

Enriquecimiento masivo del catálogo de la Biblioteca Central FCEN-UBA : carga de tablas de contenido en los registros

Cristian Nicolas Rucks, Celeste Magallanes, Virginia Teppa
Biblioteca Central, Facultad de Ciencias Exactas y Naturales, UBA
Ciudad Universitaria, Pab. II, (1428) Buenos Aires, Argentina
contacto: nrucks@bl.fcen.uba.ar

Resumen

Ante la evidente deficiencia de los puntos de acceso del catálogo, se proyecta incluir las tablas de contenido como puntos de acceso y descripción pormenorizada de los registros. El personal del sector Circulante (estudiantes de la Facultad) destina 52 horas-hombre semanales sin afectar el servicio. Un muestreo inicial estima unos 10 minutos por tabla; entre 10 y 15 meses de trabajo para cubrir 8707 títulos de la colección seleccionada.

Se da cuenta de

- criterios de prioridades en la selección del material
- criterios de carga de las tablas
- mecanismos de revisión
- indicadores de trabajo realizado

Se analiza cantidad de tablas cargadas y tiempo insumido, aporte cuantitativo de las tablas en relación a título y descriptores asignados. Se exponen limitaciones inherentes: idioma, lenguaje no controlado.

Origen de la propuesta

Detrás de este proyecto subyace la idea general de relación costo/beneficio. De alguna manera es suplir con cantidad la poca profundidad y extensión de la clasificación. Más precisamente, la imposibilidad de hacer un análisis temático profundo y exhaustivo de todos y cada uno de los materiales. No solo por falta de tiempo disponible, sino también por falta de especialistas conocedores de la amplísima gama de ciencias representadas en esta biblioteca: física, química, biología, geología, paleontología, oceanografía, meteorología, matemáticas, computación, por mencionar solamente los nombres de las carreras de la Facultad (en sus formas abreviadas).

Ante esa imposibilidad pues, se explota un recurso disponible directamente en el material: la tabla de contenido. Esto minimiza al extremo la toma de decisiones por parte del personal no profesional, llevándolo casi a la carga directa, requiriendo solamente haber fijado un número limitado de criterios claros.

Obstáculos

Los principales obstáculos o problemas encontrados son los siguientes:

- La importante cantidad de horas hombre necesarias: en nuestro caso están disponibles con una mínima reorganización de las tareas. Esto para encarar la cobertura retrospectiva del sector de Circulante.
- La ausencia absoluta de control de vocabulario: esta ausencia está compensada en buena medida por la pertinencia del vocabulario utilizado por los propios autores en las tablas, pues se trata en mayor medida de material académico-científico, que suele tener una estructura predeterminada con terminología propia y explícita.
- La problemática del idioma para la gran cantidad de material que no está en castellano, principalmente en inglés: una traducción de esa terminología está fuera de este proyecto, pues obviamente implicaría importantes costos y tiempos.

Criterios de prioridades en la selección del material

El criterio general fue dar la prioridad de carga a aquél material más susceptible de ser de interés para los usuarios. Por un lado, por una cuestión práctica nos limitamos a los materiales ya catalogados, no cubriendo materiales ingresados después de iniciado el proyecto. Esto permite encarar un conjunto cerrado y dejar para más adelante la cobertura del resto. Por otro lado, se prefirió comenzar con un conjunto de menor cantidad de obras y de mayor consulta. Nos circunscribimos entonces al material de "Circulante", que responde a ese criterio: está directamente a disposición de los usuarios, en estanterías

de libre acceso, sin necesidad de acudir al depósito y es estadísticamente el material más utilizado. Dentro de ese universo, se estableció una lista fija de carga, comenzando con los años de edición más nuevos (2014, 2013, 2012, 2011, etc.) y dentro de cada año, cubriendo todas las áreas temáticas, ordenando por los 2 primeros números de signatura de clase (CDU). De esta manera se pretende cubrir uniformemente todas las temáticas dentro de la colección de Circulante, pero comenzando por lo más nuevo.

Con la práctica, resultó evidente que las actas de conferencia constituyen un problema para este proyecto, pues sus contenidos suelen ser, naturalmente, las ponencias. La cantidad de éstas excede el criterio inicial de hasta 30 o 40 “capítulos” por tabla, por lo que se decidió dejarlas de lado en esta instancia. Se considerará su inclusión más adelante.

Establecimiento de los criterios de carga

Con el objetivo de mantener homogeneidad tanto en el tiempo como entre los operadores, se elaboró un pequeño manual de procedimientos donde se asentaron los criterios de carga; este manual ha sido sujeto a enmiendas a medida que surgieron casos no previstos, y por supuesto sigue abierto.

Los criterios siguen como pautas generales, fidelidad, completitud, exhaustividad, dentro del marco establecido, y en lo posible simplicidad.

