Implantación de un sistema *Business Intelligence* para la Revista DYNA

Implementation of a Business Intelligence tool for DYNA Journal

Aitor Goti-Elordi, Alberto de-la-Calle-Vicente, Aitor Oyarbide-Zubillaga, Elisabete Alberdi-Celaya, Borja Torres-Fuentes, Irati Unanue-Telleria, Pablo García-Bringas y Ander Bacigalupe-Gutiérrez

Universidad de Deusto (España)

DOI: http://dx.doi.org/10.6036/8609

1. INTRODUCCIÓN

El Factor de Impacto (o Impact Factor, IF) de una revista científica es una medida casi cualitativa del prestigio y visibilidad internacional de la misma [1]. Existen varios indicadores que miden dicho impacto, pero la práctica totalidad de los mismos se basa en cuántas veces son citados sus artículos. Es por tanto esencial identificar factores como tipos de artículo, temáticas o autores que sean citados con una mayor facilidad para poder tomar decisiones alineadas no sólo con la estrategia de la editorial sino también con la mejora del impacto de la revista. Así, también las herramientas de gestión y mejora internas (de apoyo a revisores, de toma de decisiones respecto a la aceptación o el rechazo de los artículos, etc.) deberían estar orientadas a mejorar el mencionado factor de impacto. El presente artículo presenta la herramienta BI diseñada para la revista Dyna con el fin de apoyar la mejora interna de la misma en términos de IF.

En referencia a la evaluación externa del impacto de una revista, existen varias listas que miden el mismo: cada una de estas listas cuenta con una serie de campos o áreas científicas en las que se clasifican las revistas y unas métricas con las que se obtienen los valores por los que éstas son posicionadas en una ubicación de mayor o menor prestigio.

De entre éstas destacan las prestigiosas listas *Journal Citation Reports* (JCR) y *Scimago Journal Rank* (SJR), en las cuales se encuentra la Revista DYNA, Ingeniería e Industria. El JCR es una lista publicada por el Instituto de la Información Científica (*Institute for Scientific Information*, ISI), mientras que el SJR fue desarrollada por el grupo SCIMAGO y se basa en la información disponible en la base de datos Scopus® de la mayor editorial de libros de medicina y literatura científica, Elsevier B.V. Ambos realizan una actualización de sus listas de revistas de impacto anualmente, y en ambos la forma de obtener su indicador principal IF es muy similar:

Así, el IF del JCR para un año, en esencia, es calculado con una base en un periodo que comprende los 2 años anteriores considerando las citas que ha obtenido la revista en ese periodo (llámese citas) en revistas de la misma lista, y las publicaciones que ha realizado en ese mismo periodo (llámese publicaciones). Así, por ejemplo, el factor de impacto en el año 2019 para una determinada publicación puede calcularse como sigue:

$$IF2019_{JCR} = \frac{\text{(Citas del 2019 para trabajos de la revista publicados en 2018 y 2017)}}{\text{(Total de trabajos citables publicados en 2018 y 2017)}} \quad (1)$$

El SJR, por su parte, es calculado de manera similar a la mostrada en la Ecuación 1, pero con las siguientes diferencias:

a) Realiza el estudio considerando los escritos de los 3 años anteriores (y no 2 como el JCR).
b) Incluye todo tipo de escritos (no únicamente artículos como en el JCR).
c) SJR da más valor a las revistas que tienen un alto prestigio (gran cantidad de citas, sin autocitas) utilizando para el cálculo el algoritmo PageRank de Google (descrito en la Ref. [2]).

Merece destacar que los valores de los indicadores IF de una revista deben ser analizados en el contexto del área de estudio, debido a que los valores de IF obtenidos por revistas de distintos campos varían ostensiblemente. Así, tanto el JCR como el SJR cuentan con una serie de áreas y subáreas en las que se posicionan las revistas según su índice de impacto. En este caso, la Revista DYNA se encuentra en la categoría y sub-categoría "Ingeniería, multidisciplinar" (Engineering, Muldisciplinary) del JCR: como se puede observar en la Fig. 1 (resaltado en rojo). Dicha categoría cuenta en la actualidad con 96 revistas, entre las cuales DYNA se posiciona en la septuagésima sexta posición.

