

43

Fecha de presentación: diciembre, 2019

Fecha de aceptación: enero, 2020

Fecha de publicación: marzo, 2020

LA LÓGICA BORROSA

Y SU IMPORTANCIA EN LA BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN EN LA WEB

BLURRED LOGIC AND ITS IMPORTANCE IN THE SEARCH AND RETRIEVAL OF INFORMATION ON THE WEB

Fernando Juca Maldonado¹

E-mail: fjucam@umet.edu.ec

ORCID: <https://orcid.org/0000-0001-7430-2157>

María Beatriz García Saltos¹

E-mail: mgarcia@umet.edu.ec

ORCID: <https://orcid.org/0000-0003-4015-0153>

Jorge Carrión González¹

E-mail: jcarrion@umet.edu.ec

ORCID: <https://orcid.org/0000-0002-8971-0628>

¹ Universidad Metropolitana. Ecuador.

Cita sugerida (APA, sexta edición)

Juca Maldonado, F., García Saltos, M. B., & Carrión González, J. (2020). La lógica borrosa y su importancia en la búsqueda y recuperación de información en la web. *Universidad y Sociedad*, 12(2), 306-311.

RESUMEN

Una de las actividades más importantes en la web es la búsqueda de información, en ese sentido las grandes empresas de tecnología han implementado varias estrategias en vías de mejorar la experiencia del usuario, una de estas es aplicar lógica borrosa en el proceso dicho proceso. El presente trabajo es una exploración descriptiva del estado del arte de la aplicación de lógica borrosa en las búsqueda y recuperación de información en la internet a través de buscadores y metabuscadores. Se realiza un análisis cualitativo y heurístico de los datos más importantes y recientes, evaluando la situación actual y futura de la aplicación de esta tecnología computacional en buscadores y metabuscadores, concluyendo su importancia e influencia tanto en hardware como en software.

Palabras clave: Internet, información, lógica borrosa, buscadores, meta buscadores.

ABSTRACT

One of the most important activities on the web is the search for information, in that sense the big technology companies have implemented several strategies in order to improve the user's experience, one of these is to apply blurred logic in the process. The present work is a descriptive exploration of the state of the art in the application of fuzzy logic in the search and retrieval of information on the Internet through search engines and meta-searches. A qualitative and heuristic analysis of the most important and recent data is made, evaluating the current and future situation of the application of this computer technology in search engines and meta-search engines, concluding its importance and influence in both hardware and software.

Keywords: Internet, information, blurred logic, search engines, meta search engines.

INTRODUCCIÓN

Hace 30 años no existía la internet como la conocemos hoy, 20 años antes solo había 130 sitios web en todo el mundo y el buscador Google no se había inventado aun y 10 años atrás Skype, Facebook, YouTube, Reddit, Twitter, Tumblr, Dropbox e Instagram no estaban disponibles. Todo esto ha cambiado vertiginosamente en los últimos años, actualmente la cantidad de usuarios, sitios web y actividad online que existe es cada vez mayor y su crecimiento es exponencial. La conversión de usuarios de la Web 1.0 a prosumidores en la Web 2.0, ha provocado la proliferación de sitios web con contenido poco fiable, valido y verás.

Así mismo, la cantidad de información aumenta cada segundo y mucha de esta información puede ser de gran utilidad, la que haya sido verificada por expertos en plataformas educativas; pero también es cierto que mucha de la información disponible en la web no siempre cuenta con ningún tipo de validación ni respaldo alguno, provocando la desinformación en la internet (Hewson, et al., 2003; Miller & Bartlett, 2012). trustworthy journalism, niche expertise and accurate facts at our fingertips than ever before, but also mistakes, half-truths, propaganda and misinformation. This article presents research on how well young people are being equipped to meet the challenge of sorting good information from bad. It reviews current literature on the subject, and presents a new poll of over 500 teachers. With analysis supplemented by additional correspondence from librarians and other IL professionals, it argues that there is strong evidence that the web is fundamental to pupils' learning and lives, but that many are not careful, discerning users of the internet. They are unable to find the information they are looking for, or they trust the first thing they see. This makes them vulnerable to the pitfalls of ignorance, falsehoods, cons and scams. The article proposes the appropriate response to be to embed digital fluency a tripartite concept constituting critical thinking, net savviness and diversity at the heart of learning, in order to create a pedagogical framework fit for the information consumption habits of the digital age. It should be noted that both authors recognise the importance of non-teaching information literacy professionals in these debates. They recognise that the poll's focus on teachers was too narrow, and have endeavoured, subsequent to the poll, to consult more widely in their research.

