

Tipo de artículo: Artículo de revisión  
Temática: Ingeniería y gestión de software  
Recibido: 20/09/19 | Aceptado: 04/01/2020 | Publicado: 06/01/2020

## **Revisión sobre los Algoritmos Fonéticos utilizados para realizar búsquedas en los sistemas de información**

### *Review of the Phonetic algorithms used to perform searches information systems*

**Selianne Labañino Urbina<sup>1\*</sup>, Camileidis Labañino Gainza<sup>2</sup>, Orlando Grabiél Toledano López<sup>3</sup>**

<sup>1</sup> Departamento de Informática. Facultad 4. Universidad de las Ciencias Informáticas. Carretera a San Antonio de los Baños, Km 2 ½. Torrens Boyero. La Habana, Cuba. [slabanino@estudiantes.uci.cu](mailto:slabanino@estudiantes.uci.cu)

<sup>2</sup> Departamento de Informática. Facultad 4. Universidad de las Ciencias Informáticas. Carretera a San Antonio de los Baños, Km 2 ½. Torrens Boyero. La Habana, Cuba. [camil@uci.cu](mailto:camil@uci.cu).

<sup>3</sup> Departamento de Informática. Facultad 4. Universidad de las Ciencias Informáticas. Carretera a San Antonio de los Baños, Km 2 ½. Torrens Boyero. La Habana, Cuba. [ogtoledano@uci.cu](mailto:ogtoledano@uci.cu).

\* Autor para correspondencia: [slabanino@estudiantes.uci.cu](mailto:slabanino@estudiantes.uci.cu)

---

#### **Resumen**

La existencia de los algoritmos fonéticos obedece a la necesidad de recuperar información que tiene semejanza sonora y cuya presentación a través de la palabra puede diferir de su pronunciación. Evidenciándose en el mundo actual con la existente tasa de error en palabras o datos escritos idénticos, debido a errores ortográficos y tipográficos, variaciones en el orden de las palabras y el uso de prefijos y sufijos. Actualmente ya existen algoritmos que han sido trabajados no precisamente para el idioma en español, pero que nos ayudarán a tener un enfoque más claro en una búsqueda comparativa que permita dar resultados sin importar errores ya descritos anteriormente, como es el caso del Algoritmo Soundex. Se realizó una revisión de la literatura científica sobre el uso de este tipo de algoritmo, identificando cuáles son las características más relevantes de los mismos, ventajas de su utilización y otros elementos de interés. Resultando de gran relevancia ya que actualmente los sistemas de información utilizan los métodos de codificación fonética para simplificar la búsqueda en las bases de datos cuando sólo se conoce la pronunciación de un nombre propio pero no su transcripción exacta, permitiendo tener un conocimiento más amplio sobre la utilidad de

estos algoritmos, propiciando una mejor toma de decisiones y proporcionando novedosas oportunidades para explotar este conocimiento y poder ser aplicado a diversos ambientes.

**Palabras clave:** *algoritmos fonéticos, Soundex, Metaphone, Double-Metaphone, NYSIIS.*

### **Abstract**

*The existence of phonetic algorithms is due to the need to retrieve information that has sound resemblance and whose presentation through the word may differ from its pronunciation. Evidenced in today's world with the existing error rate in identical words or written data, due to spelling and typographical errors, variations in word order and the use of prefixes and suffixes. Currently there are already algorithms that have been worked not precisely for the Spanish language, but that will help us to have a clearer focus in a comparative search that allows us to give results regardless of errors described above, as is the case of the Soundex Algorithm. A review of the scientific literature on the use of this type of algorithm was carried out, identifying which are the most relevant characteristics of these algorithms, advantages of their use and other elements of interest. It is of great relevance since information systems currently use phonetic coding methods to simplify the search in databases when only the pronunciation of a proper name is known but not its exact transcription, allowing to have a broader knowledge about the usefulness of these algorithms, leading to better decision-making and providing new opportunities to exploit this knowledge and be applied to various environments.*

