

El reconocimiento de las formas documentales en las búsquedas de información de calidad: una apuesta por la educación en la “docudiversidad”

The importance of document type in high-quality information searches: a proposal for education on document diversity

Cristóbal Urbano; Jesús Gascón-García

Urbano, Cristóbal; Gascón-García, Jesús (2021). “El reconocimiento de las formas documentales en las búsquedas de información de calidad: una apuesta por la educación en la ‘docudiversidad’”. *Anuario ThinkEPI*, v. 15, e15f02.

<https://doi.org/10.3145/thinkepi.2021.e15f02>

Publicado en *IweTel* el 24 de septiembre de 2021

Cristóbal Urbano

<https://orcid.org/0000-0003-0935-6436>

Universitat de Barcelona
Facultat d'Informació i Mitjans Audiovisuals
Centre de Recerca en Inf., Com. i Cultura (CRICC)
Melcior de Palau, 140. 08014, Barcelona (España)
urbano@ub.edu

Jesús Gascón-García

<https://orcid.org/0000-0002-0712-7210>

Universitat de Barcelona
Facultat d'Informació i Mitjans Audiovisuals
Centre de Recerca en Inf., Com. i Cultura (CRICC)
Melcior de Palau, 140. 08014, Barcelona (España)
gascon@ub.edu



Resumen: Los usuarios de sistemas de información no discriminan la forma y función de los recursos de información y las ignoran en sus búsquedas. Mayoritariamente utilizan buscadores web genéricos: siempre obtienen resultados y no cuestionan su calidad, procedencia ni si su forma es adecuada a los objetivos. En muchas fuentes de información tampoco aparece la forma o género como argumento de búsqueda o de filtrado. El reconocimiento del género/forma es útil en la búsqueda, pues implica expectativas del tipo de información que se quiere obtener. Los tipos documentales, especialmente en internet, han cambiado y sus denominaciones y definiciones no se han consolidado. Hay que fijar la terminología en los vocabularios documentales

y en la web semántica. La sustitución de catálogos y bases de datos por herramientas de descubrimiento contribuye a la disolución del concepto de género/forma: la búsqueda por materias no es uniforme y el usuario no percibe la diferencia de formas/géneros obtenidos, mezclándose resultados de muy diversos tipos. Es necesario mejorar los sistemas de representación y recuperación de información, incorporando la forma documental como elemento de búsqueda, pero también se requiere una alfabetización informacional (especialmente en los futuros documentalistas) que ponga en valor estas formas.

Palabras clave: Formas documentales; Géneros documentales; Búsqueda de información; Representación de la información; Recuperación de contenidos; Alfabetización informacional.

Abstract: Users of information systems do not differentiate the forms or functions of information resources and thus ignore them when carrying out searches. They use generic search engines, always obtaining results but without consideration of their quality or precedence, nor whether their form is adequate for the requirements. Moreover, many information sources lack form and genre as search terms or filters. However, the consideration of form/genre is useful in information searches as the user may expect a certain type of information. The types of documents available, especially on the internet, are changing, but their names and definitions have not been consolidated. It thus becomes necessary to define such terminology used in document vocabularies and the semantic web. As catalogs and database interfaces are replaced by web-based discovery systems, the concept of genre/form is disappearing. Subject-specific searches are not uniform, and users do not appreciate the difference among forms/genres in the results obtained, where very different types may be mixed. It is necessary to enhance the representations used by information and retrieval systems by adding document form/genre as a search term. However, informational literacy is also required, especially among future information professionals, to highlight the importance of document forms.

Keywords: Document form; Documental genre; Information search; Content retrieval; Informational literacy.

1. Introducción

La presente nota *ThinkEPI* tiene su origen en algunas conversaciones de pasillo, entre clase y clase, mantenidas entre un docente dedicado al análisis documental y la representación de la información [Jesús Gascón] y otro a los servicios de referencia y recuperación de información [Cristóbal Urbano]. Por lo tanto, responde a ese espacio común, a menudo ignorado, entre la perspectiva de la organización del conocimiento y la de los servicios de información: los usos y los usuarios. En nuestras charlas marcadas por un tono melancólico (reconozcámoslo: la edad no perdona), destacan las reflexiones sobre los cambios observados en los comportamientos de búsqueda entre los universitarios. Uno de esos cambios ha sido objeto de especial atención entre nosotros últimamente: la tendencia hacia la disolución o ignorancia de las formas y los atributos de los documentos en el marco mental de quien busca, que lleva a no tomar en consideración el potencial que ofrece la distinción entre forma y materia como determinante de las búsquedas de calidad.

Por ello, partiendo de una reflexión en la que participamos que viene de lejos (**Espelt; Gascón, 1999**) y que se basa en el reconocimiento de la importancia que tiene la "cristalización" de la información en documentos, esto es en información registrada con formatos intencionalmente diversificados, trataremos de presentar una breve panorámica sobre los comportamientos de búsqueda de los usuarios, sobre cómo se ha tratado el tema de las formas documentales conceptualmente y desde la perspectiva de los procesos técnicos orientados a mejorar los sistemas de representación/recuperación de información según su forma documental. Finalizaremos con la respuesta que creemos imprescindible a la vista del entorno y de la complejidad del problema: la apuesta por una alfabetización informacional para usuarios avanzados (entre ellos, especialmente, los profesionales de la información en sus etapas de formación) que ponga en valor las formas documentales en su marco mental a la hora de realizar búsquedas de información de calidad, en línea con lo que al respecto apunta el *Framework for information literacy for higher education* (ACRL, 2016).

"Hay una tendencia hacia la disolución o ignorancia de las formas y los atributos de los documentos en el marco mental de quien busca, que lleva a no tomar en consideración el potencial que ofrece la distinción entre forma y materia como determinante de las búsquedas de calidad"

2. Bajo el síndrome del *good enough*

En efecto, cuando los usuarios formulan una consulta en buscadores o en bases de datos que reúnen recursos con una amplia variedad tipológica, así como cuando ante la obtención de un número abultado de resultados se enfrentan al filtrado de la información recuperada, el hecho de que diversos recursos puedan tratar un mismo asunto pero con formas y grados de elaboración diversos pasa muy desapercibido, aun siendo clave para seleccionar mejor la información. Sin duda, cualquier búsqueda de información se torna mucho más significativa si el usuario es capaz de discriminar entre un post en un blog, una página web corporativa, una noticia de prensa, un póster en un congreso, un *preprint*, un artículo de revista (ya sea de investigación, de revisión sistemática o de divulgación), un libro académico, un libro blanco, una memoria oficial, un informe técnico, o un artículo de enciclopedia... entre otros

muchos tipos posibles. Más aún, que los usuarios sean conscientes de la existencia de diversos tipos de documentos con características y funciones diferenciadas es el primer paso para que con el tiempo y el dominio progresivo de su campo de conocimiento tengan la capacidad de acotar sus búsquedas a fuentes de información especializadas en el tipo de documento que necesitan (por ejemplo, bases de datos de legislación y jurisprudencia). Fuentes que pueden ser más adecuadas que las consultas realizadas en grandes buscadores web, agregadores genéricos de contenidos en acceso abierto o en herramientas de descubrimiento bibliotecarias, todas ellas caracterizadas por reunir un considerable volumen de recursos de naturaleza y procedencia heterogéneas.

La disolución de las formas documentales en el marco mental de quien busca información se podría situar en un contexto más amplio de cambio en los comportamientos de búsqueda. No descubrimos nada si decimos que hace tiempo que la navegación hipertextual por internet, el descubrimiento "aleatorio" de contenidos vía redes sociales y, de forma especial, las consultas en buscadores web como *Google* (*Google Scholar* de manera destacada en el entorno universitario), se han convertido en las principales vías de búsqueda y recuperación de información para la mayoría de las personas. Se busque como se busque, siempre hay unos resultados, con lo que la sensación de éxito se instala en quien busca, que queda suficientemente satisfecho. Casi pasa a segundo plano que los resultados encontrados sean de mejor o peor calidad o que se adecúen plenamente o no a las intenciones de la búsqueda: algo saldrá de ellos, lo que es especialmente conveniente ya que el proceso de búsqueda acostumbra a ser rápido, cómodo y toma en cuenta datos del perfil del usuario en redes sociales y buscadores web para personalizar el resultado, aunque el usuario no sea plenamente consciente de ello (Merlo-Vega, 2012; Urbano, 2018). Así pues, el comportamiento informacional de los usuarios en general (y de las generaciones más jóvenes en particular) viene marcado por lo que es más conveniente en términos del esfuerzo y del tiempo a invertir en las búsquedas, realidad que Connaway, Dickey y Radford (2011) reflejaron magistralmente en el título que dieron a un destacable trabajo de campo sobre el asunto: "If it is too inconvenient I'm not going after it': convenience as a critical factor in information-seeking behaviors".

