

ESTUDIOS E INFORMES

LINEA 2

2021



MEDICIÓN DEL ACCESO ABIERTO EN LAS
UNIVERSIDADES ESPAÑOLAS Y EL CSIC
(2016-2020)



crue

Universidades
Españolas

Red de Bibliotecas
REBIUN

**MEDICIÓN DEL ACCESO ABIERTO EN LAS UNIVERSIDADES
ESPAÑOLAS Y EL CSIC (2016-2020)**

REBIUN Línea 2. (4º P. E.).Subgrupo de Acceso Abierto.



Documento bajo licencia Creative Commons

ÍNDICE

1. INTRODUCCIÓN
2. METODOLOGÍA
3. OBSERVACIONES PREVIAS
4. RESULTADOS Y ANÁLISIS

1. INTRODUCCIÓN

En 2019 se presentó el documento “*Medición del Acceso Abierto en las universidades españolas y el CSIC (2014-2018)*”¹ (en adelante, “REBIUN OA 2019”) elaborado por el Subgrupo de Acceso Abierto de la Línea 2 de REBIUN, dando respuesta a la Acción 1: “*Hacer un diagnóstico de la situación del acceso abierto en España y un seguimiento constante de su evolución de forma que la información de que se disponga esté siempre actualizada*” establecida en el documento “*Compromisos de las universidades ante la Open Science*”², elaborado por el Grupo CRUE sobre Open Science el 19 de febrero de 2019.

Por ello, en 2021 se ha procedido a realizar un nuevo estudio de la situación de acceso abierto siguiendo la metodología ya utilizada en 2019 desarrollada por el *Observatorio del acceso abierto*³.

2. METODOLOGÍA

La antes mencionada metodología establecida por el *Observatorio del acceso abierto* permite hacer una estimación del porcentaje de publicaciones en acceso abierto de una institución y que los datos sean comparables. Actualmente la estimación se basa en los supuestos siguientes:

- Listas de publicaciones de la institución extraídas de WOS y Scopus: artículos.
- La publicación debe tener DOI.
- Utilización de *Unpaywall*⁴ para clasificar el acceso abierto.

Para llevar adelante el análisis se ha solicitado a cada biblioteca de las instituciones REBIUN que enviara la lista de los DOI de sus publicaciones en WOS y Scopus en el período 2016-2020, siguiendo el manual elaborado por el propio Subgrupo en mayo de 2019 (revisado y actualizado en 2021), en el que se explicaba detalladamente cómo se crean los listados, cómo se deduplican y cómo se envían al Subgrupo.

La recepción de las series de ficheros se ha realizado desde el 14 de mayo hasta el 16 de junio de 2021. De las 75 bibliotecas REBIUN, se han recibido 69 series de ficheros; 5 series más se han elaborado por el Subgrupo y una universidad no participa porque comunicó que no tenía producción científica propia ya que todos sus profesores pertenecen a otras universidades.

Posteriormente el Subgrupo se ha encargado de:

- Consultar en *Unpaywall* los DOI de los ficheros recibidos.
- Hacer la fusión de los resultados proporcionados por *Unpaywall* por años y posteriormente eliminar los registros duplicados (Ver: [OBSERVACIONES PREVIAS](#)).
- Analizar los resultados globales y la comparativa entre el informe “REBIUN OA 2019” y el actual (en adelante, “REBIUN OA 2021”).

¹ Estudio disponible en: <http://hdl.handle.net/20.500.11967/354>. Datos disponibles en: <https://doi.org/10.21950/UQAK4C>

² https://www.crue.org/wp-content/uploads/2020/02/2019.02.20-Compromisos-CRUE_OPENSCIENCE-VF.pdf

³ <https://bibliotecna.upc.edu/es/observatorio>

⁴ *Unpaywall* recolecta su contenido de más de 50.000 revistas y repositorios de acceso abierto, además de plataformas como PubMed Central, DOAJ, Crossref and Datacite. <https://unpaywall.org/faq>
Para agregar cualquier repositorio a estas fuentes, hay que rellenar el formulario disponible en: <https://unpaywall.org/sources>

Hay un destacado cambio metodológico respecto a “REBIUN OA 2019”. *Unpaywall* presenta la novedad del campo “oa_status” que no estaba disponible en 2019, en el que se clasifican los recursos según su ubicación y los términos de su licencia de uso como oro, híbrido, bronce, verde o cerrado (no acceso abierto)⁵. En “REBIUN OA 2019” se recurrió a una fórmula establecida por el *Observatorio de acceso abierto* para obtener esta información. Cuando *Unpaywall* comprueba un DOI no operativo, en “oa_status” no proporciona valor. Estos registros no se han incluido en el cómputo total. En cambio, con la fórmula del estudio “REBIUN OA 2019” se consideraban “No acceso abierto” y por tanto, sí se incluyeron en el cómputo global.

Debemos señalar que la metodología seguida presenta algunas limitaciones:

- Se trata de una estimación, ya que no analiza todas las publicaciones científicas de cada institución, sino únicamente los artículos que aparecen en WOS o Scopus y además tienen un DOI operativo. Sin embargo, resulta bastante representativa.
- *Unpaywall* busca los DOI seleccionados en cada una de sus fuentes registradas y clasifica el tipo de acceso según la vía dorada, verde, bronce o híbrida. Sin embargo, el resultado puede estar algo sesgado a favor de la vía dorada ya que es la primera que rastrea⁶.

