

## Sin importar el tema, los rankings tienen la misma dinámica

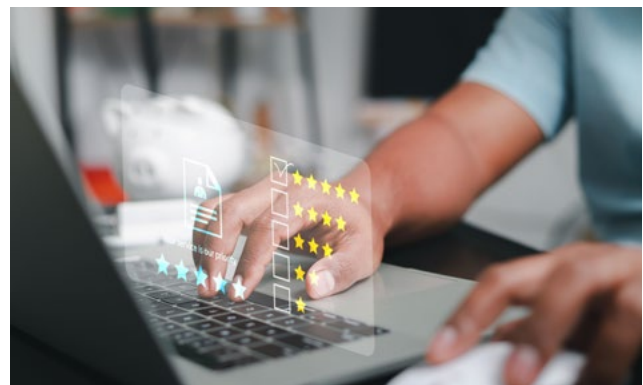
Un grupo de investigadores encontró patrones comunes en la dinámica de 30 rankings de distintos temas, lo que puede ayudar a hacer interpretaciones más cautelosas y tener mejores herramientas de clasificación.

Aleida Rueda

28 de marzo de 2022

Desde universidades clasificadas por su calidad académica, deportistas rankeados por su desempeño en alguna competencia, países enlistados por su población, estaciones de metro o regiones sísmicas, si se trata de rankings todos tienen la misma dinámica.

Es lo que ha demostrado un equipo de investigación dedicado a resolver una incógnita durante más de seis años: ¿Hay patrones comunes entre los rankings? Si es así, ¿cómo operan?



El grupo está conformado por los investigadores asociados al Centro de Ciencias de la Complejidad, C3, de la UNAM: Carlos Gershenson, del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS) y Gerardo Íñiguez, de la Universidad Europa Central (CEU), en Austria, además de Carlos Pineda, del Instituto de Física, y Albert-László Barabasi, adscrito a la CEU, la Universidad de Northeastern y la Escuela de Medicina de Harvard.

Su artículo, publicado hoy en la revista *Nature Communications*, incluye el análisis de 30 rankings en temas disímiles que incluyen sistemas naturales, sociales, económicos y de infraestructuras.

En él, muestran evidencia de que hay patrones temporales generales en la dinámica de todos los rankings, lo cual puede ayudar a hacer interpretaciones más matizadas de las clasificaciones y tener herramientas de clasificación algorítmica.

### El inicio

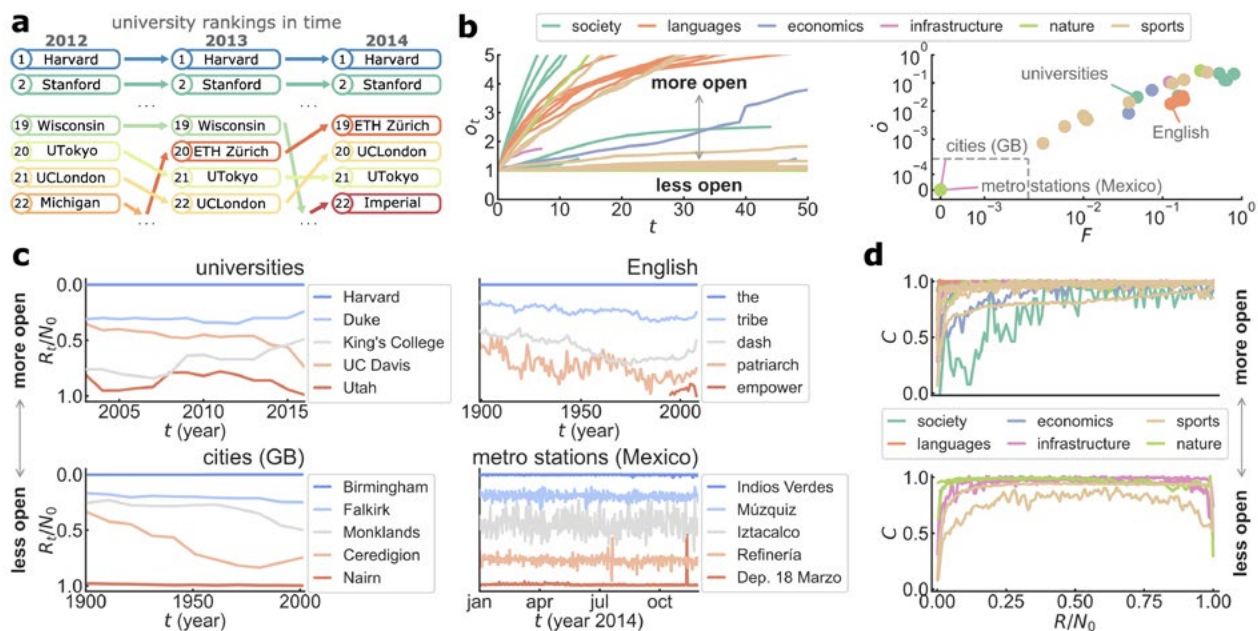
“Esta investigación se origina en una de las muchas celebraciones que le hicieron a Germinal Cocho por sus 80 años”, cuenta Gershenson desde su cubículo en el C3, en Ciudad Universitaria. Era 2013, cuando inició la colaboración entre Gershenson y Pineda con Cocho, uno de los pioneros en México en el estudio de los sistemas complejos y Jorge Flores, investigador emérito del Instituto de Física.