Es lo que en inglés se conoce como *good enough* y que, aplicado a los entornos de búsqueda de información (Hyer et al., 2009; Medeiros, 2006; Richards, 2021), viene marcando las últimas tendencias del diseño de sistemas de recuperación para hacer frente al entorno de sobreabundancia informativa. Reconociendo las muchas y positivas contribuciones a la experiencia de usuario que ese enfoque ha aportado, cabe señalar que presenta un importante problema cuando el usuario deja de preguntarse sobre la procedencia y la lógica de esos resultados. Los buscadores web han generado una "cultura de la búsqueda" (Hillis; Petit; Jarret, 2013) que crea la sensación de que casi todo lo relevante está registrado, disponible y se puede buscar y encontrar consultando una única fuente. Esto es, se asume que "todo está en internet", que la Web en singular es "la" fuente de información. Se concibe ese espacio web como un universo etéreo de información "aparentemente líquida" (Area-Moreira; Pessoa, 2012), repleto de datos que esperan que un buscador los encuentre para ser ofrecidos a quien plantea la búsqueda, plenamente confiado en la bondad del algoritmo y de la personalización del resultado.

Ahora bien, no toda la culpa de la ignorancia de las formas se puede imputar a los usuarios y su proceso de "googleización" (Haider; Sundin, 2019; Vaidhyanathan, 2011). Una primera aproximación a las interfaces de consulta de un buen número de fuentes de información bibliográfica (entre las que no dudamos en incluir los buscadores generalistas, dados los hábitos de los usuarios) muestra debilidades en su capacidad de articular, frente al contenido temático, la forma documental y su materialización en un tipo de soporte concreto. Un primer problema radica en la definición, categorización e inventario de los diversos tipos de documentos a considerar. Una segunda consideración tiene que ver con que tanto en las opciones de búsqueda como en las de filtrado, las vías para delimitar por tipo de recurso según su forma acostumbran a ser pobres, inconsistentes o pasan desapercibidas, casi deliberadamente, en el intento de obtener una interfaz simple y limpia (como en el caso de *Google*), pese a que las posibilidades técnicas para tomar en consideración el factor forma documental son actualmente muy elevadas.

"El comportamiento informacional de los usuarios en general (y de las generaciones más jóvenes en particular) viene marcado por lo que es más conveniente en términos del esfuerzo y del tiempo a invertir en las búsquedas"

"Los buscadores web han generado una 'cultura de la búsqueda' que crea la sensación de que casi todo lo relevante está registrado, disponible y se puede buscar y encontrar consultando una única fuente"

3. La naturaleza diversa de los documentos vista desde la teoría interdisciplinar de los géneros

Si bien en la comunidad bibliotecaria tradicionalmente se hablaba de "forma" para referirse a las variadas presentaciones y estructuras de los documentos, en los últimos tiempos se ha ido introduciendo un enfoque más integral que toma en consideración la palabra "género", más allá de su significado ligado a los géneros literarios, para denominar la naturaleza variada y diversa de la producción documental. **Lee y Zhang** (2013), tras analizar el tratamiento dado a los géneros en la tradición catalográfica angloamericana, afirman que las autoridades catalográficas fallaron, de forma persistente, al fijar una definición rigurosa y útil para ambos términos (*form* y *genre*) o una distinción clara y consistente entre los dos. En este sentido resultan muy significativas las críticas de **Wilson y Robinson** (1990) a los criterios de aplicación y de mantenimiento de las subdivisiones de forma de los *Library of Congress Subject Headings* en los años ochenta del siglo pasado.

Una ambigüedad que, a partir de un documento de trabajo de **Taylor** (1992), trató de gestionar el *Subject Analysis Committee* de la *American Library Association* (1993) fijando criterios de discriminación y ejemplos de lo que serían formas y géneros, entendidos como algo que se representa de una forma equivalente una vez realizado el análisis de contenido:

"Los datos del formulario son los términos y frases que designan tipos o géneros específicos de materiales. Los materiales designados con estos términos o frases pueden determinarse mediante un examen de:

- su carácter físico (ejemplo: videocasetes, fotografías, mapas, folletos);
- el tipo particular de datos que contienen (ejemplo: bibliografías, cuestionarios, estadísticas);
- la disposición de la información dentro de ellos (ejemplo: diarios, esquemas, índices);
- el estilo, técnica, propósito o audiencia prevista (ejemplo: drama, romances, dibujos animados, anuncios, obras populares);

O una combinación de los anteriores".

En la actualidad, la mención a ambos términos en este ámbito profesional sigue siendo ambigua: en ocasiones se tratan como sinónimos y en otras como cosas ligeramente diferentes pero que catalográficamente se tratan de forma equivalente. Por ello se presentan mediante la unión de las dos palabras, como sucede en el campo "655 - Index Term-Genre/Form" del formato MARC 21 (*Library of Congress*, 1999) o en el título del tesaurus destinado a cumplimentar dicho campo: *Library of Congress genre/form terms for library and archival materials (LCGFT)* (*Library of Congress*, 2021). Dado lo actualizado que está el LCGFT con sus 2.391 entradas a fecha de marzo de 2021, resulta relevante ver cómo se presenta la dicotomía género/forma en la introducción a dicha herramienta:

"... [LCGFT] es un tesaurus que describe lo que es una obra frente a lo que esta trata. [...] El tesaurus combina géneros y formas. La forma se define como una característica de las obras con un formato y/o propósito particular. Un "corto" es una forma particular, por ejemplo, al igual que "animación". El género se refiere a categorías de obras que se caracterizan por tramas, temas, escenarios, situaciones y personajes similares. Ejemplos de géneros son westerns y thrillers. En el término películas de terror, "terror" es el género y "películas" es la forma".

Ahora bien, hemos de considerar que la diversidad documental merece ser analizada desde un marco mucho más amplio que el meramente catalográfico, en el que el concepto género se debe entender como un término amplio bajo el que podemos considerar tanto formas documentales en el sentido tradicional (diccionario, enciclopedia, bibliografía, artículo científico, patente, etc.) como géneros de expresión textuales y no textuales (literarios, audiovisuales, musicales, etc.). Especialmente a partir de la diversificación documental vivida con la digitalización y el acceso web a la información, se ha prestado mucha más atención a ese concepto amplio de género por parte de las Ciencias de la Información y la Documentación y se ha situado en un plano interdisciplinar de la mano de la teoría de los géneros (**Andersen**, 2008; **Montesi**, 2008). Se trata de un concepto de género que tiene su punto de partida en el artículo seminal de **Miller** "Genre as social action" (1984) donde se conciben los géneros como el resultado de "acciones retóricas tipificadas basadas en situaciones recurrentes". En la misma línea, **Orlikowski y Yates** proponen una definición de género desde una perspectiva de documentación en las organizaciones que resulta muy clarificadora por su utilidad en todo el ámbito de la Información y la Documentación:

"Un género de comunicación organizacional, como la carta comercial, la junta de accionistas o el informe, es un tipo distintivo de acción comunicativa caracterizada por un propósito comunicativo socialmente reconocido y aspectos comunes de la forma" (**Orlikowski; Yates**, 1994).

Desde las Ciencias de la Información y la Documentación, un referente clave es **Andersen** (2008; 2015; 2017) ya que eleva la mirada y plantea los géneros como formas de comunicación y actividad humana más allá del análisis formal del mero documento como objeto, lo que permite profundizar en cómo se organiza la información en un contexto concreto de comunicación y de acción social. De esta forma apuesta por una visión más holística, en la que el reconocimiento del género es un factor importante en la comunicación del conocimiento, que a la postre resulta útil para el usuario en el momento de plantear sus búsquedas documentales, ya que el conocimiento previo de la forma/género proporciona unas expectativas de qué información puede encontrarse, de su propósito y su utilidad. Para Andersen, el género está ligado a los usos de los documentos como herramientas de comunicación tanto por parte de quienes los crean como de las comunidades donde circulan: cómo se utilizan, sus objetivos, métodos de producción, etc., definen la respuesta a una necesidad de información. Estudiar el género contribuye a situar cómo se buscan y usan diferentes tipos de recursos, partiendo de la razón de ser y el contexto en el que fueron creados.

La importancia de este reconocimiento de la forma/género por parte del usuario, que facilita la comprensión del recurso informativo, ya fue destacada con anterioridad por **Kwaśnik** y **Crowston** (2005) quienes en una introducción a un número especial de la revista *Information technology & people* dedicada a los géneros de los documentos digitales ponen de relieve la importancia del medio/soporte en el proceso de diversificación documental, lo que explica el interés que despertó la aplicación de la teoría interdisciplinar de los géneros a los intentos de clasificación automática de documentos web. En esta línea, **Montesi** (2008) hace una interesante revisión bibliográfica de las principales contribuciones sobre el género aplicado a los recursos web, apuntando ya a la utilidad de su uso en la optimización de la recuperación de la información, dado el gran potencial que su tratamiento automatizado y masivo puede tener para la mejora de los buscadores web.

4. El proceso técnico de la diversidad documental: terminología, representación y sistemas de información

Si volvemos la vista atrás, vemos que lo que se planteaba a finales del siglo pasado (**Espelt; Gascón**, 1999, por ejemplo), en relación con la relevancia de los aspectos formales como elemento de recuperación y a su representación en registros bibliográficos o en metadatos asociados a los recursos, no parece haber acabado de funcionar a pesar del tiempo transcurrido. La problemática se ha de analizar en tres campos:

- la diversidad y volatilidad de las realidades documentales y, en consecuencia, de la terminología que usamos para denominarlas;
los instrumentos y procesos que se pueden utilizar para representarlas;
- los recursos tecnológicos con los que se procesan dichas representaciones y permiten la recuperación de la información.