**Keywords:** *phonetic algorithms, Soundex, Metaphone, Double-Metaphone, NYSIIS.*

## **Introducción**

En la actualidad se desarrollan sistemas de información que almacenan un gran cúmulo de datos personales, estos cuentan con distintos módulos en los cuales es necesario implementar algoritmos que permitan obtener una información más completa y certera, contemplando un cierto grado de semejanza en datos primarios (nombres y apellidos). La codificación fonética busca representar con un mismo código aquellas palabras cuya pronunciación es similar partiendo de su forma escrita, permitiendo descubrir palabras fonéticamente similares. Con frecuencia datos que deberían estar escritos de forma idéntica no lo están debido a errores ortográficos y tipográficos, variaciones en el orden de las palabras, uso de prefijos y sufijos. Las técnicas fonéticas para detección de duplicados existentes están orientadas al idioma inglés, lo que en ocasiones dificulta la identificación y corrección de problemas como errores ortográficos en textos escritos en el idioma español.

Un algoritmo fonético es un algoritmo de indexación de palabras con su pronunciación. La utilización de estos algoritmos ofrecen soporte actualmente para un conjunto de interfaces de búsqueda avanzada, implementadas a partir de las necesidades específicas de diferentes sistemas entre los que se incluyen aplicaciones de búsqueda en bibliotecas virtuales como es el caso de AquaBrowser Library e Hibernate Search. La codificación fonética ha sido extendida también a otras especialidades, por lo que se emplea a menudo en algunos traductores potentes como el desarrollado por Google. Además, proporcionan las herramientas de corrección ortográfica de buscadores reconocidos en Internet como Yahoo y SCIRUS Scientific Information, para corregir los criterios de búsqueda cuando son ingresados erróneamente (Manuel & González, 2010).

Una manera de ilustrar esta situación es cuando una persona ingresa al país, su nombre queda archivado en las bases de datos de la Aduana, de varias maneras, debido a errores ortográficos o tipográficos, uso de abreviaturas o simplemente palabras faltantes, haciéndose necesario, la utilización de algoritmos fonéticos que ayuden a unificar todas esas soluciones encontradas, en un solo resultado, el más certero. Se han desarrollado funciones de similitud para la detección de duplicados; algunas de ellas son de tipo fonético como Soundex (Bhatti, Waqas, Ismaili, Hakro, & Soomro, 2014), Metaphone (P.Parmar & K Kumbharana, 2014), Double Metaphone (Koneru, Pulla, & Varol, 2016), Onca, y Nysiis (Gálvez, 2007). Estas técnicas, se basan en la forma como se pronuncian las palabras en un idioma en particular.

A partir de lo anterior se hace necesario realizar un estudio y comparación de algoritmos fonéticos para realizar búsquedas dentro de los sistemas de información, complementando los aspectos teóricos más relevantes de la fonetización de las cadenas de caracteres, así como el estudio de sus principales características y limitaciones, partiendo además del principio de funcionamiento de los algoritmos de normalización y codificación basados en los grupos de similitud fonética definidos para el idioma español. Logrando mostrar aproximaciones precisas de datos primarios (nombres y apellidos) y que sea lo más adaptable en el sistema que se utilice. Para realizar este estudio se utilizará el método de mapeo sistemático en la literatura científica con el objetivo de alcanzar un conocimiento más profundo sobre este tema.

## **Materiales y métodos o Metodología computacional**

Para alcanzar el objetivo propuesto, el presente estudio utilizó la metodología de mapeo sistemático, el cual es un método útil para construir clasificaciones y obtener información sobre el conocimiento existente en una temática específica; por tanto, permite identificar los vacíos y las necesidades en un área determinada, con lo que se acerca a la

definición de un nicho de investigación pertinente (Sandra J. Jácome, Armando Ordóñez, Gineth M. Cerón, 2018). Una revisión sistemática de literatura es una metodología de investigación que se desarrolla para obtener y evaluar la evidencia disponible que sea pertinente sobre un tema específico.