Esto se traduce en lo siguiente:

- Elegir el nivel de detalle más adecuado (ver más abajo);
- Cargar todos los capítulos presentes en la tabla (y solo esos), en ese nivel;
- Cargar de manera fidedigna, sin correcciones, respetando mayúsculas, acentuación, símbolos presentes;
- Distinguir únicamente la parte de “numeración” (cap. 1, cap. 2. etc.) cuando está presente, de la parte de nombramiento.
- No añadir nombres de responsables ni páginas a las que remite la tabla.

Acerca del nivel de detalle: en nuestra colección resulta bastante común que un libro no estructure su tabla de contenido en 1 solo nivel de detalle, sino en 2 o incluso 3 : por ejemplo, “Partes”, subdivididas en “Capítulos”, y en “Secciones”.

Generalmente es el nivel intermedio el que se recoge, pues resulta ser el más expresivo sin ser sobredimensionado (algunos casos tienen tablas de hasta 4 o 5 páginas).

Se decidió relevar la tabla en el nivel que mejor se aproxime a un máximo de unas 30 o 40 líneas (“capítulos”). Parece ser un nivel que ilustra bastante bien el contenido de la obra, sin hacerse demasiado extenso. Quizás porque coincida con otras tablas de un solo nivel que tienen esa cantidad de capítulos como máximo.

Criterios de carga

- Seleccionar el nivel de detalle más cercano a 30~40 capítulos
- Cargarlos todos
- No buscar en el cuerpo de la obra, solo en la tabla.
- Respetar al idéntico ortografía, mayúsculas, acentuación, etc.
- Respetar al máximo los símbolos presentes
- Incluir la secuencia o numeración, cuando la haya, separándola del título del capítulo, con la cadena “-- ” (guión, guión, espacio)
- No incluir las páginas a las que se remite
- No incluir menciones de responsabilidad
- Asentar en notas situaciones extrañas
- Consultar toda duda de casos no previstos

Rutina de trabajo

Se estableció la siguiente rutina de trabajo:

Se comienza con la búsqueda de los libros en el estante. Se sigue un listado previamente armado el cual ordena en primer nivel a los libros en orden decreciente en año de edición y en segundo nivel, por su ubicación. Una vez identificados y separados entre 10 y 15 volúmenes, se los asigna al operador en turno. Cada operador tiene turnos de 2 horas fijas destinadas a esta tarea, diferenciada de las otras tareas del sector.

Los libros se colocan a un lado del operador ordenados y accesibles, siempre en el mismo puesto de trabajo. El operador inicia su sesión en el módulo de carga y comienza con el primer libro. Con criterios ya definidos, se decide cuáles líneas irá a cargar y cuáles no de la tabla de contenidos que exhiba el libro. Una vez finalizadas las dos horas de trabajo, los libros permanecen en el sector con el objetivo de cumplir la instancia de revisión por el siguiente operador.

La revisión consiste en una inspección visual de las líneas cargadas y una comparación con lo impreso en el libro. En caso de ser necesario, se corrige.

Es importante mencionar que la revisión la realiza una persona diferente a la que carga. De esta manera se corrigen los errores que pueda haber cometido el anterior.

Finalmente, una vez cumplidas las instancias de carga y de revisión, los libros son devueltos al estante. Cabe aclarar que el operador no va a buscar libros ni los reubica, eso queda bajo la responsabilidad de los supervisores, y que mientras se dedica a esta tarea, no realiza préstamos ni devoluciones; el operador está completamente abocado a la carga y revisión de tablas de contenido.

Los números que tenemos

El objetivo principal de este enriquecimiento masivo no es la carga de datos en sí misma. Como se estableció, la idea es enriquecer los registros con puntos de accesos relevantes y obtener una mejor descripción de los ítems, llevando al mínimo posible la necesidad de análisis y toma de decisión.

Y en definitiva, a través de la presencia de las tablas de contenido en los registros, mejorar la recuperación, y así dinamizar la circulación de la colección.

Por ahora, no se podrá corroborar, como lo establece Morris, que los libros cuyas tablas de contenido están disponibles “salen más en préstamo” que los que no¹. No hasta tanto el proyecto tenga a disposición una cantidad significativa de tablas disponibles.

Dada la cantidad de tablas cargadas y la altura del año, es probable que se fije como objetivo la carga dentro del presente cuatrimestre, para arrancar el principio del año entrante, primer cuatrimestre 2016, con alrededor de 1900 tablas disponibles.

Mientras tanto, se puede dar un primer pantallazo de respuesta a las siguientes preguntas: cuánto tiempo lleva cargar una tabla, cuánto tiempo lleva revisar la carga, cuántos errores se corrigen en la revisión, y con cuántas palabras significativas contribuye una tabla de contenido, en relación a las palabras significativas del título, más descriptores.