Por último, indicar que cada biblioteca recibe un informe individualizado de su institución y que los datos están disponibles en: <https://doi.org/10.21950/2WY6YO>

3.OBSERVACIONES PREVIAS

3.1.Coautoría:

Al unificar los registros de las 74 instituciones pertenecientes a REBIUN se han obtenido registros repetidos que se corresponden a documentos en coautoría de personal investigador de dos o más instituciones. También podría darse el caso de ser registros de autores con doble filiación.

Todos estos duplicados se han eliminado para el cómputo total, excepto los indicados en el [punto 3.2.](#)

El objetivo es una medición del acceso abierto del cómputo total de registros (DOI) diferentes - menos las excepciones señaladas en [3.OBSERVACIONES PREVIAS](#) - producidos por las instituciones analizadas.

En “REBIUN OA 2019” el porcentaje de duplicados fue prácticamente aumentando desde el año 2014

⁵ <https://support.unpaywall.org/support/solutions/articles/44001777288-what-do-the-types-of-oa-status-green-gold-hybrid-and-bronze-mean->

⁶ *Unpaywall* proporciona en el campo “best_oa_location” información sobre la “mejor” ubicación del documento en Open Access (OA) utilizando un algoritmo que prioriza en primer lugar el contenido alojado por el editor (generalmente híbrida y dorada, donde el documento está alojado en el mismo dominio que el DOI resuelve), después prioriza las versiones más cercanas a la versión publicada (*PublishedVersion* sobre *AcceptedVersion*), etc. Devuelve nulo si no hay localización OA. <https://unpaywall.org/data-format#doi-object>

al 2018, con un valor entre 24-25%. En el presente estudio el porcentaje es mayor, oscilando entre 29% y 31%.

	2016		2017		2018		2019		2020	
	Total	%	Total	%	Total	%	Total	%	Total	%
Registros totales proporcionados por las instituciones	85.176		89.992		94.270		105.403		121.596	
Duplicados eliminados	25.235	29,62%	27.501	30,55%	28.777	30,52%	32.874	31,18%	37.335	30,7%
Registros sin duplicados	59.941		62.491		65.493		72.529		84.261	

3.2. Registros en coautoría con diferente clasificación de *Unpaywall*:

Dentro de los registros duplicados por coautoría se ha detectado que un número reducido de ellos presentan diferentes clasificaciones (valor del campo "oa_status") por parte de *Unpaywall*. Al ser un número muy reducido y al no poder establecer qué clasificación es la correcta sin analizar caso por caso, se han incluido en el cómputo global.

	2016	2017	2018	2019	2020
DOI duplicados con diferente "oa_status"	58	97	64	128	236
Registros duplicados con diferente "oa_status"	116	194	128	256	473

3.3. Duplicados (en años diferentes):

WOS y Scopus han proporcionado registros que aparecen en años diferentes. Una hipótesis es que son casos en los que la fecha de primera publicación es de año diferente a la fecha de volumen/número de la revista. Estos registros no han sido eliminados del cómputo global, ya que su cantidad es mínima, aunque destaca el elevado número de duplicados del período 2019-2020.

	2016	2017	2018	2019	2020
2016		109	29	5	11
2017	109		244	27	7
2018	29	244		361	66
2019	5	27	361		2.456
2020	11	7	66	2.456	

3.4.DOI erróneos o no operativos:

Este estudio solo analiza documentos con DOI, ya que es imprescindible la existencia de DOI para el análisis mediante *Unpaywall*. Si el DOI obtenido de WOS/Scopus es erróneo o no operativo (y, por tanto, no se puede “resolver” o interpretar de manera automática), *Unpaywall* no puede recuperar información sobre ese registro.

En este punto hay que mencionar el anteriormente ya indicado cambio metodológico con respecto a “REBIUN OA 2019”. En dicho estudio, la fórmula establecida para obtener la clasificación del tipo de acceso abierto consideraba que los DOI erróneos o no operativos eran “no acceso abierto”. Por lo tanto, se incluían en el cálculo total.

En “REBIUN OA 2021”, el campo “oa_status” que proporciona *Unpaywall* aparece vacío cuando un DOI es erróneo o no operativo. Por lo tanto, no se considera “no acceso abierto/closed” ni se ha incluido en el cálculo total.

	2014		2015		2016		2017		2018		2019		2020	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
DOI erróneos o no operativos. 2019. Incluidos en cómputo total.	1.284	2,32%	1.624	2,91%	1.362	2,36%	1.626	2,67%	1.593	2,54%	---	---	---	---
DOI erróneos o no operativos. 2021. No incluidos en cómputo total	---	---	---	---	1.245	2,12%	1.400	2,29%	1.453	2,26%	1.538	2,16%	1.603	1,93%

3.5.Registros en repositorios con pdf informativo:

En “REBIUN OA 2019” se detectaron registros depositados en repositorios (vía verde) en los que el acceso al documento estaba cerrado/embargado pero había accessible un pdf informativo al respecto. Aunque la fórmula de los ficheros obtenidos de *Unpaywall* los clasificaba como “verde”, en el estudio se consideraron como “no acceso abierto”.

