

UN METAMODELO OLAP PARA LA EVALUACIÓN DE AMBIENTES TEL¹

Juan Cockbaine O.^{2,3} Ignacio Casas R.²

Recibido el 30 de agosto de 2004, aceptado el 30 de noviembre de 2004

RESUMEN

Los procesos de enseñanza-aprendizaje (E/A) en las universidades pueden ser potenciados por ambientes TEL (Technology Enhanced Learning). Considerando que estos ambientes buscan provocar cambios para mejorar la efectividad educacional, es fundamental incorporar mecanismos de evaluación que provean retroalimentación respecto de su impacto en la educación. Creemos que el desarrollo de un DSS (Decision Support System), cuyos elementos de decisión estén sustentados por el análisis de múltiples fuentes de datos provenientes del propio ambiente TEL, aportará a resolver este problema. Exponemos la implementación OLAP (OnLine Analytic Processing) de CCPR_R, un metamodelo para la evaluación de ambientes TEL y el desarrollo de un modelo experimental. Concluimos que CCPR_R puede ser parte de un DSS para la evaluación de ambientes TEL, que incentiva el desarrollo de indicadores de gestión.

Palabras clave: Evaluación, Enseñanza/Aprendizaje, TEL, OLAP/DSS.

ABSTRACT

University teaching-learning (T/L) processes can be enhanced by TEL (Technology Enhanced Learning) environments. Considering that these environments goal is to provoke changes in order to improve educational effectiveness, the incorporation of evaluation tools that provide feedback is fundamental. We believe that the development of a DSS (Decision Support System), whose decision elements are supported by the analysis of a variety of sources of data that come from the TEL system itself, will contribute to the solution of this problem. We propose an OLAP (OnLine Analytic Processing) implementation of CCPR_R, a meta-model for the evaluation of TEL systems, as well as the development of an experimental model. We conclude that CCPR_R can be part of a DSS for the evaluation of TEL environments, thus promoting the development of management indicators.

Keywords: Evaluation, Teaching/Learning, TEL, OLAP/DSS.

INTRODUCCIÓN

El desarrollo de las tecnologías de información ha contribuido al cambio de enfoque del proceso de enseñanza y aprendizaje (E/A). Se observa que las universidades evolucionan paulatinamente desde un enfoque centrado en la enseñanza y la actividad del profesor a uno que enfatiza el aprendizaje y la participación proactiva del estudiante. En tal sentido, la construcción de sistemas tecnológicos de apoyo a los procesos de E/A ha sido producto de una práctica para concretar y sostener tal cambio.

Esta realidad nos permite establecer un importante requerimiento que atañe a un tomador de decisiones en una institución de educación superior, esto es, establecer indicadores que permitan medir la efectividad del cambio.

Pensamos que para atender este requerimiento es necesario evaluar [6], [14] un ambiente que reúna al menos elementos tecnológicos y de aprendizaje. De esta evaluación se podrán obtener medidas de efectividad que apoyen al tomador de decisiones en la resolución parcial de interrogantes tan complejas como ¿cuál es el nivel de impacto de las tecnologías en el aprendizaje de los estudiantes?

¹ Este trabajo es patrocinado por el proyecto MECESUP PUC0201 CITEDUC (Centro de Innovación Metodológica y Tecnológica para Mejorar la Calidad de los Procesos de Enseñanza y Aprendizaje en la Universidad).

² Pontificia Universidad Católica de Chile, Departamento de Ciencias de la Computación, Av. Vicuña Mackenna 4860, Santiago, Chile. {jqcockba, icasas}@puc.cl

³ Universidad de Santiago de Chile, Departamento de Ingeniería Informática, Av. Ecuador 3659, Santiago, Chile.

Para conceptualizar el ambiente a que nos referimos adoptamos la definición de Goodman [9] que lo concibe como un ambiente de aprendizaje potenciado por tecnología, cuya meta es mejorar la efectividad en la educación y el aprendizaje. Esta conceptualización que es referenciada como un ambiente TEL (Technology Enhanced Learning) puede incluir diversas expresiones tecnológicas, además de una serie de elementos que el autor agrupa bajo precondiciones para el cambio y procesos críticos para lograr ese cambio.

En resumen, lo que deseamos evaluar es un ambiente TEL, tarea que exige desarrollar herramientas para obtener retroalimentación de apoyo a la toma de decisiones de los gestores académicos.

En esta línea, creemos que la retroalimentación que se concreta mediante indicadores de apoyo a la toma de decisiones se debería fortalecer al utilizar un enfoque de múltiples fuentes y métodos de análisis, además de un metamodelo de evaluación que facilite la formulación de tales indicadores. Dicho de otra forma, se debe fortalecer la formulación de indicadores, no solamente basándose en los resultados obtenidos de encuestas de opinión analizadas mediante técnicas estadísticas, sino que además en otras fuentes de datos como Log de Web, monitoreos, Log de mail, análisis de páginas Web, información de los actores del proceso de E/A, etc., tratadas por diversas técnicas como la minería de datos, algoritmos ad hoc, estadística, etc.

Por otro lado, el considerar múltiples fuentes de datos abre una inmensa perspectiva de análisis que se traduce en requerimientos del tomador de decisiones. Pensamos que la implementación exitosa de estos requerimientos podrá ser canalizada por un DSS (Decision Support System) [13].

Con el objeto de apoyar las decisiones de los gestores académicos, este trabajo define un metamodelo para el desarrollo de modelos de evaluación educacional. Este metamodelo, que hemos bautizado con el nombre de CCPR_R (Contexto, Componentes, Procesos, Resultados y sus Relaciones), podrá constituirse en un componente fundamental de un DSS, permitiendo la especificación de indicadores de apoyo a la toma de decisiones, basados en diversas fuentes de datos existentes en un ambiente TEL. Luego, el objetivo de este artículo es exponer la especificación e implementación de CCPR_R, además del desarrollo de un experimento.

Basamos la especificación de CCPR_R en la estructura original del modelo CIPP (compuesto de las áreas Context, Input, Process y Product) para la evaluación

[15]. Además, para efectos de este trabajo establecemos que cada área del modelo CIPP se defina mediante la especificación de dimensiones e indicadores de evaluación, formulados en base a parámetros sustentados por hechos OLAP (http://dev.hyperion.com/resource_library/white_papers/). Esta nueva estructura corresponde al metamodelo CCPR_R, la que especificamos mediante dependencias funcionales [3] e implementamos como un modelo OLAP.

