READABILITY METRICS FOR WEB APPLICATIONS ACCESSIBILITY

Miriam Martínez, José R. Hilera and Luis Fernández-Sanz
Department of Computer Science, University of Alcalá, Alcalá de Henares, Madrid, Spain

Keywords: Readability, Understandability, Accessibility, WAI, Software metrics.

Abstract: In this work, an analysis of applicability of specific metrics to evaluation of understandability of web content expressed as text, one of the key characteristics of accessibility according to WAI, is presented. Results of application of metrics to check level of understanding of pages in English of different universities are discussed.

1 INTRODUCTION

The measurement in the field of Web Information Systems is a relatively young discipline and there is not any general consensus on the exact definition of concepts. Obviously, software measurement has a long tradition (Fenton and Pfleeger, 1997), but the problem is adaptation of the different concepts and metrics to the specific context of web applications. In general, experts in the field of documentation agree that good structure and presentation of the content (with clear idea of objectives and good adaptation to intended audience) is a key element for the quality of documents. Of course, it is also important to select the most appropriate representation format for the document or each part of it (diagrams, text, and multimedia) as well a good use of technical prose which facilitates easy understanding of the content (Edwards, 1992; Bell et al., 1994; Lehner, 1993).

Web engineering standards address quality issues and include content accessibility as a specific requirement for well engineered web sites (ISO, 2002). Moreover, certain regulations promote and enforce fulfilment of accessibility criteria for specific websites (e.g., from the beginning of 2008, all websites maintained by Public Administration in Spain must comply accessibility requirements according to international recommendations). Both standards and practitioners have adopted WAI (Web Accessibility Initiative) guidelines (W3C, 2008) as main reference for accessible design.

Within WCAG (Web Content Accessibility Guidelines) of WAI, one of the four main principles is that information and the operation of user interface must be understandable. In order to achieve this, guideline 3.1 establishes that we must make text content readable and understandable. Sadly, existing recommendations included in WCAG, although interesting and clear, are still far from being formal enough for automated evaluation. WCAG advisory techniques include avoiding unusual foreign words, limiting text column width, etc.

This work is aimed at analysing the application of existing readability metrics for documents to the web application readability problem. Section 2 reviews existing text readability metrics while Section 3 describes application of metrics to three university websites with text in English. Section 4 present results and discussion and Section 5 outlines some conclusions.

2 READABILITY METRICS

Readability of a text indicates the extent to which its content is easy to understand. In general theory of documentation, several measurable factors are identified as predictors of text readability: sentence length, word length, words specialization, number of propositions, number of monosyllables, etc. Using them, it is possible to determine, in general terms, the minimum training level required to understand the text (García, 2001).

As stated above, readability is essential for websites and applications, especially when dealing

Martínez M., R. Hilera J. and Fernández-Sanz L. (2010).
READABILITY METRICS FOR WEB APPLICATIONS ACCESSIBILITY.
In *Proceedings of the 12th International Conference on Enterprise Information Systems - Human-Computer Interaction*, pages 207-210
DOI: 10.5220/0002867902070210

with documents intended for public dissemination. For example, measurement of readability of text is used for medical texts both for patient's consent documents and for educational brochures for general public. Readability metrics are also used for the evaluation of quality of writing style in educational materials (when they are still draft documents) for primary and secondary schools (López, 1982).

Different authors have contributed to readability evaluation with indexes of text readability. In general, they tend to express complexity of reading (and subsequently of writing) as formulae which are easy to calculate. Flesh was the pioneer with his index for evaluating English-language newspapers. He presented a formula expressing the readability level in terms of average word number per sentence and average syllables per word (Flesch, 1948). The original scale interpretation was established as follows: 100 points means "easy to read" text, 65 points represents a text adequate for an average U.S. citizen and 0 points implies a document which is extremely difficult to understand.

Kincaid et al. (1981) adapted Flesch index to the educational level required to read and understand the text. This is really interesting for the evaluation of WCAG guidelines requirements (W3C, 2008) because they refer to secondary education level as upper threshold required by users to understand contents.

