LA ERA DE LAS MÁQUINAS LECTORAS¹

José Antonio Millán

<http://jamillan.com>
<http://librosybitios.com>

ABSTRACT: *Machines have been reading our texts for years, both digital and printed. His reading of what we write is already a part of the way in which we gain access to the information and create it.*

KEY WORDS: Reading, machines, digitisation, search engines.

THE ERA OF READING MACHINES

RESUMEN: Las máquinas llevan años leyendo nuestros textos, tanto digitales como impresos. Su lectura de lo que escribimos forma parte ya del modo en que accedemos a la información y la creamos.

PALABRAS CLAVE: Lectura, máquinas, digitalización, buscadores.

Un fantasma recorre el universo de los textos². Un ejército de máquinas, a las que aludimos con metáforas zoológicas (*arañas*) o mecánicas (*cosechadoras*), merodean por la Red, leen nuestros textos, e incluso atisban por encima del hombro mientras escribimos.

¿Para qué lo hacen? ¿Para espiarnos? A veces... ¿Para comprendernos mejor? Ciertamente. ¿Para ayudarnos? Eso dicen...

En el universo de la World Wide Web las máquinas (los ordenadores, o mejor dicho, sus programas) saltan constantemente de página en página a través de los enlaces, escudriñan su contenido y almacenan cada palabra y cada combinación. De esa forma, cuando les preguntamos: ¿dónde se habla de Hércules?, pueden contestarnos: *aquí*

y *allá...* Pero también leen los enlaces, y así se enteran de qué creen los autores (de páginas web, de cualquier documento accesible en la Red) que tratan las páginas a donde remiten...³.

Precisamente esa lectura de enlaces es la responsable de algunos de los hallazgos más asombrosos de los buscadores: encontrar lo que no está... Por ejemplo, la búsqueda de *gentuza* en Google lleva a esta noticia⁴ (que no contiene la palabra en cuestión)⁵:

A esta mujer, víctima del Katrina, la han dejado sin un duro de la ayuda que recibió como afectada por el huracán. Todo fue porque le hicieron una foto en primer plano con su tarjeta en la mano. Al poco tiempo de publicarse en diversos medios digitales la instantánea de la Agencia France Press, realizaron una serie de compras en Internet con su número de Mastercard.

Pero hay otras formas en las que las máquinas nos leen. Por ejemplo: cuando intermedian en los artefactos (*hardware*) que usamos para escribir. En ese caso están los *softwares* espías residentes en un ordenador (como Keystroke Spy⁶), que supervisan todas las pulsaciones del teclado, y avisarán por *e-mail* cuando su usuario teclee algo de interés⁷.

II

Ocasionalmente, las máquinas también escriben (o, para no exagerar: editan, ponen en contacto textos diversos). Ocurre, por ejemplo, cuando colocan dentro de las páginas web anuncios relacionados con su tema (que es lo que hace Google Adwords⁸).

Para ello tienen que haber leído su contenido. Por ejemplo, en una página que analiza unos carteles amenazadores⁹ aparecen estos anuncios¹⁰:

- El centro del accidentado. Ayuda jurídica para víctimas de accidentes.
- Chistes de abogados.
- Problemas con alquileres.

¿Por qué? El texto contenía términos como *amenaza*, *insulto*, *violencia*, *transgresor* o *merodeador*, junto a expresiones como "me cago en sus muertos". Los insondables

algoritmos de Google Adwords han determinado que (entre los temas de publicidad que administran) los relacionados con *accidentes*, *abogados* y *problemas* eran los más pertinentes...

(Interludio filosófico)

Este tipo de comportamientos nos podría llevar a la siguiente cuestión. Sí: las máquinas leen nuestras páginas web, pero, ¿las entienden? En realidad, esto es una variante del Test de Turing (Turing, 1950). Como se recordará, en dicha prueba un humano conectado a un terminal exclusivamente textual (tipo chat) debe determinar, sólo a través del diálogo, si al otro lado hay una máquina o un ser humano.

Uno dice "¡Gentuza!", y el buscador contesta: "Sí, como esos que estafaron a una víctima del Katrina...". Uno escribe "amenaza, violencia, transgresor", y los anuncios corean: "abogados, accidentes, problemas". ¿Nos están entendiendo las máquinas? Bueno: lo suficiente como para echarnos una mano. Y el éxito de los buscadores y de los programas de anuncios contextuales parecen indicar que lo logran...

