



Carlos Olmeda-Gómez



Carlos Olmeda-Gómez es licenciado en historia moderna y contemporánea por la *Universidad Autónoma de Madrid*, master en documentación en ciencias de la información y doctor en filología hispánica por la *Universidad Complutense de Madrid*. Profesor titular del *Departamento de Biblioteconomía y Documentación* de la *Universidad Carlos III de Madrid*. Realiza estudios de contenido bibliométrico, colaboración científica y análisis de dominios. Participa en diversas iniciativas y proyectos en el *SCImago Research Group*.
<http://orcid.org/0000-0001-5955-6423>

Universidad Carlos III de Madrid
Departamento de Biblioteconomía y Documentación
Madrid, 128. 28903 Getafe (Madrid), España
carlos.olmeda@uc3m.es

Resumen

Visión general y sintética que introduce algunos conceptos básicos y ejemplos de visualizaciones de información. Se muestran algunas visualizaciones con datos geoespaciales, relacionales y textuales de una disciplina exigente. Se exponen conceptos y los ejemplos ilustran algunas prácticas y criterios.

Palabras clave

Visualización de información, Visualización de textos, Visualización de datos, Tipos de visualización, Gráficos, Visualización.

Title: Information visualization

Abstract

This overview introduces and synthesizes some basic concepts of the broad spectrum of information visualization, providing examples with geospatial, relational and textual data from a demanding discipline. The concepts are discussed and the examples illustrate some practices and criteria.

Keywords

Information visualization, Text visualization, Data visualization, Visualization types, Graphics, Visualization.

Olmeda-Gómez, Carlos (2014). "Visualización de información". *El profesional de la información*, mayo-junio, v. 23, n. 3, pp. 213-219.

<http://dx.doi.org/10.3145/epi.2014.may.01>

1. Introducción

La visualización de información de acuerdo con **Card, Mackinlay** y **Shneiderman** (1999), se define como "el uso de representaciones visuales interactivas de datos abstractos, no físicos, para ampliar el conocimiento". En otras palabras, se centra en la extracción de estructuras de grandes conjuntos de datos, que puedan ser representadas de forma visual a los usuarios. En consecuencia, en las visualizaciones de información intervienen distintos procesos. Hay una transformación de datos brutos en abstracciones analíticas que, a continuación, se transforman en un modelo espacial-visual abstracto, para que finalmente, mediante procesos de diseño visual, el modelo visual se presente al usuario de forma gráfica y visible.

Cada proyecto de visualización difiere en sus fases o etapas, dependiendo de:

- si los datos están estructurados o no;
- las técnicas a emplear para la limpieza, selección y almacenamiento de los datos;
- los algoritmos empleados para simplificar las vistas;
- la definición de los parámetros del escenario donde se presentan; y
- el propósito de uso.

Se supone que la obtención de visualizaciones de información puede ayudar al reconocimiento de patrones, de estructuras, tendencias y relaciones que ayuden a la comprensión de los datos y del fenómeno que representen, que de otro modo serían incomprensibles por su complejidad.

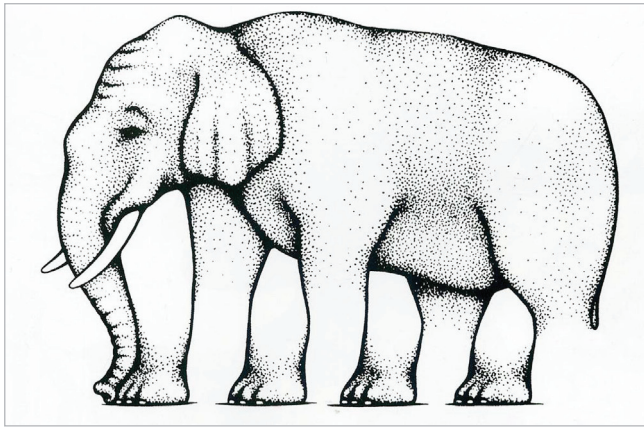


Figura 1. <http://en.wikipedia.org>

Por comprensión se puede entender la construcción de un concepto, de un modelo mental a partir de los datos o del fenómeno que se analiza. Una visualización de información captura las características esenciales de los datos, no presenta los items de forma individual y por eso el modelo es necesariamente abstracto, pero de forma simultánea, las herramientas de visualización deben facilitar la percepción de subconjuntos de datos, items o de unidades.

