

## Social Media and self-medication for weight loss: insights from Facebook in Brazil

Jaciara Carneiro\*, Denise Tsunoda\*\*, Rebeca Giacomitti\*\*\*

\* (Post-Graduation Program in Information Science, Management and Technology, Federal University of Parana, Brazil

Email: jaciara@ufpr.br)

\*\* (Post-Graduation Program in Information Science, Management and Technology, Federal University of Parana, Brazil

Email: dtsunoda@ufpr.br)

\*\*\* (Post-Graduation Program in Information Science, Management and Technology, Federal University of Parana, Brazil

Email: rebeca.suckow@gmail.com)

### ABSTRACT

Recognized as a public health issue, self-medication is employed for pain ceasing and disease prevention, or even for aesthetic purposes, although it can compromise a person's health. Part of the information that leads to self-medication is disseminated via the Internet, in online discussion groups like the ones of Facebook, the major social media platform in the world and in Brazil. This paperwork aims at analyzing if, within such groups, there are incentives to self-medication for weight loss, and how such information could be classified for better understanding. The study is a descriptive survey, with a non-experimental approach. For data treatment and analysis is employed the Descending Hierarchical Classification (DHC or Reinert's Method), using IRaMuTeQ software. Results encompass the identification of classes of weight loss-related themes, which can contribute to self-medication with said intent. It is possible to apply methods of this study to other social media sites and other platforms of interaction throughout the Internet.

**Keywords:** Descending Hierarchical Classification, Internet, IraMuTeQ, Reinert's Method, Social Media.

### I. INTRODUCTION

Health is a condition of complete physical, mental and social well-being, and not merely the absence of disease or illness [1]. Therefore, changes in natural aspects of one's body, such as weight and measurements, can bring to said person improvements in well-being and overall satisfaction with himself, a quest in which the Internet has become an undeniable source of information.

Specifically in Brazil, the country where this study is conducted, the keywords "medicine for weight loss" ("remédio para emagrecer", in the nation's official language, Portuguese) form the second most sought-after term in the Health category of the search engine Google, in the years 2013 to 2015. It stands behind only the amount of searches with the words "pain medication" ("remédio para dor", in Portuguese) [2,3,4]. However, this search for information, sometimes unrelenting, leads many people to adopt strategies that are not always related to health promotion. In such context, the problem of self-medication emerges.

#### 1.1 Self-medication in Brazil

Self-medication can be defined as the consumption of a medicine in order to treat or alleviate perceived illness symptoms [5]. Similarly, the reuse of prescriptions given for previous cases of infirmity can also characterize self-medication, considering that the medicine indicated for a past health issue may not be appropriate for a current one [6]. According to the World Health Organization (WHO), the Rational Use of Medicines (RUM) occurs only when patients receive medication appropriate to their clinical needs, in doses and periods of administration suitable to their individuality [7]. Such medicines can have prophylactic (preventive), curative, palliative and diagnostic purposes [8].

Self-medication is considered a public health issue by the Brazilian government Community of Health Agents (ACS, for its acronym in Portuguese). Although, in the country, it is commonly practiced by elders, self-medication is also registered among youngsters [9]. Medication misuse can bring consequences such as allergic reactions, addiction or, in more severe cases, can lead the individual to death [10].

Nonetheless, self-medication is a usual practice among Brazilians. In the year of 2014, a survey by the privately held Institute of Science, Technology and Quality (ICTQ, for its acronym in Portuguese) showed that 76.4% of the Brazilian population had already reported taking drugs obtained not from a medical doctor's prescription, but from recommendations of a family member or friend [11]. Despite such procedure is driven by an attempt to save time and money, taking medicines by means of a "self-diagnosis" or the recommendation of non-specialists, without the prescription and dosage by a healthcare professional, can mask the symptoms of serious diseases and complicate the picture of an individual's health. In this respect, the Brazilian National System of Toxic-pharmacological Information (Sinitox, for its acronym in Portuguese) reports 42,128 cases of human poisoning in its country in 2013, of which 11,985 (28.45%) are due to the use of medicines. Of these, 328 (2.73%) are caused by self-medication and drug abuse [12].

### **1.1.1 Self-medication and online peer recommendation in Brazil**

In what concerns to non-specialized recommendations of medicines, not only the ones given by people physically close and/or familiar to the user collaborate for self-medication. Part of the responsibility for the dissemination of health information may be attributed to the Internet and its remarkable reach in Brazil. The percentage of households with access to the World Wide Web reached 31,200,000 in 2013, concentrated in the country's Southeast, South and Midwest regions, in this order [13].

In conjunction with the dissemination of information, connections among internet users and the exchange of health information through social media platforms continually increase. According to a 2015 report released by the international consultancy agency GlobalWebIndex, 47% of Brazilians are registered in at least one of said platforms, while 29% of the world population is [14]. Still according to the same report, in comparison with other online social networks, Facebook leads the ranking of social media users, involving 25% of them.