Respecto del experimento, éste corresponde a la definición e implementación de un modelo de evaluación construido con CCPR_R. Para tal efecto, la definición de dimensiones e indicadores de evaluación se establece en base a la conceptualización de un ambiente TEL y la definición de parámetros, a partir de la arquitectura tecnológica de un sistema de enseñanza definida en [12]. Para completar el desarrollo del experimento se describe el proceso de Extracción, Transformación y Carga (ETL) [5] que explican los hechos OLAP. Tal explicación se basa en la transformación de la fuente Log de Web (<http://www.w3.org/>) mediante código SQL (<http://www.ansi.org/>) y minería de datos, específicamente técnicas de cluster [10]. Finalmente se ilustra una perspectiva de visualización de resultados.

Como campo de experimentación se ha considerado el ambiente TEL de la Pontificia Universidad Católica de Chile (PUC, <http://www.puc.cl>), el que utiliza un courseware bautizado como "Escritorio Virtual". Este ambiente u objeto de cambio se describe brevemente en la siguiente sección.

AMBIENTE TEL EXPERIMENTAL

El ambiente TEL implementado en PUC se ha fundamentado en la necesidad de transitar desde un enfoque tradicional de enseñanza a un enfoque centrado en el aprendizaje del estudiante [1], definiéndose una serie de características que han perfilado su construcción: 1) Rol del profesor, como guía y orientador más que un maestro superior, 2) Planificación Instruccional, enfocada en la naturaleza de los objetivos del curso, 3) Material Instruccional, determinado por las necesidades del estudiante respecto de su aprendizaje, 4) Dinámica Instruccional, enfatizando la relación profesor estudiante en forma bidireccional y asíncrona, 5) Ritmo Instruccional, orientado a la necesidad de otorgar mayor libertad al estudiante para decidir sus tiempos, 6) Trabajo del Estudiante, priorizando el trabajo grupal y colaborativo sin descuidar la autonomía, 7) Evaluación del Estudiante, orientado a evaluar el real aprendizaje y 8) Evaluación del Curso, enfatizando la necesidad de retroalimentación temprana y frecuente.

El elemento tecnológico de este ambiente corresponde a un courseware implementado como un catálogo de cursos en la Intranet de la universidad. Este catálogo contiene vínculos a los sitios Web de cursos por Escuelas y Departamentos Académicos (<http://cursos.puc.cl/>).

Cada sitio define una página Web dividida en tres secciones, que son material de curso, información administrativa y comunicación. Cada sección agrupa una serie de funcionalidades como material de clases, noticias, grupos de discusión, correo, entrega de trabajos, etc. Los sitios adoptan una estructura común para un grupo de cursos, permitiéndose incrustar contenido como también hacer referencia a otros contenidos tanto locales como dispuestos en Internet. Estos sitios son administrados por un sistema llamado Escri-Vir (Fig. 1), facilitando la actualización de contenido en un computador cliente y posterior transferencia de archivos al servidor Web de cursos (<http://cursos.puc.cl/escrivirweb/>).

El desarrollo de estos sitios ha sido gradual desde el año 2000 a la fecha y se ha acompañado de la aplicación de encuestas como herramienta de retroalimentación. Estas encuestas se componen de 50 preguntas en promedio, dispuestas bajo una escala Likert de 4 opciones la mayoría de ellas, alcanzándose estimados Alfa de Cronbach [11] cercanos a 0,9.



Fig. 1 Escri-Vir, Escritorio Virtual del profesor.

CCPR_R

CCPR_R corresponde a un metamodelo para el desarrollo de modelos de evaluación formativa y sumativa, mediante métodos cuantitativos [6].

CCPR_R basa su estructura en las áreas de evaluación Context, Input, Process y Product que componen el modelo CIPP de Stufflebeam. Estas áreas las hemos interpretado de la siguiente forma: 1) el contexto, corresponde al entorno en que se inserta el objeto a evaluar, 2) las entradas son los componentes, partícipes de uno o más procesos, 3) los procesos corresponden a funciones de transformación que dan valor agregado al estudiante y 4) el producto, representa los resultados obtenidos de la ejecución de los procesos. Estas áreas también se pueden precisar; por ejemplo, Stufflebeam divide la evaluación del área Product en Impact, Effectiveness, Sustainability y Transportability (<http://www.wmich.edu/evalctr/checklists/cippchecklist.pdf>).

Otros autores como De La Orden [4] y García [8] presentan interesantes propuestas para la evaluación, desarrolladas bajo una estructura similar a la del modelo CIPP.

Encontramos que el modelo CIPP es una buena base para el desarrollo de CCPR_R puesto que: 1) sus áreas permiten establecer una primera aproximación de la arquitectura del objeto de cambio, conceptuándolo como un sistema cuyos elementos de información provienen del entorno, las entradas, los procesos, las salidas y la retroalimentación, 2) no es necesario establecer completamente la evaluación para ejecutarla, lo cual otorga flexibilidad a este proceso según necesidades y prioridades y 3) cada área es factible de perfeccionar, con el objeto de especificar un modelo de evaluación en forma más rigurosa. Este último aspecto es parte de la contribución de este artículo, que se concreta en un metamodelo para la construcción de indicadores de gestión.

Especificación arquitectónica de CCPR_R

CCPR_R organiza la evaluación de ambientes TEL en macroáreas llamadas Contexto, Componentes, Procesos, Resultados y sus Relaciones (_R). A su vez, cada macroárea de CCPR_R se redefine en términos de dimensiones, indicadores, parámetros y hechos OLAP con el objeto de precisar la evaluación a efectuar (Ver Fig.2).

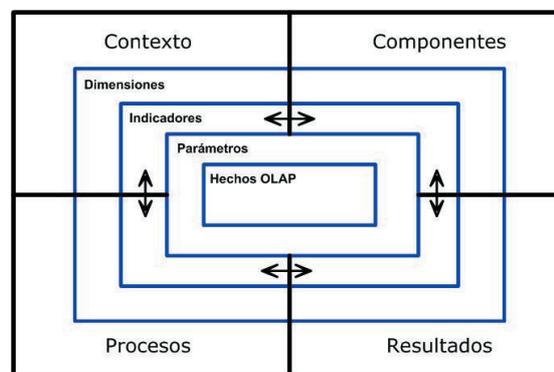


Fig. 2 Arquitectura de CCPR_R

Cada dimensión corresponde a un subconjunto de un área del modelo CCPR_R que se ha decidido evaluar. Visto de otra forma, una dimensión representa un componente del ambiente TEL, el que se ha asociado a una determinada área de CCPR_R con el objeto de definir su evaluación. A su vez, cada dimensión de evaluación se define por un conjunto propio de indicadores para la toma de decisiones. Estos indicadores corresponden a formulaciones para la generación de

retroalimentación, respecto de la dimensión de evaluación que definen. La formulación de cada indicador se basa en un conjunto de parámetros o variables que harán posible su cálculo. Cada parámetro puede formar parte de varias formulaciones, esto es, pueden ser referenciados por más de un indicador. Los hechos OLAP representan los datos asociados a cada parámetro. Estos hechos OLAP son generados mediante un proceso de transformación de las fuentes de datos existentes en un ambiente TEL.

