

ARTÍCULO DE INVESTIGACIÓN

Anatomía de un comunicado de prensa

Un censo prerregistrado de 246.777 comunicados en inglés en 1151 sitios

Émile Bolduc^{1*} 

¹ Centro de investigación PPN World, PPN World, Montreal (Quebec), Canadá

* Autor de correspondencia: ppnw@ppnsource.com · ORCID <https://orcid.org/0009-0005-7882-4741>

Cita: Bolduc, É. (2026). Anatomía de un comunicado de prensa: un censo prerregistrado de 246.777 comunicados en inglés en 1151 sitios (PPN World Research Paper No. 2, Edition 1). PPN World. <https://www.ppnworld.com/es/research/anatomy-of-a-press-release>

Registrado: 11 de octubre de 2026

Datos recogidos: 11 de octubre de 2026 – 13 de octubre de 2026

Publicado: 15 de octubre de 2026

Corregido: 17 de octubre de 2026

Derechos de autor: © 2026 PPN World. Artículo de acceso abierto distribuido según los términos de la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0), que permite el uso, la distribución y la reproducción sin restricciones en cualquier medio, siempre que se cite al autor original y la fuente.

Financiación: Este estudio fue financiado por Bollé Communications inc., que opera PPN World y PPN Source. No recibió financiación externa.

Conflicto de intereses: El autor está afiliado a PPN World (Bollé Communications inc.), que publicó este estudio, opera un servicio comercial de seguimiento de comunicados y proporcionó el archivo del que se extrajo la población. PPN Source, un distribuidor que opera la misma empresa, es uno de los sitios de la población (6 comunicados).

Resumen

Contexto. El comunicado de prensa es uno de los textos más escritos de la comunicación profesional y el recurso que los periodistas dicen encontrar más útil, pero su forma se ha medido sobre todo en muestras de unos cientos a unos miles de textos de un solo sector, o en recuentos comerciales limitados a los distribuidores.

Métodos. Nos propusimos leer cada comunicado en inglés con fecha de publicación del 1 de julio de 2026 al 6 de octubre de 2026 en un archivo de comunicados — 246.777 comunicados en 1151 sitios de distribuidores, empresas, gobiernos nacionales, regionales y locales, universidades, ligas deportivas y organizaciones sin fines de lucro — respetando el robots.txt de cada sitio, y medimos 176.492 de ellos en 877 sitios con reglas de detección registradas, junto con el código de análisis, antes de leer ninguna página, y revisadas una vez, públicamente, tras una primera prueba de fiabilidad. Las reglas se contrastaron con un codificador que leía las páginas, y se pusieron a prueba cuatro hipótesis con un bootstrap por conglomerados de sitios y una corrección de Holm.

Resultados. El comunicado se ha partido en dos. En los sitios de los distribuidores, el 64 % de los comunicados termina con una sección «Acerca de» y el 76 % empieza con una data; en los sitios propios de las organizaciones, el 3 % y el 17 % (el 4,3 % y el 11 % sin los sitios deportivos, cuyas páginas son sobre todo noticias). Los comunicados de los distribuidores son más largos en 155 palabras en la mediana (H1 confirmada), sobre todo fuera del relato: el cuerpo, de la entrada a la sección «Acerca de», no es más largo de forma detectable. Las citas de directivos llevan una palabra de emoción más a menudo que las de cargos públicos (11,8 % frente a 4,1 %; H4 confirmada). Las reglas de la data (H2) y del contacto de prensa (H3) no alcanzaron la fiabilidad exigida por el registro: esas dos hipótesis se informan, no se deciden. PARA DIFUSIÓN INMEDIATA sobrevive en el 3,2 % de los comunicados, el 77 % de ellos en un solo sitio, el del Departamento de Justicia de Estados Unidos.

Conclusiones. La sección «Acerca de», una convención de la era de la agencia, pertenece a la copia del distribuidor; la data y el contacto de prensa apuntan en la misma dirección, pero sus reglas no pudieron decidirlo. La copia que una organización publica en su propio sitio ha conservado el relato y poco del aparato que lo rodeaba.

Palabras clave: press release; public relations; genre; content analysis; census; dateline; media contact; boilerplate; newswire; pre-registration

1. Introducción

Un comunicado de prensa es una declaración que una organización publica sobre sí misma, dirigida a los medios y al público. Es una de las formas más antiguas de las relaciones públicas — el género se fecha convencionalmente en la declaración que Ivy Lee publicó para el Pennsylvania Railroad tras un accidente ferroviario en octubre de 1906 (Hiebert, 1966) —

y sigue siendo el material que los periodistas dicen encontrar más útil: en una encuesta comercial de 2025 a más de 3.000 periodistas de 19 mercados, el 72 % lo señaló como el recurso más útil que puede ofrecer un equipo de comunicación (Cision, 2025).

Cómo se escribe un comunicado se enseña en todo curso de relaciones públicas: una data, una entrada que responde a

quién, qué, cuándo y dónde, la cita de un directivo, un párrafo «Acerca de», un contacto de prensa, una marca de cierre. Los lingüistas describen el resultado como un género híbrido, informativo en su forma y promocional en su propósito (Cate-naccio, 2008), y como un texto escrito para ser copiado — «preformulado» para el artículo del periodista, con citas listas para reutilizar (Jacobs, 1999). Si esa forma sobrevive al paso de la agencia a la web, y quién la conserva, no se había medido, que sepamos, a gran escala. Los análisis de contenido del género abarcan sobre todo de unos cientos a unos miles de comunicados de un solo sector (Sumner et al., 2014; Bossema et al., 2019; McLaren y Gurău, 2005; Lassen, 2006); un corpus académico de unos 24.600 comunicados de un solo distribuidor respondía a otra pregunta — si la prensa reutiliza su texto (Nicholls, 2019); el mayor recuento de la forma en sí que hemos encontrado es comercial, el análisis de un distribuidor de 138.901 comunicados de diez agencias, limitado a los distribuidores (PR Gun, 2026). Este estudio se propone medir cada comunicado en inglés de una población definida durante un trimestre, con las hipótesis y el análisis registrados antes de leer ninguna página.