En el presente estudio se ha observado que su número es ínfimo, así que se han considerado como “verde”, tal y como indica *Unpaywall*, y no se recomienda el análisis de esta incidencia para el futuro. Una hipótesis que explicaría la elevada disminución de estos casos podría ser que se hayan superado los períodos de embargo; también podría ser que *Unpaywall* haya encontrado otras alternativas de acceso abierto para dichos registros.

	2014	2015	2016	2017	2018	2019	2020
Registros en repositorios con pdf informativo. 2019	1.330	970	741	593	227	---	---
Registros en repositorios con pdf informativo. 2021	---	---	22	9	14	25	20

4.RESULTADOS, ANÁLISIS Y CONCLUSIONES

Para entender correctamente los resultados, conviene tener en cuenta las definiciones de la clasificación que realiza *Unpaywall*⁷:

- **gold/dorado:** publicaciones en revistas totalmente en acceso abierto.
- **hybrid/híbrido:** publicaciones en revistas de suscripción que ofrecen a los autores publicar en acceso abierto de forma inmediata por medio del pago de una APC (*Article Processing Charges*) y por tanto el artículo está bajo una licencia abierta.
- **bronze/bronce:** publicaciones con lectura en abierto, pero sin licencia de reutilización.
- **green/verde:** publicaciones depositadas en repositorios, ya sean institucionales o temáticos.
- **closed/no acceso abierto:** no se encuentra en acceso abierto.

⁷ <https://support.unpaywall.org/support/solutions/articles/44001777288-what-do-the-types-of-oa-status-green-gold-hybrid-and-bronze-mean->

	2014 estudio 2019		2015 estudio 2019		2016 estudio 2019		2016 estudio 2021		2017 estudio 2019		2017 estudio 2021		2018 estudio 2019		2018 estudio 2021		2019 estudio 2021		2020 estudio 2021	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
Dorado	8.555	15,45%	9.871	17,68%	10.959	18,97%	13.337	22,72%	13.059	21,46%	16.016	26,22%	15.218	24,22%	19.061	29,76%	23.720	33,41%	31.044	37,56%
Híbrido	2.360	4,26%	2.993	5,36%	3.604	6,24%	3.057	5,21%	3.597	5,91%	3.142	5,14%	3.605	5,74%	3.337	5,21%	4.085	5,75%	5.385	6,51%
Bronce	4.483	8,09%	4.152	7,44%	3.951	6,84%	5.264	8,97%	4.060	6,67%	4.601	7,53%	2.924	4,65%	4.025	6,29%	3.742	5,27%	3.101	3,75%
Verde	8.391	15,15%	9.293	16,64%	10.300	17,83%	11.780	20,07%	9.792	16,09%	12.434	20,35%	7.720	12,29%	13.171	20,57%	13.328	18,77%	11.120	13,45%
Total acceso abierto	23.789	42,95%	26.309	47,12%	28.814	49,88%	33.438	56,97%	30.508	50,13%	36.193	59,24%	29.467	46,90%	39.594	61,83%	44.875	63,21%	50.650	61,28%
No acceso abierto	31.593	57,05%	29.523	52,88%	28.952	50,12%	25.258	43,03%	30.353	49,87%	24.898	40,76%	33.363	53,10%	24.446	38,17%	26.116	36,79%	32.008	38,72%
Total	55.382		55.832		57.766		58.696		60.861		61.091		62.830		64.040		70.991		82.658	