Si atendemos a la esfera de la terminología, podemos ver que la explotación de las formas y atributos de los documentos pasa en buena medida por disponer de un buen consenso sobre su identificación y sobre la forma que tenemos de nombrarlos, primer paso para poder utilizar representaciones normalizadas en vocabularios controlados, clasificaciones, etc. Sin duda es un problema importante en el mundo cambiante y dinámico de internet. Pongamos por ejemplo el caso de la palabra "portal": su uso es frecuente entre usuarios y profesionales, pero se asigna a realidades muy diversas, hasta el punto de que no existe una definición clara del término y, en ocasiones, se identifica simplemente con cualquier sitio web de una cierta complejidad y profundidad. Otra de las palabras más exitosas para calificar un tipo de fuente de información es la palabra "buscador", independientemente de que la fuente sea un *web search engine* tipo *Google*, un catálogo de productos o una base de datos acotada a ciertos documentos bibliográficos a texto completo. Todo recurso de información en la web que ofrece una caja de búsqueda como elemento central de la página inicial se percibe por muchos usuarios como un "buscador" de algo (estadísticas, tesis, sitios web, libros... personas).

Y es que a la dificultad clásica de inventariar formas y atributos de los documentos se ha añadido la riqueza creativa del entorno web, en el que se han generado nuevas formas, tanto las resultantes de la evolución desde el entorno analógico como otras de nuevo cuño (**Kwaśnik; Crowston**, 2005), ya se trate de recursos de información directos como de recursos indirectos, fuentes secundarias que ofrecen "información sobre la información" y que han venido a sumarse o a suplir con éxito relativo a las ya existentes. A menudo, no son más que nuevas formas adaptadas a la Web de fuentes anteriores y reciben denominaciones que, por poco consolidadas y mal definidas, no resultan claras.

De todas formas, el problema terminológico de la representación de la forma/género en lenguajes controlados viene de lejos y se agudiza en la medida en que buena parte de las listas de encabezamientos

de materias y tesauros beben de originales en inglés que se acaban traduciendo de forma muy variada y no siempre consistente dentro de un mismo ámbito lingüístico. Como ejemplo, véase el caso de la proliferación de traducciones diferentes para subdivisiones de forma en listas de encabezamientos en español (**García-Sánchez, 2013**). Un problema que se ha ido agudizando con el dominio que tiene el inglés en el desarrollo de internet.

Si atendemos al segundo aspecto, el campo de los procesos técnicos de representación, partimos de un entorno documental y profesional en el que tradicionalmente hemos basado la recuperación de la información en el uso de vocabularios que representan el contenido informativo de los recursos. Principalmente, su contenido temático, los asuntos tratados, olvidando que también afectan al contenido los aspectos formales que lo definen y lo hacen aprehensible. Especialmente en vocabularios precoordinaados, el uso de subdivisiones o auxiliares de forma que representan formas diferenciadas permite refinar la recuperación y diferenciar recursos de similar contenido temático, pero de usos bien diferentes a causa de sus características formales. En catálogos de biblioteca, una parte importante de los registros cuenta con puntos de acceso de materia donde pueden encontrarse representados aspectos formales: en el subcampo \$v de los campos 6XX de MARC21, en el campo específico de género/ forma (655), e incluso con auxiliares de clasificaciones, como las de forma o características de CDU o en valores de control de MARC21 (en etiquetas como 008 o 007), cada uno de los cuales puede contribuir a recuperar estos aspectos y que podrían ser accesibles y recuperables por los buscadores generalistas, fuera del entorno del catálogo.

Pero, aun con una taxonomía clara de formas y atributos documentales, conocida y aceptada por la comunidad profesional, el esfuerzo será inútil si ello no se traduce en representaciones de las formas/ géneros en los registros o recursos de información disponibles en bases de datos, catálogos, etc. y, sobre todo, si los sistemas no hacen recuperables esos atributos. Por ello, es indispensable que las interfaces de los sistemas de información (herramientas de descubrimiento, catálogos, buscadores, etc.) permitan ver claramente y recuperar de forma solvente estos atributos. Una duda que ya se plantearon algunos de los autores implicados en el desarrollo del campo 655 (**Hemmasi; Miller; Lasater, 1999**) y que ratifica el análisis actual de ciertas interfaces de catálogo y herramientas de descubrimiento, así como la opinión de los propios profesionales (**Bitter; Tosaka, 2020**).

Finalmente, si giramos la vista hacia la explotación automática y masiva de contenidos digitales (textuales, factuales o audiovisuales), basada en lenguaje natural (a partir de sus elementos formales y estructurales, o de sus metadatos cuando están disponibles), todo indica que la promesa de la web semántica tiene fundamentos suficientes para que tengamos fe en el potencial de la representación automática de las formas y los atributos de los recursos, aunque quizás en los últimos diez años no se

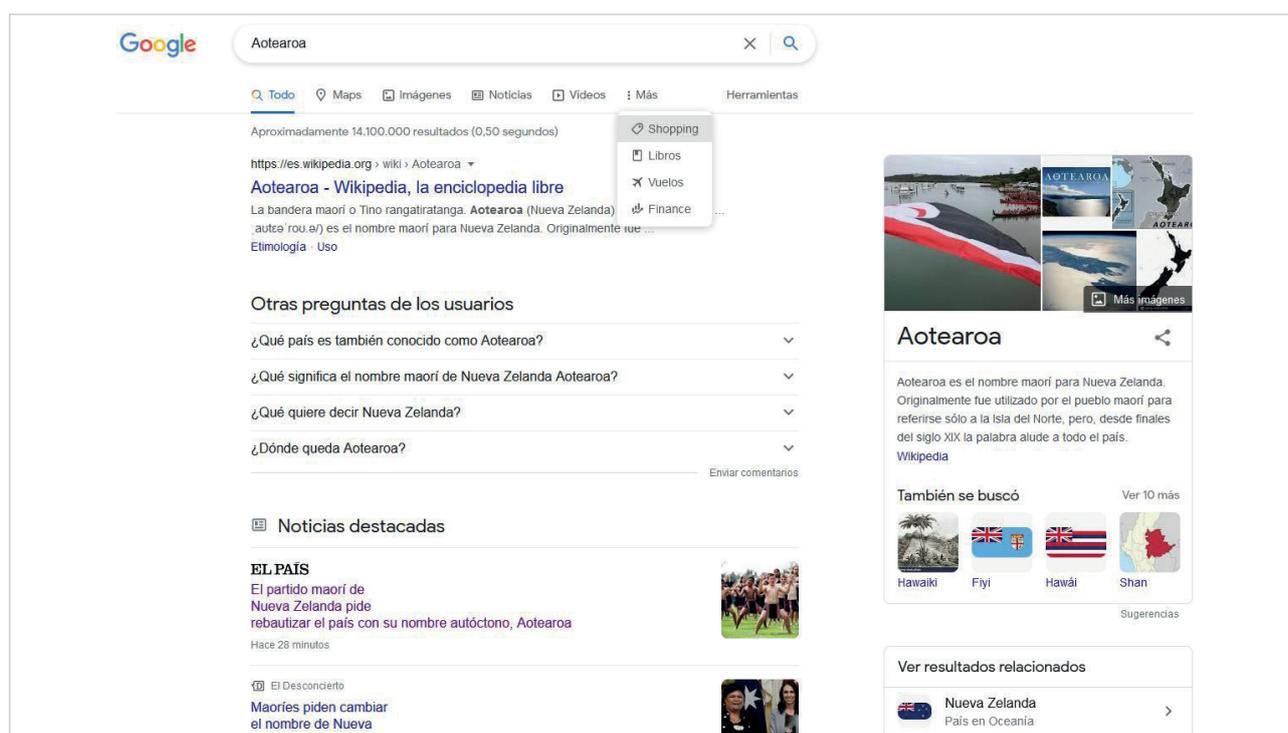


Figura 1. Snippets enriquecidos y menú con opciones de filtrado en una página de resultados de Google.

ha avanzado todo lo esperado (Codina; Pedraza-Jiménez, 2011). Más allá de lo que nos muestran hoy en día los buscadores web, se supone que, en un futuro no muy lejano, una nueva generación de buscadores será capaz de automatizar la representación del género/forma. Bien mediante técnicas de inteligencia artificial y *machine learning* para clasificar automáticamente los documentos, bien mediante la extracción de información ahora "encerrada" tanto en los documentos mismos si están bien normalizados en cuanto a estructura y metadatos (Dragon, 2020), como en catálogos y bases de datos (mediante iniciativas en funcionamiento como *Schema.org* o en desarrollo como *BibFrame*, destinado a este fin específico), ofreciendo la posibilidad de buscar aprovechando los vocabularios controlados y las relaciones que establecen y que ahora se circunscriben a esas herramientas.

