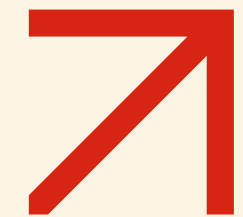




La búsqueda académica en la era de la IA: resultados de una encuesta internacional



Núria Bautista-Puig
Instituto de Filosofía (IFS), CSIC

Scholarly Information Search Habits in the Age of AI

 Enrique Orduña-Malea^{1*},  Núria Bautista-Puig²,  Eva Méndez³,  Carmen Pérez-Esparrells⁴ and  Rodrigo Costas⁵

¹ The iMetrics Lab. Department of Audiovisual Communication, Documentation and History of Art, Universitat Politècnica de València, Valencia, Spain

² Institute of Philosophy, Spanish National Research Council (CSIC), Spain

³ OpenScienceLab. Department of Library and Information Science, Universidad Carlos III de Madrid, Spain

⁴ Department of Economics and Public Finance, Autonomous University of Madrid, Spain; Research Institute for Higher Education and Science (INAEU), Spain.

⁵ Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands; DSI-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy, Stellenbosch University, South Africa

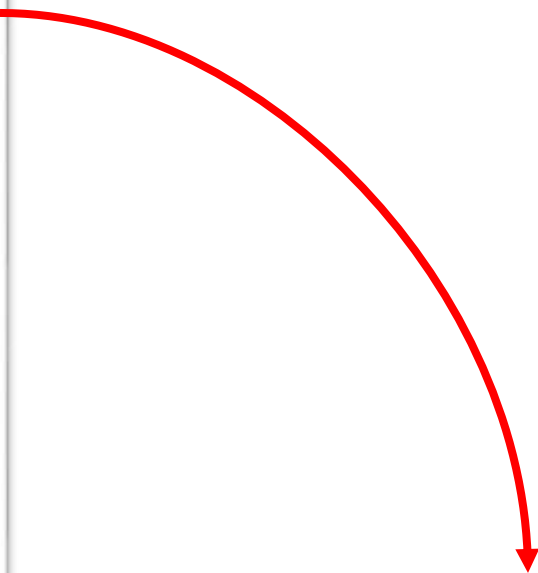
* Corresponding author: enorma@upv.es

Abstract

Scholarly information seeking has been reshaped by the growth of online scholarly content, the consolidation of academic search engines, and the proliferation of AI-based tools. Yet, evidence on how researchers discover, select, and evaluate scholarly information through online academic search remains fragmented. A survey was conducted to 3,349 worldwide researchers across disciplines, career stages, and organization types. The questionnaire is composed of four sections covering: general academic search habits, use of search tools, practices related to academic search engines; and sociodemographic variables. The results show that academic search engines dominate the initial stage of searching (76.9% often/very often), complemented by general search engines (62.7%) and publisher/journal websites (59.4%). Since 2020, 65.0% report changes in habits, most often attributed to increased AI use (51.7%) and adoption of other new tools (42.9%). Natural-language querying predominates (67.9%), while advanced commands are used by a minority (29.1%). Clicking decisions are driven primarily by full-text availability (74.1%) and terms related to the topic of searching in the title or description, whereas citation counts play a weaker role (43.1%). Findings indicate that the academic information seeking ecosystem is largely driven by search systems and access-oriented decision-making.

Keywords:

Scholarly search; academic search engines; online search habits; Google Scholar



Paper en evaluación
en JASIST

Introducción: La búsqueda académica en la era digital

La **Web** ha transformado la forma en que los investigadores localizan y acceden a la información científica (Brown, 1999; Lawrence & Giles, 1999).

Los **motores de búsqueda** han sustituido progresivamente a los sistemas tradicionales de descubrimiento proporcionados por las bibliotecas (Stelmaszewska et al., 2010).

Actualmente constituyen la principal **puerta de entrada** a la literatura científica (Haines et al., 2010; Gusenbauer & Haddaway, 2020).

Google Scholar se ha consolidado como uno de los principales intermediarios en el acceso al conocimiento científico (Du & Evans, 2011).



Introducción: Los motores de búsqueda no son neutrales

La cobertura, indexación y algoritmos de ranking determinan **qué documentos son visibles** (Lawrence & Giles, 1999).

Los algoritmos de búsqueda son en gran medida opacos para los usuarios (van Dijck, 2010; Gusenbauer & Haddaway, 2021).

En Google Scholar, las **citas** influyen significativamente en la posición de los resultados (Butler, 2004; Martín-Martín et al., 2017; Rovira et al., 2019).

Esto puede afectar qué trabajos se descubren, leen y citan (Evans, 2008; Drivas, 2025).



Introducción: ¿Qué sabemos sobre cómo buscan los investigadores?

La investigación previa se ha centrado principalmente en:

- El uso de motores de búsqueda académicos varía según disciplina, etapa profesional y otras características individuales (por ej. Jamali & Nicholas, 2008; Niu & Hemminger, 2012; Hsin et al., 2016).
- La mayoría de los estudios se han realizado en contextos institucionales o disciplinares específicos (Brindesi et al., 2013), incluso incluyendo a estudiantes (Hsin et al., 2016).
- Existe poca evidencia internacional y multidisciplinar centrada específicamente en motores de búsqueda académicos y, en muchos casos, anterior a la consolidación de plataformas como Google Scholar.

Brecha de investigación

Aún existe escasa evidencia internacional y multidisciplinar sobre cómo utilizan los investigadores los motores de búsqueda académicos y sobre el papel que desempeñan los factores sociodemográficos en estos comportamientos.

Objetivos

Analizar los **hábitos contemporáneos de búsqueda académica** online mediante una encuesta internacional a investigadores de distintas disciplinas y etapas de la carrera científica.

Objetivos específicos

-  Describir los hábitos de búsqueda de información científica.
-  Identificar las herramientas más utilizadas para localizar literatura académica.
-  Caracterizar los patrones de uso de los motores de búsqueda académicos.
-  Analizar si los hábitos de búsqueda varían según factores sociodemográficos.

Métodos: diseño de la encuesta

02

Se diseñó un **encuesta estructurada** (en español y en inglés) con cuatro bloques temáticos:

- a) Hábitos generales de búsqueda (frecuencia de búsqueda, tipo de keywords usados, uso de operadores booleanos, etc.).
- b) Herramientas de búsqueda (motores de búsqueda, preferencias, herramientas de IA)
- c) Motores de búsqueda académicos (uso de Google Scholar, estrategias de búsqueda, criterios de evaluación)
- d) Variables sociodemográficas (disciplina, etapa de la carrera investigadora, tipo de organización, país)

Filtro pregunta #6

Validación mediante un pretest con 50 investigadores.

Aprobación **ética** concedida por el Comité de Ética de la Universitat Politècnica de València (Ref. P01_26-03-2025).