2. Materiales y métodos

2.1 Población

La población es cada comunicado del archivo de PPN World con fecha de publicación del 1 de julio de 2026 al 6 de octubre de 2026, en inglés, de las fuentes que el archivo servía: 246.777 comunicados en 1151 sitios. Un sitio es un origen — esquema, host y puerto — y toma el tipo de editor de las fuentes que enlazan con él. Los distribuidores (agencias comerciales de comunicados) publicaron 142.073 de los comunicados (58 %); el resto lo publicaron gobiernos nacionales y sus organismos, gobiernos de estados, provincias y ciudades, ligas deportivas, empresas, universidades, organizaciones sin fines de lucro, organizaciones políticas y algunos otros en sus propios sitios. Los gobiernos nacionales incluyen aquí organizaciones internacionales como las Naciones Unidas y las instituciones de la Unión Europea (el 13 % del peso de ese grupo), y las ligas deportivas, sus clubes y otros organismos deportivos. Las fuentes del archivo son salas de prensa, y no todo lo que publican es un comunicado en sentido estricto: los sitios deportivos publican sobre todo noticias de sus propios redactores, algunas empresas entradas de blog, y el 92 % del peso de las organizaciones políticas es un archivo de los mensajes de Donald Trump en Truth Social. El censo conserva lo que publicaron las fuentes, como se registró; las limitaciones dan las cifras de los sitios propios sin los sitios deportivos. El estudio es un censo: se incluyó cada comunicado de la población, no se muestreó ninguno.

2.2 Lectura y extracción

Las páginas se leyeron del 11 de octubre de 2026 al 13 de octubre de 2026. Primero se leyó el robots.txt de cada sitio y se aplicó a cada comunicado (Koster et al., 2022); las solicitudes a cada sitio salían una a una, con al menos un segundo de separación, desde un cliente que se identificaba, y una vez más como navegador si era rechazado. Nunca se resolvió una prueba antibots ni se eludió un error de certificado. El texto se

extrajo con Mozilla Readability, el algoritmo del modo lectura de Firefox (Barbresi, 2021; Bevendorff et al., 2023). Un comunicado se mide si su página respondió con al menos 50 palabras en inglés que contenían el titular del archivo: 176.492 comunicados (72 %) en 877 sitios. Los demás se cuentan por motivo — sobre todo páginas rechazadas o desaparecidas (19.795), certificados no válidos (17.428, casi todos en el sitio de un mismo distribuidor), pruebas antibots (11.059) y robots.txt inaccesibles (6077). Dentro de cada sitio, los comunicados medidos se ponderan para que cada sitio conserve su parte de la población.

2.3 Medidas

Un único programa registrado lee cada página: el titular (extensión, una cifra, un verbo de anuncio); la data (un lugar en mayúsculas, una fecha o la etiqueta de un distribuidor, y luego un guion, al inicio del texto); el subtítulo; una firma; la extensión en palabras, total y del cuerpo, de la entrada a la sección «Acerca de»; las citas de ocho palabras o más con un verbo de habla, atribuidas a un directivo cuando junto al hablante figura un cargo empresarial y a un cargo público cuando figura uno público, y si contienen una palabra de emoción (pleased, excited, thrilled, delighted, proud, honored, happy, glad, elated); la sección «Acerca de»; un contacto de prensa — una indicación de contacto con un correo o un teléfono; una marca de cierre (###, -30-); PARA DIFUSIÓN INMEDIATA (FOR IMMEDIATE RELEASE); imágenes, vídeo y enlaces.

2.4 Hipótesis y análisis

Se registraron cuatro hipótesis: los comunicados de los distribuidores son más largos que los de los sitios propios de las organizaciones (H1, diferencia de las medianas ponderadas de palabras); la data es una convención de distribuidor (H2, diferencia de proporciones); la mayoría de los comunicados no nombra a nadie a quien llamar (H3, proporción con contacto, frente a la mitad); y las palabras de emoción son más frecuentes en las citas de directivos que en las de cargos públicos (H4, comunicados comerciales frente a los de organismos públicos, por cita). Los intervalos y los valores p unilaterales provienen de 2.000 réplicas bootstrap que remuestran sitios dentro de cada tipo de editor; las cuatro pruebas llevan una corrección de Holm (Holm, 1979). Una medida solo podía decidir una hipótesis si coincidía con un codificador que leía la página en al menos el 90 % de los comunicados, con un α de Krippendorff $\geq 0,800$ (Krippendorff, 2004; Lombard et al., 2002). Lo que se había visto antes del registro — los identificadores, direcciones y titulares del archivo, y dos pilotos de 100 y 1.001 comunicados fuera de la población — consta en el registro, como se recomienda para el análisis de datos existentes (Weston et al., 2019).