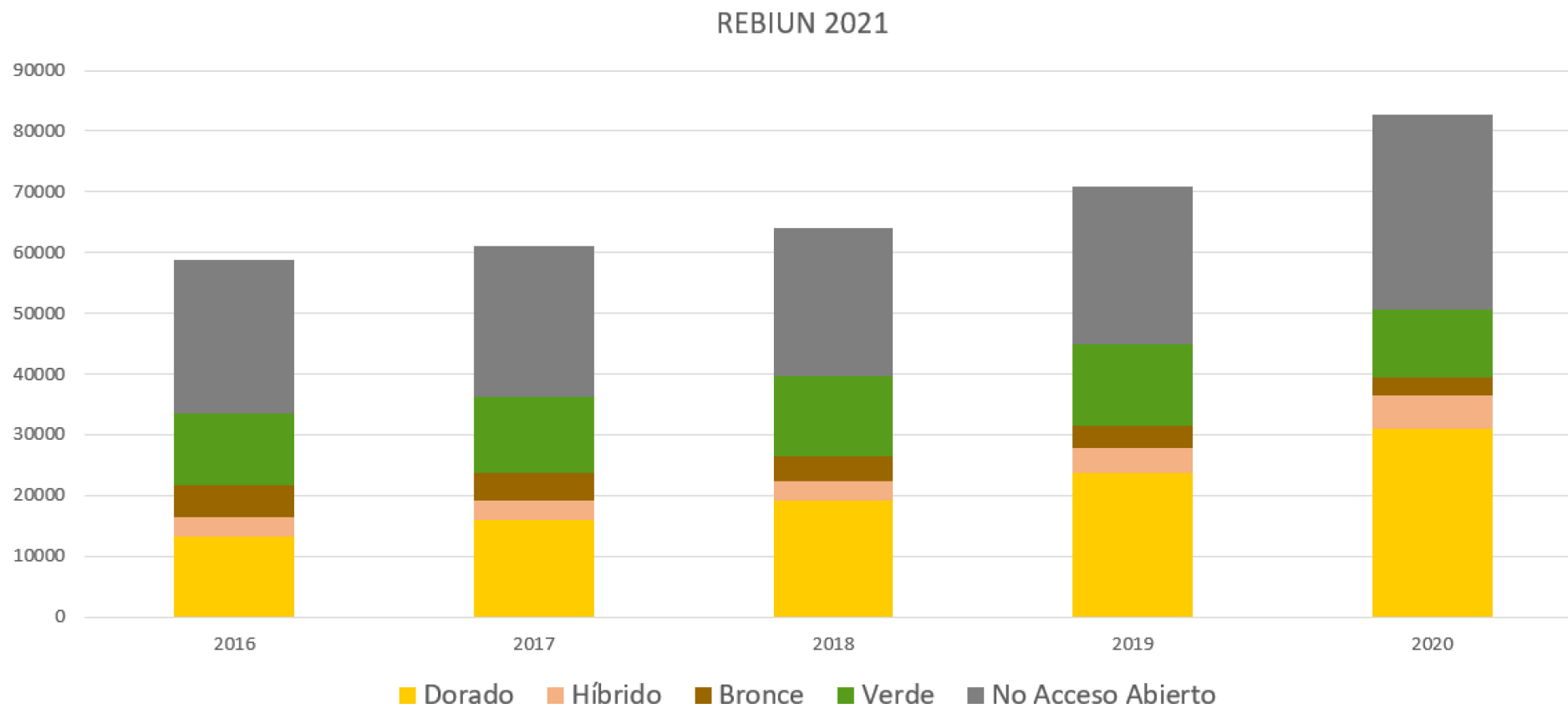


Gráfico 1: Vías de acceso abierto por número de documentos.

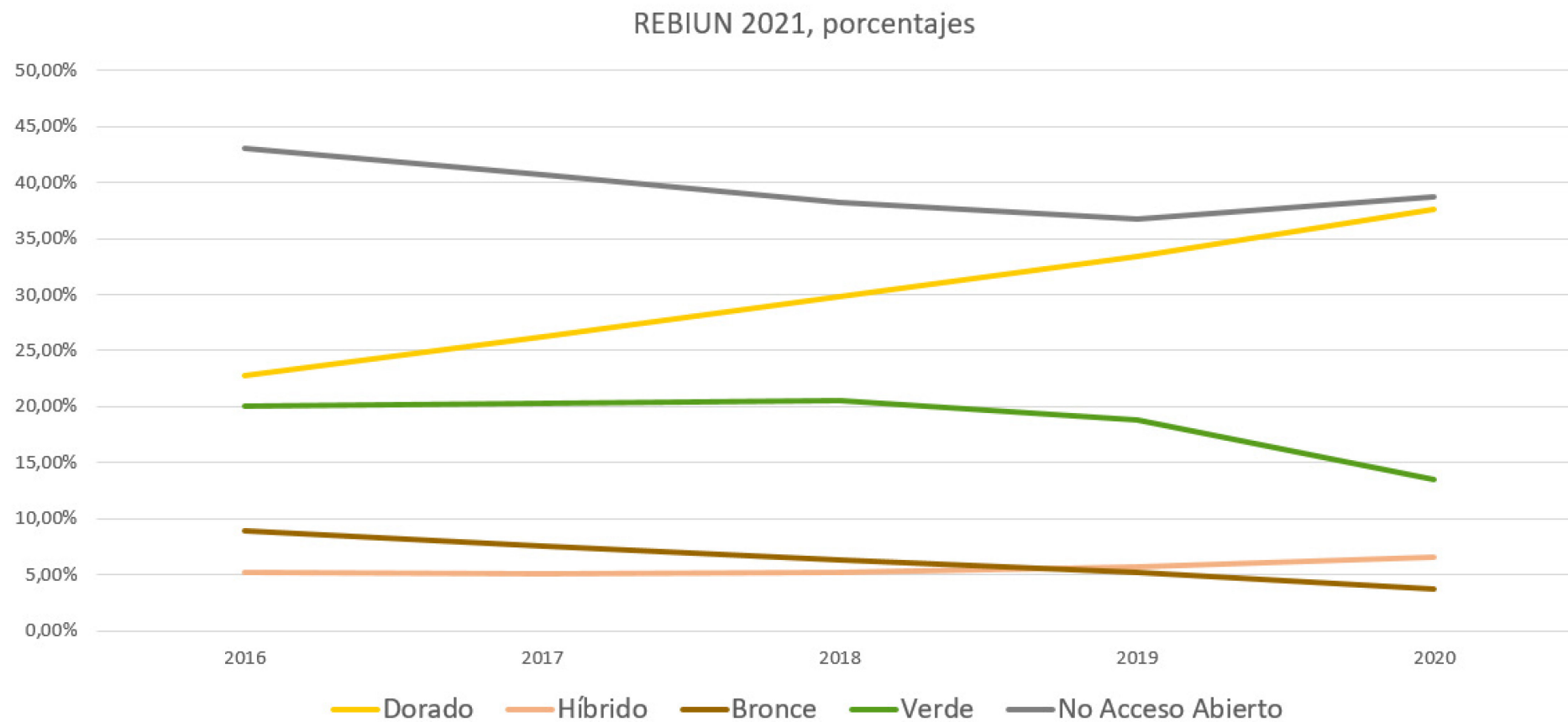


Gráfico 2: Porcentajes de vías de acceso abierto.