Métodos: distribución de la encuesta

02

- Base de contactos construida a partir de **ORCID** y **OpenAlex** (CWTS).
- Criterio: inclusión de investigadores con al menos **una publicación indexada**.
- Difusión mediante correos electrónicos, listas de distribución académicas y redes sociales.
- Periodo de recogida de datos: mayo–julio de 2025.
- Plataforma: Alchemer.

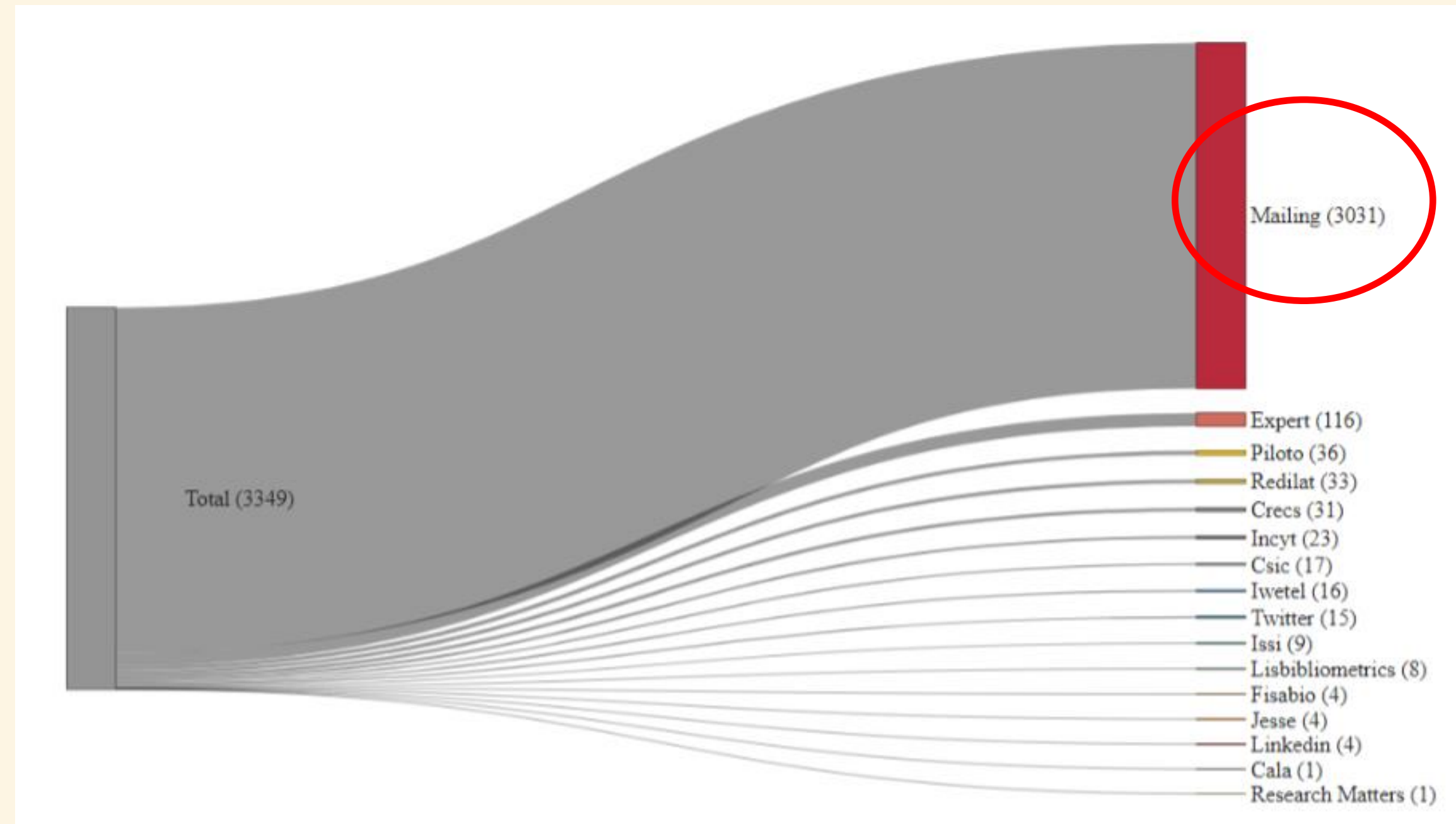


Figura 1. Diagrama Sankey de la distribución de la encuesta

Métodos: datos descriptivos

02

La muestra final (n=3.349) estuvo compuesta principalmente por investigadores afiliados a universidades (73,2%), hombres (63,5%) y con experiencia investigadora intermedia o avanzada. Las áreas más representadas fueron las Ciencias Sociales (28,2%) y las Ciencias Naturales (23,9%).

TABLE 1 Descriptive statistics of the survey sample (n=3349).

Dimension	Variable	N	%
Gender	Male	2,125	63.5
	Female	1,143	34.1
	Other	60	1.8
	Non-binary	21	0.6
Stage of research career	R3: Independent	1,227	36.6
	R2: Postdoctoral stage	1,004	30
	R1: First stage	763	22.8
	R4: Leader	355	10.6
Type of organization	University	2,452	73.2
	Public research organization	326	9.7
	Private company	146	4.4
	Health institution	132	3.9
	Government	87	2.6
	Non-university educational institution	82	2.4
	Non-profit organization	73	2.2
	Other (please specify)	51	1.5
Field of research	Social Sciences (Law, Economics, Sociology, Psychology, Communication, etc.)	945	28.2
	Natural Sciences (Biology, Physics, Chemistry, etc.)	799	23.9
	Applied Sciences (Engineering, Computer Science, etc.)	643	19.2
	Medicine and Health Sciences	526	15.7
	Humanities (Philosophy, History, Languages, etc.)	324	9.7
	Formal Sciences (Mathematics, etc.)	84	2.5
	Fine Arts (Music, Painting, Sculpture, etc.)	28	0.8
Years of experience	More than 15 years	1,171	35
	8–15 years	1,067	31.9
	4–7 years	882	26.3
	1–3 years	217	6.5
	Less than 1 year	12	0.4

Métodos: análisis de datos

02

- Limpieza y estandarización de datos.
- Estadísticos descriptivos.
- Análisis de asociaciones entre hábitos de búsqueda y variables sociodemográficas:
 - χ^2 y V de Cramér
 - Correlación de Spearman (ρ)
- Análisis en R (v.4.4.1).



Resultados: Hábitos generales de búsqueda (I)

70,8% dedica entre 1 y 10 horas semanales a buscar información científica.

40% dedica más tiempo a esta actividad que en 2020, posiblemente debido a un ecosistema de búsqueda más complejo (por ej, + literatura científica y + fuentes).

Solo el 23,4% reporta cambios sustanciales en sus hábitos de búsqueda.