La Fig. 1 también ofrece información sobre el cuartil en el que se encuentra la revista (Q4, en este caso). Tanto el JCR como el SJR tabulan la relevancia considerándola de mayor o menor impacto si se encuentran en el primero (Q1), segundo (Q2), tercero (Q3) o cuarto (Q4). Como se ha comentado anteriormente DYNA se encuentra en el Q4 de las revistas de 'Ingeniería-multidisciplinar' del JCR, mientras que se ubica en el Q3 de 'Ingeniería – varios' (*Engineering – miscellaneous*) del SJR.

La Revista DYNA, así como cualquier publicación científica, desea la mejora del impacto de su revista y su posicionamiento en cuartiles superiores. En esta línea, DYNA ha considerado como objetivo principal disponer de una herramienta que permita, en cierta medida, predecir el factor de impacto que puede llegar a tener un artículo durante el proceso de evaluación. Para ello, se ha considerado necesario implantar un sistema de análisis

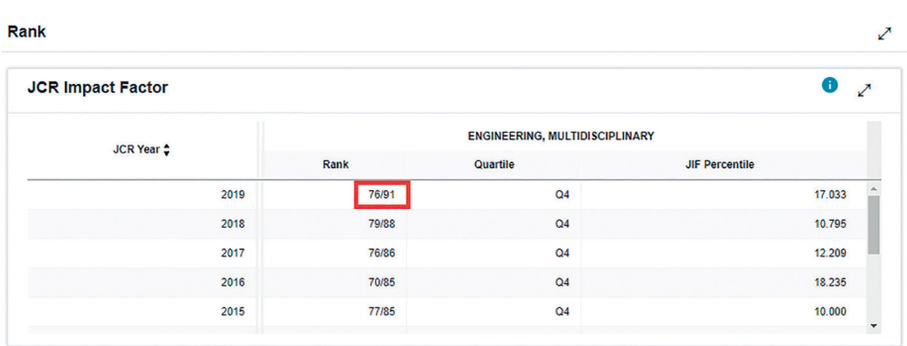


Fig. 1: Evolución de la posición de la Revista DYNA en el campo 'Ingeniería-multidisciplinar' del JCR (adaptado de: https://jcr.incites.thomsonreuters.com, consultado el 30/06/20)

de la información que sirva para detectar qué áreas, autores, revisores, etc. obtienen más citaciones y, por consecuencia, mejoran más significativamente el impacto de la misma. Para ello, y basándose en el trabajo presentado en la Ref. [3] se optó por implementar un sistema BI, siendo tal tipo de sistemas definidos por Shukla y Dhir [4] como el conjunto de técnicas y herramientas para la transformación de grandes cantidades de datos sin procesar en información visual coherente y útil para el análisis de los negocios en cuestión de segundos. Según las Refs. [5][6], los sistemas BI tienen "la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios".

Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, podemos definir BI como el conjunto de metodologías y aplicaciones y tecnológicas que permiten reunir, depurar y transformar datos de los sistemas transaccionales y de la información desestructurada (interna o externa de la compañía) en información estructurada, para su explotación directa (reportes, análisis, alertas...) o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio.

La inteligencia de negocio actúa como un factor estratégico para una empresa u organización, generando una potencial ventaja competitiva, que no es otra que proporcionar información privilegiada para responder a los problemas de negocio: entrada a nuevos mercados, promociones u ofertas de productos, eliminación de islas de información, control financiero, optimización de costes, planificación de la producción, análisis de perfiles de clientes, rentabilidad de un producto concreto, etc.

En definitiva, una solución BI completa permite:

- Observar ¿qué está ocurriendo?
- Comprender ¿por qué ocurre?
- Predecir ¿qué ocurriría?
- Estudiar ¿qué debería hacer el equipo?
- Decidir ¿qué acciones se deben tomar?

Otro de los aspectos a destacar en este trabajo es que el mismo fue realizado por 3 estudiantes de ingeniería que desarrollaron distintas soluciones para el mismo problema utilizando 3 herramientas BI distintas. Así, el trabajo sirvió, en estos casos, para desarrollar la competencia I4.0 identificada por la Ref [7] "Conocimiento y habilidades en tecnologías para Business Intelligence y Big Data".

Partiendo del siguiente enmarque, el objetivo de la presente investigación es mostrar el caso de aplicación del sistema BI en la revista DYNA. Para ello, se parte de exponer cómo se muestra la información en los visualizadores de los sistemas BI de ISI y SCIMAGO como estado del arte (apartado 2) para a continuación exponer los materiales y métodos implementados en el apartado 3. Finalmente, los apartados 4 y 5 exponen respectivamente los resultados cuantitativos y cualitativos obtenidos gracias a la implantación del BI y las conclusiones que se extraen de la investigación presentada.