Tiempos de carga, y de revisión

Al 18 de septiembre de 2015, se ha cargado un total de **954** tablas de contenido (**903** revisadas);

junio: **41** ; julio: **81** ; agosto : **425** ; septiembre (al 18 del mes) : **407**.

El tiempo promedio de carga es de un poquito menos de **10** minutos por tabla, lo que coincide con las previsiones iniciales, además, se espera que se pueda ir mejorando con el tiempo y la práctica. Sin embargo, hay que tener en cuenta que las tablas de contenido son bastante disímiles en cantidad de líneas. A la fecha, el promedio es de **21** líneas por tabla, con una dispersión importante.

Cabe aclarar que el proyecto comenzó en junio de este año 2015 con un muestro inicial de 41 tablas, y que el mes de julio consistió en aprendizaje por parte de los operadores, coincidiendo con las semanas de receso entre cuatrimestres, con lo cual la actividad de la biblioteca en general ha disminuido considerablemente, incluyendo las tareas propias de este proyecto; agosto 2015 fue el primer mes “normal” de actividades.

La estimación inicial en el papel era de unas de **624** tablas por mes: **2** horas por persona por semana, por **13** personas (12 del sector de Circulante, más otra del sector de Referencia), a un ritmo de **10** minutos por tabla. El tiempo promedio de carga resultó correctamente estimado, en cambio el tiempo requerido para la revisión, cuya importancia no había sido tenida en cuenta en un principio, se hizo evidente desde el mes de julio. Por eso se tomó la decisión de asignar, desde el 1º de agosto, **4** horas por semana y por persona al proyecto (en lugar de las 2 iniciales) repartidas en 2 turnos de 2 horas por semana y persona; es decir que con **13** personas destinadas al proyecto, se dispone de **52** horas-hombre por semana.

Una segunda estimación de **4** horas por persona por semana, por **13** personas, a un ritmo de **20** minutos por tabla (**10** de carga más **10** de revisión), compensa la duplicación de tiempo necesario por la duplicación de horas hombre disponible, manteniendo **624** tablas por mes. Significaría una proyección de

¹ Morris, R.C. (2001) Online tables of contents for books: effect on usage. *Bull Med Libr Assoc* 89(1): 29-36

unos 14 meses para cubrir los 8707 libros. No se llegó a ese ritmo en el mes de agosto, pero se mejora en el mes de septiembre, que aún no terminó. Es esperable que los números reales sean menores que los previstos, sin embargo se estima mejorar la cantidad de horas-persona trabajadas a medida que se aceite la rutina de trabajo.

Carga y Revisión

Resultó natural suponer que además de cargar las tablas iba a ser necesario revisar los datos cargados. Se estableció como rutina someter a revisión por otro operador, distinto del que cargó. Por otro lado, para minimizar el movimiento de libros, se somete a revisión el conjunto de libros que se estuvo cargando en el turno inmediatamente anterior.

Se estableció una rutina informática de comparación entre el texto ingresado durante la carga y el que quedó una vez revisado. Esa diferencia se midió siguiendo el criterio de “Distancia de Levenshtein”² que establece una distancia de una unidad por cada carácter añadido, eliminado, o intercambiado entre dos cadenas de texto. Se sumaron las distancias de todas las líneas para asignar la distancia existente entre la tabla cargada y la revisada. Por supuesto líneas completamente omitidas o añadidas contribuyen con una distancia importante. Se estima que el número así obtenido resulta relevante, siendo una cantidad medible objetiva. Al 18 de septiembre, son 839 tablas cargadas y revisadas³.

-Tal vez resulte ilustrativo que casi la mitad (48%) de las tablas cargadas hasta ahora tienen una distancia de 0 (cero), coincidiendo perfectamente lo cargado con lo revisado. Resulta imposible saber a ciencia cierta si la revisión es perfecta. Pero hay que establecer algún límite, y se confía en los operadores.

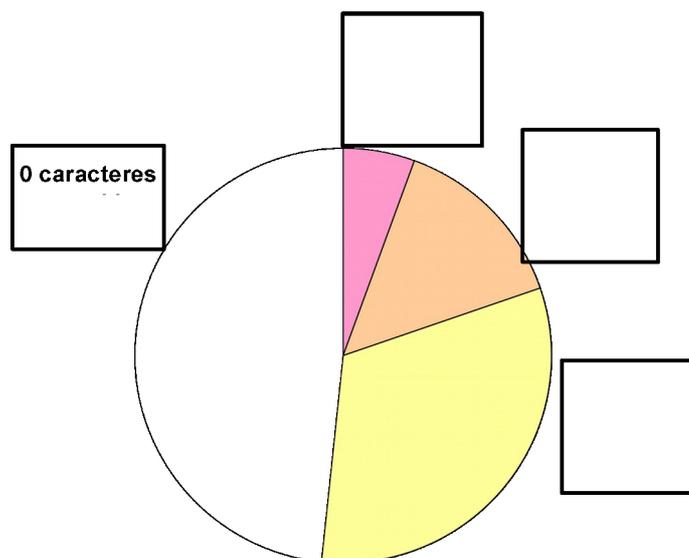
-El 32% (casi un tercio) tiene una distancia de entre 1 y 9, lo que podríamos llamar pequeños errores, de tipo quizás; aunque solo un análisis exhaustivo (fuera del alcance de este estudio) podría determinar cuántos son acumulaciones de pequeños errores y cuántas palabras olvidadas, por ejemplo.

