

La teoría de la información en bibliotecología

Carlos Izquierdo Costa¹

Resumen: Se analiza la teoría de la información con un enfoque bibliotecológico, en el que se combinan los estudios de expertos como Weaver y Shannon. Analizando esta teoría e interrelacionándola con los estudios Bradford, se introduce la investigación de Cleverdon sobre la pertinencia de los documentos y su aplicación práctica.

1. Introducción

La información surge de la necesidad de dar a conocer o transmitir las cosas abstractas generadas en el intelecto y las concretas, creadas en base al intelecto. Primitivamente o en una primera escala, surge la necesidad que tiene el ser humano de intercambiar ideas con sus semejantes.

A pesar de la simplicidad del hecho, es primordial cierta condición para alcanzar esta finalidad, y es que los participantes posean elementos o condiciones necesarias para la emisión y la recepción.

La idea primaria que se quiere informar (transmitir) debe indefectiblemente estar procesada. Este procesamiento estará en relación con el medio en que se emitirá, ya sea acústico, visual o emocional.

2. Fuentes de información

Quien crea una idea y además desea darla a conocer es una fuente de información. En la bibliotecología esta definición se perfecciona; toda fuente de información es aquel soporte de datos. A su vez, este soporte de datos puede dividirse en: materiales bibliográficos, pictográficos, sonoros, etc. Aparece así el procesamiento de libros, revistas, discos, cintas, mapas, monedas, etc.

Para Abramson (1975), dentro de los límites precisos de la ciencia de la información las fuentes pueden ser de dos tipos:

- 1) Aquellas en que la emisión se hace por grupos de símbolos y reciben el nombre de “extensiones de la fuente de memoria nula”.
- 2) Las que presentan vínculos condicionantes de cada símbolo que surge y se denominan “fuentes de Markov”.

¹ Licenciado en Bibliotecología del Centro Nacional de Tecnología y Productividad Industrial. uy33157@antel.com.uy // caiz@cnpil.gov.uy

Estas fuentes, indistintamente pueden pertenecer o responder a procesos ergódicos o no. El proceso es ergódico cuando al ser observado durante determinado tiempo emite con certeza una secuencia típica de símbolos. Por consiguiente el proceso ergódico es parte de lo normal.

3. Degradación de la información

Este es un aspecto problemático dentro de la transmisión. Luego del proceso necesario para su difusión, no siempre llega al receptor una información en su totalidad: llega deformada y a veces incomprensible. Se está entonces frente a una degradación de la información. Según Peluffo (1982), esta degradación puede ser progresiva y llegar a un punto máximo o caótico, denominada “entropía de la información”.

Para solucionar esto es necesario un proceso de regeneración y se logra tomando como base un concepto relacionado con la información misma: a la información, dividida en unidades se le pueden quitar elementos o símbolos hasta llegar a un punto mínimo, dejando sólo los fundamentales para que el mensaje sea comprendido. A esta información remanente que permite igualmente comprender la idea, se le da el nombre de “redundancia”. Teniendo en cuenta este factor de redundancia, se han desarrollado los “regeneradores” que se basan en el conocimiento del lenguaje que usa esa información y que permiten recomponer el mensaje o idea en su forma inicial.

4. El lenguaje

Al hablar de lenguaje en el campo de la información no debe entenderse éste sólo en el sentido natural. Se deben tener en cuenta aquellos lenguajes artificiales, que se adaptan a la naturaleza de las fuentes de información.

Estos lenguajes requieren estudios y preparaciones especiales para así formar códigos normalizados. Son conocidos por los bibliotecólogos y documentalistas como tesauros, listas de epígrafes, encabezamientos de materia, etc.

Como en su forma original las ideas no siempre están en el lenguaje normalizado, es necesario codificarlas. Esto implica una etapa posterior e inversa, o sea la descodificación. Dicho de otro modo: la idea original se codifica para su proceso de conservación y recuperación, y luego es descodificada por el receptor para su utilización.

5. Canales de información

La información debe transmitirse imprescindiblemente del emisor al receptor. Para ello utiliza los canales que están en función del medio que separa a ambos.