Las herramientas de visualización deben facilitar la percepción de subconjuntos de datos, items o de unidades

El observador de la visualización puede querer emplear o usar la visualización para explorar, para confirmar una hipótesis o presentar los resultados a una audiencia; para detectar anomalías, es decir datos que no se comportan de forma consistente con las expectativas; descubrir grupos o *clusters*, datos que comparten características similares de algún tipo, que de ese modo pueden indicar la presencia de un fenómeno en particular; o tendencias, es decir, datos que cambian de tal modo que puede ser característico y, por consiguiente, ser usado en modelos predictivos. El observador no es sino un *lector* que debe interpretar la visualización a partir de la forma física externa de la estructura y de los atributos de los datos que se usan para definirlos de forma visual como objetos. Fundamentalmente hay cuatro clases de atributos:

- valores visuales (posición, longitud, ángulo, dirección, forma, área, volumen, saturación y color);
- sistema de coordenadas (cartesiano, polar y geográfico);
- escala (numérica, categórica, temporal); y
- contexto (la información que proporciona el mayor entendimiento de quién, qué, cuándo, dónde y el porqué de los datos), que puede aclarar los datos a los lectores y guiarlos en la dirección correcta.

Otro componente característico de los procesos de visualización de información se refiere al modo en que se proporcionan al usuario controles para la manipulación interactiva o la selección de variables, atributos o parámetros. Las visualizaciones de información son procesos dinámicos,

donde se permite al usuario en su grado máximo, el control virtual de todas las fases de los procedimientos, desde la selección y control de los datos que se desean ver, hasta la posible selección del color o el modo de afinar las vistas. No existe una fórmula que garantice la efectividad de una visualización. Los usuarios son diferentes en su formación y grado de familiaridad, en sus habilidades de percepción y en sus preferencias, y se diferencian por las opiniones que les susciten las distintas visualizaciones. Las tareas de los usuarios también afectan a la utilidad de las visualizaciones. Un cambio en los datos puede tener consecuencias en el resultado. Por eso es obligado permitir que el usuario pueda adaptar, modificar o refinar de forma interactiva la visualización, hasta que crea que ha logrado sus objetivos, por ejemplo porque ha extraído una descripción completa, rigurosa, exacta de lo que contienen los datos o porque ha obtenido una representación clara de los patrones que los datos transmiten.

Un aspecto crítico está relacionado con las capacidades cognitivas del usuario y con las habilidades y las limitaciones propias de nuestro sistema visual. Si el objetivo de la visualización es ayudar a comprender fenómenos complejos de información mediante su transformación en objetos visuales abstractos, es esencial tener en consideración las habilidades de percepción (**Ware**, 2012) que nuestro sistema nos facilita o con el que, en ocasiones, nos traicionamos. Todo el mundo conoce la expresión “una imagen vale más que mil palabras” o el modo en que un buen dibujo puede ser estimulante para la comprensión de un fenómeno, pero si se presentan resultados de visualización de información, nadie desea ambigüedades como cuando observamos el elefante de múltiples patas de **Shepard** (1990) (figura 1).

Los usuarios interactúan con las visualizaciones basándose en lo que se ve y en lo que se interpreta

Los usuarios interactúan con las visualizaciones basándose en lo que se ve y en lo que se interpreta. Comprender cómo vemos, debería ayudarnos a producir mejores dispositivos de presentación de visualizaciones de información o al menos, no construirlas con deficiencias. En un mundo moldeado de forma creciente por algoritmos, la visualización de información juega un papel relevante para dotarle de sentido. Los autores relevantes en este aspecto de la percepción proceden de las ciencias cognitivas, la psicología, la psicofísica, la educación y el diseño gráfico. La comprensión de los procesos que están implicados en la percepción visual y de los aspectos cognitivos conexos es obligada para generar visualizaciones efectivas y eficaces (**Costa**, 1998).

