



First steps in the development of a support application for easy-to-read adaptation

Mari Carmen Suárez-Figueroa¹ · Isam Diab¹ · Edna Ruckhaus¹ · Isabel Cano²

Accepted: 24 October 2022 / Published online: 18 November 2022
© The Author(s) 2022

Abstract

The application of the easy-to-read (E2R) methodology is one of the ways to achieve cognitive accessibility and specifically, it is a path that guarantees the right of access to information of people with reading comprehension difficulties and thus improves their daily life. This methodology includes a set of guidelines and recommendations whose goal is to present clear and easily understood documents. Such guidelines are used in the manual processes of (a) adapting existing documents and (b) producing new materials. These processes are very time and human-resource consuming, due to the need of involving E2R experts as well as people with cognitive disabilities. In order to alleviate such manual processes, we are currently investigating the development of methods, based on Artificial Intelligence (AI) techniques, to support the E2R adaptation of documents in a (semi)-automatic fashion. The main goal of this research is to help E2R experts in their daily tasks of (a) assessing a particular document with respect to the E2R guidelines and (b) transforming such a document according to the E2R methodology. In this paper we present our initial efforts toward the development of an AI-based application for supporting the E2R adaptation of documents. These efforts are the elicitation of E2R needs and informal requirements and the design of an application called FACILE.

Keywords Easy-to-read methodology · Cognitive accessibility · Artificial intelligence

1 Introduction

People with disabilities have the right to participate in social activities, such as politics, education, work, and culture, in the same way as other people¹. Given that all of these activities involve the access and comprehension of written content, a way to achieve this right is to improve various aspects of content accessibility and cater to the needs of people with disabilities. In fact, there are more and more laws,² decrees³ and regulations,⁴ both local, national and European, related to accessibility, which explicitly mention the need for documents to be written in simple, plain, clear and direct language. This need is directly related to the creation of documents following the guidelines included in the easy-to-read (E2R) methodology [8, 9, 18, 19]. The E2R

methodology provides guidelines and recommendations for writing texts and making materials that are easy to understand by people with reading comprehension difficulties. However, this methodology presents two main limitations: (a) it lacks clear information on how to apply the guidelines and recommendations; and (b) it relies heavily on manual adaptation, a highly time-consuming and subjective task that naturally prevents the massive and consistent creation of easy-to-read materials. Such manual adaptation implies three key activities: (1) analysis of which E2R guidelines are fulfilled in the document; this can be seen as a kind of assessment activity; (2) transformation of the document by E2R experts based on the guidelines that are not satisfied, which are discovered during the previous assessment; and

✉ Mari Carmen Suárez-Figueroa
mcsuarez@fi.upm.es

¹ Ontology Engineering Group (OEG), Universidad Politécnica de Madrid (UPM), Madrid, Spain

² ACCEDES, Madrid, Spain

¹ Ley General de derechos de las personas con discapacidad y de su inclusión social. Versión en lectura fácil. (Diciembre 2015) (Spanish)

² <https://short.upm.es/ghaye> (Spanish) (Accessed on 28th July 2022).

³ <https://short.upm.es/r9tbv> (Spanish) (Accessed on 20th October 2022)

⁴ Convention on the Rights of Persons with Disabilities (CRPD). Article 9—Accessibility (<https://short.upm.es/s50az>) (Accessed on 28th July 2022).

(3) validation of the transformed document by a professional team comprised by people with intellectual disabilities.

The need to have understandable content and accessible information by people with learning difficulties has received increased attention in the last decade [14]. Currently, the explicit need for easy-to-read materials is mainly present in public administrations; it is expected that this need will progressively be extended to companies and private organizations in different sectors such as education, culture and leisure, transport, and any area of social life in which documents are used to present and disseminate information. In this mid-term scenario, the demand for E2R experts who adapt and/or create easy-to-read documents will increase exponentially. In order to face this demand for experts, as well as to alleviate the workload for adapting and creating easy-to-read documents, an application that supports the E2R methodology would be very useful and beneficial to help both the analysis and transformation of documents.

Previous research works on E2R analysis include semi-automatic validators that check E2R compliance in German [17, 20] and an E2R conformance checker for Spanish called Easy-to-Read Advisor [26]. Regarding E2R transformations, we include here previous works for Spanish texts, since our research is focused on that language. It is worth mentioning Simplext [24], LexSIS [4], DysWexia [22], and EASIER [15], that are based on simplification techniques. Most of these approaches pay attention only to E2R guidelines related to writing aspects (spelling, grammar, vocabulary, and style), while aspects related to the design and layout of a document have not been considered. In addition, most of the approaches are targeted to people with cognitive and learning disabilities and people with low literacy skills, but do not consider the scenario of supporting E2R experts in their daily work.

To cover these identified gaps, our research is focused on applying different Artificial Intelligence (AI) methods and techniques to (semi)-automatically perform both the E2R analysis and transformation of documents written in Spanish. The main objective of this research is to allow E2R experts to focus on complex issues and decisions when adapting or creating E2R documents, transferring the identification and transformation of simpler situations to our proposed application, called FACILE. This technological aid will imply a reduction in the effort dedicated by E2R experts to the process of creating or adapting a document. Specifically, in this paper we present our first steps toward having an application for (semi)-automatically performing (a) an E2R diagnosis of documents and (b) a recommendation on the transformations of original documents so they can comply with E2R guidelines. Such first steps are (1) the identification of needs and informal requirements involving E2R experts and (2) the creation of the initial design and mock-up of FACILE.

The rest of the paper is organized as follows: Section 2 is devoted to the state of the art on the E2R methodology and on the existing automatic approaches for checking E2R compliance and for transforming documents into easy-to-read versions. In Section 3 we describe how we gather needs and informal requirements from E2R experts. Our first attempt to design an AI-based application for supporting the E2R adaptation of reading materials is briefly presented in Section 4. Finally, Section 5 shows the conclusions and the future research lines of work.

2 State of the art

Legislation⁵ related to accessibility encourages the use of easy-to-read (E2R) guidelines as a fundamental right for equal access and inclusion. The E2R Methodology [8, 9, 18, 19] was created to improve the daily life of people with difficulties in reading comprehension. This methodology aims to present clear and easy to understand content to different sectors of the population that include people with disabilities and people with limited language or reading proficiency, among others. Other guidelines with the same goal have been recently published: (a) the first public working draft of Web Content Accessibility Guidelines (WCAG) 3.0⁶ that includes specific recommendations for making web content more accessible to users with cognitive and learning disabilities; and (b) a detailed report for providing advice on how to make content usable for people with cognitive and learning disabilities.⁷ In addition to the aforementioned guidelines, technology is required to support the labor-intensive and costly process of manually creating and adapting reading materials to make them cognitively accessible.

The rest of this section is devoted to describing the E2R methodology and to summarizing the (semi)-automatic approaches that support the assessment and transformation of materials in conformance with the E2R guidelines.