In light of the current expansion of access to digital content by the Brazilian population – and, as a consequence, of access to information on health, specifically about medication that can contribute to weight loss –, this study has the purposes of analyzing if there are peer-to-peer incentives for self-medication aimed at weight loss, throughout discussion groups on the online social network Facebook that promote conversations about the matter, as well as of exploring ways in which

such user generated content could be classified, in order to better understand it and, thus, provide subsidies for further studies on self-medication influenced by online interactions/recommendations.

To accomplish such objectives, this paper proposes, as a method of analysis, the Descending Hierarchical Classification (DHC), also known as Reinert's method, employing the free software IramuTeQ as a tool.

## **II. INTERNET USE AND DISSEMINATION OF INFORMATION ON HEALTH IN BRAZIL**

The Internet has revolutionized the realms of computing and communications. Widely used for information dissemination, the World Wide Web is a means of collaboration among individuals in which physical distances do not constitute a barrier to interaction. In addition, the Internet counts on a widespread infrastructure for information transmission and has been reaching societies as a whole, as individuals move towards the adoption of online tools for a diversity of purposes [15].

In Brazil, Internet operations began in the early 1990's [16]. Since then, possibilities of access have proven promising. As an evidence, 57.6% of the Brazilian population has been reached by the World Wide Web as of 2013. This broad coverage is due to factors such as increasing investments in technological infrastructure and the escalating purchase of mobile phones by Brazilians, which registered a 131%-growth between 2008 and 2013 [17,18].

An expanded access to technology tools leads to an expanded creation of communication channels. Social media sites are inserted in this context, for they facilitate the sharing of information and knowledge as part of the relations among members of a given community [19]. Such interaction occurs by the formation of a group of individuals who come together to share experiences or ideas of common interest [20].

Amid the rise of social media platforms that, as early as 2011, accounted for 1.1 billion users globally [21] – and their role as facilitators for information exchange, whether for personal, professional or commercial purposes, the online social network Facebook emerged. Created in 2004 by a group of Harvard University students led by Mark Zuckerberg, Facebook is deemed the largest and most accessible social media platform worldwide [22]. This platform offers, for free, the possibility of creating online discussion groups aimed at providing spaces for dialogue in which network users can talk about common interests, such as "family reunion, your after-work sports team, your book club" [23]. The privacy of information published in said groups can be configured for a

"public" or a "restricted" mode – in the former, any Internet user can view publications; in the latter, posts can be seen only by group members.

However, such free online tool presents a fragility in terms of legitimacy of the information shared, given that all the users of a group can post comments about any topic of their interest [24]. Such fragility proves especially inconvenient when health subjects are addressed.

It is by this technological scenario that an infirm yearning to learn more about diseases and their symptoms, medication, costs of treatment and other aspects, transits [25]. He faces the challenge of selecting the more coherent and legitimate information in an ocean of data made available throughout the Internet – what constitutes a complex process that requires caution [26]. Nonetheless, despite all those issues, an infirm can trust health information made public online and, by taking self-medication decisions under a false sense of knowledge on a particular subject, can worsen his overall health condition.

### III. RESEARCH METHODOLOGY

This study presents characteristics of a descriptive survey with a non-experimental approach [27], since the research involves collection of data generated by a sample of the universe of public Brazilian discussion groups on Facebook aimed at information-sharing about weight loss. Data analysis is predominantly quantitative, employing the technique of content analysis, characterized by the fragmentation of texts in order to identify themes and patterns [28].

In order to find Facebook groups whose discussions could form the corpus of analysis, eight search terms related to weight loss, in Portuguese, are employed. All of them have the Portuguese radical "*emagre*" – an equivalent of "to lose weight":

1. "*emagrecer*" (the complete form of the verb correspondent to "to lose weight", in Portuguese);

2. "*emagrecimento*" (an equivalent of "weight loss");
3. "*emagrecendo*" ("losing weight");
4. "*emagreça*" (imperative mood of "*emagrecer*" or "to lose weight");
5. "*perder peso*" (a synonym for "*emagrecer*" or "to lose weight");
6. "*perda de peso*" (a synonym for "*emagrecimento*", also an equivalent of "weight loss");
7. "*perca peso*" (a synonym for "*emagreça*" or the imperative mood of "to lose weight");
8. "*perca de peso*" (a commonly mistaken spelling of "*perda de peso*" or "weight loss").

The discussion groups must also possess the following characteristics:

- a) to have "health" as their major theme, with focus on "weight loss";
- b) to have Brazilian users as their owners/administrators;
- c) to be active by the time of data collection up to one year back – it is, with publications from 30 August, 2015 to 30 September 2016, as well as a maximum interval of one month between a publication and other;
- d) to be of public mode/access – which means that group users have implicitly agreed to make the information they share in the group available for visualization by external users. In this way, researchers can be sure that the data collection process is performed with no major risks to data privacy.

