IncluLex Hub: Accessing Language Simplification

Julia Sánchez Esquivel
Computer Science and Engineering Department
Universidad Carlos III de Madrid, Spain
julia.s.esquivel@alumnos.uc3m.es

ABSTRACT
In the Information Society, the access to information is crucial for social inclusion. Language, as a key means of communication and learning, can be a hurdle for those with intellectual disabilities, low education, or limited proficiency in the dominant language. To bridge this gap, content should be provided in Easy Reading for those with special needs, and in Plain Language for the general populace. These methods suggest the use of simple, common words and discourage technical jargon, foreign words, abbreviations, and acronyms without clear explanations. Adapting text content to these standards is becoming an increasingly recognized profession. Such adaptation is arduous, often involving the use of multiple resources like dictionaries, synonym dictionaries, and pictograms. To streamline this process, in this work a IncluLex Hub platform is proposed as a solution aiming to democratize linguistic information and make it universally understandable.

CCS CONCEPTS
Human-centered computing->Accessibility. Information systems->Data management systems. Computing methodologies->Artificial intelligence->Natural language processing.

KEYWORDS
Lexicon, accessibility, inclusion, complexity, frequency, synonym, pictogram, misinformation, technology.

1 Introduction
In today's digital era, accessing and comprehending information is crucial for all individuals. However, language can sometimes serve as a barrier rather than a gateway to knowledge, especially for those with intellectual disabilities, learning disabilities, limited education, or language proficiency. For example, in Spain, it is estimated that there are over 268,000 people with recognized intellectual disabilities, representing approximately 1% of the Spanish population. This figure encompasses individuals with varying degrees of intellectual disability [1]. These marginalized groups require content presented in Easy Reading, while the general population benefits from Plain Language. Guidelines advocate for the use of simple, familiar vocabulary and discourage the inclusion of foreign words, technical jargon, and abbreviations without clear explanations. Consequently, there is a growing demand for professionals specializing in adapting content to these accessibility standards. These professionals require many resources simultaneously, such as dictionaries of different levels of learning, word frequency dictionaries, synonym resources, pictogram resources, etc., and their work can become tedious when having to juggle numerous digital and non-digital resources to adapt the lexicon of a text to be easier from the point of understandability.

Recognizing this challenge, the IncluLex Hub is introduced as a practical tool to support social inclusion by lexical simplifying language. This platform seeks to make information more accessible by providing resources for Easy Reading and Plain Language. This poster will discuss the need for accessible language, the development of the IncluLex Hub, and how it serves as a step towards more inclusive access to information. It will explore the hub's functionality, the resources it aggregates, and its potential impact on various user groups. The IncluLex Hub aims to be more than just a utility for simplification; it is a commitment to ensuring that everyone can acquire knowledge and engage with society on equal footing.

This poster corresponds to a Final Degree Project for the bachelor's degree in computer science and engineering in the Universidad Carlos III de Madrid.

2. Proposal
The IncluLex Hub platform is introduced as an innovative solution designed to simplify the complex lexicon of our language. At its core, the platform offers an exhaustive analysis of words in an easily digestible format. For each term, it provides insights into its frequency of use within the language, evaluates its level of complexity to ensure easy comprehension, and offers a corresponding visual pictogram. Additionally, it suggests alternative synonyms that could simplify the word's meaning, supplies a precise definition, and illustrates how the word can be used across various contexts.

This multifaceted approach ensures that final users of adapted texts not only learn about the frequency and usage of a word but also understand its practical applications in real-life situations. The hub is more than a mere aggregator of linguistic

1 https://www.uc3m.es/bachelor-degree/computer-science
data; it is an advanced tool that integrates Machine Learning and Artificial Intelligence to enhance its capabilities. These technologies enable features such as the automatic detection of complex words and the generation of simpler synonyms, thereby streamlining the process of adapting texts into a easier lexicon.

3 Methodology

The development of the IncluLex Hub adopts a streamlined, engineering-focused approach. In the design phase for a responsive and accessible site. Agile development practices guide for efficient back end and database management, allowing quick adjustments based on feedback. Testing involves frequent, short user feedback cycles to promptly address issues, enhancing accessibility and usability. Deployment through iterative releases, with continuous integration, ensures the platform remains current and responsive to user needs and accessibility updates, maintaining its effectiveness.

3.1 Design and platform development

The development of the IncluLex Hub focuses on creating a website that is easy to use for as many people as possible. Using standard web tools like HTML5 for building the website structure, CSS3 for making it look good on all devices, and JavaScript to make the website interactive without needing to reload pages in web technologies based in the use of React. One of our main goals is to make sure the website is accessible to everyone, including people with disabilities. Following the Web Content Accessibility Guidelines (WCAG) [2] to achieve this. This means that the website will work well with screen readers, can be navigated using a keyboard, and will be easy to understand for people who might struggle with complex websites.

For handling the data, Python is utilized for organizing and fixing data before it is stored into a SQL database. This database is where all the information is stored so that users can search for it quickly and easily. Resources for the Spanish language are also included, like dictionaries [3], lists of common words and pictograms [4], to help Spanish speakers use the platform. References are provided to verify the reliability of this information. Lastly, Machine Learning and Artificial Intelligence [5] are added to the platform. These technologies will help automatically find difficult words and suggest easier ones. This way, the platform can help make texts simpler without much extra work. With all these parts working together, the IncluLex Hub will be a helpful resource for anyone looking to understand complex language, from students to professionals in adapting and simplifying texts, making sure that everyone has fair access to information.

3.2 Impact assessment

An impact assessment will be carried out, including several stages:

- Software Testing: Both automated and manual tests will be conducted to verify the platform’s technical performance.
- User Testing: Hands-on testing with a diverse group of users will assess the app’s usability, effectiveness, and overall experience, providing direct feedback from representative users.
- Feedback Collection: I will actively collect comments and suggestions to encourage user participation. This feedback will identify areas for improvement, highlight potential issues, and reveal positive aspects that will inform future iterations of the application. This comprehensive assessment will ensure a deep understanding of the application’s impact from multiple perspectives, guaranteeing adaptability and user satisfaction.

4. Conclusions and Expected Outcomes

The IncluLex Hub, developed for a Computer Science degree, is designed to streamline the text simplification process for professionals in the field. It provides an array of linguistic tools, including dictionaries that offer simpler synonyms, pictograms, complexity indicators, and definitions aimed at facilitating the creation of easily understandable content. This platform assists professionals who specialize in adapting texts to a more accessible language level. By providing a centralized hub of linguistic resources, the IncluLex Hub simplifies the workflow for text adaptation, making it more efficient and less time-consuming.

The goal is to support the enhancement of information accessibility, enabling users, especially those facing language challenges, to better understand and engage with content.

ACKNOWLEDGMENTS

To my tutor, Lourdes, I express my sincere thanks for her invaluable support and guidance during the development of this project. Their invaluable support, patience and dedication were critical to the success of this work.

REFERENCES


[4] ARASAAC. (s. f.). https://arasaac.org/developers/api