-El 14% (algo más de un décimo) tiene una distancia de entre 10 y 99, lo que podríamos llamar correcciones más importantes. Es probable que amerite algún análisis.

-Solo 6% tiene una distancia de 100 o más, que en algunos casos responden a errores en la aplicación de los criterios de carga; (por ejemplo no incluir la mención “Capítulo” cuando sí está presente en la tabla); se espera que este porcentaje vaya disminuyendo a medida que los operadores hagan experiencia.

Como otra conclusión parcial al respecto, este análisis aporta elementos de respuesta a la pregunta “¿vale la pena revisar?” Se estima que si bien en más de la mitad de los casos la revisión no hace más que corroborar que se cargó perfectamente, hay un tercio que pudo ser corregido, aunque no más de 9 caracteres. ¿Es eso mucho? ¿Es eso poco? Cada error es una palabra que no se recupera correctamente. Queda analizar en función de lo que cuesta (en tiempo) esa revisión. Ese dato aún no está siendo recolectado por el sistema, pero se espera que lo haga en breve.

Gráfico 1 : cantidad de caracteres corregidos al revisar



² https://es.wikipedia.org/wiki/Distancia_de_Levenshtein

³ Lamentablemente no se retuvo el texto inicial de las primeras 64 tablas cargadas, su diferencia al revisar no se puede calcular.

Alternativas al tipeo, brevemente consideradas y descartadas

Se consideró y exploró, brevemente, el acceso a las tablas de contenido disponibles en la Web, en particular a través de Google books y Amazon, sin embargo, resultaron poco satisfactorios, pues no siempre se encuentran, y cuando se encuentran, no son fácilmente copiables, pueden resultar ser solo imágenes o tablas incompletas. De manera tal que en lugar de resultar una opción de menor trabajo, aparentemente implicaría uno mayor, con resultados inciertos. No fue explorada más allá de algunos casos.

También existe la posibilidad de escaneo, pero ésta implica a su vez la corrección de OCR, la revisión de la estructura, la selección de nivel, además del propio trabajo de escaneo masivo. Parece ser, en el mejor de los casos, una opción tan trabajosa como la de tipeo integral desde cero, con desafíos diferentes.

La opción de tipeo parece, paradójicamente tal vez, la más fácil de homogeneizar y de hacer masiva, siempre que se establezcan reglas claras, se simplifiquen al máximo los pasos y se minimice tanto como sea posible la toma de decisión por parte del personal, y que los recursos humanos estén disponibles.

Conclusiones (muy) parciales

Tiempos estimados a la fecha

Con 4 horas por semana y por persona, 4 semanas al mes, 13 personas, son 208 horas por semana. Si la carga lleva unos 10 minutos más 10 de revisión, tendríamos un ritmo de 3 tablas por hora. De un total de 8707 libros de la sección de circulante, nos da una proyección de 14 meses de trabajo.

Por otro lado, queda establecida la conveniencia de contar con tiempos de revisión mensurados. Esto se hará por sistema, de modo similar al registro de tiempo de carga. Mientras tanto, es de suponer que este tiempo de revisión es una de las razones por las cuales se está por debajo del promedio de carga estimado inicialmente.

Enriquecimiento: palabras significativas

El caso de la “Glucólisis” en el ejemplo 4 más abajo, ilustra el interés en cargar las tablas de contenido. A la fecha, en el catálogo de la biblioteca, la búsqueda por “Glucólisis” solo trae una tesis de 1975. A partir de la carga de las tablas, con unas 900 tablas cargadas, quedan ya expuestos 6 libros en los que “Glucólisis” figura en algún capítulo, y que antes permanecían sin poder ser recuperados. En particular, el libro del ejemplo 4 se titula “Bioquímica”, en su capítulo 17 trata de “Glucólisis”, sin embargo solo podía ser recuperado por “bioquímica”, una palabra evidentemente demasiado general.

La cantidad de palabras significativas que contienen los títulos es muy variable de una obra a otra. Lo mismo sucede con las tablas. Algunos títulos son muy escuetos, de solo 1 o 2 palabras como por ejemplo “Bioquímica” (ver los ejemplos 3 y 4 más abajo) cuando sus tablas de contenidos aportan más de 70 palabras nuevas. En el otro extremo algunas tablas no aportan ninguna palabra que no exista ya en el título, o entre los descriptores asignados (ejemplo 2 más abajo). Algunos opúsculos ni siquiera tienen tabla, en cambio las Actas de conferencia tienen por tabla la lista de ponencias...