After applying the 12 aforementioned filters, 21 public groups are obtained, broken down by search terms, as seen on Table 1. In addition, for the composition of the corpus of analysis, a sample of 150 comments is arbitrarily set. In order to ensure that all of the 21 groups are represented in such sample, a fraction of comments to be extracted from each group is defined, taking into account each one's size (number of members). This is also shown on Table 1.

**Table 1: Classification of Comments Extracted by Search Term**  
 (white rows show the original terms in Portuguese, gray rows show their equivalent in English)

Search Terms	Nr. of public groups	Fraction of the total of groups	Nr. of members	Fraction of the total of members	Nr. of extracted comments
Terms with the radical "EMAGRE": emagreCER, emagreCIMENTO, emagreCENDO, emagreCA.	10	0.476	100,549	0.847	127
Respectively: "to lose weight"/"weight loss"/"losing weight"/"lose weight" (imperative)					
"PERDER PESO"	2	0.095	12,396	0.104	16
"to lose weight"					
"PERDA DE PESO"	3	0.143	2,620	0.022	3
"weight loss"					
"PERCA PESO"	1	0.048	1,609	0.014	2
"lose weight" (imperative)					
"PERCA DE PESO"	5	0.238	1,606	0.013	2
Misspelled form of "perda de peso" ("weight loss")					
TOTAL	21	1	118,780	1	150
<b>Total of extracted comments</b>					<b>150</b>

In order to process a database containing the extracted corpus of text, it is necessary to first standardize the ways in which a same term can be written (i.e., when an equivalent to "ppl" appears in the database, it must be changed to "people"), and to correct characters and words that can be corrupted during the extraction of comments. Therefore, noises are minimized for the analysis stage and the knowledge discovery process is made easier. For the following step of data mining and processing the software IRaMuTeQ (acronym, in French, to R Interface for Multidimensional Analysis of Texts and Questionnaires) is used. Developed in 1983 by Max Reinert, a scientist of France's National Center for Scientific Research (CNRS, for its acronym in French), IraMuTeQ is based on the R statistical language and operates as a subroutine of the statistical treatment software of the same name. This tool was chosen for this research paper because it employs a computerized technique of discourse analysis that aims at composing classes of words which represent the different discourse types that refer to the subject of interest [29]. Therefore, IraMuTeQ allows for various types of textual data analysis that range from basic lexicographic analysis to multivariate analyses, being one of them the Descending Hierarchical Classification (DHC) [30] that guides this study.

Through the technique of DHC, blocks of text are classified according to specific vocabularies (or lexicons), by which matrices are formed from the crossing between threads of texts and words [31]. From then on, as a first result, are obtained classes of text segments that feature: vocabularies similar to each other; vocabularies different from the ones of the text segments grouped in other classes [30]. Based on these matrices, the software organizes the data analysis, by means of a representation (dendrogram or

cladogram) that shows the relationships among the various classes, and can be interpreted as: content; as dynamism of the discursive movement; and as a representation of the context [32].

According to Camargo e Justo (2013) [30], IRaMuTeQ performs calculations and achieves results that enable the description of each of the classes previously found, which are characterized by their specific lexicon and by variables that appear spelled with an asterisk in the software interface. Another way of presenting the results is made possible by Factorial Analysis of Correspondence (FAC), which is also performed via DHC. In this case, IRaMuTeQ, based on previously defined classes, calculates and points out the most characteristic text segments of each class (highlighted in different colors), thus allowing the contextualization of the lexicon of each one of such classes. Still on the processing of a corpus of text by IRaMuTeQ, the parameters used for the examination of textual Context Units (CUs) are set out by the software itself (by default). It is worth adding that, for the stemming of the corpus of analysis, words characterized as adverbs and verbs are not considered as active, what leads the software to focus on the analysis of nouns and adjectives.

#### IV. RESULTS

As aforementioned, the corpus of analysis of this study consists of 150 comments, extracted from Facebook public groups about weight loss, which amount to 6,038 words. This corpus of analysis was processed by the software IRaMuTeQ, by means of the DHC technique. As a result, five classes of information shared by said groups are identified, as seen on the figures below. Fig. 1(A) shows the original results in Portuguese; Fig. 1(B), the results translated to English.

