# INFORME SOBRE SISTEMAS DE RECOMENDACIÓN BASADOS EN CONOCIMIENTO

**SERENDIPIA:** es un descubrimiento o un hallazgo afortunado, valioso e inesperado que se produce de manera accidental o casual.

## Arquitectura de un sistema basado en contenido:

**Analizador de contenido:** Cuando se trata información no estructurada, se necesita alguna clase de preprocesamiento para extraer información estructurada relevante. Este componente servirá para representar los contenidos de los objetos de forma que sea válida para los siguientes pasos del procesamiento. Los datos se analizan por extracción de características para cambiar la representación del espacio original de información al espacio objetivo. Esta representación se usa como entrada para el **Perfilador** y el **Componente de filtrado**

**Perfilador:** Este módulo recoge datos representativos de las preferencias de un usuario y trata de generalizar los datos para poder generar un perfil de usuario. La estrategia de generalización se hace normalmente a través de técnicas de aprendizaje automático que son capaces de inferir un modelo e intereses de usuario a través de gustos anteriores. Por ejemplo, el **Perfilador** de un recomendador de páginas web puede implementar un método de relevancia basado en “feedback” en el cual la técnica de aprendizaje combina vectores de ejemplos positivos y negativos en un vector prototipo que representa el perfil de usuario. Los ejemplos de entrenamiento serían el “feedback” positivo o negativo del usuario.

**Componente de filtrado:** Este módulo utiliza la información del perfil de usuario para recomendar elementos relevantes haciendo coincidir la representación del perfil con los objetos que han de ser recomendados. El resultado es una puntuación de relevancia. En el ejemplo anterior, la comparación se realizaría procesando la similitud de coseno del vector prototipo y los vectores de objetos.

La primera fase de la recomendación la realiza el analizador de contenido. Utiliza técnicas de extracción de información para obtener palabras clave, conceptos, asociaciones y otras características de los objetos del sistema y formar información estructurada que más tarde se almacena.

Para ayudar a realizar el perfil de un usuario, se almacenan las reacciones a objetos recomendados. Estas reacciones pueden ser tanto positivas como negativas, siendo la forma más simple una representación booleana de estas. También se puede solicitar al usuario que seleccione categorías que son de su interés, para realizar una primera forma de perfil que no requiere de “feedback” dado que se asume reacción positiva a las categorías seleccionadas.

Existen dos formas de captar “feedback” de un usuario. La primera, explícita, es la más directa y requiere de interacción directa con el usuario. Para este caso, se utilizan sistemas de **Me gusta / No me gusta,** Sistemas de **puntuación** y Sistemas de **comentarios.**

La segunda forma, la implícita, no requiere de interacción del usuario de manera directa, si no de un análisis de sus acciones. Una técnica para recoger este tipo de “feedback” es asignar una puntuación a cada una de las acciones de un usuario, como hacer clic en una página, guardar en favoritos o pasar una cantidad de tiempo determinada observando un objeto.

Una vez generados valores de puntuación para cada objeto y cada usuario, el perfilador se encarga de generar un modelo que se ajuste a un usuario concreto con algoritmos de aprendizaje supervisado que se almacena para ser usado por el componente de filtrado más adelante.

Una vez se tiene el resultado del perfilador para un usuario concreto, se pueden utilizar diferentes estrategias para crear un Ranking de objetos para ese usuario a través del componente de filtrado. Los objetos mejor valorados pasan a la lista de recomendación para ese usuario.

Como los gustos de un usuario pueden cambiar, o puede proporcionar nuevo feedback, es recomendable que el perfilador vuelva a generar un perfil de usuario cada cierto tiempo para mantenerse siempre al día de los gustos de este.

## VENTAJAS E INCONVENIENTES DE LOS SISTEMAS BASADOS EN CONTENIDO

**VENTAJAS**

**Independencia de los usuarios:** El perfil de cada usuario de construye de manera totalmente independiente al del resto. De esta manera, no se necesita una gran cantidad de usuarios para que estos sistemas funcionen.

**Transparencia:** De ser necesario, se pueden dar explicaciones sobre el funcionamiento de cada sistema dando una lista de características que se consideran importantes y los criterios que se siguen a la hora de generar las puntuaciones de “feedback”.

**Nuevos objetos:** Los sistemas basados en contenido son capaces de recomendar al instante nuevos objetos a los usuarios, dado que no necesitan de puntuación previa ni interacción de ningún usuario anterior. Basta con que tengan unas características definidas.

**DESVENTAJAS**

**Análisis de contenido limitado:** Normalmente se necesita de conocimiento del dominio para modelar un sistema de este tipo, dado que es importante conocer si en la realidad una característica es importante para un tipo de usuario o no y qué características son realmente destacables en un objeto dado que el número de estas es limitado.

**Sobre-especialización:** Los sistemas basados en contenido rara vez recomiendan objetos que se salen del espectro habitual del usuario. Por ello, a veces se incurre en el problema de no alcanzar la serendipia que a menudo buscan los sistemas de recomendación.

**Nuevo usuario:** Cuando un nuevo usuario entra en el sistema, es necesario que él mismo proporcione información para poder generar una primera instancia de su perfil o en caso contrario habría que esperar a que interactuase con distintos objetos para poder generar un perfil válido.

## ESTADO DEL ARTE