author: [{ "dropping-particle": "", "family": "Miller", "given": "Carl", "non-dropping-particle": "", "parse-names": false, "suffix": "" }, { "dropping-particle": "", "family": "Bartlett", "given": "Jamie", "non-dropping-particle": "", "parse-names": false, "suffix": "" }], "container-title": "Journal of Information

Literacy", "id": "ITEM-1", "issue": "2", "issued": { "date-parts": [["2012"]] }, "page": "35-55", "title": "'Digital fluency': towards young people's critical use of the internet", "type": "article-journal", "volume": "6" }, "uris": ["http://www.mendeley.com/documents/?uid=424183d5-7ce3-4f58-8a0d-7c39ec2749f1"]], "mendeley": { "formattedCitation": "(Miller & Bartlett, 2012). trustworthy journalism, niche expertise and accurate facts at our fingertips than ever before, but also mistakes, half-truths, propaganda and misinformation. This article presents research on how well young people are being equipped to meet the challenge of sorting good information from bad. It reviews current literature on the subject, and presents a new poll of over 500 teachers. With analysis supplemented by additional correspondence from librarians and other IL professionals, it argues that there is strong evidence that the web is fundamental to pupils' learning and lives, but that many are not careful, discerning users of the internet. They are unable to find the information they are looking for, or they trust the first thing they see. This makes them vulnerable to the pitfalls of ignorance, falsehoods, cons and scams. The article proposes the appropriate response to be to embed digital fluency a tripartite concept constituting critical thinking, net savviness and diversity at the heart of learning, in order to create a pedagogical framework fit for the information consumption habits of the digital age. It should be noted that both authors recognise the importance of non-teaching information literacy professionals in these debates. They recognise that the poll's focus on teachers was too narrow, and have endeavoured, subsequent to the poll, to consult more widely in their research." }, "author": [{ "dropping-particle": "", "family": "Miller", "given": "Carl", "non-dropping-particle": "", "parse-names": false, "suffix": "" }, { "dropping-particle": "", "family": "Bartlett", "given": "Jamie", "non-dropping-particle": "", "parse-names": false, "suffix": "" }], "container-title": "Journal of Information Literacy", "id": "ITEM-1", "issue": "2", "issued": { "date-parts": [["2012"]] }, "page": "35-55", "title": "'Digital fluency': towards young people's critical use of the internet", "type": "article-journal", "volume": "6" }, "uris": ["http://www.mendeley.com/documents/?uid=424183d5-7ce3-4f58-8a0d-7c39ec2749f1"]], "mendeley": { "formattedCitation": "(Miller & Bartlett, 2012).

Actualmente el aumento de la información en la internet es exponencial, pero su consumo y comprensión es lineal, debido a que no somos capaces de procesar tal caudal de datos, lo que da paso a la infoxicación, término abordado por Cornella (1999), la infoxicación se presenta cuando existe más información de la que una persona es capaz de procesar o asimilar, razón por la cual se vuelve

fundamental la búsqueda y clasificación de información válida y veraz en el vasto océano de la internet.

Bajo esta perspectiva planteada, una de las preocupaciones actuales es la recuperación de la información y que la misma sea lo más verídica posible (Nikraves, Loia & Azvine, 2002); así como el uso de técnicas nuevas que permitan que este proceso sea más efectivo y fiable.

