

Modelo conceptual para el diseño e implementación del sitio web de un museo regional*

Resumen

El presente artículo propone el modelo conceptual para la creación de un sitio Web de un museo regional multidisciplinario que no tiene objetos conocidos mundialmente, un museo común. Con base en la misión de un museo virtual regional se han desarrollado los principales modelos de grupos de usuarios y las características formales del sitio. Se ha realizado un muestreo de sitios exitosos, además del análisis de clusters de estos sitios basándose en las características que los integran y generando dos principales grupos. De acuerdo con la misión del Museo Virtual del Museo Regional de Huajuapán (MUREH), para la construcción de sus páginas Web se ha elegido el enfoque del primer grupo de sitios exitosos que está dirigido a sitios Web de museos multidisciplinarios.

1. Introducción y formulación del problema

Existen diversos sitios donde se puede encontrar información sobre la Cultura Mixteca [1,2,3], pero no existe información sistemática respecto a la construcción de estos sitios y no contienen tanta información como la que tiene el MUREH, centro de información de la Cultura Mixteca.

El objetivo de crear un sitio Web de un museo regional, es el de servir de apoyo informático del museo real y al mismo tiempo ampliar las posibilidades de difusión de la cultura Mixteca.

Tomando en cuenta esta misión, es necesario identificar el tipo de usuario que tiene interés sobre los materiales relacionados con la cultura Mixteca. Considerando las motivaciones y restricciones de acceso a Internet, se

han identificado tres tipos de modelos de usuarios: los jóvenes de la región Mixteca, adultos que tienen raíces mixtecas pero viven fuera de la región en lugares donde tienen fácil acceso a Internet e investigadores cuya especialidad está relacionada con la cultura.

Estos modelos de usuarios aunado a la misión del Museo Virtual son factores principales en la construcción del sitio Web de un museo regional, a estos factores es necesario añadirles los elementos que han influido para que un sitio Web sea exitoso.

No existe ninguna metodología o recomendación general sobre la construcción de sitios Web exitosos, especialmente para sitios que no fueron creados con fines comerciales o que no tienen objetos o herramientas que llamen la atención de los usuarios. La utilización de algunas técnicas comerciales puede atraer la atención de usuarios que no coinciden con el modelo de usuario al que está enfocado el sitio Web. Por esta razón, es necesario investigar qué atributos se deben integrar en los sitios Web exitosos de los museos regionales.

2. Sistema de atributos de sitios Web exitosos de museos regionales

Se ha elaborado un sistema de atributos para caracterizar los sitios Web de museos regionales, este sistema de atributos ha sido construido con la finalidad de revelar los atributos necesarios para la integración de un sitio Web enfocado a un museo regional.

Los parámetros establecidos en la estructura del sitio Web dependen de la misión y modelo de usuario, así como del contenido que integre al sitio Web. No se ha considerado en la lista de parámetros el contenido de los sitios analizados, debido a que realmente no se puede elegir algún contenido para el sitio Web del MUREH el cual tiene como finalidad la Cultura Mixteca.

*Trabajo parcialmente apoyado por el Consejo Nacional de Ciencia y Tecnología (CONACYT), proyecto con clave 51777

Los parámetros elegidos para la comparación de sitios Web, son parámetros simples tales como: el nombre de la máquina de búsqueda, existencia de logo, metatag, cantidad de colores usados en la primera página, cantidad de tonos, el tamaño de las letras del texto, espacio ocupado por las imágenes en la página principal, espacio libre alrededor del texto, imágenes utilizadas en las páginas de segundo nivel, texto incluido en el segundo nivel. Existen dos parámetros complejos pero importantes en el desarrollo de la investigación los cuales son: el posicionamiento asignado por las máquinas de búsqueda y la estructura de hipervínculos del sitio Web. Estos dos últimos parámetros son difíciles de medir y por esta razón se ha desarrollado un algoritmo para su análisis.

Se puede decir que los parámetros antes mencionados muestran una dependencia entre ellos, algunos con mayor claridad para su interpretación que otros, por ejemplo si se tiene mayor tamaño de texto entonces existe menos espacio libre y menos espacio para incluir imágenes.

