**IsiXhosa.click: online, open, user-friendly, and searchable isiXhosa-English dictionary software**

Cael Marquard ([cael.marquard@gmail.com](mailto:cael.marquard@gmail.com))

University of Cape Town: Department of Computer Science; Cape Town, South Africa

IsiXhosa.click[[1]](#endnote-2) is an open-source, online, easy-to-use isiXhosa-English dictionary. It supports typo-tolerant live search, allowing users to find words quickly by typing their first few letters. Word pages provide example sentences, related words, and linguistic information. The software allows users to submit corrections and new vocabulary, which are published after review. IsiXhosa.click is a community-driven, crowd-sourced dictionary project that still enforces quality standards. The website’s database and source code are freely available under open-source licenses.

IsiXhosa.click started in 2021 as an independent, student-led project by a team of home-language English speakers learning isiXhosa. Existing isiXhosa dictionaries were found to be inconvenient and costly. They use proprietary software; do not allow re-use, redistribution and modification; and do not easily permit users to suggest edits. Thus, the software for IsiXhosa.click was built expressly for the project. A key design goal was ease of use for those unfamiliar with dictionaries. The interface is simple and uses minimal jargon, preferring to use terms familiar to learners of isiXhosa. Development effort concentrated on commonly used features, particularly the ability to discover words in the dictionary.

Instead of reading an ordered list of words, users primarily find words through the site’s search feature. Live results are returned immediately as they type. Words may be searched by their full isiXhosa form, their isiXhosa stem, or their English translation. Search is provided by the Tantivy[[2]](#endnote-3) library. A fuzzy term query allows for typo-tolerance within the search. Results are ranked by how well they match the search term using the Jaro-Winkler distance (Winkler 1990). Because the search is typo-tolerant, bilingual, and stemming-enabled, searching is intuitive to users of any lexicographic skill level. Even words which users misspell may be found, which is particularly useful to learners who are new to transcribing isiXhosa.

Words may also be discovered through the “related words” sections. Once a word is found through search, clicking it brings up a page describing it in full. Here, applicable grammatical information such as part of speech, infinitive, and noun class is presented. Example sentences using the word are displayed, if available. Lastly, there is a section listing related words, hyperlinked to their entries in the dictionary. The linking method is also listed (e.g., whether they have related meanings, or whether one word is the plural form of the other). This assists users to find the most relevant entry in the dictionary.

Since IsiXhosa.click is a fully open-source project, it constitutes an Open Education Resource (OER) (Wiley & Hilton 2018). The software is released under the GNU AGPL 3.0 license and available on GitHub[[3]](#endnote-4). It may be used and adapted freely for research or other projects. The dictionary database is released under CC BY SA 4.0. Site archives in CSV format are uploaded to a public mirror[[4]](#endnote-5). This allows for interoperability between other tools, such as Microsoft Excel.

Project members initially created IsiXhosa.click’s word list by adapting resources with permission. In 2022, IsiXhosa.click received a grant through the DH OER Champions programme[[5]](#endnote-6) to expand its word list. Additional entries have been added and updated through a crowd-sourcing approach.

The software was created using an iterative, exploratory approach. When features were developed, they were published to the site for user feedback. This process guided the selection and implementation of future features of the software.

IsiXhosa.click is a website and hence accessible via all major platforms. The site is adapted for mobile use. No special software is required, so the dictionary may be accessed at any time via the Internet.

IsiXhosa.click allows users to suggest modifications to the word list. Users may click an edit button to capture missing information or correct errors and may also submit new words. Suggestions are queued for moderator review. This prevents vandalism and allows the implementation of a quality assurance process.

IsiXhosa.click is an isiXhosa-English dictionary with focus on usability. Words are easily discovered through a typo-tolerant, bilingual, live search. It is crowd-sourced, yet resistant to vandalism, allowing users to improve its database over time. Finally, it is open source, and may be freely adapted for future projects.

**References**

Wiley, D. & Hilton J. 2018. Defining OER-Enabled Pedagogy. International Review of Research in Open and Distributed Learning 19(4): 132-138.

Winkler, W. E. 1990. String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage. In: *JSM Proceedings, Survey Research Methods Section.* 1990. Washington, DC, United States of America: 354-359. Available: <http://www.asasrms.org/Proceedings/papers/1990_056.pdf>. Accessed on 21/04/2024.

1. [https://isixhosa.click](https://isixhosa.click/) [↑](#endnote-ref-2)
2. <https://github.com/quickwit-oss/tantivy> [↑](#endnote-ref-3)
3. <https://github.com/IsiXhosa-click/isixhosa_click> [↑](#endnote-ref-4)
4. <https://github.com/IsiXhosa-click/database> [↑](#endnote-ref-5)
5. The work was partially supported by the South African Centre for Digital Language Resources (SADiLaR). SADiLaR is a national centre supported by the Department of Science and Innovation as part of the South African Research Infrastructure Roadmap. [↑](#endnote-ref-6)