Queda de manifiesto que una clasificación o definición de géneros en la Web es compleja, pero que hay base suficiente para desarrollar algoritmos que puedan identificar automáticamente la forma/género de un recurso electrónico o explotar la información estructurada de los documentos mismos o de las bases de datos que los contienen, para así limitar los resultados de las búsquedas. Pese a las dificultades de su traslación al entorno más complejo de las búsquedas académicas, hoy por hoy el funcionamiento de buscadores como *Google* permite anticipar el potencial de la clasificación automática de documentos. Así se puede observar en la funcionalidad que tiene *Google* para presentar de forma directa y destacada en la pantalla de resultados ciertos tipos de datos extraídos de documentos, así como las opciones de filtrado a partir de una página de resultados (figura 1), que permiten discriminar contenidos tipológicamente diferenciados pensando en una audiencia generalista con usos de la vida cotidiana y el ocio (entradas en diccionarios o enciclopedias, mapas, oferta de vuelos en los sitios web de compañías aéreas o de comparadores de precios, noticias, libros...). Por ello cabe pensar en la viabilidad de que los buscadores académicos puedan ofrecer en el futuro opciones para discriminar entre diferentes tipos de documentos de valor académico o profesional. Se trata de algo que no se produce de forma eficiente en muchos casos cuando se consulta *Google* sin una gran elaboración de la ecuación de búsqueda, como se puede ver en el caso de la consulta sobre informes económicos de la figura 2, en la que se observan resultados que pueden ser *good enough*, pero en la que se mezclan diversos tipos documentales sin la aparición de *snippets* enriquecidos.

Si dirigimos la vista hacia las herramientas de descubrimiento, cada vez más abundantes en los sitios web de bibliotecas (Rodríguez-Yunta, 2015), se observa que han asumido el modelo de pantalla de búsqueda inicial en forma de una caja en la que formular consultas tal y como los usuarios hacen en *Google*, junto a una búsqueda combinada ("avanzada") que solo se da en segunda instancia, aunque de forma más explícita que en el caso de *Google*. Acompañan los resultados con posibilidades de filtros (figura 3) que permiten acotar los resultados en función de criterios diversos, obtenidos, sobre todo, de metadatos (como los valores MARC21 o los puntos de acceso de materia asignados en los registros de catálogo). Algunos de estos filtros corresponden a tipo de soporte (vídeos, imágenes, textos, recursos electrónicos...) y otros a materias. Según el diseño de la herramienta y, en última instancia, en función de los catálogos y bases de datos de los que extrae la información, también ofrecen filtros de formas documentales. En general, se limitan a extraer palabras de los encabezamientos de materia (incluyendo las subdivisiones de forma de los campos 6XX de MARC21) y, si se dan, encabezamientos de género y forma (del campo 655).



Figura 2. Resultados de una búsqueda poco elaborada en *Google* sobre informes de economía.

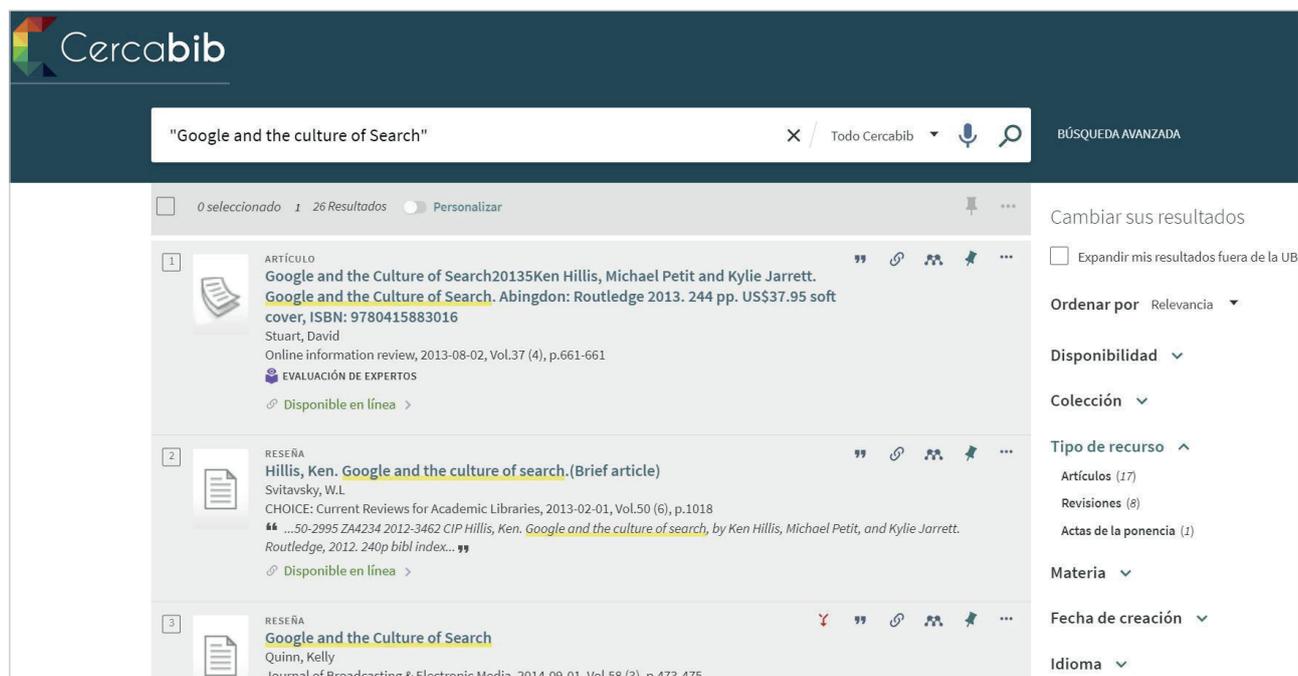


Figura 3. Búsqueda de un libro por su título con la herramienta *Primo* de *ExLibris* en una instalación bibliotecaria. No se recupera el libro en cuestión, pero sí reseñas, etiquetadas como tal en algunos casos (8) y en otros como artículo (17).

Ahora bien, mucho nos tememos que los usuarios no distinguen o no utilizan suficientemente los filtros que se ofrecen, pocos de los cuales permiten, realmente, acotar contenidos de manera consistente. Un punto débil que explica esta situación es la falta de uniformidad en los tratamientos por parte de cada una de las fuentes que alimentan la agregación de metadatos que realizan los proveedores de herramientas de descubrimiento para bibliotecas. Mientras que en registros de catálogos se puede constatar una cierta uniformidad (aun utilizando herramientas diferentes, hay criterios y procedimientos normalizados), los registros de bases de datos comerciales, indizados y tratados con criterios muy diversos no recogerán los mismos atributos y lo mismo pasará con recursos almacenados en repositorios digitales u otras plataformas de publicación en acceso abierto, donde los criterios de asignación de metadatos pueden variar enormemente.

5. Apuesta por el factor humano: educación para el reconocimiento de la diversidad documental

Nuestra percepción es que cada vez más, los usuarios tienen dificultades para discriminar la forma y función de ciertos recursos de información, esto es, la identificación tipológica y terminológica del tipo de documento. Es así que ante todas las posibilidades y realidades tecnológicas que hemos apuntado en el punto anterior, podríamos hablar de una posible "paradoja de la disolución de las formas": si se diera el caso de que en un inmediato futuro se concretaran grandes mejoras en las herramientas de representación de las formas/géneros, en el reconocimiento y clasificación automática de documentos procedentes de conjuntos heterogéneos o en la usabilidad y el diseño de las interfaces de consulta, podríamos encontrarnos con la inutilidad de dichas mejoras a causa de la incapacidad de los usuarios para percibir los diversos tipos de documentos en los resultados. Es decir, todas esas nuevas posibilidades tecnológicas requerirían de unos conocimientos de contexto y función de los documentos, de una experiencia de inmersión en el ecosistema de productos documentales del mundo académico, conocimientos consistentes con la teoría interdisciplinar de los géneros que hemos mencionado anteriormente.

De ahí que concluyamos esta nota apuntando a la necesidad de trabajar el factor humano: creemos en la necesidad inaplazable de potenciar una especie de *searching intelligence* entre profesionales y usuarios avanzados de la información, siguiendo el término *archival intelligence* que algunos autores (Janes, 2012; Yake; Torres, 2003) ponen de relieve para destacar que la complejidad de la investigación en los archivos requiere de un plus de conocimiento que va más allá de la instrucción en técnicas de uso de los instrumentos de descripción y los sistemas de información en los archivos.

Llegados a este punto, podemos observar que en esa línea se posiciona el conocido *Framework* de la ACRL, en el que se define el manejo de la información como

“el conjunto de habilidades integradas que comprenden el descubrimiento reflexivo de la información, el conocimiento acerca de cómo la información es producida y valorada, y el uso de la información en la creación de nuevo conocimiento” (ACRL, 2016).

En el marco de esa definición la educación en la “docudiversidad” que planteamos como reflexión en esta nota se podría entender como un pilar básico, dada la disolución de las formas que los usuarios experimentan cuando se observa el mundo bajo el recuadro de un navegador web en el que se hacen mayoritariamente consultas en buscadores de la familia *Google* (*Google Search*, *Google Scholar* o *YouTube*). Estamos convencidos de que como docentes del área de Información y Documentación deberíamos hacer una lectura del *Framework* pensando en la formación de nuestros estudiantes de grado y en la actualización de nuestros planes docentes, más allá de su aplicación en el diseño de la formación genérica de usuarios de bibliotecas. Se trata de una apuesta fundamental, ya que coincidimos con **Simmons** (2005) en que el camino hacia la alfabetización informacional crítica comporta potenciar el rol del profesional de la información como mediador del discurso en el que se contextualizan los documentos.