Complementariamente, CCPR_R permite la definición de relaciones entre indicadores mediante una formulación. A esta relación entre indicadores la llamamos de coherencia, la que se concreta por la formulación de indicadores que incorporan indicadores definidos en una sola dimensión o indicadores presentes en más de una dimensión. Las relaciones de coherencia formuladas con indicadores de una sola dimensión las llamamos Relaciones de Coherencia Intra Dimensión (en forma abreviada RCIntraD). Las relaciones de coherencia formuladas con indicadores provenientes desde más de una dimensión las llamamos Relaciones de Coherencia Inter Dimensión (en forma abreviada RCInterD).

Luego, la especificación de la arquitectura de CCPR_R queda formalizada por el conjunto de dependencias funcionales $D \rightarrow A$, $I \rightarrow S^{2^D}$, $I \rightarrow S^{2^P}$ y $(P,T) \rightarrow H$ que se explican a continuación:

1. La agrupación de dimensiones en áreas se representa por la dependencia funcional $D \rightarrow A$. D corresponde al conjunto de d dimensiones asociadas con A . A corresponde a las áreas de evaluación definidas, esto es $A = \{\text{CONTEXTO, COMPONENTES, PROCESOS, RESULTADOS}\}$.

2. La relación entre dimensiones y sus indicadores, relaciones de coherencia y sus indicadores se expresa por la dependencia funcional $I \rightarrow S^{2^D}$. D corresponde al conjunto de d dimensiones. I corresponde al conjunto de i indicadores. Todo indicador se asocia con un elemento de S^{2^D} según la regla $I \rightarrow S^{2^D}$. $S^{2^D} \subset 2^D$ representa dimensiones y/o relaciones de coherencia a considerar en la evaluación. 2^D es el conjunto potencia de D determinado por 2^d elementos. Como ejemplo, sea $D = \{d1, d2, d3\}$ el conjunto de dimensiones. Luego $2^D = \{\{d1, d2, d3\}, \{d1\}, \{d2\}, \{d3\}, \{d1, d2\}, \{d1, d3\}, \{d2, d3\}, \{\emptyset\}\}$. Suponiendo que cada elemento del subconjunto S^{2^D} de 2^D $\{\{d1\}, \{d2\}, \{d3\}, \{d1, d2\}, \{d1, d3\}\}$, se pueden asociar según la regla $I \rightarrow S^{2^D}$ con elementos de $I = \{i1, i2, i3, i4, i5, i6, i7, i8, i9\}$, siendo $i6, i7, i8$ e $i9$ relaciones de coherencia, entonces los elementos de S^{2^D} pueden

quedar definidos por los grupos de indicadores $\{i1\}$, $\{i2, i3\}$, $\{i4, i5, i9\}$, $\{i6, i7\}$ e $\{i8\}$ respectivamente. Las relaciones RCIntraD y RCInterD se explican al restringir $I \rightarrow S^{2^D}$ en $I \rightarrow D$ e $I \rightarrow (S^{2^D} - D)$ respectivamente.

3. La relación entre parámetros e indicadores se expresa por la dependencia funcional $I \rightarrow S^{2^P}$. P corresponde al conjunto de p parámetros. I corresponde al conjunto de i indicadores. Todo indicador se asocia con un elemento de S^{2^P} , según la regla $I \rightarrow S^{2^P}$. $S^{2^P} \subset 2^P$ representa agrupaciones de parámetros que permiten formular un indicador. 2^P es el conjunto potencia de P determinado por 2^p elementos. Dicho de otra forma, se busca establecer la dependencia multivaluada $I \rightarrow P$. Como ejemplo, sea $P = \{p1, p2, p3, p4, p5, p6, p7, p8, p9, p10\}$ el conjunto de parámetros. Luego 2^P contendrá 2^{10} elementos. Suponiendo que cada elemento del subconjunto S^{2^P} de 2^P $\{\{p1\}, \{p2\}, \{p2, p3, p4\}, \{p4, p5, p6, p7\}, \{p6, p7, p8, p9, p10\}\}$ permiten formular los indicadores $i1, i2, i3, i4$ e $i5$ respectivamente, entonces un indicador de coherencia como $i6$ basado en los indicadores $i1$ e $i3$ se podrá formular por $\{p1\} \cup \{p2, p3, p4\}$.

```

MODELO_CCPR_R
AREA_CONTEXTO ... FIN_DEFINICION
AREA_COMPONENTES ... FIN_DEFINICION
AREA_PROCESOS
    D.d1; D.d2; D.d3;
FIN_DEFINICION
AREA_RESULTADOS ... FIN_DEFINICION
INDICADORES
    I. i1=(p1);
    I. i2=(p2/2);
    I. i3=(p2*p3*p4);
    I. i4=((p4 - p5) / (p6 + p7));
    I. i5=(p6*0.1 + p7*0.2 + p8*0.3 + p9*0.3 + p10*0.1);
    I. i6=(i1/i3);
    I. i7=(i1-i2);
    I. i8=(i1+i4);
    I. i9=(i4*i5);
FIN_DEFINICION
PARAMETROS
    P. p1; ... P. p10;
FIN_DEFINICION
DIMENSIONES_INDICADORES
    R. d1= (i1);
    R. d2= (i2, i3);
    R. d3= (i4, i5);
FIN_DEFINICION
RC
    R. (d1, d2)=(i6, i7) [RCInterD];
    R. (d1, d3)= (i8) [RCInterD];
    R. (d3)= (i9) [RCIntraD];
FIN_DEFINICION
FIN_MODELO
    
```

Fig. 3 Un modelo CCPR_R.

4. La relación entre parámetros y hechos OLAP se expresa por la dependencia funcional $(P, T) \rightarrow H$. P corresponde al conjunto de p parámetros. T o conjunto de t instantes de tiempo representa la dimensión tiempo de un modelo OLAP. Cada hecho $h \in H$ se ha deducido de la transformación de fuentes de datos. Cada $p \in P$ en un instante de tiempo $t \in T$ se asocia con un elemento de H , según la regla $(P, T) \rightarrow H$.

La Fig. 3 muestra el desarrollo de un modelo CCPR_R. Se excluye la dependencia funcional $(P,T) \rightarrow H$, por representar la relación de un parámetro en el tiempo con el nivel de datos (hechos OLAP) obtenidos del proceso de transformación de las fuentes originales y que en esta etapa de definición no son relevantes para el evaluador. Cabe señalar que modelos como el desarrollado podrán ser compartidos entre evaluadores con el fin de perfeccionarlos.

Implementación de CCPR_R

Hemos implementado CCPR_R como un modelo OLAP utilizando los productos IBM DB2 Universal Database e IBM DB2 OLAP Server (<http://www-306.ibm.com/software/info/university/products/>). Esta implementación corresponde a un diseño multidimensional de base de datos, que exige la definición de variables independientes o dimensionales y variables dependientes o no dimensionales.

