

INFORME APLE2: INFORMÁTICA

Ona Domènech, M. Amor Montané, Coralie Schneider

El número inicial de candidatos a neologismos terminológicos del ámbito de la informática era de 4.532 unidades, de las que hemos seleccionado 320, que corresponden a un 7,1% del total inicial.

Candidatos	Fr. absoluta	Fr. relativa
Iniciales	4.532	100
Seleccionados	320	7,1

Tabla 1. Candidatos a neologismo terminológico

a) Grado de interferencia de otras lenguas en el ámbito: recuento de préstamos

Del total de candidatos seleccionados, 58 son préstamos. Todos proceden de la lengua inglesa (*cloud computing, copyleft, mainstreaming*), excepto uno, que es un préstamo del francés (*sousveillance*). Entre los préstamos hallamos términos monoléxicos (*pendrive, plugin, smartphone*), que son los que predominan (51 unidades), y también algunos poliléxicos (*active timeout, latent structure analysis, time to market*), aunque estos últimos son muy escasos (7 unidades).

Préstamos	Fr. absoluta	Fr. relativa	Ejemplos
Inglés	57	17,8	<i>cloud computing, copyleft, mainstreaming, pendrive, plugin</i>
Francés	1	0,3	<i>sousveillance</i>
Total	58	18,1	

Tabla 2. Préstamos

De hecho, entre los términos poliléxicos la tendencia más habitual es que solo alguno de los elementos sea un préstamo, de modo que hablamos de términos poliléxicos que contienen un préstamo. En el ámbito de la informática hallamos bastantes casos que responden a la siguiente estructura: unidad de la lengua + préstamo. Concretamente, en nuestro corpus tenemos 47 casos. El préstamo puede ser la extensión del sintagma (*potencial de big data, proveedor de hardware, lector de e-books*) o bien puede constituir su base (*software propietario, toolboxes libres, hacking negro*). El préstamo

como extensión aparece 26 veces y como base 21. En todos estos casos los préstamos proceden del inglés.

Es interesante destacar que se observa recurrencia en el uso de préstamos. Por ejemplo, las unidades *web*, *online*, *software*, *cluster* (y derivados), *hacker* (y derivados) y *plugin* son muy productivas en la formación de series:

- *web*: funciona como adjetivo (*baliza web*, *buscador web*, *cámara web*, *conferencia web*...) y como nombre (*web social*, *webs de videojuegos*);
- *online*: funciona como adjetivo (*publicidad online*, *radio online*, *rastreo online*, *revista online*...);
- *software*: funciona como nombre (*software científico*, *software privativo*, *software propietario*);
- *cluster*: funciona como nombre, curiosamente sin adaptar (*cluster de Hadoop*, *cluster de nodos*, *cluster de servidores*), y también tenemos *clustering*, de la misma familia (*clustering geodésico*);
- *hacker*: funciona como nombre, referido a personas (*hackers de sombrero blanco*, *hackers de sombrero negro*), al lado de *hacking*, de la misma familia, referido a acciones (*hacking blanco*, *hacking negro*);
- *plugin*: funciona como nombre (*plugin de agente*, *plugins de servidor*, *plugins personalizables*).

Estos préstamos que resultan recurrentes y que incluso consiguen formar familias de palabras combinándose con unidades de la lengua, también se documentan como préstamos monoléxicos. Por lo tanto, parece que se están estabilizando en la lengua: *hacker*, *on line*, *plugin*.

b) Tendencias estructurales: número de términos por estructura

De los 320 neologismos terminológicos documentados en este ámbito, 76 son términos monoléxicos y 244 son poliléxicos. El 76,3% de los primeros (58 términos) son préstamos del inglés (*plugin*, *spam*) –excepto un galicismo (*sousveillance*)–, mientras que solo 18 unidades no corresponden a formas de otras lenguas (*geomática*, *biometría*).

En cuanto a los términos poliléxicos, las estructuras predominantes son NJ, con 185 casos (*delito electrónico*, *dominio espectral*), y NPN, con 45 términos (*código de acceso*, *licencia de software*). La estructura NJ es habitual en este campo de especialidad y solos 32 casos de los 185 contienen algún préstamo (*software científico*, *vigilancia online*, *web social*). En el caso de la estructura NPN la presencia de algún préstamo es inferior: 11 términos del total de 45 contienen alguna forma de otras

lenguas, sea en la base (*plugin de agente, router de acceso*), sea en la extensión (*algoritmo de netflow, licencia de software*).

El resto de estructuras poliléxicas son muy poco habituales: tenemos 5 casos de NPNJ (*acceso de banda ancha, hackers de sombrero negro*), 3 de NPNpropio (*algoritmo de Mallat, cluster de Hadoop*), 3 de NJJ (*brecha digital internacional, procesamiento masivo paralelo, redes sociales privadas*), 1 de NN (*interfaz multilenguaje*), 1 de NJPNpropio (*modelo relacional de Codd*) y 4 casos problemáticos o dudosos pendientes de etiquetar.

