



Ingeniería. Investigación y Tecnología

ISSN: 1405-7743

iit.revista@gmail.com

Universidad Nacional Autónoma de

México

México

Alejandres Sánchez, Hugo Omar; González Serna, Juan Gabriel; González-Franco, Nimrod

Efecto de explicaciones sobre la confianza del usuario en sistemas de recomendación sensibles al contexto

Ingeniería. Investigación y Tecnología, vol. XVII, núm. 4, octubre-diciembre, 2016, pp. 419-428

Universidad Nacional Autónoma de México
Distrito Federal, México

Disponible en: <http://www.redalyc.org/articulo.oa?id=40447587001>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

Efecto de explicaciones sobre la confianza del usuario en sistemas de recomendación sensibles al contexto

Effect of Explanations on Users' Trust in Context-Aware Recommender Systems

Alejandres-Sánchez Hugo Omar

Departamento de Ciencias Computacionales

Centro Nacional de Investigación y Desarrollo Tecnológico

Correo: alejandres@cenidet.edu.mx

González-Serna Juan Gabriel

Departamento de Ciencias Computacionales

Centro Nacional de Investigación y Desarrollo Tecnológico

Correo: gabriel@cenidet.edu.mx

González-Franco Nimrod

Departamento de Ciencias Computacionales

Centro Nacional de Investigación y Desarrollo Tecnológico

Correo: nimrod@cenidet.edu.mx

Información del artículo: recibido: junio de 2014, reevaluado: mayo de 2016, aceptado: junio de 2016

Resumen

Las explicaciones en los sistemas de recomendación se han utilizado para mejorar componentes de la *experiencia del usuario* (EU). Se ha probado que las explicaciones son útiles para mejorar la percepción del usuario respecto a la satisfacción, y para persuadir al usuario de adoptar el sistema de recomendación. Este artículo se enfoca en la evaluación de la confianza del usuario de acuerdo con el sistema y sus relaciones con la comprensión y esfuerzo cognitivo que percibe, mediante el uso de explicaciones en un tipo específico de sistema de recomendación: sensible al contexto. Se realizó un experimento de usuario con un sistema de recomendación organizacional que provee recomendaciones de personas, lugares, eventos, objetos de conocimiento e infraestructura, usando tres tipos de explicación: mapas conceptuales, mapas mentales y explicaciones textuales. Los resultados muestran que, aunque los mapas conceptuales y mentales incrementan la comprensión de las recomendaciones y reducen el esfuerzo cognitivo percibido, las explicaciones textuales mejoran la confianza del usuario respecto al sistema, aun cuando requieren mayor esfuerzo cognitivo para comprenderlas.

Descriptores:

- experiencia del usuario
- evaluación de sistemas
- Findit!
- modelo de ecuaciones estructurales
- explicaciones
- estudios de usuario
- sensibilidad al contexto

Abstract

Explanations on recommender systems have been used to improve components of user experience. It has been proven that explanations are useful to enhance users' perception on satisfaction and persuade users to adopt the recommender system. This paper focuses on evaluating user trust and the trade-offs with understandability and cognitive effort by using explanations in a specific type of recommender system: context-aware. We performed an user experiment with an organizational recommender system that gives recommendations of people, places, events, knowledge objects and infrastructure, using three different types of explanations: conceptual maps, mind maps and textual explanation. The results show that although mind and conceptual maps increases users' understanding of recommendations and decreases cognitive effort, textual explanations improves the users' trust on the system even when they require much cognitive effort to be understood.

Keywords:

- user experience
- system evaluation
- Findit!
- structural equation modeling
- explanations
- user studies
- context-awareness
- user trust

Introducción

Los sistemas de recomendación (SR) han proliferado en los últimos años, por lo que actualmente tenemos presencia de ellos en múltiples dominios que van desde la recomendación de películas hasta el uso en entornos de e-learning, por ejemplo. En todos estos dominios, es necesario evaluar ya sea durante su desarrollo o una vez implementados, para lo cual normalmente se sigue un enfoque de evaluación centrado en el sistema, es decir, analizando cómo funciona el algoritmo de recomendación en relación con las recomendaciones que se emiten. Con este fin, se utilizan métricas como precisión, recuerdo, f-measure y MAE (Mean Average Error), sin embargo, con ellas solo se evalúa parcialmente un sistema de recomendación: la eficiencia del algoritmo (McNee *et al.*, 2006). Mediante el uso de estas métricas no es posible conocer si el usuario utilizará las recomendaciones en su proceso de toma de decisiones y se omiten aspectos subjetivos relacionados con la experiencia del usuario, como: ¿el usuario sabe por qué se le ofrecen las recomendaciones?, ¿en realidad el usuario confía en ellas?, ¿qué tan complicado fue para el usuario entender las recomendaciones?

Debido a ello, surgen enfoques de evaluación centrados en una de las características añadidas a los SR, que es su capacidad para ayudar a los usuarios en el proceso de toma de decisiones y, en consecuencia, mejorar su experiencia. Esto es posible añadiendo explicaciones en la presentación de las recomendaciones.

Los SR que han incorporado características de explicación se analizaron y discutieron de múltiples formas (Tintarev y Masthoff, 2015). La mayoría de los SR utilizan explicaciones tradicionales (en forma de texto) o gráficas (mediante histogramas, ratings gráficos, etcétera). Sin embargo, aún es necesario realizar una compa-

rativa entre el impacto generado por distintas formas de explicación, además de observar si estas ejercen el mismo efecto sobre la experiencia del usuario en una subclase de los sistemas de recomendación que basan su proceso de filtrado en la explotación de aquella información que sirva para caracterizar la situación de un ítem recomendable (Abowd *et al.*, 1999), es decir, los sistemas de recomendación sensibles del contexto (SRSC). En este tipo de sistemas las explicaciones se pueden generar considerando la identificación de las condiciones contextuales que influencian la recomendación de los ítems (Baltrunas *et al.*, 2012).