El uso de los Metabuscares, los cuales permiten la búsqueda en diferentes buscadores a la vez o incluso en otros Metabuscares (León González, Socorro Castro & Espinoza Cordero, 2017). conjuntamente con el uso de la lógica borrosa, la cual es una lógica multivaluada o heterogeneidad que facilita la representación matemática de la incertidumbre y la vaguedad se presenta como una alternativa para la mejora en la búsqueda y recuperación de información de manera más precisa y efectiva en la web.

Todo está cambiando vertiginosamente en los últimos años, ya que actualmente la cantidad de usuarios, sitios webs y actividad online que existe en cada vez mayor. Se estima que la cantidad de usuarios de internet a inicios del 2017 es de 3,731,973,423 aproximadamente y de 1,179,887,084 de sitios webs (Internet Live Stats, 2017) según firmas especializadas, muchos de los actuales usuarios no solo consumen información sino que también la producen, convirtiéndose en prosumidores, fenómeno producido gracias a la web 2.0, término acuñado por Tim O'Reilly en el año 2004, donde el usuario no solo es un consumidor de información, sino que también tiene la oportunidad de compartir, crear y colaborar con contenido en diferentes plataformas webs, como blogs, redes sociales, etc.

Otra fenómeno que actualmente se presenta en internet, es que según varios estudios se estima que el 75% de los sitios webs están inactivos (Hernández, 2015). Es decir, que su contenido no está actualizado. Adicional a ello, por el fenómeno antes mencionado de los prosumidores, y gracias a las propias características de la red, su accesibilidad y alcance de difusión se presenta la posibilidad de contenido falso o poco veraz, esta es una realidad que cada vez se vuelve más compleja solucionar de la cual los nativos digitales son los mayores perjudicados al ser la generación que más utiliza la internet para informarse, debatir, compartir, participar, socializar y como medio de ocio.

La cantidad de información aumenta cada segundo y mucha de esta información puede ser de gran utilidad, la que haya sido verificada por expertos en plataformas educativas; pero también es cierto que mucha de la información disponible en la web no siempre cuenta con

ningún tipo de validación ni respaldo alguno, provocando la desinformación en la internet.

Por ejemplo, existen muchas fotografías alteradas digitalmente utilizadas para distorsionar, sacar de contexto u ocultar un hecho, información mal interpretada y difundida por los usuarios de redes sociales de manera no intencional que luego se descubre cómo no relevante o veraz son situaciones que se dan a diario en sitios webs y en redes sociales.

¿Surge una pregunta cómo distinguir entre información veraz y falsa? La utilización de buscadores ofrece resultados inmediatos sobre el tema o temas de interés, pero es importante identificar y diferenciar si estos sitios son relevantes y confiables.

Otros investigadores (Piñeiros, et al., 2011) señalan que muchos usuarios dan credibilidad a la información basados en su experiencia en la búsqueda y sus propias percepciones, así como de referencias de los contactos más cercanos. Actualmente el aumento de la información en la internet es exponencial, pero su consumo y comprensión es lineal.

Esta infoxicación se presenta debido a que cada vez es más fácil y barato enviar bits a la red, es decir publicar y compartir contenido. Dado esta situación, una de las actividades que cada vez se está volviendo más importante es la gestión de la información. Es decir, la búsqueda y la clasificación de la información de manera más eficiente, en pocas palabras, filtrar es la clave al momento de consumir información en la internet.

Bajo esta perspectiva planteada, una de las preocupaciones actuales es la recuperación de la información y que la misma sea lo más verídica posible (Nikraves, et al., 2002), considerando a la información en la web como texto, imágenes, sonidos y vídeos, siendo hasta la fecha el uso del texto la forma primordial utilizada para realizar búsqueda y recuperación de información en la web, sin olvidar que el contenido de información en la internet almacenada en esta última década es tan grande como la generada por toda la humanidad en los últimos dos milenios. Es importante diferenciar el tipo de sitio web en el que se busca y recupera información, los orientados a noticias (blogs, diarios, etc.) y los orientados a los documentos o contenido (wikis, repositorios, base de datos, etc.).

