

Sinclair, Stéfán, y Rockwell, Geoffrey. Voyant Tools.

<https://voyant-tools.org>

Reseña realizada por:

Jaime Ricardo HUESCA OROZCO

jai_rik@hotmail.com

Desde las Humanidades, el trabajo de investigación mediado por computadora representa una necesidad cada vez más latente, con mayor requerimiento en un contexto pandémico donde no es posible acceder fácilmente a bibliotecas, acervos bibliográficos, prácticas de campo, etcétera. Vivimos un periodo colmado de adversidades y pérdidas, cierto, pero la generación del conocimiento ha encontrado alternativas desde la virtualidad, ya sea por medio de tutoriales, cursos online o videoconferencias, las cuales han surgido como un campo de producción fértil.

Una alternativa asequible para el trabajo académico es incursionar dentro de las Humanidades Digitales, ya que éstas ofrecen una infinidad de posibilidades para el análisis concreto de un corpus o diversos objetos a emplear para el desarrollo de una investigación. Los programas o plataformas para poner en práctica el análisis digital son diversos, aunque su grado de complejidad, de primera mano, podría intimidar a quien se adentre por primera vez en esta rama de estudio. Para una mejor experiencia en el acercamiento, es preciso acudir a entornos acordes al nivel de experiencia de su usuario; en dicho caso, Voyant Tools se posiciona como una alternativa de gran ayuda para iniciar tal proceso.

Voyant Tools (Sinclair y Rockwell, 2016) es una plataforma web de acceso abierto que procesa y analiza texto desde una lectura distante. Diseñada en 2016 por los canadienses Stéfán Sinclair y Geoffrey Rockwell, dicho espacio de análisis solamente requiere de una conexión a internet medianamente buena y texto a procesar (ya sea desde un corpus precargado de la misma plataforma, una dirección URL, texto sin formato o un documento). Cabe



mencionar que, de utilizar un corpus digital propio, es necesario curar el documento previamente, ya que dentro de VT no es posible modificar el, o los, documentos. Además, la plataforma identifica y procesa una gran cantidad de idiomas.

Luego de cargar el texto en Voyant, ingresaremos a un sitio con diversas herramientas para trabajar, el desempeño de éstas dependerá de los objetivos a conseguir en la investigación. El espacio es sencillo e intuitivo; en esta suerte de arenero se presentan por defecto cinco herramientas, cada una con sus opciones de modificación establecidas y menús propios, donde pueden sustituirse por otras funciones, exportarlas o configurarlas.

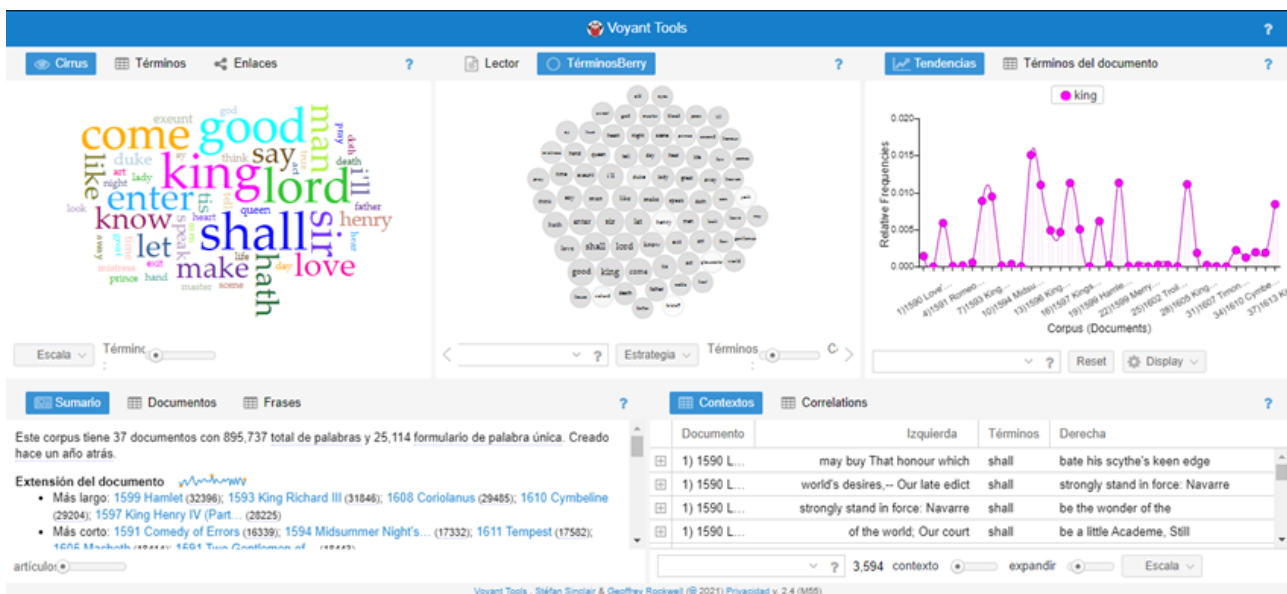


Figura 1. Elaboración propia.