En este trabajo, se presenta un estudio que aborda el efecto del uso de tres tipos de explicaciones (textuales, mediante, mapas mentales y a través de mapas conceptuales) sobre una componente de la experiencia del usuario: la confianza, en un sistema de recomendación sensible del contexto. Como parte de la experimentación, se empleó un framework de evaluación para SR que permite la incorporación de aspectos subjetivos y relaciones entre condiciones experimentales con distintos factores (Knijnenburg *et al.*, 2015). Adicionalmente, se expone la forma en que los tipos de explicaciones antes mencionados impactan en el esfuerzo cognitivo y la comprensión de los usuarios, así como su relación con la confianza y la variación de estas tres componentes de la experiencia del usuario de acuerdo con los distintos tipos de usuarios.

La organización del trabajo es la siguiente: en la sección 2 se presentan los trabajos relacionados con el uso de explicaciones en sistemas de recomendación, así como con los efectos que estas ejercen sobre la experiencia del usuario, específicamente en el componente de confianza. La sección 3 muestra los detalles del experimento de usuarios y el sistema utilizado en la experimentación. La sección 4 muestra los resulta-

dos obtenidos mediante la aplicación del experimento de usuarios. La sección 5 presenta la discusión de los resultados obtenidos y las implicaciones que estos presentan. Finalmente, la sección 6 presenta las conclusiones.

Trabajos relacionados

De acuerdo con el estado del arte, existen trabajos en los que se estudia el efecto de las explicaciones sobre componentes de la experiencia del usuario. Estos estudios se realizan sobre sistemas de recomendación, particularmente Web, cuya estrategia de presentación de ítems recomendados es acompañarlos de una explicación, en la que se expone por qué un ítem determinado se recomienda (Knijnenburg *et al.*, 2015).

Este impacto se estudia en diversas investigaciones y dentro de aquellas que consideran el componente de confianza, se siguen diferentes estrategias de análisis. Por ejemplo, en Pu y Chen (2007) se estudia, mediante el análisis a un sistema de recomendación Web, la capacidad de un sistema para explicar de manera textual las recomendaciones, de forma que se reduzca el esfuerzo cognitivo que ejercen los usuarios para comprenderlas, así como la forma en que esto influye en la confianza. En Knijnenburg *et al.* (2013) se demostró que las explicaciones pueden influir de manera simultánea en la satisfacción y la percepción que el usuario tiene sobre la ayuda recibida. Otra perspectiva interesante es la del estudio que se presenta en Henriette *et al.* (2008), donde se analiza el impacto de las explicaciones en la transparencia, la confianza y la aceptación de un sistema.

Por otra parte, el análisis presentado en Symeonidis *et al.* (2008) compara la manera en que los estilos de presentación de las explicaciones influyen en los usuarios, lo que se relaciona en este trabajo, donde se estudia la manera en que distintos tipos de explicación influyen en la confianza del usuario. Esto también se relaciona con una investigación centrada en el caso específico de los Sistemas de Recomendación Sensibles al Contexto, donde se identifica que el uso de diferentes estilos de redacción en la generación de explicaciones tipo texto impacta en la experiencia de los usuarios, particularmente en comprensión (Lim *et al.*, 2009).

De acuerdo con el análisis de la literatura, se encontró que el uso de explicaciones en los sistemas de recomendación tradicionales (Web) ejerce un efecto positivo sobre la experiencia del usuario. Sin embargo, el mayor énfasis se presenta sobre explicaciones textuales obviando, de alguna forma, modos alternativos de explicación como los mapas mentales y mapas conceptuales.

Son pocos los trabajos enfocados a sistemas de recomendación sensibles al contexto. Aquellos que existen se enfocan en una sola componente de la experiencia del usuario. Este es el caso de los trabajos presentados en Gasparic y Ricci (2015) y Elahi *et al.* (2015), donde se utilizan sistemas de recomendación sensibles al contexto para medir el efecto de los cambios en el diseño de interacción del sistema de recomendación sobre la calidad de las recomendaciones y la satisfacción del usuario, respectivamente.

Es así como esta investigación se enfoca en el análisis de distintos tipos de explicación (textual, mapa mental y mapa conceptual) y su efecto en la confianza del usuario dentro del sistema y en sistemas de recomendación sensibles al contexto, considerando una combinación de variables encontradas en la literatura como: comprensión y esfuerzo cognitivo percibido, con el fin de identificar correlación entre ellas.

Experimento de usuarios

Como se mencionó, los trabajos relacionados indican que el uso de explicaciones en un sistema de recomendación tradicional ejerce una influencia positiva sobre la experiencia del usuario. Tomando esto como base y considerando el desarrollo de procesos experimentales en Ciencias Computacionales presentados por Wohlin *et al.* (2012) y Jedlitschka *et al.* (2008), realizamos un experimento de usuarios descrito por los siguientes elementos:

Objetivo

El objetivo del estudio fue explorar la forma en que el esfuerzo cognitivo, la comprensión y la confianza de los usuarios de un SRSC varían dependiendo de la presentación de explicaciones en forma textual, mediante mapas mentales o usando mapas conceptuales.

Hipótesis

Para cumplir el objetivo anterior, se consideraron las siguientes suposiciones:

- Hipótesis Nula 1: Las diferencias entre el uso de una modalidad textual, mapas mentales o mapas conceptuales para la presentación de explicaciones son nulas para el componente de confianza.
- Hipótesis Alternativa 1: Hay un efecto sobre el componente de confianza por la presentación de explicaciones usando una modalidad textual, mapas mentales o mapas conceptuales.

