

# Nubes de palabras animadas para la visualización de información textual de Publicaciones Académicas

Felicia De Lucia Castillo, Celso Alberto Saibel Santos

feldlucyca@gmail.com, saibel@inf.ufes.br

Universidade Federal do Espírito Santo, Brasil  
Av. Fernando Ferrari, 514, Goiabeiras, CEP 29075-910  
Vitória - Brasil

**Resumen:** *La forma cómo se presenta la información tiene una fuerte influencia en la forma en que se percibe. Por lo tanto, la Técnica de Visualización elegida para presentar la información debe ser compatible con el contenido que se presenta. Una visualización fácil y rápida de los temas desarrollados por profesores de un grupo de investigación, durante un período, requiere una Técnica de Visualización dinámica, como la Nube de Palabras Animadas, presentada en este artículo. Esta técnica nos permitió utilizar los títulos de publicaciones bibliográficas de investigadores para presentar, de una manera clara y directa, la variación de frecuencia de los temas investigados a través de cambios en el tamaño de los términos más relevantes. Esta Nube de Palabras Animadas, puede ser guardada como un archivo de vídeo en formato H.264 y hacer uso de ella en cualquier momento.*

**Palabras clave:** visualización, información, texto, nubes de palabras, words cloud.

**Abstract:** *The form of how the information is presented has a strong influence in how it is perceived. Therefore, the chosen visualization technique should be compatible with the content to be presented. A quick and easy visualization of the themes developed by teachers of a research group, during certain periods of time, requires a dynamic visualization technique such as the Animated Words Cloud. This technique allowed us to use the titles of bibliographic publications of the researchers to present, in a clear and direct way, the frequency variation of the researched themes through changes in the size of the keywords. The Animated Words Cloud created can be exported and saved as a video file in H.264 format to be accessed at any time.*

**Keywords:** visualization, data, text, word clouds, cloud words.

## 1 Introducción

La Visualización de Información puede ser tratada como un campo de conocimiento bien establecido, asociado al uso de representaciones visuales de datos abstractos que tienen como fin expandir el conocimiento [Strecker12] [Card99]. La Visualización de Información es una tecnología plural que consiste en transformar datos en información semántica o en la creación de herramientas para dicha transformación, basada en la combinación de señales de naturaleza icónica (figuras) con otros de naturaleza arbitraria y abstracta (textos, estadísticas, etc.) [Cairo11]. Siendo así, la Visualización de Información es caracterizada por transformar datos brutos en información importante, permitiendo que los usuarios interactúen, transformen e interpreten esas informaciones, al mismo tiempo que adquieren conocimientos. Ella surge como respuesta a la búsqueda de herramientas que favorezcan el análisis y la comprensión de los datos abstractos, a través de la utilización de la computación gráfica interactiva y el uso de Técnicas de Visualización [Martig02] [Nualart14].

Para muchos usuarios, seleccionar, procesar, analizar y visualizar grandes colecciones de textos es una actividad necesaria, pero termina siendo una ardua tarea debido a la cantidad excesiva de información que se tiene almacenada actualmente. Es por ello que se hace indispensable la creación de nuevas Técnicas de Visualización para ese tipo de información. En muchos casos, la presentación de información es un factor importante para que los usuarios entiendan lo que realmente se quiere presentar [Nualart14].

Buscando nuevas soluciones para este problema, este trabajo presenta una nueva forma de visualizar información, generada a partir de informaciones de tipo texto permitiendo la detección de patrones, comportamientos y evidencias a partir del dinamismo en la presentación visual de los resultados. El objetivo de esta visualización es ampliar la posibilidad que informaciones escondidas sean descubiertas, mejorando la forma de examinar y comprender informaciones textuales.

En el artículo, se presentan algunas técnicas utilizadas para la visualización de información textual, así como ejemplos de uso. También se presenta nuestra solución, una Nube de Palabras Animadas, para visualizar de forma interactiva y dinámica los temas investigados y estudiados por profesores en un intervalo, presentando informaciones que puede que no sean evidentes con sólo observar el texto. También se presentan los resultados del experimento realizado sobre la información de una base de datos real.

El trabajo está dividido en secciones. La sección 2 trata sobre los trabajos relacionados, la sección 3 trata sobre la Visualización de Información y sus Técnicas; en la sección 4, es descrita la Técnica Nube de Palabras Animadas, el proceso de desarrollo y los resultados obtenidos usando la Técnica de Visualización creada; en la sección 5, se presentan las conclusiones y trabajos futuros, y finalmente, la última sección presenta las referencias bibliográficas.