Efectivamente, las corrientes más recientes de la alfabetización informacional en el entorno bibliotecario abogan por estrategias más centradas en entender el ecosistema en el que viven los documentos que en las técnicas de interrogación de buscadores y catálogos. En este sentido, resulta especialmente sugerente la propuesta de **Comor y Li** (2012) en el sentido de aprovechar la implantación de las herramientas de descubrimiento bibliotecario para mover el foco desde las técnicas de interrogación hacia lo que denominan “*critical thinking and imaginative exploration*” (“pensamiento crítico y exploración imaginativa”). Esto es, nos atrevemos a decir que, tanto entre los usuarios, como de forma especial entre los profesionales de la información, habría que reforzar los conocimientos frente a las habilidades para que se perciban las formas/géneros como resultado de un proceso, profundizando en dicho proceso. Entender la forma del recurso implica conocerlo, saber que es el resultado de un proceso (**Seeber**, 2015), saber de su existencia, de su naturaleza y alcance y, por ello, conocer su potencial informativo como opción ulterior de búsqueda en profundidad en un agregado de resultados de procedencia diversa. Al tiempo, ese enfoque permite acotar directamente la búsqueda al tipo de recurso concreto que más eficaz y eficientemente resuelva la necesidad de información.

En definitiva, en el terreno de la claridad en la delimitación y reconocimiento de las formas documentales y los atributos formales o funcionales de la información, creemos que no aplica la afirmación que el entonces presidente del gobierno hizo en TV Girona respecto a platos y vasos:

“Es muy claro y lo entiende todo el mundo: un vaso es un vaso y un plato es un plato” (*El país*, 2015).

En el entorno cambiante de la producción documental digital y en red, los formatos de los contenedores y de las piezas documentales que encontramos en ellos son muy diversos, cambiantes y, con frecuencia, responden a una naturaleza híbrida, ya que la “evolución de las especies documentales” se ha acelerado en las últimas décadas. Un plato o un vaso son objetos marcados por un contexto cultural en el que se registran unos determinados hábitos de alimentación, por lo que su distinción se basa en la capacidad de reconocimiento basada en la experiencia de su uso. Por ello siguiendo con el paralelismo de la “alimentación informativa”, en un contexto en el que la alimentación fuera totalmente procedente de uno o pocos proveedores y en formato *feed* casi parenteral, la necesidad de los “platos y vasos informativos” sería vista como remota por los usuarios y la tendencia al olvido sobre su distinción sería inevitable. El reto en nuestro ámbito es mantener la capacidad de apreciación entre los usuarios de los muchos “platos, vasos y objetos de menaje” útiles para la adquisición de información de calidad; esto es, la capacidad de identificar, de reconocer las formas documentales como productos culturales que nos definen.

Corremos el riesgo de que, si la alimentación se acepta casi exclusivamente de forma líquida, al final los platos y los vasos se confundan y sobren. La forma documental, en tanto que es reconocida por el usuario, funciona como una herramienta para orientarse, que tiene una indudable naturaleza cultural. El documento (el recurso de información, si utilizamos la terminología más genérica actualmente en uso) no deja de ser un artefacto cultural, por lo que la discriminación de su utilidad requiere de una iniciación en ese marco social y cultural al que aludía **Andersen** (2008) al aplicar la teoría interdisciplinar de los géneros al ámbito de la información y la documentación.

6. Referencias

ACRL (2016). *Framework for information literacy for higher education*. The Association of College and Research Libraries, a division of the American Library Association.
<https://www.ala.org/acrl/standards/ilframework>

Disponible traducción al español en:

ACRL (2016). *Manejo de referencia para habilidades para el manejo de la información en educación superior*. The Association of College and Research Libraries, a division of the American Library Association. https://www.ala.org/acrl/sites/ala.org/acrl/files/content/standards/Framework_Spanish.pdf

ALA (1993). *Definition of form data*. Approved by ALCTS Board of Directors, Jan. 26. American Library Association. Association for Library Collections and Technical Services. Subject Analysis Committee. <https://sites.pitt.edu/~agtaylor/lala/form-def.htm>

Andersen, Jack (2008). "The concept of genre in information studies". *Annual review of information science and technology*, v. 42, n. 1, pp. 339-367. <https://doi.org/10.1002/aris.2008.1440420115>

Andersen, Jack (2015). "What genre theory does". En: Andersen, Jack; Skouvig, Laura (ed.). *Genre theory in information studies*. Bingley: Emerald Group Publishing. ISBN: 978 1784412548 <https://doi.org/10.1108/S2055-537720140000011002>

Andersen, Jack (2017). "Genre, organized knowledge, and communicative action in digital culture". En: Andersen, Jack (ed.). *The organization of knowledge: caught between global structures and local meaning*. Bingley: Emerald Group Publishing. ISBN: 978 1 78714 531 3 <https://doi.org/10.1108/S2055-537720170000011007>

Area-Moreira, Manuel; Pessoa, Teresa (2012). "De lo sólido a lo líquido: las nuevas alfabetizaciones ante los cambios culturales de la web 2.0". *Comunicar*, v. 19, n. 38, pp. 13-20. <https://doi.org/10.3916/C38-2012-02-01>

Bitter, Colin; Tosaka, Yuji (2020). "Genre/Form access in library catalogs: A survey on the current state of LCGFT usage". *Library resources & technical services*, v. 64, n. 2, pp. 44-60. <https://journals.ala.org/index.php/lrts/article/view/7344> <https://doi.org/10.5860/lrts.64n2.44>

Cmor, Dianne; Li, Xin (2012). "Beyond boolean, towards thinking: discovery systems and information literacy". *Library management*, v. 33, n. 8/9, pp. 450-457. <https://doi.org/10.1108/01435121211279812>

Codina, Lluís; Pedraza-Jiménez, Rafael (2011). "Cerca i recuperació de la informació al web: què ha passat i què podem esperar en el futur". *BiD: textos universitaris de biblioteconomia i documentació*, n. 27. <https://doi.org/10.1344/105.000001783>

Connaway, Lynn-Sillipigni; Dickey, Timothy J.; Radford, Marie L. (2011). "'If it is too inconvenient I'm not going after it:' convenience as a critical factor in information seeking behaviors". *Library & information science research*, v. 33, n. 3, pp. 179-190. <https://doi.org/10.1016/j.lisr.2010.12.002>

Dragon, Patricia M. (2020). "Form and genre access to academic library digital collections". *Journal of library metadata*, v. 20, n. 1, pp. 22-49. <https://doi.org/10.1080/19386389.2020.1723203>

Espelt, Constança; Gascón, Jesús (1999). "La representación y la organización del conocimiento en sus distintas perspectivas: su influencia en la recuperación de la información". En: ISKO-España. Congreso (1999, Granada). *Actas del IV Congreso ISKO-España EOCONSID 99. 22-24 de abril de 1999, Granada*. Granada: Universidad de Granada. pp. 321-329. <https://dialnet.unirioja.es/descarga/articulo/1300531.pdf>

García-Sánchez, Jaquelinne (2013). *Análisis comparativo de los subencabezamientos de forma independientes en lengua española*. Tesis doctoral. Universidad Carlos III de Madrid. <http://hdl.handle.net/10016/17571>

Haider, Jutta; Sundin, Olof (2019). *Invisible search and online search engines: the ubiquity of search in everyday life*. London: Routledge. ISBN: 978 0429448546

Hemmasi, Harriette; Miller, David; Lasater, Mary-Charles (1999). "Access to form data in online catalogs". *ALCTS. Newsletter online*, v. 10, n. 4. <https://sites.pitt.edu/~agtaylor/lala/papers/HemmasiFormData.pdf>

Hillis, Ken; Petit, Michael; Jarrett, Kylie (2013). *Google and the culture of search*. New York: Routledge. ISBN: 978 0203846261 <https://doi.org/10.4324/9780203846261>

Hyer, Mark; Ivy, Helen; Pencek, Bruce; Jordan, Sharon; Tahirkheli, Sharon (2009). "Is good enough, really good enough? does algorithmic metadata search replace the need for discipline-oriented databases?". En: *Proceedings of the Charleston Library Conference*. Purdue: Purdue University. <http://doi.org/10.5703/1288284314744>

