

[Centro de Información de COVID \(CIC\): Charlas científicas de relámpago](#)



Transcripción de una presentación de Alan Porter (Search Technology, Inc. y Georgia Institute of Technology), enero de 2022

Título: [Explorando causas y curas para COVID-19 a través de un mejor acceso a la investigación biomédica](#)

[Perfil de Alan Porter en la base de datos del CIC](#)

Subvención de La Fundación Nacional de Ciencias (NSF, por sus siglas en inglés) #: [2029673](#)

[Grabación de YouTube con diapositivas](#)

[Información del seminario web del CIC de enero 2022](#)

Editora de la Transcripción: Saanya Subasinghe

Editora de la Traducción: Isabella Graham Martínez

Transcripción

Alan Porter:

Diapositiva 1

Gracias a todos, quería compartir algunos resultados de un proyecto de NSF RAPID que hicimos. Con el título aquí - y es básicamente la minería de la literatura biomédica en formas de tratar de obtener acceso convenientemente y algunas perspectivas diferentes.

Diapositiva 2

Aquí están mis antecedentes. No voy a pasar por eso. Básicamente somos observadores de tecnología - tratando de reunir datos y extraer algunos conocimientos utilizables de ella.

Diapositiva 3

Un proyecto en particular aquí estaba explorando causas y curas para COVID a través de un mejor acceso a la investigación y solo quiero tocar tres elementos. Una es [lo que] estamos llamando bibliometría inteligente para tratar de extraer información de estos textos abstractos. El segundo es un vistazo a un sistema de recomendaciones para hacer un descubrimiento basado en la literatura más allá del conjunto de datos que estamos analizando. Y la tercera es un panel para compartir resultados.

Diapositiva 4

Nuestra abreviatura para esto es 'tech mining'. Los datos a mano, aquí, están usando la principal estrategia de búsqueda COVID de la Biblioteca Nacional de Medicina. Usamos esto durante un año y medio. Y voy a estar hablando de dos subconjuntos diferentes. Uno es de unos 60.000 registros hasta principios de octubre de 2020 en el que hicimos una gran cantidad de la bibliometría inteligente tipo de trabajo y el otro es la versión más reciente del tablero de instrumentos a partir del 1 de enero - para mostrar lo que parece ahora.

Diapositiva 5

Bien, elemento número uno: la minería de tecnología es un análisis de texto de los recursos de datos de la ciencia y la innovación, generalmente en forma de registros abstractos de publicaciones de investigación o PAT, la mayoría de las veces. Algunas de las cosas que hemos hecho aquí - una es mirar dentro del cuerpo de esa investigación para qué temas muestran la atención acelerada y tratar de hacer eso de una manera oportuna. No hablaré de eso hoy. El segundo es el seguimiento de los temas a lo largo del tiempo.

Diapositiva 6

Aquí está el vistazo a que dirigido por nuestro colega Yi Zhong en la Universidad de Tecnología, Sydney. Uno dibuja mediante la agrupación del uso del término o lo que son algunos grupos de interés de actualidad en el tiempo?

Diapositiva 7

El proceso es, en esencia, tomar un período de tiempo agrupar los temas en base a la co-ocurrencia de términos, y luego hacer esto en períodos sucesivos, y luego mirar los vínculos entre esos temas. Y tengo un par de referencias aquí no voy a trabajar pero están etiquetados en la parte posterior de la presentación.

Diapositiva 8

Aquí hay un zoom en un pedacito de ella. Nuestros últimos análisis en 2020 mostraron que la rápida difusión era un tema de interés. Y que tenía antepasados en el tratamiento muy reciente de las enfermedades infecciosas, un poco antes de los estudios que notan el tiempo. Y luego volviendo a algunos de los coronavirus anteriores - algunos trabajan con PCR en 2006 concentración. Remontándose al SARS en 2003. Uno puede hacer algunos enlaces sobre temas y con el tiempo.

Diapositiva 9

Nuestro segundo elemento - el descubrimiento basado en la literatura - y aquí lo hemos utilizado solo para un sistema de recomendación muy simple. Vea si hay un tema que le interese dentro de la literatura COVID. ¿Podría haber alguna investigación útil para usted fuera de esa literatura?

Diapositiva 10

Y el proceso aquí - generamos un conjunto de datos como una simple ilustración de unos 30 o más temas co-morbilidades con uno que acabamos de elegir aquí. Y luego dentro de la literatura de COVID, en este caso son 60.000 registros, buscando los documentos más relacionados con eso. A continuación, hacer algunos análisis de texto allí para limpiar la literatura que detecta que, y buscando qué términos tienen especialmente alta frecuencia en esos documentos, y por el contrario especialmente baja frecuencia, llamando a que un modelo de conocimiento. Luego salir y mirar los cálculos de la frecuencia de la firma [y] la frecuencia inversa del documento para esos términos en el conjunto de datos completo de PubMed Medline. Esos no son idénticos [inaudible]. E identificar resúmenes de artículos que podrían ser de interés para alguien que persigue un tema en particular, como la comorbilidad.

Diapositiva 11

Aquí hay una pequeña ilustración de tres de los temas de comorbilidades donde hemos añadido en un recuento de citas como un segundo criterio. Si usted está interesado en ser citado y está utilizando las palabras similares al uso de comorbilidad dentro de la literatura COVID, [esto] podría ser de su interés. Tal vez echar un vistazo a estos y podría ser tres artículos, o podríamos ir todo el camino hasta los cargadores en los 33 millones, que no creo que nadie querría.

Diapositiva 12 - <https://sites.google.com/searchtech.com/covidproject/home>

El último elemento por el que quería pasar era echar un vistazo a nuestro panel de control. Y aquí está el tablero del proyecto y dentro de eso solo estoy mencionando una diana de vacunas que nuestros colegas de BizInt armaron, que es interesante rastrear a lo largo del tiempo. Voy a echar un vistazo a nuestro tablero PubMed. Y eso ahora está poniendo de relieve los 153 - 150.000 registros abstractos a partir del 1 de enero. Y hay una demostración allí, un par de minutos, para explicar cómo usarlo. Aquí hay algunos de los diferentes campos de datos disponibles. Y déjame hacer un rápido pop-in aquí hay una versión donde podemos ir y podríamos detectar algo de interés y yo solo voy a - tratar de destacar. Y creo que estamos recibiendo una respuesta lenta, aquí. La investigación en Irán - podríamos ir a través de eso y ver lo que algunas personas de allí están haciendo con respecto a las comorbilidades. E incluso podríamos ver un artículo en algún lugar aquí que atrae nuestro interés. Y aquí podemos abrir ese disco abstracto. Para hacer las cosas rápidamente, solo hay otra ilustración. Si quisiéramos destacar el trabajo más reciente y que es de diciembre del pasado 2021. Esto está incompleto porque la codificación, la categorización todavía está sucediendo, la catalogación, etc., pero podríamos entrar aquí de nuevo y elegir lo que podría ser de interés. [Vídeo de congelación] En Columbia - y hay algunos de los 46 artículos de - asociado con un autor de Columbia en el último mes. Esa es nuestra demo- si alguien quiere ir al panel de control nos encantaría. Todavía es un trabajo en progreso.

Diapositiva 13

Y básicamente he terminado. Ahí vamos - Recursos y Final. ¡Gracias por su atención!