El uso de los buscadores en la web se ha convertido en uno de los servicios más utilizados (Nikraves, et al., 2002) al momento de tratar de recuperar información en la internet, tanto es así que actualmente se realizan más de 50.000 búsquedas en un segundo a través del buscador

de Google (Internet Live Stats, 2017); estos motores de búsqueda deben ser eficientes y adaptarse al significado semántico de las peticiones de los usuarios y no tratar las solicitudes como una búsqueda estructurada en una base de datos. Actualmente existen varias herramientas principales para la búsqueda de información en la web:

- Directorios, que organizan la información de manera jerárquica por temas.
- Buscadores, o motores de búsqueda que indexan el contenido de la web.
- Metabusadores, permiten la búsqueda en diferentes buscadores a la vez o incluso en otros metabuscadores.

El estudio se centrará en la utilización de estos últimos, los metabuscadores, junto con lógica borrosa, lo cual ha permitido mejorar sus características y aunque no poseen una base de datos propia indexada, permiten realizar las peticiones de los usuarios en múltiples buscadores y obtener un resultado más diverso.

La computación clásica binaria ha sido usada con éxito desde sus inicios, pero muchas de los nuevos retos tecnológicos no pueden ser abordados con la misma lógica ya que esta no siempre suele ajustarse a requerimientos en un lenguaje más natural.

En este sentido González Morcillo (2011), menciona que básicamente la lógica borrosa es una lógica multivaluada o heterogenea, que permite valores intermedios que facilita la representación matemática de la incertidumbre y la vaguedad (que extiende a la lógica clásica a través de herramientas formales para su tratamiento, es decir, que permite asignar diferentes grados de certeza.

Ya lo indica Zadeh (1973), *cuando aumenta la complejidad, los enunciados precisos pierden su significado y los enunciados útiles pierden precisión*, que puede interpretarse como “los árboles no dejan ver el bosque”. (González Morcillo, 2011)

Es decir, que la lógica difusa es una técnica de inteligencia computacional que permite procesar datos con alto grado de imprecisión a diferencia de la lógica convencional que trabaja con datos precisos. Siempre se ha tratado de lograr la precisión, intentando simular el ambiente real en modelos matemáticos, los cuales son precisos y fijos, a través de una lógica binaria. La lógica difusa trata de flexibilizar estas leyes para que se adapte más a la realidad incierta.

Tomando el ejemplo de Guzmán & Castaño (2006), quienes mencionan que debido a las diferentes formas que las personas tenemos de percibir los conceptos o situaciones tales como alto, bajo, ruidoso, silencio, dulce,

amargo, caro, barato, etc., para una persona de la montaña puede ser caliente la noche en la playa, para alguien de la costa le puede parecer fresco, por lo que los conceptos de caliente, tibio y frío puede ser un conjunto borroso, es decir, que un conjunto borroso es un conjunto con límites que no están “muy bien definidos”.

En un estudio previo Bermúdez, et al. (2005), describen que inicialmente el profesor Zadeh deseaba representar de manera más formal la imprecisión expresada en el lenguaje humano, pero sorprendentemente la lógica borrosa se incluyó en el campo del control de procesos automáticos, específicamente se dio gracias al interés del gobierno japonés y de varias universidades y compañías privadas que a fines de los ochentas desarrollaron nuevas tecnologías aplicando lógica borrosa, en contraposición a occidente que estaba más reacio a cambiar la lógica binaria impuestas siglos atrás desde Aristóteles.

Como ejemplo se puede mencionar, el caso del primer metro controlado por lógica borrosa, controles para maniobrar aviones, sensores de cámaras, lavadores de ropa de la Cia. Bosh que autorregulan la cantidad de jabón, aires acondicionados Mitsubishi, sistemas de foco en las cámaras, optimización de control industrial, sistemas de reconocimiento de escritura, mejoramiento del uso eficiente del combustible en motores, sistemas expertos, base de datos difusas (FSQL), modelamiento con redes neuronales y lógica difusa. Es decir, que el uso de lógica borrosa no es un mero experimento sino una realidad presente en mucha tecnología actual donde no se puede aplicar un modelo matemático convencional o exacto.

En la lógica clásica, los resultados pueden concluirse partiendo de hechos conocidos o implicaciones que se basan en dichos hechos. En la lógica borrosa, se trabaja con datos parcialmente verdaderos.