## 2 Teoría del dominio y trabajos previos

Actualmente, existen diversas herramientas disponibles para el análisis y la visualización de datos textuales. Esas

herramientas posibilitan, desde la obtención de informaciones simples sobre un texto, hasta la identificación de patrones y tendencias que llevan a *insights* significativos [Viegas08].

Una herramienta conocida y muy utilizada para la visualización de información es *Many Eyes* [Viegas07]. *Many Eyes* es una herramienta colaborativa, creada en el año 2007 por un grupo de investigación de la empresa IBM. Esta herramienta permite que los usuarios realicen una carga de datos para crear diversas visualizaciones a partir de ellos. Uno de sus principales objetivos es transformar la información en gráficos. Entre las muchas Técnicas de Visualización de Información que *Many Eyes* tiene disponible, se encuentra *Tag Cloud*. *Tag Cloud* fue creada para atender a aquellos usuarios que tienen datos no estructurados como entrada y deseaban contar la frecuencia de las palabras presentadas. Los creadores notaron que las nubes de palabras atraían a un nuevo tipo de usuario, interesados en datos textuales en vez de datos numéricos [Viegas07].

*Tagul*<sup>1</sup> es una herramienta para crear Nubes de Palabras *online*, sin la necesidad de la instalación de algún *software* en el computador [Santos14]. Las Nubes de Palabras creadas con *Tagul* tienen interactividad y una función interesante capaz de agregar un hipervínculo a cada palabra.

*Similar Diversity*<sup>2</sup> es una Técnica de Visualización creada para trabajar textos que presenta las similitudes y diferencias que existen entre cinco libros religiosos. Para generar el gráfico, la herramienta cuenta todas las palabras de los libros y crea grandes arcos visuales en torno a las cuarenta principales palabras más frecuentes. También identifica y presenta en grande los nombres de los personajes que tienen un papel importante dentro de esos libros. Abajo de las palabras principales, es creado un gráfico de barras con los verbos ya filtrados que están relacionados con los personajes principales identificados, presentando así las actividades que esos personajes hicieron o lo que hablaron. El tamaño de la fuente varía de acuerdo con la frecuencia de esas actividades en los libros. Esas visualizaciones son generadas de forma digital para ser utilizadas comúnmente en museos y en exposiciones abiertas al público.

*WordSwarm*<sup>3</sup> es un proyecto desarrollado en Python de código abierto, licenciado con GPLv3 y disponible para su utilización. El proyecto fue originalmente concebido para analizar las tendencias de variación en el área financiera, en intereses científicos y en nombres populares en un intervalo, a través de una Técnica de Visualización dinámica. Por ser un proyecto interesante, contactamos al autor, le presentamos nuestro trabajo y recibimos sugerencias para nuestro desarrollo.

*Pygame*<sup>4</sup> es un conjunto de módulos de Python que permiten la creación de aplicaciones multimedia o interfaces gráficas con animaciones. Es también utilizado

para crear juegos en dos dimensiones de una forma simple. La animación creada en este desarrollo fue basada en los módulos de *Pygame*.

### 3 Visualización de información

La Visualización de Información puede ser vista como una representación visual capaz de transformar datos simples en informaciones [Card99]. Su principal objetivo es transmitir las informaciones a los usuarios de forma clara y eficaz a través de gráficos. Para expresar el conocimiento de forma eficaz, tanto la dimensión funcional como la estética deben trabajar en conjunto [Friedman07].

La Visualización de Información es útil para realizar un análisis rápido de una situación específica cuando se tienen grandes cantidades de información. Un caso práctico que exige un análisis y toma de decisión rápida, es en un desastre natural [Kumar11].

Un diseño adecuado de herramientas de visualización simplifica la búsqueda de información, mejora la detección de patrones, aumenta los recursos accesibles, entre otros [Card99]. Entonces, el hecho de contribuir en el área de Visualización de Información significa crear una representación visual innovadora que sea comprensible y que genere conocimientos a los usuarios [Cairo11].