2. ESTADO DEL ARTE

Tanto ISI para el JCR como SCIMAGO para el SJR cuentan con cubos BI diseñados para que los interesados puedan analizar la información que consideren necesaria. Mediante ellos es posible obtener infinidad de información como por ejemplo y entre otros, la comparación de trayectorias de revistas o analizar el nivel de auto-citado (citado de artículos recientes de la misma revista en la que se ha publicado el artículo) de las mismas. Por ejemplo, la Fig. 2 muestra como ejemplo una captura de pantalla de una de las funcionalidades del cubo de ISI en la que se puede analizar desde qué revistas se cita a DYNA.

Pese a que los cubos de ISI y SCIMAGO disponen de infinidad de funcionalidades, no cuentan con datos que ayuden a obtener información del funcionamiento interno de las revistas. Por ejemplo, no disponen de quiénes han sido los revisores de las mismas, ni del nivel o nota que han otorgado los mismos a un artículo cuando lo han revisado. Ante esta carencia, parece que algunas editoriales han optado por desarrollar sus propios sistemas BI. Una editorial que parece cuenta con un sistema de este tipo es la "Multidisciplinary Digital Publishing Institute" (https://www.mdpi.com/). Merece destacar que esta editorial cuenta con un sistema de sugerencias de artículos para autores que envían sus trabajos. Concretamente, independientemente de los resultados de las revisiones de los artículos, la página web de la revista devuelve al autor una lista de potenciales artículos de la editorial que podrían encajar en el escrito que ha mandado a revisar.

Ante el potencial de la herramienta se concluyó que era necesario integrar la información referente a citaciones disponible en las listas de impacto con la información de gestión interna de los artículos. Así, se creyó que, a partir de integrar la información de los artículos citados, así como cuántas veces y por quién habían sido citados, junto con la información de funcionamiento interno en un BI se lograría extraer información interesante no obtenida hasta la actualidad.

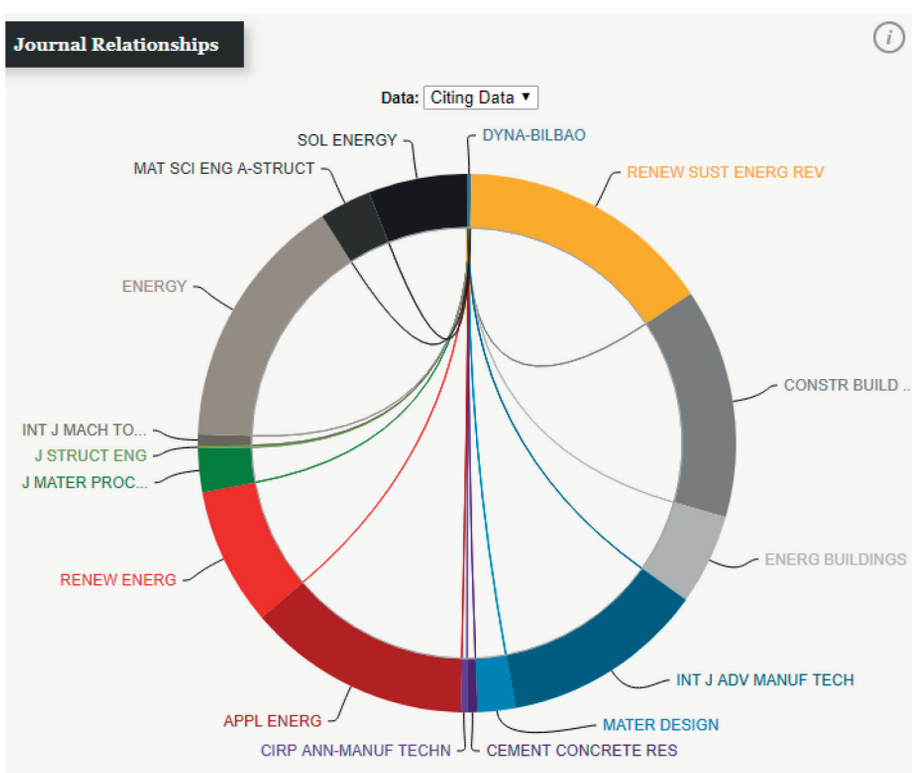


Fig. 2: Ejemplo de funcionalidades del cubo BI de ISI, con visualización de las revistas del JCR que citan a DYNA (fuente: https://jcr.incites.thomsonreuters.com, consultado el 10/02/2020)

3. MATERIALES Y MÉTODOS

Los métodos y materiales utilizados en el caso son respectivamente presentados en las secciones 3.1 y 3.2, para exponer en la sección 3.3 en un ejemplo cómo han sido combinados estos métodos y materiales para obtener, entre otras, las funcionalidades que se muestran en la sección 3.4.