2.5 Fiabilidad de las medidas

Fiabilidad. El codificador A, un modelo de lenguaje (Claude Opus 5.5, Anthropic) — modelos de este tipo han igualado o superado a trabajadores de plataformas en tareas de anotación comparables (Gilardi et al., 2023) —, codificó 200 comunicados extraídos al azar — al menos 60 de distribuidores y 60 de organismos públicos, como se registró; la extracción dio 106, 70 y 24 de otros editores — a partir de la propia página,

con el libro de códigos registrado y sin ver la salida del detector. Un segundo codificador, humano, estaba registrado para un subconjunto y se retiró tras cinco comunicados; el estudio descansa, por tanto, en un solo codificador, un modelo de lenguaje, y lo dice. Donde las reglas registradas fallaron, se revisaron sobre esos 200 comunicados y se probaron una sola vez, sin cambios, en 100 nuevos extraídos de la misma manera (al menos 30 y 30; la extracción dio 56, 37 y 7) y codificados a ciegas (figura 2). La extracción devolvió el comunicado completo en el 78 % de las páginas de la prueba, y una parte en la mayoría de las demás.

2.6 Desviaciones del registro

1. Ritmo. El paso de medición se ejecutó en paralelo y se leyeron más sitios a la vez de lo registrado; las reglas de cortesía no cambiaron, y volver a medir el segundo piloto en paralelo dio resultados idénticos. El ordenador de investigación durmió durante parte del censo; los comunicados que fallaron mientras dormía se volvieron a leer con el ordenador despierto (se informan ambos recuentos).
2. Los rechazos 429 Too Many Requests se volvieron a pedir una vez, con al menos 30 segundos de separación por sitio; las hipótesis se deciden con el recuento registrado y el recuento repetido es una fila de robustez.
3. El segundo codificador, humano, estaba registrado para 50 comunicados y doce preguntas, luego para 20 y cuatro; se detuvo tras cinco, y no se calcula acuerdo entre codificadores.

4. Las reglas registradas del contacto y de la cita de directivo no superaron la prueba de fiabilidad en los 200 primeros comunicados. La regla del contacto lee ahora también la página alrededor del texto extraído; los cargos directivos se leen con cualquier uso de mayúsculas y se vinculan a un hablante nombrado después solo por su apellido; las comillas se emparejan dentro de un párrafo. Las páginas de GOV.UK que no son comunicados (avisos estadísticos, consultas, publicaciones) se apartaron: 3655 páginas, contadas en la población entre las no medidas. Las reglas revisadas se congelaron antes de una segunda prueba, a ciegas, con 100 comunicados nuevos, cuyo resultado se consignó antes de volver a medir el censo. Los resultados registrados se informan junto a los revisados.

3. Resultados

3.1 Dos comunicados

La figura 1 lee un comunicado de arriba abajo en los dos tipos de sitio. Las piezas que la agencia volvió estándar — la data arriba, la sección «Acerca de» y el contacto abajo — son la norma en los sitios de los distribuidores y raras en los de las organizaciones: una sección «Acerca de» cierra el 64 % de los comunicados de los distribuidores frente al 3 %, una data abre el 76 % frente al 17 %, un contacto aparece en el 65 % frente al 14 %. La cita es la pieza que ambos conservan: el 59 % de los comunicados de los distribuidores y el 42 % de los demás citan a alguien.

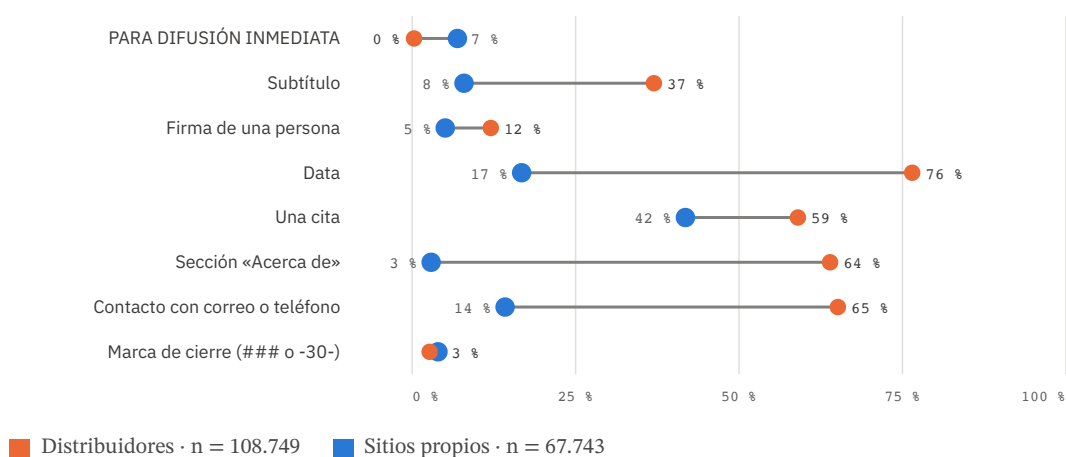


Figura 1. Dos comunicados. Proporción de comunicados que llevan cada elemento, en los sitios de los distribuidores y en los sitios propios de las organizaciones, ponderada por los comunicados que no pudieron leerse. Elementos en el orden en que se lee un comunicado. Las reglas del subtítulo, la marca de cierre y la firma no alcanzaron el umbral de fiabilidad (figura 2): esas filas son orientativas.