2.1 The easy-to-read methodology

The easy-to-read (E2R) methodology is a crucial element to enhance cognitive accessibility. The main goal of the E2R methodology [8, 9, 18, 19] is to present clear and easily understood documents. This methodology includes a set of

⁵ Convention on the Rights of Persons with Disabilities (CRPD). Article 9—Accessibility (<https://short.upm.es/s50az>) (Accessed on 28th July 2022).

⁶ <https://www.w3.org/TR/wcag-3.0/> (Accessed on 28th July 2022).

⁷ Making Content Usable for People with Cognitive and Learning Disabilities (<https://www.w3.org/TR/coga-usable>) (Accessed on 28th July 2022).

guidelines and recommendations that affect writing of texts, supporting images, design and layout of documents, and final editing format. Oscar García made a comprehensive compilation based on what has been published so far about the easy-to-read methodology [8]. This document presents a collection of E2R guidelines that serves as a reference for E2R adaptors. Recently, the Spanish Association for Standardization (UNE) has published the experimental standard “UNE 153101 EX: Easy to read. Guidelines and recommendations for the elaboration of documents”⁸, which is the first technical norm related to the E2R methodology. This UNE standard was created due to the expansion of Easy-to-Read, with the goal of having a technical regulation that would agree on the guidelines and recommendations published to date and modernize them so that they would be adapted to the current reality. The norm reorders and structures the guidelines and recommendations that were already collected in reference publications [8, 9, 18]. In this regard, this UNE standard is consistent with existing initiatives [9, 18]. It is worth mentioning that such a technical norm is related to Federal Plain Language Guidelines [21] with respect to their main objectives that in both cases are to have documents easy to read and to understand. However, the UNE standard is focused on people with cognitive and learning difficulties as well as people with reading comprehension problems, while plain language guidelines are used to reach a general audience.

The E2R guidelines refer to (a) the use of simple and short sentences; (b) the non-inclusion of metaphors, technicalities, abbreviations, acronyms, among others (as far as possible); (c) the use of suitable typography; (d) the inclusion of images that reinforce the message and clarify the content; and (e) the selection of an editing format that is easy-to-use and suitable for the expected use of the document; among others. Such guidelines are organized in a set of categories [8]: in the field of writing the following types of guidelines have been included: spelling, grammar, lexicon, and style; and in the design and layout the following type of guidelines have been established: illustrations, typography, text composition and pagination.

The E2R guidelines and recommendations are used in the manual processes of (a) adapting existing documents and (b) producing new materials. These processes are laborious and consume a large amount of human resources due to the need to involve both E2R experts and people with cognitive disabilities. The process of manually adapting existing documents is cyclic and implies three activities: (1) the analysis of which guidelines of the methodology are satisfied and which are not in the document to be adapted; (2) the

transformation of the material by E2R experts, based on the deficiencies detected in the analysis activity with respect to the methodological guidelines; and (3) the validation of the transformed document carried out by a team of people with intellectual disabilities trained for this purpose, called validators. Analysis, transformation, and validation activities are complementary activities to achieve high quality reading materials that are effective in terms of cognitive accessibility.

The E2R methodology presents two main limitations: (a) it lacks clear information on how to apply the guidelines and recommendations; and (b) it heavily relies on manual adaptation, a highly time-consuming and subjective task that naturally prevents the massive and consistent adaptation of texts.

It should be stressed that our research is based on the fact that E2R is a priori a methodology that addresses the needs of people with reading comprehension difficulties, and that it is based on needs that have been determined through real-life work experience with people with these difficulties. However, there do exist studies developed to assess the effectiveness of the methodology on the population with reading-comprehension difficulties [10].

2.2 Automatic approaches for checking E2R compliance and transforming documents

A small number of efforts exist to develop automatic approaches both for checking E2R compliance and for transforming documents into easy-to-read versions. Before presenting automatic approaches for checking E2R compliance and transforming documents, it is worth mentioning some manual efforts to know whether a document is compliant with the E2R methodology. In this regard, there are two main manual techniques: (a) using checklists, as for example the one⁹ created in the PUZZLE project, and (b) involving groups of people who discuss the E2R requirements satisfied in a document. However, manual approaches are labour-intensive and costly. In order to decrease both time and resources spent for checking E2R guidelines and transforming documents, semi-automatic tools can be used.

Research on automatically **checking E2R guidelines compliance** in reading materials is quite scarce. It is worth mentioning a couple of tools [17, 20] that review German texts with respect to E2R guidelines, and a tool for Spanish called Easy-to-Read Advisor [26]. The tool presented in [17] is based on an approach to empirically evaluate E2R guidelines in German documents; while VisRA [20] has the aim of supporting the writer in the task of revising a German text

⁸ <https://goo.gl/Gmsmwj> [19] for Spanish, English and French. (Accessed on 28th July 2022).

⁹ http://www.puzzle-project.eu/docs/EN/IO1/IO1_EtRLearningMaterial_Checklist.pdf (Accessed on 28th July 2022).

with respect to E2R guidelines. The Easy-to-Read Advisor is an E2R conformance checker for assessing educational documents in the form of slides written in Spanish; these slides have been created in HTML. Easy-to-Read Advisor provides as output a detailed report on the E2R guidelines that are satisfied (marked in green) and the ones not satisfied (marked in red). This advisor, developed as a proof of concept, implements 22 E2R guidelines: ten of them are related to design and layout and twelve of them are related to writing. To the best of our knowledge, there were no specific E2R validators for Spanish documents before the development of Easy-to-Read Advisor. Furthermore, there is a writing assistant called LanguageTool¹⁰ that checks texts in several languages for style and grammar issues. LanguageTool includes some rules for E2R assessment (e.g., text uses capital letters according to the general rule, and text is made up of short sentences), but it is not geared exclusively to E2R checking. Similarly, there are two interesting and relevant applications for Spanish that are not focused on E2R but on plain language: Clara¹¹ and arText.¹² On the one hand, Clara analyzes the clarity of Spanish texts applying computational linguistics and machine learning techniques; on the other hand, arText is a writing assistant for texts in plain language.

Regarding the idea of **transforming reading materials based on E2R guidelines** in an automatic way, a relevant approach that can be applied is the so-called text simplification. Text simplification is the “task of modifying the content and structure of a text in order to make it easier to read and understand, while retaining its main idea and approximating its original meaning” [1]. In general terms, we can divide the E2R guidelines in which the approaches have focused on, into two groups: (a) lexical simplification and (b) syntactic simplification. In the first group, E2R guidelines such as avoiding the use of complex words, abbreviations and acronyms, and avoiding anaphoric expressions can be mentioned, while E2R guidelines like avoiding complex sentences and the use of the passive voice can be treated by syntactic simplification. Concerning lexical simplification, most of the approaches agree on replacing difficult words by synonyms using different linguistic resources such as WordNet [2], Cornetto [5], CanooNet [27], or TeP 2.0 [25]. In order to determine whether a particular word is difficult, its level of complexity is measured by its frequency of use [2, 5, 25], by using vocabulary lists [11, 12] or by considering the number of characters [27]. This idea of applying lexical simplification for obtaining E2R materials can be illustrated by

the following tools: LexSIS [4], DysWexia [22], and EASIER [15]. LexSIS and DysWexia are focused on synonym substitution in Spanish and their target are Spanish-speaking persons with dyslexia; EASIER¹³ is a web application that provides synonyms, definitions and pictograms for the complex words detected in textual content in order to improve understanding and readability. With respect to syntactic simplification to avoid complex sentences, all the approaches opt for the division into simpler ones by means of removing relative pronouns (e.g., *who*, *which*) and discursive markers (e.g., *however*, *nevertheless*, *despite*) [6, 11–13, 25, 27–29]. Another kind of syntactic modification is related to the transformation of sentences with passive voice into sentences with active voice [11, 23, 25].