- Hipótesis Nula 2: Las diferencias entre el uso de una modalidad textual, mapas mentales o mapas conceptuales para la presentación de explicaciones son nulas para el esfuerzo cognitivo.
- Hipótesis Alternativa 2: Hay un efecto sobre el esfuerzo cognitivo por la presentación de explicaciones usando una modalidad textual, mapas mentales o mapas conceptuales.
- Hipótesis Nula 3: Las diferencias entre el uso de una modalidad textual, mapas mentales o mapas conceptuales para la presentación de explicaciones son nulas para el componente de comprensión.
- Hipótesis Alternativa 3: Hay un efecto sobre el componente de comprensión por la presentación de explicaciones usando una modalidad textual, mapas mentales o mapas conceptuales.
- Hipótesis Nula 4: Los componentes comprensión y esfuerzo cognitivo tienen un efecto nulo sobre el componente de confianza.
- Hipótesis Alternativa 4: Hay un efecto sobre el componente de confianza por parte de variaciones en los componentes comprensión y esfuerzo cognitivo.

Variables

Para aceptar o rechazar las hipótesis se analizaron tres variables dependientes medidas en una escala ordinal con valores del 1 al 5 de acuerdo con el promedio de intensidad con que se percibía cada componente de EU: la variable ECP para el estudio del esfuerzo cognitivo percibido, la variable Cm para examinar la comprensión, y finalmente la variable Cn para conocer la confianza.

Diseño

El diseño experimental implicó la asignación aleatoria de sujetos a 12 grupos de control, cada uno sometido a una condición de interacción definida por el cruce de dos variables independientes nominales: la variable TU (*tipo de usuario*), utilizada para indicar el perfil de un individuo mediante uno de tres posibles valores (Estudiante, Profesor o Empresario) y la variable IE (*interface de explicación*), empleada para indicar el modo en que las explicaciones se presentaron a un individuo mediante uno de tres posibles valores: texto, mapas mentales o mapas conceptuales.

Para determinar el tamaño de la muestra de participantes, fue necesario realizar un análisis de poder estadístico. Esto se origina por la necesidad de detectar una magnitud de efectos bien definidos, es decir, se requiere detectar un efecto de determinada magnitud en una

muestra considerando que el efecto en realidad está presente (Cohen, 1988). Generalmente, en los sistemas de recomendación se presentan efectos de pequeña a mediana magnitud, con diferencias de 0.2 a 0.5 desviaciones estándar, respectivamente (Knijnenburg, 2015). Para calcular el tamaño de la muestra se consideró 85% de poder estadístico, p-value de 0.05 y una magnitud de efecto de 0.5 con una análisis de tipo A priori. La muestra óptima para el experimento correspondía a 118 sujetos de prueba, de los cuales se contó con la participación de 119. Los participantes no contaban con experiencia en el uso del objeto de experimentación y pertenecían a distintos centros de investigación, empresas e instituciones de educación superior. Solo se permitió participar a usuarios que formaran parte de alguno de los tres perfiles de usuario: estudiantes, profesores y empresarios. Del total de participantes, 33 fueron mujeres y 86 hombres. Respecto al rango de edades, 42 participantes se encontraron entre 18 y 24 años, 39 entre 25 y 30, 31 entre 31 y 40 y 7 mayores a 41 años.

Objetos experimentales

Como parte del estudio, los participantes utilizaron un sistema de recomendación organizacional sensible al contexto denominado FindIt! Este sistema se diseñó para clientes móviles y actualmente atiende estudiantes, profesores y empresarios que visitan, por ejemplo, un centro de investigación, institución de educación superior u oficina de gobierno. FindIt! almacena toda la información a procesar en un modelo semántico y como parte de su funcionamiento ofrece recomendaciones sobre personas, objetos de conocimiento, lugares, actividades, eventos, recursos tecnológicos y servicios -basados en competencias organizacionales- que resulten interesantes para ellos. Estas recomendaciones se obtienen considerando las características del perfil de los usuarios y las condiciones contextuales presentes al momento en que llegan a una organización, tales como información temporal, espacial y disponibilidad de los ítems recomendables.

Al momento de desplegar las recomendaciones, es posible acceder a una explicación por cada ítem recomendado que puede visualizarse de diferentes formas de acuerdo con las preferencias de los usuarios, esta tiene por objetivo indicar la razón detrás de la recomendación. Para ello, se siguen tres estrategias de acuerdo con los distintos tipos de explicaciones disponibles en el sistema, conforme a lo siguiente:

- Explicaciones textuales: Se explota la información almacenada en el modelo semántico para llenar una

plantilla de texto explicativo por cada ítem recomendado, posteriormente incorpora diferentes términos y conectores lingüísticos según el tipo de usuario al que se muestre la explicación.

- Explicaciones mediante mapas mentales: Se despliegan los ítems recomendados en torno a un nodo central, indicando las interconexiones entre ellos con arcos etiquetados conforme a las relaciones no taxonómicas especificadas en el modelo semántico.
- Explicaciones mediante mapas conceptuales: En este esquema se emplean arcos etiquetados para indicar los tipos de ítems recomendados, los cuales se agrupan en torno a un nodo central con una oración descriptiva.

En la figura 1 se pueden observar los distintos tipos de explicación utilizados.