La representación OLAP de CCPR_R se basa en su estructura relacional de base de datos cuyas tablas forman parte de un DataWarehouse. Esta estructura relacional la componen las tablas AREAS que representa áreas de CCPR_R, DIMENSIONES o subconjuntos de evaluación, INDICADORES de cada dimensión, PARAMETROS que sustentan cada indicador, IP o asociaciones del tipo N-M producto de la dependencia funcional $I \rightarrow S2^p$ y la tabla HECHOS que almacena los hechos determinados por algún proceso de transformación de fuentes de datos.

En base a estas tablas se establecen las dimensiones del modelo OLAP. Tales dimensiones se han definido con los nombres CCPR_R correspondiente al meta-modelo CCPR_R, TIME que representa la dimensión tiempo, y ACCOUNTS para totalizaciones. La Fig. 4 muestra el diseño OLAP del metamodelo CCPR_R.

La dimensión OLAP CCPR_R permite definir múltiples modelos de evaluación con sus propias dimensiones e indicadores, basados en parámetros y hechos existentes. La recuperación de datos al interior de esta dimensión

corresponde a un JOIN relacional (concatenación de datos en base a atributos comunes) entre las tablas que componen la dimensión. La visualización de los datos en esta dimensión se establece a través de la definición de una “jerarquía” [2] o nexa lógico que comprende los atributos AREAS.NOMBRE, DIMENSIONES.NOMBRE, INDICADORES.NOMBRE y PARAMETROS.NOMBRE.

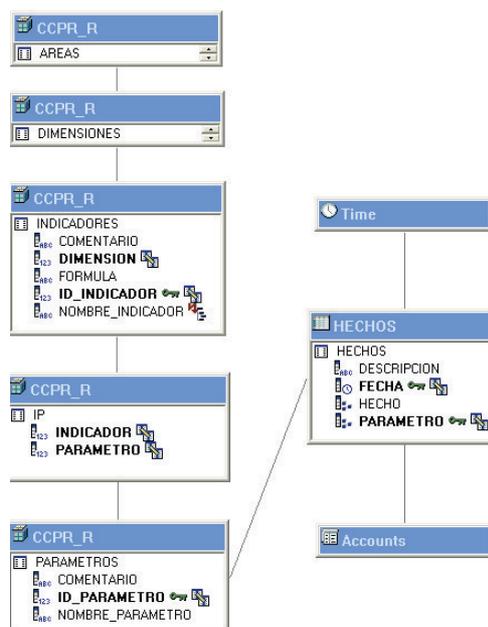


Fig. 4 Diseño OLAP del metamodelo CCPR_R.

También definimos que el servidor OLAP trate las duplicaciones de parámetros producidas por asociaciones N-M de la tabla IP como “miembros compartidos”. Este tratamiento permitirá compartir y visualizar un mismo parámetro en más de un indicador. La formulación de indicadores es implementada automáticamente, al especificar que el indicador sea calculado en base al valor (fórmula) depositado en INDICADORES.FORMULA. La dimensión OLAP TIME implementa una jerarquía sobre el momento en que se ha determinado el hecho, permitiendo visualizar los datos en forma agregada, por ejemplo agrupados por año, mes, semana, día. La dimensión OLAP ACCOUNTS implementa totalizaciones respecto de los hechos existentes.

El metamodelo OLAP diseñado define el DataWarehouse que almacenará modelos CCPR_R incluyendo sus hechos en el tiempo. La visualización de estos modelos se puede efectuar mediante las herramientas de administración de aplicaciones del servidor OLAP, por ejemplo, Fig. 5 muestra el detalle de la dimensión

OLAP CCPR_R, correspondiente al modelo de evaluación desarrollado en la sección 4. En la figura cada área detalla sus dimensiones, indicadores y parámetros asociados. Las relaciones de coherencia son

generalizadas por el elemento “RC” dispuesto a nivel de área, el que a su vez define un elemento representativo de las dimensiones relacionadas y un indicador formulado en base a otros indicadores.

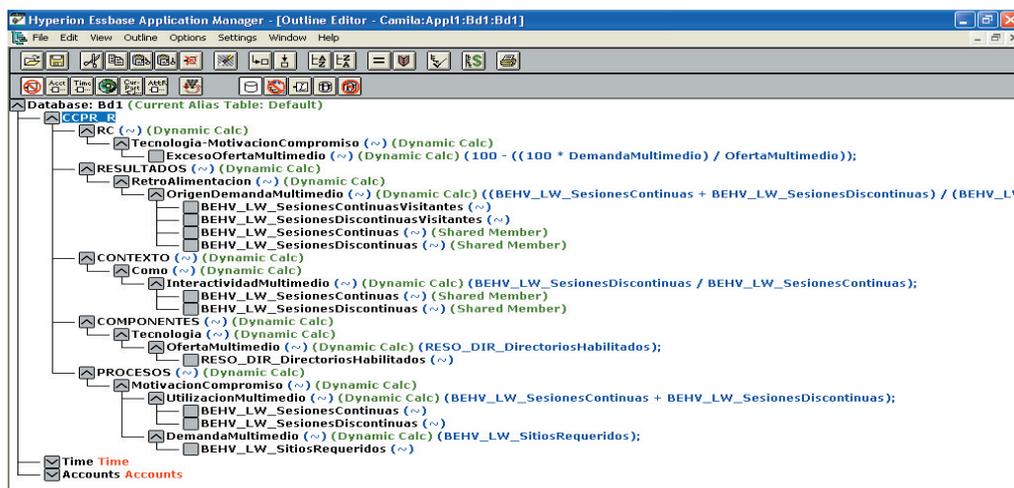


Fig. 5 Detalle de la dimensión OLAP CCPR_R.

DESARROLLO DE UN MODELO DE EVALUACIÓN UTILIZANDO CCPR_R

CCPR_R guía la definición de un modelo de evaluación; sin embargo, para lograr su implementación recomendamos enfocar tal definición desde una perspectiva top down que permita caminar desde lo más abstracto como la determinación de dimensiones, indicadores y relaciones de coherencia a lo más específico como la determinación de parámetros y sus hechos. En esta sección presentamos una forma de abordar la definición de un modelo de evaluación utilizando CCPR_R.

Definición de dimensiones e indicadores

Un modelo CCPR_R requiere establecer dimensiones de evaluación y perfilar sus indicadores de acuerdo a la conceptualización de un ambiente TEL. Para tal efecto adoptamos el enfoque de Goodman [9] que se resume en el siguiente párrafo.