The previous approaches only address either lexical simplification or syntactic simplification. Some researchers have worked on both forms of simplification; this is the case of (a) Simplext (Spanish) [24] and YATS (English) [7] that uses Natural Language Processing (NLP) techniques for semantic simplification and rule-based techniques for syntactic simplification; and (b) Lee and colleagues [12] who developed a customizable text simplification browser-based editor that facilitates user post-editing, for example in choosing candidate substitutions or undoing sentence splits. Another related example is the capito application,¹⁴ which translates German texts¹⁵ into three language levels: A1 (short and simple), A2 (easily understandable) and B1 (colloquial language); these levels allow to have texts in an easy-to-understand language.

As shown, several efforts have been made with the purpose of having automatic systems that analyze and transform texts in an easier to read and understandable format. Most of the research groups focus on E2R guidelines concerning writing aspects (vocabulary, spelling, grammar and style), instead of paying attention also to design and layout considerations included in the E2R methodology. Only one group [27] expands its scope and offers a broader perspective considering not only content and language guidelines, but also transforming the text at the design and layout levels¹⁶. In addition, most of the approaches are targeted to help people with cognitive and learning disabilities and people with low literacy skills, but do not consider the scenario of supporting E2R experts in their daily work. In order to cover these identified gaps, our research work emphasises on (a) all the aspects mentioned in the E2R methodology, (b) both the

¹⁰ <https://languagetoolplus.com/http-api/#/default> (Accessed on 28th July 2022).

¹¹ <https://clara.comunicacionclara.com/> (Accessed on 28th July 2022).

¹² <http://sistema-artext.com/> (Accessed on 28th July 2022).

¹³ <http://163.117.129.208:8080/> (Accessed on 28th July 2022).

¹⁴ <https://www.capito.eu/en/the-capito-app/> (Accessed on 19th October 2022).

¹⁵ Application is focused on German, but other languages for selected documents are also provided (English, Italian, Spanish, Polish, and French).

¹⁶ They base the transformation on the guidelines proposed by Maaß [3] for simplified German.

analysis and transformation of documents, and (c) the target group of E2R experts.

3 Requirements identification: involving E2R experts

As mentioned in Sect. 1, our research work is focused on applying Artificial Intelligence (AI) to (semi)-automatically perform both the analysis and the transformation of documents with respect to the E2R Methodology. With the goal of improving and extending the proof of concept Easy-to-Read Advisor [26], and thus, expanding our research work, we decided to involve stakeholders, in particular E2R experts, in the process of gathering needs and requirements for our application called FACILE. With this in mind, we launched a survey to obtain the opinion of E2R professionals on (a) the layout and writing guidelines included in the proof of concept Easy-to-Read Advisor that could be reused in FACILE, and (b) further E2R guidelines that are deemed necessary for their inclusion in FACILE. This section is organized as follows: first we describe the questionnaire; then information about survey participants is provided; then, the survey data are described; finally, we provide some interpretations and conclusions over the gathered data.

3.1 Questionnaire design

The questionnaire is divided into three main parts: (1) a part that includes questions related to the usefulness of an application for checking E2R conformance and for suggesting E2R adaptations; this part also contains questions about missing E2R guidelines in the proof of concept Easy-to-Read Advisor¹⁷ that are also needed in the proposed application FACILE; (2) a part with questions regarding which functionalities or features should be additionally implemented in FACILE; likewise, participants were asked for their particular comments, opinions, and feedback; and (3) a final part with questions related to the participants' knowledge, background and experience (data gathered in this last part are summarized in Sect. 3.2).

The questionnaire is written in Spanish and implemented as a Google Form.¹⁸ It was launched in July 2020 through mailing lists of E2R experts in Spain.

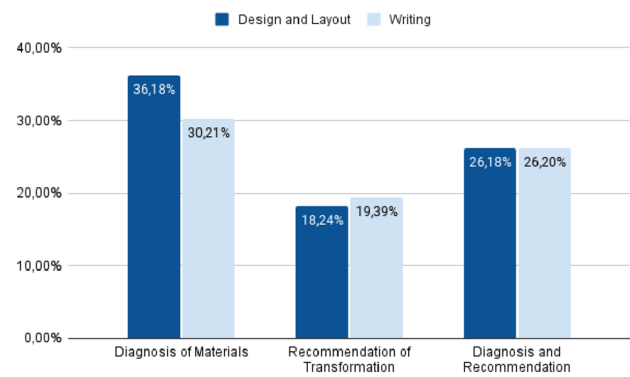


Fig. 1 Percentages of responses about the usefulness of an application for checking E2R conformance and for suggesting E2R adaptations

3.2 Survey participants

A total of 34 participants, mostly Spanish (94.1%), provided responses to the questionnaire. The sample consisted of professionals in the field of cognitive accessibility (70.59%) and communication (29.41%). Most of the participants (82.4%) were women, whereas only 14.70% were men. Regarding the age range, half of the participants ranged from 31 to 45 years old, 38.2% from 46 to 60, and the remaining 11.8% were over 60 years old. With respect to their experience in the adaptation of reading materials, 38.2% worked more than 5 years in this field, 17.6% between 3 and 4 years, and the rest of the participants had less than 3 years of experience.

3.3 Survey data

The two parts regarding the E2R experts' opinions, mentioned in 3.1, are described in detail in this section.

Questionnaire Part 1. In the first place, we collected opinions and comments on two main questions: (a) how useful would an application be that indicates which E2R guidelines a given document does not meet; and (b) how useful would an application be that indicates which transformations can be developed to make a document compliant with E2R guidelines. After analysing the collected data, 76.47% of the participants (see the Appendix A¹⁹) selected a score of 5 (strongly agree) in both questions, so we can say that there is a clear favourable trend toward having an application both for E2R guidelines assessment and for the recommendation of transformations.

¹⁷ Participants are provided with the set of E2R guidelines, both on writing and layout, considered in Easy-to-Read Advisor.

¹⁸ <https://doi.org/10.5281/zenodo.7229161>.

¹⁹ <https://doi.org/10.5281/zenodo.7229161>.

