

# **Plataforma Colaborativa en Escenarios Educativos para el Aprendizaje Colectivo mediante KDD**

**Jhon Haide Cano Beltrán**

Universidad Cooperativa de Colombia, Cali, Valle, Colombia, jhon.canob@campusucc.edu.co

## **RESUMEN**

La inteligencia colectiva ha permitido resolver más problemas a nivel de grupo que si lo hicieran de manera individual, ésta afirmación conlleva a que los estudiantes y profesores dejan de ser actores pasivos para convertirse en actores activos, las técnicas de filtro colaborativo permiten inferir recomendaciones y estrategias para la generación de conocimiento colectivo. El objetivo principal es analizar mediante minería de datos, todas las interacciones en la plataforma que conduzcan a un verdadero aprendizaje colectivo. Este conocimiento colectivo tiene dos premisas, la primera es que el conocimiento colectivo se da entre dos o más personas y segundo que exista interacciones entre ellas, aunado a estas dos premisas la web provee un espacio fascinante que cumplen con las condiciones anteriores y además genera espacios de participación y colaboración, lo que posibilita la construcción de redes de tipo colaborativo, de esta manera se tiene la red como plataforma y el software como un servicio, los sistemas de recomendación y la minería de datos permitieron guiar a los usuarios en la obtención de información de allí que se puedan obtener patrones y recomendaciones sobre un tema específico.

**Palabras Clave:** Plataforma Colaborativa, Inteligencia Colectiva, KDD, Motor de Sugerencias, Educación.

## **ABSTRACT**

Collective intelligence has solved more problems at group level than if they were individually, this statement entails that students and teachers from passive to become active actors, collaborative filtering techniques allow inferring recommendations and strategies for the generation of collective knowledge. The main objective is to analyze by data mining, all interactions on the platform leading to a true collective learning. This collective knowledge has two premises, the first is that the collective knowledge exists between two or more people and second that there is interaction between them, these two premises together with the web provides a fascinating space that meet the above conditions and generates spaces participation and collaboration , enabling the construction of networks of collaborative type , so there is the network as platform and software as a service, recommendation systems and data mining allowed to guide users in obtaining information there that they can get patterns and recommendations on a specific topic.

**Keywords:** Platform Collaborative Collective Intelligence, KDD, Motor Suggestions, Education.

## **1. INTRODUCCION**

Las entidades educativas en Colombia, enfrentan un reto frente a la sociedad en red (Castells, 2001), y en especial con las Tics (Tecnologías de la Información y la Comunicación), crear y organizar mediante inteligencia colectiva grupos con actividades de aprendizaje similar, propende para que la información y el conocimiento se conviertan en elementos generadores de aprendizaje, de esta manera proveer una plataforma colaborativa en el que los estudiantes y profesores puedan obtener sugerencias dependiendo de su área de interés y que de esas sugerencias puedan tomar decisiones se convertiría en una plataforma de colaboración útil en los escenarios educativos.

El Centro para la Inteligencia Colectiva (CIC) en cabeza de su director Tom Malone del MIT (Instituto Tecnológico de Massachusetts) define lo colectivo como un grupo de actores individuales, incluyendo por ejemplo personas, agentes computacionales y organizaciones; y por inteligencia al comportamiento colectivo del grupo que presenta características tales como la percepción, el aprendizaje, el juicio o la resolución de problemas. Actualmente el CIC adelanta investigaciones orientadas a las organizaciones entre ellas la colaboración distribuida, teniendo como variables las culturas, estructuras y procesos organizacionales.

Las instituciones educativas actualmente se han convertido en escenarios de carácter colaborativo, es decir, las individualidades son propensas a crear limitantes o barreras mentales, por ésta razón han dejado de ser relevantes para los entornos educativos y poco a poco esas individualidades se han estado convirtiendo en grupos de trabajo colaborativo, estos modelos colaborativos aplican entre muchos otros a entornos educativos, permitiéndole crear espacios visibles en donde los estudiantes y profesores trabajen conjuntamente y de manera colaborativa.