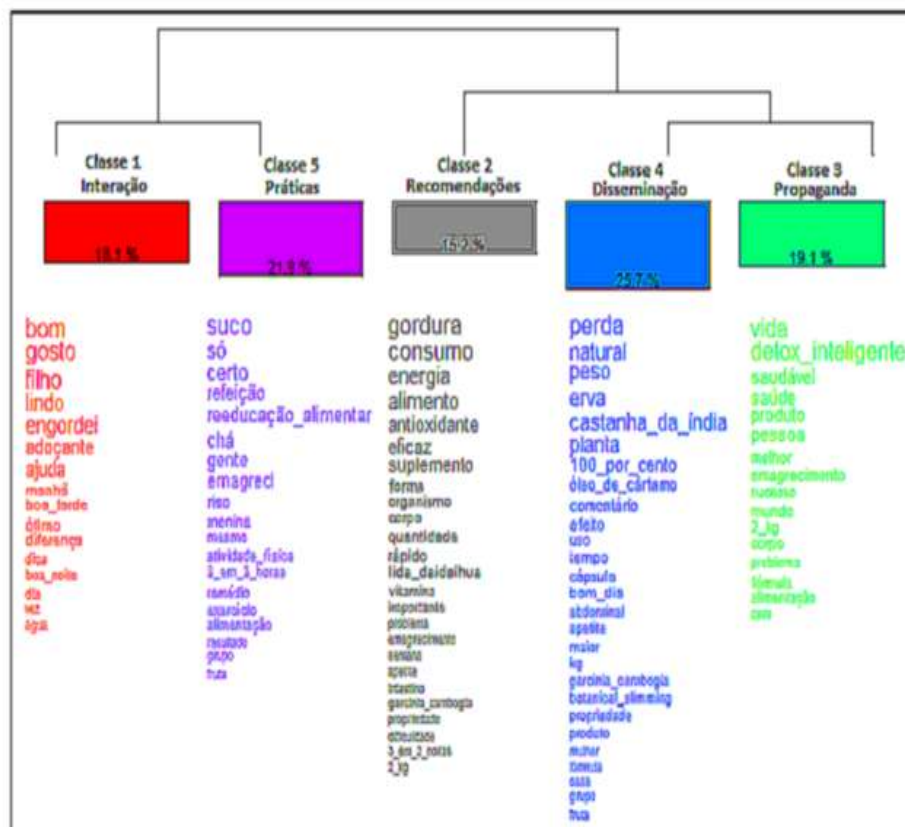


Figure 1(A): Cladogram of descending hierarchical classification via software IRaMuTeQ (in Portuguese)

It is possible to observe that the majority of data (25.7%) is located in Class 4 (blue) – named by the authors of this study as *Dissemination Class*. Here, is shown the exchange of information on indications/contraindications of substances that may help the weight loss process – what can lead to self-medication by members of Facebook public groups that deal with the subject. In the *Dissemination Class*, can be found sentences like this: “I have never seen so many comments that are negative or related to contraindications about a homemade medicine for weight loss like the ones on horse chestnut”.

Second in the amount of data is Class 5 (fuchsia), name as *Practices Class*. In this one, are registered 21.9% of the occurrences of the corpus of text. It is characterized by weight loss “tools”, such as “juice”, “meal”, “nutritional education”, “physical activity” and “exercise”. Sentences like: “the right thing to do is to have [dietary] shakes only if you don't have a way to eat a true meal. They do work, but it will depend a lot on how you

are eating”, illustrate the class' main thematic focus.

Class 3 (green), labeled as *Propaganda Class*, represents the third largest group of data, aggregating 19.1% of occurrences. It groups verbs that denote product advertisement, in phrases such as “(...) it [product 'X'] is zero cholesterol, does not contain gluten, lactose or fructose. Choose your favourite flavour and try the new formula with collagen”. Content like this can also act as a trigger for self-medication among members of Facebook public discussion groups about weight loss.

With 18.1% of the corpus of text's occurrences comes Class 1 (red), named *Interaction Class*. It has as its main theme the social interactions among weight loss groups' members. Greetings like “good afternoon” and adjectives such as “gorgeous” and “great”, represent the relation present in this class, as shown in the example: “can you help me to become so very pretty just like you?”

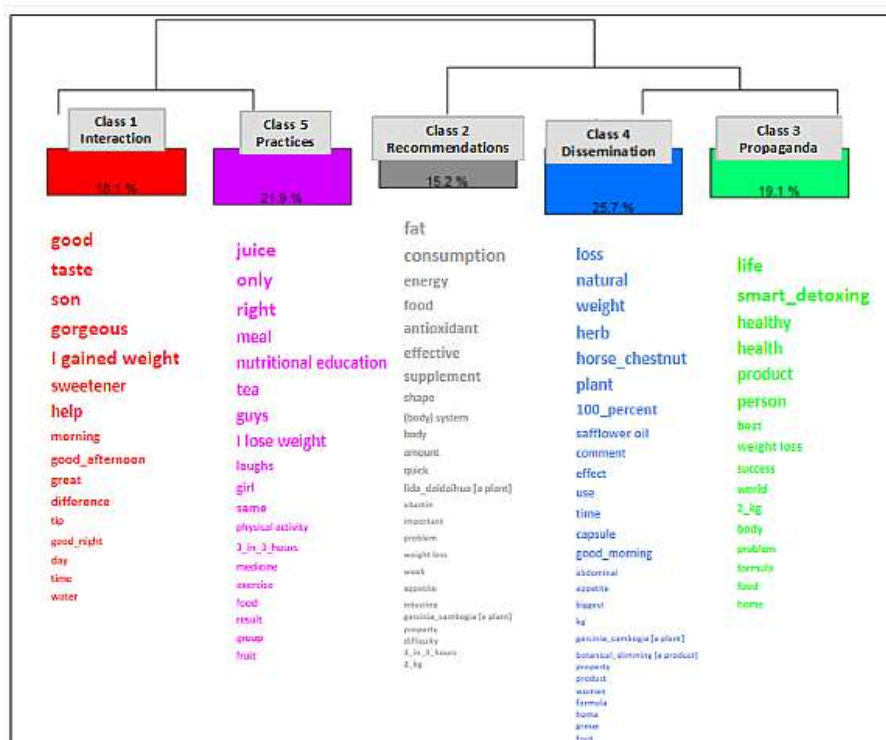


Figure 1(B): Cladogram of descending hierarchical classification via software IRaMuTeQ (English version)