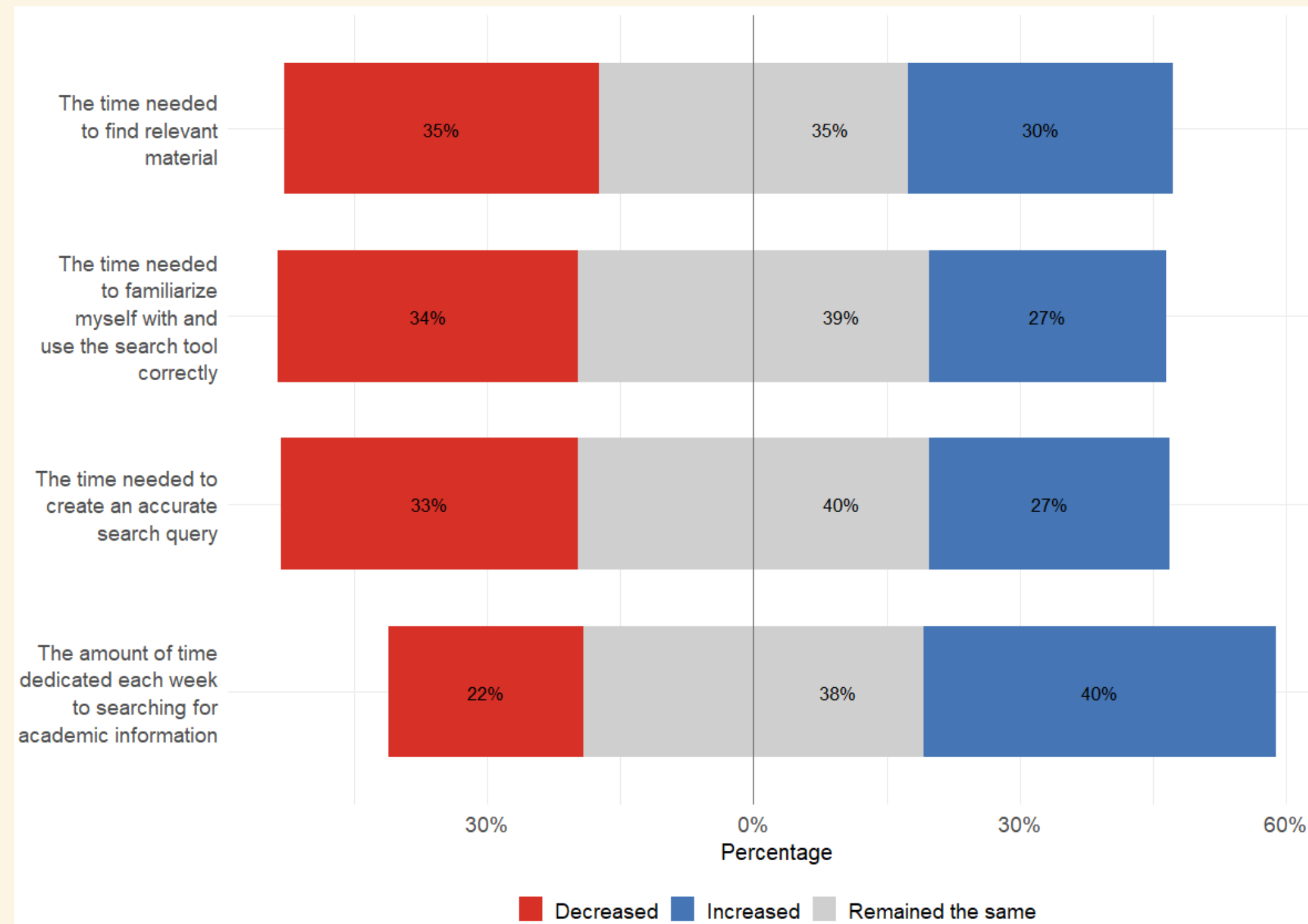


Figura 2. Cambios en el tiempo dedicado a actividades relacionadas con la búsqueda de información desde 2020.

Resultados: Hábitos generales de búsqueda (II)

¿Qué ha impulsado los cambios?

- Uso de herramientas de IA (51,7%)
- Nuevas herramientas de búsqueda (42,9%)
- Recomendaciones de colegas (24,6%)

¿Para qué buscan información?

- Mantenerse actualizado (78,5%)
 - Revisiones bibliográficas/sistemáticas (72,2%)
 - Propuestas de investigación (67,0%)
- ...La búsqueda académica responde a necesidades continuas de actualización/revisión del conocimiento.