Más allá de esos casos, hasta ahora nuestras estadísticas muestran que en promedio, una tabla de contenido aporta más de 8 veces la cantidad de palabras que tiene el título, aunque con muy considerable dispersión.

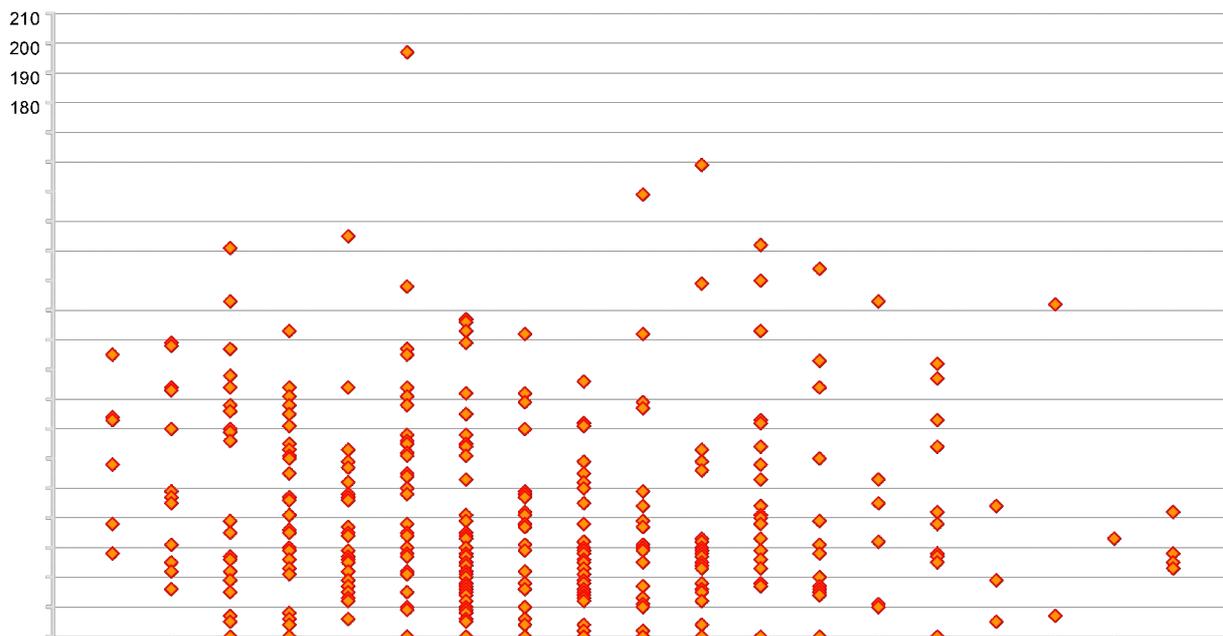
Un caso ilustrativo son las obras con 7 palabras entre título y descriptores que tienen tablas que aportan entre 0 y 107 palabras significativas únicas, mientras que las obras con 5 palabras entre título y descriptores tienen tablas que aportan entre 6 y 135.

Para estos cálculos se establecen las siguientes condiciones: se ignoran diferencias de acentuación y mayusculación, y no se tienen en cuenta palabras “vacías” generales como artículos (el la los), ni típicas de las tablas de contenido como “Prefacio”, “Prólogo”, “Introducción”, “Bibliografía”, etc.. Asimismo se ignoran repeticiones de palabras, tanto en el conjunto título+descriptores como en la tabla.

En el gráfico 2 se puede ver la cantidad de palabras nuevas aportadas por cada libro, ordenando por cantidad de palabras ya existentes entre título y descriptores. Se puede apreciar las dispersiones tanto en cantidad de palabras preexistentes (título y descriptores) como aportadas por las tablas.

