

# El canon de Google

**Miguel Espigado**  
**2007**



Para la construcción de un canon tradicional, un erudito o un conjunto de eruditos indagan en su conocimiento de la literatura para establecer el grupo de obras y autores más relevantes de una época, una corriente, una región, o bien de la universalidad al completo. El problema al que estas listas quieren dar respuesta es común a toda labor de difusión cultural. Ante la imposibilidad de contar con *todo* el tiempo, y *todos* los medios, los agentes culturales se ven obligados a difundir una selección y no una totalidad. En dicha selección siempre está presente la tensión entre lo que cada cual entiende por iluminar el mundo, desde su pretensión objetiva y filantrópica, y los compromisos, conscientes o inconscientes, que los agentes adquieren con los grupos a los que pertenecen.

En último término, se trata de decisiones personales o poco consensuadas. El sistema de trabajo imperante demanda de los cargos superiores una actividad resolutoria constante, que impide la participación del colectivo y a la vez incentiva el egocentrismo en la toma de decisiones. En el terreno de la Teoría, también predomina la categoría de autor frente a la de agrupación. En concreto, es el texto de autor lo que predomina sobre cualquier otra cosa, siendo el texto teórico, habitualmente un artículo publicado en una revista, un canal con evidentes carencias para el trabajo en equipo. Afortunadamente, los avances tecnológicos y sociales han permitido la generalización de foros virtuales y congresos, cuya proliferación es ya un indicativo del esfuerzo del colectivo por rellenar dichas carencias. El intercambio ágil de ideas acelera el estado de la cuestión y, sobre todo, democratiza las conclusiones. Su resultado textual debe recoger la opinión de todos, ya sea como divergencia o como consenso, limando particularidades e intereses individuales que alejen el discurso del objetivo bien común.

Es lógico pensar que los teóricos deben redoblar sus esfuerzos en esta dirección cuando sus decisiones vinculan a la sociedad. Al contrario que muchos otros debates del ramo, de cuya trascendencia solo participan los especialistas, el canon es la más poderosa forma de hacer política que la Teoría tiene a mano. De ahí la enorme controversia que suscita la cuestión a todos los niveles territoriales y artísticos. Lo que está en juego es nada menos que la selección que va a ser exportada a la cultura de masas, a través de exposiciones, ciclos, ediciones, programas educativos, campañas publicitarias, etc; una oportunidad casi única de romper el cerco minoritario al que las expresiones artísticas elevadas se ven sometidas.

Desde el espíritu de estas ideas se ha concebido el presente trabajo. Una herramienta tan sencilla y accesible como es Google, el buscador de Internet más utilizado del mundo, va a permitirnos establecer el índice de relevancia de un grupo

de autores que, en este caso concreto, se limita a los narradores en lengua española que estaban vivos en 1980. Cuando un usuario introduce un término en Google, un motor de búsqueda rastrea unas ocho mil millones de páginas Web y ofrece en forma de lista de enlaces todas y cada una de las páginas donde el término aparece, así como el número aproximado de las páginas implicadas. La cifra en sí misma es un cuantificador incontestable de la abundancia en Internet de un concepto, por tanto, de su importancia como contenido en la Red. De la comparación de la cifra relativa a cada miembro de un grupo, resulta una lista que ordena a esos miembros, del más importante- el más abundante-, al más secundario, es decir, escaso.

Lo que aquí se ha hecho ha sido anotar el número de apariciones en Google del nombre exacto de 2070 escritores en lengua española de todos los tiempos. La lista de autores ha sido extraída de Wikipedia, una enciclopedia libre *on line* que en su versión española cuenta con 184.250 artículos. La elección de la fuente no es fortuita. Wikipedia es un proyecto de creación colectiva donde cualquiera puede redactar un nuevo artículo para su publicación, o bien modificar uno existente. Eso quiere decir que cualquier omisión importante es rápidamente detectada y subsanada y, teniendo en cuenta que lo que aquí se pretende es extraer de la larga lista exclusivamente a los 75 escritores más relevantes, no importa tanto que ésta sea incompleta como que no haya ausencias significativas en los niveles superiores. Precisamente por su forma de creación, esto último queda descartado.