En el caso del parámetro complejo para la medición del criterio de éxito del sitio Web se ha establecido el enfoque de que para el muestreo solamente se consideren sitios que tienen el posicionamiento más alto en los resultados generados por las máquinas de búsqueda, resultados que se obtienen mediante la formulación adecuada de palabras claves de búsqueda. Las palabras utilizadas en las búsquedas han sido realizadas en diferentes idiomas, usando una traducción literal de éstas y que tengan el mismo sentido. La frase "museo regional" no tiene una traducción literal en el idioma ruso que corresponda al mismo sentido, por lo tanto en la búsqueda se ha utilizado la frase "museo de estudios de región" (KRAYEVEDCHESKY MUZEY) que es lo más adecuado como traducción de la frase y con el mismo sentido.

Se han elegido máquinas de búsqueda que tienen el mismo principio de posicionamiento de los sitios Web, este criterio es el que establece el grado de éxito del sitio Web mediante la cantidad de hipervínculos con otros sitios, es decir, el número de sitios que tienen un enlace dirigido hacia el sitio evaluado.

Estos principios de posicionamiento lo utilizan entre otros las máquinas de búsqueda Google, Yandex y Yahoo, por esta razón se han buscado candidatos al muestreo de sitios Web generados por estas tres má-

quinas de búsqueda, búsquedas realizadas en tres idiomas: inglés, español y ruso. Los sitios Web en inglés y español, han sido considerados como candidatos a miembros del muestreo en una forma más abstracta por los usuarios principales del modelo del sitio Web del MUREH, debido a que no muchos futuros usuarios del sitio Web a desarrollar entenderán ruso e inglés. El idioma ruso no es tan importante en este caso por que todos los sitios incluidos en el muestro tienen una versión de su sitio Web en inglés y algunas veces en un tercer idioma.

Los resultados generados por las máquinas de búsqueda han sido integrados en una lista común, en el caso de los sitios encontrados en dos o más máquinas de búsqueda para cada uno de ellos se acepta el nivel máximo de sus posicionamientos, para otros sitios que se encuentran solamente en una máquina se acepta el nivel que corresponde a la posición de este sitio en comparación con el posicionamiento de los sitios encontrados en diversas máquinas.

Con respecto al segundo parámetro complejo, el desarrollo de la estructura o el número de hipervínculos que integran el sitio Web en el primer y segundo nivel, se ha medido utilizando un método basado en el Edit Tree Algorithm [4]. La idea principal de este algoritmo es medir la distancia entre dos estructuras de tipo árbol, donde la página principal es la raíz, las aristas son las páginas de segundo y tercer orden, etc. La medición de la distancia entre esta estructura de tipo árbol con respecto a la estructura de tipo árbol ideal propuesto para el MUREH, se realiza con la finalidad de establecer la discrepancia en las estructuras, en las diferentes opciones que proponen en los diferentes niveles de los sitios Web.

Como estructura ideal del sitio Web de un museo regional, se propone una distribución de páginas en forma de árbol donde los nodos del primer y segundo orden tienen un grado de salida de 13, es decir de cada nodo de primer y segundo orden salen 13 aristas hacia el nivel inferior.

Para medir esta distancia se considera la discrepancia existente en estos árboles, diferencia de nodos hijos, nodo padre, hijos con nombres iguales, nodos en subpáginas con el mismo nombre en un nivel padre y en un nivel hijo, es decir si en un nivel algunos nodos hijos son padre en otro nivel, esta situación indicaría la discrepancia de posicionamiento del nodo en la estructura.

Este enfoque ha sido simplificado, debido a que los nombres de los nodos del tercer nivel por lo general son muy diferentes y dependen del contenido de cada sitio Web, contenido que todavía no es considerado para el desarrollo de la investigación.

Por otra parte no fueron descubiertos tantos casos en los cuales los nodos cambien de nivel en algunas subpáginas, por esta razón este criterio se redujo sólo al cálculo de la cantidad de nombres de las opciones integradas en la primer página de la página ideal y de la página analizada, así como la suma de grados de nodos del segundo nivel (nodos hijos).