Lastly, containing 15.2% of the data, appears Class 2 (grey), or *Recommendations Class*. It brings together advice and “recipes” of methods for the weight loss process. Such a class is characterized, therefore, by counseling like “don’t take this supplement together with other medicines. Carefully read all instructions before use”. To facilitate the elucidation of the contexts exposed in classes, the IRaMuTeQ software provides a representation by Cartesian plane, which results from the technique of Factorial Analysis of Correspondence (FAC). Thus, IraMuTeQ performs a crossing between the corpus of text’s vocabulary (taking into account the frequency of occurrence of its terms) and its classes, what enables the visualization of oppositions and affinities among such classes [31]. In Fig. 2, relationships among the five classes detected in this study are represented. This graphic representation is provided only in

Portuguese, given the majority of terms shown in such FAC graph correspond to the ones already translated in Fig. 1(B).

In the graph of Factorial Analysis of Correspondence shown in Fig. 2, it is possible to see with greater clarity, in Quadrant 1 (Q1), the exchange of information about compounds regarded as “slimming products” – such as horse chestnut and other plant, *garcinia cambogia*, besides safflower oil –, that are characteristic of the *Dissemination Class* (in blue – which, not incidentally, occupies practically the whole Q1, being the class that concentrates just over 1/4 of the occurrences of the corpus of text). It turns out that such information is strongly related to the terms grouped in the *Propaganda Class* (in green). Thereby, Q1 is the quadrant that best represents the role of Facebook public groups about weight loss as facilitators of self-medication.

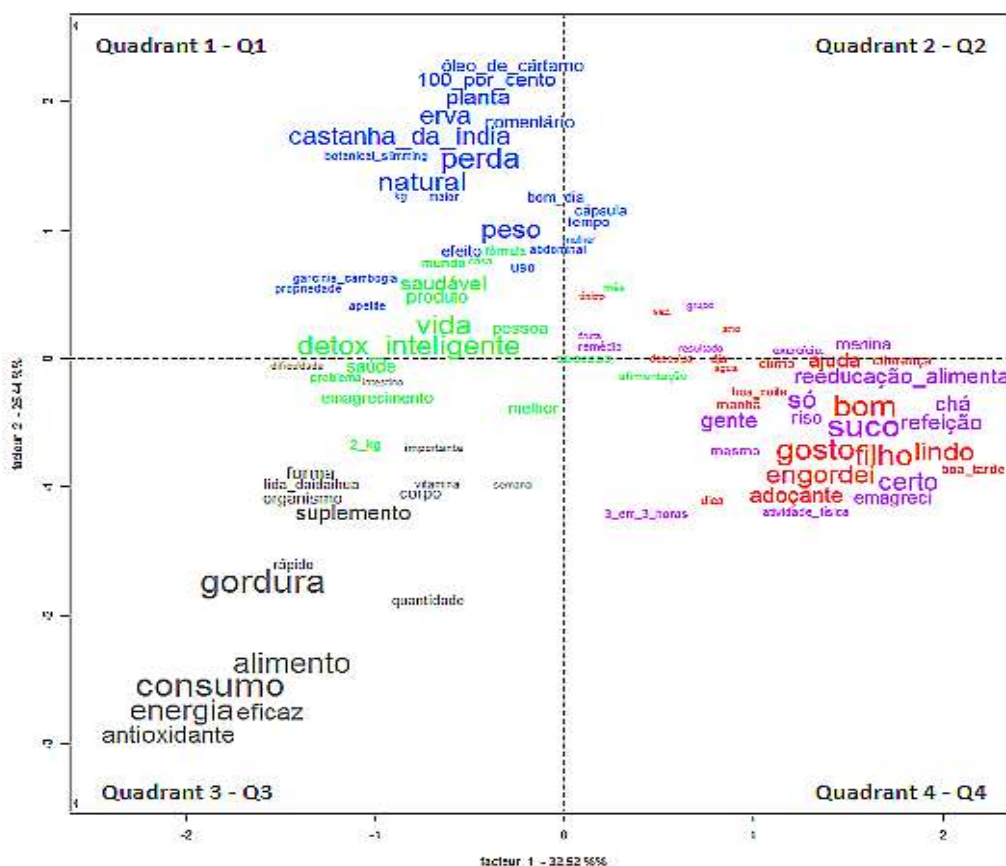


Figure 2: FAC - Cartesian plane of analysis of correspondence among classes

It can be considered yet more robust the relationship between the discourses found in the classes of *Interaction* (in red) and *Practices* (in fuchsia), both concentrated and “mixed” in Quadrant 4 (Q4). Thus, it turns out that, within what is expected of the dynamics of online discussion groups (like the ones on Facebook), the exchange of information on practices for weight loss emerges, primarily, from interactions among members of the groups in this study.