En ese tiempo, Cocho y Flores, quienes fallecieron en 2019 y 2020, respectivamente, tenían la idea de estudiar cómo se distribuían las palabras en diferentes idiomas a través del tiempo. Así que el equipo, que también incluyó a Sergio Sánchez, en ese tiempo de la Facultad de Ciencias de la UNAM, utilizó la base de datos de Google Books, con registros de miles de libros escritos desde 1500 hasta 2018, para identificar las palabras más usadas.

Pero la colaboración no paró allí. Vinieron otros estudios con clasificaciones de otros temas, desde lenguajes hasta deportes, y en todos había una coincidencia: el rango superior, compuesto por los elementos que ocupaban los primeros sitios de las listas, solía variar menos que los elementos ubicados más adelante en la clasificación. Así que les surgió el interés de saber si este comportamiento ocurría en todos los rankings.

Uno de los primeros pasos fue buscar diversos conjuntos de datos ordenados que les permitieran identificar su evolución a lo largo del tiempo. “Algunos rankings eran de siglos, como la frecuencia anual de cientos de miles de palabras en seis idiomas desde el año 1600, otros eran granulares, como el número de viajeros en 636 entradas del metro de Londres cada 15 minutos durante una semana”, explica Íñiguez.

Al final juntaron 30 clasificaciones de temas tan diversos como ciudades, animales, transportes o deportes. Y se dedicaron a estudiar su dinámica.



Ranking lists in nature and society show generic patterns in their dynamics.

Algunas de las gráficas de rankings de diversos temas.

Imagen de [Nature Communications](#)

## El modelo nulo

En una estancia sabática en Boston entre 2015 y 2016, Carlos Gershenson empezó a desarrollar lo que llamaría un modelo nulo, un modelo teórico sencillo basado en mecanismos aleatorios y diseñado para reproducir la estadística en todos esos conjuntos de datos. Durante varios años, juntos, Gershenson, Pineda e Íñiguez fueron puliendo ese modelo hasta llegar a la versión final.

El modelo nulo permite tener dos probabilidades: de desplazamiento y de remplazo. La **dinámica de desplazamiento** es cuando toman un elemento al azar de un ranking para colocarlo en otra posición

al azar (que representa, por ejemplo, cuando un jugador de tenis de cierta posición sube o baja a otra); mientras que la dinámica de remplazo sucede cuando un elemento es sustituido por otro (como cuando, por ejemplo, en los rankings de personas, unas mueren y son sustituidas por otras).

“Hicimos ajustes de los datos con este modelo para ver qué probabilidades tenía cada conjunto de datos y resulta que, aunque algunos rankings tienen más remplazo y otros, más desplazamiento, todos caen en un mismo patrón”, explica Gershenson. Un patrón en el que los rangos más altos suelen variar menos que los que están en medio.

Pero, además, el modelo deriva en una ley de potencia, una relación matemática entre dos cantidades, en la que un cambio relativo en una es proporcional a un cambio relativo en la otra. “Resulta que las clasificaciones del mundo real se sitúan en algún lugar del espectro caracterizado por cantidades inversamente proporcionales de desplazamiento y sustitución de elementos, independientemente de los detalles del sistema”, dice Íñiguez.

Esto significa que, si en un ranking hay mucho remplazo entonces habrá poco desplazamiento, y viceversa. Por ejemplo: en un ranking de empresas clasificadas por sus ganancias, en el que hay constantes cambios porque surgen unas y quiebran otras, habrá menos cambios en los rangos superiores e inferiores del ranking, es decir las empresas ubicadas en los primeros y últimos lugares variarán poco.

“Lo que me sorprende es que sean así de similares”, dice Íñiguez. “Que a pesar de las diferencias obvias (...), una vez que saltamos al espacio de las clasificaciones y nos preocupamos solo por las listas de cosas ordenadas, aparezcan patrones comunes de estabilidad temporal”.