## **2. FUNDAMENTO TEÓRICO**

El advenimiento de la Web 2.0 privilegia las plataformas colaborativas como un nuevo enfoque de participación colectiva en la que los usuarios y la información en sus diferentes formatos, se presentan como una nueva forma de conocimiento dirigido y centrado en los temas que atañen a cada comunidad en particular, privilegiando las ontologías que son la conceptualización formal de un dominio en particular que sea de interés para un grupo de trabajo.

La Inteligencia Colectiva permite que la colaboración entre individuos se destaque muy por encima de la individualidad, las organizaciones intelectuales como los grupos de investigación y centros de investigación de desarrollo tecnológico necesitan plataformas de colaboración que le permita procesar información de diferentes fuentes y de esa forma trabajar en temáticas afines con su objeto de estudio.

Ahora bien, la inteligencia colectiva posibilita la construcción de redes de tipo colaborativo haciendo que la Web se comporte como una plataforma y el software como un servicio, esto le permitirá a estudiantes y profesores o a las organizaciones intelectuales crear relaciones de trabajo e investigación, propiciando que esos grupos humanos puedan constituirse en colectivos inteligentes y así mantenerse altamente competitivos en su quehacer académico y empresarial, aunado a este trabajo colaborativo la minería de datos se convierte en un agente catalizador en la búsqueda y conceptualización de la información referente a un tema de estudio por parte de los estudiantes y profesores.

“Las personas han utilizado la frase Inteligencia colectiva por décadas, y se ha convertido cada vez más popular y más importante con el advenimiento de las nuevas tecnologías de la comunicación.” (Segaran, 2007). La inteligencia colectiva existía mucho antes que la aparición del internet, pero ahora la habilidad de recolectar información de miles de millones de personas en la Web ha abierto nuevas posibilidades en este campo.

“El ideal de la inteligencia colectiva implica la valoración técnica, económica, jurídica y humana de una inteligencia repartida en todas partes con el fin de desencadenar una dinámica positiva del reconocimiento y de la movilización de las competencias.” (Levy, 2004), es por ésta razón que “Una nueva generación de tecnologías tiene como objetivo incursionar en las organizaciones” (Grasso, 2012), de allí que el internet juega un papel muy importante en la consecución de las metas propuestas de una organización, pero no solo el internet como red de redes, es importante hacer mención de la web 2.0 en la que se soporta la inteligencia colectiva, las instituciones de educación hacen evidente que las relaciones académicas en un ámbito descentralizado ayuda al intercambio de conocimiento y hacen colectivas las decisiones, es decir, se está empezando a crear espacios en el que la misma organización se vuelve un consumidor de su información.

Pero qué hacer con toda esa información relevante e importante para una institución educativa, es allí en donde la inteligencia colectiva se convierte en pieza clave, entre sus múltiples aplicaciones se encuentran las recomendaciones, grupos de aprendizaje, búsqueda y posicionamiento, optimización, filtros, modelamiento y decisiones. Los modelos estadísticos casi siempre están en el campo de las predicciones y es más acertado, que personas expertas en un campo de acción determinado, pero si las personas tienen la habilidad de obtener los datos y organizarlos, de allí la pregunta ¿Cómo se pueden hacer predicciones en acciones o comportamientos en sistemas sociales complejos? (Nagar, 2011), expuesta la pregunta es importante decir que la inteligencia colectiva en sus aplicaciones aborda este tipo de problemáticas, las sociedades al igual que las organizaciones son estructuras complejas, que necesitan modelos computacionales y personas que interactúen con esos modelos, ahora bien, los sistemas de recomendación tienen dos aspectos a tratar el primer aspecto tiene que ver con la generación de posibles soluciones y el segundo aspecto la evaluación de esas posibles soluciones, lo que hace supremamente complejo la toma de decisiones, actualmente los sistemas autogestionados toman mayor campo de acción en las industrias, ya que estos sistemas pueden resolver situaciones en un menor tiempo que si lo hiciera un humano experto.

