

# Capítulo 1

## Introducción

Este proyecto de tesis presenta una interfaz de visualización para el entendimiento de grandes colecciones digitales. El modelo de visualización está centrado en la red neuronal conocida como mapas auto-organizados y la visualización de categorías mediante diagramas de Voronoi.

El objetivo principal de la interfaz es presentar al usuario un panorama general de las colecciones, facilitando la comprensión de la información presentada al descubrir relaciones entre los elementos de las colecciones. Esto, con la finalidad de ayudarlo en sus actividades de aprendizaje, descubrimiento de conocimiento e investigación.

### 1.1 Motivación y contexto

El manejo y la visualización de grandes cantidades de información son dos de los principales retos de las áreas de interacción humano-computadora (IHC), graficación, diseño visual y psicología. En especial, esta tesis se enfoca en el área de IHC.

Definida como *“el uso de representaciones visuales e interactivas de datos abstractos, soportadas por un computador para amplificar el conocimiento”* (Card, Mackinlay & Shneiderman, 1999), la visualización de información no es más que una entrada al mundo abstracto de la información. Permite formar imágenes mentales de conceptos de naturaleza abstracta, como las relaciones existentes entre los documentos de una colección, el grado de colaboración entre investigadores o la cercanía semántica entre dos líneas de investigación.

Uno de los principales objetivos de la visualización de información, es facilitar la comprensión de la información presentada (Card, et al., 1999). Y a pesar de que la visualización de información requiere de complejos procesos computacionales, algoritmos y sofisticadas técnicas de diseño, la finalidad última debe ser proveer al usuario una manera de entender los contenidos a los que se refiere la representación visual.

La velocidad con la que las redes científicas, académicas y tecnológicas generan información hace imposible su organización y clasificación manual. Es por ello que surge la necesidad de mecanismos automáticos que realicen esta tarea, dando como resultado información organizada en grandes colecciones digitales. A menudo, estas colecciones son heterogéneas y se encuentran almacenadas de manera distribuida.

Un aspecto relevante de la visualización es su potencial para mostrar características de la información que de otra forma no serían evidentes. Tales características, presentadas de manera apropiada, fomentan el entendimiento de colecciones. Es decir, que al mostrar los elementos de una colección a través de una adecuada visualización, el usuario entiende su contenido de manera general.

En lo particular, esta tesis se desarrolló debido a la necesidad de entender colecciones de documentos a partir del descubrimiento de las relaciones y la cercanía semántica entre tales documentos. Sin embargo, la aplicación del esquema de visualización presentado en este trabajo no se limita a la visualización de colecciones de documentos, sino también puede ser aplicado a la visualización de imágenes, de páginas web y de personas (redes sociales).

## **1.2 Hacia el entendimiento de colecciones**

La razón primaria de cualquier visualización de información es proveer precisamente de información al usuario. Tradicionalmente, la disciplina que ha satisfecho esta necesidad de información de los usuarios es la recuperación de información (RI), a través de modelos de búsqueda por consulta.

El modelo tradicional de RI ha funcionado y sigue funcionando gracias a la amplia investigación de algoritmos de recuperación que se adaptan a la creciente cantidad de información (p. ej. Google<sup>1</sup>). Sin embargo, los usuarios poseen diferentes maneras de adquirir nueva información de acuerdo a sus necesidades. Estas necesidades se tipifican en tres actividades: búsqueda simple, aprendizaje e investigación (Marchionini, 2006). Mientras que en la búsqueda simple se intenta dar una respuesta específica a una pregunta (consulta) específica del usuario, en el aprendizaje y en la investigación se intenta que el usuario adquiera conocimiento a través del descubrimiento, la exploración y la presentación de un panorama general de la información.

---

<sup>1</sup> GOOGLE: <http://www.google.com> recuperado el 1 de mayo de 2013

Existen trabajos (Marchionini, 2006; Wilson et al., 2010; Mirizzi & Di Noia, 2010; Nolan, 2008; Dörk et al., 2012; Tvarozek & Bieliková, 2008) que indican que la evolución de RI se dirige de la búsqueda simple al entendimiento complejo. Ello, combinando los esfuerzos de IHC y RI en interfaces interactivas para aprendizaje y descubrimiento exploratorio (Marchionini, 2006).

La combinación de RI e IHC es denominada por Marchionini (2006) búsqueda exploratoria. En otras palabras, sería la combinación de estrategias de búsqueda por consulta (perspectiva RI) y visualización para el entendimiento de colecciones (perspectiva IHC). A pesar de que ambos enfoques, tanto búsqueda por consulta y entendimiento de colecciones se contraponen en un sentido, su propósito conjunto es el de proveer al usuario de la información necesaria para cubrir sus necesidades de búsqueda, aprendizaje e investigación.

Por un lado, el entendimiento de colecciones se refiere a tener una idea general de ellas a partir de la visualización de los elementos que las constituyen desde una perspectiva amplia, permitiendo al usuario ganar una perspectiva global de la información y crear su propio entendimiento al derivar relaciones entre los elementos de las colecciones (Chang et al., 2004). Por el otro, la búsqueda por consulta provee de elementos específicos cuando el usuario ya tiene un conocimiento previo de la información.

