Luca (2014): Ansätze einer halbkollaborativen Lexikographie. In: Abel, Andrea/Klosa, Annette (eds.): *Ihr Beitrag bitte! – Der Nutzerbeitrag im Wörterbuchprozess. 3. Arbeitsbericht des wissenschaftlichen Netzwerks "Internetlexikografie"*. Mannheim, 26–47.
- Meyer, Christian M. (2013): Wiktionary: The Metalexicographic and the Natural Language Processing Perspective. Dissertation, Darmstadt: Technische Universität Darmstadt. <http://tuprints.ulb-tu-darmstadt.de/3654/> [last access: May 2, 2024].
- Meyer, Christian M./Andrea, Abel (2017): User Participation in the Internet Era. In: Fuertes-Olivera, Pedro A (ed.): *The Routledge Handbook of Lexicography*. London/New York: Routledge, 735–753.
- Meyer, Christian M./Gurevych, Iryna (2012): Wiktionary: A new rival for expert-built lexicons? Exploring the possibilities of collaborative lexicography. In: Granger, Sylviane/Paquot, Magali (eds.): *Electronic Lexicography*. Oxford: Oxford University Press, 259–291.
- Meyer, Christian M./Gurevych, Iryna (2014): Methoden bei kollaborativen Wörterbüchern. In: *Lexicographica* 30, 187–212.
- Müller-Spitzer, Carolin/Möhrs, Christine (2008): First ideas of user-adapted views of lexicographic data exemplified on OWID and ELEXIKO. In: Zock, Michael/Huang, Chu-Ren (eds.): *Proceedings of the COLING Workshop on 'Cognitive Aspects on the Lexicon'*. Manchester, 39–46.
- Müller-Spitzer, Carolin/Wolfer, Sascha/Koplenig, Alexander (2015): Observing Online Dictionary Users: Studies Using Wiktionary Log Files. In: *International Journal of Lexicography* 28:1, 1–26.
- Naber, Daniel (2005): OpenThesaurus: ein offenes deutsches Wortnetz. In: Fisseni, Bernhard, et al. (eds.): *Sprachtechnologie, mobile Kommunikation und linguistische Ressourcen: Beiträge zur GLDV-Tagung*. Frankfurt, 422–433.
- Rafaeli, Sheizaf/Ariel, Yaron (2008): Online Motivational Factors: Incentives for Participation and Contribution in Wikipedia. In: Barak, Azy (ed.): *Psychological Aspects of Cyberspace: Theory, Research, Applications*. Cambridge: Cambridge University Press, 234–267.
- Rautmann, Karin (2014): Duden online und seine Nutzer. In: Abel, Andrea/Klosa, Annette (eds.): *Ihr Beitrag bitte! – Der Nutzerbeitrag im Wörterbuchprozess. 3. Arbeitsbericht des wissenschaftlichen Netzwerks "Internetlexikografie"*. Mannheim, 48–61.
- Rosenberg, Louis B. (2015): Human Swarms: a real-time paradigm for collective intelligence. In: Andrews, Paul, et al. (eds.): *Proceedings of the European Conference on Artificial Life 2015 (ECAL)*. Cambridge: MIT Press, 658–659.

- Rundell, Michael (2012): 'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical. In: Vatdedt Fjeld, Ruth/Torjusén, Julie Matilde (eds.): *Proceedings of the 15th EURALEX International Congress*. Oslo, 47–92.
- Schoonheim, Tanneke, et al. (2012): Dictionary Use and Language Games: Getting to Know the Dictionary as Part of the Game. In: Vatdedt Fjeld, Ruth/Torjusén, Julie Matilde (eds.): *Proceedings of the 15th EURALEX International Congress*. Oslo, 974–979.
- Storrer, Angelika (1998): Hypermedia-Wörterbücher: Perspektiven für eine neue Generation elektronischer Wörterbücher. In: Wiegand, Herbert Ernst (ed.): *Wörterbücher in der Diskussion III*. Tübingen: Niemeyer, 107–135.
- Storrer, Angelika (2010): Deutsche Internet-Wörterbücher: Ein Überblick. In: *Lexicographica* 26, 155–164.
- Storrer, Angelika/Freese, Katrin (1996): Wörterbücher im Internet. In: *Deutsche Sprache* 24:2, 97–153.
- Surowiecki, James (2005): *The Wisdom of Crowds*. New York: Anchor Books.
- Thier, Katrin (2014): Das Oxford English Dictionary und seine Nutzer. In: Abel, Andrea/Klosa, Annette (eds.): *Ihr Beitrag bitte! – Der Nutzerbeitrag im Wörterbuchprozess. 3. Arbeitsbericht des wissenschaftlichen Netzwerks "Internetlexikografie"*. Mannheim, 62–69.
- Verlinde, Serge/Binon, Jean (2010): Monitoring Dictionary Use in the Electronic Age. In: Dykstra, Anne/Schoonheim, Tanneke (eds.): *Proceedings of the 14th EURALEX International Congress*. Ljouwert, 1144–1151.
- Wang, William Yang/McKeown, Kathleen (2010): "Got You!": Automatic Vandalism Detection in Wikipedia with Web-based Shallow Syntactic-Semantic Modeling. In: Huang, Chu-Ren/Jurafsky, Dan (eds.): *Proceedings of the 23rd International Conference on Computational Linguistics*. Beijing, 1146–1154.
- Wiegand, Herbert Ernst, et al. (2010) (eds.): *Wörterbuch zur Lexikographie und Wörterbuchforschung/ Dictionary of Lexicography and Dictionary Research*, vol. 1: A–C. Berlin/New York: De Gruyter.