Goodman define una serie de elementos agrupados bajo precondiciones y procesos de cambio de los que podemos obtener retroalimentación. Las precondiciones o elementos contextuales comprenden: 1) definición de “features” generadores de inercia en la universidad, 2) identificación de ambientes de aprendizaje existentes en base a los elementos dimensionales dónde, cuándo, qué y cómo se aprende y 3) considerando las precondiciones

anteriores, definición de la forma de cambio que se quiere producir. Los procesos que se introducen son: 1) planificación, para establecer stakeholders y alinear la organización, tecnología e individuos para el cambio, 2) implementación de conceptos, métodos y nuevos ambientes de aprendizaje. Este proceso se divide en motivación-compromiso o estímulos para moverse de un estado a otro, socialización que responde a cómo encajar el tipo de conocimiento requerido con los mecanismos para impartir tal conocimiento y retroalimentación-rediseño enfocado en la construcción de medidas respecto de comportamientos y consecuencias y 3) institucionalización para lograr que el cambio persista en el tiempo. Este proceso se divide en motivación-compromiso, socialización, difusión y retroalimentación-rediseño enfocados en sostener el cambio. Difusión responde al cambio gradual que se extiende por la organización legitimándolo y generando infraestructura.

El ambiente TEL descrito nos permite establecer que cada uno de los elementos agrupados bajo precondiciones y procesos de cambio corresponda a las dimensiones de un modelo CCPR_R, tal como lo presenta la Fig. 6.

Observamos que la necesidad de retroalimentación expuesta en los procesos de implementación e institucionalización se puede materializar al establecer indicadores y relaciones de coherencia RCIntraD y RCInterD. La flexibilidad del metamodelo CCPR_R también permite establecer indicadores respecto del

proceso de planificación y precondiciones, complementando aquella proveniente de las dimensiones

motivación-compromiso, socialización, difusión y rediseño. Por otra parte el área “Resultados”

<p>MODELO_CCPR_R AREA_CONTEXTO D. Features; D. Donde; D. Cuando; D. Que; D. Como; D. FormaCambio; FIN_DEFINICION AREA_COMPONENTES D. Stakeholders; D. Organizacion; D. Tecnologia; D. Individuo; FIN_DEFINICION</p>	<p>AREA_PROCESOS D. MotivacionCompromiso; D. Socializacion; D. Difusion; D. Rediseño; FIN_DEFINICION AREA_RESULTADOS D. Retroalimentación FIN_DEFINICION RC R. TecnologiaMotivacionCompromiso= (ExcesoOfertaMultimedia); FIN_DEFINICION</p>	<p>PARAMETROS ... FIN_DEFINICION DIMENSIONES_INDICADORES R. Como= (InteractividadMultimedia); R. Tecnologia= (OfertaMultimedia); R. MotivacionCompromiso= (DemandaMultimedia, UtilizacionMultimedia); R. Retroalimentacion=(OrigenDemandaMultimedia); FIN_DEFINICION FIN_MODELO</p>
<p>INDICADORES [Un multimedia como el sitio Web de curso, es un componente tecnológico del ambiente TEL y agrupa medios de apoyo como el material de clases, noticias, foros, etc.] I. DemandaMultimedia= (BEHV-LW-SitiosRequeridos) [Cantidad de multimedios demandados por estudiantes de la universidad (≥ 0)]; I. OfertaMultimedia= (RESO-DIR-DirectoriosHabilitados) [Cantidad de multimedios dispuestos a servir de apoyo al proceso de E/A (≥ 0)]; I. UtilizaciónMultimedia= (BEHV-LW-SesionesContinuas + BEHV-LW-SesionesDiscontinuas) [Nivel de utilización del multimedia basado en sesiones de trabajo del estudiante (≥ 0)]; I. InteractividadMultimedia= (BEHV-LW-SesionesDiscontinuas / BEHV-LW-SesionesContinuas) [Grado de interactividad entre el estudiante y el multimedia (≥ 0). Valores cercanos a 0 indica interacciones puntuales generalmente de corta duración. Valores superiores a 1 indica mayor interactividad con el multimedia en el tiempo]; I. OrigenDemandaMultimedia= ((BEHV-LW-SesionesContinuas + BEHV-LW-SesionesDiscontinuas) / (BEHV-LW-SesionesContinuasVisitantes + BEHV-LW-SesionesDiscontinuasVisitantes)) [Explica el origen la demanda por multimedios ($0 \geq \text{Origen} \leq 1$). Valor igual a 1 representa demanda 100% originada por visitantes autenticados, generalmente estudiantes]; I. ExcesoOfertaMultimedia= (100 - ((100 * DemandaMultimedia) / OfertaMultimedia)) [Relación de coherencia que establece el exceso de oferta de multimedios ($0\% \geq \text{exceso} \leq 100\%$)]; FIN_DEFINICION</p>		

Fig. 6 Dimensiones e indicadores de un modelo de evaluación.

de CCPR_R podrá definir otros indicadores que ayuden a responder interrogantes del tipo ¿en qué medida el cambio ha contribuido a aumentar el nivel de empleo de los estudiantes?

Encontramos técnicamente factible formular algunos indicadores para las dimensiones, como, Tecnología, MotivacionCompromiso y Retroalimentación. La Fig. 6 expone una primera versión del modelo logrado (comentarios entre “[“ y “]”) cuyo conjunto de parámetros se explican en la siguiente sección.

Definición de parámetros

La definición de dimensiones e indicadores formulados en base a parámetros exige establecer las fuentes de datos a transformar. Decidimos utilizar las fuentes de datos Log de Web cuya conceptualización Entidad Relación se presenta en la Fig. 7 y la estructura de directorios que habilita sitios en un servidor Web de cursos.

La pertinencia de estas fuentes fue analizada de acuerdo al modelo arquitectónico de sistemas de aprendizaje

soportados por tecnologías LTSA, el cual se encuentra definido en la tercera capa del estándar IEEE P1484.1/ D11 [12]. Este estándar se ha declarado neutral respecto de contenidos, pedagogías, culturas, implementaciones y plataformas.

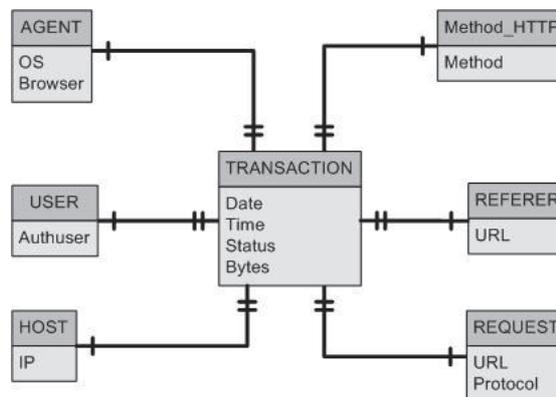


Fig. 7 Conceptualización E-R del Log de Web.

...
PARAMETROS

[Generalidades:

- 1) La fuente Log de Web corresponde a un archivo plano que registra según formato Common Log Format o Extended Common Log Format, la actividad producida al resolver automáticamente los vínculos de una página Web, estableciéndose una comunicación entre computadores cliente y servidores, basada en el protocolo HTTP.