- Janes, Andrew** (2012). "Do you have archival intelligence?". *The national archives. Blog*, Oct 15th.
<https://blog.nationalarchives.gov.uk/do-you-have-archival-intelligence>
- Kwaśnik, Barbara H.; Crowston, Kevin** (2005). "Introduction to the special issue: Genres of digital documents". *Information technology & people*, v. 18, n. 2, pp. 76-88.
<https://doi.org/10.1108/09593840510601487>
- Lee, Hur-Li; Zhang, Lei** (2013). "Tracing the conceptions and treatment of genre in Anglo-American cataloging". *Cataloging & classification quarterly*, v. 51, n. 8, pp. 891-912.
<https://doi.org/10.1080/01639374.2013.832457>
- Library of Congress* (1999). *MARC 21 format for bibliographic data*. Washington, DC: Library of Congress, Network Development and MARC Standards Office. Update 32, June 2021.
<https://www.loc.gov/marc/bibliographic/>
- Library of Congress* (2021). *Library of Congress Genre/Form Terms for library and archival materials (LCGFT)* (2021 edition). Washington DC: Library of Congress.
<https://www.loc.gov/aba/publications/FreeLCGFT/freelcgft.html>
- Medeiros, Norm** (2006). "Good enough is good enough: cataloging lessons from the University of California Libraries". *OCLC systems & services: international digital library perspectives*, v. 22, n. 3, p. 155-158.
<https://doi.org/10.1108/10650750610686199>
- Merlo-Vega, José-Antonio** (2012). "Rebelarse contra la relevancia: cómo encontrar lo que buscamos y no lo que Google cree que buscamos". *Anuario ThinkEPI*, v. 6, pp. 293-296.
<https://thinkepi.profesionaldelainformacion.com/index.php/ThinkEPI/article/view/30440>
- Miller, Carolyn R.** (1984). "Genre as social action". *Quarterly journal of speech*, v. 70, n. 2, pp. 151-167.
<https://doi.org/10.1080/00335638409383686>
- Montesi, Michela** (2008). "Géneros web: líneas de investigación". *El profesional de la información*, v. 17, n. 5, pp. 551-557.
<https://doi.org/10.3145/epi.2008.sep.10>
- Orlikowski, Wanda J.; Yates, JoAnne** (1994). "Genre repertoire: the structuring of communicative practices in organizations". *Administrative science quarterly*, v. 39, n. 4, pp. 541-574.
<https://doi.org/10.2307/2393771>
- El país* (2015). "Rajoy: 'Está muy claro: un vaso es un vaso y un plato es un plato'". *El país*, 24 septiembre.
https://elpais.com/politica/2015/09/24/actualidad/1443082572_575030.html
- Richards, Maureen** (2021). "Is 'just googling it' good enough for first-year students?". *College & undergraduate libraries*.
<https://doi.org/10.1080/10691316.2021.1894295>
- Rodríguez-Yunta, Luis** (2015). "Servicios de descubrimiento basados en un índice centralizado: su expansión en las bibliotecas académicas españolas y futuras líneas de investigación". *Anuario ThinkEPI*, v. 9, pp. 49-55.
<https://doi.org/10.3145/thinkepi.2015.09>
- Seeber, Kevin-Patrick** (2015). "Teaching 'format as a process' in an era of web-scale discovery". *Reference services review*, v. 43, n. 1, pp. 19-30.
<https://doi.org/10.1108/RSR-07-2014-0023>
- Simmons, Michelle-Holschuh** (2005). "Librarians as disciplinary discourse mediators: using genre theory to move toward critical information literacy". *Portal: Libraries and the academy*, v. 5, n. 3, pp. 297-311.
<https://doi.org/10.1353/pla.2005.0041>
- Taylor, Arlene G.** (1992). *Definition of "Form", "Genre", and "Physical Characteristics"*. Paper written for the purpose of discussion by the SAC Subcommittee on Form Data, January 1992.
<https://sites.pitt.edu/~agtaylor/lala/form-mss.htm>
- Urbano, Cristóbal** (2018). "Sobre la autodeterminación del usuario de información en tiempos de acceso social a los contenidos digitales". *Anuario ThinkEPI*, v. 12, pp. 236-241.
<https://doi.org/10.3145/thinkepi.2018.35>
- Vaidhyanathan, Siva** (2011). *The Googlization of everything*. Updated ed. Berkeley: University of California Press. ISBN: 978 0520948693
<https://doi.org/10.1525/9780520948693>
- Wilson, Patrick; Robinson, Nick** (1990). "Form subdivisions and genre". *Library resources & technical services*, v. 34, n. 1, pp. 36-43.

Yakel, Elizabeth; Torres, Deborah (2003). "AI: archival intelligence and user expertise". *The American archivist*, v. 66, n. 1, pp. 51–78.
<https://doi.org/10.17723/aarc.66.1.q022h85pn51n5800>

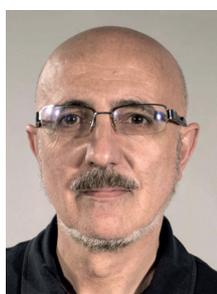
Cristóbal Urbano
Universitat de Barcelona
urbano@ub.edu

Jesús Gascón García
Universitat de Barcelona
gascon@ub.edu

* * *

La "docudiversidad" en la curación de contenidos

Javier Guallar



La nota de Urbano y Gascón recupera y pone sobre la mesa de debate una perspectiva que, si bien no es nueva, quizás no haya recibido en los últimos años la atención que merece: la importancia de la "docudiversidad" o diversidad documental en un proceso de recuperación de información. Sin entrar en las cuestiones terminológicas que se presentan en dicho trabajo, seguiré en un sentido amplio los términos que emplea el propio Urbano en un artículo anterior (**Urbano**, 2016), de "continente" y "contenido", para sumarme a la visión de que se trataría de considerar el continente y no solamente el contenido, como un factor de calidad en cualquier proceso de selección o recuperación de información.

Las reflexiones que siguen se sitúan estrictamente desde el punto de vista de la curación de contenidos. Desde esta perspectiva el debate es igualmente pertinente en tanto que se entiende la curación como un proceso de selección de información con aportación de valor dirigida a un público determinado. En el artículo previo de Urbano ya mencionado, el autor lanzaba una primera reflexión en esta línea cuando se preguntaba:

"¿Se puede aplicar el concepto *curation* al continente de igual forma que se aplica al contenido?" (**Urbano**, 2016).

Y es que en todos (o en la mayoría) de planteamientos sobre *content curation*, hasta donde creo recordar, el foco o énfasis se ha puesto principalmente en el contenido y se han tenido mucho menos en cuenta, los diversos tipos de continentes (por ejemplo, en un repaso no exhaustivo: **Bhargava**, 2009; **Good**, 2010; **Guallar**; **Leiva-Aguilera**, 2013; **Deshpande**, 2015).

Un ejemplo de ello es que cuando se considera como criterio de calidad de una curación (tanto en las propuestas de cómo realizarla como en las de cómo analizarla) la diversidad de las fuentes documentales empleadas como punto de partida para hacer la selección de contenido, no se ahonda suficiente en la diversidad de las formas documentales de dichas fuentes. Así, en la habitual representación gráfica de un proceso de curación de información mediante un embudo, se suele enfatizar que la parte superior de dicho embudo debería ser lo más amplia posible (alta diversidad de fuentes documentales), pero con esta idea nos referimos casi siempre a la diversidad de contenidos, no a la diversidad de continentes. La inclusión por tanto de la "forma documental" como un elemento sustancial a considerar en la calidad de una curación es pertinente dado que es fácilmente constatable que no ha sido suficientemente puesta en valor.

El camino a seguir, a partir de las reflexiones actuales de Urbano y Gascón, y las previas de **Urbano** (2016), desde el punto de vista de la curación, pasaría a mi entender por profundizar más, al menos, en dos tipos de acciones:

- Explicitar e incorporar el papel de las diversas formas documentales en la ejecución de un proceso de curación, como elemento a tener en cuenta en cualquiera de los métodos de curación existentes, por ejemplo, en las 4S's (**Guallar**; **Leiva-Aguilera**, 2013). Aquí hay sin duda terreno por recorrer y por mi parte queda recogido el guante de Urbano y Gascón, para tenerlo en cuenta en nuevas propuestas.
- Incorporar el estudio de las formas documentales de un proceso de curación dentro de los sistemas de análisis y evaluación de curación de contenidos. En esta línea, en mi último trabajo junto a los investigadores Pedraza-Jiménez, Pérez-Montoro y Anton, (**Guallar et al.**, 2021), sí que empezamos a

introducir esta perspectiva como uno de los parámetros de calidad de una curación: es lo que denominamos "Fuentes de información atendiendo a su morfología" (p. e., páginas web, blogs, redes sociales, fuentes secundarias...), en el que consideramos que a una mayor diversidad de estas formas documentales en un producto de curación, mayor puntuación obtendrá el mismo en el sistema de evaluación.

En curación de contenidos, por tanto, empezamos a considerar más el continente, y no solo el contenido.