La representación gráfica de información cuantitativa tiene también raíces profundas que se basan en las historias que contienen la cartografía temática, la estadística gráfica y la visualización de datos. Se pueden reconocer antecedentes en la generación de diagramas geométricos, de los mapas y las cartografías que ayudaron a la navegación y a la exploración en los siglos de los descubrimientos o de la *Cosmo-*

graphia de Ptolomeo. También en las técnicas e instrumentos precisos para la observación y la medición de cantidades, así como de las notaciones y fórmulas necesarias para su expresión. Otros antecedentes se sitúan en la invención de la geometría analítica, la estadística demográfica y la aritmética política, todos ellos orientados para tratar de comunicar con mayor facilidad o facilitar la inspección visual (Tufte, 1990).

2. Visualización de datos geoespaciales

Los datos geoespaciales son diferentes de cualquier otro tipo de dato, ya que describen objetos o fenómenos con una localización específica en el mundo real. Los objetos de información susceptibles de ser ubicados en un espacio, pueden proceder de múltiples aplicaciones, como las transacciones financieras, solicitudes de créditos, comunicaciones en *Twitter*, registros medioambientales, censos demográficos o votaciones políticas. Son datos que pueden ser potencialmente geolocalizados por las denominaciones de los lugares, direcciones o códigos postales y ser vinculados a una ubicación geoespacial en la pantalla. Con mucha frecuencia los datos geolocalizables se ubican en mapas temáticos. Los mapas constituyen la forma más familiar de representar objetos de información geolocalizable, que una vez ubicados en el espacio pueden revelar patrones y relaciones difíciles de descubrir o de ser apreciados de otro modo.

Los mapas pueden subdividirse basándose en las propiedades de los datos: cualitativos frente a cuantitativos; dis-

cretos o continuos, y por las propiedades de las variables gráficas: puntos para localizar, líneas para interconectar, tamaños, volúmenes. Ello da como resultado la generación de mapas simbólicos, de mapas de puntos, mapas de área y cartogramas, mapas coropléticos (con colores según densidad) e isopléticos (con líneas de igual valor), mapas de línea o de superficie (figura 2).

En la visualización de objetos de información geolocalizable las proyecciones elegidas juegan un papel crítico. Las proyecciones en los mapas tienen que ver con la transposición de la localización de los objetos de información en un globo, que es esférico, a una superficie plana como es la pantalla que es interactiva. Las proyecciones de los mapas se pueden clasificar dependiendo del tipo de superficie sobre el que la esfera se proyecta. Las superficies de proyección en la pantalla son cilíndricas, planas y cónicas y en otros contextos, cartesianas, polares o hiperbólicas (figura 3).