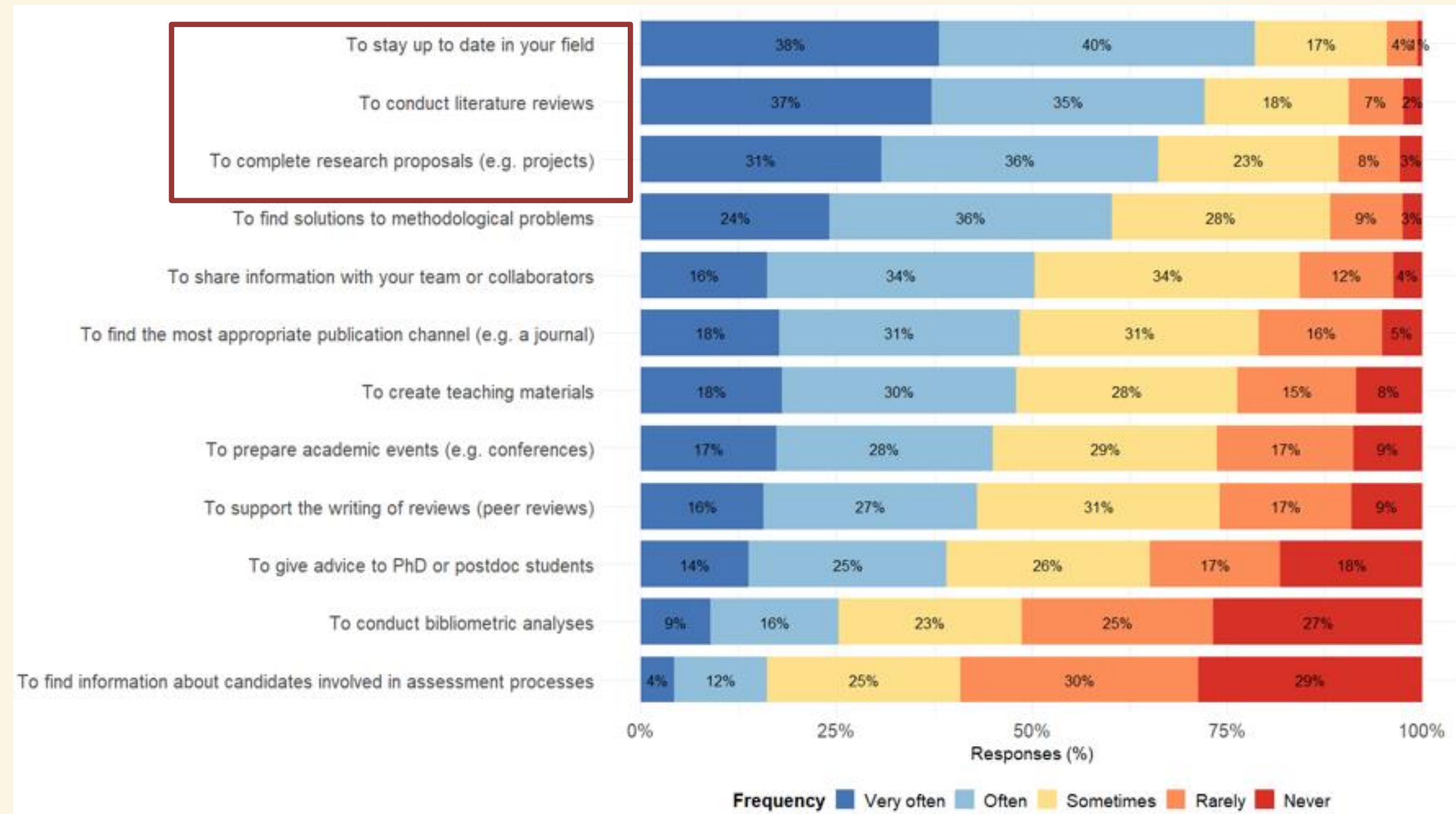


Figura 3. Frecuencia de búsqueda de información científica según el propósito de la búsqueda.

Resultados: Herramientas de búsqueda

02

Principales herramientas de búsqueda

- Los **motores de búsqueda académicos** **dominan** las prácticas de búsqueda (76,9%), por delante de los motores de búsqueda generales (62,7%) y las webs de revistas y editoriales (59,4%). Por el contrario, los repositorios (institucional/temático) y las herramientas basadas en IA muestran una adopción mucho menor.
- **Google Scholar** es la herramienta preferida para iniciar una revisión bibliográfica (37,5%), seguida de Google (13.7%).
central entry points into the scholarly information ecosystem
- Existen **diferencias disciplinares** en el uso de bases de datos especializadas.

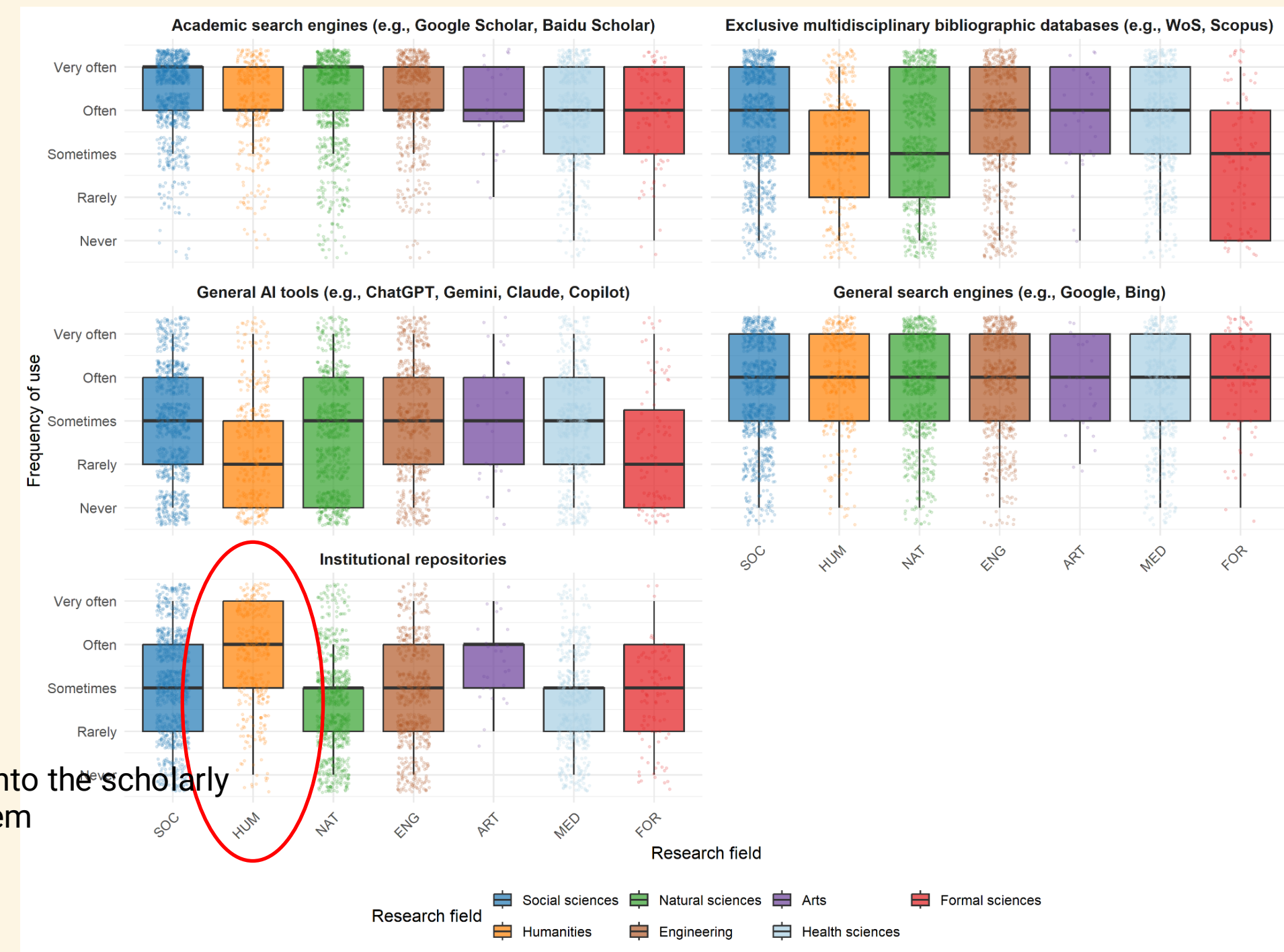


Figura 4. Frecuencia de uso de herramientas específicas de búsqueda según el área.

Resultados: Herramientas de búsqueda

02

Figura 5. Distribución de las herramientas de búsqueda académica según el área



Resultados: Herramientas de búsqueda

02

Selección de herramientas

- El **acceso gratuito o institucional** es el criterio más importante al elegir una herramienta (66,5%).
- La relevancia de los resultados (62,7%) es más importante que las funcionalidades avanzadas (export, alert, citation data).

Retos

- La principal dificultad es la **falta de acceso** al texto completo debido a barreras de pago/suscripciones (64,0%), seguida de la sobrecarga de información y las limitaciones de tiempo.

... sin embargo, las barreras relacionadas con las *habilidades de búsqueda* son poco frecuentes (2,2%).