Instrumento de evaluación

Se creó un cuestionario para la medición de las variables ECP, Cm y Cn, este se forma por 20 reactivos agrupados en tres categorías que analizan los factores de interés para la investigación. Para validar el cuestionario, las respuestas de los reactivos se sometieron a un *Confirmatory Factor Analysis* (CFA) con indicadores categóricos, aplicando *Weighted Least Squares* (WLS) como

técnica de estimación. En el resultado de este proceso se obtuvo lo siguiente:

- Respecto al factor *esfuerzo cognitivo percibido*, 3 reactivos resultaron apropiados para medir la variable ECP, pues se obtuvo un α de 0.76 y un AVE de 0.678. Estos reactivos se adaptaron de Chen y Pu (2010) y corresponden a afirmaciones cortas del tipo: "Tengo que revisar más de una vez la explicación para entender la recomendación".
- De acuerdo con el factor *comprensión*, 5 reactivos resultaron apropiados para medir la variable Cm, pues se obtuvo un α de 0.87 y un AVE de 0.666. Estos reactivos se adaptaron de Knijnenburg *et al.* (2012) y corresponden a afirmaciones cortas del tipo: "Las explicaciones proporcionadas son claras".
- Asimismo, con el factor *Confianza*, 5 reactivos resultaron apropiados para medir la variable Cn, pues se obtuvo un α de 0.83 y un AVE de 0.616. Estos reactivos se adaptaron de Chen y Pu (2010) y corresponden a afirmaciones cortas del tipo: "No tengo razones para confiar en el ítem recomendado".

Se excluyeron 7 reactivos del análisis debido a que presentaban una baja comunalidad, cargas factoriales cruzadas o correlaciones residuales. Para los factores indicados, los valores de Cronbach's Alpha y Average

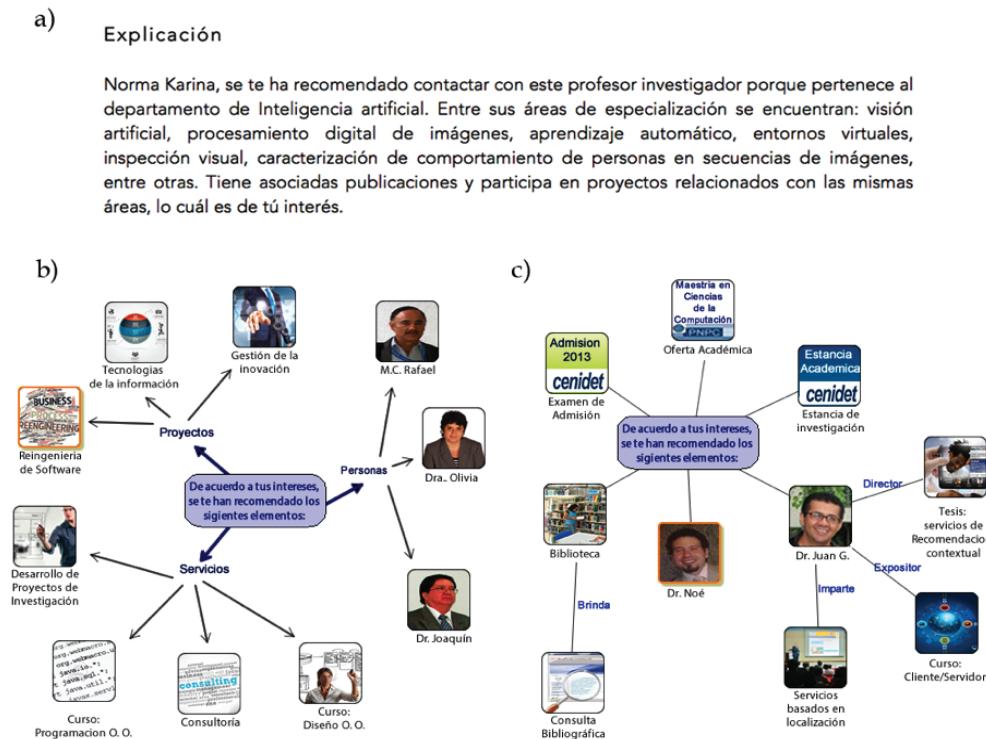


Figura 1. Tipos de explicación (condiciones experimentales),
a) explicación textual,
b) mapa mental, c) mapa conceptual

Variance Extracted (AVE) se mostraron altos ($\alpha > 0.70$, AVE > 0.50), indicando validez convergente. Aunado a esto, la raíz cuadrada de la AVE fue mayor que el factor de correlación para todos los factores, indicando validez divergente.

Procedimiento

Para iniciar la recolección de datos, se pidió a los participantes contestar un formulario básico (nombre, fecha de nacimiento, nivel académico, área de conocimiento y áreas de interés). Una vez hecho esto, se les presentaron las recomendaciones de los ítems que podrían ser de su interés, acompañadas con sus respectivas explicaciones de acuerdo con la condición a la cual se asignaron.

En la tabla 1 se puede observar la asignación de sujetos a los grupos de experimentación, además se muestra la especificación de dos conjuntos de estudiantes. El manejar dos grupos de estudiantes tuvo el objetivo de determinar si existe alguna influencia por efecto de la variabilidad entre tipos de ítem, es decir, cada conjunto de usuario estudiante recibirá distintos tipos de ítems.

Es importante señalar que cada grupo de usuarios recibió recomendaciones de ítems con base en conjunto de tipo de ítems, por ejemplo, el grupo Estudiante recibe recomendaciones de 3 profesores, 3 eventos, 2 objetos de conocimiento y 2 proyectos, mientras que el grupo Estudiante' recibe recomendaciones de 2 profesores, 3 servicios, 3 proyectos y 2 lugares. Los tipos de ítems que reciben son los mismos, sin embargo, esto no significa que reciban los mismos ítems, ya que estos se adecuan al perfil del usuario.