Definiciones propias:

- 1) Sesión. Agrupación de transacciones HTTP cercanas en el tiempo para un Visitante.
 - 2) Visitante. Mismo IP-SITIO-LOGIN.
 - 3) Autorizado. Visitante autorizado a ingresar a un sitio por parte del servidor Web según su autenticación.
 - 4) Continuidad. La continuidad se mantiene en una secuencia de n transacciones HTTP $t_1, t_2, t_3, \dots, t_n$ si la diferencia de tiempo en que ocurre t_i y t_{i+1} es a lo más de 5 minutos para un Visitante, caso contrario (diferencia sobre 5 minutos) la secuencia se define como discontinua]
- P. BEHV-LW-SesionesContinuasVisitantes [Sesiones continuas generadas por todos los visitantes];
P. BEHV-LW-SesionesDiscontinuasVisitantes [Sesiones discontinuas generadas por todos los visitantes];
P. BEHV-LW-SesionesContinuas [Sesiones continuas generadas por visitantes autorizados];
P. BEHV-LW-SesionesDiscontinuas [Sesiones discontinuas generadas por visitantes autorizados];
P. RESO-DIR-DirectoriosHabilitados [Directorios de sitios Web de cursos habilitados];
P. BEHV-LW-SitiosRequeridos [Sitios consultados por visitantes autorizados];
P. BEHV-LW-SitiosRequeridosVisitantes [Sitios consultados por visitantes];

FIN_DEFINICION

...

Fig. 8 Parámetros de un modelo de valuación.

Luego, el modelo CCPR_R que fuera iniciado en la sección anterior se podrá completar mediante la definición de parámetros construidos sobre la fuente Log de Web caracterizada por el componente LTSA BEHV (behavior) y la estructura de directorios de sitios Web caracterizada por el componente LTSA RESO (learning resources). Cabe señalar que LTSA BEHV representa el comportamiento de un LTSA LENT (learner entity), o abstracción de un “learner” humano que aprende en forma individual o colaborativa y que interactúa con los medios en el contexto de su experiencia de aprendizaje.

Una vez que se ha conceptualizado y entendido la estructura de datos de cada fuente, es factible especificar un conjunto de parámetros para la formulación de los indicadores que se hayan definido. Fig. 8 expone los parámetros que complementan el modelo CCPR_R iniciado.

Cabe señalar que desde el punto de vista de la implementación, cada parámetro representa un proceso de transformación algorítmico de alguna fuente de datos, resultando en un conjunto de hechos OLAP en el tiempo. La siguiente sección expone el proceso de transformación de datos necesario para el desarrollo de los parámetros más relevantes.

Transformación de datos

La implementación de parámetros como los especificados anteriormente implica la ejecución de un proceso de Extracción, Transformación y Carga (ETL)

respecto del componente LTSA BEHV Log de Web y RESO directorios. Esto es, una vez que el proceso de extracción deja disponible la fuente de datos para posterior manipulación, es necesario transformarla y cargar los resultados de la transformación en el DataWarehouse definido. En el tratamiento del Log de Web, este proceso utiliza un área de trabajo en la que se depositan temporalmente datos como IP, login, timestamp de fecha y hora expresada en segundos, método HTTP, sitio (multimedia), funcionalidad (medio), referer, bytes, extensión del archivo requerido, código de retorno HTTP.

En la extracción, los datos son llevados desde las fuentes originales, que contienen principalmente transacciones HTTP GET, al área de trabajo descrita. Esta función se concreta mediante un servicio escrito en PERL (<http://www.perl.com/>), el que formatea un archivo de Log de Web comprimido correspondiente a un día de Log y produce un Log de Web formateado de acuerdo a la estructura de datos del área de trabajo. El archivo original es analizado línea a línea por el servicio, identificando mediante una expresión regular (<http://www.perldoc.com/>) los datos a extraer.

Respecto de la transformación de datos, definimos que este proceso corresponda a reducir datos donde sea posible o extraer otros nuevos a partir de datos existentes, definición que también es utilizada en [7]. Es así que la determinación de hechos OLAP para cuatro de los parámetros relacionados a LTSA BEHV Log de Web corresponde a una secuencia de acciones algorítmicas basada en scripts SQL y un modelo de

cluster demográfico (Ver Anexo Técnicas), de la siguiente forma:

1. BEHV-LW-SesionesContinuasVisitantes
SCRIPT1; SCRIPT2; SCRIPT3; SCRIPT4.
2. BEHV-LW-SesionesDiscontinuasVisitantes
SCRIPT1; SCRIPT2; SCRIPT3; SCRIPT5;
CLUSTER.
3. BEHV-LW-SesionesContinuas
SCRIPT6; SCRIPT2; SCRIPT3; SCRIPT4.
4. BEHV-LW-SesionesDiscontinuas
SCRIPT6; SCRIPT2; SCRIPT3; SCRIPT5;
CLUSTER.

La carga de hechos OLAP al DataWarehouse, producto de la secuencia algorítmica definida, se efectúa mediante SQL. Cabe señalar que los hechos OLAP fueron establecidos al analizar una muestra de 30 archivos de Log de Web, correspondiente a septiembre del año 2003.

Visualización de resultados

En la visualización de resultados se ha utilizado un “add-in” en MS Excel, permitiendo “navegar” por el modelo OLAP y graficar información de apoyo a la interpretación de indicadores.

A	B	C	D
		Accounts	
2003-Sep	ExcesoOfertaMultimedia	49,26 %	
	OrigenDemandaMultimedia	0,93 (0-1)	
	InteractividadMultimedia	0,42 (>= 0)	
	OfertaMultimedia	2.488 Sitios	
	UtilizacionMultimedia	9.246 Sesiones	
	DemandaMultimedia	1.262 Sitios	

Fig. 9 Indicadores del mes de septiembre.

La Fig. 9 muestra los indicadores de la dimensión OLAP CCPR_R agrupados a nivel del elemento “2003-Sep” de la dimensión OLAP TIME. Fig. 10 detalla el elemento “2003-Sep-Sem1-Dia3”. La gráfica en Fig. 11 se basa en la completa expansión de la dimensión OLAP TIME y los indicadores de la dimensión OLAP CCPR_R. Información como la recopilada en estas figuras deberá ser interpretada por el tomador de decisiones en el contexto de las áreas y dimensiones de evaluación del modelo (días 18 y 19 corresponden a Fiestas Patrias en Chile).

CONCLUSIONES

Este artículo describió la implementación OLAP de CCPR_R, un metamodelo para el desarrollo de modelos CIPP de evaluación, que incorporan la definición de dimensiones e indicadores, además de parámetros y hechos OLAP.