Realizamos la búsqueda y se nos muestra a los escritores ordenados por el número de veces que su nombre aparece en la parte de Internet que indexa Google, que es la gran mayoría. De entre ellos, se selecciona a los 75 con mayor número de entradas, los vivos en 1980 cuya dedicación fundamental sea la prosa literaria, ya sea en forma de narrativa o ensayo. Con ello se pretende proponer un canon: escritores en lengua española de prosa literaria de la actualidad (últimos 25 años). Como es lógico, cualquier otro corte hubiera sido posible como, por ejemplo, poetas españoles de los últimos cien años, escritores ingleses de la última década, etc. La lista resultante es la siguiente:

<b>Autores</b>	<b>nº entradas</b>	<b>Autores</b>	<b>nº entradas</b>
Gabriel García Márquez	2010000	Antonio Muñoz Molina	188000
Jorge Luís Borges	1780000	Tomás Eloy Martínez	182000
Isabel Allende	1550000	Carmen Martín Gaité	175000
Mario Vargas Llosa	1340000	Juan Carlos Onetti	174000
Julio Cortázar	1240000	Álvaro Mutis	172000
Octavio Paz	1170000	Guillermo Cabrera Infante	162000
Alejandro Jodorowsky	1120000	Juan Marsé	160000
Carlos Ruiz Zafón	1020000	Zoé Valdés	159000
Carlos Fuentes	1010000	Juan Manuel de Prada	155000
Camilo José Cela	906000	Augusto Monterroso	154000
Arturo Pérez-Reverte	864000	Enrique Vila-Matas	153000
Mario Benedetti	744000	Elvira Lindo	141000
Juan Rulfo	615000	José Donoso	138000
Alejo Carpentier	497000	Carmen Posadas	134000
Ernesto Sabato	397000	Alfredo Bryce Echenique	133000
Fernando Savater	375000	Félix de Azúa	131000
Javier Marías	365000	José Antonio Marina	129000
Eduardo Mendoza	345000	José Luis Sampedro	129000
María Zambrano	344000	Alberto Fuguet	123000
Manuel Vázquez Montalbán	310000	Fernando Vallejo	121000
Roberto Bolaño	307000	Ana María Matute	120000
Miguel Delibes	307000	Álvaro Cunqueiro	116000
Laura Esquivel	296000	Miguel Serrano	116000
Antonio Gala	294000	Gregorio Martínez	115000
Juan Goytisolo	282000	Gonzalo Torrente Ballester	110000
Francisco Ayala	267000	José Emilio Pacheco	109000
Elena Poniatowska	265000	Juan José Arreola	106000
Rosa Montero	255000	Mempo Giardinelli	106000
Manuel Rivas	249000	Ray Loriga	101000
Carlos Monsiváis	240000	Ramón J. Sender	98600
Francisco Umbral	230000	Lorenzo Silva	94900
Jorge Bucay	228000	Luis Antonio de Villena	91900
Manuel Puig	228000	Oswaldo Soriano	91800
Juan José Millás	227000	Jorge Volpi	91600
Adolfo Bioy Casares	223000	Francisco Coloane	90600
Álvaro Pombo	209000	Manuel Vicent	90500
Reinaldo Arenas	205000	Juan Villoro	90400
Almudena Grandes	189000		

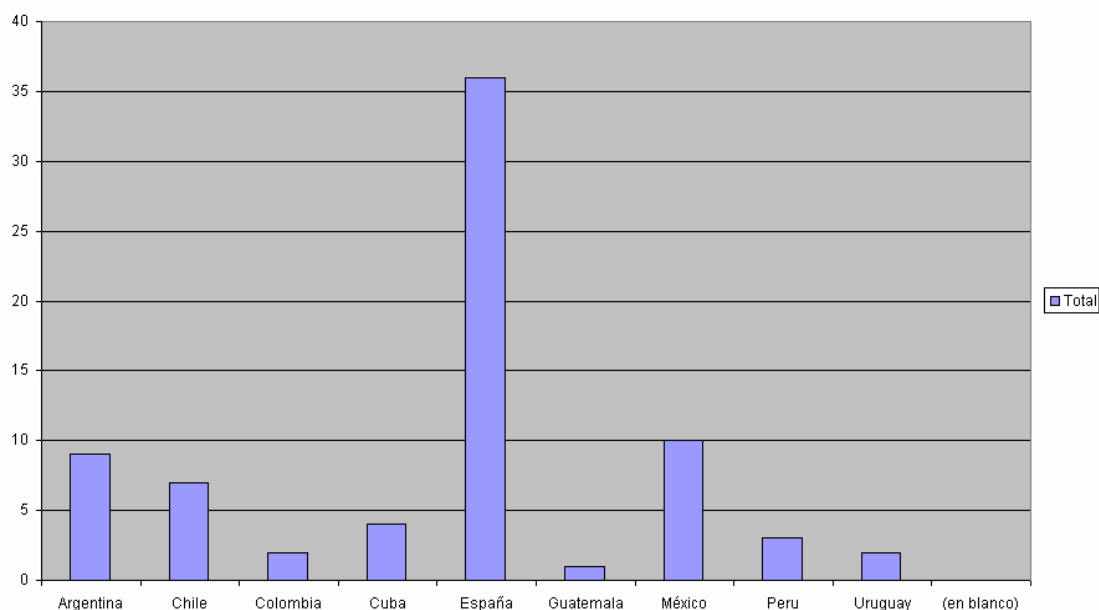
Antes de tomar en cuenta estos datos como un canon objetivo, conviene señalar algunas de las características y limitaciones del procedimiento. La objeción principal es técnica y se fundamenta en lo que se llama "ruido" de la búsqueda. Por ruido entendemos, entre otras cosas, todas las veces que Google indexa una página que contiene el término exacto de la búsqueda, pero dicho término alude a un referente diferente al deseado. En otras palabras; si buscamos "Juan Ruiz", nombre de un hipotético escritor, Google indexa todas la veces que aparece "Juan Ruiz" en Internet, sin que exista posibilidad de discernir entre los "Juan Ruiz" escritor y los

