

# Generación automática de documentos

## Automatic document generation

Autor 1: Javier Mauricio Gil Cardona

*Ingeniería de sistemas y computación, Universidad Tecnológica de Pereira, Pereira, Colombia*

Correo-e: mauriciogi@utp.edu.co

**Resumen**— Internet está lleno de documentos, noticias, papers que hablan del mismo tema, cualquiera que sea y estos son usados para crear a su vez nuevos documentos, lo que se quiere para este paper es mostrar cómo se puede hacer un generador de documentos automático con inteligencia artificial y web scraping, esto con el fin de disminuir tiempos a la hora de realizar dichos documentos.

**Palabras clave**— *Generador de documentos, Inteligencia artificial, Web scraping.*

**Abstract**— The Internet is full of documents, news, documents that talk about the same subject, whatever it is and these are used to create new documents once, what you want for this paper is to show how you can make an automatic document generator with artificial intelligence and web scraping, this in order to reduce the time when making these documents.

**Key Word** — Document generator, Artificial intelligence, Web scraping.

### I. INTRODUCCIÓN

Gran cantidad de documentos se encuentran en internet que hablan del mismo tema, esto es de gran utilidad para la construcción automática de otros documentos. Los motores de búsqueda usan algoritmos similares para encontrar alguna página web

El área de investigación de generación automática de documentos apunta a ayudar a optimizar el tiempo en la escritura de distintos tipos de documentos.

Aun con los avances que tenemos hoy en el ámbito informático concretamente en inteligencia artificial es muy difícil hacer un software de generación automática de documentos debido a las limitaciones en inteligencia artificial que aun poseemos.

La implementación de este software al menos en sus inicio no pretende reemplazar a las personas que se encargan de hacer todo este tipo de documentos, debido a que es un software basado en inteligencia artificial, es decir; el documento puede

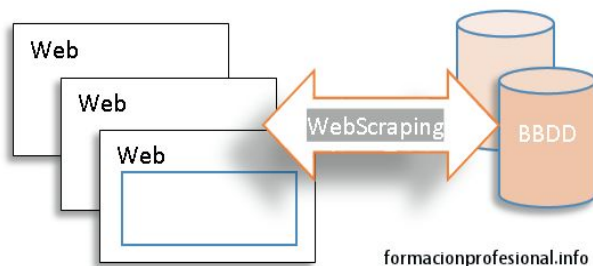
no estar al nivel de uno hecho por un columnista o escrito hábil, este software en cambio será de mucha ayuda y optimización para estas personas debido a que se generará un documento de cualquier tema y ello simplemente lo modificarán a su antojo.

Este software puede ayudar a personas que se destaquen en ambientes donde la escritura de documentos sea necesaria, como un periodista, escritor, columnista, etc.

### II. CONTENIDO

El sistema de generación de documentos automáticos está pensado para que funcione mediante **web scraping** lo que hace esta técnica es mediante programas extrae información de sitios web haciendo como una especie de escaneo y filtrando en estos. Esta técnica es muy usada en los motores de búsqueda debido a su gran potencial en analizar y extraer informaciones de sitios web.

La idea es que mediante un diccionario se almacenen palabras clave de los textos a los cuales se les aplicó la técnica de web scraping y estos mediante filtros cada vez más explícitos empiecen a generar un documento en base al tema requerido donde su contenido sea acorde y tenga sentido lógico



### III. CONCLUSIONES

Este sistema es complejo de resolver debido a que tiene bastantes variables y es demasiado extenso, depende también de muchos filtrados y decisiones, pero tampoco es imposible de resolver, aunque la dificultad ya estaría en dejar el documento perfectamente redactado, como lo haría un buen columnista.

### REFERENCIAS

<https://www.docpath.com/art-document-production-software/?lang=es>  
[https://es.wikipedia.org/wiki/Web\\_scraping](https://es.wikipedia.org/wiki/Web_scraping)