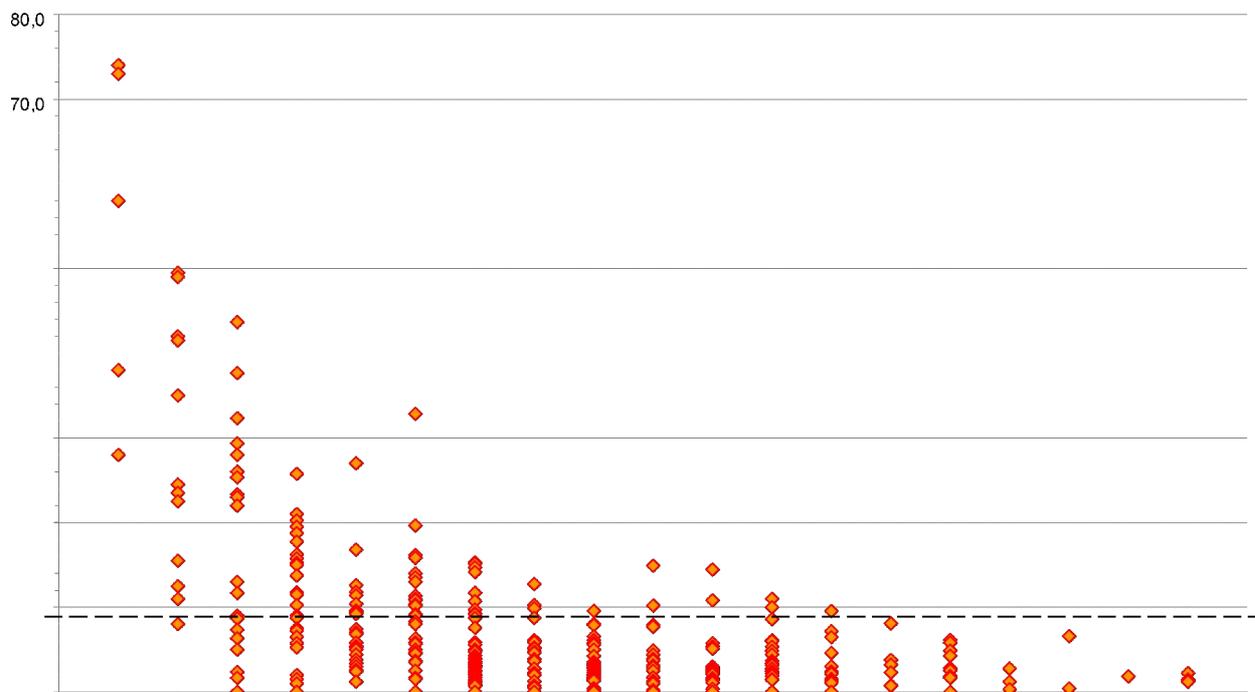
Gráfico 2 : cantidad de palabras nuevas en la Tabla de contenido, agrupados por cantidad de palabras en título más descriptores

Palabras nuevas en la TDC



En el gráfico 3 se muestran los mismos libros, graficando la relación (palabras nuevas)/(palabras ya existentes). El promedio de 8 veces está marcado con un línea punteada. Por supuesto los libros con una sola palabra en el título tienen relaciones muy superiores, pues sus tablas aportan comparativamente mucho más.

Gráfico 3 : relación cantidad de palabras nuevas en la Tabla de contenido/cantidad de palabras en título más descriptores



Con solo 3 meses de avance en la ejecución del proyecto, resulta imposible sacar conclusiones definitivas. Solo primeras impresiones. Por de pronto, no cabe duda de que, en términos de palabras puntos de acceso, efectivamente se está enriqueciendo el catálogo, en una muy importante relación: unas 8 veces las la cantidad de palabras del conjunto título+descriptores.

Por otro lado, el tiempo de carga parece haber sido relativamente bien estimado, unos 10 minutos, no así el de revisión, que tendrá que ser mensurado. Dicho esto, pensamos que a pesar del tiempo que requiere, el tiempo de revisión aporta una mejora de calidad significativa, siendo que casi la mitad de las tablas han tenido por lo menos una corrección.

Veremos con el tiempo si estos números mejoran.

Ejemplo de cálculo de la relación palabras del título + descriptores / palabras de la tabla

Para estos cálculos solo se consideran obras en castellano, a efectos de tener en cuenta la presencia de descriptores.

“Hidrogeología de llanuras”.

Auge, Miguel Pedro . Buenos Aires: Universidad de Buenos Aires, 2003

Descriptores: HIDROGEOLOGIA ; AGUA SUBTERRANEA ; LLANURAS

Tabla:

LLANURAS
 LLANURAS ALUVIALES
 CARACTERES HIDROGEOLÓGICOS
 MÉTODOS y TÉCNICAS PARA LOS ESTUDIOS HIDROGEOLÓGICOS
 BIBLIOGRAFÍA

Palabras significativas y únicas del título: **2**

Hidrogeología	1
de	palabra no significativa
llanuras	2

Palabras significativas y únicas de los descriptores: **4**

Palabras significativas y únicas de los descriptores que no están ni en el título: **2**

AGUA	1
HIDROGEOLOGIA	Ya existe en el título
LLANURAS	Ya existe en el título
SUBTERRANEA	2

Palabras significativas y únicas de título + descriptores: **2+2=4**

Palabras significativas y únicas de la tabla: **7**;

Palabras significativas y únicas de la tabla que no están ni en el título ni en los descriptores: **6**

ALUVIALES	1
BIBLIOGRAFÍA	palabra no significativa propia de tablas
CARACTERES	2
ESTUDIOS	3
HIDROGEOLÓGICOS	4
LLANURAS	Ya existe en el título
LOS	palabra no significativa
MÉTODOS	5
PARA	palabra no significativa
TÉCNICAS	6
y	palabra no significativa

Esta tabla de contenidos aporta **6** palabras nuevas en relación a un título+descriptores de **4**.