Las herramientas presentes en VT son variadas y se complementan una con otra. A mi modo de ver, "Cirrus" es el punto de partida por su alcance interpretativo. Básicamente, es una nube de palabras que determina la frecuencia bruta de recurrencias en el corpus, el tamaño de cada término es proporcional al número de sus repeticiones. Por medio de su visualización, es posible identificar los términos más utilizados a lo largo del documento, cuestión que podría demostrar, por ejemplo, tendencias discursivas, el lenguaje común empleado en determinado contexto sociohistórico, campos semánticos, etcétera. A su vez, otra herramienta similar es "TérminosBerry", la cual expone también los términos más representativos del corpus pero, a diferencia de "Cirrus", su interacción muestra correlaciones por cercanía con otras palabras.

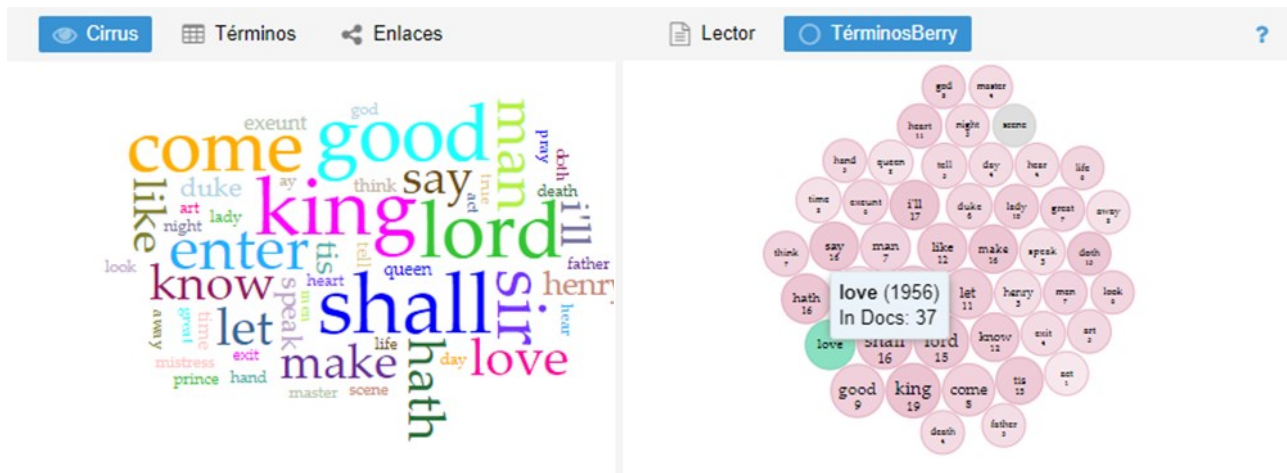


Figura 2 y 3. Elaboración propia.

Una de las grandes ventajas de VT es la posibilidad que ofrece para incorporar reflexiones de corte cuantitativo en la Humanidades, donde sus ramas de conocimiento, generalmente, se enfocan en lo cualitativo por cuestiones de practicidad. Para los estudios de lingüística o literatura, disciplinas donde me especializo, el soporte de VT facilita el estudio de grandes cantidades de texto, cuestión que sería imposible llevar a cabo de forma manual o, al menos, no de manera inmediata; por otro lado, gracias a su lectura distante es posible visualizar relaciones intertextuales ocultas para la visión del investigador.