3.1. MÉTODO

El método seguido en la investigación fue el Real Case Solving (RCS) o Resolución de Casos Reales [8]. El RCS es una variante de la metodología educativa de Aprendizaje Basado en Proyectos o Problem Based Learnig (PBL), en la que estudiantes resuelven casos reales. Concretamente, el enfoque tradicional de PBL se complementa con sesiones puntuales con empleados de empresas que vienen a clase para buscar posibles soluciones a un problema real [9]. Por su parte, el PBL [10] es un método ampliamente utilizado en el mundo de la educación, en el que uno de los principales objetivos es mejorar las habilidades de "aprendizaje autodirigido", al modificar el papel conferenciante habitual de los profesores hacia un papel de mentor. El PBL comprende básicamente las fases de a) la identificación del problema a resolver, b) la identificación (por sí mismos) de lo que los estudiantes deben aprender, y c) aprender y aplicar el conocimiento aprendido para resolver el problema. Así, el esquema resumen del RCS se muestra en la Fig. 3.

El principal atractivo inherente del RCS [8] es que la realidad del caso y la presencia de personal de la empresa como apoyo motiva a los estudiantes más que otras metodologías. La presente implementación de un sistema BI se realizó dentro de un contexto educacional, pues fue uno de los proyectos de implementación de BI para empresas reales ofrecidos a una serie de alumnos. Éstos tuvieron que desarrollar prototipos de soluciones a empresas de sectores diversos como el de la alimentación, el portuario o el de las publicaciones científicas. En cada uno de estos casos cada alumno dispuso de un set de datos con el que desarrollar la herramienta junto con un conjunto de especificaciones marcadas por el cliente, el cual revisó en hitos puntuales la trayectoria del proyecto para mejorar el ajuste de la solución que ofreciese la solución al problema real.

3.2. MATERIALES

La Revista DYNA marcó una tabla de requerimientos técnicos a la hora de encaminar la selección del tipo de producto BI. En concreto, la solución debía:

- Ser fácil de emplear y de aprender, apoyando la forma de acercarse a los datos que se pretenden manipular y obteniendo las relaciones asociativas pertinentes para crear el modelo que ayudará a tomar la decisión pertinente.
- Ser una solución gratuita.
- Posibilitar el establecimiento de conexiones a orígenes muy distintos entre sí para poder recoger las bases de datos de servidores como BBDD Access o ODBC y así poder conectarse a cualquier servidor tanto de forma offline como online. De este modo, se debía poder obtener información tanto de ficheros del tipo SQL ubicados en servidores como otro tipo de ficheros planos que se encuentren en otras ubicaciones (.xls, csv u otro tipo de ficheros de texto).
- Contar con un lenguaje de programación conocido en caso de que sea necesario alterar el código de la solución desarrollada en un futuro.
- Presentar de forma muy visual los cuadros de mando, es decir, el sistema que recoge los principales indicadores a analizar gracias a los distintos objetos de hoja como: tablas, gráficos, contenedores y menús de filtros.
- Personalizar el cuadro de mando de la forma que se desee, dando al usuario libertad total a la hora de gestionar la información que desea analizar.
- Proveer de respuestas de negocio inmediatas y una rápida valoración por medio de informes que se pueden tanto exportar a ficheros excel en formatos de tablas o gráficas como en otro tipo de documentos impresos.
- Visualizar desde diferentes dispositivos online como offline, dotando al usuario de una alta independencia a la hora de utilizar la herramienta.

El equipo de proyecto seleccionó un software BI que se ajustase a los requerimientos marcados por la revista y, tras validar la elección de la herramienta con la editorial, procedió al desarrollo de la solución.