Quadrant 3 (Q3), in turn, is dominated by terms from the *Recommendations Class* (in grey), which presents the greatest dispersion of the chart. This denotes that advices and recipes for weight loss do not have a strong relationship with information found in the other classes. In addition, Quadrant 2 (Q2) is the one that has the lowest concentration of information, containing almost exclusively themes that come from 4 of the 5 classes.

Examples of words shown in Q2 are “fruit” (“*fruta*”, in Portuguese), “food” (“*alimentação*”), “healthy” (“*saudável*”), “weight loss” (“*emagrecimento*”), “success” (“*sucesso*”) and “life” (“*vida*”) – besides “product” (“*produto*”) and “medicine” (“*remédio*”), all terms

that can be regarded as relevant in the “universe” of self-medication. On the other hand, words positioned the furthest from the center point of the Cartesian plane are more linked to the context of their own class and, thus, do not appear in the others.

In order to facilitate the understanding of the context of the groups in this study, it is possible to extract, via IRaMuTeQ, a word cloud representing the frequency of occurrence of the corpus of text's terms, as shown in the figures below. Fig. 3(A) presents the original word cloud in Portuguese; Fig. 3(B) shows the most frequent (biggest in size) terms translated to English. .

It is possible to identify that the words “weight” (“*peso*”, in Portuguese) and “fat” (“*gordura*”) are the most frequent in the construction of the research corpus – each features 30 occurrences, according to the statistics provided by IRaMuTeQ. Shortly thereafter, in order of frequency, appear the words “product” (“*produto*”), “person” (“*pessoa*”) and “day” (“*dia*”), with 25, 24 and 19 occurrences, respectively. Such frequencies can be explained by the fact that the Facebook groups analyzed in this study are focused on weight loss as a theme.



Figure 3(A): Word cloud of comments from Facebook public groups about weight loss in Brazil (original)

However, next to these words, in frequency of occurrences, appear the terms “natural” (“*natural*”) and “medicine” (“*remédio*”), with 16 cases each, and “capsule” (“*cápsula*”), with 15. Since these are words related to the “universe” of medication, it is possible to infer that, in the corpus of text, not only there are classes of terms related to self-medication, but also such terms have a significant representativity in relation to the whole. In addition, names of products such as “supplement” (“*suplemento*”) and “horse chestnut” (“*castanha-da-india*”) indicate the promotion of these and similar elements for weight loss.

## V. DISCUSSION

Within Facebook public groups about weight loss in Brazil, incentives to the use of products and to self-medication for weight loss purposes happen, and are manifested above all in textual classes 3, of *Propaganda* (shown in green in Figs. 1 to 2), and 4, of *Dissemination* (shown in blue in Figs. 1 to 2). There is a consistent exchange

of information, among users of the groups in study, about recommendations, use, indications, contraindications and posologies, of natural and synthetic products. It can also be inferred that stories of personal experiences, concentrated on Class 1, of *Interaction* (shown in red in Figs. 1 to 2), are of minor concern, being the focus of interest on knowledge about weight loss in itself, transmitted from a user to another.

Unlike the dissemination of knowledge by experience among groups' members, comments of commercial character do not seem to have significant repercussions. Therefore, posts by weight loss products' advertisers exert little or no effect on discussions and rapidly “vanish”, amidst the aspiration of groups' members for information on weight loss routines and procedures that brought satisfactory results for the others. It can be assumed that information deemed as relevant comes from group peers, to the detriment of those published by members whose business interest in the discussions overlaps personal goals of weight loss.





exclusion of its affixes or suffixes [34]. Stemming makes information retrieval easier both in searches throughout the Internet as a whole, as in libraries and documents available online and offline. The way in which some of the Facebook public groups that provide the corpus of text for this study are selected, from the search for the Portuguese stem or radical “*emagre*” (an equivalent of “to lose weight”), via the social media platform’s own search engine, can serve as an example of the use of stemming, for future studies.

Furthermore, the influence of information discussed and shared through social media platforms, about behaviors, attitudes and individual decisions, seems to offer so many possibilities of research as the technological innovations that impact the human communication. In this respect, it is assumed that a natural deployment of this research paper would be the inclusion, in the corpus of text in study, of data collected on Facebook pages about weight loss. Another possibility is to apply the same Descending Hierarchical Classification adopted here to content collected on Instagram, a social media platform that has established itself as a place of convergence of an audience interested in health, well-being, fitness and weight loss, counting on well-known users from a variety of areas – athletes, physical education professionals and media celebrities – who stand as references in such topics.