Para la toma de decisiones se puede hablar de tres enfoques: Los Alcances, Valor Agregado y Auto-organización (Bonabeau, 2011), los sistemas de información en su más simple estructura siempre están apuntando a la evaluación de alternativas, obtener la más adecuada y de una u otra manera auto-organizarse, en otras palabras esto no es más que el ciclo básico de un sistema (Entrada – Proceso - Salida).

Las TICs (Tecnologías de la Información y la Comunicación ) permiten representar conocimiento de tal manera que los estudiantes y profesores puedan colaborar con otros actores, consultar expertos, compartir experiencias de aprendizaje y resolver problemas complejos, estos elementos cambian los modos de aprendizaje y la manera en que las personas procesan la información, lo que se podría denominar ecosistemas digitales en los que se intercambia información teniendo en cuenta el contexto en un aprendizaje conectado.

Existen tres tipos de inteligencia en la web (Alag, 2009), la inteligencia explícita, en la que un amigo o una persona hace una recomendación u opiniones y ésta puede llegar a tener un alto impacto en el mismo grupo o en otros grupos, inteligencia implícita, la información que se obtiene de los usuarios proviene de fuentes diversas, muchas veces la información no se encuentra estructurada (foros, mensajes, blogs), el último tipo de inteligencia en la web es la inteligencia derivada, las aplicaciones que implementan este tipo de inteligencia, por lo general utilizan minería de datos, motores de recomendación, búsquedas inteligentes y análisis predictivo para obtener información, por lo general son algoritmos complejos.

El Instituto Tecnológico de Massachusetts (MIT) cuenta con un centro para la inteligencia colectiva<sup>1</sup>, este centro tiene varios proyectos de investigación entre los que se encuentra el estudio de la inteligencia colectiva en las organizaciones hoy en día, y se trabaja en temas como la colaboración distribuida, organizaciones sensibles, redes colaborativas de innovación y tendencias a través de análisis online social media.

Haciendo un recorrido por diferentes aplicaciones surgidas en este tema se podría mencionar PHOAKS (Teryeen, 1997) (People Helping One Another Know Stuff), sistema experimental para compartir recomendaciones, trabajaba reconociendo, relacionando e intercambiando recomendaciones de recursos web extraídos de mensajes de grupos de noticias, Fab (Balabanovic, 1997) que es un sistema híbrido, basado en contenido y filtrado colaborativo, se basa en determinar similitudes entre usuarios y haciendo un cruce de los perfiles realiza una recomendación colaborativa y ReferralWeb (Kautz, 1997) otro sistema de recomendación interactivo para la construcción, visualización y búsqueda de redes sociales, se centraba en la concurrencia de nombres en páginas web utilizando un buscador web. No se intentaba crear nuevas comunidades sino hacer un uso más eficiente de las redes existentes.

---

<sup>1</sup> Disponible en: <http://cci.mit.edu/>

Las computadoras actualmente tienden cada vez más a dejar de ser dispositivos aislados y se están convirtiendo en puntos de red para intercambio de información y transacciones comerciales, lo que permite una interacción entre los datos, la información y el conocimiento que son claves en la consecución de información para un dominio establecido.

Las ontologías precisamente permiten tener una comprensión compartida y un dominio en común entre las personas y los sistemas de aplicación, basándose en estándares actuales de la web, por esta razón las ontologías permiten apoyar procesos de intercambio de información en diferentes áreas del conocimiento, hoy en día se tienen a las ontologías como tema de investigación en ingeniería del conocimiento, procesamiento de lenguaje natural, inteligencia artificial y representación del conocimiento.