3. Muestreo de sitios Web exitosos

Basándose en el conjunto de atributos antes mencionados se han medido los parámetros del muestreo integrado por:

PALABRAS DE BÚSQUEDA	MAQUINAS DE BÚSQUEDA	
museo regional	Google.mx (30 sitios)	Yahoo.mx (30 sitios)
regional museum	Google (30 sitios)	Yahoo (30 sitios)
и краеведческий музей (regional study museum)	Google (30 sitios)	Yandex.ru (30 sitios)

La suma total de los sitios Web es menor a 180 debido a que existen varias repeticiones de sitios Web exitosos, por otro lado algunos sitios Web incluyen en su contenido las palabras clave utilizadas en las búsquedas pero no son sitios de museos, son solamente sitios donde hay una página dedicada a un museo regional, por esta razón se reunieron 100 sitios para llevar a cabo el análisis detallado.

Entre estos sitios también existen sitios Web cuyo contenido también está integrado sólo por una única página, un sitio Web muy básico de un museo regional y que no se puede considerar como un ejemplo a seguir y fueron descartados, por lo tanto se obtiene finalmente una tabla de 46 sitios Web con 15 atributos cada uno.

4. Resultados del análisis de clusters del muestreo

Con los sitios Web y los atributos de cada uno se ha integrado la matriz objetos-tributos. La matriz contiene atributos lógicos (con valores 1 y 0) y atributos numéricos los cuales fueron normalizados utilizando valores comprendidos entre 1 y 0 utilizando la siguiente fórmula:

$$ValorNormalizado = \frac{ValorOriginal - ValorMínimo}{ValorMáximo - ValorMínimo} \quad (1)$$

Los valores de la fórmula 1 fueron elegidos del total de valores de cada atributo, generando la matriz normalizada con todos sus valores comprendidos entre 0 y 1. Esta matriz ha sido analizada con métodos de análisis de clusters [5], con los cuales se agruparon los parámetros en grupos que corresponden a los clusters de sitios y en el que cada cluster tiene alguna característica especial del conjunto de parámetros considerados en el análisis.

Del total de parámetros, es necesario que algunos de éstos no sean considerados debido a que no proporcionan ninguna información importante por ejemplo, al realizar el análisis de clusters el primer clustering genera tres clusters que coinciden con la lengua materna utilizada en el desarrollo de los sitios Web, lo cual no es interesante ya que el idioma de cada sitio puede identificarse al ver el nombre del sitio; por otro lado fueron identificados otros parámetros que tienen una distribución casi uniforme en todos los clusters de los sitios, estos parámetros también fueron omitidos.

Como resultado del análisis de los clusters se han obtenido dos grupos de sitios que tienen rasgos opuestos, en el espacio de los factores 2 y 3 que son ortogonales y son los más grandes (en el caso del factor 1 el cual genera solamente los valores promedio para cada atributo y por esta razón que no se ha considerado), el factor 2 es la combinación lineal de la cantidad de colores con signo (-), la cantidad de tonos con signo (+) y el espacio de texto en la pagina principal con signo (-), los otros parámetros casi no participan en la separación de los grupos. En el caso del factor 3, es la combinación de los atributos: tamaño de letras con signo (-), espacio de imagen en la primer página con signo (+) y la distancia como grado de desarrollo de la estructura del sitio con signo (-).

En la Figura 1 se muestra con la línea continua el factor y con la línea punteada el representante del cluster, se puede ver la representación gráfica del factor 2 en la parte superior y el factor 3 en la parte inferior con dos sitios Web que pertenecen al primer grupo "Tomsk" y "Mogilev" y dos sitios Web que pertenecen al segundo grupo "NewCastle" y "Southbend".