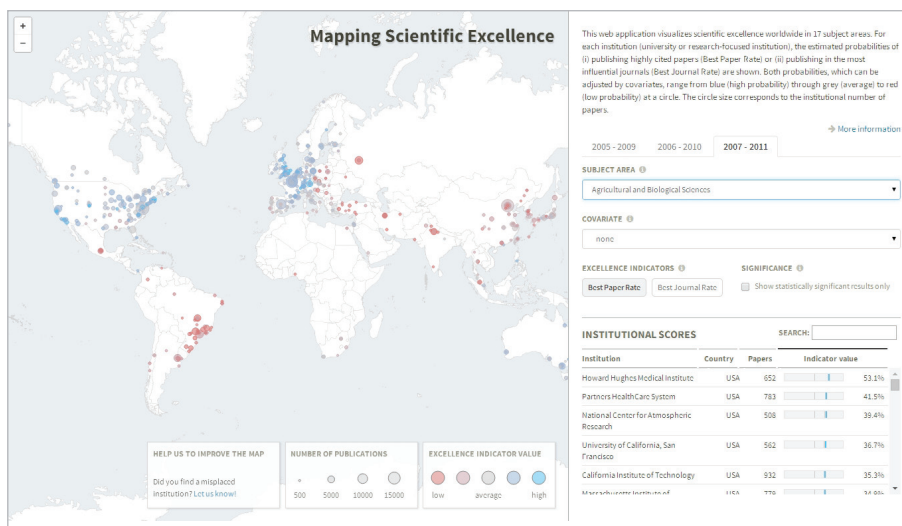
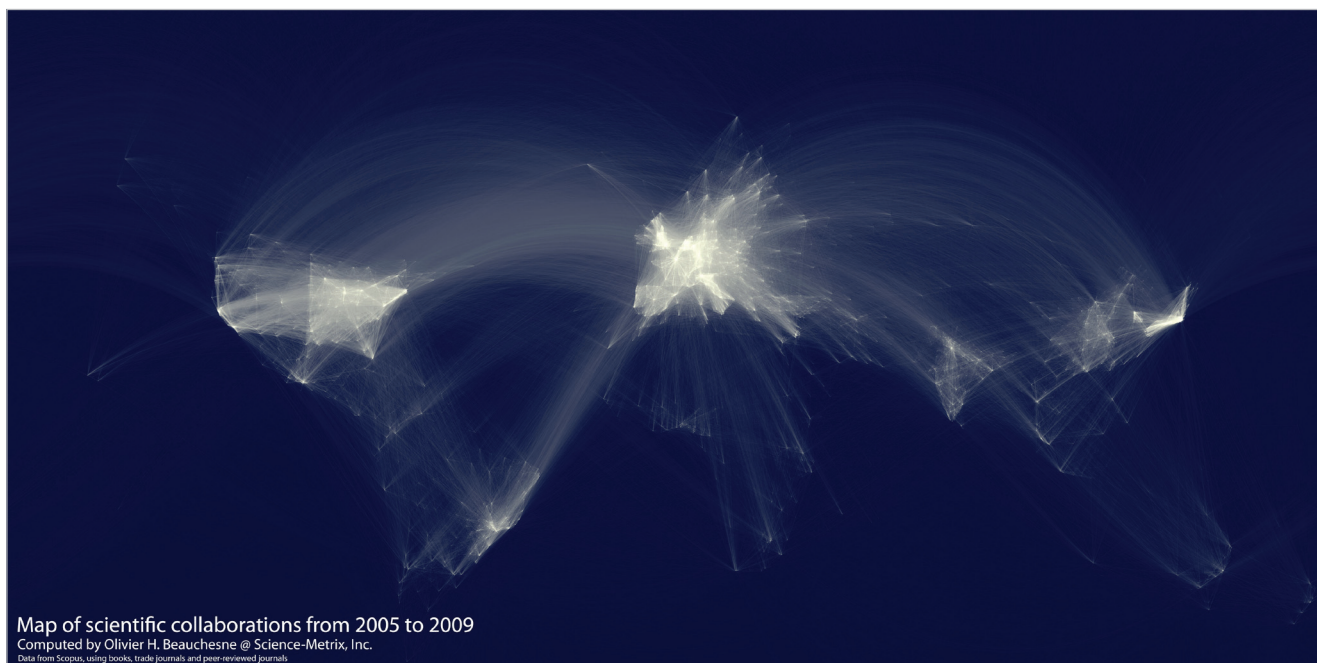


Figura 2. <http://www.excellencemapping.net>



Map of scientific collaborations from 2005 to 2009
 Computed by Olivier H. Beauchesne @ Science-Metrix, Inc.
 Data from Scopus, using books, trade journals, and peer-reviewed journals

