VoTec terminographic environment over the years: brief overview

Guilherme Fromm[1*] e Joel Victor Reis Lisboa[2]

[1]Universidade Federal de Uberlândia, Avenida João Naves de Ávila, 2121, 38408-144, Uberlândia, Minas Gerais, Brasil. [2]Programa de Pós-Graduação em Estudos Linguísticos, Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil. *Author for correspondence. E-mail: guifromm@ufu.br

ABSTRACT. VoTec (Vocabulário Técnico Online) is a web environment for terminographic management created by Fromm (2007) with the aim of training undergraduate and graduate students for terminographic and translation practices. Since 2007, it has been used in several research and projects aimed at researchers training and monolingual and bilingual reference works production. This paper aims to present an overview of the projects and research carried out through VoTec between 2007 and 2023, as well as of the updates made on its database and query page. To do so, we describe its initial version (Fromm, 2007), the Terminology in Fiction (Fromm, 2011a) and the Linguistics Vocabulary (Fromm & Yamamoto, 2013) projects, as well as the following terminographic works/proposals: Vocabulary of Historical Linguistics (Yamamoto, 2015), Vocabulary of Business and Events Tourism (Carvalho, 2015), Vocabulary of Theology (Cardoso, 2017), Bilingual Vocabulary of Linguistics (Yamamoto, 2020), Vocabulary of Digital Marketing (Mateus, 2022) and Vocabulary of Angiology and Vascular Surgery (Santos, 2023). This panorama highlights: (i) VoTec's productivity for research and projects in the scope of Terminology, Translation, monolingual, bilingual and multilingual Terminography and Pedagogical Terminography; (ii) its continuous updates allow VoTec to follow latest trends and technologies on online reference works, becoming increasingly robust and multimodal, as well as to approach the needs of current users of online reference works; (iii) due to the possibilities of customizing fields and functionalities, VoTec easily adapts to different types of research and is capable of generating terminographic products with different purposes and target audiences.

Keywords: VoTec; web environments for terminographic management; terminography.

O ambiente terminográfico VoTec ao longo dos anos: breve panorama

RESUMO. O VoTec (Vocabulário Técnico Online) é um ambiente web de gerenciamento terminográfico criado por Fromm (2007) com o objetivo inicial de treinar alunos de graduação e pós-graduação para a prática terminográfica e de tradução. Desde 2007, ele já foi utilizado em diversas pesquisas e projetos voltados à formação de pesquisadores e à produção de obras terminográficas monolíngues e bilíngues. Este artigo objetiva apresentar um panorama dos projetos e pesquisas realizados por meio do VoTec, bem como das atualizações feitas no banco de dados e na página de consulta entre os anos de 2007 e 2023. Para tanto, descrevemos o projeto inicial do VoTec (Fromm, 2007), os projetos Terminologia em Ficção (Fromm, 2011a) e Vocabulário de Linguística (Fromm & Yamamoto, 2013), bem como as seguintes obras/propostas terminográficas: Vocabulário de Linguística Histórica (Yamamoto, 2015), Vocabulário de Turismo de Negócios e Eventos (Carvalho, 2015), Vocabulário de Termos da Teologia (Cardoso, 2017), Vocabulário Bilíngue de Linguística (Yamamoto, 2020), Vocabulário do Marketing Digital (Mateus, 2022) e Vocabulário de Angiologia e Cirurgia Vascular (Santos, 2023). O panorama apresentado no artigo evidencia os seguintes pontos: (i) a produtividade do VoTec em pesquisas e projetos no escopo da Terminologia, Tradução, Terminografia monolíngue, bilíngue e multilíngue e Terminografia Pedagógica; (ii) as contínuas atualizações permitem que o VoTec siga novas tendências e tecnologias referentes a obras de referência on-line, tornando-se cada vez mais robusto e multimodal, bem como que aproxime-se das necessidades dos atuais usuários de obras de referência on-line; (iii) devido às possibilidades de customizar campos e funcionalidades, o VoTec se adapta facilmente a diferentes tipos de pesquisa e é capaz de gerar produtos terminográficos com finalidades e públicos-alvo distintos.

Palavras-chave: VoTec; ambientes web de gerenciamento terminográfico; terminografia.

Introduction

Web environments for terminographic management (WETM) are online platforms which bring together, in the same environment, a variety of tools and resources necessary for the management of terminographic data and the production of reference works (Oliveira, 2009; Bocorny, Villavicencio, Kilian, & Wilkens, 2010; Wilkens, Bocorny, Kilian, & Villavicencio, 2012). These web environments optimize and facilitate terminographic work due to the following advantages: (i) they accelerate and make the terminographic management process more efficient; (ii) as they are online, they favor cost reduction and remote group work; (iii) they (semi)automatize processes that, if done manually, are exhaustive and time-consuming; therefore, they increase productivity in terminographic work (Oliveira, 2009; Wilkens et al., 2012). There are many examples of WETM produced over the years, like Corpógrafo (Sarmento, Maia, & Santos, 2004), e-Termos (Oliveira, 2009), TermWeb (Interverbum Technology, 2023), TermWiki Pro (CSOFT International, 2023), WebTerm (STAR Group, 2023), etc., but in this paper we will be focusing on VoTec.

VoTec (***Vo****cabulário **Téc***nico Online*) is a WETM created by Fromm (2007) for his PhD dissertation. It was developed with the aim of: (i) increasing the contact of in-training translators with reference works (specifically dictionaries and vocabularies); (ii) promoting greater familiarity with environments and tools to aid terminographic and translation work; (iii) training undergraduate and graduate students to use and produce reference works; (iv) serving as an environment through which new reference works, especially terminographic ones, can be produced and made available.