Algunos ejemplos ilustrativos

Ejemplo 1• Caso tipo: registro de un título con 4 palabras significativas, más descriptores que aportan 3 palabras nuevas (total 7), es enriquecido por una tabla de 22 capítulos o líneas aportando 42 palabras nuevas. Por ejemplo en este caso, “tecnologías” e “información” ya existen en el título y no son consideradas aportantes; tampoco las palabras no significativas típicas de tablas que son “prólogo” y “prefacio”.

“Gobierno de las tecnologías y los sistemas de información”.

Piattini Velthuis, Mario; Hervada Vidal, Fernando. Madrid: Ra-Ma, 2007

Descriptor: SEGURIDAD INFORMATICA ; TECNOLOGIAS DE INFORMACION ; SISTEMAS OPERATIVOS ; SISTEMAS DE INFORMACION

Relación Tabla/Título+Descriptor = 42/(4+3) = 6

Tabla:

AUTORES
PRÓLOGO
PREFACIO
CAPÍTULO 1-- EL GOBIERNO DE LAS TECNOLOGÍAS Y SISTEMAS DE INFORMACIÓN
CAPÍTULO 2-- TÉCNICAS Y PLANIFICACIÓN PARA EL GOBIERNO DE TSI
CAPÍTULO 3-- EL VALOR Y EL RETORNO DE LA INVERSIÓN DE LAS TSI
CAPÍTULO 4-- ASPECTOS JURÍDICOS
CAPÍTULO 5-- VISIÓN GENERAL DEL GOBIERNO DE LA SEGURIDAD
CAPÍTULO 6-- IMPLANTACIÓN DEL GOBIERNO DE LA SEGURIDAD DE LAS TSI
CAPÍTULO 7-- GESTIÓN DE RIESGOS DE SEGURIDAD
CAPÍTULO 8-- CUADROS DE MANDO PARA EL GOBIERNO DE TSI
CAPÍTULO 9-- MODELOS DE MADUREZ PARA SGSI DESDE UN ENFOQUE PRÁCTICO
CAPÍTULO 10-- PANORÁMICA GENERAL SOBRE NORMAS DE SEGURIDAD DE TECNOLOGÍAS DE LA INFORMACIÓN
CAPÍTULO 11-- NORMA ISO 27001- GESTIÓN DE LA SEGURIDAD
CAPÍTULO 12-- LAS MÉTRICAS DE SEGURIDAD DE LA INFORMACIÓN
CAPÍTULO 13-- NORMA ISO 17799. CÓDIGO DE BUENAS PRÁCTICAS PARA LOS SGSI
CAPÍTULO 14-- COSO
CAPÍTULO 15-- COBIT 4
CAPÍTULO 16-- ITIL Y LA NORMA UNE-ISO/IEC 20000
CAPÍTULO 17-- SARBANES-OXLEY
CAPÍTULO 18-- BASILEA II: RETO Y OPORTUNIDAD PARA LA AUDITORÍA DE SISTEMAS
ÍNDICE ALFABÉTICO

Ejemplo 2• Caso extremo sin aporte. Con un título de 9 palabras significativas diferentes, la tabla de 4 capítulos, (por lo menos existe, hay casos sin tabla), no aporta ninguna palabra que no exista ya en el título! La tabla de contenido de 4 líneas contribuye con 0 palabras nuevas.

“Para hacer valer tus derechos: Misión, estructura, labor, historia”

Pierini, Alicia. Buenos Aires: Defensoría del Pueblo de la Ciudad de Buenos Aires, 2011.

Descriptor: DERECHOS DEL CIUDADANO

Relación Tabla/Título+ Descriptor = 0/(8+1) = 0

Tabla:

Misión
Estructura
Labor
historia

Ejemplo 3• Caso extremo con aporte. Con un título de solo 2 palabras significativas y “MICROBIOLOGIA” por único descriptor, que no enriquece el título, la tabla de 38 capítulos o líneas aporta 98 palabras nuevas (únicas).

“Introducción a la Microbiología”.

Tortora; Funke; Case .Buenos Aires: Médica Panamericana, 2007.