El reparto no enfrenta a empresas y gobiernos. Las empresas que publican en su propia sala de prensa conservan más de ese aparato que los organismos públicos — el 21 % tiene sección «Acerca de» y el 24 % un contacto — pero mucho menos que el mismo tipo de texto en el sitio de un distribuidor (tabla

- 1). Los gobiernos regionales y locales son la excepción entre los organismos públicos: el 46 % de sus comunicados nombra un contacto. Las ligas deportivas, el mayor grupo de sitios propios medido, casi nunca lo hacen (0,6 %): sus páginas son sobre todo noticias, no comunicados.

Tabla 1. Anatomía de un comunicado, por editor. Comunicados medidos, extensión mediana en palabras y proporción que lleva cada pieza, ponderada por los comunicados que no pudieron leerse. Los distribuidores son las agencias comerciales; todas las demás filas publican en sus propios sitios. Los gobiernos nacionales incluyen organizaciones internacionales; las ligas deportivas, clubes y otros organismos deportivos, cuyas páginas son sobre todo noticias; la mayor parte del peso de las organizaciones políticas es un archivo de los mensajes de Donald Trump en Truth Social.

Editor	Medidos	Palabras (mediana)	Data	«Acerca de»	Contacto	Cita	Para difusión inmediata
Todos los comunicados	176.492	588	51 %	38 %	43 %	52 %	3 %
Distribuidores	108.749	637	76 %	64 %	65 %	59 %	0 %
Empresas	4120	718	17 %	21 %	24 %	36 %	0 %
Gobiernos nacionales	23.152	401	9 %	2 %	19 %	35 %	15 %
Gob. regionales y locales	7106	405	21 %	4 %	46 %	43 %	13 %
Universidades	4134	678	2 %	2 %	17 %	66 %	0 %
Ligas deportivas	26.362	594	27 %	1 %	1 %	47 %	0 %
Organizaciones sin fines de lucro	1621	513	15 %	10 %	11 %	46 %	0 %
Organizaciones políticas	873	128	2 %	0 %	0 %	5 %	0 %
Todos los sitios propios	67.743	482	17 %	3 %	14 %	42 %	7 %

3.2 La extensión

El comunicado mediano tiene 588 palabras, con un nivel Flesch–Kincaid de 15 (Kincaid et al., 1975) — escrito para un lector universitario. H1 se confirma: los comunicados de los distribuidores son más largos que los de los sitios propios en 155 palabras en la mediana (intervalo del 95 %: 55 a 244; p ajustada de Holm = 0,005).

El registro hacía la misma pregunta solo para el cuerpo, de la entrada a la sección «Acerca de», al contacto o a la marca de cierre (R5). Ahí la diferencia casi desaparece: 468 palabras en los sitios de los distribuidores, 465 en los de las organizaciones (diferencia de 3, intervalo del 95 %: –84 a 96). Sin los sitios deportivos, cuyas noticias son largas, es de 57 palabras (intervalo del 95 %: –2 a 117; una comprobación no registrada), frente a 204 para el comunicado entero. El comunicado de un distribuidor es más largo sobre todo por lo que rodea al relato — data, sección «Acerca de», bloque de contacto, avisos legales —, y no, de forma detectable, porque el relato lo sea.

3.3 A quién llamar

El bloque de contacto es la única pieza de un comunicado escrita solo para el periodista. H3 preguntaba si la mayoría de los comunicados ya no nombra a nadie a quien llamar. Según la medición, el 43 % de los comunicados lleva un correo o un teléfono bajo una indicación de contacto (intervalo del 95 %: 24 % a 57 %) — el 65 % en los sitios de los distribuidores, el 14 % en los de las organizaciones.

La medida no pudo decidir. En 100 comunicados extraídos para ponerla a prueba, la regla coincidió con un codificador que leía la página en el 89 % de los casos (α ,779), por debajo del 90 % registrado. Rara vez se equivoca cuando encuentra un contacto (precisión 96 %), pero pasó por alto el 16 % de los que encontró el codificador, sobre todo contactos bajo una fórmula que no conoce («Questions?», «Address for clarification») o que la extracción dejó fuera. La proporción del censo es, por tanto, un mínimo, lo bastante cerca de la mitad como

para que la pregunta registrada siga abierta: H3 se informa, no se decide. Omisiones a ese ritmo no cerrarían, sin embargo, la brecha entre los dos tipos de sitio.

3.4 La data

Una noticia enviada desde fuera empieza por el lugar donde se escribió, en mayúsculas (Associated Press, 2026); los distribuidores aplican la convención a los comunicados, a menudo con la fecha y su propio nombre. El 51 % de los comunicados empieza así: el 76 % en los sitios de los distribuidores y el 17 % en los de las organizaciones, una diferencia de 60 puntos (intervalo del 95 %: 23 a 73).

H2 se informa, no se decide. La regla de la data superó la prueba de fiabilidad en los 200 primeros comunicados codificados (90,5 % de acuerdo, α ,810) y no la superó en los 100 extraídos para probar las reglas revisadas (87 %, α ,739); el registro aplica la prueba a la segunda extracción. Sus errores van sobre todo en un sentido: cuando encuentra una data, casi siempre acierta (precisión 98 %); pasó por alto el 20 % de las que encontró el codificador.

3.5 Citas y emociones

El 52 % de los comunicados cita a alguien en una frase de ocho palabras o más. H4 compara a los directivos citados en comunicados comerciales con los cargos públicos citados en los de organismos públicos. Las citas de directivos llevan una palabra de emoción — pleased, excited, thrilled, delighted, proud, honored, happy, glad, elated — en el 11,8 % de los casos (132.328 citas); las de cargos públicos, en el 4,1 % (13.313). H4 se confirma: la diferencia es de 7,6 puntos (intervalo del 95 %: 4,9 a 8,8; p ajustada de Holm = 0,002).