Referencias

Bhargava, Rohit (2009). "Manifiesto for the content curator: the next big social media job of the future?". *Rohit-Bhargava.com*, September 30.
<https://www.rohitbhargava.com/2009/09/manifiesto-for-the-content-curator-the-next-big-social-media-job-of-the-future.html>

Deshpande, Pawan (2015). "The definitive guide to content curation". *The curata blog*, January 15.
<http://www.curata.com/blog/the-definitive-guide-to-content-curation/>

Good, Robin (2010). "Real-time news curation - The complete guide. Part 4: Process, key tasks, workflow". *Master newmedia*, September 29.
<http://www.masternewmedia.org/real-time-news-curation-the-complete-guide-part-4-process-key-tasks-workflow/>

Guallar, Javier; Leiva-Aguilera, Javier (2013). *El content curator. Guía básica para el nuevo profesional de internet*. Barcelona: Editorial UOC. ISBN: 978 84 9064 018 0

Guallar, Javier; Pedraza-Jiménez, Rafael; Pérez-Montoro, Mario; Anton, Laura (2021). "Curación de contenidos en periodismo. Indicadores y buenas prácticas". *Revista española de documentación científica*, v. 44, n. 2, e296.
<https://doi.org/10.3989/redc.2021.2.1742>

Urbano, Cristobal (2016). "Control y evaluación de recursos de información en clave de "continent" curation". *Anuario ThinkEPI*, v. 10, pp. 137-141.
<https://doi.org/10.3145/thinkepi.2016.29>

Javier Guallar
Universidad de Barcelona
jguallar@ub.edu

* * *

Los productos de la curación de contenidos pueden analizarse desde características formales

Cristóbal Urbano

Me parecen muy interesantes tus reflexiones. Abundando en ellas, quisiera añadir que, como bien explicas en uno de tus últimos trabajos (**Guallar et al.**, 2021), los productos resultantes de la curación de contenidos son diversos y pueden analizarse según características formales que pueden ser indicadores de situaciones en las que claramente se dan funciones comunicativas y documentales diversas con el fin de alcanzar el propósito perseguido. En este sentido, quiero destacar que la propuesta de dimensiones e indicadores formales que en dicho artículo habéis denominado CAS (*Curation Analysis System*) constituye una base muy interesante para un posible estudio de los diversos productos resultantes de la curación de contenidos desde el punto de vista de los géneros.

En efecto, dado que los productos resultantes de la curación de contenidos mayoritariamente circulan vía web se podrían estudiar aprovechando el marco de una tradición de trabajos dedicados a dicho medio (**Askehave; Nielsen**, 2005; **Montesi**, 2008). Creo que mercería la pena esa aproximación específica para todo lo que para vosotros encaja bajo la denominación "curación de contenidos", tomando como referentes estudios dedicados específicamente a situaciones en el web, pero también a tipos muy concretos de documentos pre-web, como por ejemplo el artículo científico (**Bazaerman**, 1988) o las comunicaciones internas dentro de una organización (**Orlikowski; Yates**, 1994). También se podrían contemplar ejemplos de trabajos recientes dedicados al análisis de las interfaces de consulta y las pantallas de resultados en buscadores web (**Andersen**, 2017) o a las interacciones de los usuarios en entornos web (**Andersen**, 2021).

Tanto los contenidos que se seleccionan/listan/comentan/citan en los productos resultantes de un proceso de curación de contenidos, como esos mismos productos, se pueden estudiar como fenómeno comunicativo con su contexto y propósito en el marco de unos determinados actores implicados. Esto es, como decía **Miller** (1984), los géneros se han de entender como acciones sociales, tipificadas desde el punto de vista retórico y basadas en situaciones recurrentes, de ahí que se podría complementar vuestro CAS (*Curation Analysis System*) en clave de otras dimensiones de naturaleza social y cultural junto a las formales de vuestra propuesta, en tanto que, como dicen **Orlikowski** y **Yates** (1994), un género es un tipo de acción comunicativa diferenciado que se caracteriza por un propósito de comunicación socialmente reconocido y aspectos comunes de forma.

En fin, toda esta línea de reflexión sobre los géneros en relación con la curación de contenidos creo que puede ayudar a que desde vuestra "escuela" de curación de contenidos afloren y se resignifiquen fragmentos como el siguiente, que a mi entender constituyen una declaración *avant la lettre* de los principios de la curación de contenidos que están presentes en tu respuesta a nuestra nota *ThinkEPI*:

"... Nos enfrentamos a opciones que pueden homogeneizar la información o mantener la visibilidad de los documentos humanos (ya sean materiales o virtuales), las situaciones de las que surgieron y las situaciones en las que se utilizarán una vez que se acceda a ellos. Por supuesto, se obtienen grandes beneficios mediante la búsqueda rápida de amplios recursos electrónicos, y podemos aprender mucho de las piezas atomizadas reunidas en un solo espacio, pero para muchos otros propósitos necesitamos comprender la procedencia, el género, el contexto de la actividad y las estructuras sociales e institucionales, de donde surge la información y en la que está destinada a ser utilizada. A medida que desarrollamos herramientas, sistemas y conceptos para unir piezas más heterogéneas de circunstancias más heterogéneas, también necesitamos desarrollar herramientas, sistemas y conceptos para ver la información en sus circunstancias particulares de uso". (**Bazerman**, 2012)

Referencias

- Andersen, Jack** (2017). "Genre, organized knowledge, and communicative action in digital culture". En: Andersen, Jack (ed.). *The organization of knowledge: caught between global structures and local meaning*. Bingley: Emerald Group Publishing. ISBN: 978 1 78714 531 3
<https://doi.org/10.1108/S2055-537720170000011007>
- Andersen, Jack** (2021). "Genre as digital social action: the case of archiving, tagging and searching in digital media culture". *Journal of documentation*, ahead-of-print.
<https://doi.org/10.1108/JD-01-2021-0023>
- Askehave, Inger; Nielsen, Anne-Ellerup** (2005). "Digital genres: a challenge to traditional genre theory". *Information technology & people*, v. 18, n. 2, pp. 120-141.
<https://doi.org/10.1108/09593840510601504>
- Bazerman, Charles** (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison: University of Wisconsin Press
- Bazerman, Charles** (2012). "The orders of documents, the orders of activity, and the orders of information". *Archival science*, v. 12, pp. 377-388.
<https://doi.org/10.1007/s10502-012-9178-1>
- Guallar, Javier; Pedraza-Jiménez, Rafael; Pérez-Montoro, Mario; Anton, Laura** (2021). "Curación de contenidos en periodismo. Indicadores y buenas prácticas". *Revista española de documentación científica*, v. 44, n. 2, e296.
<https://doi.org/10.3989/redc.2021.2.1742>
- Montesi, Michela** (2008). "Géneros web: líneas de investigación". *El profesional de la información*, v. 17, n. 5, pp. 551-557.
<https://doi.org/10.3145/epi.2008.sep.10>
- Orlikowski, Wanda J.; Yates, JoAnne** (1994). "Genre repertoire: the structuring of communicative practices in organizations". *Administrative science quarterly*, v. 39, n. 4, pp. 541-574.
<https://doi.org/10.2307/2393771>

Cristóbal Urbano
Universitat de Barcelona
urbano@ub.edu

* * *

¿Dónde catalogar y dónde aportan valor las catalogaciones al distribuirlas? Tomás Saorín



Un excelente trabajo que deberíamos leer dos o tres veces, y de paso replantearnos algunas cosas que damos por sentadas.

¿De verdad sigue algún bibliotecario catalogando recursos en un lugar donde alguien lo busca?

En caso de ser así, ¿está catalogando las propiedades que harán que el usuario entienda los resultados de búsqueda?

¿Captan nuestros sistemas la realidad transmedia no solo de las franquicias de acción sino también las de la publicación científica transmedia?

Algún trabajo reciente sugiere que en plataformas de libro electrónico simplemente se cargan los metadatos que vienen de la editorial y para adelante (**Gil-Leiva et al.**, 2020), y por algún otro lado se sugiere que hay que catalogar (enriquecer con metadatos que expandan las posibilidades de identificación y descubrimiento) allí donde el usuario buscará, y no en nuestro estrecho e invisible catálogo (**Scott**, 2017).

Me pregunto también si tal vez cada vez que hablamos de *information literacy* no da un poco la sensación de que cargamos sobre el usuario la responsabilidad de hacer bien la búsqueda, seleccionar, discriminar... y dejamos un poco de lado lo poco que ayudan nuestros entornos de descubrimiento a hacer natural la tarea. Los *featured snippets* de *Google* van transformando cada vez más la búsqueda en una recomendación o, en realidad, en una conversación que ofrece resultados empaquetados listos para su consumo, porque entienden las "formas documentales" y se preocupan de presentar cada cosa de forma que pueda ser entendida y usada en su contexto. Claro, lo hacen con el propósito quizá banal de darnos la previsión del tiempo, los próximos estrenos y vender zapatillas, pero refleja un esfuerzo en el que una experiencia de usuario persuasiva y detallista, haga fácil las cosas en el medio digital. Los autores usan la analogía con la alimentación y me evocó al principio el gag de *Tiempos modernos* y su comida-matic para obreros, pero al final se impuso el escenario pasivo y gomoso de *Wall-e*. ¿En qué condiciones hay espacio para gastrobares de información frente a las cadenas de comida rápida? ¿Por qué los bares de carretera son así en internet?

Me gustaría plantear a la reflexión dos ideas: una sobre dónde catalogar y otra sobre dónde aportan valor las catalogaciones al distribuirlas.