Table 1 E2R guidelines implemented in easy-to-read advisor [26]

Design and layout guidelines	Writing guidelines
1. The text font belongs to the accepted styles	1. The size of the lines is correct
2. The font size must be large enough	2. Large numbers are expressed correctly
3. Text avoids italics	3. Text avoids the use of special characters
4. Text includes moderately bold words	4. Text avoids the use of ordinal characters
5. Text includes moderately underlined words	5. Text is made up of short sentences
6. Text avoids typographic effects	6. Dates are written in full
7. Text uses capital letters according to the general rule	7. The use of pronouns is correct
8. The contrast between text and background is correct (focused on text color)	8. Text avoids the use of Roman numbers
9. The contrast between text and background is correct (focused on background)	9. Text is written in the second person
10. The amount of words in the text is correct	10. Text avoids passive voice
	11. Sentences have a subject
	12. Text is made up of simple sentences

In addition, Fig. 1²⁰ shows the percentage of participants that consider useful to have in an application the option of automatically checking E2R compliance in the design and layout of the document, i.e., 36.18%. The percentage that considers E2R assessment useful for writing guidelines is 30.21%. Regarding the issue of having the option of a (semi)-automatic recommendation for transforming documents into an easy-to-read version, 18.24% of the participants consider this useful for design and layout guidelines, while 19.39% think that this option is useful for writing guidelines. Additionally, regarding the question of including guidelines in both tasks, assessment and suggestion of transformations, percentages for both groups of guidelines are similar, 26.18% for design and layout guidelines and 26.20% for writing guidelines.

Participants were also asked which E2R guidelines could be useful for their daily work and are missing in the former proof of concept Easy-to-Read Advisor [26] (either related to diagnosis of documents or recommendations of transformations). E2R experts were provided with the set of implemented guidelines in Easy-to-Read Advisor (see Table 1). The main goal of this part of the questionnaire is to gather which E2R guidelines could be interesting to include in FACILE. For this purpose, a controlled proposal was set up for the participants, composed of: (a) a set of 10 design and layout guidelines, and (b) a collection of 22 writing guidelines (see Table 2). These guidelines were selected from [19] since this technical norm is currently being used in Spain by E2R experts as a base document. The selection of the list of guidelines was performed in collaboration with an E2R expert. In addition, it is worth mentioning that this list has no overlap with the implemented guidelines in

Easy-to-Read Advisor. For each guideline in the aforementioned controlled proposal, participants were asked to select whether they would include the guideline in the diagnosis (or assessment) task, in the recommendation of transformation task, or in both tasks.

After analysing participants' responses, we have created the rankings of the selected guidelines with respect to the three tasks: diagnosis of materials, transformation, or both tasks at the same time.²¹ Such rankings are shown in Figs. 2, 3 and 4, respectively, and will be explained more specifically in Sect. 3.4. In general terms, these three rankings indicate the guidelines to which professionals pay most attention and thus should lead the development of FACILE. Such guidelines will be presented in more detail in Sect. 4.

Apart from collecting E2R experts' opinions using a controlled set of guidelines, survey participants had the freedom of commenting which other E2R guidelines should be implemented in FACILE for supporting the E2R adaptation of reading materials. After collecting all the comments the main outcomes are the following:

- Regarding Design and Layout Guidelines, 50% of the participants emphasised the treatment of images: there must be a hidden description for people who have visual impairments, writing inside images should be avoided, and images should have a plain background; 33.33% refer to the format of the text: using colours for the division of topics, avoiding the use of italics and bold, or excluding the use of underlining; the rest of the comments propose slight modifications to the guidelines already included in the controlled set of guidelines such as text justification or page number style;

²⁰ It is important to note that several participants did not choose any of the three proposed options (guidelines to be included in the diagnosis task, guidelines to be included in the recommendation of transformation task, or guidelines to be included in both tasks). For this reason, percentages do not necessarily sum up to 100.

²¹ See Appendix A to consult the tables upon which the analysis is based and an illustrative example of the rankings depicted as scatter plots (<https://doi.org/10.5281/zenodo.7229161>).

Table 2 Controlled proposal of E2R guidelines envisaged to be included in FACILE

Design and layout guidelines	Writing guidelines
DL1. Text should be left-aligned	W1. Do not capitalise words or phrases, except in the case of acronyms
DL2. Text justification should be avoided	W2. Initial capital letters should be used at the beginning of a paragraph or title, after a full stop or in proper nouns
DL3. Avoid hyphenating the words at the end of a line	W3. The full stop must be replaced by a new paragraph or by a conjunction
DL4. All pages of the document must have the same orientation (landscape or portrait)	W4. A colon (:) must be used when entering a list that enumerates more than three items or when entering a dialogue
DL5. Pictures should be placed close to the text	W5. The semicolon (;) should not be used
DL6. Pictures must agree with the text and concretely depict what is described in the text	W6. Avoid the use of etcetera or ellipses
DL7. Pictures cannot be separated between pages	W7. Inverted commas should not be used (Spanish «» or English “ ”)
DL8. Text written in portrait orientation should be avoided	W8. Adverbs ending in -mente (-ly in English) should be avoided
DL9. The page number must be included on the outside of the page, at the bottom or at the top	W9. Superlatives should be avoided
DL10. Minimum line spacing should be 1.5"	W10. The use of words with undefined content such as: thing, something or subject should be avoided
	W11. The telephone numbers must be separated by blocks
	W12. The use of fractions and percentages should be avoided
	W13. Avoid writing the time in 24-hour format
	W14. Passive reflexive should be avoided
	W15. If possible, the use of sentences with gerunds should be avoided
	W16. The use of two or more consecutive verbs should be avoided, except for periphrases with the modal verbs
	W17. The use of complex connectors between sentences, such as therefore, nevertheless or consequently should be avoided
	W18. The text should be made up of short, simple sentences
	W19. The use of pronouns must be correct
	W20. The text must be written in the second person
	W21. Passive voice should be avoided
	W22. Sentences must have a subject

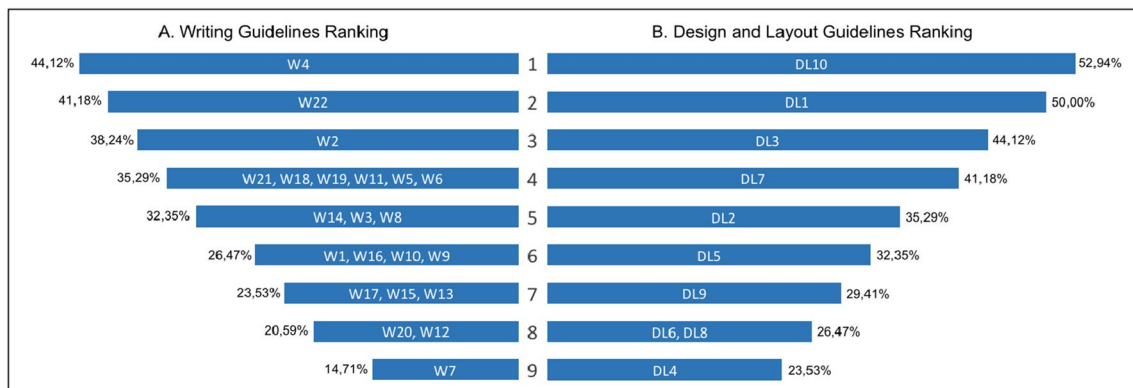


Fig. 2 Ranking of the guidelines in the diagnosis of materials choice. **A** shows the ranking for writing guidelines, while **B** shows the ranking for design and layout guidelines

– With respect to Writing Guidelines, most of the comments (37.5%) refer to avoiding the use of abbreviations in texts; the remaining opinions allude to the occasional

use of subordination -since a text written only in short sentences can be robotic-, respecting spelling rules, or avoiding the use of foreign words.