3.3. DESARROLLO DE LA SOLUCIÓN

En el desarrollo de la solución se combinaron principalmente dos fuentes de datos: por una parte, se dispone de una Base de Datos de gestión de artículos de la revista la cual cuenta con tablas específicas que guardan información referente a autores, artículos, disciplinas y subdisciplinas, quiénes han revisado el artículo, valoración media de su trabajo, puntuación otorgada al artículo en originalidad, utilidad, claridad, repercusión, conclusiones y nivel de madurez tecnológica (TRL). Por otra parte, se cuenta con una tabla con todos los artículos de DYNA que habían sido citados en el JCR (e información adicional que incluye, por ejemplo, quién lo ha citado, en qué revista, cuándo, etc.). Fruto de la combinación de ambos orígenes de datos fue posible desarrollar modelos relacionales o no relacionales como el mostrado como ejemplo en la Fig. 4.

A partir de los modelos desarrollados para la solución ha sido posible obtener la serie de funcionalidades descrita en el siguiente apartado.

3.4. FUNCIONALIDADES OBTENIDAS

Como se muestra en las siguientes Figs., el desarrollo realizado permite elaborar distintos tipos de análisis de datos. Por ejemplo, la Fig. 5 muestra una visua-

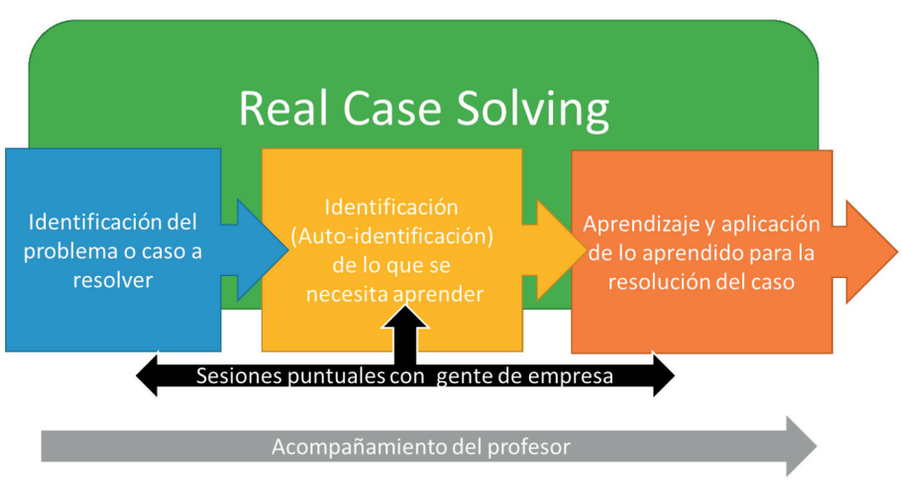


Fig. 3: Esquema descriptivo de la metodología RCS

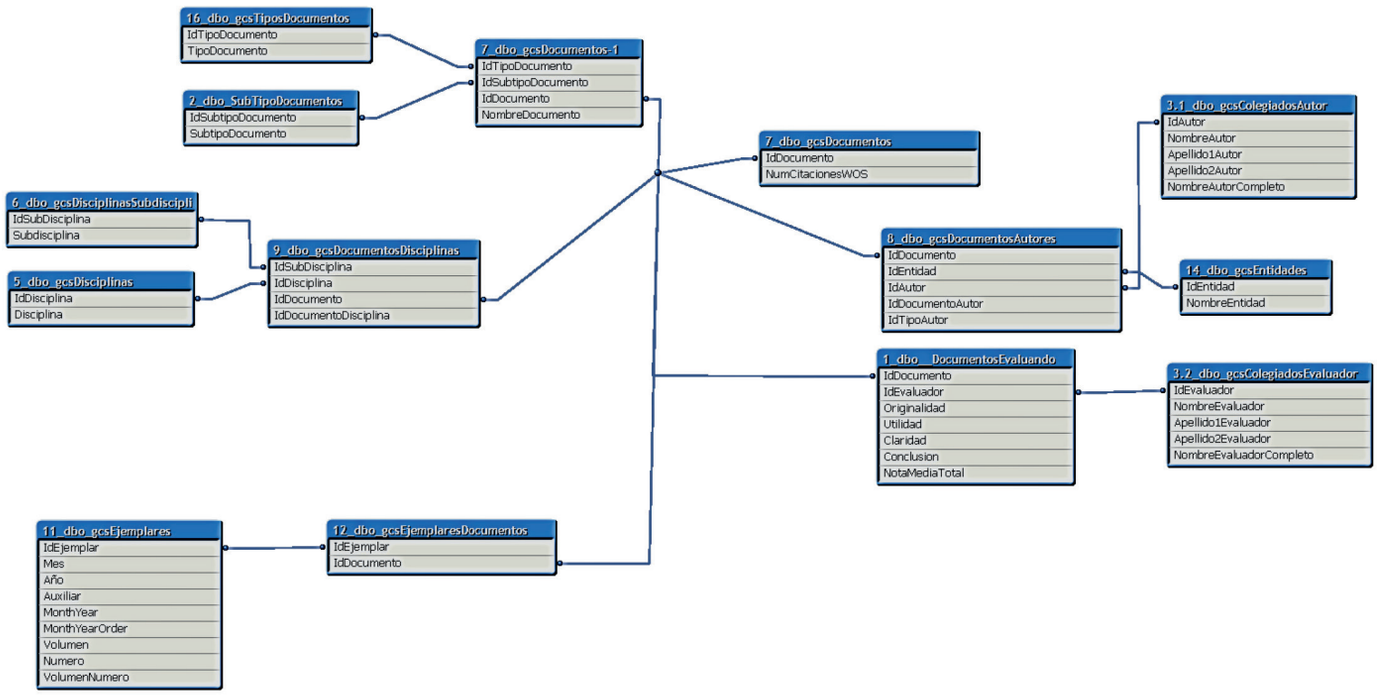


Fig. 4: Ejemplo de modelo relacional del BI de DYNA

lización de las disciplinas y subdisciplinas que más citaciones JCR han recibido, en la que una mayor superficie indica una mayor cantidad de citaciones recibidas.

Como se aprecia en la Fig. 6, la superficie de cada color es identificativo del impacto de cada disciplina de la revista, y asimismo al área cubierta por cada recuadro de un mismo color representa la aportación al impacto de cada sub-disciplina. Este análisis visual también puede ser realizado en modo tabla, para una más sencilla identificación del impacto de todos estos campos, como se muestra en la siguiente captura del BI.

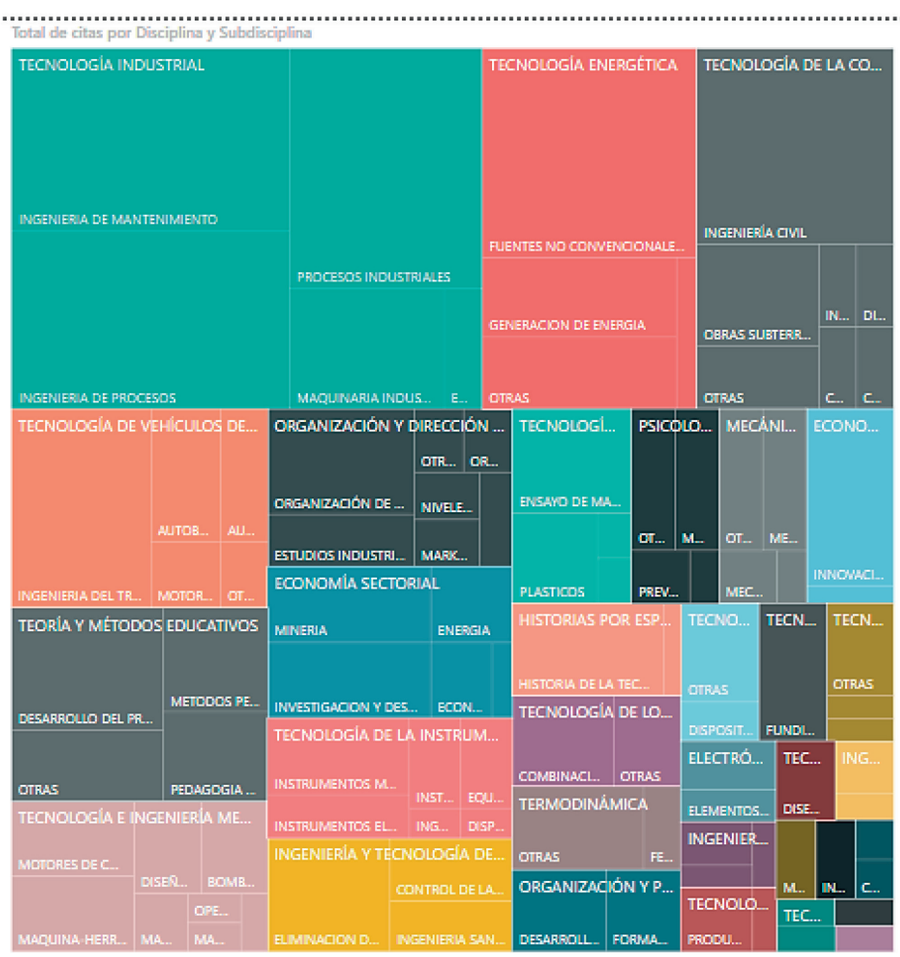


Fig. 5: Visualización del impacto JCR obtenido por cada una de las disciplinas y subdisciplinas de la Revista DYNA