El mapeo se define como un proceso y una estructura de informe que permite categorizar los resultados que han sido publicados hasta el momento en un área específica. A pesar de ser poco exhaustivo permite obtener una visión científica e identificar tendencias sobre el empleo de los algoritmos fonéticos en determinados sistemas. Para ello el mapeo sistemático determina realizar primero preguntas de investigación para obtener estudios primarios, luego se definen criterios de análisis y finalmente se discuten los resultados (Toledano López, 2019).

Cabe notar que no existen, prácticamente, revisiones sobre el mismo tema que aborda este trabajo. No obstante, en uno de los trabajos primarios encontrados (González-Cam, 2008) se revisan algunos elementos de la aplicación de los algoritmos fonéticos en un sistema de marcas y signos distintivos, pero no se aborda sobre las principales diferencias entre ellos. Además, en (Alejandro & Barragán, 2009) los autores llevan a cabo una revisión sistemática de literatura para capturar el estado del arte de estos algoritmos en los sistemas de información. A pesar de esto, solo se focaliza en las técnicas de recuperación de la información, quedando por analizar alcance de la codificación fonética. Finalmente en (Manuel & González, 2010) se hace un resumen sobre estos algoritmos encaminados a darle solución a una problemática en particular (Sistema Aduana en Cuba), aunque en este estudio no se abordan los temas de codificación ni los principios de implementación a tener en cuenta para la aplicación de los algoritmos. De esta forma, la contribución de esta investigación contempla el establecimiento de la línea base del estado del arte sobre la aplicación de los algoritmos fonéticos en los sistemas de información, a partir de la cual, la comunidad científica puede desarrollar propuestas de mejora.

Se plantean varias macropreguntas para resolver a través del mapeo sistemático:

1. ¿Cuáles son los algoritmos fonéticos que se conocen en la actualidad así como sus principales características?
2. ¿Cómo se realizan las particiones de datos hechas por los algoritmos?
3. ¿Cuáles son las principales diferencias entre los algoritmos analizados?

En esta investigación se seleccionaron las bases de dato *Google Scholar*, *Web of Science*, *Springer* que son algunas de las que pueden contener artículos referidos al tema de la investigación. También se realizaron búsquedas

oportunistas en internet, limitándose a artículos de publicaciones académicas que mostraran texto completo, también entre las referencias de los artículos seleccionados y otros artículos ya identificados. El período de las búsquedas comprendió desde el año 2009, inclusive, a febrero del 2019.

Para conformar la cadena de búsqueda en las bases de datos de publicaciones se tuvo en cuenta el foco central de la investigación que eran los algoritmos fonéticos aplicados a los sistemas de información. En el caso de la palabra clave "algoritmos fonéticos " se seleccionaron, además conceptos normalmente utilizados en el área de investigación, tales como: codificación fonética y similitud fonética.

Para la selección de estudios se consideraron los siguientes criterios de inclusión/ exclusión:

- Son elegibles publicaciones científicas relacionadas con los algoritmos de codificación fonética donde se propusieran, discutieran o evaluaran los principales algoritmos o algunas variaciones de los mismos.
- No hubo restricciones sobre el área de aplicación de las publicaciones, siempre y cuando, el enfoque fuera los algoritmos de codificación fonética.
- Se incluyeron estudios cualitativos y cuantitativos realizados por profesionales en el tema.
- Se excluyeron los artículos que solo tratan el uso específico de los algoritmos fonéticos, sin especificar un proceso real donde estuvieran incluidos.
- Si se encuentran varios artículos del mismo grupo de investigadores que tratan la misma propuesta desde diferentes perspectivas, se considerará el de mayor completitud.
- Se excluyeron todos aquellos estudios en los que su enfoque principal se alejara de los algoritmos fonéticos.

## **Resultados y discusión**

1. ¿Cuáles son los algoritmos fonéticos que se conocen en la actualidad?

La codificación basada en la similitud fonética de los nombres personales se aplica principalmente a los nombres y apellidos para reducirlos a una forma común. La mayoría de estos sistemas se desarrollaron originariamente para el idioma inglés. Entre los procedimientos de codificación fonética más conocidos se encuentran los sistemas *Soundex* desarrollado en 1918, *Daitch-MokotoffSoundex* de 1985, *Metaphone* de 1990 y *NYSII* en 1970.