CCPR_R podrá ser utilizado como un componente fundamental de un DSS de apoyo a la toma de decisiones, en el contexto de un ambiente TEL.

A	B	C
		Accounts
Sem1-Dia1	CCPR R	
Sem1-Dia2	CCPR R	
Sem1-Dia3	ExcesoOfertaMultimedia	43,73
	Tecnologia.MotivacionCompromiso	
	RC	
	BEHV LW SesionesContinuasVisitantes	8.114
	BEHV LW SesionesDiscontinuasVisitantes	2.664
	BEHV LW SesionesContinuas	7.140
	BEHV LW SesionesDiscontinuas	2.495
	OrigenDemandaMultimedia	0,89
	RetroAlimentacion	
	RESULTADOS	
	BEHV LW SesionesContinuas	7.140
	BEHV LW SesionesDiscontinuas	2.495
	InteractividadMultimedia	0,35
	Como	
	CONTEXTO	
	RESO DIR DirectoriosHabilitados	2.488
	OfertaMultimedia	2.488
	Tecnologia	
	COMPONENTES	
	BEHV LW SesionesContinuas	7.140
	BEHV LW SesionesDiscontinuas	2.495
	UtilizacionMultimedia	9.635
	BEHV LW SitiosRequeridos	1.400
	DemandaMultimedia	1.400
	MotivacionCompromiso	
	PROCESOS	
	CCPR R	
Sem1-Dia4	CCPR R	
Sem1-Dia5	CCPR R	
Sem1-Dia6	CCPR R	

Fig. 10 Indicadores del 3 de septiembre.

La independencia en la formulación de un indicador respecto de los parámetros en que se basa permite que un modelo CCPR_R pueda ser compartido entre evaluadores, sin perder la información obtenida de los procesos de transformación de fuentes de datos, cuando sea necesario reformular algún indicador. Además CCPR_R puede reunir en un indicador la información extraída al utilizar más de una técnica sobre múltiples fuentes de datos presentes en un ambiente TEL.

La experimentación nos llevó a concluir que la definición de un modelo de evaluación debe basarse en la conceptualización de un ambiente TEL, aspecto que ayuda a precisar la definición de dimensiones y facilita la determinación de indicadores. Esta conceptualización, que es necesaria para establecer lo que se entiende por un ambiente TEL, aportará un lenguaje común de entendimiento cuando el modelo de evaluación deba ser

compartido. Adicionalmente, la determinación de la arquitectura del componente tecnológico del ambiente TEL ayudará a precisar las fuentes de datos sobre las que se desarrollarán los parámetros CCPR_R. De este análisis también se podrá desprender la necesidad de ajustar el componente tecnológico que debiera acompañar al cambio que se quiere institucionalizar. Estos ajustes se podrán traducir en nuevos desarrollos para atender las funcionalidades básicas que demanda el modelo arquitectónico requerido, incluyendo la generación de fuentes de retroalimentación para concretar la evaluación.

Si bien el experimento se concentró en el mes de septiembre, algunas conclusiones pueden ser obtenidas de la información que proveen los indicadores desarrollados. Por ejemplo, el indicador DemandaMultimedia muestra un típico comportamiento durante una semana de trabajo y un fin de semana. El indicador OfertaMultimedia, que permanece constante durante el semestre, refleja una política de administración que deja disponibles sitios para profesores y estudiantes en función de la cantidad de cursos que se dictan. Es así que el indicador ExcesoOfertaMultimedia muestra valores cercanos a un 49%. Esta situación ha sido corregida al rediseñar

sitios que atienden cursos especiales, por ejemplo aquellos que abarcan más de un período académico. El indicador OrigenDemandaMultimedia con valores cercanos a 0.9 muestra que si bien son estudiantes los que requieren multimedia, el resto son visitantes externos buscando material. Este indicador nos ha motivado a revisar las políticas de administración, flexibilizándolas gradualmente, para permitir que los visitantes en general accedan a los multimedios dispuestos en la Intranet universitaria. El indicador UtilizaciónMultimedia muestra periodicidad respecto del uso del multimedia y es consistente con el indicador DemandaMultimedia. El indicador InteractividadMultimedia con valores cercanos a 0.4 muestra que los estudiantes realizan interacciones puntuales con el multimedia, como lectura de noticias, consulta de calificaciones y descarga de material.

Como trabajo futuro, se continuará potenciando el modelo de evaluación experimental planteado, definiendo nuevos indicadores, perfeccionando los existentes e incluyendo nuevas fuentes de datos. Con esta acción se espera brindar un mejor apoyo a la toma de decisiones respecto de un ambiente TEL al interior de PUC y por tanto contribuir al mejoramiento de los procesos de enseñanza y aprendizaje potenciados por tecnología.

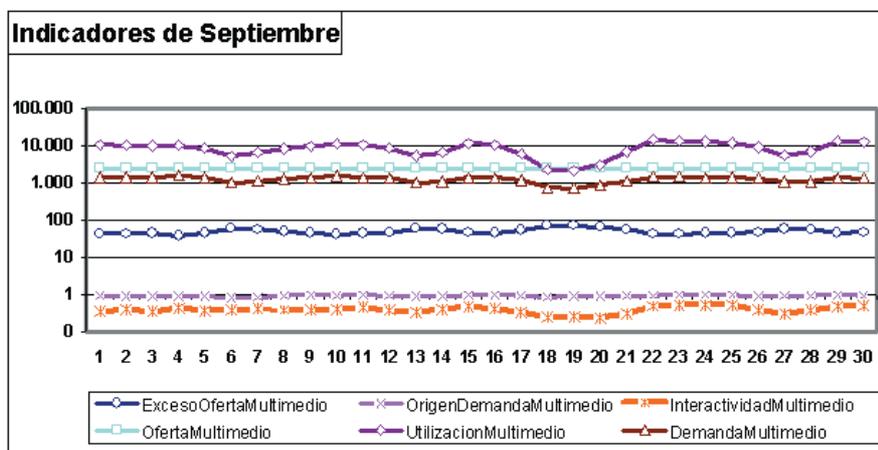


Fig. 11 Representación gráfica de indicadores.

REFERENCIAS

[1] I. Casas, A. Vergara. "CursoWebPUC Una plataforma de apoyo a la educación presencial". Actas del II Congreso Europeo de Tecnologías de la Información en la Educación y la Ciudadanía, Barcelona, España & Actas del II Congreso Internacional "Docencia Universitaria e Innovación", Tarragona, España, 2002.

[2] S. Chaudhuri, U. Dayal. "An overview of data warehousing and OLAP technology". ACM Sigmod Record, Vol. 26(1), pp. 65-75, 1997.

[3] C.J. Date; "An introduction to database systems"; Addison-Wesley, 2000.