El modelo desarrollado también sirve para identificar fácilmente los autores más citados, así como cuáles son los parámetros de valoración de la evaluación que más correlación tienen con el IF, cómo pueden predecir los evaluadores el resultado del artículo, disciplinas publicadas con más repercusión en el IF, etc. La Fig. 7 expone, como ejemplo, una versión anonimizada de los 10 autores y evaluadores cuyos artículos redactados o revisados han conseguido más citas.

4. RESULTADOS

Como se ha podido observar en el apartado anterior, se han desarrollado infinidad de paneles de visualización que han servido para detectar, entre otros: a) las áreas de conocimiento que son más o menos susceptibles a recibir más citas, b) los autores que son más o menos susceptibles a recibir más citas, c) los revisores que evalúan favorablemente un artículo que a posteriori recibe más citas. Valiosa información que ayuda a la revista, a predecir el factor de impacto que puede llegar a tener un artículo durante el proceso de evaluación. Asimismo, el BI sirve para identificar visualmente disciplinas y subdisciplinas de mayor o menor impacto; a este respecto merece destacar, por ejemplo, que la disciplina 'Organización y Dirección de Empresas' sea la quinta disciplina que más citas genere, pese a que la

IdDisciplina	Subdisciplina	Total de citas
3310	INGENIERIA DE MANTENIMIENTO	33
3310	INGENIERIA DE PROCESOS	32
3310	PROCESOS INDUSTRIALES	30
3322	FUENTES NO CONVENCIONALES DE ENERGIA	29
3305	INGENIERÍA CIVIL	25
3317	INGENIERIA DEL TRAFICO	18
3327	SISTEMAS DE TRANSITO URBANO	15
3305	TECNOLOGÍA DEL HORMIGÓN	14
5801	DESARROLLO DEL PROGRAMA DE ESTUDIOS	12
3310	MAQUINARIA INDUSTRIAL	12
3305	TRANSMISION DE CALOR EN LA EDIFICACION	12
3322	GENERACION DE ENERGIA	10
5306	INNOVACION TECNOLOGICA	10
5311	ORGANIZACIÓN DE LA PRODUCCIÓN	10
3312	PROPIEDADES DE LOS MATERIALES	10
3308	ELIMINACION DE RESIDUOS	9
5506	HISTORIA DE LA TECNOLOGIA	9
3322	OTRAS	9
3312	ENSAYO DE MATERIALES	8
5312	INVESTIGACION Y DESARROLLO	8
5312	MINERIA	8
3305	OBRAS SUBTERRANEAS	8
3306	TRANSMISION Y DISTRIBUCION	8
2209	TRATAMIENTO DIGITAL DE IMÁGENES	8
3311	INSTRUMENTOS MEDICOS	7
5801	METODOS PEDAGOGICOS	7
2213	OTRAS	7
5801	OTRAS	7
Total		**680**

Fig. 6: Captura de pantalla del BI de la Revista DYNA en la que se expone la tabla que muestra las citas logradas por cada subdisciplina, en orden descendente

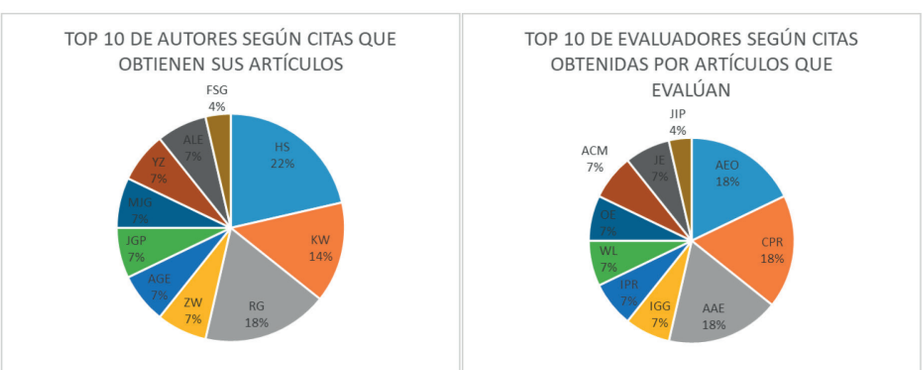


Fig. 7: Visualización de los 10 autores y 10 evaluadores cuyos artículos han obtenido más citas (anonimizado)

editorial cuente con otra revista específica para artículos relacionados con este ámbito (Dyna Management).