Dictionaries

- ANW = *Algemeen Nederlands Woordenboek*. Leiden: Instituut voor Nederlandse Lexicologie. <http://anw.inl.nl> [last access: May 2, 2024].
- BAB.LA = *Bab.La. Online Wörterbuch für 24 Sprachen*. Hamburg: bab.la GmbH. <http://bab.la> [last access: May 2, 2024].
- BLF = *Base lexicale du français*. Leuven: Katholieke Universiteit. Archived at <https://web.archive.org/web/20131025210135/http://ilt.kuleuven.be/blf/> [last access: May 2, 2024].
- CANONET = *Deutsche Wörterbücher und Grammatik*. Basel: Canoo Engineering. Archived at <https://web.archive.org/web/20181001082324/http://www.canoo.net:80/> [last access: May 2, 2024].
- DDO = *Den Danske Ordbog (The Danish Dictionary)*. København: Det Danske Sprog- og Litteraturselskab. <https://ordnet.dk/ddo> [last access: May 2, 2024].
- DICT.CC = *dict.cc. Deutsch-Englisch Wörterbuch*. Wien: dict.cc GmbH. <http://www.dict.cc> [last access: May 2, 2024].
- DICTIONARY.COM = *Dictionary.com*. Oakland, CA: Dictionary.com. <http://www.dictionary.com> [last access: May 2, 2024].
- DUDEN ONLINE = *Duden*. Berlin: Bibliographisches Institut/Dudenverlag. <http://www.duden.de> [last access: May 2, 2024].
- DWDS = *Digitales Wörterbuch der deutschen Sprache*. Berlin: Berlin-Brandenburgischen Akademie der Wissenschaften. <http://dwds.de> [last access: May 2, 2024].
- ELDIT = *Elektronisches Lernerwörterbuch Deutsch-Italienisch*. Bozen: Europäische Akademie. <http://eldit.eurac.edu/> [last access: May 2, 2024].

- GLOSBE = *Glosbe – das mehrsprachige Online-Wörterbuch*. Warschau: Cloud Inside. <http://glosbe.com> [last access: May 2, 2024].
- GOOGLE DICTIONARY = *Google Dictionary*. Mountain View: Google. Archived at <https://web.archive.org/web/20101127085248/http://www.google.com:80/dictionary> [last access: May 2, 2024].
- KAMUSI = *The Kamusi Project, Global Online Living Dictionary*. Genf: Kamusi Project International/Delaware: Kamusi Project USA. <http://kamusi.org> [last access: May 2, 2024].
- LEO = *LEO*. München: LEO GmbH. <http://dict.leo.org> [last access: May 2, 2024].
- MACMILLAN DICTIONARY ONLINE = *Macmillan Dictionary Online*. London: Macmillan Publishes Ltd. Archived at <https://web.archive.org/web/20220111230227/https://www.macmillandictionary.com/> [last access: May 2, 2024].
- MACMILLAN OPEN DICTIONARY = *Macmillan Open Dictionary*. London: Macmillan Publishers Ltd. Archived at <https://web.archive.org/web/20211129035202/https://www.macmillandictionary.com/open-dictionary/latestEntries.html> [last access: May 2, 2024].
- MERRIAM-WEBSTER ONLINE = *Merriam-Webster Online*. Springfield, MA: Merriam-Webster. <http://www.merriam-webster.com> [last access: May 2, 2024].
- MERRIAM-WEBSTER OPEN DICTIONARY = *The Open Dictionary*. Springfield, MA: Merriam-Webster. Archived at <https://web.archive.org/web/20180624155024/http://nws.merriam-webster.com/opendictionary/> [last access: May 2, 2024].
- NITTY GRITS = *Nitty Grits: International Culinary Dictionary* (Suzy Oakes, ed.). New Orleans: Southern Food and Beverage Museum, 2011. <http://www.nittygrits.org> [last access: May 2, 2024].
- OED = *Oxford English Dictionary* (John A. Simpson/Edmund S.C. Weiner, ed.). 2nd edition. Oxford: Oxford University Press 1989.
- OED ONLINE = *Oxford English Dictionary Online*. Oxford: Oxford University Press. <http://www.oed.com> [last access: May 2, 2024].
- OMEGAWIKI = *OmegaWiki, a dictionary in all languages*. San Francisco: Wikimedia Foundation. Archived at <https://web.archive.org/web/20230709193706/http://www.omegawiki.org/> [last access: May 2, 2024].
- OPENTHESAUROS = *OpenThesaurus, Synonyme und Assoziationen*. Potsdam: Daniel Naber. Online: <http://www.openthesaurus.de>.
- RAP DICTIONARY = *The Rap Dictionary*. Nijmegen: Patrick Atoon. <http://www.rapdict.org> [last access: May 2, 2024].
- SPRACHNUDEL = *Sprachnudel.de*, Wörterbuch der Jetztsprache. Berlin: WEB'arbyte. <http://www.sprachnudel.de> [last access: May 2, 2024].
- SZENESPRACHENWIKI = *Duden Szenesprachenwiki*. Mannheim: Bibliographisches Institut/Dudenverlag. <http://szenesprachenwiki.de> [now offline].
- TECHDICTIONARY = *TechDictionary, the Online Computer Dictionary*. Chesterbrook: techdictionary.com. Offline [archived at <https://web.archive.org/web/20190605111602/http://www.techdictionary.com/>].
- URBAN DICTIONARY = *Urban Dictionary*. San Francisco: Urban Dictionary. <http://www.urbandictionary.com> [last access: May 2, 2024].
- WEBSTER = *Webster's Revised Unabridged Dictionary* (Noah Porter, ed.). Springfield: G & C. Merriam Co., 1913.
- WIKTIONARY = *Wiktionary, das freie Wörterbuch*. San Francisco, CA: Wikimedia Foundation. <http://www.wiktionary.org> [last access: May 2, 2024].
- WORDNIK = *Wordnik. All the words*. San Mateo, CA: Wordnik. <http://www.wordnik.com> [last access: May 2, 2024].