Para dar a cada usuario una experiencia comparable, se limitó el número de recomendaciones a ocho elementos. Al terminar de consultar las recomendaciones, se pidió a los participantes contestar el instrumento de evaluación para plasmar su experiencia sobre la interacción con el sistema.

Análisis de resultados

Las figuras 2, 3, 4 y 5, muestran la distribución y simetría de los resultados obtenidos tras la aplicación de los

Tabla 1. Distribución de participantes por condición experimental

Tipo de usuario	Textual	Mapa mental	Mapa conceptual
Estudiante	10	11	9
Estudiante'	10	10	9
Profesor	10	9	10
Empresario	10	11	10

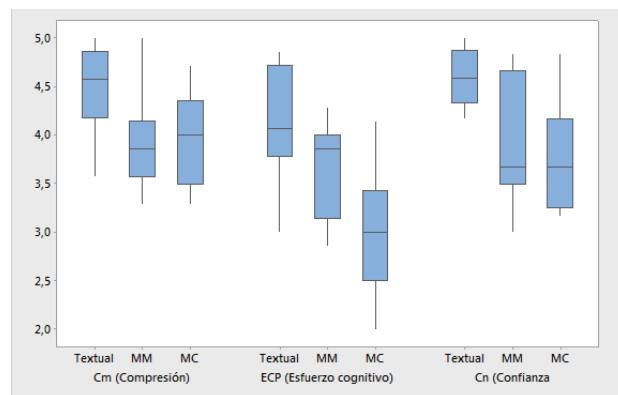


Figura 2. Diagrama de caja de los datos obtenidos por los participantes profesores

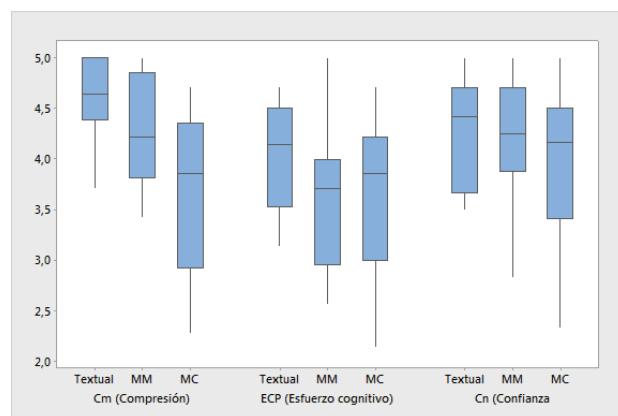


Figura 3. Diagrama de caja de los datos obtenidos por los participantes empresarios

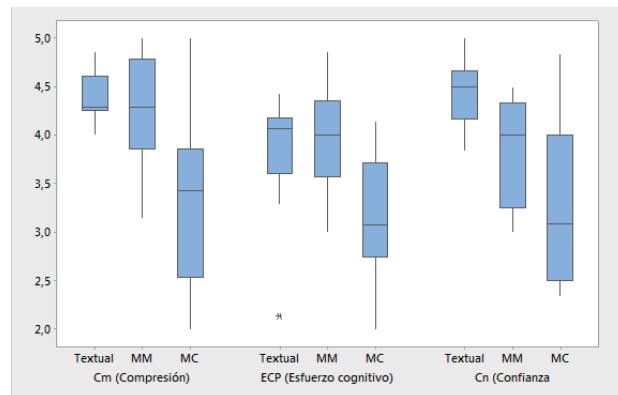


Figura 4. Diagrama de caja de los datos obtenidos por los participantes estudiantes

cuestionarios y como puede observarse, los valores de las tres variables dependientes analizadas no tuvieron la misma desviación respecto a la media.

Para buscar correlaciones entre los componentes de la Experiencia del Usuario y las condiciones experi-

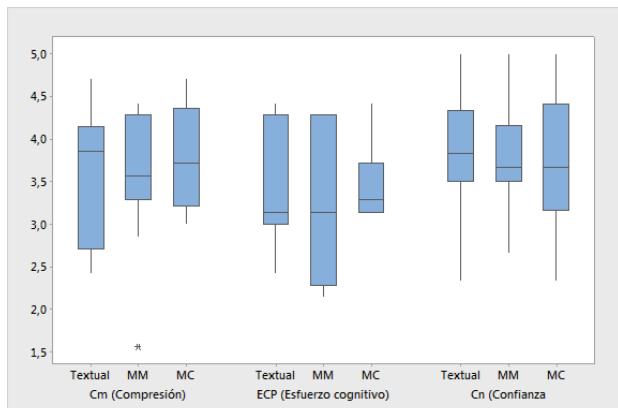


Figura 5. Diagrama de caja de los datos obtenidos por los participantes estudiantes'

mentales, los datos se sometieron a un modelo de ecuaciones estructurales, el cual incluye el modelo de evaluación de los factores y las relaciones estructurales entre los factores y otras variables.

El modelo mostrado en la figura 6, se creó considerando el *framework* descrito en Knijnenburg *et al.* (2015) que incluye la interfaz de explicación como el aspecto objetivo del sistema y que presenta las condiciones experimentales: explicación textual, mapa mental y mapa conceptual. De igual manera se consideran características personales del usuario para indicar qué tipo de usuario participó. La inclusión de este último aspecto permite conocer si los efectos analizados varían respecto a los distintos perfiles de usuario de los participantes.

Asimismo, se incluyen las componentes de la experiencia del usuario sobre las cuales se evaluarán los efectos de las condiciones experimentales: esfuerzo cognitivo, comprensión y confianza. En la figura 6, se pueden observar los arcos que indican las magnitudes de los efectos ejercidos y la significancia estadística de cada uno de ellos. El modelo presenta un ajuste correcto¹ con los datos obtenidos, que fueron los siguientes: $p < 0.01$; RMSEA = 0.0071; 90% CI: [0.047, 0.092]; CFI = 0.971, TLI = 0.964.

