Pict2Text 2.0. Translating messages with pictograms into text

Final project
Course 2021–2022

Autors
Gasan Mohamad Nazer
and
Veronika Borislavova Yankova

Directors
Virginia Francisco Gilmartín
and
Susana Bautista Blasco

Bachelor's Degree in Software Engineering
Faculty of Informatics
Complutense University of Madrid

Pict2Text 2.0. Translating messages with pictograms into text

Bachelor's Degree Final Project in Software Engineering

Autors
Gasan Mohamad Nazer
and
Veronika Borislavova Yankova

Directors
Virginia Francisco Gilmartín
and
Susana Bautista Blasco

Convocation: *June 2022*

Bachelor's Degree in Software Engineering
Faculty of Informatics
Complutense University of Madrid

January 10, 2021

Contents

Chapter 1

Introduction

This chapter will display the motivation behind this project and its objectives. In the end, we present the document structure with the different sections within it and a brief description of them.

1.1. Motivation

Communication is a pillar in interpersonal relationships, a fundamental need in our society. However, for some people communication requires a lot of effort. Their differences create an uncrossable barrier and almost impossible human connection using the traditional way of communication. To remove this barrier an alternative approach should be used, for example the use of alternative ways of communication as pictograms. These graphic images, representing an object or a concept, have helped people with special needs to communicate. However, the majority of the population does not understand pictograms.

In order to include pictograms into the communication between people with disabilities and those without, we can use modern technology and create a software to be a mediator between the two sides.

Currently, multiple tools allow transforming natural language into pictograms, one of which is the first version of this project - Pict2Text, but there is no such tool that does the inverse - from an image of pictograms to text. Pict2Text translates pictograms into natural language, but to create the text with pictograms, the pictograms must be selected manually searching for every single one in a search engine by entering the words associated with the pictograms. This is not good enough as people with disabilities who need the pictograms, cannot think about the words that compose the sentence and search for them in the search engine. This problem can be solved, giving those people the option to upload or take a picture of a sentence

constructed with pictograms. By creating a new version of Pict2Text, we aim to reduce the exclusion of those people with special needs from the society and break the communication barrier, providing the aforementioned functionality. Another objective of ours is the improvement of the translation since currently only simple sentences can be translated.

The beneficiaries of this software are people who don't understand pictograms but want to communicate with people who are using them. This project will help people understand each other better, creating a more equal society, independently of the ability to use the natural language in our communication.

1.2. Goals

The main goal of our project is to improve Pict2Text. To do that, we have two objectives - to build an application that can recognize pictograms on uploaded or taken at the moment pictures and improve the translation that Pict2Text provides given a set of pictograms.

We use services-oriented application architecture, constructing web services, and microservices, to increase maintainability and scalability. For the same reasons, we also follow industry standards during the whole process.

Last but not least, we want to consolidate and expand the knowledge acquired during the Software Engineering Degree.

1.3. Document structure

The document is structured as follows. Chapter 2, State of the Art, covers the state of the first version of Pict2Text, the functionalities we will include to increase its usefulness, and the general idea of the machine learning algorithms and models we will be using. Chapter 3, Software development methodology, describes the methodology we have chosen to use, it's characteristics, and things specific to our application of it.

Chapter 2

State of the Art

In this chapter, we have briefly defined Augmentative and Alternative Systems of Communication (section 2.1) and pictograms (section 2.2). Also, in section 2.3, we review Pict2Text 1.0, which is the tool that serves as the basis for our work. In section 2.5. we present the machine learning models we have analyzed to solve our problem and their characteristics.

2.1. Augmentative and Alternative Systems of Communication (AASC)

The Augmentative and Alternative Systems of Communication (AASC) are communication systems, alternative to the natural language, that don't use spoken or written words but can transmit information. They are used by people, who cannot use natural language, or it is not sufficient for them to express themselves. They are created to increase the communication capabilities of the people who use them.

The AASCs do not arise naturally, but they need previous knowledge. There are two types of AASCs - those who need additional equipment such as objects, pictures, pictograms, etcetera, and those that do not need any equipment, depending on the needs, as they are often personalized to match the needs of its user.

They include different systems of symbols: graphic and gestural. The gestural symbols can vary widely from mimicry to hand signs. The graphic symbols can be used by both people with an intellectual disability or with a motor disability. Examples of graphic symbols are drawings and pictures, as well as pictograms, which will be better explained in the next chapter.

