

LaMa, a thematic labelling webapp

placeholder1

Eindhoven University of Technology

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright
and release the work under a
Creative Commons Attribution 4.0
International License ([CC BY 4.0](#))

Summary

Qualitative analysis of data is relevant for a variety of domains, among which empirical research studies. When analysing large data sets such as StackOverflow, Apache Mailing Lists, and GitHub issues qualitatively, condensing pieces of data to a set of terms or keywords simplifies analysis, and helps in obtaining useful insight. This condensation of data can be achieved by associating keywords “labels” with text fragments “artifacts”. To achieve greater accuracy, and limit bias, collaboration and consensus building is essential. LaMa (Labelling Machine) is an open-source web application for thematic labelling of qualitative data. The application facilitates cooperation through simultaneous analysis by different researchers, as well as consensus building through conflict resolution techniques. The source code is available at: <https://github.com/muctadir/lama>

Statement of need

As mentioned previously, LaMa is a thematic labeling tool. The tool builds upon already existing software of the Labeling Machine ([Muctadir, 27 Apr. 2022](#)) and expands already existing functionality, as described in the Key features section.

Thematic analysis is a widely used technique with several popular tools already on the market, such as Atlas.ti (“[ATLAS.ti](#),” 2022) and maxQDA (“[All-in-One Qualitative & Mixed Methods Data Analysis Tool](#),” 2022). Analysing qualitative data has been proven to be a labor intensive and time consuming ([Pope et al., 2000](#)) due to its nature. Thematic analysis is a ‘powerful yet flexible method for analysing qualitative data’ ([Kiger & Varpio, 2020](#)). Analysing textual data through this method allows the user to understand experiences and thoughts, as well as emotions and behaviors throughout a data set. Due to the flexibility of this analysis method the users are not bound to using only one paradigmatic perspective but within different data sets can use different ones ([Braun & Clarke, 2006](#)).

LaMa facilitates thematic analysis through its unique features, making it stand out from the previously developed tools. Our tool is aimed at academic usage, with a possibility of expanding it. To tackle the limitations of other tools, LaMa introduces features that enable teamwork, allow the user to create themes to group labels and facilitate dynamic construction of the data set. With the help of this tool thematic analysis becomes less time consuming and more structured due to the flow that the tool enforces.

Key features

LaMa is in its core functionality similar to comparable labelling tools for qualitative analysis, however a number of key features ensure greater consensus and collaboration between users, as well as provide additional insight compared to existing labelling tools.

- Conflict resolution: To ensure consensus during collaboration automatic conflict detection has been implemented. Users can view the disagreement and are provided with a variety

- of methods to resolve the conflict.
- Dynamic entities: Complex input data and user defined constructs can be broken up into parts, or merged to ensure simplicity of analysis as well as coherence.
 - Themes: Lama allows users to group labels into themes, thereby providing a way of classifying data and the analysis of the data. In additional themes categorized hierarchically further aiding in the analysis and classification process.
 - Visualizing hierarchy: To obtain insight into how artifacts are mapped to labels, and labels ordered into themes, a visualization of the hierarchy can be constructed on demand.

Acknowledgements

The students of Eindhoven University of Technology, V. Bogachenkova, T. Bradley, E.J. Costa Martins, V. Fürst, B. Henkemans, J.L.O. Jansen, L.C. Lavin, L.V. Nguyen, and A. Olteniceanu, for designing the tool from scratch. The client of the project, H.M. Muctadir, and the supervisor of the project, A. Serebrenik.

All-in-one qualitative & mixed methods data analysis tool. (2022). In *MAXQDA*. <https://www.maxqda.com/>

ATLAS.ti. (2022). In *ATLAS.ti - The Qualitative Data Analysis & Research Software*. <https://atlasti.com/#>

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77–101. <https://doi.org/10.1191/1478088706qp063oa>

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE guide no. 131. *Medical Teacher*, 42(8), 846–854. <https://doi.org/10.1080/0142159X.2020.1755030>

Muctadir, H. M. (27 Apr. 2022). *Labeling machine tool*. <https://github.com/muctadir/labeling-machine>

Pope, C., Ziebland, S., & Mays, N. (2000). Analysing qualitative data. *BMJ*, 320(7227), 114–116. <https://doi.org/10.1136/bmj.320.7227.114>