otros cientos de miles de "Juan Ruiz" que aparezcan en Internet como nombre de otras personas. Cuanto más vulgar sea el nombre de un escritor, más distorsionada al alza quedará la cifra que aquí se tiene en cuenta. "Ruido" para nuestros propósitos son también las veces que un escritor aparece por labores no estrictamente literarias. Si, por ejemplo, Mario Vargas Llosa hubiera ganado unas elecciones, su número de apariciones en Internet aumentaría masivamente al ser citado, además de cómo escritor, como presidente de la República Chilena en su legislatura. Para solventar este *handicap* se ha depurado la lista de algunos nombres cuya posición era altamente sospechosa por esta o la otra razón. Es innegable que en todas y cada una de las búsquedas hay un porcentaje de ruido, lo que obliga a tomar con cautela la posición de cada autor en la lista, no tanto su presencia. La precisión jerárquica disminuye cuanto menor sea el número de búsquedas que distancian a un escritor del adyacente, y por lo mismo, resulta muy fiable cuanto mayor. Cuando comparamos las entradas de Jorge Luís Borges con las de Carlos Fuentes, y nos encontramos una diferencia de 700.000 entradas entre uno y otro, no hay posibilidad de error. No así con, pongamos por caso, Ray Loriga y Gonzalo Torrente Ballester, cuya diferencia de entradas es demasiado pequeña para ser tenida en cuenta. Una convención extremadamente prudente sería fijar las 100000 entradas como diferencia a partir de la cual puede afirmarse la preeminencia en Internet de un escritor respecto a otro, y tratar en bloque a grupos o parejas de autores cuyas diferencias recíprocas fueran inferiores. Por todo ello puede decirse que tanto la posición como la presencia de los autores de la parte más inferior de la lista es más cuestionable que la del resto, y que tanto la presencia como la posición de los autores en la parte superior es irrefutable. En conjunto se trata una gama de datos que necesitan interpretación, pero datos al fin y al cabo, por tanto, empíricos. En la vida diaria aceptamos que un personaje es tanto más importante cuanto más se hable de él. En la Teoría, cuánto más se cite, cuánta más crítica genere a su alrededor. Eso es exactamente lo que aquí se está midiendo. Internet no es más que un gigantesco corpus de datos cuyos elementos pueden ser rastreados e inventariados. Aún reconociendo todas las imprecisiones, quedan al descubierto ciertas verdades gruesas que merecen nuestra atención.