#### 3.1. Técnicas de visualización de información

Las Técnicas de Visualización de Información son utilizadas para aprovechar la capacidad humana de extraer patrones a partir de imágenes y ayudarles a comprenderlos rápidamente. En la búsqueda de conocimientos en grandes colecciones de información, las técnicas permiten capturar estructuras intelectuales percibidas en un dominio de conocimiento específico [Chen02]. Las Técnicas de Visualización de Información también pueden resaltar características importantes de un conjunto de informaciones, transformando datos abstractos en formas visuales. Por ejemplo, un conjunto grande de imágenes extraídas de la red social *Instagram* relacionadas a un tema específico, pueden ser organizadas en mosaicos a partir de colores, brillo y saturación, para hacer posible la percepción de patrones visuales y analizar de otra forma lo que estas imágenes podrían representar [Honorato14].

#### 3.2. Visualización de textos

La cantidad de información a la que tenemos acceso crece cada día y la mayoría de esa información está en formato texto [Nualart14]. La información de tipo texto puede presentar contenido subjetivo y tener una estructura abstracta, lo que hace difícil su análisis y la extracción de conocimiento de ellos.

Los investigadores Fernanda Viegas y Martin Wattenberg, enfocados en el área de Visualización de Información, relataron hace algunos años en una entrevista con Jeff Heer, que ellos creían realmente promisoría la Visualización de Textos. Durante mucho tiempo, el texto fue ignorado como herramienta de Visualización de Información, a pesar de que grandes cantidades de información importante se encuentran en

<sup>1</sup> <https://tagul.com/>

<sup>2</sup> <http://similardiversity.net/>

<sup>3</sup> <http://www.thisismikekane.com/>

<sup>4</sup> <http://www.pygame.org/hifi.html>

formato de texto [Heer10]. Recientemente, la explosión de las redes sociales y de los contenidos producidos y compartidos por los usuarios en ellas, generaron investigaciones relacionadas con la captura, procesamiento y visualización de informaciones textuales [Araujo13] [Feldman13].

El objetivo del análisis de textos o minería de textos [Feldman06] es detectar informaciones que están representadas dentro de textos y que no pueden ser efectivamente descritas a simple vista [Hearst03]. Por lo tanto, la Visualización de Textos lleva a la comprensión de textos de una forma diferente, interactiva, modificando el contenido original que es convertido en un conjunto de informaciones no estructuradas. La información de tipo texto completa no es usada directamente para la visualización, es quebrada en pequeños pedazos representativos a partir de un proceso de transformación. Este proceso puede ser hecho a partir de una reducción de texto, transformándolo en una lista de palabras ordenada de acuerdo con su frecuencia [Nualart14]. Esta lista de palabras puede ser representada de forma gráfica, transformando los textos en información utilizable, interactiva y más fácilmente perceptible para el análisis visual. Un ejemplo común de representación visual de información de tipo texto, es la Nube de Palabras.

### 3.3. Nubes de palabras

Las Nubes de Palabras, *Word Cloud* o *Tag Cloud*, son una combinación de varios tamaños de fuentes en una única visualización, teniendo sus primeros indicios hace unos 90 años, en la era del constructivismo soviético [Viegas08]. Además del estilo visual, las Nubes de Palabras tienen una finalidad específica, presentar una descripción visual de una colección de datos tipo texto [Viegas08] [Viegas07] [McNaught10].

Las Nubes de Palabras pueden ser utilizadas para fines analíticos, comunicar patrones de textos y presentar retratos de personas, por ejemplo. Esta Técnica de Visualización es flexible, pudiendo presentar grandes cantidades de textos, teniendo rápidamente una visión general de los términos más relevantes. La creciente demanda de las Nubes de Palabras indica la necesidad del usuario de visualizar una importante clase de información, información de tipo texto [Viegas08] [Cui10].

## 4 Experimentos y resultados

La representación visual de palabras sirve principalmente para identificar las palabras más utilizadas en un texto, conocer los contenidos enfatizados y para analizar la densidad de esas palabras clave identificadas. Esta Técnica de Visualización de Información ofrece un elemento gráfico tipográfico interesante que presenta una característica que la distingue del resto de técnicas, no es necesario que el texto sea estructurado, siendo posible comprender las informaciones rápidamente con una visión general.

Una Nube de Palabras Animada presenta las informaciones de acuerdo con un atributo temporal, de forma dinámica y diferente. En este sentido, se torna una técnica atrayente para representar, por ejemplo, informaciones sobre la evolución de los temas

investigados en el tiempo en un programa de estudio, o detectar el tema de investigación que se ha mantenido consistente durante un intervalo, indicando la consolidación de este tema para un área de investigación.