Descriptor: MICROBIOLOGIA

Relación Tabla/Título+ Descriptor = 98/(2+0) = 49

Tabla:

1-- El mundo microbiano y usted
2-- Principios de química
3-- Observación de los microorganismos a través del microscopio
4-- Anatomía funcional de las células procariontes y eucariontes

5-- Metabolismo microbiano
 6-- Crecimiento microbiano
 7-- Control del crecimiento microbiano
 8-- Genética microbiana
 9-- Biotecnología y DNA recombinante
 10-- Clasificación de los microorganismos
 11-- Procariontes: dominios Bacteria y Archaea
 12-- Eucariontes: hongos, algas, protozoos, y helmintos
 13-- Virus, viroides y priones
 14-- Principios de enfermedad y epidemiología
 15-- Mecanismos de patogenicidad microbianos
 16-- Inmunidad innata: defensas inespecíficas del huésped
 17-- Inmunidad adquirida: defensas específicas del huésped
 18-- Aplicaciones prácticas de la inmunología
 19-- Trastornos asociados con el sistema inmunitario
 20-- Fármacos antimicrobianos
 21-- Enfermedades microbianas de la piel y de los ojos
 22-- Enfermedades microbianas del sistema nervioso central
 23-- Enfermedades microbianas de los sistemas circulatorio y linfático
 24-- Enfermedades microbianas del aparato respiratorio
 25-- Enfermedades microbianas del aparato digestivo
 26-- Enfermedades microbianas de los aparatos urinario y genital
 27-- Microbiología ambiental
 28-- Microbiología aplicada e industrial
 Apéndice A-- Clasificación de las bacterias según el Manual Bergey
 Apéndice B-- Métodos para la obtención de muestras clínicas
 Apéndice C-- Vías metabólicas
 Apéndice D-- Exponentes, logaritmos exponenciales y tiempo de generación
 Apéndice E-- Guía taxonómica para las enfermedades
 Apéndice F-- Respuestas de las preguntas de revisión y de opciones múltiples del cuestionario de estudio
 Glosario
 Créditos
 Índice analítico

Ejemplo 4• Caso extremo con aporte. Con un título de solo 1 palabra significativa y “BIOQUIMICA” por único descriptor, la tabla de 36 líneas aporta 75 palabras nuevas (únicas).

“Bioquímica” 3a ed.

Voet; Voet. Buenos Aires: Editorial Médica Panamericana, 2006

Descriptores: BIOQUIMICA

Relación Tabla/Título+ Descriptores = $75/(1+0) = 75$

Tabla:

Capítulo 1-- La vida
 Capítulo 2-- Soluciones acuosas
 Capítulo 3-- Principios de termodinámica: revisión
 Capítulo 4-- Aminoácidos
 Capítulo 5-- Ácidos nucleicos, expresión génica y tecnología del DNA recombinante
 Capítulo 6-- Técnicas de purificación de proteínas y ácidos nucleicos
 Capítulo 7-- Estructuras covalentes de las proteínas y los ácidos nucleicos
 Capítulo 8-- Estructura tridimensional de las proteínas
 Capítulo 9-- Plegamiento, dinámica y evolución estructural de las proteínas
 Capítulo 10-- Hemoglobina: función proteica en el microcosmo
 Capítulo 11-- Azúcares y polisacáridos
 Capítulo 12-- Lípidos y membranas
 Capítulo 13-- Introducción a las enzimas
 Capítulo 14-- Velocidades de las reacciones enzimáticas
 Capítulo 15-- Catálisis enzimática
 Capítulo 16-- Introducción al metabolismo
Capítulo 17-- Glucólisis
 Capítulo 18-- Metabolismo del glucógeno
 Capítulo 19-- Transducción de señales
 Capítulo 20-- Transporte a través de membranas
 Capítulo 21-- Ciclo del ácido cítrico
 Capítulo 22-- Transporte de electrones y fosforilación oxidativa
 Capítulo 23-- Otras vías metabólicas de los carbohidratos
 Capítulo 24-- Fotosíntesis

Capítulo 25-- Metabolismo de los lípidos
Capítulo 26-- Metabolismo de los aminoácidos
Capítulo 27-- Metabolismo energético: integración y especialización de los orgánulos
Capítulo 28-- Metabolismo de los nucleótidos
Capítulo 29-- Estructura de los ácidos nucleicos
Capítulo 30-- Replicación, reparación y recombinación del DNA
Capítulo 31-- Transcripción
Capítulo 32-- Traducción
Capítulo 33-- Virus: paradigmas para la función celular
Capítulo 34-- Expresión génica de los eucariontes
Capítulo 35-- Fisiología molecular
ÍNDICE ANALÍTICO

Listado de palabras no significativas generales (castellano)

A; AL; ANTE; BAJO; CABE; CON; CONTRA; DE; DEL; DESDE; DURANTE; EL; EN; ENTRE; HACIA; HASTA; LA; LAS; LES; LO; LOS; MEDIANTE; O; PARA; POR; SEGUN; SIN; SO; SOBRE; TRAS; UN; UNA; UNAS; UNOS; Y;

Listado de palabras no significativas propias de tablas de contenido (castellano)

(Las mayúsculas y tildes son indistintos para el matcheo; “\?” marca la presencia o ausencia indistinta del caracter anterior).

Prefacio; Prologo; Presentacion; Palabras; Agradecimientos\?; Introduccion; General; Glosario; Referencias\?; Bibliografic\?as\?; Notas\?; Anexos\?; Apendices\?; Indices\?