It comprises a databank and a query page. The databank is restricted to administrators and researchers and was developed to cover many possible fields necessary for the creation of entries in reference works. The query page, with free online access, was developed to have a simple and user-friendly interface, as well as to be, as far as possible, self-explanatory, that is, so that it did not require prior and time-consuming training.

Since 2007, VoTec has been used in several monolingual and bilingual terminographic projects and research. Thus, this paper aims to present an overview of VoTec over the years (2007-2023) and of the projects and research carried out through this WETM.

VoTec over the years

This section presents the projects and research developed through VoTec between 2007 and 2023. Simultaneously, we describe the three major updates made in the databank and on the query page over this period. In short, we respectively present:

(i)   the original development of VoTec (Fromm, 2007) and its use in the following projects: Terminology in Fiction (Fromm, 2011a), Linguistics Vocabulary (Fromm & Yamamoto, 2013), Vocabulary of Historical Linguistics (Yamamoto, 2015) and Vocabulary of Business and Events Tourism (Carvalho, 2015);

(ii)   VoTec adaptations and updates for the construction of the Vocabulary of Theology (Cardoso, 2017);

(iii) VoTec adaptations and updates for the construction of the Bilingual Vocabulary of Linguistics (Yamamoto, 2020) and their use for the Vocabulary of Digital Marketing (Mateus, 2022);

(iv) VoTec adaptations and updates for the construction of the Vocabulary of Angiology and Vascular Surgery (Santos, 2023).

Original version

Fromm (2007) created VoTec[1] with the aim of increasing the familiarity and contact of Translation students with reference works, specifically dictionaries and terminographic vocabularies, given their usefulness and relevance for professional translators. He also aimed to expand the contact of students with computerized tools which aid the translation process. Furthermore, VoTec was developed not only to serve in-training translators, but also professional ones, since they could rely on different terminographic vocabularies hosted on VoTec in the future or even produce their own. Its main principle is that all work should be corpus-driven or corpus-based, that is, definitions should be entirely produced from corpus data, without researcher's intervention, or corpus data should serve to exemplify the definitions, which should be formulated based on the data but not solely confined to it.

VoTec comprises a databank and a query page. In the databank, the main researcher of the platform has access to the control page, where new projects can be created, as shown in Figure 1.

---

[1] The computational work was developed, in partnership with the researcher, by ICMC Junior (University of São Paulo/São Carlos).

Figure 1. VoTec's control panel.
Source: VoTec's databank.

Its databank follows a lexicographic approach, that is, it seeks to cover as many fields as possible for creating entries in reference works, except for diatopic, diaphasic, diastratic and diachronic variations, given its terminographic purposes. The databank fields work as a terminographic record and are filled only if the corpus provides the information required. They cover the following data:

1.   Linguistic contexts of the target-term extracted from the corpus (with source and retrieval date), in order to analyze its properties (grammar, collocations etc.);

2.   An initial concept, created by the researcher, for each context inserted, which will be used during semantic trace analysis;

3.   Data about the target-term: ontology, part of speech, number, gender, acronyms, term in full (if the entry is an acronym), morphosyntactic variants, meanings, and corpus information (wordlist position and number of occurrences of the term);

4.   Distribution of distinctive semantic traces, taken from the linguistic contexts, to aid the definition process[2];

5.   Semantic information: if the term is already dictionarized (its definition and the source can be inserted to help the researcher with possible meanings), hypernyms, hyponyms, co-hyponyms, synonyms and antonyms, and also a field for notes about researcher's semantic doubts;

6.   Equivalent Term[3]: in bilingual projects, every term is created twice in the databank, one version in Portuguese and another in English, and they are correlated in this field. It is important, though, to explain that the definitions are not exactly the same for both languages – they are created using the information from comparable (not parallel) corpora[4], which means that the definitions can be different. Therefore, two monolingual entries in contrast (one for each language) are presented in the query page;

---

[2] The semantic distinctive traces are taken from the contexts (with explanatory and definitory information, Aubert, 2001) selected for each term. We use the Componential Analysis model proposed by Ilari (2003) in a manual way, extracting the basic semes of each example and creating a definition by the genus proximun, differentia specifica model.

[3] The equivalence between languages, in most part of the platform projects, is conceptual. We seek the equivalence comparing the term candidates and their examples in both languages and verify if the conceptual grid is equivalent. The equivalence not always delivers a simple translation of a term from a language to another, as in 'links patrocinados' and 'PPC' (pay-per-click; digital marketing vocabulary), or, sometimes, we realize that a term is not translated from a language to the other one, as in mouse (computing vocabulary).

[4] According to Aijmer (2008, p. 276), "[…] parallel corpora consist of a source text and its translation into one or more languages. They can be further characterised in terms of the direction of the translation. […] A comparable corpus on the other hand does not contain translations but consists of texts from different languages which are similar or comparable with regard to a number of parameters such as text type, formality, subject-matter, time span, etc.".

7. Cross-references: this field links terms related to the one being described. These related terms must be already inserted in the databank so the system can link them;

8. Encyclopedic Information: information taken from free online sources, like Wikipedia, to supply the entry with an encyclopedic definition. The fields here are: definition, name of the article, source and link[5];

9. A final concept (including the information gathered from the semantic trace analysis) and a definition (which will be shown in the query page).

Figure 2 illustrates steps 1 and 2, while steps 3 to 9 can be partially seen in Figure 3.