Hay en marcha sistemas todavía más sofisticados. Por ejemplo: un programa que analiza, en un foro sobre valores bursátiles, cuál es la opinión generalizada sobre cuáles van a subir y cuáles a bajar. Es el Community Sentiment de Yahoo¹¹. Un análisis de este estilo exige manejar un número considerable de variables semánticas y pragmáticas.

Pero hay que tener en cuenta que las máquinas no sólo están leyendo nuestras páginas web: también leen nuestros diarios (o blogs), nuestro propio correo (en sistemas como Gmail¹²), nuestras notas personales (Google Bloc de notas¹³), nuestra escritura manuscrita en una agenda electrónica (a través de programas como PenReader¹⁴). Si además parece que están enterándose, ¿no supone esto un problema?

La verdad es que sí, pero también nuestros secretarios (o secretarías) leen nuestra correspondencia, y a ellos dictamos nuestras cartas¹⁵. Digamos que quien confía en ayudas externas (ya sean de carne y hueso o de código) debe atenerse a las consecuencias...

III

Y en este momento nos surge un tema de especial interés. Si las máquinas nos leen, ¿no habrá que tenerlas en cuenta cuando escribimos? La respuesta es claramente que sí: el autor o editor de cualquier material en la Web tiene que favorecer que le lean las máquinas, so pena de comprometer su propia difusión.

Un ejemplo particularmente ilustrativo es el de las licencias Creative Commons. Cada una de ellas tiene tres versiones:

- el resumen, legible por humanos¹⁶. Dice cosas como:

Usted es libre de: copiar, distribuir y comunicar públicamente la obra.

- el código legal, legible por abogados¹⁷; éste es su comienzo:

LA OBRA (SEGÚN SE DEFINE MÁS ADELANTE) SE PROPORCIONA BAJO LOS TÉRMINOS DE ESTA LICENCIA PÚBLICA DE CREATIVE COMMONS ("CCPL" O "LICENCIA"). LA OBRA SE ENCUENTRA PROTEGIDA POR LA LEY ESPAÑOLA DE PROPIEDAD INTELECTUAL Y/O CUALESQUIERA OTRAS NORMAS RESULTEN DE APLICACIÓN. QUEDA PROHIBIDO CUALQUIER USO DE LA OBRA DIFERENTE A LO AUTORIZADO BAJO ESTA LICENCIA O LO DISPUESTO EN LAS LEYES DE PROPIEDAD INTELECTUAL.

- el código digital, legible por máquinas¹⁸

```
<rdf:RDF xmlns="http://creativecommons.org/ns#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"><License rdf:about="http://creativecommons.org/licenses/by/2.5/es/"><permits rdf:resource="http://creativecommons.org/ns#Reproduction"/><permits rdf:resource="http://creativecommons.org/ns#Distribution"/><requires rdf:resource="http://creativecommons.org/ns#Notice"/><requires rdf:resource="http://creativecommons.org/ns#Attribution"/><permits rdf:resource="http://creativecommons.org/ns#DerivativeWorks"/></License></rdf:RDF>
```

Este caso refleja bien las complejidades de la autoría/edición en la Web. Una norma general de redacción es tener en cuenta al destinatario de un texto para ajustar su nivel, y eso justifica la diferencia entre el "código legal" y el "resumen": sus receptores son, respectivamente, el abogado y el lego en la materia. Pues bien, además de estas cuestiones (comunes con cualquier texto, esté o no en la Red), tenemos otra añadida.

El código digital está destinado a ser leído por sistemas automáticos. En el caso de Creative Commons, se ha incluido para informar a los buscadores que quieran localizar contenidos con determinados tipos de licencia. Las personas no tienen por qué entenderlo, y ni siquiera leerlo: el texto no está visible en la página.

Al igual que este código, en las páginas web hay muchos elementos cuyos destinatarios son las máquinas: los ficheros *robots.txt*¹⁹, los metatags, y las palabras que se incluyen para forzar el *spamdexing* (Millán, 2000). El *webmaster* (de ingenio aunque tramposo) que llena de palabras prometedoras la parte inferior de su página, y las escribe en el mismo color del fondo para que no se lean en la pantalla, sabe bien que no busca lectores humanos... En el terreno de la edición científica, los metadatos son un elemento básico del texto.

Y una última, pero importante consecuencia, para aquellos que escriben o editan en la Web: cada enlace es un voto a una página. Y mediante el texto específico que enlazamos estamos diciendo algo sobre la página de destino no sólo a nuestros lectores humanos, sino, sobre todo, a las máquinas.