Those systems provide various benefits for their users. They prevent or decrease the isolation of people with disabilities, helping with the improvement of social and communication abilities. Also, AASCs are relatively easy to

learn and apply, and adapted for modern technology.

!!!LINKS!!!

2.2. Pictograms

Pictograms are written signs representing objects from the real world, ideas, actions, etcetera. In general, they represent anything that a person would want to express, without the need for verbal or written language. Pictograms are used in the day to day life at hospitals, malls, airports, etcetera because they are not strictly related to the spoken language. Pictograms are widely used by people with special needs, to help with communication and social integration.

As pictograms are not universal, various systems exist, such as Blissymbolics, CSUP, Minspeak, ARASAAC, and more.

!!!PICTURE OF A SINGLE PICTOGRAM!!!

2.2.1. Blissymbolics

Blissymbolics, created in 1949, is an ideographic language consisting of several hundred basic symbols, each representing a concept. Each symbol is represented by basic forms (circles, triangles) and universal signs (numbers, punctuation signs) and uses colour codification to mark the grammar category. They can be combined to generate new symbols that represent new concepts. Blissymbolics characters do not correspond to the sounds of any spoken language and have their use in the education of people with communication difficulties. !!!LINK!!! !!!PICTURE!!!

2.2.2. CSUP

The Communication System Using Pictograms (CSP), developed in 1981 by Mayer-Johnson, is one of the systems that use pictograms to support interactive non-verbal communication. This system uses pictograms not only for objects but with events as well. It is designed in a way that it can be used between a person with a disability and a non-disabled person, child and adult, people speaking different languages, and so on. !!!LINK!!! !!!PICTURE!!!

2.2.3. Minspeak

Minspeak (Semantic Compaction Systems, s.f.) is a pictographic system created by Bruce Baker in 1992. Unlike other systems, this system is based on multi-meaning icons whose meaning is determined by the speech therapist and the user. Usually, two or three icons are combined, determined by rule-driven patterns, to code vocabulary. !!!LINK!!! !!!PICTURE!!!

2.2.4. ARASAAC

The ARASAAC system is the most used pictogram system in Spain. The ARASAAC project was created in 2007, and it currently consists of more than 30.000 pictograms, including complex pictograms with already constructed phrases, in more than 20 languages. The pictograms are separated into five groups - coloured pictograms, black and white, photographs, and sign language videos and pictures. Unlike other pictogram systems, ARASAAC makes a difference between singular and plural and gender differentiation.

One word can be represented by various pictograms.

Verbs come with a different pictogram for every conjugation, and the tense is determined by pictograms representing yesterday, today, and tomorrow.

ARASAAC is free to use, internationally recognized, and provides a wide variety of pictograms. For those reasons, we decided to use the pictograms we can obtain from their website. !!!LINK!!! !!!PICTURE!!!

2.3. Communication based on pictograms

Communication via pictograms happens with the help of a board, where the person points to the pictograms one by one to form a sentence. The complexity of the phrases is limited, usually consisting only of subject, verb, and object. Often, only the most significant words are used, although ARASAAC has pictograms for determinatives and prepositions. !!!LINK!!! !!!PICTURE!!!

2.4. Pict2Text 1.0

As described previously, Pict2Text 1.0 is the initial state and the base of our project. The first version of this project is a web application that permits the translation of sentences written using pictograms to natural language (Spanish).

2.4.1. User perspective

Using the user interface we can write and search a specific word from the ARASAAC pictogram database and display it into a panel on the right part of the web page. After that, we can include the chosen one into the pictogram sentence panel, from where later the message with pictograms will be translated into natural language. The following images and descriptions present a simple flow of actions a user can do to achieve the above-mentioned behavior. When entering the website[1] the user can see on the left part, a big panel, the pictogram sentence panel, with a caption “Pictograms” above