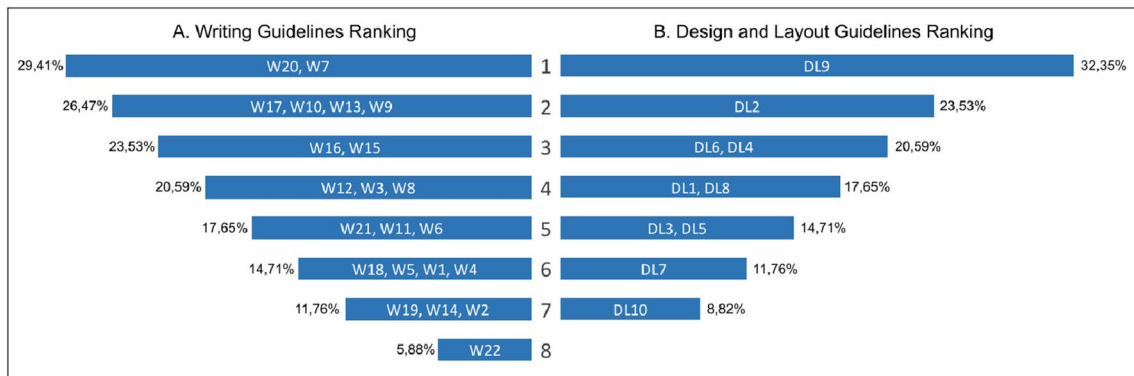


Fig. 3 Ranking of the guidelines in the recommendation for transformation choice. **A** shows the ranking for writing guidelines, while **B** shows the ranking for design and layout guidelines

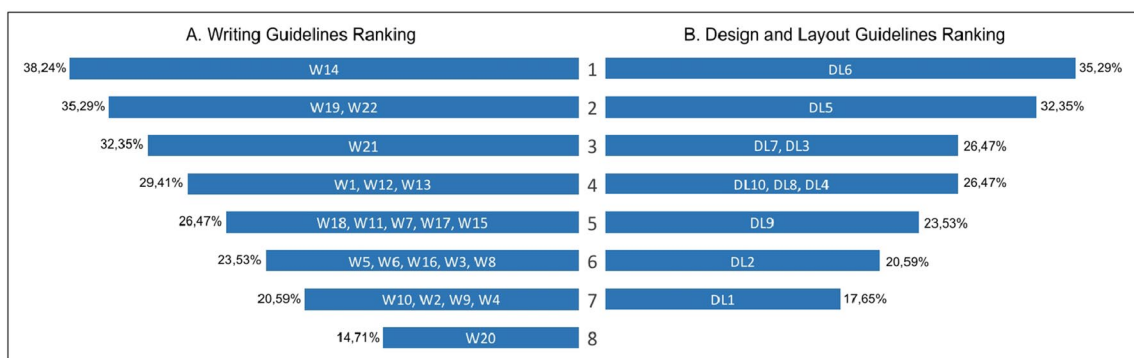


Fig. 4 Ranking of the guidelines for the option of having both the E2R analysis and the subsequent suggestion of transformations. **A** shows the ranking for writing guidelines, while **B** shows the ranking for design and layout guidelines

Questionnaire Part 2. Participants were asked which functionalities or features could be worthwhile to include in FACILE. As a general comment, almost half of the respondents (43.33%) agree that the main feature is to have an effortless application. They stress that the application should be accessible to any user, intuitive, dynamic and easy to use. They also mention that it should be able to accept any type of document as input such as slide presentations, PDF files or text documents. In addition, it is worth mentioning that 36.67% of the participants highlight that such an application should analyse which are the E2R guidelines the text does not comply with, and suggest a transformation. In contrast, 6.67% of the comments refer to an automatic correction of texts that are non-compliant with the guidelines.

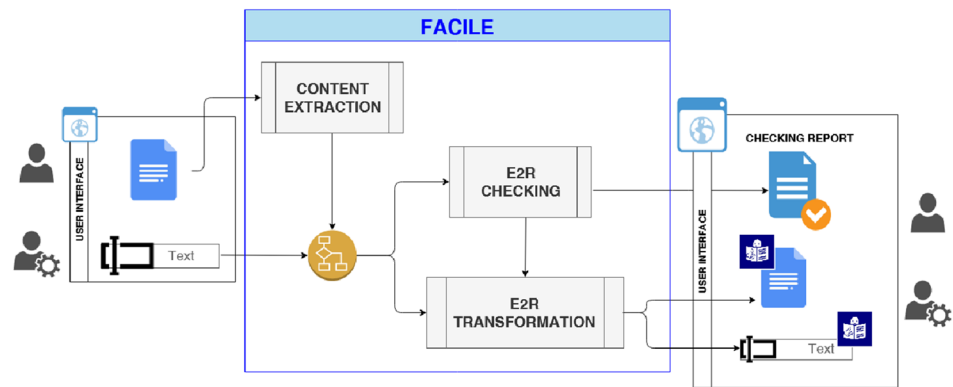
3.4 Survey data interpretation and conclusions

Considering the high scores obtained on the two first questions set out at the first part of the questionnaire (76.47% of the participants selected a score of 5, as mentioned in Sect. 3.3), our main outcome is that an

application that supports the process of adapting materials to E2R guidelines is very useful. In general, professionals prefer an application that identifies the guidelines that the document does not comply with and provides suggestions of E2R transformations, instead of having an application that directly performs an automatic transformation.²² Nevertheless, for those E2R guidelines which do not necessarily require extensive human intervention, survey participants prefer the application to recommend a transformation (e.g., DL9 in design and layout guidelines, and W7 regarding writing guidelines). For the rest of the guidelines, those that imply more elaborated modifications, participants prefer the application to inform them about the situation and let them manually transform the materials in relation to the information provided by the application. This means that E2R experts are more in favour of having a support application that performs

²² Such statement is based on the verbatim opinions of the participants in the questionnaire: 36.67% of the comments agree that the application should suggest possible options from which the experts select the transformation that best suits their needs.

Fig. 5 FACILE flowchart



diagnosis and recommendation for transformations in the context of the E2R adaptation process.

Furthermore, as remarkable points when looking at the rankings presented in Figs. 2, 3 and 4 we can observe the contrast between some of the guidelines in relation to the three tasks. For instance, W22 in Fig. 3 (transformation) came last in the rankings according to the participants' preferences, while in Fig. 2 (diagnosis of materials) it came second. The same applies to the DL10 guideline, which ranks last in the transformation task, while it is the most chosen in diagnosis. However, continuing with the first example, we see how participants tend to prefer the W22 guideline to be treated under both tasks (as we see in Fig. 3, the guideline came second). Thus, we can notice that guidelines with a low score related to the recommendation of a transformation have a high score when chosen for both diagnosis and transformation (e.g., W14 is ranked 8th (11.76%) concerning transformation, whereas it is the most selected guideline (38.24%) for diagnosis and transformation). Therefore, taking these results into account, we find that these guidelines, selected for both diagnosis and transformation, are the ones on which we should focus on, regarding the development of our application FACILE.