Desde mi punto de vista, dos aspectos fundamentales para desarrollar el estudio en VT son las recurrencias y correlaciones entre términos, pues gran parte de las herramientas del sitio dependerán del conteo y visualización de esta data. Así mismo, otro factor de relevancia es la distinción espacial de unidades léxicas dentro del documento, ya sea visualmente o a partir de su contexto, pues con ello puede interpretarse la influencia de un término en algún pasaje narrativo establecido, por ejemplo, el contraste de aparición entre varias palabras, sus variaciones semánticas, etcétera. Para lo anterior hay diversas opciones de visualización que facilitan la actividad mencionada, algunas de ellas son “Líneas de burbuja”, “Contextos” “Tendencias”:

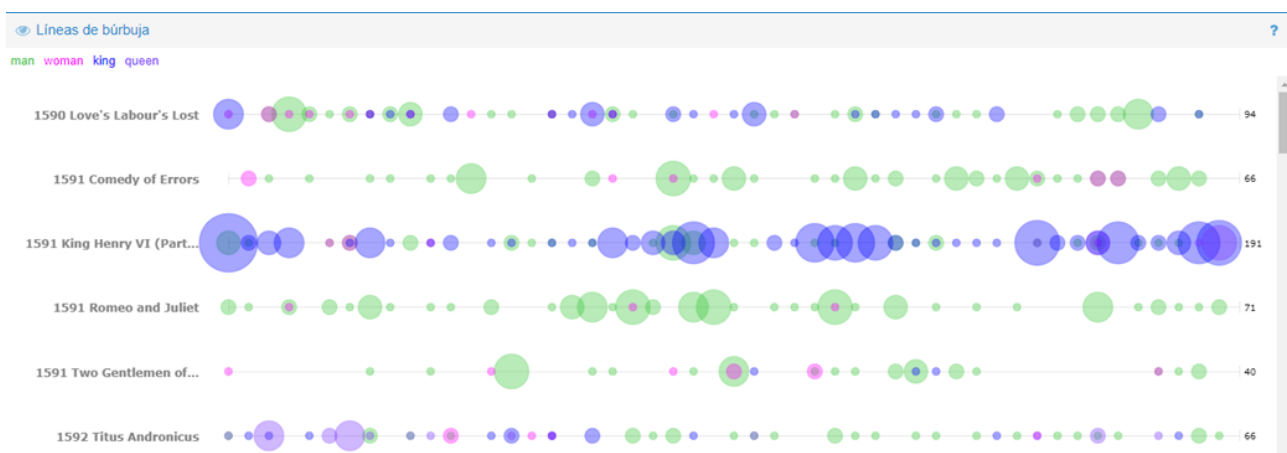


Figura 4. Elaboración propia.

Contextos			
Documento	Izquierda	Términos	Derecha
1) 1590 L...	study here three years. But there are other strict observances; As, not to see a	woman	in that term, Which I hope well is not enrolled there; And one day in
1) 1590 L...	name. FERDINAND How well this yielding rescues thee from shame! BIRON [Reads] 'Item, Th...	woman	shall come within a mile of my court.' Hath this been proclaimed? LONGAVILLE Four days
1) 1590 L...	dangerous law against gentility! [Reads] 'Item, If any man be seen to talk with a	woman	within the term of three years, he shall endure such public shame as the rest
1) 1590 L...	sir, for the manner.--it is the manner of a man to speak to a	woman	: for the form.-- in some form. BIRON For the following, sir? COSTARD As it shall
1) 1590 L...	a child of our grandmother Eve, a female, or, for thy more sweet understanding, a	woman	. Him I, as my ever-esteemed duty pricks me on, have sent to thee, to
1) 1590 L...	gates. I am in love too. Who was Samson's love, my dear Moth? MOTH A	woman	. master. DON ADRIANO DE ARMADO Of what complexion? MOTH Of all the four, or the
1) 1590 L...	this damsel, I must keep her at the park: she is allowed for the day-	woman	. Fare you well. DON ADRIANO DE ARMADO I do betray myself with blushing. Maid! JAQUEN...
1) 1590 L...	noise abroad. Navarre hath made a vow. Till painful study shall outwear three years, No	woman	may approach his silent court: Therefore to's seemeth it a needful course, Before we enter
1) 1590 L...	Exit] LONGAVILLE I beseech you a word: what is she in the white? BOYET A	woman	sometimes, an you saw her in the light. LONGAVILLE Perchance light in the light. I
1) 1590 L...	like a tumbler's hoop! What, if I love! I sue! I seek a wife! A	woman	. that is like a German clock. Still a-repairing, ever out of frame, And never

Figura 5. Elaboración propia.