La emoción es una fórmula más que la regla: la mayoría de las citas de directivos no contiene ninguna de las nueve palabras. Cuando aparece una, pertenece al registro que describe Jacobs (1999): una cita escrita por el equipo de comunicación, para trasladarse entera a un artículo.

Tabla 2. Las cuatro hipótesis registradas. Estimaciones con las reglas de detección revisadas, intervalos bootstrap del 95 % por conglomerados y valores p ajustados de Holm; la última columna da la estimación con las reglas registradas. H2 y H3 descansan en medidas que quedaron por debajo de la regla de fiabilidad registrada y no se deciden.

Hipótesis	Estimación	Intervalo del 95 %	p (Holm)	Veredicto	Registrada
H1 · Los comunicados de los distribuidores son más largos (palabras, mediana)	155	55 a 244	0,005	Confirmada	169
H2 · La data es una convención de distribuidor (diferencia, puntos)	60	23 a 73	0,002	No decidida	60
H3 · La mayoría no nombra a nadie a quien llamar (proporción con contacto)	43 %	24 % a 57 %	0,185	No decidida	29 %
H4 · Más palabras de emoción en los directivos (diferencia, puntos)	7,6	4,9 a 8,8	0,002	Confirmada	8,1

3.6 Las marcas de la era del fax

PARA DIFUSIÓN INMEDIATA indicaba a una redacción que un comunicado en papel podía publicarse enseguida. En una página web ya no tiene a quién instruir, y es rara: la lleva el 3,2 % de los comunicados, el 0,3 % en los sitios de los distribuidores. Sobrevive en los gobiernos, y sobre todo en un sitio: el Departamento de Justicia de Estados Unidos la imprime en cada comunicado medido, y sus comunicados son el 77 % de todos los que la llevan — una estimación basada en los 365 de sus 4862 comunicados que pudieron leerse (casi todos los demás estaban tras una prueba antibots), que las ponderaciones hacen valer por todos. La proporción de los gobiernos nacionales es del 14,9 % con él y del 0,5 % sin él; los gobiernos regionales y locales la mantienen en el 13,0 % de sus comunicados. La marca de cierre — ### o -30-, «fin del texto» — aparece en el 3,2 % de los comunicados, en una cuarta parte de los de gobiernos regionales y locales (26 %) y en pocos más; la regla de la marca de cierre, como las del subtítulo y la firma, no alcanzó el umbral de fiabilidad (figura 2), así que estas proporciones son orientativas.

Los titulares son largos. El mediano tiene 81 caracteres, 93 en los sitios de los distribuidores y 67 en los de las organizaciones; el 75 % de todos los titulares y el 86 % de los de distri-

buidores superan los 60 caracteres, cerca del 90,2 % que halló en las agencias el recuento de un distribuidor (PR Gun, 2026). Un verbo de anuncio (announces, launches, unveils...) abre el 31 % de los titulares de los distribuidores y el 8 % de los demás.

3.7 Comprobaciones de robustez

Robustez. Ponderar igual cada sitio, excluir los cinco sitios más grandes, excluir los artículos firmados (tal como los encuentra la regla de la firma, que no alcanzó el umbral de fiabilidad), las páginas leídas solo tras renderizarlas o las que el cliente identificado no pudo leer, quitar las ponderaciones, y contar los comunicados rechazados primero con 429 Too Many Requests (1853 medidos más) o dejar como se registraron primero los comunicados perdidos mientras el ordenador de investigación dormía (917 medidos menos): cada variante repite las cuatro pruebas. H1 y H4 se confirman en todas ellas; la diferencia de H1 va de 94 a 163 palabras, y la de H4 de 5,5 a 7,7 puntos. Con las reglas registradas se midieron 179.606 comunicados y las estimaciones fueron 169 palabras (H1), 60 puntos (H2), 29 % con contacto (H3) y 8,1 puntos (H4); cada prueba registrada quedaba entonces nominalmente confirmada.