Figure 2. Linguistic contexts and initial concepts for the term 'Processor'.
Source: VoTec's databank – Computing Vocabulary (Fromm, 2007)[6].

VoTec's query page (available at: http://treino.votec.ileel.ufu.br) can be configured according to some given possibilities:

1. 'Normal Exhibition': it mirrors the same microstructure formats found in printed dictionaries (in a simple paragraph), as seen in Figure 4. Note that the definitions and other fields in the microstructure are not the same in both languages, since the source corpora are comparable (not parallel) and the information they provide to be filled in the databank are not the same, generating different structure possibilities. What holds the terms in both languages together in the same screen is their shared concept.

2. 'Extended Exhibition': it displays the same content of the 'Normal Exhibition' but presents each microstructure field in a different line (Figure 5). In this case, all databank fields are presented, even those that have not been filled (in cases where the corpus did not provide enough information).

---

[5] Fromm (2007) tried to insert the possibility to include audio, video and pictures in this version of VoTec, but due to time and resource constraints, this was not possible at that time.
[6] During his doctoral research, Fromm (2007) produced two partial vocabularies to test VoTec: the Computing Vocabulary and the Linguistics Vocabulary.

Figure 3. Information fields for the term 'processor'.

Source: VoTec's databank – Computing Vocabulary (Fromm, 2007).

Figure 4. Normal exhibition for the entry 'computer/*computador*'.

Source: VoTec's query page – Computing Vocabulary (Fromm, 2007).

3. As we can see in Figures 4 and 5, 'Normal' and 'Extended exhibitions' are displayed with the 'Total search' activated, which means that all databank fields are shown in the query page.

4. Besides 'Total search', users can also choose the 'Translators' search' (Figure 6). It displays specific fields that better suit translating practices, according to research held by Fromm (2007). Users can also use 'Translators' Search' with 'Extended Exhibition'.
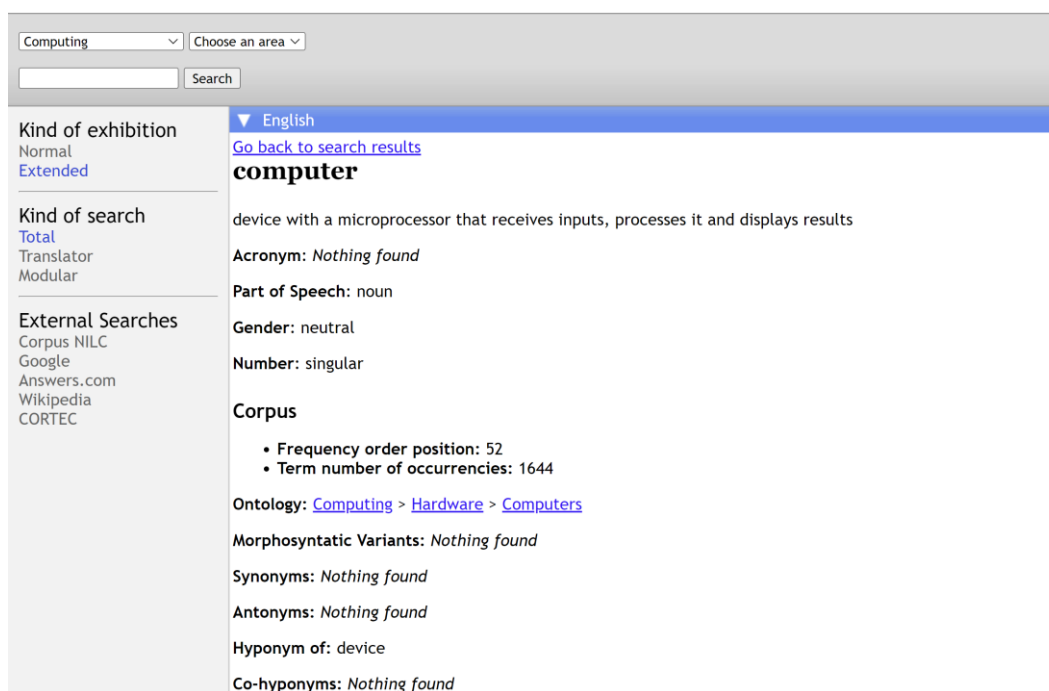
Figure 5. Extended exhibition for the term 'computer'.

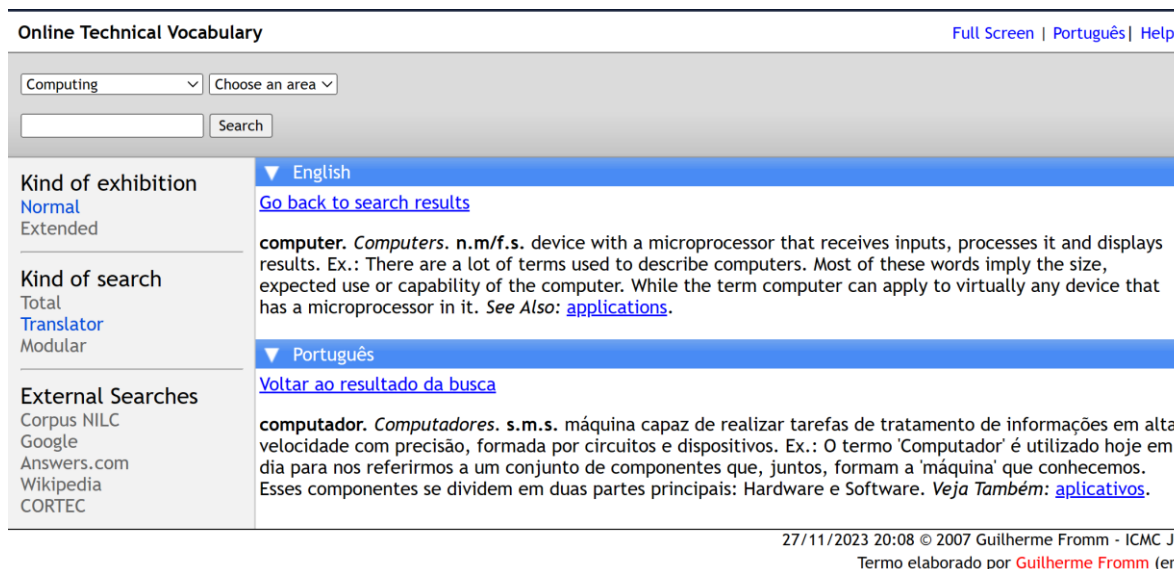Source: VoTec's query page – Computing Vocabulary (Fromm, 2007).