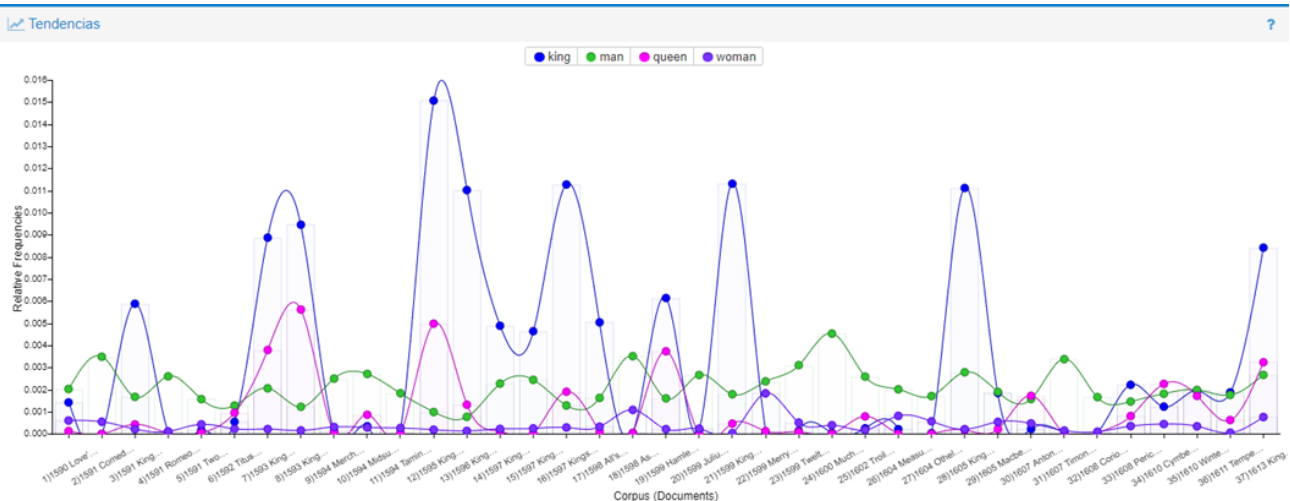


Figura 6. Elaboración propia.

Por otro lado, debo advertir que los resultados obtenidos también dependerán de un sustento teórico que el investigador proponga: la página web cuantifica la información, pero el resultado se interpretará a partir de la visión y marcos conceptuales del usuario. De lo contrario, la palabra se vuelve número y pierde su sentido verdadero. De ser así, para un resultado óptimo es prioritaria una metodología mixta donde el dato cuantificado se entrecruce con las reflexiones cualitativas de quien ejerce la lectura densa; la plataforma no es capaz de realizar toda la carga de trabajo, VT es solamente una extensión metodológica más, igual que toda metodología, presenta alcances y límites.

Si bien VT posee herramientas sofisticadas para visualizar data, ésta cuenta con ciertas restricciones, por ejemplo, la imposibilidad de modificar el *corpus* dentro de la plataforma, como ya mencioné; además, respecto al documento, es necesaria una herramienta que permita la codificación o etiquetación dentro de éste para elaborar estadística más allá del conteo de términos sino de ideas u otros elementos pragmáticos que destaque quien investiga. Tales herramientas se encuentran presentes, por ejemplo, en *Recogito* o *MAXQDA*. Por otro lado, aunque existe una función que determina tendencias por tópicos en el *corpus*, considero que ésta se encuentra en un desarrollo primario y los resultados se basan en selecciones léxicas aleatorias, cuestión que no brinda seguridad para su consideración. De ser una función viable, el desarrollo del análisis daría un gran salto interpretativo y, por otro lado, permitiría el debate entre ambos tipos de lectura.

Aparte, debo mencionar que VT se sustenta en el análisis de contenido y, como mencioné, esta situación produce una limitante cuando se trata de avanzar a un nivel de análisis enunciativo. Sin embargo, otro gran punto a favor de la plataforma es que posiblemente las carencias anterio-

res se encuentren dentro de la agenda de mejoras, pues ésta es perfectible y permanece en constante actualización, hecho visible en varias de sus herramientas, las cuales aún se hallan en un proceso experimental. Además, sería injusto identificar a VT por sus contadas desventajas, pues sigue en crecimiento, su abanico de opciones es suficiente para la elaborar una investigación, no requiere de ninguna instalación de software y, sobre todo, es de libre acceso.

Reconocer que Voyant necesita del apoyo complementario de otros entornos digitales tampoco es cuestión que demerite su eficacia, como señalé al inicio de este comentario: es el umbral para practicar el estudio desde las Humanidades Digitales y después adentrarse a opciones más complejas para el iniciado en la materia. Contrariamente, otra opción válida es aprovechar toda la riqueza que ofrece Voyant Tools, donde también está implicada la creatividad y originalidad de quien opera a través de la pantalla.

REFERENCIAS BIBLIOGRÁFICAS

Sinclair, S., y Rockwell, G. (2016). *Voyant Tools*. <https://voyant-tools.org>