Gunning (1968) proposed another index in his book about techniques of clear writing in English language. It uses the words average per sentence and the number of words known as "hard words" /the ones which are not used daily by people) as parameters for calculating the readability factor. The result is the minimum formal education level required to easily read the text. Specific adaptations to different languages have appeared. In the case of the Spanish language, Spaulding (1951) presented the first metric. Fernández-Huerta adapted the Flesch formula to the Spanish language and López-Rodríguez contributed with a series of readability metrics (Fernández-Huerta, 1959).

There are two Flesh-Kincaid indexes: the "reading easiness" and "educational level" (Kincaid et al., 1981). The first is basically a formula to measure if a text is easy or difficult to read depending on the number of syllables, words and sentences. The basic premise is that more readable texts contains generally less complex sentences and, subsequently, less words on average and less over-elaborated words, with less syllables on average.

In general most of existing readability metrics are based on determining the amount of significant lexical and syntactic elements which appear in the text (syllables, words, sentences, etc.) and combining these values with some coefficients obtained empirically. As a summary, the Table 1 shows the exact calculation formulae for the metrics used in this work.

Table 1: Readability metrics used in this work.

Author/year	Expression
Flesch (1948)	$206.85 - 0.846 \cdot \bar{n}_s - 1.105 \cdot \bar{n}_p$
Farr *et al.* (1951)	$1.599 \cdot p_1 - 1.015 \cdot \bar{n}_p - 31.517$
Gunning (1968)	$0.4 \cdot (\bar{n}_p + l)$
Smith and Kincaid (1970)	$\bar{n}_p + 9 \cdot \bar{n}_l$
Kincaid *et al.* (1981)	$0.39 \cdot \frac{n_p}{n_f} + 11.8 \cdot \frac{n_s}{n_p} - 15.59$

The meaning of the symbols, which appear in these formulas, is the following:

$\bar{n}_s$: Average word length (average number of syllables per word);

$\bar{n}_p$: Average sentence length (average number of syllables per word);

p_1: Percentage of words in the text with only one syllable;

l: Percentage of long words in the text (words with three or more syllables);

n_p: Number of words in the text;

n_f: Number of sentences in the text;

n_s: Number of syllables in the text;

$\bar{n}_l$: Average words length (average number of letters per word);

n_l: Number of letters in the text;

n_{pd}: Number of different words in the text.

These metrics are intended to evaluate the content complexity of a text: in the three first indexes, the higher value calculated is, the easier the text is understood. Analogously, low values in the first two metrics and large values in the last three suggest the text is difficult to understand. In most cases, the authors of these indexes recommend applying the corresponding calculation not to the full text but to texts chunks between 100 and 200 words.

3 APPLICATION OF METRICS TO THREE UNIVERSITY WEBSITES

As case studies of the application of readability metrics, we choose university websites pages in English from three different countries. Three universities websites has been chosen to calculate the 5 readability metrics to evaluate results. Chosen Universities were:

- University of Alcalá (Spain) http://www.uah.es/idiomas/ingles/Little_history.shtm
- University of Coimbra (Portugal) http://www.uc.pt/en/informacaopara/visit/hist
- University of Oxford (England) http://www.ox.ac.uk/visitors_friends/visiting_the_university/

4 RESULTS

Tables 2 to 6 show values corresponding to each of the three universities.

Table 2: Flesch Index.

	UNIVERSITY OF ALCALA	UNIVERSITY OF OXFORD	UNIVERSITY OF COIMBRA
$\overline{n}_s$	63	24.44	22.44
$\overline{n}_p$	32.8	14.25	12.31
Index	*117.31*	*170.43*	*174.26*

Table 3: Farr et al. Index.

	UNIVERSITY OF ALCALA	UNIVERSITY OF OXFORD	UNIVERSITY OF COIMBRA
p_1	48.17	51.75	46.19
$\overline{n}_p$	32.8	14.25	12.31
Index	*12.21*	*36.77*	*29.84*

As the Table 2 shows, the Flesh index values for University of Oxford is very similar to the ones of Coimbra while they are significantly higher than the ones for University of Alcalá. This indicates the first two ones are easier to understand than the later.

Table 4: Gunning Index.

	UNIVERSITY OF ALCALA	UNIVERSITY OF OXFORD	UNIVERSITY OF COIMBRA
$\overline{n}_p$	32.8	14.25	12.31
l	93	39	57
Index	*50.32*	*21.30*	*27.73*

Table 5: Smith and Kincaid Index.