REBIUN 2021

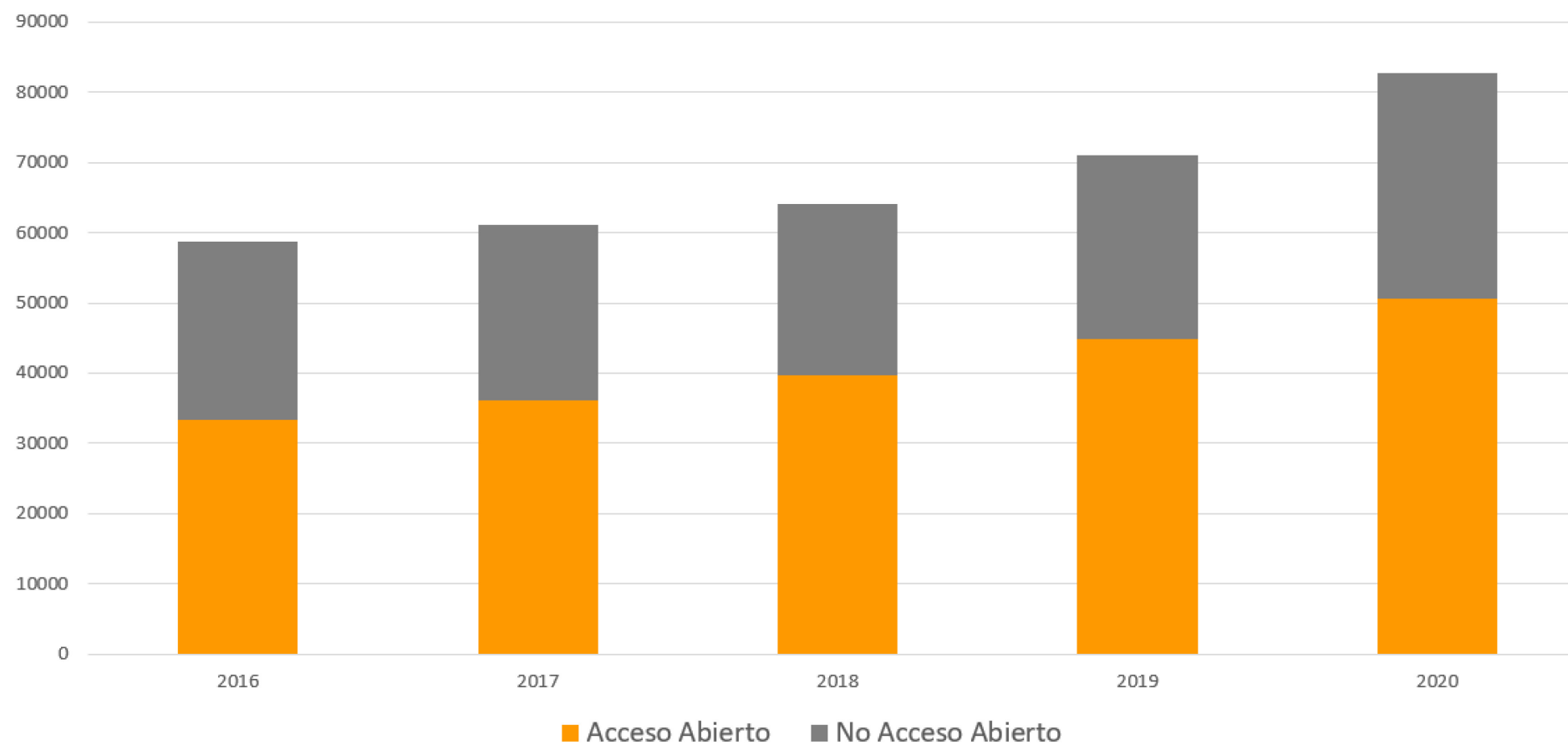


Gráfico 3: Acceso abierto/No acceso abierto por número de documentos.