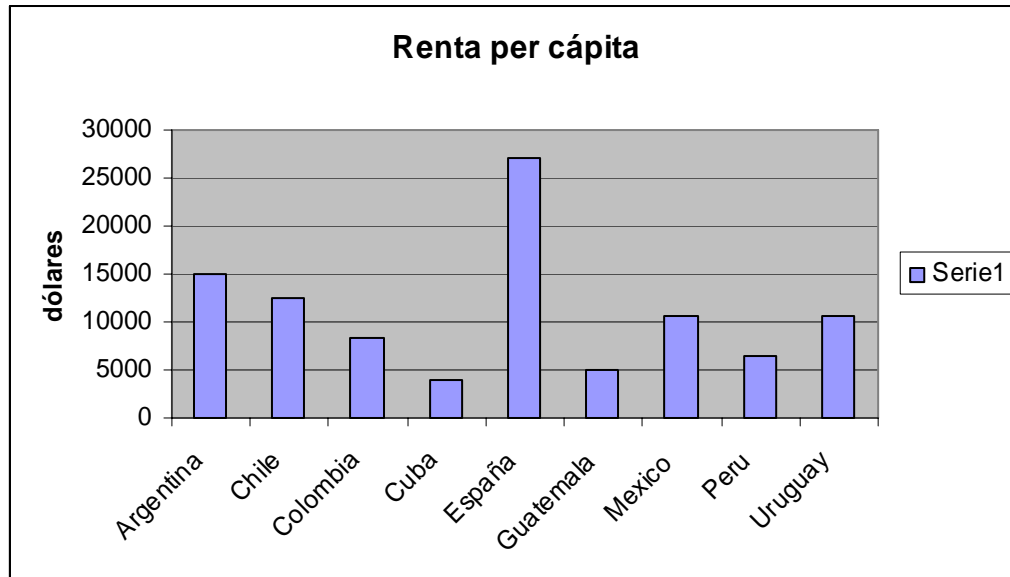
Como ya se ha dicho, lo que aquí se ha medido es la importancia de los autores según su número de apariciones en el Google. La lógica invita a pensar que un escritor es tanto más importante cuanto más se hable de él (no, desde luego, cuánto más se le lea, o los más importantes serían los autores de best-sellers y

libros de auto ayuda). Pero ¿Qué significa ser importante en Internet? Desde luego, no significa ser importante para la humanidad. Los autores más importantes de Google no son aquellos cuya obra tiene el valor de ejercer la influencia más positiva y benigna a los que los leen. Sin desmerecer necesariamente en lo anterior, estos escritores se diferencian del resto por aparecer más, literalmente, por estar más presentes y poseer en consecuencia mayor capacidad de influencia. Otras definiciones de importancia son subjetivas y, por tanto, no pueden ser cuantificadas. A efectos reales, y para lo que aquí nos ocupa, carecen de valor, pues cualquier canon se impone desde la cantidad; desde la cantidad de agentes que lo respalda, desde la cantidad de copias de un determinado texto que son asimiladas positivamente, etc, ya sea por imposición o de forma libre y espontánea. El Canon con mayúsculas de una nación o corriente es el que más se difunde, en el que se implican mayor número de agentes culturales y receptores. Dicha difusión y participación deja un rastro textual cuya porción accesible a Google forma parte de la estadística en la que aquí nos basamos. Al indexar no uno, ni varios, sino todos esos rastros textuales, el resultado es un auténtico canon de cánones, que resultaría incuestionable si pudieran salvarse las imprecisiones arriba señaladas. Ya quedó razonado en las primeras líneas cuál era la importancia de que el canon fuera fruto del mayor consenso posible. El presente trabajo trata, a su modo, de trabajar positivamente para este fin. El primer análisis de los datos arrojados indica, sin embargo, que este método no evita toda injusticia latente en los cánones tradicionales, pues se nutre de ellos. Las siguientes gráficas lo ilustran con bastante claridad:

**Número de autores por países en el top 75 de Google**



¿Existen razones extra culturales para explicar, entre otras cosas, el gigantesco saldo positivo de autores españoles respecto a los de cualquier otro país? Otro gráfico puede valer más que mil palabras.