En este artículo, se presenta el desarrollo de Nubes de Palabras Animadas para presentar informaciones sobre la producción bibliográfica asociada a varios grupos de investigadores, pertenecientes a un programa de estudio de una universidad, teniendo como base la variación de frecuencia de las palabras clave, de las publicaciones producidas por estos docentes que conforman los grupos de investigación.

Para desarrollar las Nubes de Palabras Animadas, deben ser entendidos los retos que representa, por ejemplo: entender cuáles informaciones pueden ser acompañadas, entender el significado de los cambios observados en el intervalo presentado y cómo interpretar la dinámica de las informaciones visualizadas.

Las Nubes de Palabras Animadas presentan diferencias de tamaño en el texto para evidenciar la cantidad de trabajos producidos sobre un determinado tema en un determinado año. La animación de estos tamaños, evidencia el momento del surgimiento, los años que se sostuvo el tema en desarrollo y su relevancia, o la finalización de un tema en las investigaciones del grupo.

### 4.1. Método de captura de los datos

Las informaciones de las publicaciones de los profesores pertenecientes a los grupos de investigación considerados en este artículo fueron obtenidos directamente de la plataforma *CNPq-Lattes*<sup>5</sup>. La plataforma *CNPq-Lattes* mantiene una base de datos de los *currículos* de investigadores, así como también sobre las publicaciones producidas por estos. La idea principal para la captura de datos fue realizar un *Web Scraping* a la página de *Lattes* de cada profesor, sin embargo, esto no fue posible porque recientemente la plataforma comenzó a aplicar el control de acceso con el uso de un *Captcha*. La solución encontrada fue acceder, de forma individual, a cada currículo y descargar el archivo XML que genera la plataforma, correspondiente a cada profesor.

Para el almacenamiento de los datos, fue desarrollado un programa en PHP, con el fin de manipular los datos de los archivos descargados en formato XML. Este programa interpreta la estructura de los archivos XML, captura sólo la información requerida, hace la conexión y almacena los datos en la base de datos MySQL creada. En la base de datos, se almacenan todos los datos de las publicaciones bibliográficas de cada profesor: artículos en periódicos, libros y capítulos de libros, trabajos y presentaciones en certámenes. Específicamente, los datos utilizados para generar las Nubes de Palabras Animadas fueron: palabras clave de las publicaciones, año de publicación y grupo de investigación.

### 4.2. Método de creación de la visualización

Las palabras clave de las publicaciones bibliográficas son términos descriptivos de su contenido principal. Las

<sup>5</sup> <http://lattes.cnpq.br/>

palabras clave indican específicamente el área en la que se desarrolló la publicación, sirven de ayuda para clasificarlas correctamente, pudiendo así llegar a un mayor número de investigadores.

Partiendo de esta premisa y una vez que la información fue almacenada en la base de datos, se creó un *script* para clasificar las palabras clave por grupo de investigación y para contar el número de veces que esas palabras clave fueron utilizadas para describir las publicaciones bibliográficas por año, indicando así su relevancia.

A partir de estos datos y considerando las palabras clave que tienen incidencia mayor o igual a dos, fue posible generar un archivo CSV a partir de un programa en PHP, que contiene las palabras clave y el número de apariciones de cada una de ellas por año. Este archivo es utilizado como entrada para la segunda parte del proceso de la creación de las Nubes de Palabras Animada.

El componente responsable por la creación de las Nubes de Palabras Animadas, es basado en la plataforma *WordSwarm* (ver sección 2). *WordSwarm* fue creada inicialmente para realizar *Web Scraping* directamente de un *website* y extraer las informaciones relevantes, generando un archivo CSV con las informaciones filtradas. Como no fue necesario realizar ese proceso, la aplicación original fue modificada para permitir que la entrada fuese el archivo CSV generado anteriormente. Entre las ventajas de utilizar esta plataforma, está que permite reducir el tiempo de procesamiento de las informaciones y genera los vídeos de las animaciones, después que se realizara el trabajo de filtrar las informaciones irrelevantes para el análisis en el proceso de creación del archivo CSV.

Para la creación de las animaciones finales en Python, es utilizada la biblioteca *PyGame* (ver sección 2). Con esta biblioteca, son generados varios *frames* en formato PNG con dimensiones de 1920 x 1080 (*full HD*). Estos son, posteriormente, convertidos en un archivo de vídeo codificado en H.264 con el uso de *FFmpeg*<sup>6</sup>. El proceso completo para la creación de las Nubes de Palabras Animadas de nuestro trabajo, es mostrado en la Figura 1.