Un modelo formal de datos basado en tripletas como lo es RDF (Resource Description Framework), objetos, propiedades y valor, fue desarrollado por el World Wide Web Consortium (W3C), y su creación se debe a la gran cantidad de información que se puede encontrar en internet, es por ésta razón que los motores de búsqueda deban apuntar de manera perentoria la forma en que se recuperan la estructura y la semántica de ese conjunto de elementos (Información), adicionalmente la W3C ha propuesto nuevas tecnologías que facilitan la incorporación de semántica a las páginas Web con el fin de imponer la necesidad de una restricción estructural para proporcionar métodos inequívocos de expresión semántica, las grandes empresas de software buscan desarrollar e implementar ésta infraestructura de descripción de recursos, dado que hoy en día la red está catalogada como la plataforma de los sistemas de información.

Las redes sociales abren aún más el espectro del verdadero significado de la representación del conocimiento en la web semántica, el análisis de las redes tiene una visión global sobre las estructuras sociales en la creencia de que los tipos y patrones de relaciones surgen de la conectividad individual, en este sentido el estudio de la red se centra en la interdependencia de la comunidad, de allí que se desprenda el descubrimiento de patrones, tendencias y regularidades para poder determinar ciertas características en las variables asociadas y comprender mejor los datos de las estructuras sociales.

La minería de datos en la web permite el descubrimiento de recursos y extracción de la información en los que se encuentran todo tipo de documentos en la web y consulta a bases de datos, para luego descubrir y validar patrones que permitan un acercamiento al objeto de estudio, en los que se puede encontrar hábitos, preferencias de los usuarios o contenidos a partir de las sesiones que crean los usuarios, esto presenta un enfoque diferente a los fenómenos sociales, lo que exige un nuevo conjunto de conceptos y nuevos métodos para la recolección y análisis de datos, proporcionando modelos formales en la captura de las propiedades comunes para la plataforma colaborativa.

Aunado a lo anterior la inteligencia colectiva ha permitido resolver más problemas a nivel de grupo que si lo hicieran de manera individual, esta afirmación conlleva a que los usuarios dejan de ser actores pasivos para convertirse en actores activos de los cuales se puede obtener información relevante y de interés para grupos especializados, de esta manera las técnicas de filtro colaborativo permiten inferir recomendaciones y estrategias que permitan la generación de conocimiento colectivo, vale la pena resaltar que el ambiente de un entorno colaborativo pasa indudablemente por las interacciones entre los actores participantes dentro de la plataforma de colaboración, por ende, se podría afirmar que se está evolucionando a un aprendizaje conectado.

Este conocimiento colectivo tiene dos premisas importantes, la primera es que el conocimiento colectivo se da entre dos o más personas y segundo que exista interacciones entre ellas, aunado a estas dos premisas la web provee un espacio fascinante que cumplen con las condiciones anteriores y además genera espacios de participación y colaboración, no solo de personas sino también de recursos tales como servicios web, peer to peer, redes sociales entre otros, lo que posibilita la construcción de redes de tipo colaborativo, vínculos internos y vínculos externos, de esta manera se tiene la red como plataforma y el software como un servicio, Toby Segaran en su libro Programming Collective Intelligence aclara que “los precios no se establecen por un individuo sino por

el comportamiento comercial de muchos independientes que actúan y crecen por su propio interés”, de allí que se puedan obtener patrones y recomendaciones sobre un tema específico aplicando la inteligencia colectiva.

Las tecnologías semánticas desacoplan las aplicaciones de datos mediante un modelo simple y abstracto para la representación del conocimiento, lo que permite centrarse en el comportamiento de las aplicaciones en lugar del procesamiento de datos, lo que conlleva a comunicar con suficiente significado el resultado de una acción. La esencia de los modelos de datos semánticos son esquemas flexibles en donde las relaciones son descritas por los propios datos (Watson, 2009), de allí que los metadatos representan el conocimiento de cómo los datos pueden ser utilizados. Las ontologías precisamente permiten tener una comprensión compartida y un dominio en común entre las personas y los sistemas de aplicación, basándose en estándares actuales de la web, por esta razón las ontologías permiten apoyar procesos de intercambio de información en diferentes áreas del conocimiento, hoy en día se tienen a las ontologías como tema de investigación en ingeniería del conocimiento, procesamiento de lenguaje natural, inteligencia artificial y representación del conocimiento.