El propósito esencial sería que los metabuscadores aprovechen la lógica borrosa para realizar búsquedas con aproximaciones semánticas, no sintácticas como actualmente sucede. Se puede mencionar cuatro ventajas de usar metabuscadores:

- Amplía el radio de búsqueda: por la extensión y cantidad de información almacenada en la web, un solo buscador probablemente no contenga toda la información indexada, por lo que usar los resultados de varios buscadores maximiza las posibilidades de mejores resultados.
- Escalabilidad de búsquedas: incremento de información temporal o permanente que puede darse en un sitio web y por ende la adaptación de los metabuscadores para incrementar su rango de resultados.

- Fácil acceso a múltiples buscadores: de manera simultánea se puede acceder a los resultados más relevantes de varios buscadores con una sola consulta.
- Mejora la efectividad de recuperación: Por dar como resultados documentos de buscadores especializados, evita el desvío entre los típicos resultados no relevantes que se pueden obtener en los buscadores tradicionales.

La calidad de los resultados ofrecidos por los metabuscadores dependerá de los buscadores que se encuentren detrás de él (León González, et al., 2017). Es decir, los buscadores en los cuales el metabuscador hace la consulta y como considera Nikravesch et al., (2002), agregar lógica borrosa a un buscador es una necesidad para lograr deductividad en la búsqueda, cosa de la cual carecen actualmente, es decir, tratar de razonar semánticamente para aproximarse a un mejor resultado.

Actualmente buscadores y metabuscadores lo único que hacen es el completar o sugerir la palabra o frase que se utiliza en la búsqueda, pero no aplican lógica borrosa en el proceso de búsqueda en sí mismo; tanto la lógica borrosa, como las redes neuronales u otros procesos de inteligencia artificial se muestran como opciones viables al momento de obtener resultados más precisos.

Ya existen investigaciones de aplicaciones de lógica borrosa, como por ejemplo el de la aplicación de procesos borrosos AHP, el cual refleja el pensamiento humano, mediante el uso de cuantificadores lingüísticos al hacer comparaciones en lugar de números precisos, así los juicios exactos se tornan juicios borrosos. Este tipo de proceso se puede utilizar para realizar las consultas del metabuscador hacia los buscadores.

En la línea de la búsqueda de información, también existen trabajos como el de Ruiz Morilla (2016), y su modelo de polirepresentación, el cual aplica la lógica borrosa para la recuperación de grandes cantidades de información de la internet a través de la generación de un grupo de consultas que se reformulan para el proceso de búsqueda siguiendo las condiciones dadas por el usuario en la consulta original.

También cabe mencionar el uso de ontologías para mejorar las búsquedas, como menciona Olivas (2005), a Gruber *“una ontología es una especificación explícita y formal sobre una conceptualización compartida”*; siendo las ontologías una de las áreas más utilizadas en la web semántica, ya que entre los diferentes estándares y herramientas para la creación de ontologías, la lógica borrosa permite constituir varios grados de relación entre los conceptos de una ontología.

En el campo de la medicina, se abre paso a la aplicación de la lógica borrosa en diversos campos. Por ejemplo, el sistema Neuro-difuso desarrollado por la Universidad del Mar de Plata en Argentina, para la detección de envejecimiento arterial o para el cálculo institucional de los costos generados por los pacientes diabéticos internados. La Universidad de Valencia investiga en la línea de la aplicación de lógica borrosa en el uso de la geometría estocástica en el campo del procesamiento digital de imágenes.

La aplicación de lógica borrosa está ampliándose cada vez más, desde procesos tales como análisis de recursos y capacidades internas de las empresas hasta su uso sensores de todo tipo como el de celdas solares, sin olvidar las mejoras que aún se espera en el campo de las búsquedas en el internet cuando tecnologías como la web semántica, la inteligencia artificial, redes neuronales entre otras maduren con la evolución de nuevos algoritmos y avances del hardware y software.