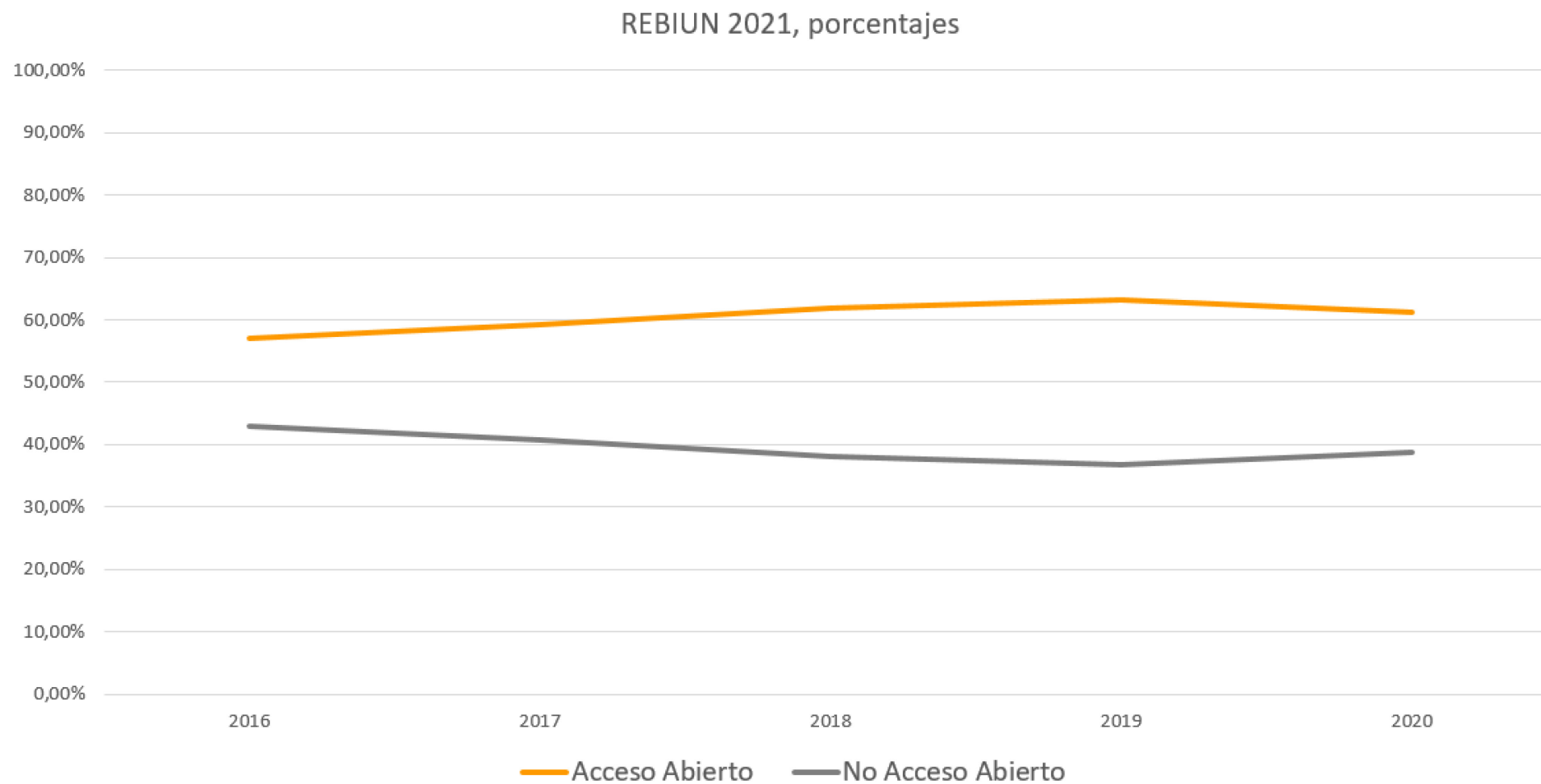


Gráfico 4: Porcentajes Acceso abierto/No acceso abierto.

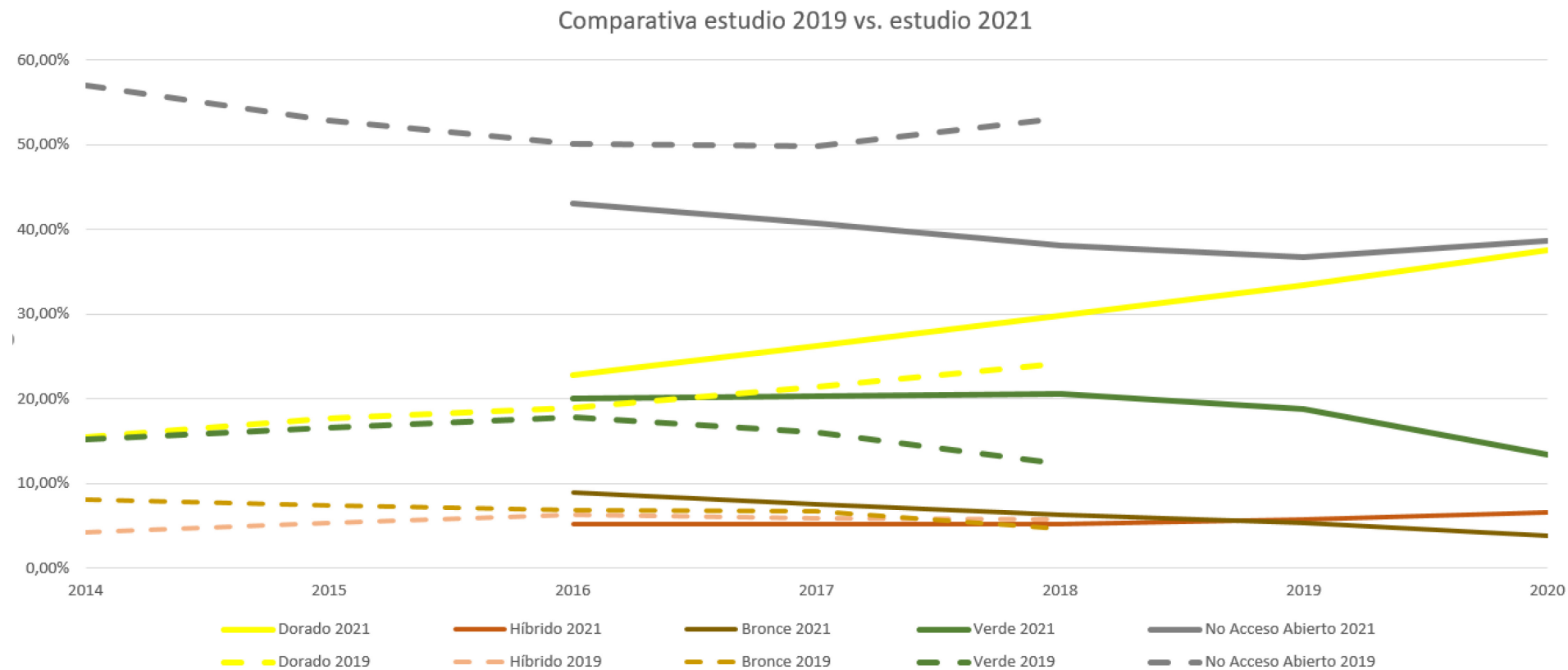


Gráfico 5: Comparativa de porcentajes de vías de acceso abierto entre los estudios “REBIUN OA 2019” y “REBIUN OA 2021”.