- [4] A. De La Orden, I. Asensio, R. Carballo, J. Fernández, A. Fuentes, J. M. García, S. Guardia. “Desarrollo y validación de un modelo de calidad universitaria como base para su evaluación”. RELIEVE Revista Electrónica de Investigación y Evaluación Educativa, Vol. 3(1), http://www.uv.es/RELIEVE/v3n1/RELIEVEv3n1_2.htm, 1997.
- [5] W. Eckerson, C. White. “The Data Warehousing Institute (TDWI) Report Series, Evaluating ETL and Data Integration Platforms”. <http://www.dw-institute.com>, 2003.
- [6] T. Escudero. “From tests to current evaluative research. One century, the XXth, of intense development of evaluation in education”. RELIEVE Revista Electrónica de Investigación y Evaluación Educativa, Vol. 9(1), http://www.uv.es/RELIEVE/v9n1/RELIEVEv9n1_1.htm, 2003.
- [7] A. Fong, S. Hui, G. Jha. “Data mining for decision support”. IEEE IT Pro, March/April, 2002.
- [8] L. García. “Indicators for the evaluation of the teaching in a distance university”. Revista Iberoamericana de Educación, Vol. 1(1), 1998.
- [9] P. S. Goodman. “Creating organizational and technological change”. En P. S. Goodman, “Technology enhanced learning: Opportunities for change”; Lawrence Erlbaum Associates Inc., 2002.
- [10] J. Han, M. Kamber. “Data mining concepts and techniques”; Academic Press, 2001.
- [11] B. Hayes. “Cómo medir la satisfacción del cliente – diseño de encuestas, uso y métodos de análisis estadístico”; Oxford University Press México, S.A., 1999.
- [12] IEEE P1484.1/D11. “Draft standard for learning technology – Learning Technology Systems Architecture (LTSA)”, <http://jtc1sc36.org/doc/36N0388.pdf>, 2002.
- [13] R. McLeod. “Sistemas de información gerencial”, 7ª ed.; Prentice-Hall Hispanoamericana S.A., 2000.
- [14] M. J. Rubio. “Focus and models of evaluations of the e-learning”. RELIEVE Revista Electrónica de Investigación y Evaluación Educativa, Vol. 9(2), http://www.uv.es/RELIEVE/v9n2/RELIEVEv9n2_1.htm, 2003.

- [15] D. L. Stufflebeam. “The CIPP model for evaluation”. Annual Conference of the Oregon Program Evaluators Network (OPEN), <http://www.wmich.edu/evalctr/pubs/CIPP-ModelOregon10-03.pdf>, 2003.

ANEXO TÉCNICAS

SCRIPT1. Reduce la generación de sesiones innecesarias dada la cercanía en el tiempo (1 minuto) de transacciones del grupo HTTP 400 y transacciones autenticadas (generalmente con código HTTP 200):

```
UPDATE ODS.LOGWEB X SET LOGIN = (SELECT
MAX(LOGIN) FROM ODS.LOGWEB Y WHERE X.IP = Y.IP
AND X.SITIO = Y.SITIO AND ABS(X.TIMESTAMP -
Y.TIMESTAMP) BETWEEN 0 AND 60) WHERE STATUS >=
400 AND STATUS <= 499;
```

SCRIPT2. Transacciones con diferencia de 5 minutos a lo más, se agrupan:

```
DELETE FROM ODS.LOGWEB X WHERE EXISTS (SELECT
* FROM ODS.LOGWEB Y WHERE X.IP = Y.IP AND X.SITIO
= Y.SITIO AND X.LOGIN = Y.LOGIN AND
ABS(X.TIMESTAMP - Y.TIMESTAMP) BETWEEN 1 AND 300
AND X.TIMESTAMP < Y.TIMESTAMP);
```

SCRIPT3. Eliminación de algunas transacciones con igual TIMESTAMP para un mismo visitante:

```
DELETE FROM ODS.LOGWEB X WHERE EXISTS (SELECT
* FROM ODS.LOGWEB Y WHERE X.IP = Y.IP AND X.SITIO
= Y.SITIO AND X.LOGIN = Y.LOGIN AND X.TIMESTAMP =
Y.TIMESTAMP AND X.STATUS < Y.STATUS);
```

SCRIPT4. Informa sobre agrupaciones continuas (sesiones):

```
SELECT IP CONCAT SITIO CONCAT LOGIN FROM
ODS.LOGWEB GROUP BY IP, SITIO, LOGIN HAVING
COUNT(*) = 1;
```

SCRIPT5. Eliminación de agrupaciones continuas desde el ODS. Sobre las agrupaciones restantes, CLUSTER determinará la cantidad de sesiones para un mismo IP-SITIO-LOGIN en un día de log:

```
DELETE FROM ODS.LOGWEB WHERE IP CONCAT SITIO
CONCAT LOGIN IN (SELECT IP CONCAT SITIO CONCAT
LOGIN FROM ODS.LOGWEB GROUP BY IP, SITIO, LOGIN
HAVING COUNT(*) = 1);
```

SCRIPT6. Eliminación de transacciones grupo HTTP 400 y la posibilidad de determinar sesiones de este tipo:

```
DELETE FROM ODS.LOGWEB X WHERE X.STATUS >= 400
AND X.STATUS <= 499;
```

MODELO DE CLUSTER. Establece sesiones de trabajo en torno a un multimedia. Su formulación se basa en el algoritmo de cluster demográfico implementado por IBM Intelligent Miner (<http://www-306.ibm.com/software/data/iminer/library.html>):

- Maximum Passes= 2 (mejoramiento de la calidad del cluster obtenido en función del número de pasos).
- Maximum Clusters= 99999 (número máximo de clusters a crear).
- Accuracy improvement = 2 (el proceso de iteración termina cuando el mejoramiento de la calidad entre dos pasos es menor que el % especificado).
- Similarity Threshold = 0,75 (limita los valores aceptados en un cluster. Si el % de similitud especificado es alto se generaran más clusters).
- Active Field (selección de campos a considerar en el minado): TIMESTAMP (Continuous), Concatenación IP, SITIO, LOGIN (Categorical).

- Supplementary field (selección de campos informativos): HORA (Time).
- Field Weight (asignación de más o menos peso a una relación entre campos): TIMESTAMP = 1, IP_SITIO_LOGIN= 1.
- Units / Distance Factor (Distancia expresada en la unidad especificada. Permite determinar inclusión o no de un registro en un cluster por diferencia absoluta de los valores comparados. Distancias pequeñas implica mayor número de cluster):
TIMESTAMP= Absolute number / 2.700 (45 minutos).

Las medidas de similitud respecto de un elemento con otro elemento en un cluster (Condorcet) fueron 0.74 valor mínimo, 0.88 valor máximo, 0.85 valor moda. Un valor igual a 1 en el rango 0 a 1, indica que todos los elementos en un cluster son idénticos y que no existen elementos con las mismas características fuera del cluster.