Figura 3. <http://olihb.com>

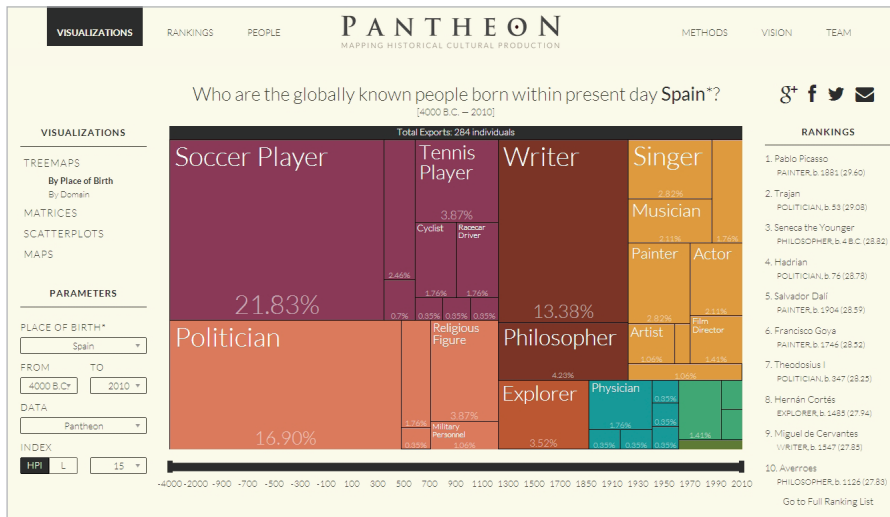


Figura 4. <http://pantheon.media.mit.edu>

Los mapas se usan de múltiples formas. Por ejemplo para proporcionar información específica sobre posiciones concretas, para generar información general sobre patrones espaciales o para comparar patrones en múltiples mapas. Por ello es obligado que las variables visuales reflejen las propiedades de los datos espaciales. Las variables visuales para datos espaciales son el tamaño de los símbolos utilizados, la anchura de las líneas o los tamaños de las áreas; la forma de los símbolos individuales o los patrones de las líneas y las áreas. Otras variables son el brillo, el color, la textura, la orientación, el tipo de perspectiva y la disposición de los patrones en el seno de los símbolos individuales (por ejemplo, distribuciones regulares o aleatorias de símbolos).

En sistemas de geovisualización exploratorios la interacción con los mapas es crucial. En contraste con la cartografía tradicional, la clasificación y el mapeo de datos pueden adaptarse para que sea el usuario quien interactúe con los mapas, al mismo tiempo que puede filtrar, ojear y formular preguntas desde la interfaz o manipular la presentación en la pantalla, por ejemplo mediante zoom. De este modo se

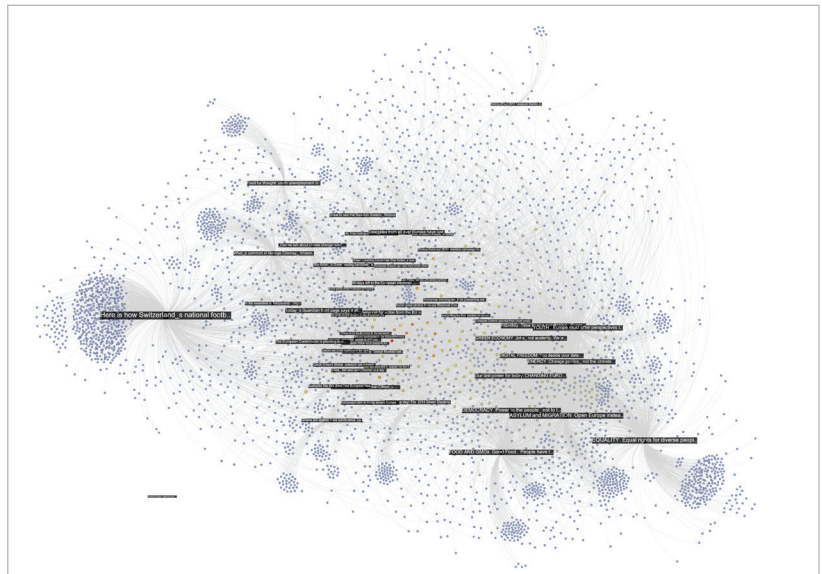


Figura 5. <http://www.theguardian.com/uk>

pueden lograr mapas interactivos de objetos de información. Los diseñadores de mapas interactivos parten de un mapa base que constituye el soporte, sobre el que se superponen capas de datos, que llevan incorporados los valores de coordenadas y de otro tipo.

3. Visualización de estructuras jerárquicas y de redes

La visualización de grafos, simples, elegantes, directos, pueden informarnos, sin conocer nada sobre los datos subyacentes y, sin ni siquiera intentarlo, se descubren relaciones y patrones. Los nodos y los enlaces pueden ser conexiones

entre miembros de *Facebook*, *Twitter*, redes de distribución de datos, llamadas a centralitas telefónicas, mensajes de correo electrónico o interconexiones en gramáticas. Todas son redes, entendidas como entidades interconectadas, y todas ellas se pueden analizar mediante generación de visualizaciones de grafos, pudiéndose disponer como interfaz visual de acceso a bibliotecas digitales, ya que la diferencia entre visualización e interfaz, se está borrando.

La visualización de redes se centra en las interrelaciones de la información relacional. Las interrelaciones pueden adoptar distintas formas:

- jerárquicas (padre/hijo; partes/sub-partes);
- conectividad (interconexiones entre lugares; redes de ordenadores);
- derivaciones, a partir de una secuencia de pasos o de fases (redes de citación);
- clasificaciones compartidas;
- semejanza entre valores o en atributos (por ejemplo, espaciales, temporales).