## REFERENCES

- [1] WHO, Preamble to the Constitution of the World Health Organization *as adopted by the International Health Conference*, New York, 1946, 19-22
- [2] Google. (2016, October 04). *Google Trends*, 2016. Retrieved from [http://www.google.com/trends/explore?cat=45&date=2013-01-01%202013-12-31&geo=BR&q="remédiao%20para"](http://www.google.com/trends/explore?cat=45&date=2013-01-01%202013-12-31&geo=BR&q=)
- [3] \_\_\_\_\_ (2016, October 04). *Google Trends*, 2016. Retrieved from [http://www.google.com/trends/explore?cat=45&date=2014-01-01%202014-12-31&geo=BR&q="remédiao%20para"](http://www.google.com/trends/explore?cat=45&date=2014-01-01%202014-12-31&geo=BR&q=)
- [4] \_\_\_\_\_ (2016, October 04). *Google Trends*, 2016. Retrieved from [http://www.google.com/trends/explore?cat=45&date=2015-01-01%202015-12-31&geo=BR&q="remédiao%20para"](http://www.google.com/trends/explore?cat=45&date=2015-01-01%202015-12-31&geo=BR&q=)
- [5] Revista da Associação Médica Brasileira. Automedicação (Self-medication), *Revista da Associação Médica Brasileira*, 47(4), 2001, 269-270.
- [6] L. G. Paulo, A. C Zanini. Automedicação no Brasil. (Self-medication in Brazil), *Revista da Associação Médica Brasileira*, 34, 1988, 69-75.
- [7] I. Franceschet-de-Sousa et al., Uso racional de medicamentos: relato de experiência no ensino médico da Unesc, Criciúma/SC (Rational use of medicines: case studies in the medical education of Unesc, Criciúma, State of Santa Catarina, Brazil). *Revista Brasileira de Educação Médica*, 34(3), 2010, 438-445.
- [8] ANVISA (Brazilian National Health Surveillance Agency). *Caderno de educação e promoção da saúde no contexto escolar: o contributo da Agência Nacional da Vigilância Sanitária para o uso racional de medicamentos (Textbook of education and health promotion in the school context: the contribution of the Brazilian National Health Surveillance Agency for the rational use of medicines)*. Brasília, Brazil: ANVISA, 2007.
- [9] G.P.Kauling et al., Utilização de medicamentos: limites e possibilidades das orientações dos agentes comunitários de saúde às famílias (Use of medicines: limits and possibilities of the guidelines given by community health agents to families). *Mundo da Saúde*, 37(1), 2013, 44-45.
- [10] Brazilian Society of Endocrinology and Metabolism (2016), Os perigos da automedicação (The dangers of self-medication). Retrieved from: <http://www.endocrino.org.br/os-perigos-da-automedicacao/>
- [11] M. König (2015, April 05), Medicamentos intoxicam 27 mil pessoas por ano no Brasil (Medicines intoxicate 27 thousand people per year in Brazil). Retrieved from <http://www.gazetadopovo.com.br/vida-e-cidadania/medicamentos-intoxicam-27-mil-pessoas-por-ano-no-brasil-0gqwphwirhjk2u6x6hgb9qay4>
- [12] Brazilian Ministry of Health, Fiocruz, Sinitox (2016), Dados de intoxicação – casos registrados de intoxicação humana por agente tóxico e circunstância (Intoxication data – recorded cases of human poisoning by toxic agent and circumstance). Retrieved from <http://sinitox.icict.fiocruz.br/dados-de-agentes-toxicos>
- [13] Brazilian Institute of Geography and Statistics – IBGE (2015), Acesso à internet e à televisão e posse de telefone móvel celular para uso pessoal (Internet and TV access and possession of mobile phones for personal use). Retrieved from <http://biblioteca.ibge.gov.br/visualizacao/livros/liv99054.pdf>
- [14] R. Caixeiro (2016, June 09), Quais são as redes sociais mais acessadas do Brasil? (Which are the most accessed online social networks in Brazil?). Retrieved from <http://www.edialog.com.br/listas/quais-sao-redes-sociais-mais-acessadas-brasil>