Figure 6. Normal exhibition + Translators' search for the entry 'computer/*computador*'.

Source: VoTec's query page – Computing Vocabulary (Fromm, 2007).

5. 'Modular search': it allows users to choose (among all available databank fields) how the microstructure is shown. As we can see in Figure 7, the part of speech, gender, number, definition, examples and cross-references fields were chosen (displayed in blue on the left). The idea here is that users can build the microstructure according to their search needs.

All these display possibilities and the external searches[7] (presented in Figures 4 to 6) were intended to give users more choices in how to use the query page.

Original version, new projects

After becoming a professor at the Federal University of Uberlândia (UFU), Fromm decided to continue working with the VoTec platform. Some projects, as follows, used the original programming.

---

[7] Not all of them are still functional, since some websites are no longer available.

Figure 7. Normal exhibition + Modular search for the entry 'computer/*computador*'.

Source: VoTec's query page – Computing Vocabulary (Fromm, 2007).

Terminology in Fiction project

The first version of VoTec was adapted for the Terminology in Fiction Project, created and coordinated by Fromm at UFU.[8] This project was first introduced by Fromm (2011a, 2011b). Through terminographic practice, it aims to train undergraduate research scholars and undergraduate and graduate students to work with Bilingual Terminography and language description. Students are introduced to the theoretical-methodological framework of Terminology, Terminography and Corpus Linguistics, and, at the end, build their own prototypes of bilingual terminographic vocabularies.

The aim is to bring knowledge produced in Linguistics closer to topics of interest to students. Due to the great popularity of fictional audiovisual productions for academic and non-academic audiences, they are the core of this project. In addition, these audiovisual productions rely on highly structured conceptual and terminological systems for the construction of fictional universes and are therefore of interest for terminological and terminographic studies.

Thus, this project works with subtitle corpora of television series in the English/Portuguese language pair. Under the guidance of the professor, students, in pairs, choose an audiovisual production, compile the corpora from online subtitle repositories and carry out the analysis using lexical analysis software. After that, they produce and make available their own vocabulary prototypes on VoTec. In this process, students acquire theoretical and technical knowledge about the entire process of researching and producing terminographic works and have their work hosted online.

---

[8] After 6 years with this project, Fromm inserted only one minor change to the VoTec system: the names of the students who developed each entry are shown at the bottom of the query page.

The prototypes present, in general, ten entries of idiosyncratic terms from each audiovisual production (5 in each language). In Figure 8 we can see an example of a bilingual entry from this project.

Figure 8. Entry for 'true alpha/*alfa genuíno*' from the Teen Wolf TV series.

Source: VoTec's query page – Terminology in Fiction Project.

There are 30 prototypes of bilingual terminographic vocabularies hosted on VoTec so far, covering television series and, more recently, reality shows. The results of this project can be accessed at: http://ic.votec.ileel.ufu.br/. The 12-year experience in the development of this project provided the theoretical and methodological foundations for what Fromm (2020) conceptualizes as Pedagogical Terminography.

Linguistics Vocabulary project

The Linguistics Vocabulary was the second project developed in this version of VoTec. According to Fromm (2013) and Fromm and Yamamoto (2013, 2021), it was a collaborative project which involved students of an undergraduate course and two graduate courses at UFU from 2010 to 2018.

This project aimed to formulate a corpus-based/corpus-driven taxonomy of Linguistics and to produce and make available bilingual entries (Portuguese/English) of terms from the area of Linguistics on VoTec. The entries produced in this project would be useful for teachers and students of Language and Translation courses, as well as for the general public interested in the area. Some of them are available at: http://treino.votec.ileel.ufu.br/. Figure 9 presents an example of the entries created in this project.

In pairs, in the case of undergraduate students, or individually, in the case of graduate students, the students who took part in the project carried out the following procedures: (i) compilation of bilingual (Portuguese/English) academic corpora for a specific Linguistics subarea; (ii) computational corpora analysis; (iii) selection of term candidates; (iv) production and availability of 10 entries (5 in each language) on VoTec. The corpora compiled in this project were partially used in the construction of VoBLing – Bilingual Vocabulary of Linguistics (Yamamoto, 2020), another terminographic product originated from VoTec. The adequacy and recycling procedures for these corpora reuse were detailed in Fromm and Yamamoto (2021).

In addition to terminographic training for undergraduate and graduate students and the benefit of having their work hosted for free online access on VoTec, which favors the popularization of their academic production, the students who participated in this project had the opportunity to get closer to the terminology of their own scientific field, especially with regard to students of the Languages course.