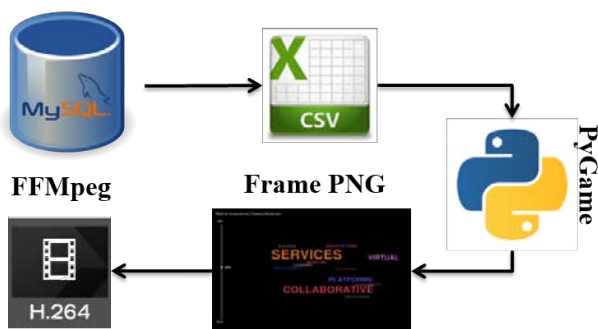


Figura 1. Proceso de creación de la Nube de Palabras Animadas.

El vídeo generado al final del proceso presenta las características de las investigaciones realizadas por un grupo a lo largo de los años. De forma simple e intuitiva, es posible observar, por ejemplo, los temas mayormente

investigados en un determinado año, así como también las evoluciones y consolidaciones de los temas investigados.

La técnica de Nube de Palabras utilizada comúnmente, sólo presenta una especie de retrato de los datos en un momento determinado, pero no la dinámica de esos datos. Es exactamente para cubrir esa laguna que nuestra técnica es utilizada.

### 4.3. Visualización de los resultados

Para la visualización de los resultados, se aplicó la Técnica de Visualización de Información desarrollada a dos grupos de investigación.

En el caso del primer grupo de investigación seleccionado, se observó que son varios los temas que han sido desarrollados a lo largo de los años. Se pueden visualizar diversas áreas de trabajo y no una única tendencia de un asunto principal que sea consistente en las investigaciones del grupo.

Uno de los temas de investigación relevante para el grupo es *Multimedia*, siendo desarrollado por primera vez en el año 1996 (Figura 2). Se evidencia que este tema es relevante en ese momento, siendo el único tema de investigación presente. Durante el año 2000, el tema continúa siendo investigado con un nivel de relevancia similar a otros que también fueron desarrollados en ese momento, como: *Hypermedia*, *Semantic Verification*, entre otros (Figura 3). Durante el año 2001, se puede observar que el área de *Multimedia* captó la atención de los investigadores, siendo *Multimedia Distributed System* el tema con mayor relevancia, sin embargo, también se puede observar que *Adaptive Multimedia* está entre los temas también desarrollados pero con menor relevancia (Figura 4). Durante el año 2003, continuaron las investigaciones del tema *Multimedia*, con la misma relevancia que tuvieron otros temas como: *Hypermedia* y *Collaborative Virtual Environments* (Figura 5). Su evolución continuó, durante el año 2008 vuelve a ser uno de los principales temas de investigación junto a *Information Retrieval* y *Context-Aware* (Figura 6). Durante el año 2015, ya *Multimedia* no está dentro de las investigaciones más relevantes del grupo, como puede observarse en la Figura 7.

Además de la evolución del tema *Multimedia*, fue posible observar otros temas importantes también desarrollados en esos años, tales como: *RT-LOTOS*, *Requirements Engineering*, *Collaborative Systems*, *Multimedia Journalism*, *Mobile Networks*, *Digital Video*, *Semantic Annotations*, *Digital TV*, *Web Services*, *Model Checking*, *Synchronization*, entre otros. Estos temas caracterizan las investigaciones del grupo en el intervalo de tiempo considerado para la presentación de los resultados.

<sup>6</sup> <https://ffmpeg.org/>



Figura 2. Tema *Multimedia* desarrollado en el año 1996.



Figura 3. Tema *Multimedia* desarrollado en el año 2000.

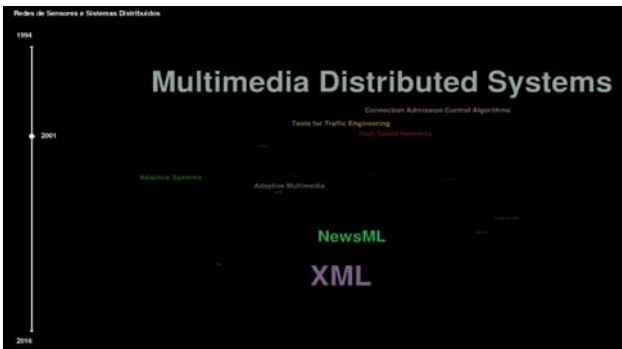


Figura 4. Tema *Multimedia* desarrollado en el año 2001.