Dicho modelo muestra que el uso de explicaciones mediante mapas mentales y mapas conceptuales respecto a las explicaciones textuales, ejerce un efecto negativo sobre el esfuerzo cognitivo percibido por el usuario, excepto en el caso de los profesores, es decir, para la mayoría de los casos se reduce el esfuerzo cognitivo percibido, sin embargo este efecto no es significativo, por lo que la Hipótesis Nula 2 se rechaza.

1 En Hu y Bentler (1999) se proponen los valores mínimos de los parámetros estadísticos para un ajuste correcto del modelo: CFI > 0.96, TLI > 0.95, RMSEA < 0.05 y el límite superior de 90% CI por debajo de 0.10.

También se puede observar que al utilizar distintos tipos de explicación y de usuario se reduce la comprensión sobre las recomendaciones presentadas a los usuarios. Esto se debe a que, en conjunto, las tres condiciones experimentales ejercen distintos efectos sobre la comprensión, es decir, el efecto de la explicación textual puede aumentar la comprensión de las recomendaciones mientras que el efecto de los mapas conceptuales es decremental, por lo que la Hipótesis Nula 3 se rechaza.

Además, en cuanto al esfuerzo cognitivo percibido y la comprensión, se observa una relación significativa entre ellos, es decir, a menor esfuerzo cognitivo, mayor comprensión. Por otro lado, la comprensión influye en la confianza del usuario respecto al sistema, ya que presenta un efecto positivo muy significativo. El efecto que se ejerce sobre la confianza por el esfuerzo cognitivo percibido es, además de no significativo, irrelevante, por lo que la Hipótesis Nula 4 se rechaza parcialmente.

Una vez observados los efectos, correlaciones y sus significancias, en la figura 7 se presenta la comparación de la magnitud de los efectos de cada condición experimental (tipos de explicaciones) sobre los factores analizados (esfuerzo cognitivo percibido, comprensión y confianza). Se puede observar que tanto mapas mentales y mapas conceptuales (figura 7b-c) reducen significativamente el esfuerzo del usuario para comprender las recomendaciones que el sistema emite. Esto se debe a que ambos presentan relaciones gráficas entre los ítems, lo que permite al usuario ejercer un menor esfuerzo cognitivo para visualizar la utilidad de los mismos con base en otros ítems.

Por otra parte, la condición experimental que presenta un efecto mayor sobre la comprensión es la explicación textual (figura 7a). Sin embargo, también esta condición es la que presenta un mayor esfuerzo cognitivo por parte del usuario. Esto significa que los usuarios tienden a comprender mejor una explicación en forma de texto, pero al hacerlo se requiere aplicar un esfuerzo cognitivo mayor que con el resto de las condiciones.

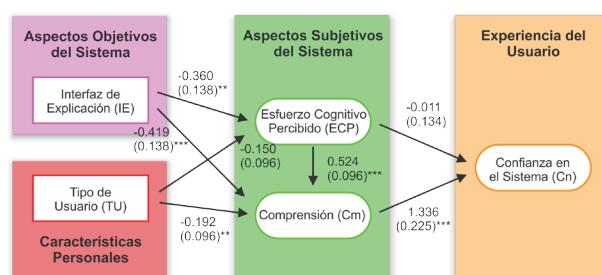


Figura 6. Modelo de ecuaciones estructurales para los datos del experimento.

Niveles de significancia: *** $p < 0.001$, ** $p < 0.01$, no significativo $p > 0.05$

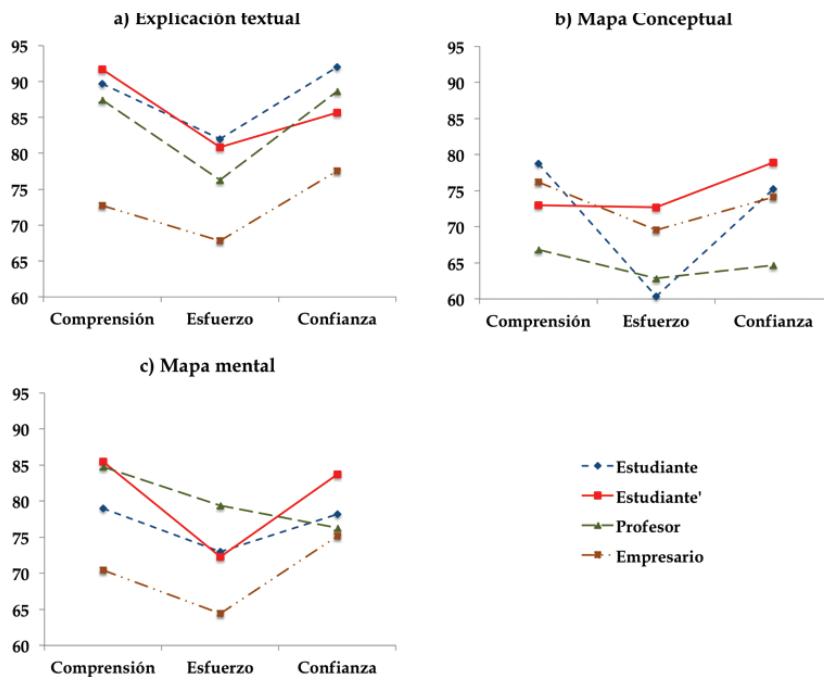


Figura 7. Efectos ejercidos sobre comprensión, esfuerzo y confianza por las tres condiciones experimentales, a) explicación mediante texto, b) explicación mediante mapas conceptuales, c) explicación mediante mapas mentales. División por tipo de usuario

ciones, ya que la explicación textual no presenta gráficamente relaciones con otros ítems.