El canal de información será un procesador (manual o mecánico) con su entrada (in-put) y su salida (out-put). Para la entrada al receptor es menester que la información esté codificada como se ha mencionado anteriormente. Por lo tanto, la demanda de información debe realizarse en ese mismo lenguaje codificado, la descodificación corre por cuenta del receptor.

Es esta etapa es fundamental la actividad del técnico en la ciencia de la información (bibliotecólogo, documentalista). Será él quien conozca a fondo los lenguajes de codificación para ingresar y recuperar luego la información.

Un canal óptimo o perfecto es aquel en el cual la información no se degrada; sin embargo, podría decir que es normal que un canal no sea óptimo en su totalidad.

Weaver y Shannon (1975), llamaron “canal ruidoso” al canal no perfecto. Profundizando sus investigaciones, arriban a teoremas fundamentales:

- 1) En un canal de información hay un valor máximo teórico de la cantidad de información que puede cursar.
- 2) Por un canal ruidoso es posible transmitir cierta cantidad de información sin error.

Este último es considerado como “Teorema fundamental de la información”.

Bradford (1961), utiliza una idea similar, empleando los términos ruido y silencio. Cuando se seleccionan documentos (ideas) pueden extraerse aquellos que son pertinentes junto a otros no pertinentes. Al porcentaje de los no pertinentes llama “ruido” y a los pertinentes “silencio”. O sea que cuanto más silencioso sea el mensaje, más pertinente u óptimo será.

También Cleverdon (1963) tiene en cuenta el ruido y el silencio, pero va más allá de la simple mención. Crea una fórmula que permite medirlos. Para ello utiliza cuatro factores que están relacionados dos a dos y a la vez son contrarios: 1) Exhaustividad; 2) Economía; 3) Pertinencia; 4) Rapidez.

Realiza entonces el siguiente cuadro para los documentos procesados:

	Pertinentes	No pertinentes
Documentos extraídos	A	B
Docum. No extraídos	D	C

Es sabido que cuando una información es requerida no siempre se extraen todos los elementos existentes, por eso los divide en documentos extraídos y no extraídos. Los no extraídos se deben posiblemente a errores producidos en la codificación previa a su proceso.

En base al cuadro, se obtiene la siguiente fórmula para medir el nivel de optimización:

$$\frac{a}{a + b}$$

También se puede medir el nivel de exhaustividad de la información según la siguiente fórmula:

$$\frac{a}{a + d}$$

Generalmente el resultado será inferior a 1. En una primera apreciación se consideraría este un factor negativo; pero teniendo en cuenta que el segundo teorema de Shannon, comprobamos que se está frente a una posibilidad que cabe dentro de los límites normales.

6. Acumulación de la información

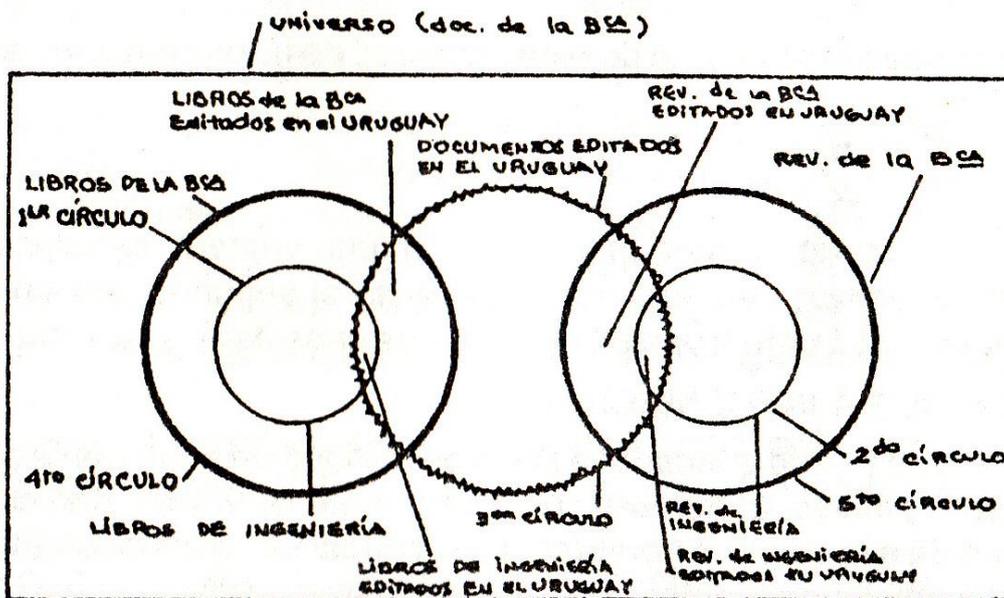
La información no puede concebirse sin la posibilidad de acumularla para su recuperación en el momento preciso. El hombre como integrante de la sociedad crea información que a la vez está unida a la de otros hombres. Este caudal se va almacenando en la cultura de un pueblo. Los medios para contenerlas son las “memorias” que son de diversos tipos: libros, revistas, discos, cintas magnetofónicas, etc.