### **Soundex**

Este algoritmo fue desarrollado en 1918 por Robert Russell and Margaret Odell y patentado finalmente en 1922. Una variante de este algoritmo fue llamado *American Soundex*, y utilizando en 1930 para el análisis retrospectivo de los censos de Estados Unidos de 1890–1920. *Soundex* es un sistema mediante el cual los valores se asignan a los

nombres de tal manera que los nombres que suenan similares obtienen el mismo valor. Estos valores se conocen como codificaciones de Soundex. Una aplicación de búsqueda basada en Soundex no buscará un nombre directamente, sino que buscará la codificación correspondiente. Al hacerlo, obtendrá todos los nombres que suenan como el nombre que se busca (Bhatti et al., 2014).

Este algoritmo (su versión estadounidense) coincide con palabras del índice numérico como A126. Su principio de funcionamiento se basa en la partición de consonantes en grupos con números ordinales, que luego se compilan al valor resultante (González-Cam, 2008).

B P F V	1
C S K G J Q X Z	2
D T	3
L	4
M N	5
R	6

**Original Soundex**

B P	1
F V	2
C K S	3
G J	4
Q X Z	5
D T	6
L	7
M N	8
R	9

**Refine Soundex**

Tabla 1: Particiones del algoritmo *Soundex*. Las letras que no están enumeradas en la tabla (todas las vocales y algunas consonantes) son ignoradas. Letras adyacentes o letras del mismo grupo separadas por H o W están escritas como una letra. El resultado se trunca a 4 caracteres. Las posiciones perdidas contienen ceros.

Ejemplos: Cielo C400                      Giraldo C643

Es claro, que *Soundex* no está orientado al español ya que ni siquiera contempla el juego de caracteres (“ñ”, “ll”). Asimismo, la dependencia de la letra inicial, la agrupación por punto de articulación del idioma inglés y el estar limitado a cuatro caracteres implica que no es eficiente para detectar errores ortográficos comunes en el idioma español (Amón, Moreno, & Echeverri, 2012).

**Metaphone - 1990**

En 1990 Lawrence Phillips publicó un artículo que describía un sistema Soundex más avanzado al que llamó *Metaphone*. El mismo intenta producir su codificación en función de cómo se pronuncia un nombre en lugar de como se deletrea. Pero está basado únicamente en la pronunciación del inglés. Al igual que *Daitch-Mokotoff*, también usa secuencias de letras en lugar de solo letras. Y a diferencia de todos sus predecesores, basa sus codificaciones en el nombre completo en lugar de truncar después de considerar solo una parte inicial del nombre. No siguió el ejemplo de *Daitch-Mokotoff* de permitir que un nombre tenga más de una codificación, sino que genera una codificación única para cada nombre como se hizo en los sistemas anteriores (Koneru et al., 2016).

El algoritmo de Metaphone tiene una eficiencia algo mejor. Tiene un enfoque diferente al proceso de codificación: transforma la palabra original usando reglas de pronunciación en inglés, por lo que las reglas de conversión son mucho más complicadas. El algoritmo pierde mucha menos información, ya que las letras no están divididas en grupos. El código final es un conjunto de caracteres 0BFHJKLMNPRSTWXY, pero al comienzo de una palabra también pueden aparecer vocales del conjunto AEIOU.

#### Algoritmo de cálculo del código Metaphone (Gálvez, 2007)

1. Elimine todas las letras vecinas que se repiten excepto la letra C.
2. El comienzo de la palabra debe transformarse usando las siguientes reglas: KN → N GN → N PN → N AE → E WR → R
3. Quite la letra B al final, si está después de la letra M.
4. Reemplace C usando las reglas a continuación: Con X: CIA → XIA, SCH → SKH, CH → XH Con S: CI → SI, CE → SE, CY → SY Con K: C → K