IV

Además de las búsquedas, que antes veíamos, las máquinas también están leyéndonos para ayudarnos con distintas tareas...

Los servicios de alertas, como Yahoo Alerts²⁰, rastrean la prensa y otras páginas web para avisarnos de cuándo aparece alguna de las palabras clave que les hemos indicado. Resulta muy útil para tener controlada a una empresa rival, conocer los movimientos de una determinada persona, o sencillamente, ver qué dicen de nosotros (*ego surfing*).

Los lectores o agregadores de RSS (que suministran el contenido nuevo de sitios web, como Google Reader²¹) leen los sitios que escogemos para enviarnos sus titulares junto con una porción mayor o menor de texto.

Los detectores de plagios, como Damocles²², comparan el texto que les sometamos con muchos otros dispersos por la Web, con el objeto de determinar si se han utilizado (sin citar) partes de otras obras.

Los sintetizadores de voz (como SodelsCot²³) leen los textos que les proponemos. A veces su lectura no es muy buena, como los lectores que, ante un texto sin puntuación en la antigüedad clásica, leían "mutilando los pensamientos y pronunciando imperfectamente" (Aulo Gelio, citado en Saenger, 1997, 11), pero en ocasiones leen con mucha fidelidad. Tenemos testimonios de cómo los *lectores* de la antigüedad clásica servían, entre otras cosas, de ayuda a personas con problemas en la visión (Starr, 1991), y ése es uno de los usos actuales de los conversores texto-habla.

Sin olvidar a los programas traductores (como SoftCatalà²⁴, del catalán al castellano y viceversa), que leen nuestros textos para traducirlos.

Y por último, el sistema de espionaje anglosajón ECHELON²⁵ (gobernado por Estados Unidos, Canadá, Gran Bretaña, Australia, y Nueva Zelanda) o el sistema Carnivore²⁶ del gobierno de los Estados Unidos (FBI) escrutan las comunicaciones (correos electrónicos, por ejemplo) a la búsqueda de términos o nombres. Lo bajo de sus fines no debe hacernos olvidar la magnitud de la tarea que afrontan.

V

Hasta aquí nos hemos movido en un dominio, el digital, que posibilita que las máquinas nos lean directamente. En la página web los humanos vemos formas, desciframos signos y por último leemos palabras. Las máquinas también las leen, pero no por el dibujo que pintan en la pantalla (el cual puede cambiar según las preferencias de nuestro navegador), sino porque acceden al código que les representa. Por ejemplo: la H tiene el código hexadecimal 48, y el fragmento de código

%48%E9%72%63%75%6C%65%73

se leería *HERCULES*²⁷. El sintetizador de voz que lee el documento de procesador de textos y el programa espía que supervisa nuestro correo acceden también al código de las letras.

En caso de contradicción entre el mensaje visual y el código los humanos seguimos, por supuesto, lo que nos dicen nuestros ojos. Por eso en los años ochenta, para burlar la censura que supervisaba las BBS (tabloneros de anuncios electrónicos²⁸), los usuarios escribían sustituyendo letras por otros signos con los que tenían cierto parecido (pero que no compartían su código)²⁹. Por ejemplo, para escribir *similar* se usaba la siguiente secuencia de caracteres:

51m1L4R

Lamentablemente, ya hay programas que leen también estas escrituras...

VI

Pero aparte de este acceso directo al código, las máquinas están leyendo cada vez más las publicaciones impresas.

La película *Los tres días del cóndor* de Sydney Pollack (1975)³⁰ se iniciaba con unas oficinas de la CIA en las que una máquina iba pasando páginas de los periódicos bajo el ojo escrutador de una cámara. Han transcurrido más de un cuarto de siglo, y los programas ya son muy buenos leyendo impresos.

Pero hay dos formas en que las máquinas pueden tratar nuestros textos impresos. Una es fotografiando sencillamente el texto, es decir, describiendo píxel a píxel la traza de sus letras.

La Fig. 1 muestra las tres letras iniciales de la palabra *Hercules* en el facsímil JPEG de la primera edición del *Quijote* en la Biblioteca Virtual Cervantes³¹.