Vocabulary of Historical Linguistics

Yamamoto (2015) has also used this same version of VoTec to produce the Vocabulary of Historical Linguistics for his M.A. thesis. It is a bilingual (Portuguese/English) reference work based on comparable

academic corpora which presents contrastive monolingual definitions in both languages for terms from the area of Historical Linguistics. It was produced to meet the needs of translators, as well as Languages and Linguistics students, professors and researchers. It comprises 34 entries (17 for each language), which can be accessed at: http://treino.votec.ileel.ufu.br/. An example can be seen in Figure 10.

Figure 9. 'Morphology/*morfologia'* entry.

Source: VoTec's query page – Linguistics Vocabulary Project.

Figure 10. 'Clitic/*Clítico'* entry.

Source: VoTec's query page – Vocabulary of Historical Linguistics (Yamamoto, 2015).

The author justifies using VoTec due to the following reasons: (i) as it is online and free to access, VoTec makes reference works available worldwide; (ii) VoTec was built under the principle of being corpus-based or corpus-driven, which meets the methodology followed in the production of this vocabulary; (iii) as the definitions on VoTec are produced from the identification and systematization of semantic traces in corpora, it is possible to reduce, as far as possible, the researcher's subjectivity and strengthen objectivity; (iv) VoTec offers different ways of entries visualization, so it meets different user needs; (v) VoTec online databank makes terminographic data organization easier and faster.

Vocabulary of Business and Events Tourism (VocTur)

The Vocabulary of Business and Events Tourism (VocTur) is a vocabulary prototype developed by Carvalho (2015) for her M.A. thesis, in order to demonstrate the theoretical-methodological procedures that can be replicated for the construction of a more comprehensive reference work in the field of Tourism. It comprises 24 terms (12 in each language) and its target audiences are Translation and Tourism undergraduate students. It can be accessed at: http://treino.votec.ileel.ufu.br/.

VocTur is a terminographic product which is guided and based on comparable corpora of scientific papers, as well as of texts from websites of Business Tourism and Events companies and promotional and government materials related to this field. Like other bilingual vocabularies presented throughout this paper, since it was developed using VoTec, it also presents monolingual entries (Portuguese/English) in contrast, as can be seen in the example presented in Figure 11.

Figure 11. 'Incentive/*Incentivo'* entry.
Source: VoTec's query page – VocTur (Carvalho, 2015).

The author listed two main reasons for using VoTec in the construction of VocTur: (i) personalized querying possibilities, which can meet specific needs of different users; (ii) ease of access to the vocabulary, since it is hosted online.

This was the last reference work produced in the first version of VoTec.

First update: version 1.1

The researcher and his advisees realized, some years later, that the original version of VoTec was not exhaustive enough regarding the fields available in the databank. It was decided, then, that new Terminography PhD research would study and propose new fields/functionalities for both the databank and query page.

Vocabulary of theology (TermosTeo)

Developed by Cardoso (2017) for her PhD dissertation, the Vocabulary of Theology (TermosTeo) was the first project to propose some changes in the original version of VoTec. It is available at: http://teo.votec.ileel.ufu.br/.

In practice, TermosTeo comprises three monolingual (Portuguese) terminographic vocabularies: (i) Theological Vocabulary of *Faculdade Shalom de Ensino Superior* (Uberlândia), with 36 entries; (ii) Theological Vocabulary of *Faculdade Católica de Uberlândia*, with 33 entries and (iii) Contrastive Theological Vocabulary of the two previous ones, with 32 entries. This reference work allows users to analyze how the same term is differently conceptualized depending on the theological current in focus.

TermosTeo is based on corpora of theoretical and academic texts used by students and professors from the two aforementioned colleges. Therefore, Theology students and professors are its target audiences. Cardoso (2017) presents three main reasons for using VoTec in her research: (i) free access; (ii) different entries display according to users' needs; (iii) ease of making updates and changes, as it is online.

In the development of TermosTeo, the author proposed some changes both in the VoTec's databank and query page. These changes aimed to adapt VoTec for this specific project, given that it differs from other previously produced terminographic works. Some of the adjustments were proposed due to results obtained in pilot studies with students from the two focused colleges, which represented a sample of the target audience for which the work was intended. In summary, changes are subdivided into additions, deletions[9] and modifications.

The fields/resources added to VoTec were the following: (i) notes field for inserting encyclopedic definitions; (ii) word etymology field; (iii) audio with pronunciation. The following fields/resources were deleted from the updated version for TermosTeo: (i) acronym; (ii) term in full; (iii) number of meanings; (iv) hypernymy, hyponymy and co-hyponymy; (v) translators' search. Modifications were made in the following fields/resources: (i) equivalence term: it is now monolingual, that is, it links terms from different vocabularies in the same language; (ii) interface languages: only Portuguese, whereas the original version query page was available in Portuguese and English; (iii) external queries: reduced to Google and Wikipedia, whereas the original version featured five: NILC, Google, Answers.com, Wikipedia and CORTEC.[10]

Figure 12 presents one of TermosTeo's entries.

Figure 12. Entries for the term 'Concupiscência'.

Source: VoTec's query page – TermosTeo (Cardoso, 2017).

---

[9] We highlight that although these features were excluded from the TermosTeo project, they are still available in other versions of VoTec.

[10] The external query websites were reduced as they experienced periods of downtime or instability at certain moments over the years.

This first update demonstrated that the continuous development of the VoTec platform was possible, and Fromm's advisees could redesign it according to their own project needs.