5. DISCUSIÓN / CONCLUSIONES

El presente artículo presenta un caso de implantación de un sistema BI en la revista DYNA que ha tenido como objetivo disponer de una herramienta que le ayude a mejorar su factor de impacto seleccionando más adecuadamente los trabajos a aceptar y rechazar. Concretamente, el trabajo realizado ha servido para identificar, entre otros, temáticas con mayor aceptación, revisores cuyos criterios no se encuentran excesivamente alineados con los autores cuyos artículos revisan, etc. en definitiva, para estimar el impacto que pueden llegar a tener todos los parámetros de análisis de los que dispone la revista.

A nivel pedagógico cabe destacar que el caso desarrollado, junto con otros casos reales, ha servido para aplicar un proceso PBL al diseñar soluciones BI de casos reales en los grados de Ingeniería en las clases que los autores imparten.

En cuanto a las líneas futuras, éstas se pueden diferenciar en líneas a corto y largo plazo. Las acciones a corto plazo se enfocan en desarrollar otros dos cubos BI para la Revista DYNA, uno para el análisis de los subscriptores (de cara a maximizar los ingresos vía suscripciones) y otro para realizar el mismo trabajo pero en el área de la publicidad. A medio largo plazo se planifica la implementación de una cantidad mayor de características predictivas. Aunque algunas de ellas se utilizan en la actualidad, no muestran ningún tipo de correlación con el IF (p.ej. la utilidad y claridad del artículo) y otras no han sido aún consideradas (p.ej. se trata o no de un asunto polémico, la temática es o no nuclear de DYNA, etc.).

REFERENCIAS

[1] M. Bordons, M. Fernández, and I. Gómez, "Advantages and limitations in the use of impact factor measures for the assessment of research performance," Scientometrics, vol. 53, no. 2, pp. 195–206, 2002.
[2] I. Rogers, "The Google Pagerank Algorithm and How It Works," Http://Iprcom.Com/Papers/Pagerank, 2005.
[3] A. Goti-Elordi, A. De-La-Calle-Vicente, M. J. Gil-Larrea, A. Errasti-Opakua, and J. Uradnicek, "Aplicación de un sistema business intelligence en un contexto big data de una empresa industrial alimentaria," Dyna, vol. 92, no. 3, pp. 347–353, 2017. DOI: http://dx.doi.org/10.6036/8008
[4] A. Shukla and S. Dhir, "Tools for Data Visualization in Business Intelligence: Case Study Using the Tool Qlikview," 2016, pp. 319–326.
[5] Ibis-Computer, "Business Intelligence." [Online]. Available: https://www.ibiscomputer.com/software-erp/accesorios/business-intelligence. [Accessed: 27-Feb-2020].
[6] Gisa Consultores, "GISA Consultores – Business Intelligence." [Online]. Available: http://www.gisaconsultores.com/business-intelligence/. [Accessed: 27-Feb-2020].
[7] J. Horrillo Tello and J. Triado Aymerich, "CARENCIAS FORMATIVAS DE LOS GRADOS DE INGENIERÍA PARA LA INDUSTRIA 4.0 EN ESPAÑA. UNA PROPUESTA DE ACTUACIONES," DYNA, vol. 93, pp. 365–369, 2018.
[8] S. B. Caldwell and T. D. Gedeon, "Boldly Going where no Higher Educators Have Gone Before: A Review of the 1st International Conference on Higher Education Advances, Valencia, Spain," Procedia - Soc. Behav. Sci., vol. 228, pp. 348–355, Jul. 2016.
[9] J. Larmer, Mergendoller J., and Boss S, Setting the Standard for Project Based Learning: A Proven Approach to Rigorous Classroom Instruction. VA USA: ASCD Alexandria, 2013.
[10] M. Zubizarreta, "Innovation of the teaching-learning process in the degree of technical engineering in electronics: a case study," Bilbao, 2006.

AGRADECIMIENTOS

Nuestro más sincero agradecimiento a José María Hernández Álava, Director y editor de la Revista DYNA por el apoyo en el desarrollo de la herramienta.
Nuestro agradecimiento también para la Fundación BBK, Idom, Accenture, la Fundación Telefónica y el Grupo Inzu, Patronos de la Cátedra Deusto en Industria Digital.