La primera sería catalogar (describir) en *Wikidata* (o plataformas por el estilo), en un *dataset* colaborativo y abierto, que pueda alimentar la cadena de descubrimiento de información allá donde suceda, en *Google*, *Amazon*, *Netflix*, *Hispana*, etc. Dejar los metadatos, pulir los metadatos colaborativamente, incluso reflejando diferentes puntos de vista, en entornos abiertos donde puedan enriquecer cualquier catálogo-buscador-fuente. Constatemos que *IMDB* tiene mayor expresividad y compromiso con los usuarios digitales (léase reales) que el catálogo de la *BNE*. Los metadatos de *IMDB* se integran en *Prime Video* para potenciar experiencias digitales. Y convertir los entornos de *Web Discovery* en *Netflix*. Habrá que montar una buena cadena de comida rápida. O distribuir información-metadatos de calidad a quien se atreva a montar su propio bar y conozca a sus clientes, y entienda cómo y dónde comunicarse con ellos.

Y me gusta citar esta idea tan básica de Dan Scott:

"En lugar de centrarse en mejorar directamente nuestros propios silos de repositorios de datos locales (por ejemplo, catálogos de bibliotecas, exposiciones digitales), las bibliotecas y los archivos deberían invertir sus recursos limitados en enriquecer *Wikidata*, un repositorio de datos centralizado, para maximizar la visibilidad de esas entidades y el reutilización de esos datos en el mundo en general... y luego recuperar esos datos en nuestros repositorios locales para enriquecer nuestras pantallas y la integración con el mundo más amplio de los datos" (**Scott**, 2017)

Por último, si hay metadatos útiles, podrán ser útiles en la plataforma de distribución, mediante creatividad y esfuerzo en "el lado del usuario". Un éxito de público como el préstamo de libros electrónicos *eBiblio*, contiene metadatos y posibilidades de descubrimiento banales, muy limitadas. ¿Se debe el éxito meramente a que da acceso gratuito a libro electrónico y el usuario, en el contexto del top-manta legal se conforma con lo que hay, que menos es nada? Porque, bajo mi punto de vista, las posibilidades de descubrimiento de obras y contenidos son muy muy limitadas. Algunas bibliotecas se esfuerzan en compensar con selección editorial en los carruseles de la *home* (mayor disponibilidad para préstamos, autores locales, etc.) pero es poco sugerente tratar de buscar o navegar por el fondo. Poca catalogación, poco enriquecimiento bibliotecario.

Finalizo, ya sí, volviendo a recomendar el estupendo libro de Ana Vukadin, *Metadata for transmedia resources* (**Vukadin**, 2019) y la idea de que el catálogo debería transformarse en una herramienta de

mediación, “un mecanismo de mapeo o interoperabilidad, entre descripciones distribuidas”, un actor fiable y de valor en las cadenas que nos alimentan de información digital.

Referencias

Gil-Leiva, Isidoro; Lopes-Fujita, Mariângela-Spotti; Díaz-Ortuño, Pedro-Manuel; Akama-dos-Reis, Daniela-Majorie (2020). “The abandonment of the assignment of subject headings and classification codes in university libraries due to the massive emergence of electronic books”. *Knowledge organization*, v. 47, n. 8. <http://eprints.rclis.org/4098111/Gil-LeivaFujitaDiazMajorie2020.pdf>

Vukadin, Ana (2019). *Metadata for transmedia resources*. Chandos Publishing. ISBN: 9780081013007

Scott, Dan (2017). “Wikidata, Canada 150, and music festival data”. *Dan Scott*, June 2017. <https://coffeecode.net/wikidata-canada-150-and-music-festival-data.html>

Tomás Saorín
Universidad de Murcia
tsp@um.es

* * *

Reflexiones y retos frente a la cruda realidad Cristóbal Urbano

Tiene mucha intensidad y valor lo que dices: abres muchas preguntas nuevas y llamas al debate, que justo es lo que se persigue por parte de los editores del *Anuario ThinkEPI* con la dinámica de publicar las notas en la lista *IweTel*.

Así que mi respuesta será breve: me has hecho pensar (creo que a todos), pero debería meditar las respuestas y comprobar bibliografía que las pueda apoyar. Creo que hay que invitar a la audiencia a volver a tus preguntas, pues quizás sea bueno esperar a ver la opinión de otras personas suscritas a la lista que están más implicadas que yo en las decisiones sobre procesos técnicos que tú planteas; quizás puedan responderte tomando en consideración la cruda realidad en materia de gestión de servicios, recursos humanos, costes de oportunidad, relación coste-beneficio, estadísticas de uso y de usuarios, etc., que pueden contextualizar lo que se está haciendo (aunque quizás no justificarlo abiertamente), todo ello en el marco de aceptación del *good enough* de una forma pragmática.

Estaría bien saber de las experiencias, penas, trabajos y alegrías de quienes están lidiando en bibliotecas con la catalogación de recursos electrónicos, catalogación por copia, descarga masiva de registros, adquisiciones desencadenadas por los usuarios (*PDA: Patron Driven Acquisitions*) y por último, pero no menos importante, con la externalización de la catalogación de recursos electrónicos fuera del propio catálogo mediante la configuración de las herramientas de descubrimiento bibliotecario (software tipo *discovery tool*) que agregan metadatos suministrados por editores primarios y secundarios, mediante acuerdos con la empresa proveedora del software de la solución de descubrimiento. Todo bastante complejo y de resultados inciertos, pero que se hace al parecer con la intención de evitar las barreras que genera el efecto silo de los diversos actores del control bibliográfico clásico. El problema es que en ciertas situaciones el efecto silo es una bendición, siempre que uno sepa lo que busca y dónde buscarlo.

En definitiva, lo que planteas mercería una nota *ThinkEPI* sobre el rol que deberían jugar los catálogos de bibliotecas y otros instrumentos de control bibliográfico clásicos (como las bases de datos de indexación y resumen) en un entorno en el que el descubrimiento bibliográfico se ha desplazado de forma muy notoria hacia buscadores web generales o académicos [aquí estaría de más usar el término genérico “buscador”, casi podríamos decir directamente *Google* y *Google Scholar*, pese a los cantos de sirena que llegan de nuevos actores como *Dimensions* o *Lens.org*] y, en menor medida, hacia las *discovery tools* que ofrecen algunas bibliotecas, o hacia portales agregadores de contenidos procedentes de repositorios digitales académicos.

Respecto al tema de la alfabetización informacional creo que nuestra nota no quería “cargar la responsabilidad sobre los usuarios”. Justamente hemos tratado de subrayar la importancia de pasar del “entrenamiento” en las técnicas de búsqueda a la adquisición de conocimiento sobre la realidad de la producción documental y los procesos sociales subyacentes. No se trata en primera instancia de que los usuarios sepan plantear las claves de búsqueda relacionadas con el género/forma como algo instrumental para superar los defectos de los sistemas de recuperación de información, sino empoderarlos en sus conocimientos sobre la existencia de una diversidad documental que es fruto de procesos sociales.

Con enfoque literario y filosófico, el escritor y ensayista **George Perec** (1985) destaca en la pieza "Los esquimales", de la compilación *Pensar-clasificar*, que para determinadas sociedades el conocimiento de su entorno es muy importante para su vida cotidiana y eso se refleja en su capacidad para reconocer la diversidad de determinados objetos o realidades y en su riqueza lingüística para nombrarlos:

"Afirmar que los esquimales no tienen nombre genérico para designar el hielo; tienen varias palabras (he olvidado el número exacto, pero creo que son muchas, alrededor de una docena) que designan específicamente los diversos aspectos que cobra el agua entre su estado líquido y sus diversos grados de congelación. Es difícil, por cierto, encontrar un ejemplo equivalente en francés; tal vez los esquimales solo tengan una palabra para designar el espacio que separa los iglús, mientras nosotros tenemos por lo menos siete en nuestras ciudades [...] y los ingleses por lo menos veinte..."

Nuestra apelación a la alfabetización informacional tenía que ver con la necesidad de que las formas documentales sean para los estudiantes universitarios algo así como las clases de hielo para los esquimales según Perec. Ahora bien, el reto es importante, ya que en nuestra nota también se aludía al reto que en el presente tenemos de denominar y representar consensuadamente la realidad documental, tanto por la dinámica cambiante de dicha realidad como por la obsolescencia no programada de algunas de nuestras herramientas, como clasificaciones, tesauros, vocabularios, etc. Por ello, hojeando el librito de Perec, uno no puede dejar de sonreír y al tiempo identificarse con la pieza titulada "Cómo clasifico":

"Mi problema con las clasificaciones es que no son duraderas; apenas pongo orden, dicho orden caduca. Como todo el mundo, supongo, tengo a veces un frenesí del ordenamiento; la abundancia de cosas para ordenar, la casi imposibilidad de distribuirlas según criterios verdaderamente satisfactorios, hacen que a veces no termine nunca, que me conforme con ordenamientos provisorios y precarios, apenas más eficaces que la anarquía inicial".

Referencias

Perec, George (1985). *Pensar-clasificar*. Barcelona: Editorial Gedisa, 2017. Traducción del original de 1985.

Cristóbal Urbano
Universitat de Barcelona
urbano@ub.edu

Dialnet | métricas
Nueva edición 2020

Índice Dialnet de Revistas e Investigadores

Revistas Investigadores

Título de la revista

IDR

Indicadores Dialnet

Fundación Dialnet

dialnet.unirioja.es/metricas