Se espera realizar un nuevo estudio para ampliar conceptos y temas en esta área, profundizando ya en el uso de nuevas metodologías o paradigmas de búsqueda como el modelo de polirepresentación propuesto por Ruiz Morilla (2016), lo cual permita aportar nuevos análisis en este interesante campo de la lógica borrosa.

CONCLUSIONES

La cantidad de información que se genera en la internet continuará siendo cada vez mayor, por lo cual se demandan nuevas opciones y metodologías para la búsqueda y recuperación de información, de forma rápida y veraz.

Una de las alternativas viables es el uso de metabuscadores para ampliar así el campo de acción de búsqueda y la aplicación en este proceso de lógica borrosa, permitirá efectivizar dicho proceso y elevar el grado de precisión en los resultados.

Las investigaciones en esta área aún están en sus inicios, investigadores como el Dr. Zadeth que iniciaron este campo y como Olivas (2005), que continúa en sus innumerables investigaciones, así como otros muchos, permitirán adelantos y mejoras en las búsquedas y recuperación de la información en la web, a través de procesos como el de la lógica borrosa y otras técnicas de vanguardia.

En procesos donde la lógica binaria solo nos brinda dos opciones, extiende las posibilidades de multivaluar las posibilidades a través del uso de la lógica borrosa.

REFERENCIAS BIBLIOGRÁFICAS

- Bermúdez, A., Barriga, A., Baturone, I., & Sánchez, S. (2005). *Logica difusa*. Instituto de Microelectrónica de Sevilla - Centro Nacional de Microelectrónica.
- Cornella, A. (1999). Cómo sobrevivir a la infoxicación. <https://esthermusicasecundaria.files.wordpress.com/2014/02/reflexic3b3n-sobre-infoxicac3b3n.pdf>
- González Morcillo, C. (2011). Lógica Difusa, una introducción práctica. *Softcomputing*. https://www.esi.uclm.es/www/cglez/downloads/docencia/2011_Softcomputing/LogicaDifusa.pdf
- Guzmán, D., & Castaño, V. M. (2006). La Lógica difusa en ingeniería: Principios, aplicaciones y futuro. *Ciencia y Tecnología*, 24(2).
- Hernández, A. (2015). Internet parece enorme, pero hay menos de 1.000 millones de webs activas. Eldiario.es. http://www.eldiario.es/hojaderouter/shortcut/Internet-enorme-millones-mayoria-inactivas_6_439266075.html
- Hewson, C., Yule, P., Laurent, D., & Vogel, C. (2003). *Internet research methods: A practical guide for the social and behavioural sciences*. Sage.
- Internet Live Stats. (2017). Total number of websites in real time. Internet Live Stats. Retrieved from <http://www.internetlivestats.com/watch/websites/>
- León González, J. L., Socorro Castro, A. R., & Espinoza Cordero, C. X. (2017). Uso de la información científica y tecnológica en la investigación y la innovación. Universo Sur.
- Miller, C., & Bartlett, J. (2012). "Digital fluency": towards young people's critical use of the internet. *Journal of Information Literacy*, 6(2), 35–55.
- Nikraves, M., Loia, V., & Azvine, B. (2002). Fuzzy logic and the Internet (FLINT): Internet, World Wide Web, and search engines. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 6(5), 287–299.
- Olivas, J. A. (2005). Las técnicas de Soft Computing y el futuro de Internet. *ÁGORA - Papeles de Filosofía*, 24(2), 117–127.
- Piñeiros, C., Gutiérrez, L. M., Gómez Díaz, J., Pérez, A. M., Salgado Cardona, A., Mora, M. L., Grijalva, N., Estrada, J., & Ramírez, M. J. (2011). Credibilidad percibida del periodismo a través de la Internet. *Revista Diversitas-Perspectivas en Psicología*, 7(2), 225–238.
- Ruiz Morilla, J. J. (2016). *Polirepresentación de consultas borrosas para la extracción masiva de documentos en Internet 5 de Febrero de 2016*. (Tesis Doctoral). Universidad de Castilla La Mancha.
- Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *Systems, Man and Cybernetics, IEEE Transactions on*, (1), 28–44.