7. Lo aportes de Boole y Venn

Ya que la información y particularmente la teoría de la información utiliza el producto intelectual, puede decirse que se relaciona con la filosofía y la lógica. Boole (1970) lo comprobó en 1854 analizando el razonamiento humano y convirtiéndolo en expresiones matemáticas. Fundamentó su trabajo en el sistema de razonamiento lógico que se había desarrollado a lo largo de los siglos, desde la época de Aristóteles. Uno de los puntos fundamentales de este sistema de silogismo o método de razonamiento basado en la obtención de una conclusión derivada de dos premisas relacionadas. Por ejemplo: <<todos los triángulos tienen tres lados. Esta figura es un triángulo. Luego, esta figura tiene tres lados>>.

Venn (1970), perfeccionó el método de Boole creando un método gráfico para representar las ideas y sus relaciones. Hoy se conocen como “diagramas de Venn”, y tienen gran utilidad para establecer conclusiones lógicas. La forma común del diagrama se realiza dibujando primeramente un rectángulo que simboliza el conjunto total o universal. Por ejemplo, si se trata de documentos existentes en la biblioteca. Un círculo en su interior, representa, por ejemplo, todos los libros de ingeniería, y un segundo círculo a todas las revistas de ingeniería. Obviamente ambos círculos no se cortan pues no hay miembros comunes en ambos conjuntos (un conjunto es de libros y el otro de revistas). Sin embargo si se desea simbolizar con un tercer círculo a todos los documentos editados en el Uruguay, este último deberá cortar a los dos primeros ya que este conjunto contiene miembros de los dos anteriores. Si por otra parte se desea representar con un cuarto círculo a todos los libros de la biblioteca (sin tener en cuenta el tema), éste deberá incluir completamente al primero y análogamente, un quinto círculo que simbolice a todas las revistas deberá incluir a todo el segundo círculo de las revistas de ingeniería y a una parte del tercero, las revistas de ingeniería editadas en el Uruguay, no incluidas en el círculo de libros.

Los diagramas de Venn Permiten comprobar en forma gráfica las relaciones lógicas que son las que mentalmente realiza el bibliotecólogo o documentalista para satisfacer las necesidades de información de la comunidad en la cual está inmerso.



Bibliografía

- Abramson, N. 1975. Information theory and coding. New York, Mac-Graw Hill.
- Bradford, S.C. 1961. Documentação. Río de Janeiro, Fundo de Cultura.
- Boole, Venn. 1970. El álgebra. Enc. Mente Sagaz, Barcelona, (5): 1781-1784.
- Cleveron, C.W., Milss, J. 1963. The analysis of index language devices. ADI. Annual Meeting. Proceedings, Part 3: 451-454. Washington.
- Peluffo Cánepa, L.N. 1982. Teoría de la información y sus implicancias. Rev. Ingeniería, Montevideo (36):21-29, dic.
- Weaver, W., Shannon, C.E. 1975. Theorie mathématique de la communication: preface de Abraham A Moles, Paris, Retz.