Thanks to the data extracted from the comments and opinions of these professionals, we can create a plan to be followed in further development steps. Since the users prefer a first evaluation of the guidelines before performing any transformation, we may underline that our application is a support tool oriented to help the manual process of adapting the materials to the E2R methodology.

4 A proposal for supporting the E2R adaptation of reading materials

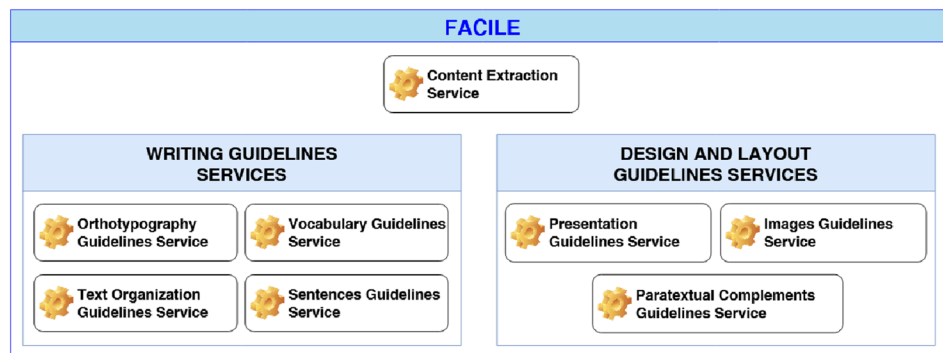
The ultimate goal of our research is to develop an application that performs a (semi)-automatic process of adapting documents to E2R guidelines. Thus, our proposal consists

of creating a scalable and extensible web application that allows E2R experts to obtain (a) a diagnosis about the level of E2R compliance of a specific reading material, and (b) a set of recommendations on the transformations of original reading materials so they can comply with E2R guidelines.

The proposed web application, called FACILE, is an extension and improvement of Easy-to-Read Advisor [26]. Our proposal is that FACILE not only checks E2R guidelines compliance in reading materials as Easy-to-Read Advisor, but also provides suggestions of transformations in a (semi)-automatic fashion. Developing FACILE as a diagnosis and recommendation application was decided based on the survey findings presented in Sect. 3.4.

The conception we have of FACILE is that its users are able to choose between (a) introducing a document or (b) entering a text; in both cases the language is Spanish. Additionally, users can decide whether to (a) perform an assessment of E2R guidelines or (b) request suggestions for E2R modifications. Another key point in our conception is that FACILE allows users to decide which E2R guidelines are taken into account in both the assessment and suggestion processes. Figure 5 shows the general flowchart of FACILE. The input in FACILE is either a document or a text; if the input is a document, the application should make a content extraction process whose goal is to gather all the presented elements in the input document (design and layout elements as well as content parts, which include text and graphical elements). Using the input text or the content extraction information, FACILE can perform (a) an E2R analysis of the input that provides as output a report with the set of satisfied and non-satisfied guidelines, (b) a suggestion of transformations that returns different possibilities of how to modify the original document or text to be compliant with E2R guidelines, or (c) both processes at the same time. The specific option is carried out based on the user selection.

Fig. 6 FACILE general service-based architecture



We have designed FACILE following a microservice-based architecture,²³ as shown in Fig. 6. Each microservice has a very specific purpose and responsibility, significantly different from the rest, related to a particular type of E2R guideline. Thus, FACILE's architecture is highly cohesive. Each service refers to a different E2R guideline category from those presented in [19]; hence, the FACILE architecture is composed of seven services: orthotypography, vocabulary, sentences, text organization (these four services are related to writing guidelines), presentation, images, and paratextual complements (these last three services are related to design and layout guidelines). Each service is composed of two methods per E2R guideline: one for E2R checking, which performs the assessment of the input (both text or document) with respect to a specific E2R guideline; and the other for proposing E2R transformations, in charge of providing suggestions for the transformation of the original input, non-compliant with a particular E2R guideline, into an adapted output which is easy-to-read.

In addition to the aforementioned collection of E2R services, the architecture of FACILE contains a service for performing the content extraction for those cases in which the input is a document, as shown in Fig. 6. This service deals with all the operations related to obtaining the whole information about the input document, that is, data about the design and layout of the document as well as about its content (including text and graphical elements). All these services communicate with each other through APIs (Application Programming Interfaces), whose responses are based on the JSON format [16]. We have designed FACILE services based on Artificial Intelligence (AI) techniques such as Rules, Natural Language Processing, and Machine Learning.

In order to clarify which types of E2R guidelines are included in each of the aforementioned services, we provide examples of writing guidelines in the different services:

- Orthotypography: Do not capitalise words or phrases, except in the case of acronyms (W1) or Inverted commas should not be used (W7);
- Vocabulary: The use of fractions and percentages should be avoided (W12), Avoid writing the time in 24-hour format (W13), or The telephone numbers must be separated by blocks (W11);
- Sentences: Passive reflexive should be avoided (W14) or The text should be made up of short, simple sentences (W18);
- Text organization: The text must be written in the second person (W20).

In addition, we list examples of design and layout guidelines for the three services on such a dimension:

- Presentation: The font size must be large enough and the contrast between text color and background color must be correct; these guidelines are already implemented in Easy-to-Read Advisor [26]; additional examples are Avoid hyphenating the words at the end of a line (DL3) and Minimum line spacing should be 1.5" (DL10);
- Images: Pictures must agree with the text and concretely depict what is described in the text (DL6) and Pictures should be placed close to the text (DL5);
- Paratextual complements: Notes describing difficult-to-understand words must be provided [19].

FACILE is conceived as an improvement of Easy-to-Read Advisor [26], since the latter is based on a client–server architecture with two layers (presentation and application), while FACILE is modelled following a microservice-based architecture. This decision implies a modular application that is flexible in its use as well as in its maintenance. Furthermore, FACILE is an extension of Easy-to-Read Advisor [26], since our plan is that the E2R guidelines already implemented in Easy-to-Read Advisor will be integrated in FACILE. For this purpose, we need to perform a re-engineering process in order to transform the Easy-to-Read Advisor code

²³ <https://martinfowler.com/articles/microservices.html> (Accessed on 28th July 2022).