Las relaciones pueden ser simples o complejas, directas o bidireccionales, ponderadas o sin ponderar, ciertas o con grados de incertidumbre. Las relaciones pueden proporcionar más información que la que contienen los registros de las bases de datos tomados de uno en uno. Hay información en las conexiones. A simple vista se identifican nodos que atraen un número superior de enlaces, los nodos se pueden arracimar y constituir distintos subgrupos o se pueden detectar nodos aislados que carecen de conexiones.

Los árboles o jerarquías son unas de las estructuras más comunes que contienen información relacional. La visualización de jerarquías es una de las ramas más maduras y con mayor actividad de la visualización de información. Los conjuntos de datos de este tipo, proceden de estruc-

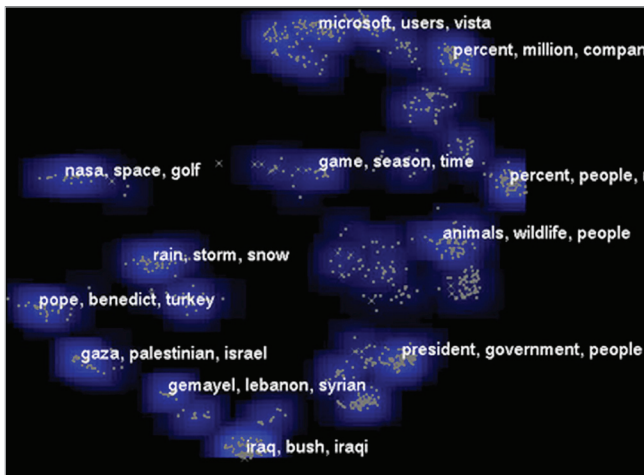


Figura 9. <http://in-spire.pnnl.gov>

4. Visualización de textos

La visualización de textos es una clase de técnica de visualización de información que transforma objetos contenidos en los *corpora* o colecciones de documentos, en un formato visual. Estos objetos pueden ser palabras, sentencias, párrafos, atributos y metadatos, a nivel individual o en colecciones. Si se identifican entidades importantes en el seno de los textos, se puede analizar su contenido en tres niveles de representación: en un nivel léxico, sintáctico o semántico. Por ejemplo mediante la obtención de frecuencias o distribuciones de palabras; el contenido semántico y repeticiones; el tema o los temas que definen conjuntos de documentos o los temas centrales de los documentos. También pueden detectarse semejanzas mediante conexiones, cambios o fenómenos de difusión de la información.

La visualización de un solo documento permite conocer los temas centrales del contenido y las características del texto de forma rápida. La forma más sencilla de visualización se conoce como nubes de textos o palabras, y no son sino representaciones de la frecuencia en la que aparecen términos, en los que se usan vectores de los términos y algún algoritmo de diseño para que el usuario modifique la fuente, el color o la configuración (figura 7).

En los análisis de la literatura es un lugar común analizar un texto o una colección de textos, extrayendo una concordancia: un índice de todas las palabras que contiene un texto, en el que se muestran las palabras en el contexto en que aparecen. Es un modo similar a presentar palabras claves en contexto (KWIC, *keyword in context*). Permite comprender la frecuencia de términos y su distribución, o la creación de estructuras de conocimiento en bases de datos de naturaleza léxica o de otro tipo. Otro tipo de visualización puede reflejar el contenido semántico, es decir la representación semántica asociada al lenguaje. Normalmente se extraen tripletas sujeto-verbo-atributos del objeto, después de pre procesar el texto, extraer entidades, resolver anáforas (elementos que dependen del contexto) y normalizar de forma semántica. Por últimos se les dota de apariencia visual, en forma de redes semánticas (figura 8).

Buscamos texturas de datos con valor, en contextos comparables, que ayuden a comprender lo que la complejidad revela, con economía de medios

En la mayoría de los casos, el objetivo de la visualización de documentos es ubicar documentos semejantes cerca unos de los otros, después de computar su semejanza entre pares de documentos y definir un diseño de la vista. Se trata de mapear matrices de términos muy complejas y traducirlas en representaciones de menores dimensiones que retengan tanto significado como sea posible. Normalmente se emplean métodos de descomposición y de factorización, como el de descomposición de valores singulares (*singular value decomposition, SVD*), análisis de componentes principales o mapas autoorganizativos (Vargas-Quesada; De-Moya-Aneón, 2007).