Este proyecto de tesis ha dirigido sus esfuerzos hacia el entendimiento de colecciones, con la finalidad de contribuir en el mejoramiento de interfaces de visualización de información en la dirección resaltada por Marchionini (2006): de la búsqueda al entendimiento. Si bien, el alcance de esta tesis enfatiza el aspecto IHC de entendimiento de colecciones, se promueve su futura integración en interfaces de búsqueda exploratoria debido a su objetivo en común: proporcionar información al usuario a través de una visualización que fomente su aprendizaje y al mismo tiempo propicie el descubrimiento y entendimiento de nuevo conocimiento en tópicos complejos o desconocidos para él (Mirizzi & Di Noia, 2010).

### 1.3 Resultados y contribuciones

A continuación se presentan los resultados y principales contribuciones de este proyecto de tesis:

**Estado actual en interfaces de visualización de información:** Se realizó un estudio de las interfaces de visualización de información desde el punto de vista entendimiento de colecciones. De este estudio, se concluyó que las interfaces de visualización van en dirección de la búsqueda hacia el entendimiento de información. Sin embargo, en ese camino existen aspectos abiertos que pueden ser mejorados, tales como:

- La presentación de un panorama general que incluya los cuatro aspectos fundamentales del espacio de información: contenido, estructura, relaciones y cambios.
- La integración de navegación interactiva para promover la exploración y entendimiento de la información.

Considerando su relevancia para el enfoque del esquema de visualización propuesto, en este estudio se incluye investigación acerca de interfaces de visualización que hacen uso de mapas auto-organizados y de diagramas de Voronoi. De los mapas auto-organizados se destaca su utilidad para el entendimiento de colecciones y su exploración, además de que sintetizan y clasifican dinámicamente la información, por otro lado los diagramas de Voronoi destacan por su capacidad para segmentar el espacio y mejorar la transmisión de información visual.

**Esquema de visualización de información:** Con base en los aspectos abiertos del estudio del estado actual, se diseñó el esquema de visualización denominado VOROSOM.

El objetivo de VOROSOM, es proveer un panorama general de las colecciones, por medio de la integración de elementos visuales que faciliten al usuario la comprensión del contenido, la estructura, las relaciones y los cambios en un espacio de información (p. ej. colecciones digitales). Esto se logra, incorporando elementos interactivos que ayuden a navegar y explorar el espacio de información.

VOROSOM toma ventaja **de la clasificación dinámica y de la preservación de las relaciones y la cercanía semántica entre los elementos como originalmente se encuentran en el espacio de información**, características proporcionadas por los mapas

auto-organizados; y también hace uso de la **visualización de categorías** con regiones de Voronoi.

**Prueba de concepto:** Para demostrar el potencial y la viabilidad del esquema de visualización VOROSOM, se realizó una implementación prototípica que dio como resultado una interfaz web para visualizar colecciones de documentos. Posteriormente, se realizó un estudio de usabilidad con un grupo de ocho usuarios en sesiones de interacción constructiva.

A partir de los resultados del estudio de usabilidad realizado sobre la implementación prototípica, se puede concluir que VOROSOM es útil para el **entendimiento de colecciones de documentos ya que muestra un panorama general completo y propicia la recuperación de información semánticamente relacionada**. Esto, a través de la navegación jerárquica de las colecciones de una manera fácil e intuitiva (aprovechando la metáfora de mapas).

**Principal contribución:** El principal aporte de este proyecto de tesis, es la generación de un esquema de visualización que ha dirigido sus esfuerzos hacia el entendimiento de colecciones y comprensión de información. Tomando como base, las observaciones del estado actual en interfaces de visualización. El esquema (VOROSOM) integra los conceptos de panorama general, navegación y exploración, relaciones semánticas y visualización de categorías en una visualización interactiva, sencilla e intuitiva.

## 1.4 Organización del documento

El presente documento está organizado en seis capítulos, el presente capítulo es la parte introductoria y los subsecuentes detallan aspectos del desarrollo de esta tesis:

**Capítulo 2:** Este capítulo presenta un estudio del estado actual en interfaces de visualización desde el enfoque de entendimiento de colecciones. De igual manera, presenta un estudio de trabajos de visualización que incluyen mapas auto-organizados y diagramas de Voronoi, todo esto, con la finalidad de tener una perspectiva general del área y entender la motivación de esta tesis.

**Capítulo 3:** En este capítulo se detalla el diseño del esquema de visualización VOROSOM. Se presentan los objetivos de diseño, así como los principales conceptos que sustentan el esquema. Se exponen también las razones por las cuales se escogieron

mapas auto-organizados (SOMs) y diagramas de Voronoi como base para el diseño de la interfaz VOROSOM.

**Capítulo 4:** Este capítulo presenta cada uno de los pasos y herramientas que fueron utilizados para la implementación prototípica de VOROSOM. Se explica de manera general la arquitectura de la interfaz, cuáles fueron las herramientas utilizadas, y la integración de la interfaz a la Red Mexicana de Repositorios Institucionales (REMEDI).

**Capítulo 5:** en este capítulo se reportan los hallazgos de un estudio de usabilidad realizado sobre la implementación prototípica de VOROSOM. El estudio se realizó sobre un grupo de ocho usuarios en sesiones de interacción constructiva. Los usuarios realizaron una serie de tareas en donde fueron verbalizando sus opiniones acerca de la interfaz, y además, llenaron un cuestionario previo y un cuestionario posterior en donde plasmaron sus opiniones y sugerencias para la mejora de la interfaz.

**Capítulo 6:** en este capítulo se presentan las conclusiones generales obtenidas en este proyecto de tesis. Asimismo, se expone el trabajo en proceso y el trabajo futuro para la mejora de la interfaz de visualización VOROSOM.