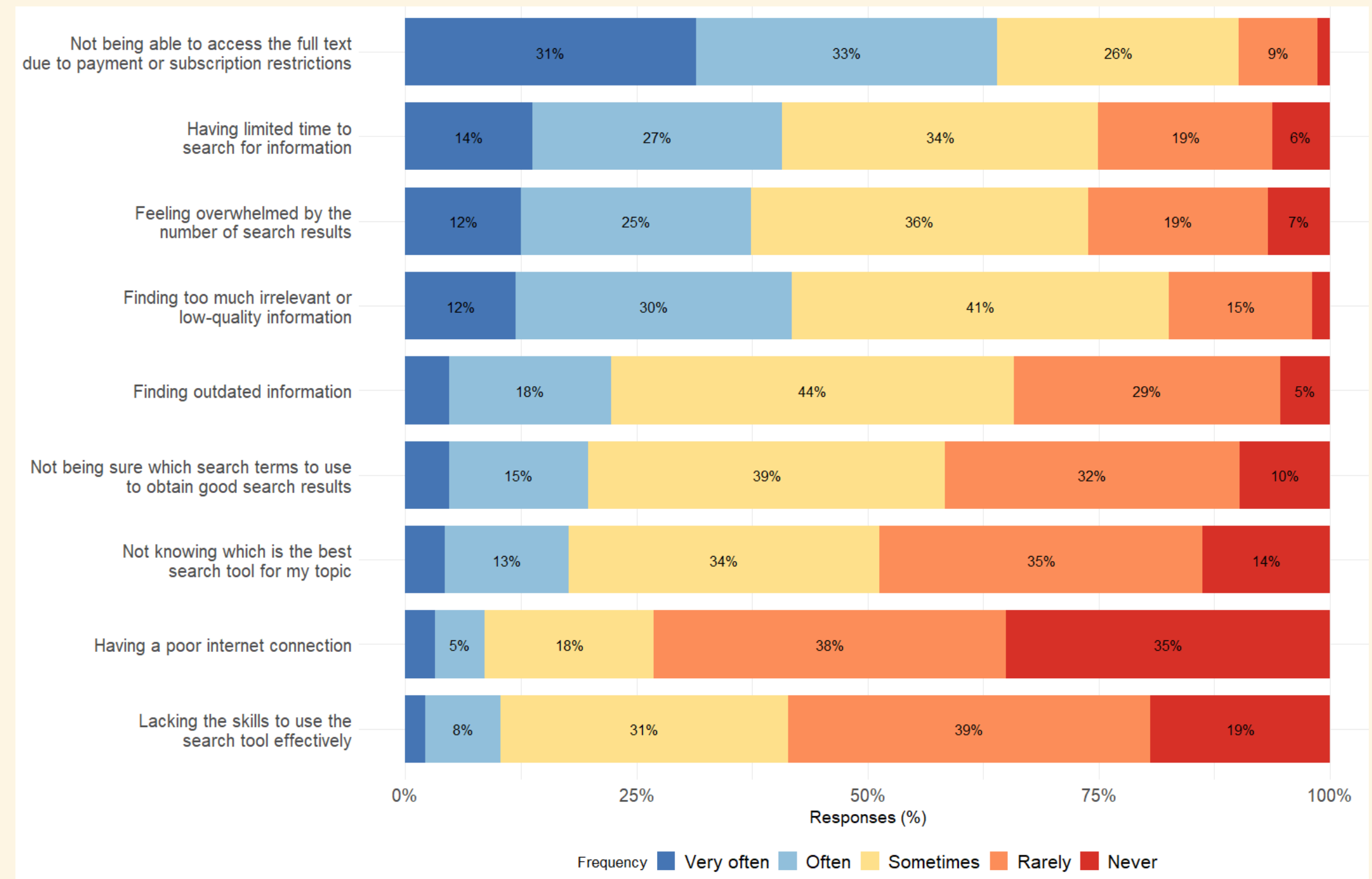


Figura 7. Frecuencia de los desafíos encontrados durante la búsqueda de información científica.

Resultados: Motores de búsqueda académicos

02

Estrategias de búsqueda en motores académicos

- Predomina el uso de lenguaje natural (67,9%) frente a las búsquedas avanzadas (29,1%).
- Las consultas se basan principalmente en palabras clave (87,2%), seguida de copiar/pegar referencia (56,8%). **Búsquedas sencillas**
- Entre quienes realizan **búsquedas avanzadas**, las técnicas más utilizadas son:
 - Comillas para frases exactas (68,7%)
 - Filtros y limitadores (68,1%)
 - Operadores booleanos (48,7%)
 - ... las estrategias más complejas (truncamiento) tienen una adopción limitada.