Estructura	Fr. absoluta	Fr. relativa	Ejemplos
N	76	23,8	<i>plugin, spam, geomática, biometría, sousveillance</i>
NJ	185	57,8	<i>delito electrónico, dominio espectral, software científico, vigilancia online, web social</i>
NPN	45	14,1	<i>algoritmo de netflow, código de acceso, licencia de software, plugin de agente, router de acceso</i>
NPNJ	5	1,6	<i>acceso de banda ancha, hackers de sombrero blanco, hackers de sombrero negro, licencia de software libre</i>
NJJ	3	0,9	<i>brecha digital internacional, procesamiento masivo paralelo, redes sociales privadas</i>
NPNpropio	3	0,9	<i>algoritmo de Mallat, cluster de Hadoop, tuplas de Linda</i>
NJPNpropio	1	0,3	<i>modelo relacional de Codd</i>
NN	1	0,3	<i>interfaz multilenguaje</i>
D	1	0,3	<i>on line</i>
Total	320	100	

Tabla 3. Estructuras

c) Cuantificar los neologismos del ámbito: ocurrencias en Google Académico

Resultados de Google Académico	Fr. absoluta	Fr. relativa	Ejemplos
0	3	0,9375	<i>call centers asistenciales, clustering geodésico, intercomunicador web</i>
1	7	2,1875	<i>algoritmo de netflow, colector remoto, hacking negro, petaprocesamiento, toolboxes libres</i>
2-10	26	8,125	<i>delincuentes digitales, filtros interpolatorios, inactive timeout, interfaz multilenguaje, memoria de flujos</i>
11-50	29	9,0625	<i>hackers de sombrero negro, hacking blanco, interfaz esclava, lectores de e-books, protocolo de exportación</i>
51-100	32	10	<i>escritorio remoto, programa privativo, teclado externo, terminal simple, consola web</i>
101-1.000	93	29,0625	<i>microblog, diseño colaborativo, hardware libre, nickname, editor de archivos</i>
más de 1.001	130	40,625	<i>pendrive, licencia de software, big data, netbook, domótica</i>
Total	320		

Tabla 4. Neologicidad

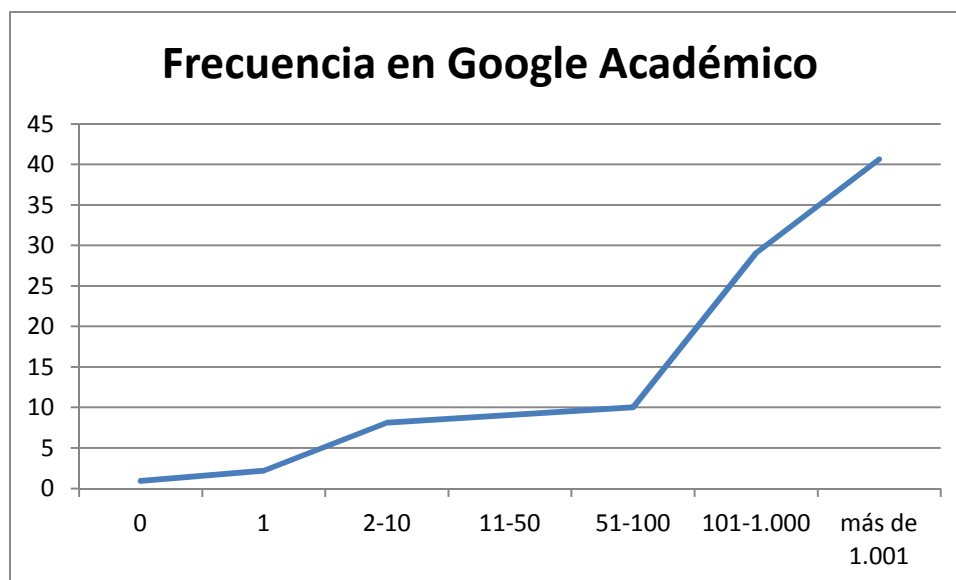


Gráfico 1. Curva del índice de neologicidad

Entre todos los neologismos documentados en el corpus (320), solo 3 no se documentan en Google Académico y, por tanto, podemos considerar que se trata claramente de neologismos: *call centers asistenciales, clustering geodésico, intercomunicador web*. Bastantes unidades se documentan 1 vez en este buscador académico (7) y el número

aumenta cuando observamos los que se documentan menos de 10 veces (26): *bytecode intermedio* (2 ocurrencias), *plugin de agente* (3 ocurrencias), *memoria de flujos* (4 ocurrencias), *ordenador esclavo* (5 ocurrencias), *zettabytes* (6 ocurrencias), *delincuente digital* (7 ocurrencias), *interfaz multilinguaje* (8 ocurrencias), *atacantes maliciosos* (9 ocurrencias). A causa de la frecuencia baja, consideramos que las unidades que se documentan menos de 10 veces se pueden considerar neologismos del área. De hecho, seguramente los términos que tienen menos de 50 ocurrencias se podrían tratar como neologismos dentro del área de la informática. Encontramos 65 que cumplen estas características (si contamos también las que no se documentan).

Otras 32 unidades presentan entre 51 y 100 ocurrencias. En este grupo encontramos unidades que nos pueden resultar relativamente familiares, si aplicamos un criterio psicológico para medir su condición neológica, por ejemplo: *escritorio remoto* (96 ocurrencias), *conferencia online* (61 ocurrencias), *edición de ficheros* (84 ocurrencias).

Entre 101 y 1000 ocurrencias hay términos como *petabyte*, *proveedor de hardware*, *conferencia web*, *portal de búsqueda*, *buscador web*, *microblog*, *toolbox*; son un total de 93, que tienen una frecuencia de uso suficientemente elevada como para dejar de considerarlos neologismos.

Finalmente, hay 130 términos que se documentan más de 1001 veces en Google Académico. El que más ocurrencias presenta es *on line* (112.000), seguido de *copyright* (111.000) y *blog* (66.400 en singular y 37.400 en plural). Otros de los términos más frecuentes en el ámbito de la informática son los siguientes: *red social*, *tweet*, *brecha digital*, *bitácora*, *usabilidad*, *dispositivo móvil*, *aula virtual*...