En la Figura 2 se puede observar que los sitios "Tomsk" y "NewCastle" casi no tienen componentes en el factor 3 pero tienen características opuestas en el factor 2; al mismo tiempo "Southbend" casi no tiene componentes del factor 2 pero representa valores po-

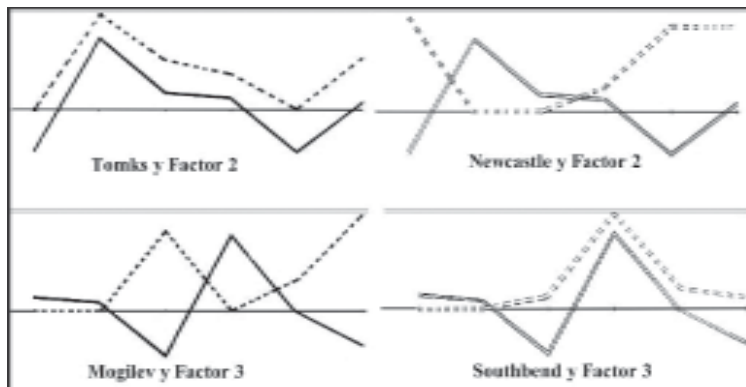


FIGURA 1. FACTOR 2 Y FACTOR 3 CON REPRESENTANTES DE CLUSTERS

sitios grandes del factor 3 y "Mogilev" casi no tiene componentes del factor 2 pero representa valores negativos más bajos del factor 3.

En base a los signos de los componentes del factor 2 y factor 3 y por la disposición de los objetos en el plano f2 y f3 en los sectores negativos y positivos, los sitios del primer grupo por ejemplo "Tomsk" y "Mogilev" donde f2 es positivo y f3 es negativo, se puede interpretar que los sitios que integran este primer grupo tienen más tonos, menos colores, menos espacio de texto en la primera página, más tamaño de letras, menos tamaño de imagen en la primera página y una estructura más desarrollada.

De acuerdo con los signos de los componentes de los factores y con la posición de los sitios Web, se puede observar que los sitios del segundo grupo en el plano f2 y f3 tienen menos tonos, más colores, más espacio de texto en la primera página, menos tamaño de letras,

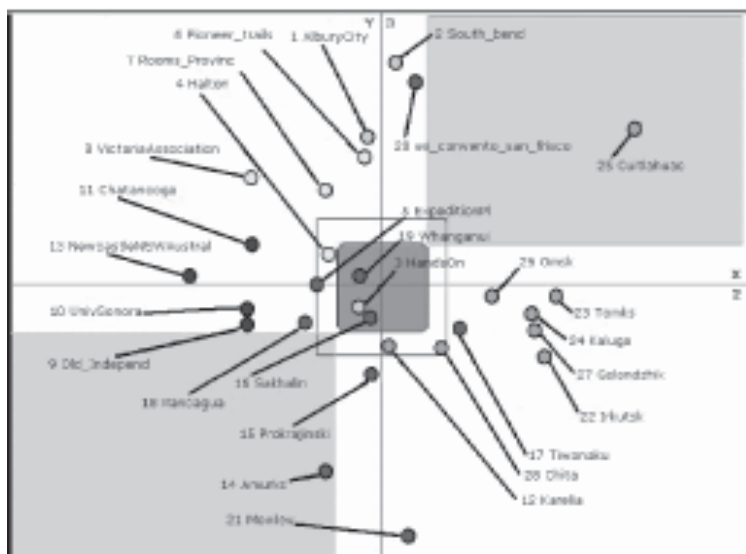


FIGURA 2. CLUSTERS FINALES

más tamaño de imagen de la primera página y una estructura más simple.

Cada sitio Web es único pero con características comunes compartidas con los sitios del primer grupo, los cuales son sitios más multidisciplinarios que los sitios del segundo grupo, y con respecto a los sitios del segundo grupo son sitios de temas específicos y pequeños. Por ejemplo, el sitio Aurora está enfocado únicamente a los bomberos, es un sitio de un tema específico, mientras que el sitio Victoria Association es un sitio Web cuyos temas principales son la colección de pinturas de un pintor y la casa de un famoso político de la ciudad, otros sitios incluyen no más de tres temas especiales en el desarrollo del sitio Web.