<p>5. Reemplace D usando las siguientes reglas:</p> <p>Con J: DGE → JGE, DGY → JGY, DGI → JGI</p> <p>Con T: D → T</p>
<p>6. Reemplace GH → H, excepto que está al final o antes de una vocal.</p>
<p>7. Reemplace GN → N y GNED → NED, si están al final.</p>
<p>8. Reemplace G usando las siguientes reglas</p> <p>Con J: GI → JI, GE → JE, GY → JY</p> <p>Con K: G → K</p>
<p>9. Elimine toda H después de una vocal pero no antes de una vocal.</p>
<p>10. Realice las siguientes transformaciones usando las reglas a continuación:</p> <p>CK → K</p> <p>PH → F</p> <p>Q → K</p> <p>V → F</p> <p>Z → S</p>
<p>11. Reemplace S con X:</p> <p>SH → XH</p> <p>SIO → XIO</p> <p>SIA → XIA</p>
<p>12. Reemplace T usando las siguientes reglas</p> <p>Con X: TIA → XIA, TIO → XIO</p> <p>Con O: TH → O</p> <p>Eliminar: TCH → CH</p>
<p>13. Transformar WH → W al principio. Elimina W si no hay vocal después de eso.</p>
<p>14. Si X está al principio, entonces reemplace X → S, de lo contrario reemplace X → KS</p>
<p>15. Elimina todas las Y que no están delante de una vocal.</p>
<p>16. Elimina todas las vocales excepto la vocal al comienzo de la palabra.</p>

### Doble Metaphone - 2000

El nombre proviene del hecho de que produce dos codificaciones para cada nombre. Por lo tanto, no tiene la solidez de *Daitch-Mokotoff* que puede tener muchas codificaciones, pero sin duda es más robusto que los sistemas anteriores que tienen solo una codificación por nombre. Otra característica nueva del *Double Metaphone* es que incluye pronunciaciones extranjeras, pero agrupa todas las reglas extranjeras juntas y no distingue qué regla corresponde a qué idioma. Phillips dejó de lado una de las mejoras de su Metaphone anterior, es decir, la codificación se volvió a restringir a la parte inicial del nombre (González-Cam, 2008).

Este algoritmo produce una clave secundaria junto con una palabra primaria codificada para identificar el nativo más común en cuanto a la pronunciación. Doble Metaphone tiene una codificación extensa para letra "C", "G" y "S", ya que tienen mayor diferenciación en pronunciaciones al combinar con otras vocales y consonantes (P.Parmar & K Kumbharana, 2014). El algoritmo retiene solo el primer sonido de vocal carácter "A" mientras que todos los demás sonidos de vocales son omitidos. Posteriormente se realizan otras transformaciones en las letras restantes basadas en las letras presentes en el índice sucesor y el índice predecesor.

## NYSIIS

Otro código fonético fue el propuesto por Taft en 1970 y desarrollado por *The New York State Division of Criminal Justice*. El sistema de codificación presentado por Taft se denomina New York State Identification and Intelligence Systems (NYSIIS) y se basa en la reducción de los nombres a un código de hasta 6 letras (Koneru et al., 2016).

Las reglas utilizadas por el algoritmo NYSIIS para la codificación fonética son las siguientes:

1. El primer carácter de la clave fonética corresponde al primer carácter del nombre.
2. Traduce los primeros caracteres del nombre MAC → MCC PH → FF KN → NN K → C SCH → SSS

3. Traduce los últimos caracteres del nombre

EE → Y

IE → Y

DT, RT, RD, NT, ND → D

Si el último carácter es S, eliminar.

Si el último carácter es A, eliminar.

Si los últimos caracteres son AY, sustituir por Y.

En su trabajo, Taft compara el NYSIIS algoritmo con Soundex y concluye que NYSIIS tiene una ratio de precisión del 98.72%, mientras la precisión de Soundex es del 95.99%. Sin embargo, Taft pone de manifiesto que tanto Soundex como NYSIIS sólo tratan las variantes de los nombres producidas por errores fonéticos. En 1998 the New York State Division of Criminal Justice sustituye el sistema NYSIIS por el producto NameSearch, por medio del cual no sólo se identifican las variantes fonéticas sino las producidas por errores de transcripción, formas abreviadas, o variantes originadas por la distinta ordenación de las secuencias de los componentes que forman los nombres personales (Carmen Gálvez, 2016).