	UNIVERSITY OF ALCALA	UNIVERSITY OF OXFORD	UNIVERSITY OF COIMBRA
$\overline{n}_p$	32.8	14.25	12.31
$\overline{n}_l$	5.07	5.01	5.06
Index	*78.43*	*59.34*	*57.85*

Table 6: Kincaid et al. Index.

	UNIVERSITY OF ALCALA	UNIVERSITY OF OXFORD	UNIVERSITY OF COIMBRA
Nf	10	16	16
Ns	630	391	359
Np	328	228	197
Index	*23.87*	*14.20*	*14.72*

In Table 3, the Farr index for the University of Oxford and Coimbra, although somewhat different, are higher than the index for University of Alcalá.

Table 4 shows the Gunning index for the University of Oxford and for Coimbra that, while different, are lower than the index for University of Alcalá. In Table 5, the Smith and Kincaid index for the University of Oxford and Coimbra, really similar, are lower than the index for the University of Alcalá, a fact which confirms the trend of results from the previous tables.

Finally, as Table 6 shows, Kincaid index for University of Oxford and for Coimbra is very similar and smaller than 20, which indicates a high readability text; however, index value (higher than 20) for University of Alcalá is very high so it is assumed that text is difficult to understand. So, according to the results obtained, the English text from University of Alcalá is more difficult to understand than the ones for University of Oxford and for University of Coimbra, especially because it includes a greater number of words.

5 CONCLUSIONS

This work has shown why text readability is an important factor for quality of web applications and websites today although it is a factor frequently forgotten in the daily professional work in this area. Its essential role within the content accessibility aspect of web pages is highlighted by the fact that accessibility is currently an important requirement due to both de facto guidelines like WCAG (W3C, 2008) and some governmental regulations, at least

for public Administration websites. As usually happens when Web application development teams are uniquely covered by IT staff with a software engineering profile, some of these aspects tend to be forgotten or poorly managed. So it is important that web development teams get a multidisciplinary flavour including experts in content management, edition and quality as well as a good group of graphic design profiles. This is one of the principles of the so-called Web Engineering discipline (Deshpande and Hansen, 2001). As an ongoing line of action, we are analyzing a broader sample of academic web pages involving multinational teams of students (in order to check perceived quality of this audience) as well as experts in content edition from editorial industry (in order to check possible complementary methods to evaluate readability).

REFERENCES

Fenton, N. E., Pfleeger, S. L., 1997. *Software metrics: a practical and rigorous approach*, Chapman & Hall. London.

Edwards, P. A., 1992. Elements of quality user documentation. In *Information management: strategy systems and technologies*, Auerbach Publications.

ISO, 2002. *ISO/IEC 23026-IEEE Std 2001-2002, ISO/IEC Software Engineering — Recommended Practice for the Internet — Web Site Engineering, Web Site Management, and Web Site Life Cycle*, Internartional Standards Organization, Geneve.

W3C, 2008. *Web Content Accessibility Guidelines (WCAG) 2.0*, World Wide Web Consortium. http://www.w3.org/TR/2008/REC-WCAG20-20081211/.

García J. A., 2001. Legibilidad de los folletos informativos. *Pharm Care Esp*, 3, 49-56.

López, N., 1982. *Cómo valorar textos escolares*. Cincel. Madrid.

Flesch, R. E., 1948. *The Art of Readable Writing*. Harper & Brothers. New York.

Kincaid, J. P., Aagard, J. A., O'Hara, J. W. y Cottrell, L. K., 1981. Computer readability editing system, *IEEE transactions professional communication*, 24(1), 38-41.

Gunning, R., 1968. *The Technique of Clear Writing*. McGraw-Hill. New York.

Spaulding, S., 1951. Two Formulas for Estimating the Reading Difficulty of Spanish. *Educational Research Bulletin*, 30, 117-124.

Fernández-Huerta, 1959. Medidas Sencillas de Lecturabilidad, *Consigna*, 214, 29-32.

Deshpande, Y., Hansen, S., 2001. Web Engineering: Creating a Discipline among Disciplines, *IEEE Multimedia*, 8(2), 82-87.