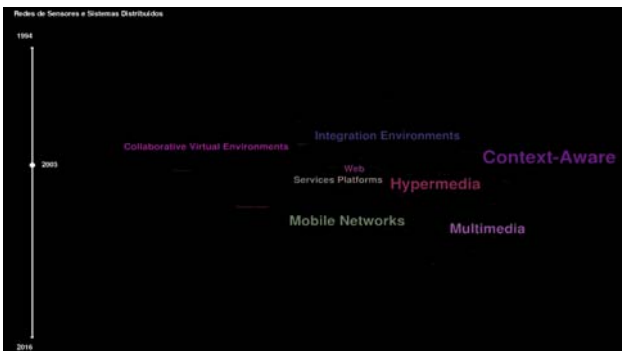


Figura 5. Tema *Multimedia* desarrollado en el año 2003.



Figura 6. Tema *Multimedia* desarrollado en el año 2008.



Figura 7. Tema *Multimedia* no presente en el año 2015.

En el caso del segundo grupo de investigación seleccionado, se pudo observar que los temas desarrollados están relacionados principalmente con: *Learning, Education, Intelligent Tutoring, Intelligent Agent, Multi-Agent System, Ontology*, entre otros.

Uno de los temas más resaltante es *Learning*, que se visualiza en casi todos los años de la Nube de Palabras Animadas, acompañado por otras investigaciones. Esto indica que este grupo tiene consolidado un tema de investigación, ya que este se mantiene relevante durante todos los años considerados en el estudio. La primera aparición fue en el año 1999 (Figura 8), siendo uno de los temas más destacados en ese año junto a *Multi-Agent System* y *Knowledge Systems*. Durante los años 2000 y 2001 (Figura 9), *Learning* se mantuvo presente en las investigaciones, destacándose los temas *Distance Learning, Cooperative Learning, Collaborative Learning* y *Online Learning*. Así como los temas *Knowledge Management, Software Development, AORML* y otros fueron importantes, *Learning* fue igual de relevante en el año 2004 (Figura 10). Durante el año 2008 (Figura 11) y durante el año 2013 (Figura 12), el tema *Learning* continuó presente en las investigaciones, además otros temas importantes empezaron a surgir, como: *Measurement, Software, LIBRAS, Ontology, Software Testing*, entre otros. Se puede inferir que el crecimiento de los temas de investigación corresponde con el momento de la llegada de nuevos profesores al grupo de investigación. Durante el año 2015 (Figura 13), el tema *Learning* continuó estando presente en los temas desarrollados pero con menor relevancia en comparación con los temas: *Inclusive Education* y *Digital Inclusion*, que tuvieron mayor relevancia ese año y que mantienen el interés del grupo actualmente.



Figura 8. Tema *Learning* desarrollado en el año 1999.



Figura 12. Tema *Learning* desarrollado en el año 2013.

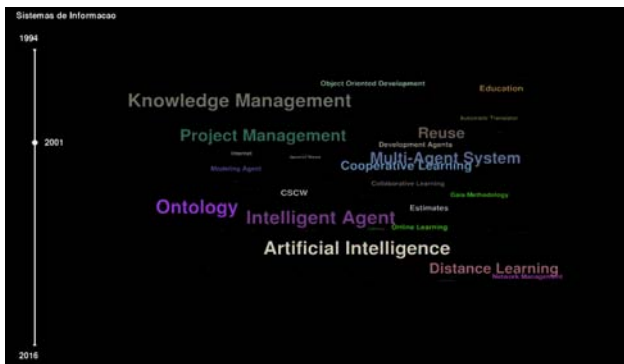


Figura 9. Tema *Learning* desarrollado en el año 2001.

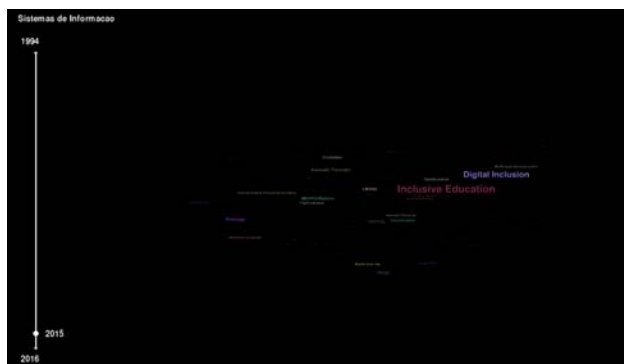


Figura 13. Tema *Learning* desarrollado en el año 2015.



Figura 10. Tema *Learning* desarrollado en el año 2004.