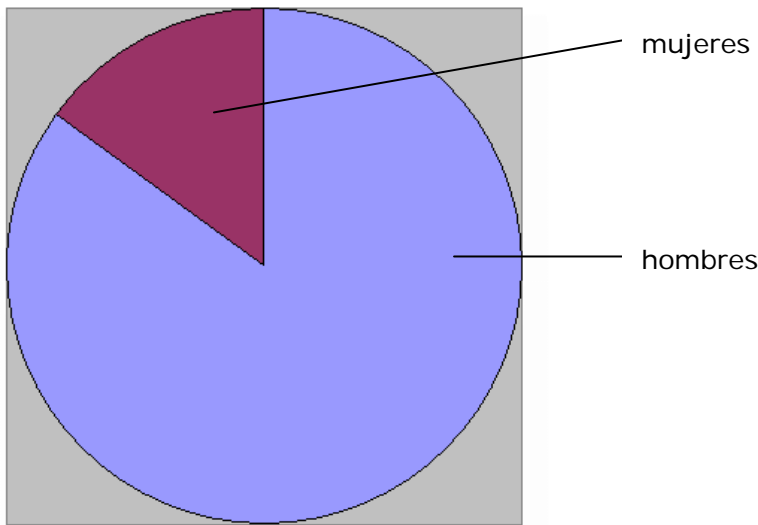
La similitud entre ambos gráficos es casi exacta. Si hacemos caso a estos datos, hay una relación inequívoca entre la situación económica de un país y el peso cuantitativo de su cultura en Internet. No es de extrañar si tenemos en cuenta que Internet exige una infraestructura estatal, empresarial y doméstica, y cuanto más desarrollada está la infraestructura, mayor número de ciudadanos accede Internet y tanto más se incrementan los contenidos y servicios a los que éstos son adictos. El canon de Google se revela entonces no como una consulta democrática, donde cada participante cuenta con un voto, sino más bien como una oferta pública de compra de acciones, donde las carteras con más liquidez se aseguran una mayor participación y, por tanto, influencia, en la empresa global que podemos llamar Cultura.

Debemos tener en cuenta que los contenidos de Internet, en su mayor parte, no son más que el tentáculo virtual de algo que está actuando y ocurriendo en la vida física (pocas actividades son exclusivamente virtuales, por ahora). En ese sentido, Internet representa la actividad visible y pública del mundo, aquello que se muestra. Parece entonces lógico dar cierto crédito al canon de Google como representación no solo de lo que ocurre en Internet, sino de lo que ocurre.

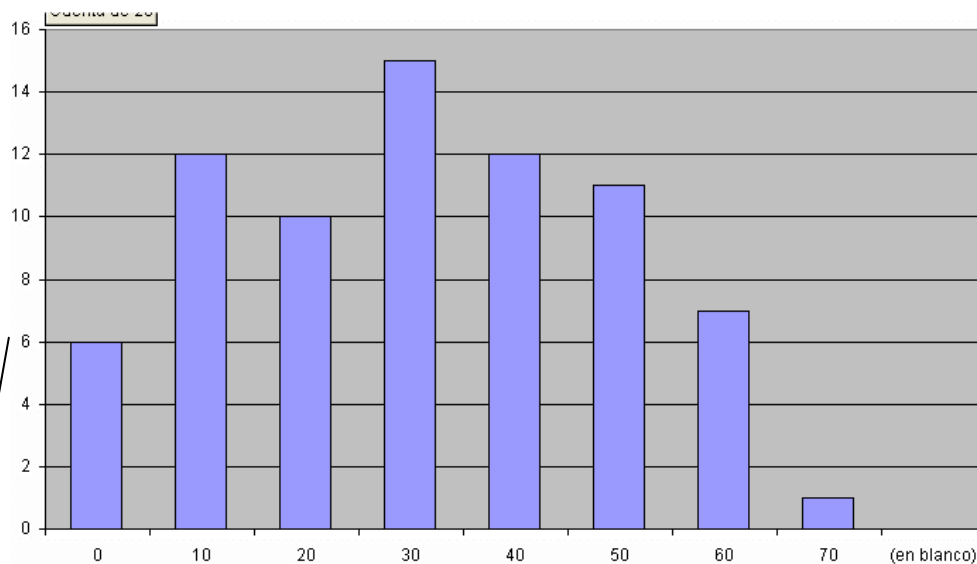
Para finalizar, me gustaría dar al término "canon de Google", y a lo que aquí se ha dicho sobre sus posibilidades, una dignidad mayor de la que se desprende de mi

lista de autores. La elaboración casera de la misma pretende solo esbozar un camino que podría culminar en la creación de valiosas herramientas de medición del impacto cultural a través del rastreo de Internet. Esta naturaleza rudimentaria, primeriza, del método tampoco debe ser óbice para negar en bloque los resultados de la lista; aquí se ha argumentado razonablemente la fiabilidad de unos cuantos. Quedan otros muchos en el aire y la invitación implícita a jugar con la lista y sacar conclusiones propias. Como ejemplo, incluyo al final otras dos gráficas que corresponden a la variable edad y a la variable sexo de los escritores.

**Porcentaje por sexo en el top 75 de Google**



**Escritores por edades en el top 75 de Google**



Número de escritores

Década del SXX en la que nacieron