### Una dinámica de rangos general

Los rankings sirven para reducir los sistemas complejos a listas ordenadas que permiten extraer información o tomar decisiones. Cada ranking está organizado en función de puntajes, que van desde los más amplios: la diferencia de población entre una ciudad y otra, por ejemplo, hasta los más pequeños: la diferencia en la velocidad de dos corredoras en una carrera de 100 metros.

Lo que advierte el equipo de investigadores es que, por un lado, pueden emerger patrones similares en los rankings si, en lugar de enfocarnos en sus detalles, como la especificidad de sus puntajes, los analizamos a partir de conjuntos de datos, sus rangos. “Cuando pasamos de puntajes a rangos, normalizas todas esas diferencias”, dice Gershenson.

Pero, además, al aplicar el modelo nulo, logran confirmar su hipótesis: “La dinámica de rangos no depende del mecanismo preciso de cómo cambian los rangos ni tampoco depende del mecanismo para asignar puntajes, que son muy variables. En realidad, hay una dinámica general que aplica a todos ellos”, dice Gershenson.

De manera que todos los rankings se comportan igual. Sin importar si hablan de animales o de tenistas, el cambio de rango es bajo en la parte superior de una clasificación. Y dependiendo de qué tanto desplazamiento o sustitución de elementos haya, el resto de la clasificación será más variable.

## Los límites de la predicción

Saber que todos los rankings funcionan igual tiene muchas implicaciones; una importante es empezar a verlos como conjunto de cambios aleatorios más que determinantes absolutas de “éxito” o “fracaso”.

“Una comprensión más matizada de las características genéricas de la dinámica de los rankings podría ayudarnos a limitar el agotamiento de los recursos en entornos competitivos (...) mediante mejores herramientas de clasificación algorítmica”, concluyen los autores en el artículo.

Tener mejores herramientas algorítmicas, por ejemplo, podría ayudar a identificar que la presencia de elementos en los primeros lugares de una lista puede ser el resultado de la interacción entre dinámicas del ranking más que de una absoluta superioridad.

Lo explica Íñiguez: “Teniendo en cuenta que las clasificaciones están vinculadas a la forma en la que asignamos recursos en la sociedad moderna (puestos de trabajo, becas, adquisiciones, et.), quizá merezca la pena observar que su dinámica puede captarse simplemente mediante un cambio de clasificación aleatorio. Tal vez la elección de los rankings, con sus picos naturalmente estables como señales de excelencia, no sea la mejor manera de luchar contra la desigualdad sistemática y el afianzamiento de las élites”.

En cuanto a la posibilidad de predicción, Carlos Gershenson dice que este estudio no es una forma de predecir cómo se comportarán los rankings. “Nuestro modelo confirma lo que mucha gente ya sabe: que los que están mejor rankeados son los que tienen mayor probabilidad de ganar un juego, por ejemplo. No es nada nuevo. Lo que sí podemos identificar, por ejemplo, es qué deporte, en un conjunto de rankings, será más fácil o difícil de predecir”.

Eso tiene muchas aplicaciones, especialmente para tomar decisiones: desde compañías de seguros a las que les interese medir riesgos, hasta las universidades que buscan saber qué tan difícil les resultará escalar posiciones en un ranking.

## Ligas de interés

- Artículo original en *Nature Communications*

[https://www.nature.com/articles/s41467-022-29256-x.epdf?sharing\\_token=wGNdpLctMvctvZhu1EhYKtRgN0jAjWel9jnR-3ZoTv0NqWzthxoCkpVSuNKNbRryaAfp\\_KGafduWW5hmFNUHBDp1nRhjec\\_AL0Heql\\_2teyDP1oEmHcwJwEUj29m0VAcl-TvEAqmNddJ5mO4B6wqXCltT1S8HHfj6Hcxrosj1beq8%3D](https://www.nature.com/articles/s41467-022-29256-x.epdf?sharing_token=wGNdpLctMvctvZhu1EhYKtRgN0jAjWel9jnR-3ZoTv0NqWzthxoCkpVSuNKNbRryaAfp_KGafduWW5hmFNUHBDp1nRhjec_AL0Heql_2teyDP1oEmHcwJwEUj29m0VAcl-TvEAqmNddJ5mO4B6wqXCltT1S8HHfj6Hcxrosj1beq8%3D)

- Con ecuaciones, físicos describen la evolución de las palabras

<https://www.fisica.unam.mx/es/noticias.php?id=2143>

- ¿Cómo ha cambiado el uso de las palabras en el tiempo?

<https://www.fisica.unam.mx/es/noticias.php?id=1050>