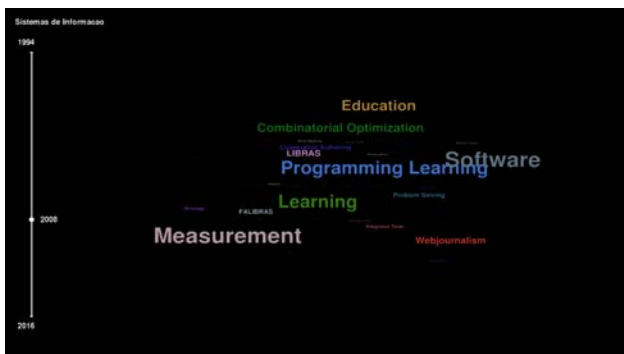


Figura 11. Tema *Learning* desarrollado en el año 2008.

#### 4.4. Observaciones de los resultados

El acompañamiento de la evolución de los términos asociados a la producción académica de un grupo de investigación, puede evidenciar la consistencia de una determinada línea a lo largo del tiempo. Además, cuando uno o más términos aparecen destacados, es decir, la palabra más grande, en un año específico, y en los años siguientes se tornan pequeñas o desaparecen, es un indicio que no fueron temas recurrentes o fue un tema que atrajo la atención momentánea de los investigadores durante un cierto periodo y después perdió interés por diferentes motivos que pueden ser: la conclusión de un proyecto, interés comercial, entre otros. Siguiendo esta línea, cuando uno o más términos se muestran semejantes a lo largo del tiempo, esto puede indicar que la línea de investigación del grupo es consistente y bien definida, como se evidenció en el grupo 2.

La cantidad de palabras presentadas con tamaños destacados a lo largo de la línea de tiempo, también puede revelar puntos importantes sobre el perfil del grupo y de sus investigadores. Si el número de temas relevantes es constante, puede indicar que el grupo posee un perfil conservador. Por lo contrario, si los temas relevantes son variables, puede indicar que el grupo incorpora nuevos tópicos a sus intereses de investigación a lo largo del tiempo.

Los vídeos utilizados para presentar los resultados, se encuentran disponibles online<sup>7</sup>.

7

[https://www.youtube.com/channel/UChmc0a\\_fYzakmTADpmswMeQ](https://www.youtube.com/channel/UChmc0a_fYzakmTADpmswMeQ)

## 5 Conclusiones y trabajos futuros

El tema abordado en el presente artículo fue la extracción de conocimiento a partir de la Visualización de Información por medio de Nubes de Palabras Animadas. Diferente a otros trabajos relacionados con la visualización de grandes volúmenes de informaciones textuales, nuestra propuesta tiene interés en la dinámica de esas informaciones a lo largo del tiempo. En este sentido, agregamos diferentes tecnologías para crear un proceso de desarrollo de Nubes de Palabras Animadas, que permite evaluar la evolución de un conjunto de términos durante un periodo. A partir de esa evolución, son muchos los indicios de esas informaciones que pueden ser levantados, facilitando la extracción de conocimiento.

La técnica propuesta fue aplicada a una base de datos que contiene los datos de las publicaciones bibliográficas de dos grupos de investigación de una universidad en un intervalo de 10 años (1995-2015). Cabe destacar, esta Técnica de Visualización de Información puede ser utilizada en cualquier conjunto de características, individuales o colectivas que se desee evaluar.

Como trabajo futuro, pretendemos aplicar sincronización a los vídeos generados para observar la relación entre los datos que pueden estar escondidos y que no son evidentes con la visualización individual de las Nubes de Palabras Animadas de cada grupo de investigación.

Para finalizar, otros puntos de interés para la evolución de nuestro trabajo son: el acompañamiento individual de los miembros con relación a su grupo de investigación, desarrollar otras técnicas de Visualización de Información para permitir observar diferentes dimensiones de las publicaciones de los miembros de los grupos de investigación, como cantidad de publicaciones por año, índice de impacto de las publicaciones, clasificación de las publicaciones dentro de la comunidad científica, entre otros.