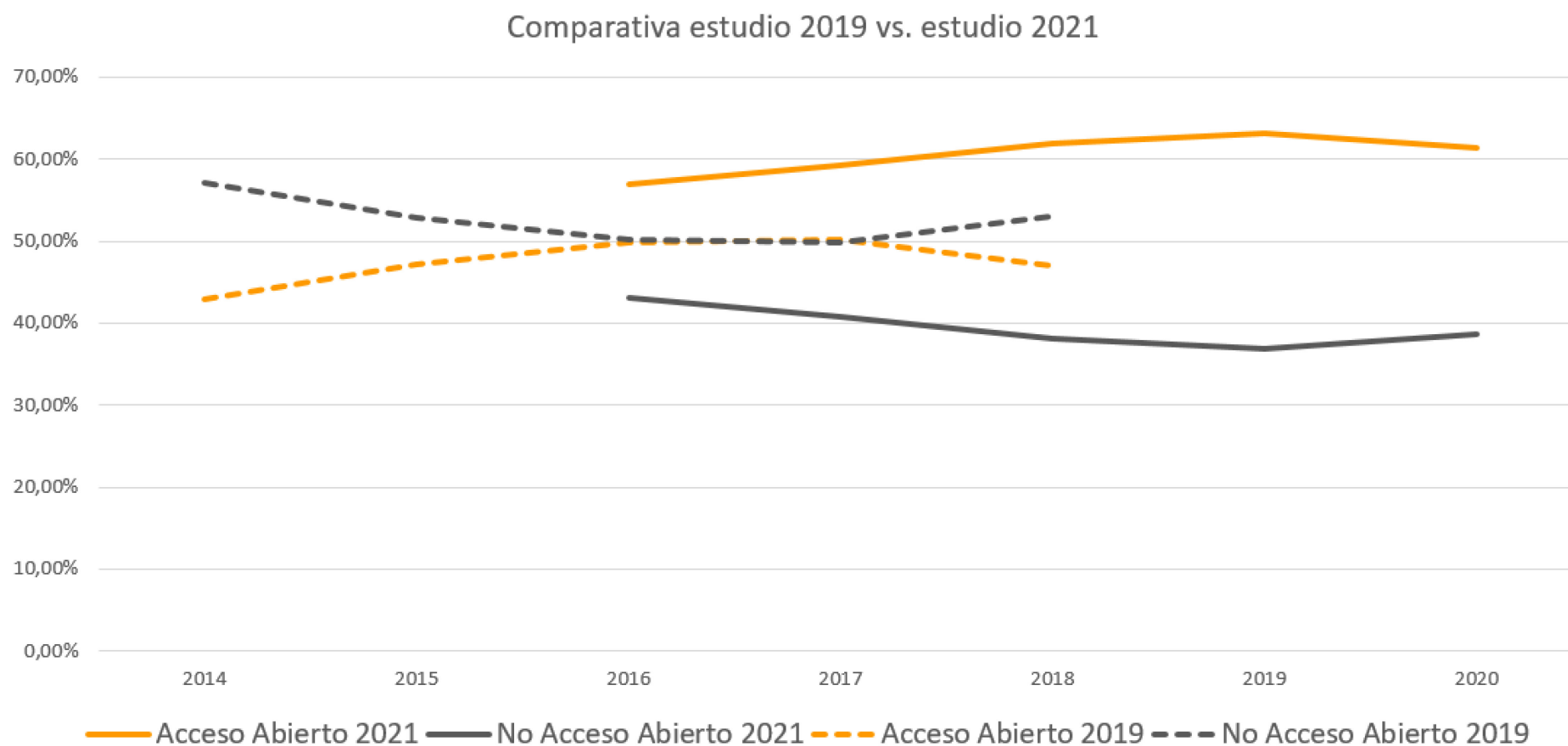


Gráfico 6: Comparativa de porcentajes Acceso abierto/No acceso abierto entre los estudios “REBIUN OA 2019” y “REBIUN OA 2021”.

Conclusiones:

- En primer lugar hay que mencionar que en el estudio “REBIUN OA 2019” se analizó el período 2014-2018 y en el estudio “REBIUN OA 2021”, el período 2016-2020. Por lo tanto, solo los años 2016-2017-2018 son coincidentes en los dos estudios.

Se han vuelto a generar los ficheros de DOI del período 2016-2018, por lo que se pueden apreciar diferencias en el número de registros, probablemente debido a:

- las actualizaciones de las bases de datos WOS y Scopus y por las solicitudes realizadas por las bibliotecas/autores para mejorar filiación, etc.;
- cambio en la metodología. Ya no se aplica la fórmula establecida por el *Observatorio de acceso abierto*, tal y como hemos comentado anteriormente, ya que ahora *Unpaywall* proporciona los valores de las diferentes vías en el campo “oa_status”. Por ello, los DOI con resultado vacío (no operativos) no se han incluido en el cómputo (con la fórmula anterior se consideraban “No acceso abierto”).