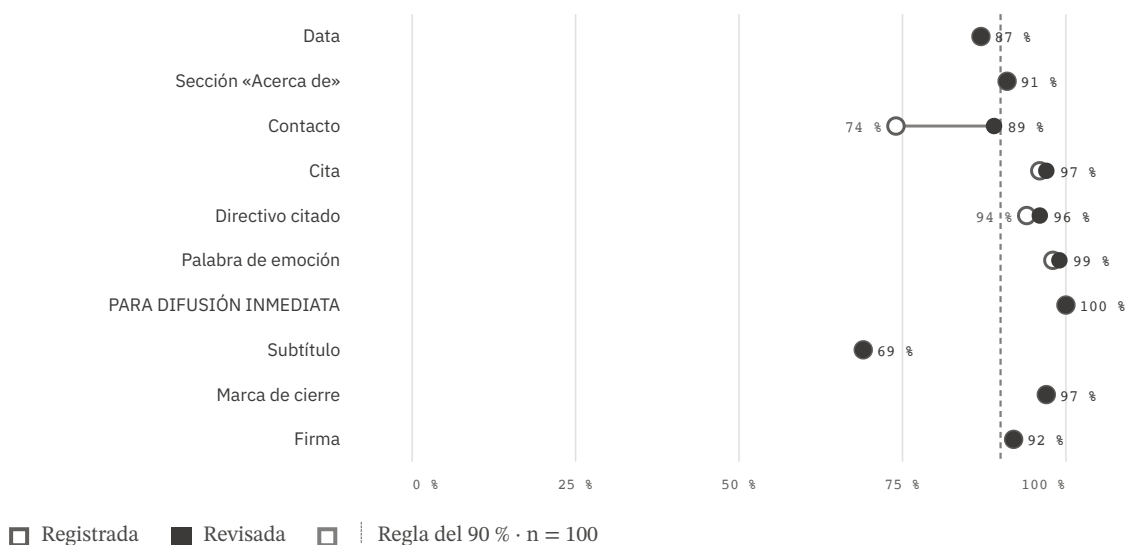


Figura 2. Hasta dónde es fiable cada medida. Acuerdo entre la regla de detección y un codificador que lee la página, en 100 comunicados extraídos después de revisar las reglas (hueco: la regla registrada; lleno: revisada). Una medida decide una hipótesis solo con un 90 % de acuerdo y un α de Kripendorff $\geq 0,800$.

4. Discusión

4.1 Hallazgos principales

1. **La sección «Acerca de» es de los distribuidores.** El 64 % de los comunicados publicados en los sitios de los distribuidores termina con un párrafo estándar que describe a la organización; el 3 % de los publicados en los sitios propios de las organizaciones. El mismo reparto vale para la data (76 % frente a 17 %).
2. **Los comunicados de los distribuidores son más largos — sobre todo alrededor del relato (H1).** En la mediana, un comunicado tiene 637 palabras en el sitio de un distribuidor y 482 en el de la organización (intervalo del 95 % de la diferencia: 55 a 244; H1 confirmada). El cuerpo, de la entrada a la sección «Acerca de», no es más largo de forma detectable: 468 frente a 465 palabras (468 frente a 411 sin las noticias de los sitios deportivos). La diferencia está sobre todo en el aparato que lo rodea.
3. **El comunicado típico tiene 588 palabras; el de una empresa, 718; el de un gobierno nacional, 401.** En la mediana, un comunicado tiene 588 palabras: 637 en los sitios de los distribuidores y 482 en los de las propias organizaciones. Según quién lo publica, el de una empresa tiene 718 palabras, el de una universidad 678, el de una liga deportiva 594, el de un gobierno regional o local 405 y el de un gobierno nacional 401. La diferencia está sobre todo en el texto mismo: el cuerpo del comunicado de una empresa tiene 637 palabras, el de un gobierno nacional 387.
4. **Los directivos se dicen encantados más a menudo que los cargos públicos (H4).** Una palabra de emoción — pleased, excited, thrilled, delighted, proud — aparece en el 11,8 % de las citas de directivos en comunicados comerciales y en el 4,1 % de las citas de cargos públicos en comunicados de organismos públicos (diferencia de 7,6 puntos, intervalo del 95 %: 4,9 a 8,8; H4 confirmada).
5. **PARA DIFUSIÓN INMEDIATA sobrevive en un sitio federal — y en los gobiernos regionales y locales.** La línea escrita para el fax sobrevive en el 3,2 % de los comunicados, y el 77 % de ellos está en un solo sitio: el Departamento de Justicia de Estados Unidos la imprime en cada comunicado que pudimos leer (365 de 4862). Sin él, la proporción de los gobiernos nacionales cae del 14,9 % al 0,5 %, y la global al 0,8 %. Los gobiernos regionales y locales la mantienen en el 13,0 % de los comunicados; los distribuidores, en el 0,3 %.
6. **Los titulares son largos, sobre todo en los distribuidores.** El titular mediano tiene 81 caracteres — 93 en los sitios de los distribuidores, 67 en los de las organizaciones. El 75 % de los titulares supera los 60 caracteres, el 86 % en los distribuidores.

4.2 Qué significa para los comunicadores

¿Qué extensión tiene un comunicado de prensa típico? 588 palabras en la mediana, sobre 176.492 comunicados en inglés medidos en 2026: 637 en los sitios de los distribuidores, 482 en los de las organizaciones. El relato en sí — de la entrada a la

sección «Acerca de» — tiene 468 palabras en los sitios de los distribuidores y 465 en los de las organizaciones.

¿Un comunicado sigue necesitando una data? En el sitio de un distribuidor es la norma: el 76 % de los comunicados empieza con una. En el sitio de la organización, el 17 % (el 11 % sin los sitios deportivos). Son estimaciones: la regla de la data no alcanzó la fiabilidad exigida por el registro, y deja escapar datas más a menudo de lo que las inventa. La data pertenece a la copia que sale por la agencia.

¿Debe un comunicado incluir un contacto de prensa? Es la única pieza escrita solo para el periodista, y la copia del sitio de la organización suele omitirla: según la medición, el 14 % lleva un correo o un teléfono (el 23 % sin los sitios deportivos), frente al 65 % en los sitios de los distribuidores. Son estimaciones: la regla del contacto no alcanzó la fiabilidad exigida por el registro. Una sala de prensa sin un nombre al que llamar envía al periodista a la copia del distribuidor — o a otra parte.

¿Se sigue usando PARA DIFUSIÓN INMEDIATA? Rara vez: la lleva el 3,2 % de los comunicados, y el 77 % de ellos procede de un solo sitio, el del Departamento de Justicia de Estados Unidos. Sin él, la proporción cae al 0,8 %, y los que más la mantienen pasan a ser los gobiernos regionales y locales (13,0 %); en los sitios de los distribuidores casi no aparece (0,3 %).