Internet sources

AMAZON MECHANICAL TURK = <https://www.mturk.com/mturk/welcome> [last access: May 2, 2024].

CANOONET-Sprachblog = archived at <https://web.archive.org/web/20160724003445/http://canoo.net/blog/> [last access: May 2, 2024].

CREATIVE COMMONS = <http://de.creativecommons.org/> [last access: May 2, 2024].

CROWDFLOWER = <http://www.crowdflower.com/>.

GOOGLE ANALYTICS = https://www.google.com/intl/de_de/analytics/ [last access: May 2, 2024].

LEO-Sprachblog = <https://blog.leo.org/> [last access: May 2, 2024].

LMF = *Lexical Markup Framework*. <http://www.lexicalmarkupframework.org/> [last access: May 2, 2024].

MACMILLAN-DICTIONARY-Blog = <http://www.macmillandictionaryblog.com/> [last access: May 2, 2024].

MATOMO = <https://matomo.org/> [last access: May 2, 2024].

TEI = *Text Encoding Initiative*. www.tei-c.org [last access: May 2, 2024].

WIKIPEDIA = *Wikipedia, The Free Encyclopedia*. San Francisco, CA: Wikimedia Foundation. www.wikipedia.org [last access: May 2, 2024].

Images

- Fig. 8.1 (top left)** “Ballot paper for the 2021 United Kingdom local elections (Coventry, Westwood ward)”, licensed under the “Creative Commons Attribution 4.0 International” licence (CC BY 4.0) by Wikimedia Commons user domdomegg. <https://commons.wikimedia.org/w/index.php?oldid=815943269> [last access: September 9, 2023].
- Fig. 8.1 (top right)** “Eine Debatte im Plenarsaal des Bayerischen Landtages. Fotografiert im Rahmen des Landtagsprojektes Bayern 2012.”, licensed under the “Creative Commons Attribution-Share Alike 3.0 Unported” licence (CC BY-SA 3.0) by Wikipedia user Tobias Klenze. <http://commons.wikimedia.org/w/index.php?oldid=74775027> [last access: July 18, 2012].
- Fig. 8.1 (bottom left)** “A demonstration in Erlangen, Germany, against tuition fees”, licensed under the “Creative Commons Attribution-Share Alike 2.0 Germany” licence (CC BY-SA 2.0) by Stefan Wagner (<http://trumpkin.de>). <http://commons.wikimedia.org/w/index.php?oldid=112937438> [last access: December 31, 2013].
- Fig. 8.1 (bottom right)** “LA Times”, licensed under the “Creative Commons Attribution-Share Alike 2.0 Generic” licence (CC BY-SA 2.0) by Flickr user Daniel R. Blume. <http://commons.wikimedia.org/w/index.php?oldid=117616015> [last access: February 28, 2014].