En cuanto a la confianza, se puede observar que las explicaciones textuales son aquellas que generan la mayor confianza del usuario con el sistema, por lo que la Hipótesis Nula 1 se rechaza. Esto nos lleva a confirmar la significancia de la relación entre la comprensión y la confianza. Con una nota de precaución, debido a la falta de significancia de algunos resultados (aquellos relacionados a esfuerzo cognitivo percibido), se puede interpretar que ante un incremento de la comprensión de las explicaciones se presentará el mismo efecto sobre la confianza del usuario con el sistema.

Con base en la introducción intencional de un segundo grupo de Estudiantes (Estudiantes'), que se presenta en la tabla 1, se realizó un análisis χ^2 sobre los efectos obtenidos de ambos conjuntos sobre los tres factores analizados para determinar si puede existir variabilidad entre tipos de ítem. Con $\chi^2=0.147$, valor menor al valor de referencia con 2 grados de libertad $\chi^2=5.99$, se concluyó que la heterogeneidad de los ítems recomendados no presenta un efecto significativo sobre los factores analizados.

Discusión

Con los resultados obtenidos se ratifica lo que se muestra en la literatura: el uso de explicaciones incrementa la calidad de la experiencia del usuario. En este trabajo, se comprueba específicamente que el uso de explicaciones

incrementa la confianza del usuario en sistemas de recomendación sensibles al contexto y que el tipo de explicación a utilizar dependerá del factor que se desee potenciar. La naturaleza de un SRSC proviene de la ayuda en el proceso de toma de decisiones. A diferencia de los SR tradicionales (Web), un SRSC posee la capacidad de adaptarse al contexto en el cual se encuentra, razón por la cual es posible usarlos como sistemas de guiado, de búsqueda de lugares, etcétera.

Respecto a la posibilidad de generalizar los resultados que se obtuvieron con otras poblaciones o situaciones similares, se calculó el tamaño de la muestra probabilística con base en Cohen (1988), lo que permitió aplicar el principio de transferibilidad (Martínez, 2012) a las siguientes inferencias:

- Las explicaciones ejercen un efecto positivo sobre la confianza del usuario en el sistema, además el efecto se encuentra relacionado con el aumento de la comprensión de las recomendaciones que emite el sistema. Debido a ello, se identificó que si se desea reducir el esfuerzo cognitivo percibido por el usuario, se deberá utilizar algún tipo de explicación que muestre gráficamente la relación entre distintos tipos de ítems recomendados, como el caso de los mapas mentales y mapas conceptuales.
- Si se busca aumentar la confianza del usuario en el sistema sin importar el esfuerzo cognitivo que esto implique al usuario, es necesario utilizar una explicación textual de las recomendaciones. Sin embar-

go, hay que enfatizar que si bien el efecto es positivo sobre la confianza, aún no se ha comprobado su interacción con otras componentes de la experiencia del usuario.

Dado que el sistema de prueba FindIt! contaba con heterogeneidad de ítems, fue importante demostrar que esta característica no presenta influencia alguna sobre el resto de los factores evaluados. Los resultados del análisis χ^2 muestran que la heterogeneidad de ítems no ejerce efecto significativo sobre el esfuerzo cognitivo, la comprensión o incluso la confianza del usuario respecto al sistema.

El modelo utilizado para el análisis de los datos es útil considerando escenarios donde se requiera evaluar componentes de la experiencia del usuario, como confianza y satisfacción, mediante la modificación de la interfaz de usuario del sistema de recomendación. Consideramos que el modelo puede mejorarse mediante la selección de nuevas componentes de la experiencia del usuario. Sin embargo, será necesario verificar su influencia sobre el resto de las variables definidas previamente. Se trata de observar si los efectos de las explicaciones sobre la confianza permanecen en relación con nuevos factores. La característica de sensibilidad al contexto es muy importante. Se requiere que los sistemas de recomendación, sobre los cuales se aplique el modelo propuesto, cuenten con alguna característica como personalización con base en tiempo, ubicación o perfil del usuario, entre otras.

Finalmente, con los resultados obtenidos es posible crear una serie de recomendaciones para los diseñadores/programadores de sistemas de recomendación sensibles al contexto para la mejora de la experiencia del usuario.

Conclusiones

Tradicionalmente, el paradigma de evaluación de los sistemas de recomendación se centra en el algoritmo de recomendación. Este artículo presenta un estudio que utilizó un paradigma centrado en el usuario, donde se ha realizado poco trabajo. De esta forma, se aporta al estado del arte de la evaluación de sistemas de recomendación, mediante la descripción del proceso de evaluación de la confianza del usuario, así como su relación con la comprensión y el esfuerzo cognitivo mediante el uso de explicaciones, en un tipo específico de sistema de recomendación: sensible al contexto. Los resultados muestran que las explicaciones textuales ayudan a incrementar la confianza del usuario hacia el sistema, aun cuando implique un esfuerzo mayor para

interpretarlas. Por otro lado, si lo que se desea es reducir el esfuerzo que ejerce el usuario para entender las explicaciones, es necesario utilizar explicaciones gráficas mediante mapas mentales o conceptuales.