Fig. 7 Screenshot of the FACILE mockup that shows one option in which the user can provide the input

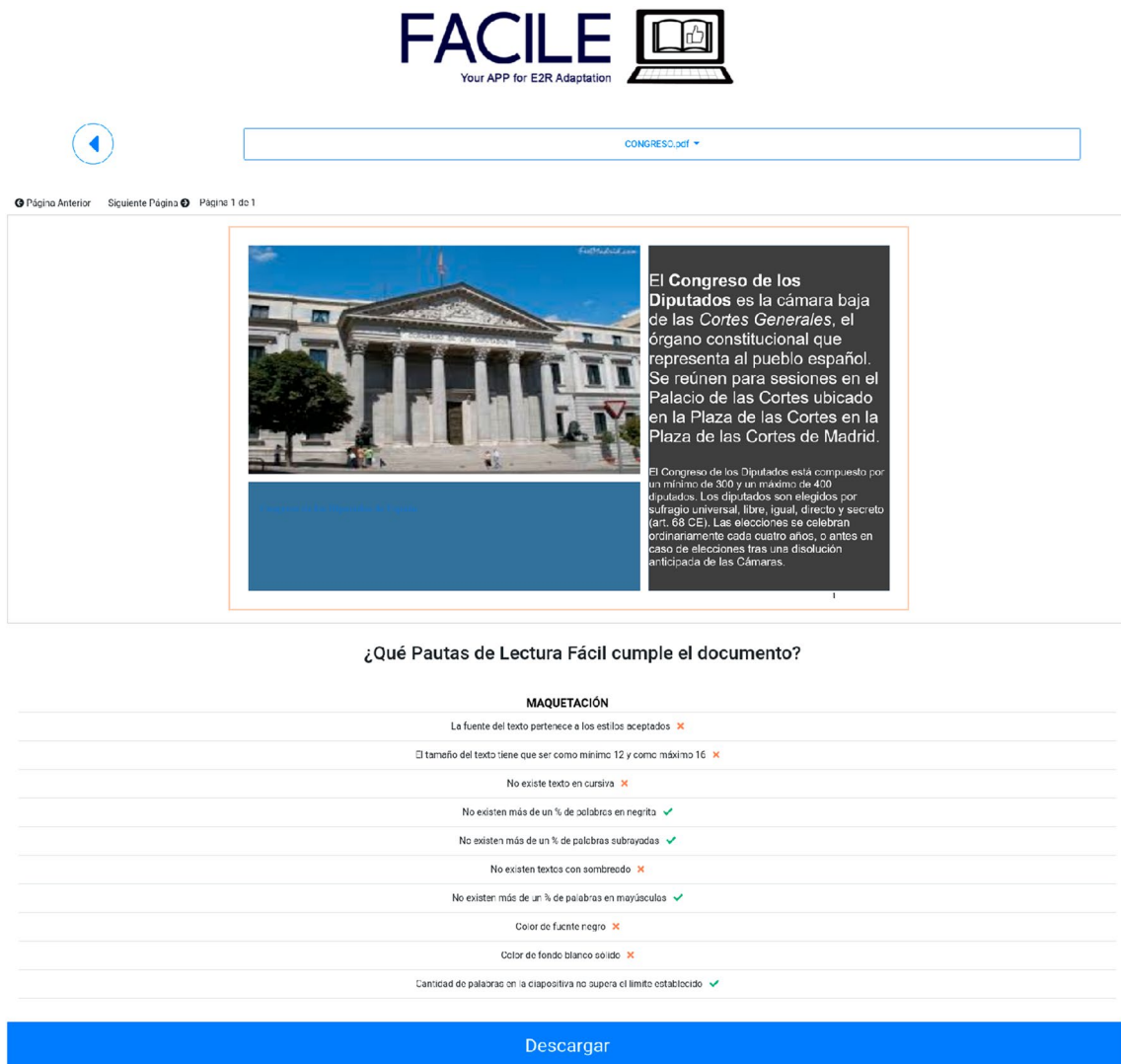


Fig. 8 Screenshot of the FACILE mockup that presents the output of the E2R analysis of a particular input

for identifying which guidelines are not fulfilled into a set of services with the same functionality.

Currently, we have developed an interactive mockup of FACILE with a user interface in Spanish. Figure 7 shows how an E2R expert can provide a specific document as input for FACILE, while Fig. 8 presents how the output of the E2R analysis performed by FACILE is shown to the final user. This figure shows (a) the input provided by the E2R expert, as a kind of reminder; in this case the input was a slide in which the definition and composition of the Congress of Deputies in Spain are presented and (b) the output presented by FACILE regarding the E2R assessment of the input; this output is produced as a report in which satisfied and not satisfied E2R guidelines are shown.

Regarding FACILE functionalities, we are currently (a) re-engineering the Easy-to-Read Advisor code in order to develop individual services to be integrated in FACILE and (b) developing services associated to E2R guidelines according to the survey results presented in Sect. 3.4. Based on these results, we decided to prioritize the following guidelines in the services development process:

- Regarding **Design and Layout Guidelines**, both methods for document assessment and suggestion of transformations are planned to be developed for DL6, DL5, DL7, DL3, DL10, DL8, and DL4, based on the ranking presented in Fig. 4. The development of an E2R assessment method for DL1 and an E2R suggestion method for DL9 and DL2 is also planned, according to the rankings shown in Figs. 2 and 3. Indeed, we currently have an early prototype for the service associated to DL6;
- With respect to **Writing Guidelines**, both the assessment and the suggestion of transformations methods are planned to be developed for W14, W19, W22, and W21, based on the ranking presented in Fig. 4; while E2R assessment methods for guidelines W4, and W2, and E2R suggestion methods for W20, W7, W17, W10, W13, and W9 are also planned to be developed, according to the rankings shown in Figs. 3 and 4. Indeed, we currently have early prototypes for services associated to W10, W13, W17, and W21 as well as a working prototype for W9 service.

Our next steps in FACILE development process are as follows: (1) to validate the mockup by a group of E2R experts. The aim of this evaluation is to decide the most appropriate user interface for the target group of FACILE as well as the most convenient interactions; (2) to analyse data gathered from the user-based evaluation; (3) to develop the changes and improvements suggested by the evaluators; (4) to perform the integration between FACILE UI and the functional E2R services; and (5) to perform different evaluation activities over the FACILE application.

5 Conclusions and future work

The demand for experts in the E2R Methodology who carry out the manual adaptation of documents is currently increasing in an exponential way. In this scenario, it would be very useful and beneficial to have a support application that performs a (semi)-automatic process of adapting documents to E2R guidelines. We envision that with this kind of technological support, in general: (a) E2R experts will have more capacity to meet the growing demand for adaptations of materials to the E2R methodology; (b) E2R experts will focus on complex issues and decisions, transferring the identification and transformation of simpler situations to the support application, thus, there will be a reduction in the effort dedicated by the experts to a particular adaptation work; and (c) the E2R adapted documents will be more rigorous and of higher quality. With this situation in mind, we have as a main research goal to develop a support application, called FACILE, based on Artificial Intelligence (AI) techniques, to support the work of E2R experts. In order to follow a human-in-the-loop approach in our AI development, we consider of key importance the involvement of E2R experts from the very beginning of the project. In the first phase of our AI development we gathered needs and informal requirements from E2R experts via questionnaires. The collected data help us to understand what kind of support application the experts consider more useful and what E2R guidelines should be prioritized in the development. After this collecting phase, we designed the FACILE architecture based on microservices and developed an interactive mockup of FACILE in Spanish based on needs, informal requirements and general comments and suggestions obtained from E2R experts. Currently, we are performing the re-engineering of the Easy-to-Read Advisor code and the development of services associated to E2R guidelines according to the survey results. In addition, we have planned a series of user studies with E2R experts to evaluate both FACILE UI and services. Taking into account the results of these user validations, we will update the mockup in order to implement the web application, which will be available for all.