- [15] B. M. Leiner et al., Brief history of the Internet, Website of the Internet Society, 1997. Retrieved from [http://www.internetsociety.org/sites/default/files/Brief\\_History\\_of\\_the\\_Internet.pdf](http://www.internetsociety.org/sites/default/files/Brief_History_of_the_Internet.pdf)
- [16] M. S. R. M. Carvalho, A trajetória da internet no Brasil: do surgimento das redes de computadores à instituição dos mecanismos de governança (The trajectory of Internet in Brazil: from the rising of computer networks to the institutionalization of governance mechanisms), master's diss., UFRJ (Federal University of Rio de Janeiro), Rio de Janeiro, Brazil, 2006.
- [17] International Telecommunication Union – ITU) and United Nations Educational, Scientific and Cultural Organization – UNESCO, The state of broadband 2015: broadband as a foundation for sustainable development (Geneva: ITU, UNESCO, 2015).
- [18] Presidency of the Republic of Brazil (2015, April 29), Número de brasileiros com celular cresceu mais de 131% em oito anos (Number of Brazilians with cell phones grew more than 131% in eight years). Retrieved from <http://www.brasil.gov.br/economia-e-emprego/2015/04/brasileiros-passaram-a-ter-mais-aparelhos-celulares-entre-2005-e-2013>
- [19] M.I.Tomaél, A.R. Alcará, I. G. Di Chiara, Das redes sociais à inovação (From social networks to innovation). *Ciência da Informação*, 4(2), 2005, 93-104. Retrieved from <https://dx.doi.org/10.1590/S0100-19652005000200010>
- [20] R. M. M. Marteleto, Análise de redes sociais: aplicação nos estudos de transferência da informação (Social network analysis: application in studies on information transfer). *Ciência da Informação*, 30(1), 2001, 71-81.
- [21] ComScore, Inc. (2011, September 20) A ascensão das redes sociais na América Latina: como a mídia social está definindo o cenário digital da América Latina (The rise of social networks in Latin America: how social media is defining Latin America's digital scenario). Retrieved from [http://www.comscore.com/por/Insights/Apresentacoes-e-documentos/2011/The-Rise-of-Social-Networking-in-Latin-America?cs\\_edgescape\\_cc=BR](http://www.comscore.com/por/Insights/Apresentacoes-e-documentos/2011/The-Rise-of-Social-Networking-in-Latin-America?cs_edgescape_cc=BR)
- [22] S. Phillips (2007, July 25), A brief history of Facebook. Retrieved from <https://www.theguardian.com/technology/2007/jul/25/media.newmedia>
- [23] Facebook (2016) About: groups. Retrieved from [https://www.facebook.com/help/1629740080681586?helpref=hc\\_global\\_nav](https://www.facebook.com/help/1629740080681586?helpref=hc_global_nav)
- [24] \_\_\_\_\_.(2016) What is public information? Retrieved from <https://www.facebook.com/help/20380546632373>
- [25] H. B. R. Garbin, A. F. Pereira Neto, M. C. R. Guilam, The internet, expert patients and medical practice: an analysis of the literature. *Interface – Comunic., Saúde, Educ.*, 12(26), 2008, 579-588.
- [26] D. Giustini, How Google is changing medicine, *British Medical Journal*, 331(24-3), 2005, 1487-1488.
- [27] A. C. Gil, Como elaborar projetos de pesquisa (How to elaborate research projects) (São Paulo, Brazil: Atlas, 2002).
- [28] L.Bardin, Análise de conteúdo (Content analysis), (Lisbon, Portugal: Edições 70, 2009).
- [29] N. Kronberger and W. Wagner, Palavras-chave em contexto: análise estatística de textos (Keywords in context: statistical analysis of text features). In: M.W. Bauer and G. Gaskell (Orgs.), *Pesquisa qualitativa com texto, imagem e som: um manual prático (Qualitative researching with text, image and sound: a practical handbook for social research (Translated to Portuguese by P. A. Guareschi) (Petropolis, Brazil: Vozes, 2003) 416-444.*
- [30] B. V. Camargo; A. M. Justo, IRaMuTeQ: um software gratuito para análise de dados textuais (IRaMuTeQ: a free software for the analysis of textual data). *Temas em Psicologia*, 21(2), 2013, 513-518.
- [31] A. R. A. Nascimento and P. R. M. Menandro, Análise lexical e análise de conteúdo: uma proposta de utilização conjugada (Lexical analysis and content analysis: a proposal for conjugate utilization), *Estudos e Pesquisas em Psicologia*, 6(2), 2006, 72-88. Retrieved from: <http://www.revispsi.uerj.br/v6n2/artigos/pdf/v6n2a07.pdf>
- [32] M. Reinert, Alceste, une méthode statistique et sémiotique d'analyse de discours; application aux "Rêveries du promeneur solitaire" (Alceste, a statistical and semiotic method of discourse analysis; application to "Reveries of the solitary walker"). *La Revue Française de Psychiatrie et de Psychologie Médicale*, 5(39), 2001, 32-36.
- [33] J. F. R. D. Souza, C. L. C. Marinho and M. C. Guilam, Consumo de medicamentos e internet: análise de uma comunidade virtual (Medicine consumption and the Internet: analysis of a virtual community). *Revista da Associação Médica Brasileira*, 54(3), 2008, 225-231.
- [34] R.S. Minani. (2014). Pré-processamento de documentos (Preprocessing of documents) [PowerPoint slides]. Retrieved from [http://www.facom.ufu.br/~miani/site/teaching\\_files/ori/T%C3%B3pico%20de%20Pr%C3%A9-processamento%20de%20documentos.pdf](http://www.facom.ufu.br/~miani/site/teaching_files/ori/T%C3%B3pico%20de%20Pr%C3%A9-processamento%20de%20documentos.pdf)