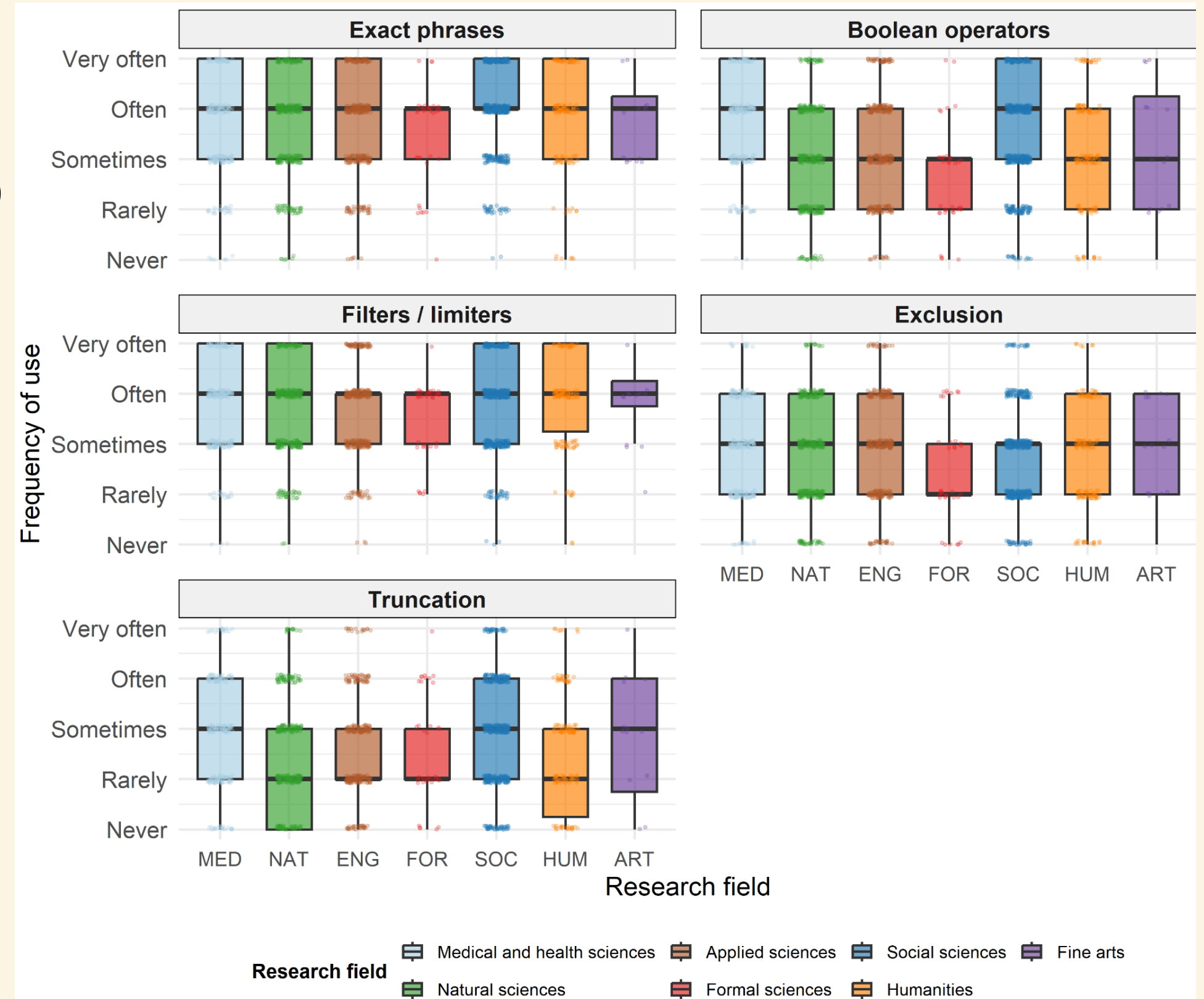


Figura 8. Uso de técnicas avanzadas de búsqueda por área.

Resultados: Motores de búsqueda académicos

02

Estrategias de búsqueda en motores académicos

- Predomina el uso de lenguaje natural (67,9%) frente a las búsquedas avanzadas (29,1%).
- Las consultas se basan principalmente en palabras clave (87,2%), seguida de copiar/pegar referencia (56,8%). **Búsquedas sencillas**
- Entre quienes realizan **búsquedas avanzadas**, las técnicas más utilizadas son:
 - Comillas para frases exactas (68,7%)
 - Filtros y limitadores (68,1%)
 - Operadores booleanos (48,7%)
 - ... las estrategias más complejas (truncamiento) tienen una adopción limitada.

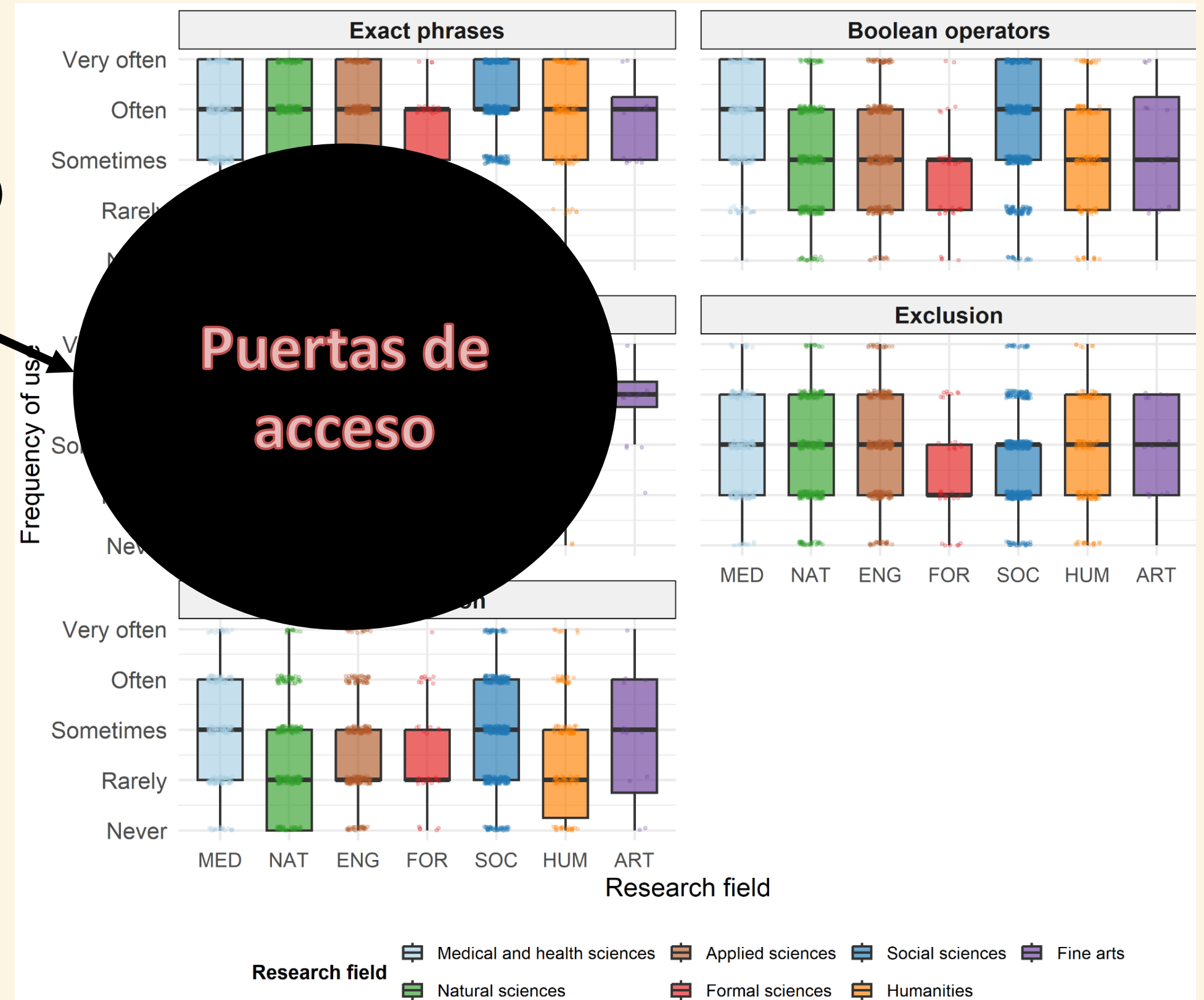


Figura 8. Uso de técnicas avanzadas de búsqueda por área.

Resultados: Motores de búsqueda académicos

02

Estrategias de búsqueda en motores académicos

- Predomina el uso de lenguaje natural (67,9%) frente a las búsquedas avanzadas (29,1%).
- Las consultas se basan principalmente en palabras clave (87,2%), seguida de copiar/pegar referencia (56,8%). **Búsquedas sencillas**
- Entre quienes realizan **búsquedas avanzadas**, las técnicas más utilizadas son:
 - Comillas para frases exactas (68,7%)
 - Filtros y limitadores (68,1%)
 - Operadores booleanos (48,7%)
 - ... las estrategias más complejas (truncamiento) tienen una adopción limitada.