In addition, further work, includes investigating the best way to present the different suggestions of transformation in the documents to the E2R experts. Moreover, we plan to provide FACILE's user interface, functionalities and E2R services not only in Spanish, but also in other natural languages such as English. Finally, we have also the intention to directly help people with reading comprehension difficulties by providing them with an easy-to-use tool that obtains easy-to-read versions of complicated texts.

Acknowledgements This research work has been partially funded by the project Datos 4.0 (TIN2016-78011-C4-4-R) (Agencia Estatal de Investigación del MINECO y Fondos FEDER). We would like to thank Iván Martínez for his help on the architecture selection and description; Arminda Moreno for her advice in the analysis of the survey data; Alvaro González for his work on FACILE; and last but not least, Olga Berrios from Plena Inclusión España and all the E2R experts who answered our questionnaire and provide us with informal requirements for FACILE.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alva-Manchego, F., Scarton, C., Specia, L.: Data-driven sentence simplification: survey and benchmark. *Comput. Linguist.* **46**(1), 135–187 (2020). https://doi.org/10.1162/coli_a_00370
- Bengoetxea, K., Gonzalez-Dios, I., Aguirregoitia A.: LagunTest: a NLP based application to enhance reading comprehension. In: READI. (2020)
- Blumesberger, S.: MaaB, Christiane: Leichte Sprache. Das Regelbuch. Berlin: LIT Verlag 2015 (Barrierefreie Kommunikation. Herausgegeben von der Forschungsstelle Leichte Sprache, Universität Hildesheim, Band 1), ISBN 978-64312907-9, 184 S. Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, 69(1), 155-158 (2016). <https://doi.org/10.31263/voebm.v69i1.1408>
- Bott, S., et al.: Can Spanish Be Simpler? Lex- SiS: Lexical Simplification for Spanish. In: Proceedings of COLING 2012. pp. 357–374. (2012)
- Bulté, B., Sevens, L., Vandeghinste, V.: Automating lexical simplification in Dutch. *Comput. Linguist. Neth. J.* **8**, 24–48 (2018)
- Evans, R., Orasan, C.: Identifying signs of syntactic complexity for rule-based sentence simplification. *Natl. Lang. Eng.* **25**(1), 69–119 (2019)
- Ferrés, D., et al.: YATS: yet another text simplifier. In: Natural language processing and information systems, pp. 335–342. Springer, Cham (2016)
- Muñoz, Ó. García: Lectura fácil: Métodos de redacción y evaluación. Real Patronato sobre Discapacidad, (2012)
- Inclusion Europe. Information for All. European standards for making information easy to read and understand. German. OCLC: 838005460. Brüssel: Inclusion Europe, 2009. ISBN: 9782874601101. <https://short.upm.es/yuh77>
- Inclusión, Plena: Universidad Autónoma de Madrid, and Fundación ONCE. Estudio sobre la comprensión lectora. Una escalera para comprender el mundo. Ed. by Plena Inclusión. (2019)
- Koptient, A., Grabar, N.: Fine-grained text simplification in French: steps towards a better grammaticality. In: ISHIMR Proceedings of the 18th International Symposium on Health Information Management Research. Kalmar, Sweden, (2020)
- Lee J., Zhao W., Xie W.: A customizable editor for text simplification. In: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations, pp. 93-97. (2016)
- Martin, L., et al.: MUSS: multilingual unsupervised sentence simplification by mining paraphrases. In: [arXiv:2005.00352](https://arxiv.org/abs/2005.00352) [cs.CL] (2021)
- Matausch, K., Nietzio, A.: Easy-to-read and plain language: defining criteria and refining rules. In: Easy-to-Read on the Web. Online Symposium (2012)
- Moreno, L., Alarcón, R., Martínez, P.: EASIER system. Language resources for cognitive accessibility. In: The 22nd International ACM SIGACCESS Conference on Computers and Accessibility. ASSETS '20. Virtual Event, Greece: Association for Computing Machinery, pp. 1–3. (2020)
- Namiot, D., Sneps-Snepe, M.: On micro-services architecture. *Int. J. Open Inf. Technol.* (2014)
- Nietzio, A., Scheer, B., Bühler, C.: How long is a short sentence? A linguistic approach to definition and validation of rules for easy-to-read material. In: Miesenberger, K. (ed.) Computers helping people with special needs, pp. 369–376. Springer, Heidelberg (2012)
- Nomura, M., et al.: Guidelines for easy-to-read materials. English. OCLC: 643236274. The Hague: IFLA Headquarters, (2010)
- Asociación Española de Normalización y Certificación. Lectura Fácil. Pautas y recomendaciones para la elaboración de documentos (UNE 153101: 2018 EX). (2018)
- Oelke, D., et al.: Visual readability analysis: how to make your writings easier to read. *IEEE Trans. Vis. Comput. Graph* **18**(5), 662–674 (2012). <https://doi.org/10.1109/TVCG.2011.266>
- National Archives Office of the Federal Register and Records Administration. Public Law 111-274-Plain Writing Act of 2010. Ed. by U.S. Government Printing Office. (2010). <https://www.govinfo.gov/app/details/PLAW-111publ274>
- Rello, L., Baeza-Yates, R., Saggion, H.: DysWexia: Textos más accesibles para personas con dislexia. *Proces. Leng. Natl.* **51**, 205–208 (2013)
- Rennes E., Jönsson, A.: A tool for automatic simplification of Swedish texts. In: Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015). Vilnius, Lithuania: Linköping University Electronic Press, Sweden, pp. 317–320. (2015)
- Saggion, H., et al.: Text simplification in simplext. Making text more accessible. *Proces. Leng. Natl.* **47**, 341–342 (2011)
- Scarton, C., et al.: SIMPLIFICA: a tool for authoring simplified texts in Brazilian Portuguese guided by readability assessments. In: NAACL. (2010)
- Suárez-Figueroa, M.C., et al.: Towards the assessment of easy-to-read guidelines using artificial intelligence techniques. In: Miesenberger, K. (ed.) Computers helping people with special needs, pp. 74–82. Springer, Cham (2020)
- Suter, J., Ebling S., Volk M.: Rule-based automatic text simplification for German. In 13th Conference on Natural Language Processing (KONVENS 2016), Bochum, Germany, 19 September 2016–21 (2016). <https://doi.org/10.5167/uzh-128601>
- Widyastuti, R., Yusliani, N., Fachrurrozi, M.: Simplification complex sentences in Indonesia language using rule-based reasoning. (2014)
- Wilkens, R., Oberle, B., Todirascu, A.: Coreference-based text simplification. Proceedings of the 1st Workshop on Tools and Resources to Empower People with READING Difficulties (READI). pp. 93–100. (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.