[1] https://holstein.fdi.ucm.es/tfg-pict2text

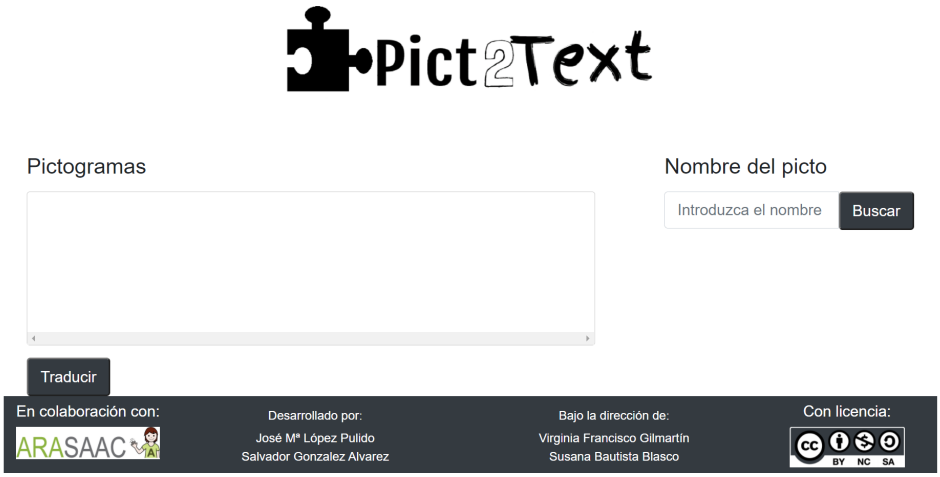

Figure 2.1: Pict2Text version 1 website

it, and a button "Traducir" below it. On the right part, an input box with a caption "Nombre del picto", and a button "Buscar" on the left of it. (See picture 2.1)

To translate pictograms to natural language, the user should first search for a pictogram. To do that, they should write the world they are looking for in the input box of the right side and click the button "Buscar". In image 2.2 it is shown a search of the word "Hombre" and the corresponding pictogram.

Having the pictogram, the next action needed is to include it into the left panel with pictograms by clicking the button "Añadir". In image 2.3 it is shown the pictogram corresponding to the word "Hombre" included in the pictogram sentence panel.

Repeating the previous steps with other words, a sentence can be formed. In image 2.4 it is shown a translation of a sentence written with the pictogram corresponding to the searched words "Hombre", "Comer", "pizza".

2.4.2. Implementation

For the front-end of the project, Angular (put reference) was used. As the website itself is a SPA (Single-Page Applications), which needs to respond fast, a framework like Angular fulfills this performance requirement.

As all of the different functionalities from the project were implemented as web services, most of them in Python, the team of 'Pict2Text' version 1 have decided to use the framework Django (reference) for integration and intercommunication between them.

The core of Pict2Text 1.0 is the API of ARASAAC (reference). It provides the searching mechanism used to match words to pictograms, the

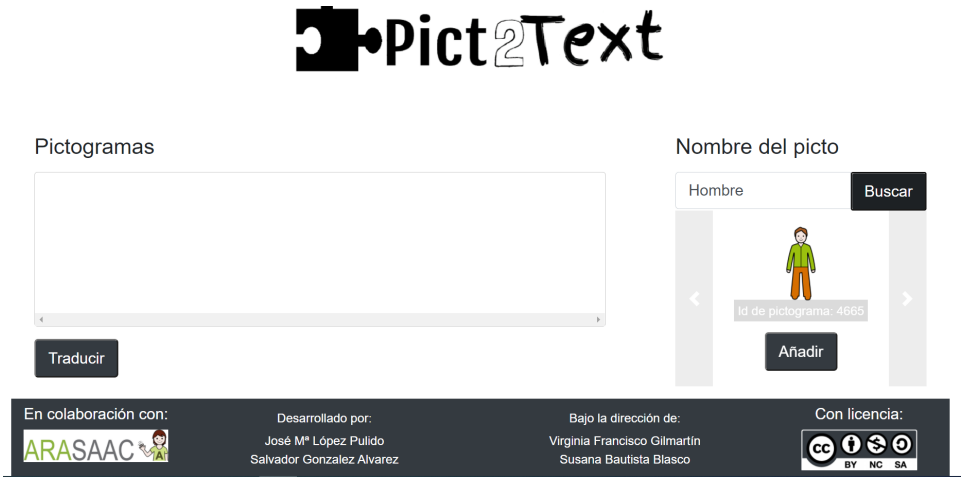

Figure 2.2: Searching the word "Hombre"

graphical images of pictograms, and additional information about them.

A generator of grammatically correct phrases in Spanish was needed. SimpleNLG (reference) was used in Pict2Text 1.0, a Java library for natural language generation. This library permits the creation of simple and complex phrases. To do that, it requires sentence structure- subject, verb, adjectives, gender, and number (singular or plural) of every word in the formed sentence. With this information, SimpleNLG can generate grammatically correct sentences.