Aplicaciones de los algoritmos fonéticos (P.Parmar & K Kumbharana, 2014)

- La aplicación de estos algoritmos se puede incorporar en sistemas de reconocimiento de voz para identificar la palabra correcta entre palabras fonéticamente similares.
- En la corrección ortográfica ayuda a producir más de una palabra correcta, con pronombre similares.
- Las aplicaciones de búsqueda pueden proporcionar el conjunto de búsquedas relacionadas cuando se le dan palabras de error ortográfico al encontrar palabras codificadas similares.
- Útil para que los niños aumenten su vocabulario de inglés incluyendo palabras y homófonos que tienen similitudes y que se pronuncian igual pero con un significado diferente.

La comparación de todos estos algoritmos implica un análisis de tiempo, cuestiones de rendimiento, precisión e implementación. La complejidad del tiempo no es el principal problema de desempeño debido a los avances en la programación, mejoras en el lenguaje y en el hardware. Para tal fonética, la precisión del algoritmo es el factor crítico

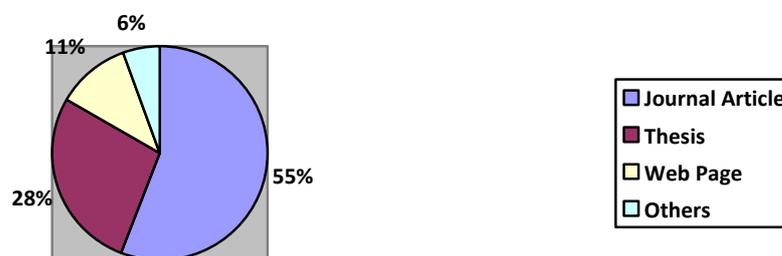
para la corrección del mismo, porque debido a que el idioma inglés es demasiado grande con estructura ortográfica compleja, tiene un significado diferente con pronunciación similar (Moreno & Echeverri, 2012).

**Tabla resumen de los principales elementos de los algoritmos**

Algoritmo fonético	Inconvenientes del algoritmo	Consideraciones importantes
Soundex (Bhatti et al., 2014)	<ul style="list-style-type: none"> <li>• Soundex no está orientado al español.</li> <li>• La dependencia de la letra inicial, la agrupación por punto de articulación del idioma inglés y el estar limitado a cuatro caracteres implica que no es eficiente para detectar errores ortográficos comunes en el idioma español.</li> </ul>	<ul style="list-style-type: none"> <li>• La longitud de la salida no excederá los cuatro caracteres. Las posiciones perdidas contienen ceros.</li> <li>• Elimina cualquier repetición consecutiva de caracteres.</li> </ul>
Metaphone (Gálvez, 2007).	<ul style="list-style-type: none"> <li>• Las reglas de conversión son mucho más complicadas.</li> </ul>	<ul style="list-style-type: none"> <li>• Basa sus codificaciones en el nombre completo en lugar de truncar después de considerar solo una parte inicial del nombre.</li> </ul>
Double Metaphone (González-Cam, 2008).	<ul style="list-style-type: none"> <li>• Incluye pronunciaciones extranjeras, pero agrupa todas las reglas extranjeras juntas y no distingue qué regla corresponde a qué idioma.</li> </ul>	<ul style="list-style-type: none"> <li>• Produce una clave secundaria junto con una palabra primaria codificada para identificar el nativo más común en cuanto a la pronunciación.</li> </ul>
NYSIIS(Manuel & González, 2010)	<ul style="list-style-type: none"> <li>• Solo trata las variantes de los nombres producidas por</li> </ul>	<ul style="list-style-type: none"> <li>• Traduce los primeros y los últimos caracteres del nombre, respondiendo a las reglas de</li> </ul>

	errores fonéticos.	normalización y codificación fonética. <ul style="list-style-type: none"> <li>Solo trata las variantes de los nombres producidas por errores fonéticos</li> </ul>
--	--------------------	---

### Tipología de la literatura consultada



El resultado mostrado se fundamenta en una selección de acuerdo a varios puntos de gran interés, como es la regla de normalización utilizada que debe estar basada en las reglas del idioma, dependientemente de la fonética para las cuales fue desarrollado estos algoritmos. Las posibilidades que brindan, permiten que sea fácilmente adaptable a las reglas de normalización fonética del idioma español.