### Referencias bibliográficas

- [Araujo13] Araújo, Matheus, et al. Métodos para análise de sentimentos no Twitter. Proceedings of the 19th Brazilian symposium on Multimedia and the Web (WebMedia'13), 2013.
- [Cairo11] Cairo, A. El arte funcional: infografía y visualización de información. Alamut, 2011. ISBN: 9788498890679.
- [Card99] Card, S. K.; Mackinlay, J. D.; Shneiderman, B. Readings in information visualization: using vision to think. Morgan Kaufmann, 1999. ISBN-10: 1558605339.
- [Chen02] Chen, Chaomei. Visualization of knowledge structures. Handbook of software engineering and knowledge engineering, 2002, vol. 2, pp. 700.
- [Cui10] Cui, W.; Wu, Y.; Liu, S.; Wei, F.; Zhou, M. X.; Qu, H. Context preserving dynamic Nuvem de Palavras visualization. IEEE Pacific Visualization Symposium, 2010, pp. 121-128. URL:

<http://research.microsoft.com/en-us/um/people/weiweicu/images/wordpv.pdf>

- [Feldman06] Feldman, R; Sanger, J. The text-mining handbook: Advanced approaches in analyzing unstructured data. Cambridge Univ. Press, 2006. ISBN: 13 978 0 521 83657 9.
- [Feldman13] Feldman, R. Techniques and applications for sentiment analysis. Commun.of the ACM 56, 2013, vol. 4, pp. 82-89.  
DOI= <http://dx.doi.org/10.1145/2436256.2436274>
- [Friedman07] Friedman, V. Data visualization: Modern approaches. Smashing Magazine, 2007, vol. 2. URL: <http://www.theartstory.org/critic-fried-michael.htm>
- [Hearst03] Hearst, M. A. What is text mining? SIMS, UC Berkeley, 2003. URL: <http://people.ischool.berkeley.edu/~hearst/text-mining.html>
- [Heer10] Heer, J. A conversation with Jeff Heer, Martin Wattenberg, and Fernanda Viégas. Queue, 2010, vol. 8, n. 3, pp. 10.  
DOI= <http://doi.acm.org/10.1145/1737923.1744741>
- [Honorato14] Honorato, J. I.; Carreira, L. S.; Goveia, F. G. Análise de Big Data pelos Parâmetros de Características Visuais. Anais do XIX Cong. de Ciências da Comunicação na Região Sudeste, 2014. URL: <http://www.portalintercom.org.br/anais/sudeste2014/resumos/R43-1279-1.pdf>
- [Kumar11] Kumar, S.; Barbier, G.; Abbasi, M.A.; Liu, H. TweetTracker: An Analysis Tool for Humanitarian and Disaster Relief. Proc. 5th Int. AAAI Conf. on Weblogs and Social Media. ICWSM, 2011, pp. 661-662.
- [Martig02] Martig, S. R.; Castro, S. M.; Di Luca, S. Interacción en la visualización de información. IV Workshop de Investigadores en Ciencias de la Computación, 2002. URL: [http://sedici.unlp.edu.ar/bitstream/handle/10915/21829/Documento\\_completo.pdf?sequence=1](http://sedici.unlp.edu.ar/bitstream/handle/10915/21829/Documento_completo.pdf?sequence=1)
- [McNaught10] Mcnaught, C.; Lam, P. Using Wordle as a supplementary research tool. The qualitative report, 2010, vol. 15, n. 3, pp. 630. URL: <http://www.nova.edu/ssss/QR/QR15-3/mcnaught.pdf>
- [Nualart14] Nualart V., J.; Pérez M., M.; Whitelaw, M. Cómo dibujamos textos: Revisión de propuestas de visualización y exploración textual. El profesional de la información, 2014, vol. 23, n. 3, pp. 221-235.
- [Santos14] Santos, R.; De Próspero, R. P.; Gianordoli, G. Estudo sobre a abordagem quantitativa para visualização de dados qualitativos: processo e ferramentas. Blucher Design Proc, 2014. URL: <http://pdf.blucher.com.br.s3-sa-east-1.amazonaws.com/designproceedings/cidi/CIDI-141.pdf>

- [Strecker12] Strecker, J. Data visualization in review: summary. Tech. Report IDRC, 2012. URL: <http://idlbnc.idrc.ca/dspace/bitstream/10625/49286/1/IDL-49286.pdf>
- [Viegas07] Viégas, F.; Wattenberg, M.; Van Ham, F.; Kriss, J.; McKeon, M. Manyeyes: a site for visualization at internet scale. IEEE Transactions on Visualization and Computer Graphics, 2007, vol. 13, n 6, pp. 1121-1128.
- [Viegas08] Viégas, F.; Wattenberg, M. Timelines tag clouds and the case for vernacular visualization. Interactions, 2008, vol. 15, n. 4, pp. 49-52.