La visualización de los temas que se contienen en grandes colecciones persigue descubrir temas concretos o las relaciones entre diversos temas. Se emplea en el descubrimiento de disciplinas emergentes, evoluciones, detección de ten-

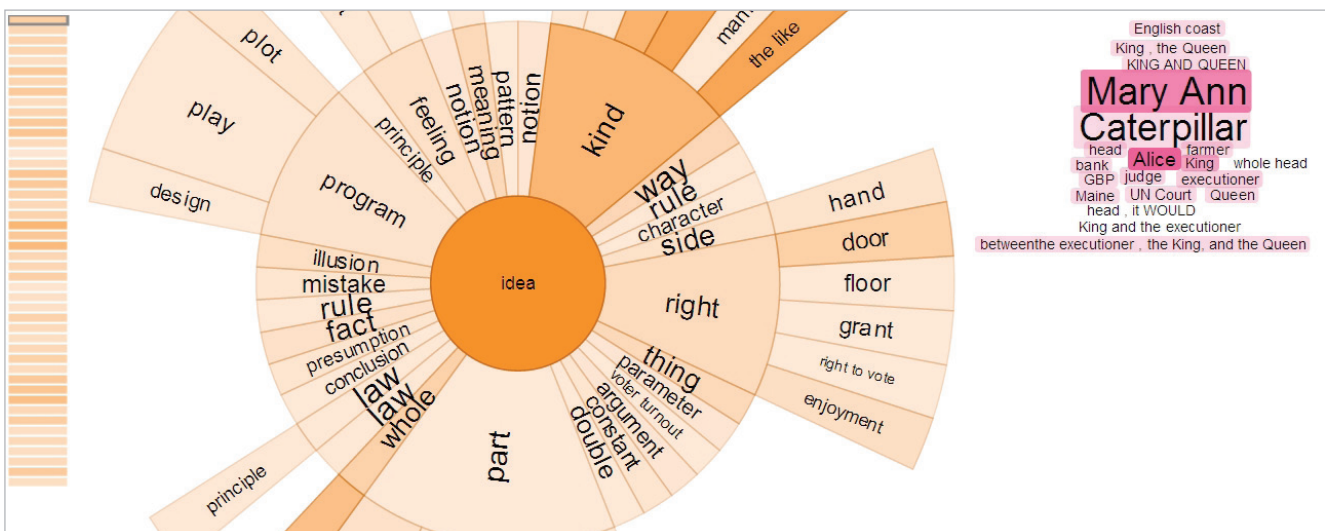


Figura 10. <http://vialab.science.uoit.ca>

dencias o rupturas de paradigmas (Chen, 2006). Son resúmenes de *corpora* que emplean paisajes en 2D o 3D, en el que el peso y el color se usan para representar densidades de documentos semejantes. Otros sistemas emplean redes, tablas o mapas geoespaciales (figura 9).

Los textos electrónicos cambian en sus distintas versiones, y visualizar las diferencias entre las múltiples adaptaciones puede ser objeto de un tipo de visualización específica. Las visualizaciones deben asumir que los datos de entrada progresan a lo largo del tiempo. Los diseños visuales de datos textuales temporales muestran las diferencias entre las múltiples versiones, las correcciones efectuadas, los patrones de cooperación en el caso de textos electrónicos con multiautoría (*Wikipedia*), el predominio o no del anonimato frente a las autorías reconocidas, las formas de negociación de contenidos, las zonas de texto estables o inalteradas, las procedencias geográficas de los lectores... (figura 10).

La visualización de literaturas científicas y las relaciones entre citas de los autores que escriben artículos científicos, es un campo sumamente amplio de análisis. La visualización de las formas en las que los autores, las revistas y los temas citan y son citados da origen a la generación de grafos de coautoría, mapas de la ciencia o sistemas interactivos con interfaces visuales para acceder a información bibliométrica (figura 11).

En un mundo moldeado de forma creciente por algoritmos, la visualización de información juega un papel relevante para dotarle de sentido

5. Final

La representación de estructuras de información multidimensional mediante sistemas de información no es un asunto trivial. Los procesos de diseño requieren tanto métodos de razonamiento analítico, como espaciales y visuales, además de múltiples habilidades en codificación. La creación obliga a codificar, el uso a decodificar. Las visualizaciones de información generadas por ordenador, dinámicas e interactivas ya son una realidad tal y como señalaba Edward Tufte, y sin duda, nos acompañarán en el porvenir, porque “capturan una parte valiosa de la sutileza del itinerario humano”.