Relacionado con el tema anterior las redes sociales abren aún más el espectro del verdadero significado de la representación del conocimiento en la web semántica (Segaran, 2009), el análisis de las redes tiene una visión global sobre las estructuras sociales en la creencia de que los tipos y patrones de relaciones surgen de la conectividad individual, en este sentido el estudio de la red se centra en la interdependencia de la comunidad, de allí que se desprenda el descubrimiento de patrones, tendencias y regularidades para poder determinar ciertas características en las variables asociadas y comprender mejor los datos de las estructuras sociales. La minería de datos en la web permite el descubrimiento de recursos y extracción de la información en los que se encuentran todo tipo de documentos en la web y consulta a bases de datos, para luego descubrir y validar patrones que permitan un acercamiento al objeto de estudio, en los que se puede encontrar hábitos, preferencias de los usuarios o contenidos a partir de las sesiones que crean los usuarios.

La creación del motor de sugerencias integrado al KDD, logra que las recomendaciones o sugerencias tengan en realidad un significado, es decir, las instituciones educativas se convierten en generadoras de conocimiento en red y la información que se obtiene pasa de ser la agrupación de datos a unas consultas semánticas con la habilidad de sugerir o recomendar actividades o estrategias en los campos de acción de cada estudiante o profesor.

### **3. METODOLOGÍA**

La plataforma colaborativa tiene como objetivo principal proveer un espacio en la web en la que los usuarios puedan generar ese ecosistema académico ideal para centrarse en un tema de interés particular o grupal, y de esa manera recibir sugerencias de los demás usuarios que han accedido a la plataforma.

Inicialmente las interacciones generadas por los usuarios en la plataforma son identificadas y organizadas, permitiendo su posterior análisis mediante el algoritmo de vecinos cercanos en una estructura tipo árbol llamada Kd-Tree. Una vez se ejecuta el algoritmo el usuario visualiza mediante su navegador que usuarios están trabajando en la misma área de interés (Sistema de Recomendación), generando relaciones que propenden en un aprendizaje colaborativo destacando los ecosistemas digitales como fundamento del intercambio de información.

Para la extracción de los datos mediante KDD, se tuvieron en cuenta algunas variables ponderadas por su importancia en la plataforma colaborativa y obviamente por la interacción de los mismos, también se filtran entradas que no aporten al descubrimiento de información. El objetivo principal es encontrar reglas con una alta confianza restringiendo el espacio de soluciones.

### **4. RESULTADOS**

Tomando la pregunta ¿Cómo se pueden hacer predicciones en acciones o comportamientos en sistemas sociales complejos?, durante los inicios de ésta investigación, se hizo mucho énfasis en los algoritmos de filtrado colaborativo, estos algoritmos unos más rápidos que otros presentan sus ventajas y desventajas, es racional

afirmar que el volumen de datos presentes en la plataforma colaborativa es alto y crítico, por ende es necesario implementar un algoritmo que permita obtener información de los usuarios y de las acciones que ejecuta dentro de la plataforma con el fin de ir haciendo sugerencias de los temas a fin o de las contribuciones hechas por otros usuarios; en un primer acercamiento al modelo se optó por un filtro colaborativo basado en el algoritmo de vecinos cercanos con JSON (JavaScript Object Notation), primero porque permite interacción asíncrona y segundo JSON es un formato ligero de intercambio de datos, de esta manera lograr mayor interactividad natural dentro de la plataforma colaborativa.

La plataforma colaborativa utiliza el framework Django el cual permite una organización del proyecto, maneja una estructura limpia para archivos y carpetas, posee componentes de seguridad y demás propiedades de un framework maduro.