Second update: version 1.5

Bilingual Vocabulary of Linguistics (VoBLing)

The second project to propose significant changes to VoTec's databank and query page was the Bilingual Vocabulary of Linguistics – VoBLing (Yamamoto, 2020). This is a bilingual (Portuguese/English) terminographic vocabulary prototype which presents monolingual entries in contrast to terms from the area of Linguistics.

This vocabulary was produced with the aim of detailing the theoretical framework and methodological procedures which will be replicated in future and increasingly comprehensive developments of VoBLing. Its target audiences are newly enrolled students in Languages courses, professional and in-training translators, as well as Linguistics professors and researchers.

VoBLing is based on comparable academic corpora in both languages. Part of these corpora was compiled along the Linguistics Vocabulary Project detailed earlier in this paper. At the moment, 56 entries (28 in each language) are available in this vocabulary, which can be accessed at: http://vobling.votec.ileel.ufu.br/. An example of one of the VoBLing's bilingual entries is shown in Figure 13.

Figure 13. 'Attestations/*Atestações'* entry.
Source: VoTec's query page – VoBLing (Yamamoto, 2020).

According to the author, among the reasons for choosing VoTec to produce this vocabulary are: (i) free online access; (ii) possibility of adding multimodal resources; (iii) pedagogical possibilities offered by VoTec; (iv) due to the distinctive features systematization and visualization layout, it is easier to produce corpus-based definitions.

As previously noted, VoBLing brought some changes to VoTec. The first was the modernization of the databank and query page interfaces. In addition, the following fields/resources were added: (i) total occurrences of phraseologies; (ii) lemma field, for inserting spelling variations; (iii) multimedia field, where it is possible to insert different types of files to make the entries more informative and pedagogical; (iv) pop-up cross-references, so that users do not need to switch pages to access them. Three fields/features were excluded from VoBLing in relation to the original version of VoTec: (i) acronyms; (ii) number of meanings; (iii) translators' search. Finally, only Google was left as an external query resource.

Vocabulary of Digital Marketing

The penultimate terminographic product developed and hosted on VoTec is the Vocabulary of Digital Marketing (Mateus, 2022), available at: http://pos.votec.ileel.ufu.br/. This is a bilingual (Portuguese/English)

bidirectional reference work prototype driven by and based on comparable corpora of academic and instructional texts. It is aimed at translators and undergraduates in Business courses.

The same updated version used for the construction of VoBLing (version 1.5) was used in the development of this vocabulary. Currently, 18 entries (9 in each language) are available. They cover both the definition of the Digital Marketing area, as well as each of its 8 subareas. Figure 14 shows one of its entries.

Figure 14. 'PPC/*Links patrocinados*' entry.
Source: VoTec's query page – Vocabulary of Digital Marketing (Mateus, 2022).

According to the author, among other WETM, VoTec was chosen due to the following reasons: (a) it is free to use; (ii) its flexibility and usefulness for different types of terminographic research has already been proven over the years; (iii) as VoTec allows web hosting, it is easier for users to access the Vocabulary.

Third update: version 2.0

After many years of bilingual projects, a possibility of multilingualism was added to the core of VoTec.

Vocabulary of Angiology and Cardiovascular Surgery (AngCiV)

The first project developed in this third update of the platform (and current one) was Santos' PhD dissertation (2023) about different ways of creating a definition according to the level of specialization of the target audience. With the aim of presenting possibilities for accessible scientific writing and text simplification, based on the bibliography on textual simplification (Paraguassu, 2018) and terminological accessibility (Finatto, 2022), the author made available on VoTec a vocabulary prototype of the area of Angiology and Vascular Surgery composed of 9 terms. It is available at: http://votec2.ileel.ufu.br/.

Four different Portuguese corpora were compiled for her research: (i) a specialized corpus with texts written by specialists to other specialists; (ii) a specialized corpus of scientific dissemination made up of texts written by specialists to lay audiences; (iii) a journalistic corpus of scientific dissemination with texts written by journalists to lay people; (iv) a lay corpus of scientific dissemination composed of texts and video subtitles produced by lay people to lay audiences. Based on these corpora, AngCiV presents four entries for each term, following the level of specialization of the target audiences to which the texts of the aforementioned corpora were written, as shown in Figure 15.