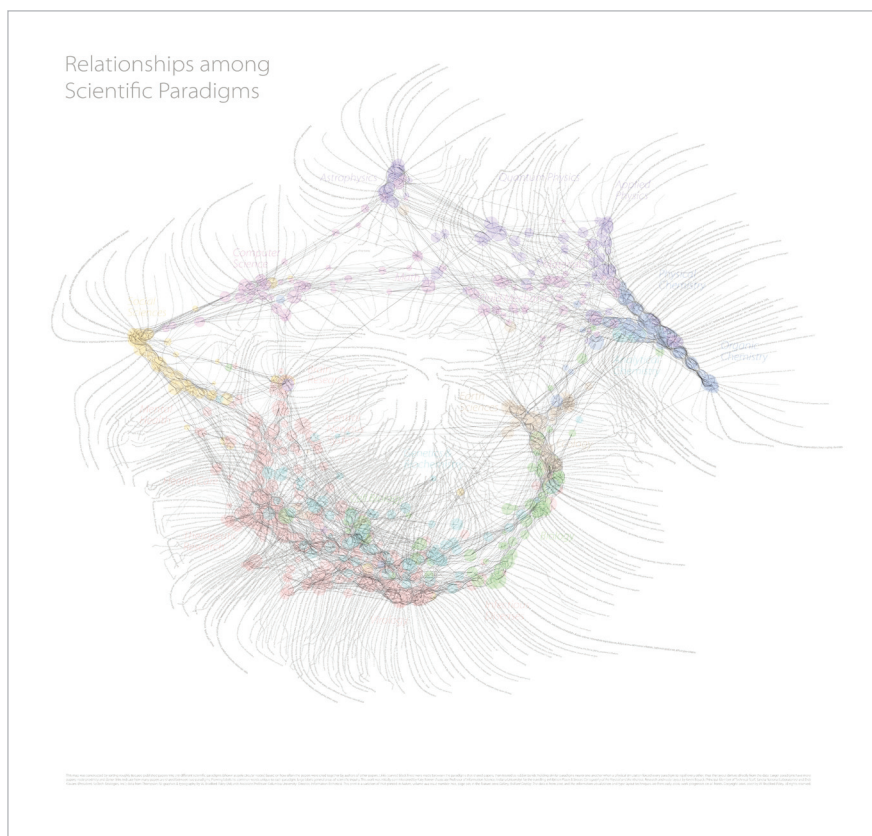


Figura 11. <http://wbpaley.com>

6. Bibliografía

Card, Stuart K. (1999). *Readings in information visualization, using vision to think*. Stuart K. Card, Jack D. Mackinlay, Ben Shneiderman (Eds.). San Francisco: Morgan Kaufmann. ISBN: 15 586 0533 9

Chen, Chaomei (2004). *Information visualization: Beyond the horizon*. London: Springer. ISBN: 18 523 3789 3

Costa, Juan (1998). *La esquemática. Visualizar la información*. Barcelona: Paidós. ISBN: 84 493 0611 6

Shepard, Roger N. (1990). *Mind sights: Original visual illusions, ambiguities, and other anomalies. With a commentary on the play of mind in perception and art*. New York: WH Freeman & Co. ISBN: 9780 716 72133 8

Shneiderman, Ben; Plaisant, Catherine (2006). *Diseño de interfaces de usuario: estrategias para una interacción persona-computadora efectiva*. Madrid: Pearson Addison-Wesley. ISBN: 9788 420 54803 6

Tufte, Edward R. (1990). *Envisioning information*. Cheshire: Graphic Press.

Vargas-Quesada, Benjamín; De-Moya-Anegón, Félix (2007). *Visualizing the structure of science*. Berlin: Springer. ISBN: 9783 540 69727 5

Ware, Colin (2012). *Information visualization: perception for design. 3rd ed.* Waltham (MA): Morgan Kauffmann. ISBN: 15 586 0511 8

Colección de libros de bolsillo

El profesional de la información (Editorial UOC)



El profesional de la
información

 **EDITORIAL UOC**

Más información:

<http://www.elprofesionaldelainformacion.com/libros.html>