Dado el análisis de estas características, se puede deducir que los sitios del primer grupo en la primera página no tienen tanto texto y el que se incluye es de un tamaño más grande y además tienen una estructura más desarrollada dado que su misión multidisciplinaria les exige este enfoque; por otro lado los sitios Web con temas específicos no necesitan mucho espacio para desplegar su información y por esta razón usan recursos de Internet de una forma más económica, incluyen más textos y menos espacio vacío en la primera página, incluyen menos opciones en el menú y el tamaño del texto no es tan amigable para los usuarios; por lo general no hay hechos o eventos nuevos relacionados con los temas específicos del sitio Web y por lo tanto no tienen una frecuente actualización y renovación de sus sitios Web.

Ambos grupos de sitios Web, son grupos de museos exitosos, por esta razón se necesita elegir uno de estos modelos para el desarrollo del modelo ideal, no se puede realizar una síntesis de estos dos grupos de sitios Web debido a que los atributos que los integran son opuestos.

5. Modelo de creación del sitio Web exitoso del MUREH

Las características de los materiales que se utilizarán en el Museo Virtual del MUREH, tienen características similares a los sitios Web del primer grupo de museos exitosos, y por lo tanto se debe considerar este enfoque en el desarrollo del modelo ideal, características tales como, letras de tamaño grande, característica que es bien recibida por los usuarios jóvenes, una estructura desarrollada recomendable en sitios con metas multidisciplinarias, utilizar más tonos y menos


colores en el desarrollo del sitio Web, para el caso de usuarios de tipo científico se debe añadir como parámetro obligatorio el metatag.

El modelo del sitio Web propuesto para el MUREH en su primera aproximación se planteó con 13 botones en primer nivel y 13 para cada opción para el segundo nivel, realmente este esquema tiene discrepancias con los sitios del primer grupo por que estos sitios tienen un sistema no muy desarrollado, no incluyen tantas opciones. Por esta razón se ha revisado el esquema propuesto y se han agrupado algunas opciones, disminuyendo de esta forma las opciones del primer nivel para enriquecer las opciones del segundo nivel, al mismo tiempo algunas opciones como, por ejemplo, retroalimentación o visita rápida mantienen sus hipervínculos que son diferentes a los incluidos en los botones. Finalmente, como resultado se ha obtenido una estructura más flexible y que conserva su principio de balanceo, acercándose a la estructura de los sitios exitosos del primer grupo.

6. Conclusiones

Con base en la misión del museo y las restricciones de infraestructura de Internet, se ha logrado determinar los principales modelos de usuarios del sitio Web del Museo Regional de Huajuapán, así como la integración del conjunto de atributos de sitios Web de museos exitosos y la formación del muestreo de estos sitios. Además, se ha realizado el análisis de clusters del conjunto de sitios Web exitosos en base a sus atributos.

El resultado de este análisis muestra dos grupos principales de sitios Web con características opuestas, esta deducción ha sido confirmada analizando directamente el contenido de los sitios Web que integran estos grupos.

Se ha elegido desarrollar el sitio Web del Museo regional de Huajuapán basándose en el primer grupo de sitios exitosos, por lo tanto se puede decir que la elección de los parámetros que integrarán este sitio Web, han sido elegidos en una forma objetiva indicando que cada elección es adecuada y que guiará a un sitio exitoso 

Referencias

1. *El mundo mixteco*, <http://virtual.utm.mx/mixteca/>
2. Noticias de Oaxaca, <http://www.noticias-oax.com.mx>
3. Mixteca las noticias en Internet, <http://www.mixteca.com.mx>
4. CASTRO DE, D., ET AL.
2004 *Automatic Web News Extraction Using Tree Edit Distance*, Department Federal University of Minas Gerais, Belo Horizonte, Brasil.
5. MAKAGONOV, P., SBOYCHAKOV, K.
1998 "Man-machine methods for solution of weakly formalized problems in humanitarian and natural fields of knowledge (visual heuristic cluster analysis)", International Computer Symposium CIC'98 "la computación: Investigación, desarrollo y aplicaciones", Conferencia Magistral por Invitación, CIC-IPN MEXICO.

*Pavel Makagonov
Celia B. Reyes Espinoza
Universidad Tecnológica de la Mixteca*