¿Son las citas de directivos más emotivas que las de cargos públicos? Sí, aunque no es la regla: el 11,8 % de las citas de directivos en comunicados comerciales contiene pleased, excited, thrilled, delighted, proud o una palabra similar, frente al 4,1 % de las citas de cargos públicos en comunicados de organismos públicos.

4.3 Limitaciones

- La población son las fuentes de un archivo, elegidas para un servicio comercial de seguimiento, durante un trimestre, en inglés. Los distribuidores aportan más de la mitad de sus comunicados; las organizaciones que solo publican en sus propios sitios, y las que el archivo no sigue, están infrarrepresentadas. Los resultados se dan por tipo de editor para que ningún grupo oculte a otro.
- No todo lo que publican las fuentes del archivo es un comunicado en sentido estricto. Los sitios deportivos aportan el 38 % del peso de los sitios propios y publican sobre todo noticias; sin ellos, el 4,3 % de los comunicados de los sitios propios lleva sección «Acerca de», el 11 % una data y el 23 % un contacto (frente al 3 %, el 17 % y el 14 %), el 52 % de todos los comunicados lleva un contacto (intervalo del 95 %: 33 % a 64 %), y la diferencia de H1 es de 204 palabras (122 a 269). El reparto entre los dos tipos de sitio se mantiene. Estas comprobaciones se hicieron después de la publicación y no estaban registradas.
- Dos sitios de distribuidores concentran el 68 % del peso de los distribuidores, así que las cifras de los distribuidores los describen sobre todo a ellos; sin los cinco sitios más grandes (R4), H1 y H4 siguen confirmadas.
- Los comunicados tras pruebas antibots, tras certificados no válidos, rechazados a nuestro cliente o en sitios cuyo robots.txt no pudo leerse se cuentan, no se adivinan. Las ponderaciones devuelven a cada sitio su parte, pero no

pueden devolver un sitio del que no se midió nada — el 17 % de los comunicados de la población está en sitios así: el 19 % de los de los distribuidores, casi todos en dos sitios, y el 20 % de los de los gobiernos nacionales, la mitad en un sitio de noticias militares —, y cuando pudieron leerse pocos comunicados de un sitio, esos pocos valen por todos.

- Los tipos de editor son los del archivo, y algunos sitios están mal clasificados: el Gobierno de Alberta y la Oficina Australiana de Estadística entre las empresas (el 5 % del peso de ese grupo), el CERN y el SANS Internet Storm Center entre las universidades. Una página a la que se llega con y sin [www](#) cuenta en dos sitios (245 páginas medidas). Ambas cosas se dejan como se registraron.
- Las medidas son reglas. Las de la data y el contacto no alcanzaron la fiabilidad exigida en la extracción de prueba, y las hipótesis que dependen de ellas no se deciden.
- El único codificador es un modelo de lenguaje. Sus codificaciones se depositarán con los datos para que cualquiera pueda contrastarlas con las páginas.
- La extracción es imperfecta: en una de cada cinco páginas de prueba, aproximadamente, el texto extraído era solo una parte del comunicado.
- El estudio lo financia y publica una empresa que vende seguimiento de comunicados; los resultados presentan a los distribuidores en conjunto y no nombran ninguno de sus sitios, salvo el del editor.

5. Conclusiones

Ahora hay dos comunicados. La copia publicada en el sitio de un distribuidor conserva las piezas de la agencia: una data, una sección «Acerca de», un contacto al que llamar. La copia publicada en el sitio de la organización, en general, no. La diferencia de longitud está sobre todo en ese aparato, no en el relato.

Prerregistro. Las hipótesis, la población, las medidas y los análisis se registraron el 11 de octubre de 2026, antes de leer ninguna página de la población (commit 4b769927). El manifiesto de las huellas de los archivos registrados se envió a OpenTimestamps al empezar el censo; el bloque de Bitcoin que lo ancla (970.932) se minó 21 minutos después de solicitar la primera página, y el manifiesto y su prueba se depositarán con los datos. El registro, su adenda, el libro de códigos y el registro de desviaciones se publican sin modificar. <https://www.ppnworld.com/research/anatomy-of-a-press-release/registration/press-release-anatomy.md>

Disponibilidad de los datos. La tabla de anatomía por tipo de editor se publica en CSV con licencia CC BY 4.0; sus grupos conservan los nombres del archivo (wire: distribuidores; government: gobiernos nacionales; municipality: gobiernos regionales y locales; league: deportes), y sus columnas de data, contacto, subtítulo, firma y marca de cierre llevan los límites de fiabilidad expuestos arriba. Los datos por comunicado — cada comunicado de la población con su dirección, sitio, tipo de editor, resultado, ponderación y medidas, sin texto de página, citas ni datos de contacto — se depositarán en Zenodo junto con el código de cada paso; mientras tanto, el Centro de investigación los envía a petición. El registro, su adenda, el libro de códigos, las instrucciones del codificador A y el registro de desviaciones se publican sin modificar. <https://www.ppnworld.com/research/anatomy-of-a-press-release.csv>

Declaración ética. El estudio leyó páginas web públicas y no involucró a participantes humanos; no se solicitó revisión ética. Respetó el robots.txt de cada sitio, se identificó primero y reintentó una vez como navegador si

era rechazado (una fila de robustez excluye esas páginas), limitó su ritmo, y nunca resolvió una prueba antibots ni eludió un error de certificado. Los bloques de contacto contienen datos personales: solo su presencia y su tipo entran en los datos; las páginas guardadas para repetir las mediciones no se publican.