El modelo propuesto se destaca por dos lineamientos centrales, el primero actúa sobre consultas y preferencias que hacen los usuarios a un tema en particular y la segunda se hace mediante recomendaciones o sugerencias que se obtienen de otros usuarios que trabajan o investigan en el mismo tema de afinidad, éste modelo simula un ecosistema digital en donde las interacciones de los usuarios crean ese ambiente colaborativo.

## 5. CONCLUSIONES

La inteligencia colectiva ha permeado en gran medida organizaciones que buscan mejorar significativamente procesos realmente complejos a la luz de la tecnología, las instituciones educativas no son ajenas a este tipo problemáticas, constantemente los profesores buscan estrategias para compartir el conocimiento y lograr que los estudiantes lo asimilen como propio. Es por eso que uno de los objetivos principales de ésta investigación es la creación del motor de sugerencias utilizando filtros colaborativos en el que un grupo de usuarios en este caso estudiantes y profesores participen de manera colaborativa.

La plataforma colaborativa la cual usa tecnología asíncrona, le permite a los usuarios estar recibiendo y enviando información en todo momento, haciendo que el uso de la aplicación sea más natural, y propicie esos espacios de colaboración de los cuales se ha venido discutiendo.

Las actividades colaborativas y la minería de datos representan en gran medida una tendencia positiva en la apropiación del conocimiento por parte del estudiante, en donde los usuarios organizan sus herramientas para que el sistema de información mediante los algoritmos de filtrado colaborativo, haga las recomendaciones con respecto al tema de interés, en este caso el algoritmo de vecinos cercanos utilizado con JSON, permitió procesar mayor información para hacer las recomendaciones a los usuarios.

El acelerado impulso que están ofreciendo las comunicaciones en donde la red es la plataforma y el software se presenta como un servicio, se ha evidenciado que la red y las tecnologías de la información son elementos claves en el futuro de la educación, en donde las redes colaborativas en su aprendizaje autónomo permitan crear tópicos especializados para grupos de interés académico.

## REFERENCIAS

- Alag, S. (2009). *Collective Intelligence in Action*. Manning.
- Balabanovic, M. and Shoham, Y. (1997). Fab: Content-based, Collaborative Recommendation. *Communications of the ACM*. vol. 40. no. 3.
- Bonabeau, E. (2011). *Decisions 2.0. The Power of Collective Intelligence*. MIT Sloan Management Review. Vol.50 No.2.
- Castells, M. (2001). *Internet y la Sociedad Red*.

- Grasso, A. and Convertino, G. (2012). *Collective Intelligence in Organizations. Tools and Studies*. Springer.
- Kautz, H., Selman, B. and Shah, M. (1997). ReferralWeb: Combining Social Networks and Collaborative Filtering. *Communications of the ACM*. vol. 40. no. 3.
- Levy, P. (2004). *Inteligencia Colectiva, por una Antropología del Ciberespacio*. Biblioteca Virtual Em Saúde.
- Nagar, Y. and Malone, T. (2011). *Combining Human and Machine Intelligence for Making Predictions*. Center for Collective Intelligence, Massachusetts Institute of Technology Cambridge, MA 02142.
- Segaran, T. (2007). *Programming Collective Intelligence, Building Smart Web 2.0 Applications*. USA. O'Reilly Media.
- Segaran, T., Evans, C. and Taylor, J. (2009). *Programming the Semantic Web*. USA. O'Reilly Media.
- Teryeen, L., Hill, W., Amento, B., Mcdonald, D. and Creter, J. (1997). PHOAKS: A System for Sharing Recommendations. *Communications of the ACM*. vol. 40. no. 3.
- Watson, M. (2009). *Scripting Intelligence. Web 3.0 Information Gathering and Processing*. USA. Apress.

### ***Authorization and Disclaimer***

*Authors authorize LACCEI to publish the paper in the conference proceedings. Neither LACCEI nor the editors are responsible either for the content or for the implications of what is expressed in the paper.*