### Conclusiones

Al término de este trabajo se arriban a las siguientes conclusiones:

- La codificación fonética es útil en la tarea de recuperación de información en documentos orales.
- Existen diversas funciones de similitud sobre cadenas de texto pero no consideran las particularidades del idioma español, lo que las hace inapropiadas en tareas de detección de duplicados originadas por errores ortográficos en este idioma.
- La mayoría de los algoritmos fonéticos se han desarrollado para su uso con el idioma Inglés. La utilización de estos ofrecen soporte actualmente para un conjunto de interfaces de búsqueda avanzada, implementadas a partir de las necesidades específicas de diferentes sistemas.

## Referencias

- Alejandro, M., & Barragán, R. (2009). *Similitud fonética entre palabras para mejorar la Recuperación de Información en Documentos Orales*. Instituto Nacional de Astrofísica, Óptica y Electrónica.
- Amón, I., Moreno, F., & Echeverri, J. (2012). Algoritmo fonético para detección de cadenas de texto duplicadas en el idioma español. *Revista Ingenierías. Universidad de Medellín*, 11(20), 127–138. Retrieved from <http://www.scielo.org.co/pdf/rium/v11n20/v11n20a11.pdf>
- Bhatti, Z., Waqas, A., Ismaili, I. A., Hakro, D. N., & Soomro, W. J. (2014). Phonetic based SoundEx&ShapeEx algorithm for Sindhi Spell Checker system. *Advances in Environmental Biology*, 8(4), 1147–1155.
- Carmen Gálvez, P. (2016). *PERSONAL NAME IDENTIFICATION THROUGH PHONETIC CODIFICATION SYSTEMS*. Retrieved from <http://encontros-bibli-blog.blogspot.com/>
- Gálvez, C. (2007). Identificación de nombres personales por medio de sistemas de codificación fonética. *Departamento de Biblioteconomía Y Documentación*, 11(22), 105–116. <https://doi.org/10.5007/1518-2924.2006v11n22p105>
- González-Cam, C. (2008). Algoritmos fonéticos en el desarrollo de un sistema de información de marcas y signos distintivos. *Biblios: Revista Electrónica de Bibliotecología, Archivología Y Museología*, (32), 8.
- Koneru, K., Pulla, V. S. V., & Varol, C. (2016). Performance evaluation of phonetic matching algorithms on english words and street names comparison and correlation. - *Proceedings of the 5th International Conference on Data Management Technologies and Applications*, (DATA), 57–64. <https://doi.org/10.5220/0005926300570064>
- Manuel, C., & González, R. (2010). *Implementación de un Motor de Codificación y Búsqueda Fonética para el Sistema de Gestión Integral de la Aduana*. Universidad de las Ciencias Informáticas.
- Moreno, F., & Echeverri, J. (2012). *ALGORITMO FONÉTICO PARA DETECCIÓN*. 11(20), 127–138.
- P.Parmar, V., & K Kumbharana, C. (2014). Study Existing Various Phonetic Algorithms and Designing and Development of a working model for the New Developed Algorithm and Comparison by implementing it with Existing Algorithm(s). *International Journal of Computer Applications*, 98(19), 45–49. <https://doi.org/10.5120/17295-7795>

Sandra J. Jácome, Armando Ordóñez, Gineth M. Cerón, A. F. V. (2018). *Mapeo sistemático del uso de las tecnologías de la información y la comunicación en la diabetes tipo 2*. Retrieved from [scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S2307-21132018000400005](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2307-21132018000400005)

Toledano López, O. G. (2019). *Revisión sobre el impacto de las bases de datos NoSQL en el contexto de los Datos Educativos Masivos*.