Referencias

- Abowd G.D., Dey A.K., Brown P.J., Davies N. Towards a better understanding of context and context-awareness, en: 1st Proceedings of the International Symposium on Handheld and Ubiquitous Computing, 1999, Londres, UK.
- Baltrunas L., Ludwig B., Peer S., Ricci F. Context relevance assessment and exploitation in mobile recommender systems. *Personal and Ubiquitous Computing*, volumen 16 (número 5), 2012: 507-526.
- Chen L. y Pu P. User evaluation framework of recommender systems, en proceedings of the joint workshop on interfaces and human decision making for recommender systems, Hong Kong, China, 2010.
- Cohen J. *Statistical power analysis for the behavioral sciences*, 2a ed., Lawrence Erlbaum Associates, 1988.
- Elahi M., Ge M., Ricci F., Fernández-Tobías I., Berkovsky S., Massimo D. Interaction design in a mobile food recommender system, en proceedings of workshop on social recommender systems, Viena, Austria, 2015.
- Gasperic M., Ricci F. Modeling context-aware command recommendation and acceptance in an IDE, en: Proceedins of the second international workshop on context for software development, Nueva Jersey, 2015.
- Henriette C., Evers V., Ramlal S., Someren M.V., Rutledge L., Stash N., Aroyo L., Wielinga B. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction*, volumen 18 (número 5), 2008: 455-496.
- Hu L., Bentler P.M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, volumen 6 (número 1), 1999: 1-5.
- Jedlitschka A., Ciolkowski M., Pfahl D. Reporting experiments in software engineering, en: Shull F., Singer J., Sjøberg D.I.K. *Guide to advanced empirical software engineering*, 1a ed., Springer London, 2008, pp. 201-228.
- Knijnenburg B., Bostandjiev S., O'Donovan J., Kobsa A. Inspectability and control in social recommenders, en: 6th Proceedings of the ACM conference on recommender systems, Dublin, 2012, pp. 43-50.
- Knijnenburg B., Kobsa A. Making decisions about privacy: information disclosure in context-aware recommender systems. *ACM Transactions on Interactive Intelligent Systems*, volumen 3 (número 3), 2013: 20.
- Knijnenburg B., Willemsen M. Evaluation recommender system with user experiments, en: Ricci F., Rokach L., Shapira B. *Re-*

- commender Systems Handbook*, 2a ed., Springer, 2015, pp. 309-352.
- Lim B.Y., Dey A.K., Avrahami D. Why and why not explanations improve the intelligibility of context-aware intelligent systems, en: Proceedings of the SIGCHI conference on human factors in computing systems, ACM, 2009, pp. 2119-2128.
- Martínez-Salgado C. El muestreo en investigación cualitativa. Principios básicos y algunas controversias. *Revista Ciencia & Saude Coletiva*, volumen 17 (número 3), 2012: 613-619.
- McNee S.M., Konstan J.A., Riedl J. Being accurate is not enough: How accuracy metrics have hurt recommender systems, en: Extended abstracts of the ACM conference on human factors in computing systems, Montreal, Canada, 2006, pp. 1097-1101.
- O'Donovan J., Smyth B. Trust in recommender systems, en: 10th Proceedings of the international conference on intelligent user interfaces, San Diego, 2005, pp. 167-174.
- Pu P. y Chen L. Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems*, volumen 20 (número 6), 2007: 542-556.
- Symeonidis P., Nanopoulos A., Manolopoulos Y. Justified recommendations based on content and rating data, en: Workshop on Web mining and Web usage analysis, Las Vegas, 2008.
- Tintarev N. y Masthoff J. Explaining recommendations: Design and evaluation, en: Ricci F., Rokach L., Shapira B. *Recommender Systems Handbook*, 2a ed., Springer, 2015, pp. 353-382.
- Wholin C., Runeson P., Höst M., Ohlsson M.C., Regnell B., Wesslén A. *Experimentation in software engineering*, 1a ed., Springer-Verlag Berlin Heidelberg, 2012, pp. 99-36, 55-72.

Este artículo se cita:

Citación estilo Chicago

Alejandres, Sánchez, Hugo Omar, Juan Gabriel González-Serna, Nimrod González-Franco. Efecto de explicaciones sobre la confianza del usuario en sistemas de recomendación sensibles al contexto. *Ingeniería Investigación y Tecnología*, XVII, 04 (2016): 419-428.

Citación estilo ISO 690

Alejandres-Sánchez H.O., González Serna J.G., González-Franco N. Efecto de explicaciones sobre la confianza del usuario en sistemas de recomendación sensibles al contexto. *Ingeniería Investigación y Tecnología*, volumen XVII (número 4), octubre-diciembre 2016: 419-428.

Semblanzas de los autores

Hugo Omar Alejandres-Sánchez. Es maestro en ciencias de la computación por el Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), México. Es profesor-investigador de la División Académica de Tecnologías de la Información de la Universidad Tecnológica Emiliano Zapata del Estado de Morelos. Entre sus líneas de especialización se encuentran la evaluación de la experiencia del usuario, Web semántica y sistemas de recomendación semánticos sensibles al contexto.

Juan Gabriel González-Serna. Es doctor en ciencias computacionales por el Centro de Investigación en Computación (CIC) del Instituto Politécnico Nacional (IPN), México. Miembro del Sistema Estatal de Investigadores (SEI) en Morelos, México. Es profesor-investigador del área de Sistemas Distribuidos en CENIDET. Entre sus líneas de especialización se encuentran los Servicios Basados en Localización (LBS), sistemas de recomendación semánticos sensibles al contexto, realidad aumentada, tecnologías WiFi, Bluetooth, RFID y QRCodes.

Nimrod González-Franco. Es maestro en ciencias de la computación por CENIDET, México. Actualmente cuenta con una beca otorgada por CONACYT para la realización de estudios doctorales en CENIDET. Entre sus líneas de especialización se encuentran los sistemas de recomendación semánticos sensibles al contexto, la ingeniería ontológica y el modelado organizacional.