Spacy (reference) is the tool that gives the previous word characteristics. It is a python library with a high accuracy used for advanced natural language processing.

2.4.3. Conclusions

Although the first version of Pict2Text translates the pictograms into natural language, it requires the user to manually select the pictograms they want to use in the construction of their sentence. Writing the words is impossible for people with disabilities who need pictograms to communicate if it was not, they would have used the natural language in the first place.

This problem can be solved, giving those people the option to upload a picture of a sentence constructed with pictograms. The functionality we are building will be able to separate the different pictograms from the original image and later translate the phrase using the implementation in Pict2Text 1.0.

Another issue that exists in this project is the translation of pictograms to natural language. The algorithm used allows the correct translation of simple phrases with one subject and one verb. One of our goals is to improve that

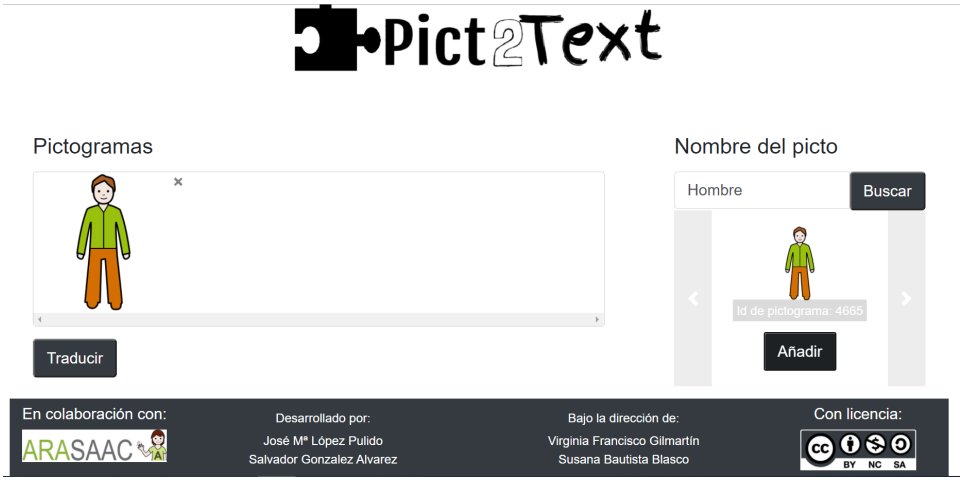

Figure 2.3: Adding the pictogram "Hombre" to the pictogram sentence panel

performance and provide the opportunity to construct complex sentences with a high level of assertion.

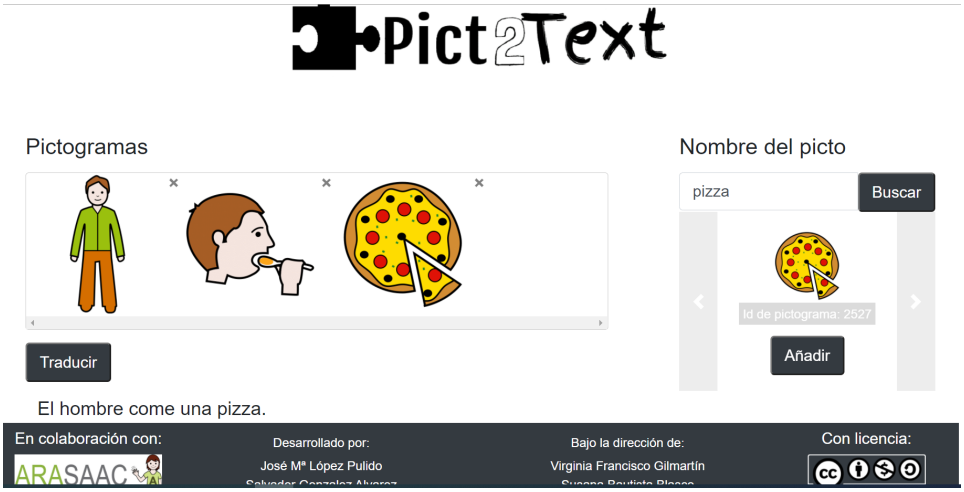

Figure 2.4: Translating the sentence "El hombre come una pizza."