This latest version of the platform intends to be plurilingual and pluridirectional, which means that projects like TermosTeo (same Portuguese term twice in contrast), AngCiV (same Portuguese term four times in contrast) and new ones can be built contrasting different languages or the same language many times, going beyond VoTec's previous versions of bilingual bidirectional contrasts (English/Portuguese). The control panel was also updated to reflect the new features of this version, such as the multilingual options, as can be seen in Figure 16.
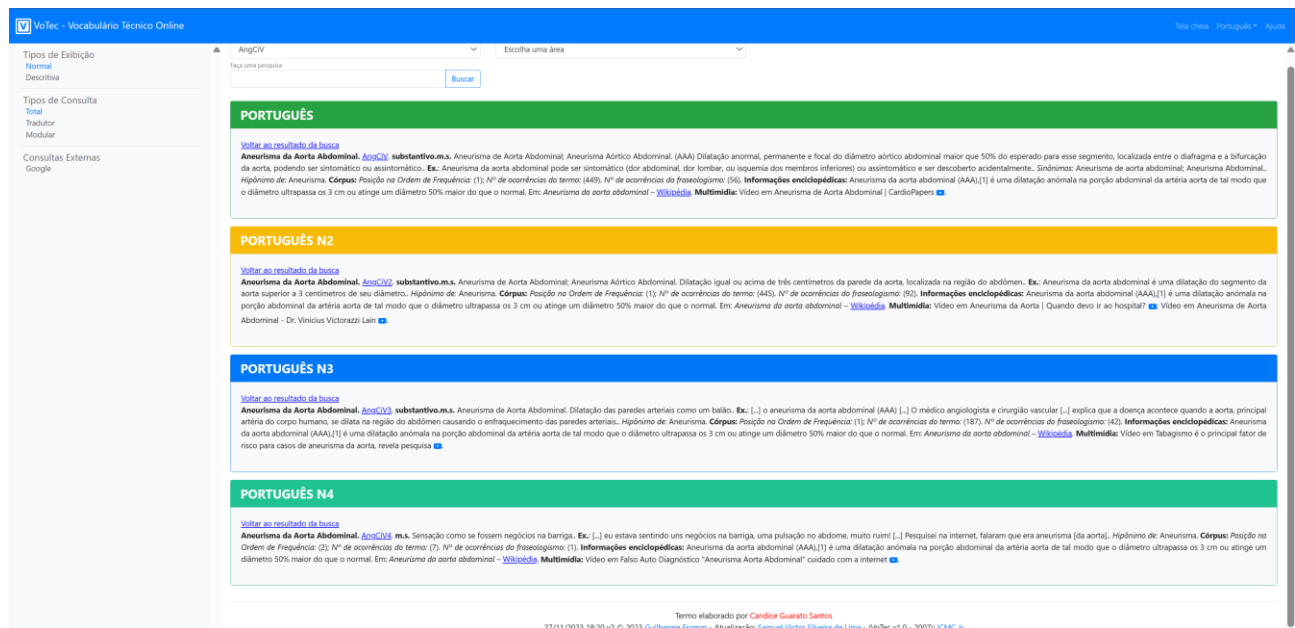
Figure 15. Entries for the term *Aneurisma da Aorta Abdominal*.
Source: VoTec's query page – AngCiV (Santos, 2023).
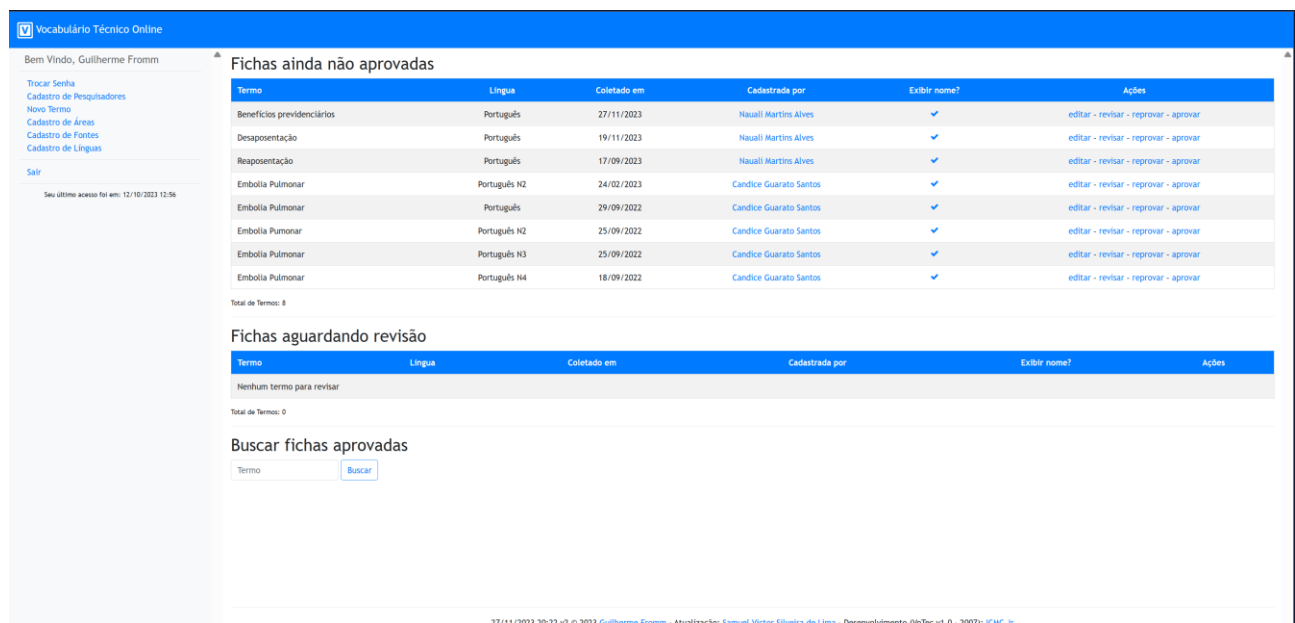
Figure 16. VoTec's version 2 control panel.
Source: VoTec's databank.

Final remarks

In this paper, we presented the projects and research developed through VoTec between 2007 and 2023. This overview demonstrates the usefulness and productivity of VoTec for research and projects in the scope of Terminology, Translation, and Monolingual, Bilingual, Multilingual and Pedagogical Terminography.

The panorama presented here indicates that, unlike some other WETM that were discontinued or remained static, VoTec has been constantly updated for years, which allows it to follow latest trends and technologies in online terminographic works, as well to get closer to the needs of current users of online reference works. In addition, due to the ease of updating, it is increasingly robust and multimodal.

This overview also illustrates the plasticity of VoTec, as it easily adapts to different research and is capable of producing terminographic products with different purposes and target audiences. As we detailed throughout this paper, it is possible to add, delete and modify fields and functionalities in order to customize both the databank and the query page according to the target product and/or audience.