Como conclusión, hay que indicar que el número de documentos que finalmente han sido analizados para determinar su vía de acceso abierto/no acceso abierto sigue experimentando un progresivo aumento a lo largo del período del estudio “REBIUN OA 2021”, tal y como se constató en el estudio anterior.

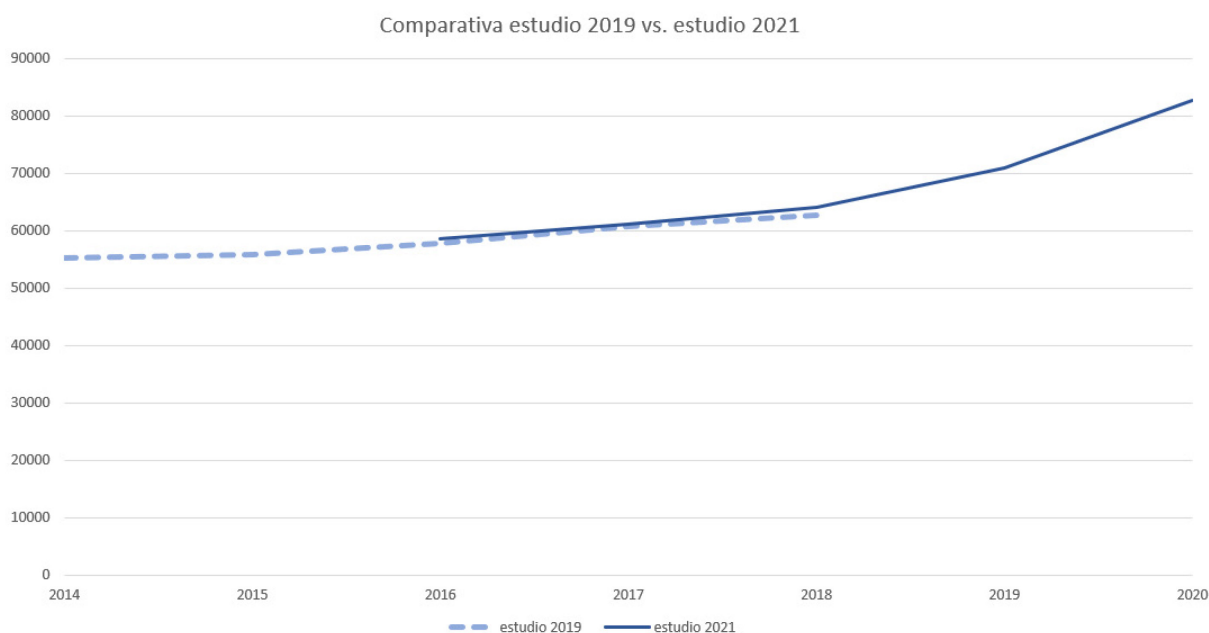


Gráfico 7: Número de documentos.

- En el presente estudio el conjunto de documentos en acceso abierto supera a los documentos que no lo están. Este es el cambio más significativo con respecto al estudio “REBIUN OA 2019”.
- Analizando las diferentes vías de acceso abierto, la predominante sigue siendo la vía dorada, siempre en ascenso. En el período de análisis coincidente (2016-2018) su porcentaje ha aumentado. Podemos apreciar un 3,75% de aumento en 2016; 4,76% en 2017 y 5,54% en 2018. Sería interesante

profundizar en las circunstancias de este aumento. Como hipótesis de partida podrían ser: posibles actualizaciones/revisiones de *Unpaywall*, publicaciones de vía híbrida o bronce que hayan redefinido su política editorial, revistas de acceso cerrado que hayan pasado a ser de vía dorada, etc.

- Las vías híbrida y bronce siguen trayectorias similares entre los valores 3-8%, aunque al finalizar el período de estudio (2020) la vía híbrida supera a la bronce.
- Respecto a la vía verde, también ha aumentado su porcentaje en relación al estudio “REBIUN OA 2019”. Se percibe que el descenso observado en el último período de dicho estudio (años 2017-2018) - probablemente debido a períodos de embargo y a la no inmediatez de la inclusión de los registros en los repositorios - se ha superado en el estudio actual, produciéndose similar descenso en el año 2020, achacable a las posibles mencionadas causas. La vía verde en el año 2018 ha aumentado de 12,29% en el estudio “REBIUN OA 2019” hasta 20,57% en el actual.

Como se ha mencionado anteriormente en la [Metodología](#), *Unpaywall* prioriza la vía dorada frente a la verde⁸, por lo que podemos deducir que ha de ser mayor el número de registros del período analizado depositados en repositorios. Por lo tanto, este estudio no es útil para determinar cuántos documentos están depositados en los repositorios de las instituciones pertenecientes a REBIUN ya que *Unpaywall* prioriza la vía dorada y, además, puede haber encontrado el documento en un repositorio que no pertenezca a REBIUN.

⁸ En: Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, Farley A, West J, Haustein S. (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ* 6:e4375 <https://peerj.com/articles/4375/> denominan a los registros verdes que han quedado relegados frente a otras vías como “shadowed Green”.

