

## IV Congreso de la Sociedad Internacional de Humanidades Digitales Hispánicas

Toledo, 23-25 de octubre de 2019

### Pathfinder: sistema para el análisis y visualización de datos sobre catálogos y exposiciones artísticas

Nuria Rodríguez-Ortega (UMA)  
Antonio Cruces Rodríguez (UMA)

El objetivo de esta comunicación breve es presentar el sistema Pathfinder (<http://pathfinder-art.es>) adoptado el formato de presentación «demo». Pathfinder (PF) es una aplicación utilitaria concebida como herramienta para el análisis gráfico y textual de la información registrada en ExpoFinder ([www.expofinder.es](http://www.expofinder.es)), un repositorio de datos estructurados y semánticamente relacionados sobre exposiciones artísticas (Rodríguez-Ortega & Cruces Rodríguez, 2018). Se trata de un desarrollo *software* realizado por el grupo de investigación iArtHis\_Lab, del Departamento de Historia del Arte de la Universidad de Málaga (España), en el marco de los proyectos de investigación *Catálogos artísticos: gnoseología, epistemologías y redes de conocimiento. Análisis crítico y computacional* (HAR2014-51915-P), financiado por el Ministerio de Economía y Competitividad, y *Metodologías de datos aplicadas al análisis de las exposiciones artísticas para el desarrollo de la economía creativa* (PRY128-17), financiado por and el Centro de Estudios Andaluces de la Junta de Andalucía.

El objetivo principal de PF es hacer visible y materializar las posibilidades que aporta el tratamiento estadístico de los datos para el estudio de fenómenos complejos, como es el caso de las exposiciones artísticas y de los ecosistemas socioculturales que se generan en torno a ellas.

Técnicamente, PF basa su desarrollo en un sencillo grupo de requisitos:

1. Tecnología de código abierto bajo licencia Creative Commons.
2. Código fuente de otros desarrollos y datos de terceros de acceso libre.
3. No emplea código propietario.
4. Está concebida como una aplicación en línea, siguiendo el modelo cliente-servidor, pero situando el núcleo de trabajo computacional en el lado del cliente y reduciendo al mínimo la participación del servidor.
5. Está desarrollado con ECMAScript/JavaScript y funciona sobre un navegador que deberá estar actualizado para que sea eficiente, rápida y poco ávida de recursos.
6. Utiliza un número reducido de bibliotecas de código externas, por lo que la mayoría de las rutinas de programación son de desarrollo propio.
7. La base de datos es un archivo con formato JSON que contiene los datos y se descarga del servidor ExpoFinder (EX), previa identificación. Se conserva en disco local con un cifrado mínimo.

De acuerdo con el modelo de datos de EX, hay seis tipos principales de registros primarios: actores (artistas, comisarios, etc.), obras de arte, catálogos, empresas, instituciones y exposiciones. Dado que el sistema de datos de EX sigue el modelo de una ontología, estas entidades se encuentran conectadas entre sí mediante múltiples relaciones y están descritas mediante un sistema enriquecido de metadatos en el que se contemplan numerosos atributos. La granularidad y complejidad del sistema de datos incrementa las posibilidades de extracción de conocimiento a través de las funcionalidades implementadas en PF.

Desde el punto de vista de la funcionalidad, PF basa su diseño en cuatro principios esenciales.

1. Flexibilidad y alta precisión en las funciones de búsqueda. Junto a las funcionalidades de búsqueda tradicionales, PF incluye un poderoso sistema de filtrado basado en la combinación de múltiples *queries*, lo que hace posible obtener subconjuntos específicos de datos utilizando tantas condiciones como se requieran. Así, por ejemplo, un investigador podría recuperar el listado de exposiciones celebradas en Madrid, desde 2010 hasta 2015, comisariadas por comisarios no españoles, financiadas por entidades privadas, y en las que hubiesen participado con obras pictóricas artistas mujeres de entre 25 y 35 años.

2. Optimización de los sistemas de análisis para la extracción de conocimiento basado en procesos cuantitativos mediante la exploración de conceptos estadísticos potencialmente significativos para la investigación humanística. En este sentido, se han implementado funcionalidades de estadística descriptiva avanzada, haciendo posible la incorporación de categorías hasta ahora no exploradas en el análisis de ecosistemas de exposiciones artísticas, como las nociones de «entropía» –medida mediante el índice o coeficiente de Jaccard- y «valores atípicos (*outliers*)» extraídos mediante el test de Tukey. Los valores atípicos, por ejemplo, suelen representar un inconveniente cuando se analizan series largas de datos pero, al mismo tiempo, son una oportunidad para la investigación humanística porque nos permiten detectar elementos que caen fuera de los límites de la «normalidad», es decir, nos lleva a focalizar la atención en aquellos elementos menos homogéneos respecto del conjunto total. Por ejemplo, frente a la homogeneidad que presentan las exposiciones artísticas en relación a su duración temporal, es posible detectar aquellas que nos responden al parámetro general de duración establecido en 5 meses. Labor del historiador del arte o del analista cultural será interpretar la mayor o menor relevancia de este hallazgo.

Asimismo, también se han implementado otras funcionalidades de análisis significativas para el campo que nos ocupa, como las basadas en coocurrencias o matriz de coincidencias entre valores extremos en las relaciones mutuas, lo que nos permite indagar, por ejemplo, sobre la frecuencia con la que los actores X e Y coinciden en exposiciones según su tipología y, por tanto, la tendencia de determinados artistas a coaparecer en unas mismas exposiciones.

3. Énfasis en las visualizaciones como herramientas hermenéuticas e interpretativas. De acuerdo con esta idea, PF se ha implementado con una multiplicidad de herramientas de visualización configurables por el usuario con el objetivo de hacer la información lo más heurística posible. Así, junto con los tradicionales mapas, histogramas y redes, algunos tipos de visualizaciones han sido diseñadas *ad hoc* para PF, como las que hemos denominado «geogramas» y «taxogramas», que trazan un gráfico de dispersión o de radar según el tipo de metadato seleccionado. Por su parte, la sección geoanalítica está compuesta de mapas que, además de la información georreferenciada usual y de los flujos entre actores, incluye el cálculo basado en la triangulación de Delaunay y los polígonos de Thiessen o de Voronoi, lo que nos ayuda a detectar posibles áreas de influencia a partir de la proximidad entre diferentes grupos de elementos.

4. Amplio rango de usabilidad e interoperabilidad. Finalmente, PF puede operar con cualquier conjunto de datos, sea extraído de EX o no, siempre y cuando este sea conforme con el modelo de estructura JSON disponible en github (<https://github.com/antoniocruces/pathfinder>). Esto significa que PF puede ser utilizado por cualquier grupo o investigador interesado en utilizar las oportunidades de exploración, análisis y visualización que PF ofrece. Al mismo tiempo, cualquier subconjunto de datos obtenido mediante el sistema de filtrado de PF puede ser exportado en formatos estándares para ser procesado con en otras plataformas o con otros *software*.

Durante la presentación, se realizará una demostración de todas estas funcionalidades mediante ejemplos concretos.

### **Referencia**

Rodríguez-Ortega, N. and Cruces Rodríguez, A. (2019), «Development of Technological Ecosystems for Cultural Analysis: The Case of Expofinder System and Art Exhibitions», *Digital Scholarship in the Humanities*, 34, n. 2, 423-448. fqy018, Oxford University Press.