References

Aijmer, K. (2008). Parallel and comparable corpora. In A. Lüdeling, & M. Kytö (Eds.), *Corpus linguistics: an international handbook* (p. 275-292). Berlin, BE: Walter de Gruyter.

Aubert, F. H. (2001). *Introdução à metodologia da pesquisa terminológica bilíngue*. São Paulo, SP: FFLCH/CITRAT.

Bocorny, A., Villavicencio, A., Kilian, C., & Wilkens, R. (2010). A construção de um glossário bilíngue (inglês/português) multimeios online colaborativo para aprendizes baseado em corpus especializado da área de Relações Internacionais. *Revista Trama, 6*(12), 9-25.

Cardoso, S. A. F. (2017). *TermosTeo: a elaboração de vocabulários monolíngues de termos da Teologia em um estudo conduzido por corpus* (Tese de Doutorado). Universidade Federal de Uberlândia, Uberlândia.

Corpógrafo V5. (2023). *Linguateca*. Retrieved from https://www.linguateca.pt/corpografo/

CSOFT International (2023). *TermWiki pro: terminology made easy*. Retrieved from https://pro.termwiki.com/

e-Termos. (2023). *Ambiente colaborativo web de gestão terminológica*. Retrieved from https://www.etermos.cnptia.embrapa.br/

Carvalho, D. A. (2015). *VocTur: proposta de vocabulário bilingue bidirecional do Turismo* (Master's Thesis). Universidade Federal de Uberlândia, Uberlândia.

Finatto, M. J. B. (2022). Acessibilidade textual e terminológica, o que é isso? In M. J. B. Finatto, & L. B. Paraguassu (Orgs.), *Acessibilidade textual e terminológica* (p. 16-40). Uberlândia, MG: EDUFU.

Fromm, G. (2007). *VoTec: a construção de vocabulários eletrônicos para aprendizes de tradução* (Tese de Doutorado). Universidade de São Paulo, São Paulo.

Fromm, G. (2011a). Ficção, Tradução, Terminografia e Linguística de Corpus: confluências. In *Anais do 3º Simpósio Internacional de Letras e Linguística* (p. 1-16). Uberlândia, MG.

Fromm, G. (2011b). Linguística de Corpus e ensino de Terminografia para alunos de Letras e Tradução: uma proposta. *Revista do SELL*, *3*(1), 285-302.

Fromm, G. (2013). A questão da taxonomia num corpus colaborativo para construção de um vocabulário na área de Linguística. In *Anais do 4º Simpósio Internacional de Letras e Linguística* (p. 1-9). Uberlândia, MG.

Fromm, G. (2020). Por uma Terminografia Pedagógica. *Estudos Linguísticos*, *49*(2), 761-776.

Fromm, G., & Yamamoto, M. I. (2013). Terminologia, Terminografia, Tradução e Linguística de Corpus: a criação de um vocabulário bilíngue sobre Linguística. In S. E. O. Tagnin, & C. R. Bevilacqua (Orgs.), *Corpora na Terminologia* (p. 129-151). São Paulo, SP: Hub Editorial.

Fromm, G., & Yamamoto, M. I. (2021). Compilação, reciclagem e padronização de um Corpus Colaborativo de Linguística: percursos metodológicos. *Revista de Estudos da Linguagem*, *29*(3), 2041-2078.

Ilari, R. (2003). *Introdução ao estudo do léxico: brincando com as palavras*. São Paulo, SP: Contexto.

Interverbum Technology (2023). *TermWeb*. Retrieved from https://termweb.store/

Mateus, L. A. (2022). *Da Administração ao Marketing Digital: um estudo taxonômico-terminográfico bilíngue baseado em corpus* (Dissertação de Mestrado). Universidade Federal de Uberlândia, Uberlândia.

Oliveira, L. H. M. (2009). *e-Termos: um ambiente colaborativo web de gestão terminológica* (Doctoral Dissertation). Universidade de São Paulo, São Carlos.

Paraguassu, L. B. (2018). *Tradução especializada acessível (TEA): revisão do tema e proposta de disciplina para cursos de graduação em tradução* (Dissertação de Mestrado). Universidade Federal do Rio Grande do Sul, Porto Alegre.

Santos, C. G. (2023). *Propostas de simplificação de definições de termos para a divulgação científica: um estudo contrastivo por meio de corpora* (Tese de Doutorado). Universidade Federal de Uberlândia, Uberlândia.

Sarmento, L., Maia, B., & Santos, D. (2004). The corpógrafo – a web-based environment for corpora research. In *Proceedings of the 4th International Conference on Language Resources and Evaluation* (p. 449-452). Lisboa, PT.

STAR Group. (c2023). *WebTerm: web-based terminology*. Retrieved from https://www.star-group.net/en/downloads/webterm

Wilkens, R., Bocorny, A., Kilian, C., & Villavicencio, A. (2012). Ambientes web de gestão terminológica para a criação de produtos terminológicos on-line. *Debate Terminológico, 8*(1), 1-7.

Yamamoto, M. I. (2015). *Linguística Histórica e Linguística de Corpus – caminhos que se cruzam para desvelar a história da linguagem: um vocabulário bilíngue português-inglês* (Dissertação de Mestrado). Universidade Federal de Uberlândia, Uberlândia.

Yamamoto, M. I. (2020). *VoBLing: vocabulário bilíngue de Linguística, português-inglês, direcionado por corpus* (Tese de Doutorado). Universidade Federal de Uberlândia, Uberlândia.