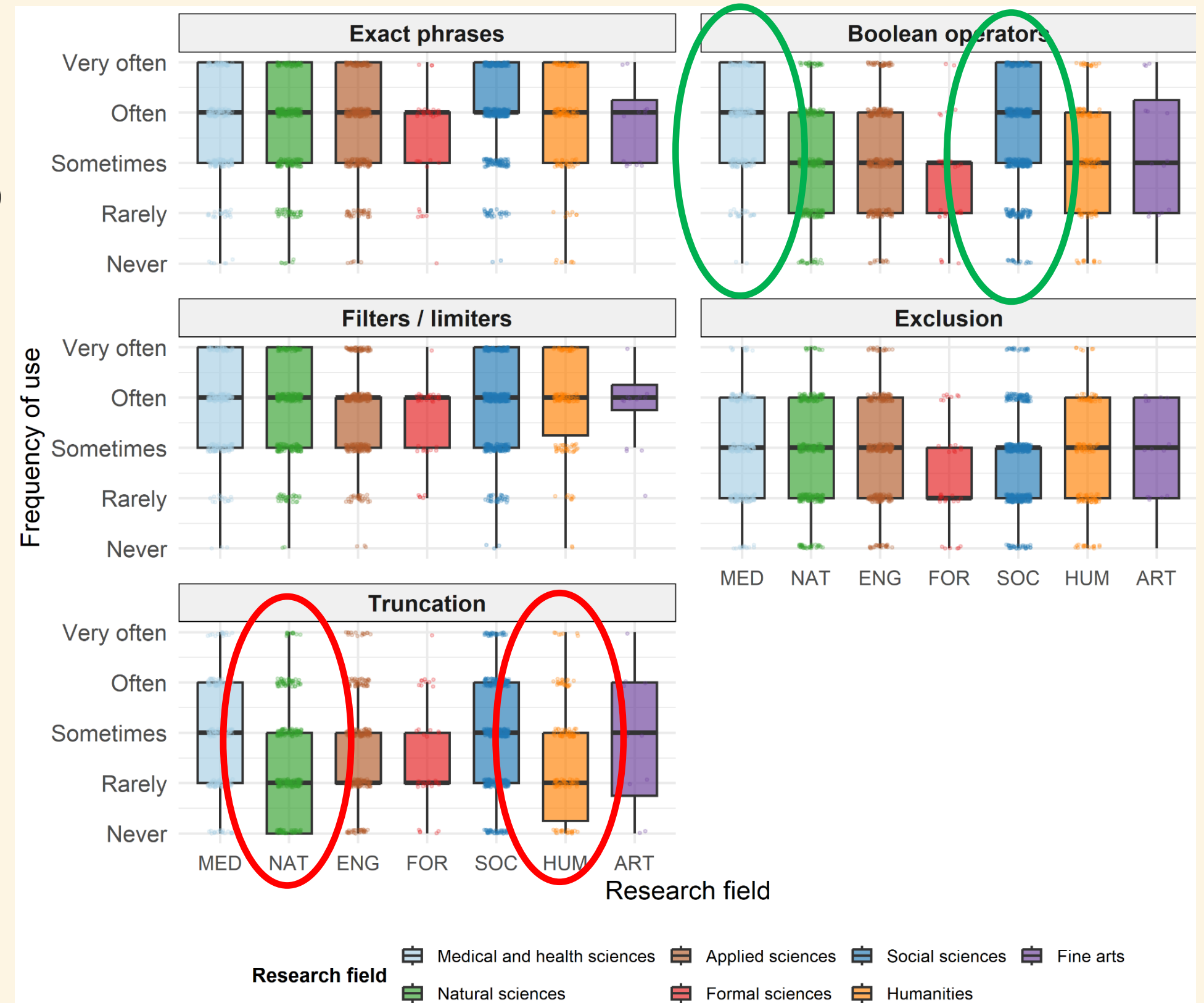


Figura 8. Uso de técnicas avanzadas de búsqueda por área.

Resultados: Motores de búsqueda académicos

02

Factores que influyen en la selección de resultados

- La disponibilidad de acceso directo al texto completo es el principal factor de decisión (74,0%).
- La presencia de términos de búsqueda en el título o descripción también desempeña un papel central.
- La reputación del autor o de la revista tiene una influencia moderada y el número de citas tiene un peso relativamente limitado (43,1%).

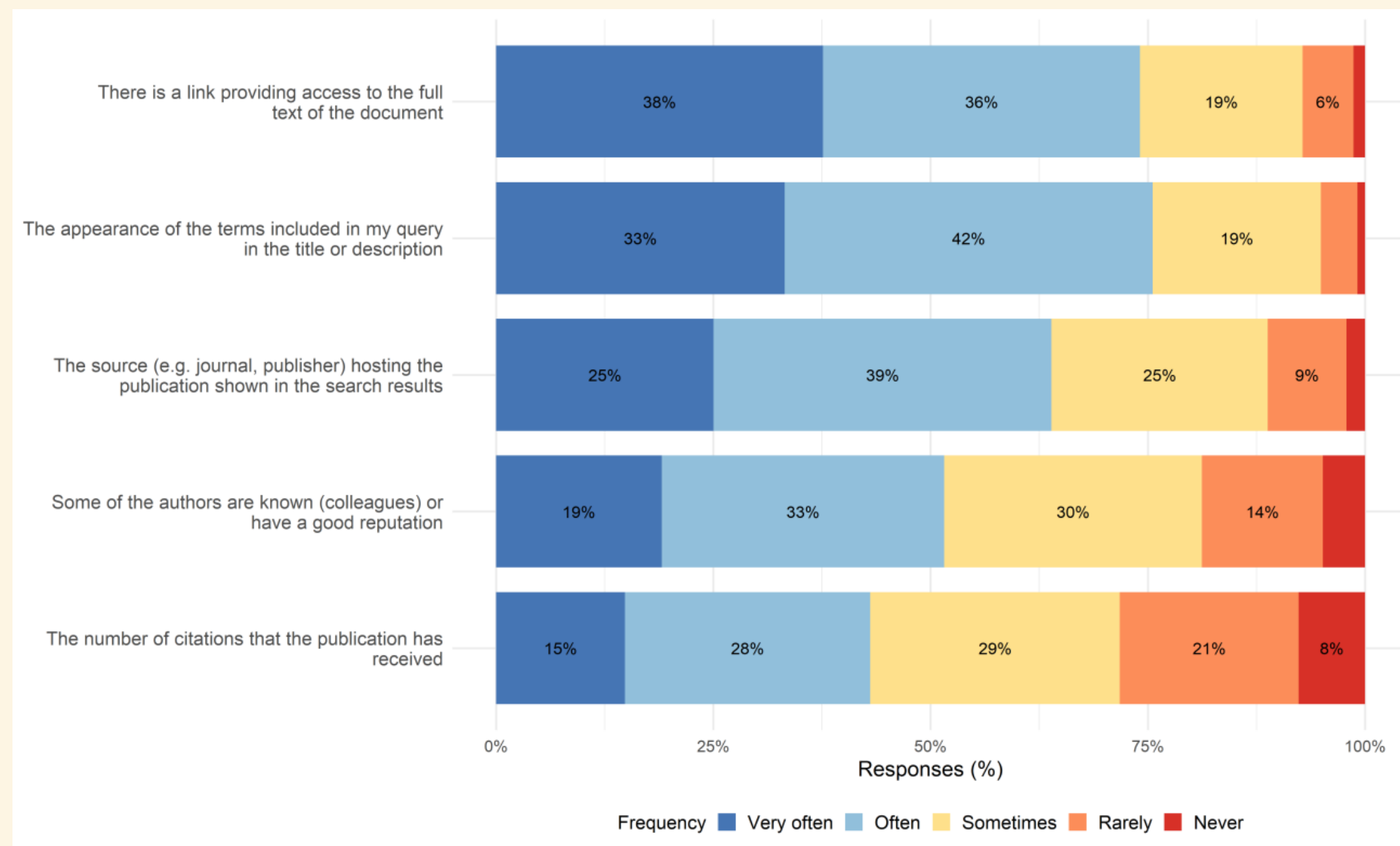


Figura 9. Factores que influyen en el clic sobre un resultado de búsqueda.