Figura 1. Descomposición en píxeles de un fragmento de imagen con tres letras

Describir la forma de los signos alfabéticos no es un comportamiento muy sofisticado. Es lo que hacía en el siglo V a. C. un pastor iletrado con las letras griegas que constituían el nombre TESEO (ΘΗΣΕΥΣ) en la tragedia perdida de Eurípides³²:

"No soy habilidoso en las letras, pero diré sus formas y claros signos. Hay un círculo, como trazado a compás; éste tiene una clara marca en el centro. La segunda letra tiene primero dos líneas, y otra las separa en el centro. La tercera es como un rizo de cabello, mientras que la cuarta, de nuevo, tiene una línea hacia arriba y tres que se apoyan en ella. La quinta no es fácil de explicar: hay dos líneas separadas, pero se encuentran en un soporte. La última letra es como la tercera."

El público ático del siglo V a. C. (ya parcialmente alfabetizado) podía reconocer las letras por las formas transmitidas. Pero el lector actual puede, más cómodamente, leer en la alineación de píxeles (Fig. 2).

industria de Hercu-

Figura 2. Palabras tal y como aparecen en la pantalla (ampliación)

Ahora bien, ¿sabía leer el pastor de Eurípides? ¿Sabe leer el escaneador de páginas ante el que desfilaron las páginas del *Quijote*? Claramente, no. El portador de formas de letras no lee.

VII

Para que las máquinas lean de verdad hay que ir un paso más allá: Google Libros³³ (por poner un ejemplo bien conocido) está digitalizando libros de las bibliotecas. Pero además de fotografiar sus páginas les aplica un programa de reconocimiento óptico de caracteres (OCR).

A través de ese procedimiento, la máquina reconocerá la forma que "tiene primero dos líneas, y otra las separa en el centro" como una hache mayúscula (si el texto está en alfabeto latino) o como una eta mayúscula (si está en griego). Y así sucesivamente. Por ejemplo, sometamos el

archivo con las palabras del *Quijote* de la Fig. 2 a un OCR accesible por línea³⁴. Nos dará este resultado:

indufhia de Hercu-

Como vemos, puede haber errores. En este caso, la tipografía del XVII tiene ligaduras (como la que une s y t) que el programa no reconoce: enseguida veremos cómo lidiar con ellos. Pero en casos más modernos o claros la máquina puede leer todo el texto satisfactoriamente.

Al final del proceso, el OCR habrá extraído del "cuerpo" del libro (el papel y la tinta) su "alma", el conocimiento de la secuencia de caracteres que lo constituyen: la "acertada disposición del impresor y corrector"³⁵...

El OCR hace que los impresos se fundan en el *continuum* digital del que ya formaban parte las páginas web y otros archivos accesibles por Internet. Y cuando preguntemos: ¿en qué obras se encuentra la palabra "Hércules"?, acudirán a respondernos no sólo las páginas web, sino también las de los libros.

Al entrar en el universo de la imprenta, las máquinas pueden hacer averiguaciones suplementarias: por ejemplo, sus programas pueden rastrear qué pasajes de cada obra están citados en otros libros. Así, podemos enterarnos, por ejemplo, de que el famoso pasaje de la *Exposición al Libro de Job* de Fray Luis de León que reza³⁶:

"las escrituras que por los siglos duran nunca las dicta la boca, del alma salen, adonde por muchos años las compone y examina la verdad y el cuidado."

está citado en 13 obras más (de las que están en el fondo digitalizado por Google)³⁷.

VIII

Por último, veamos cómo los humanos estamos enseñando a las máquinas a perfeccionar su lectura.

Captcha³⁸ es el sistema mediante el que un sitio web con intervención del público se defiende de los programas que

se dedican a introducir *spam*, proponiendo a los usuarios que tecleen el texto de una secuencia de letras deformada o borrosa que se les ofrece (Fig. 3).

Escriba los caracteres que vea en la imagen de arriba.

Figura 3. Captura del captcha de un blog

Esta tarea exige (al menos por el momento) un ser humano, y en ese sentido es un test de Turing³⁹.

Pues bien: ha nacido reCaptcha⁴⁰ (Fig. 4). Su peculiaridad es que el texto que propone para interpretación proviene del escaneado de libros: son palabras que el reconocimiento óptico de caracteres no acierta a interpretar (como *industria*, que veíamos anteriormente). El programa de OCR detecta una palabra problemática y reCaptcha la ofrece como clave de acceso, emparejada con otra palabra cuya interpretación se conoce (y que sirve de control).

Las palabras dudosas se ofrecen cierto número de veces, hasta que la lectura se confirma.