Uso de herramientas de IA. Claude (Anthropic) fue el codificador A de la validación y, usado a través de Claude Code con Claude Opus 5.5, preparó el software de recogida y análisis, borradores del texto y esta maquetación, bajo la supervisión del autor. El autor es responsable del contenido del estudio.

Correcciones. Corregido el 17 de octubre de 2026, tras la publicación del 15 de octubre de 2026; las correcciones se enumeran, fechadas, en la página web del estudio, y ninguna estimación de las cuatro hipótesis cambió. <https://www.ppnworld.com/es/research/anatomy-of-a-press-release>

Referencias

- Associated Press. (2026). The Associated Press stylebook 2026–2028 (58th ed.). Basic Books.
- Barbaresi, A. (2021). Trafilatura: A web scraping library and command-line tool for text discovery and extraction. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 122–131. <https://doi.org/10.18653/v1/2021.acl-demo.15>
- Bevendorff, J., Gupta, S., Kiesel, J., & Stein, B. (2023). An empirical comparison of web content extraction algorithms. Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. <https://doi.org/10.1145/3539618.3591920>
- Bossemma, F. G., Burger, P., Bratton, L., Challenger, A., Adams, R. C., Sumner, P., Schat, J., Numans, M. E., & Smeets, I. (2019). Expert quotes and exaggeration in health news: A retrospective quantitative content analysis. Wellcome Open Research, 4, 56. <https://doi.org/10.12688/wellcomeopenres.15147.2>
- Catenaccio, P. (2008). Press releases as a hybrid genre: Addressing the informative/promotional conundrum. Pragmatics, 18(1), 9–31. <https://doi.org/10.1075/prag.18.1.02cat>
- Cision. (2025, May 7). 2025 State of the Media report [Survey of more than 3,000 journalists in 19 markets].
- Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. Proceedings of the National Academy of Sciences, 120(30), e2305016120. <https://doi.org/10.1073/pnas.2305016120>
- Hiebert, R. E. (1966). Courtier to the crowd: The story of Ivy Lee and the development of public relations. Iowa State University Press.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics, 6(2), 65–70. <https://www.jstor.org/stable/4615733>
- Jacobs, G. (1999). Preformulating the news: An analysis of the metapragmatics of press releases. John Benjamins. <https://doi.org/10.1075/pbns.60>
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas for Navy enlisted personnel (Research Branch Report 8-75). Naval Technical Training Command.
- Koster, M., Illyes, G., Zeller, H., & Sassman, L. (2022, September). Robots Exclusion Protocol (RFC 9309). Internet Engineering Task Force. <https://doi.org/10.17487/RFC9309>
- Krippendorff, K. (2004). Content analysis: An introduction to its methodology (2nd ed.). Sage.
- Lassen, I. (2006). Is the press release a genre? A study of form and content. Discourse Studies, 8(4), 503–530. <https://doi.org/10.1177/1461445606064830>
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. Human Communication Research, 28(4), 587–604. <https://doi.org/10.1111/j.1468-2958.2002.tb00826.x>
- McLaren, Y., & Gurău, C. (2005). Characterising the genre of the corporate press release. LSP and Professional Communication, 5(1), 10–30.

Nicholls, T. (2019). Detecting textual reuse in news stories, at scale. *International Journal of Communication*, 13, 4173–4197.

PR Gun. (2026, August 25). [Analysis of 138,901 press releases from ten newswires, 2021–2026] [Press release].

Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Venetis, C. A., Davies, A., Ogden, J., Whelan, L., Hughes, B., Dalton, B., Boy, F., & Chambers, C. D. (2014). The association between exaggeration in health

related science news and academic press releases: Retrospective observational study. *BMJ*, 349, g7015. <https://doi.org/10.1136/bmj.g7015>

Weston, S. J., Ritchie, S. J., Rohrer, J. M., & Przybylski, A. K. (2019). Recommendations for increasing the transparency of analysis of preexisting data sets. *Advances in Methods and Practices in Psychological Science*, 2(3), 214–227. <https://doi.org/10.1177/2515245919848684>

Apéndice. Datos de las figuras

Cada tabla da las cifras representadas en la figura del mismo número.

Tabla A1. Dos comunicados

Elemento	Distribuidores	Sitios propios
PARA DIFUSIÓN INMEDIATA	0,3 %	6,9 %
Subtítulo	37,0 %	7,9 %
Firma de una persona	12,0 %	5,1 %
Data	76,5 %	16,7 %
Una cita	59,0 %	41,8 %
Sección «Acerca de»	63,9 %	2,9 %
Contacto con correo o teléfono	65,1 %	14,2 %
Marca de cierre (### o -30-)	2,7 %	4,0 %

Tabla A2. Hasta dónde es fiable cada medida

Medida	Acuerdo, regla registrada	Acuerdo, regla revisada	α , revisada	Puede decidir
Data	87 %	87 %	,739	solo se informa
Sección «Acerca de»	91 %	91 %	,811	decide
Contacto	74 %	89 %	,779	solo se informa
Cita	96 %	97 %	,940	decide
Directivo citado	94 %	96 %	,910	decide
Palabra de emoción	98 %	99 %	,928	decide
PARA DIFUSIÓN INMEDIATA	100 %	100 %	1,000	decide
Subtítulo	69 %	69 %	,326	solo se informa
Marca de cierre	97 %	97 %	,558	solo se informa
Firma	92 %	92 %	,162	solo se informa