Resultados: Motores de búsqueda académica

Comportamiento ante resultados insatisfactorios

- Los investigadores suelen reformular la consulta (72,1%) o explorar nuevas páginas de resultados (71,8%).
- El 74,2% modifica la consulta más de dos veces antes de abandonar la búsqueda.
- El 49,0% explora más allá de la tercera página de resultados.

.... mayor **tolerancia a la exploración**

- El 73,6% se siente confiado o muy confiado en sus capacidades de búsqueda.

Casi la mitad de los investigadores (48,9%) utiliza con frecuencia el botón “Citar” de los motores
...y más usado por mujeres.

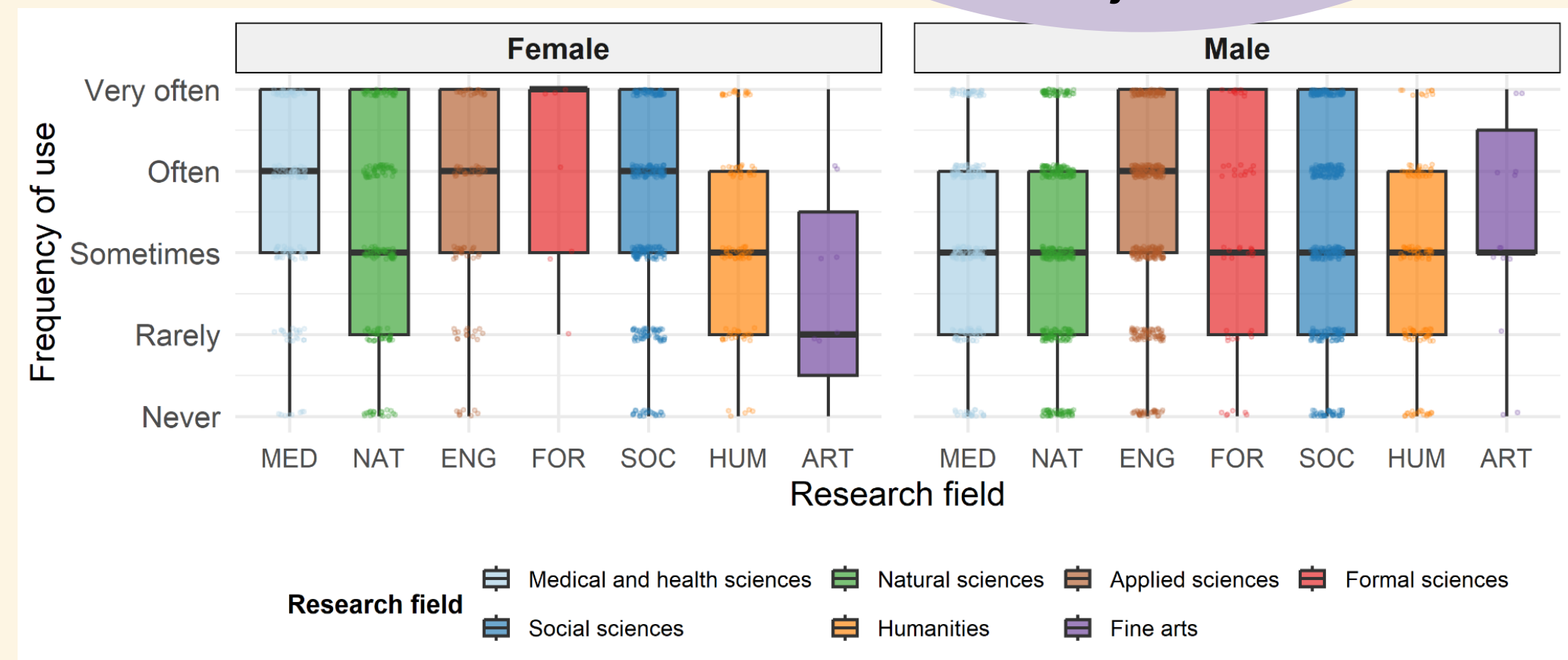


Figura 10. Uso del botón “Citar” según el área de investigación y el género.

Conclusiones (I)

02

- Este estudio ofrece una **visión actualizada** de cómo los investigadores buscan información científica en línea y del papel que desempeñan los motores de búsqueda académicos en los procesos cotidianos de investigación.
- La búsqueda académica está orientada principalmente a la **actualización** y **consolidación del conocimiento**, más que a la resolución de problemas puntuales.
- Los motores de búsqueda se han convertido en la principal **puerta de entrada** a la literatura científica, aunque predominan las búsquedas simples frente a estrategias avanzadas.

Conclusiones (II)

02

- La **accesibilidad** al texto completo influye más en la selección de información científica que los indicadores bibliométricos o las funcionalidades avanzadas de búsqueda.
- La **inteligencia artificial** emerge como uno de los principales factores de cambio en los hábitos de búsqueda académica.
- Las **diferencias sociodemográficas** son limitadas, lo que sugiere una creciente homogeneización de los hábitos de búsqueda académica, impulsada por plataformas como Google Scholar.

Limitaciones y estudios futuros

02

Limitaciones:

Los resultados deben interpretarse con cautela, ya que se basan en comportamientos autodeclarados, por lo que no pueden considerarse plenamente representativos de la población investigadora global.

Implicaciones

- **Formación:** reforzar competencias de búsqueda que combinen estrategias intuitivas y recuperación sistemática de información, promoviendo el uso de herramientas y plataformas abiertas.
- **Investigación:** comprender mejor cómo los algoritmos de descubrimiento condicionan la visibilidad y circulación del conocimiento científico.
- **Políticas científicas:** mejorar la accesibilidad y difusión de la investigación financiada con fondos públicos.

Futuras líneas de investigación

- Analizar 'comportamientos reales' mediante estudios experimentales.
- Analizar longitudinalmente si las tecnologías emergentes (por ej. IA), están transformando los hábitos de búsqueda académica.

**¡GRACIAS POR
VUESTRA ATENCION!**

nuria.bautista@cchs.csic.es