ReCaptcha está funcionando por el momento como una ayuda para las digitalizaciones del Open-Access Text Archive⁴¹. Teniendo en cuenta que cada día se resuelven 60 millones de Captchas, que llevan de media 10 segundos, su suma daría 150.000 horas de trabajo al día, que reCaptcha pondría al servicio de la digitalización de libros.

Figura 4. ReCaptcha

y IX

... Y éste es el panorama: ejércitos de autómatas rastreando el ciberespacio y hordas de máquinas leyendo las bibliotecas. Programas que descifran letras y humanos que les ayudan, porque así se ayudan a sí mismos.

Más círculos: humanos que preguntan a las máquinas dónde están las cosas que les interesan, para luego escribir

textos que leerán las máquinas para a su vez contarle a otros humanos de qué tratan.

Este espacio simbiótico de personas y máquinas, este *continuum* digital de textos y códigos es el caldo de cultivo de la cultura actual.

NOTAS

- 1 Original entregado en diciembre de 2007.
- 2 Este texto comenzó como una conferencia en el Seminario *Litterae* de septiembre del 2007. Gracias a Antonio Castillo, Vanessa de Cruz y Emilio Torné por su invitación a participar. Gracias a Javier Candeira por su ayuda.
Está sujeto a la licencia de Creative Commons Reconocimiento-No comercial-Compartir bajo la misma licencia 2.5 España (<http://creativecommons.org/licenses/by-nc-sa/2.5/es/deed.es>).
- 3 El mejor texto sobre El *modus operandi* de los buscadores (o sea, de Google) sigue siendo el de Javier Candeira, "La Web como memoria organizada", en *Revista de Occidente* (Madrid), marzo de 2001 (versión electrónica en http://jamillan.com/para_can.htm).
- 4 Búsqueda realizada el 14 de noviembre del 2007. *20 minutos*, "Revientan la tarjeta de una víctima del Katrina tras salir su foto en los medios", 14 de octubre del 2005 (<http://www.20minutos.es/noticia/56158/0/tarjeta/victima/katrina/>).
- 5 El hecho de que Google se dejara guiar por el contenido de los enlaces llevó al fenómeno de la "Google bomb", desgraciadamente hoy inhabilitada por el buscador (http://en.wikipedia.org/wiki/Google_bomb).
- 6 <http://www.spytech-web.com/keys-troke-spy.shtml>.
- 7 Los programas de espionaje también pueden registrar el ruido del teclado, y a partir de él deducir qué se está escribiendo (<http://www.securitysolutionsmagazine.com/Articles/SSM38.pdf>, p. 22).
- 8 <http://adwords.google.com/>.
- 9 <http://jamillan.com/florarma.htm>.
- 10 A 13 de noviembre del 2007: los anuncios pueden cambiar cada vez que se accede a la página.
- 11 <http://finance.yahoo.com/newfp>.
- 12 <http://www.gmail.com>.
- 13 <http://www.google.com/notebook/>.
- 14 <http://www.penreader.com/pocket-pc-software/PenReader.html>.
- 15 El rey de Aragón Pedro IV el Ceremonioso (siglo XIV) instituyó por primera vez la comunicación escrita para un gobierno peninsular, y también el archivo de los documentos reales. Se rodeó de una red de secretarios y escribanos, e incluso uno de ellos debía dormir por la noche en sus aposentos, "por si le venía la necesidad de escribir". Pero cuando quiso mantener secreto un tratado con el rey de Castilla lo escribió de su puño y letra. Véase Gimeno Blay (2006).
- 16 <http://creativecommons.org/licenses/by/2.5/es/deed.es>.
- 17 <http://creativecommons.org/licenses/by/2.5/es/legalcode.es>.
- 18 Contenido en el código fuente de <http://creativecommons.org/licenses/by/2.5/es/deed.es>.
- 19 <http://www.webrecursos.com/pages/promo/promobot.htm>. Curiosamente, una de las funciones de los *robots.txt* puede ser la indicación "No me leas".
- 20 <http://alerts.yahoo.com>.
- 21 <http://www.google.com/reader>.
- 22 <http://viper.csse.monash.edu.au/damocles/about/>.
- 23 <http://www.sodels.com/>.
- 24 <http://www.softcatala.org/traductor/>.
- 25 Véase el insustituible libro *Echelon. La red de espionaje planetario* (Barcelona, Melusina, 2007). Para un resumen, el artículo de la *Wikipedia*: <http://es.wikipedia.org/wiki/ECHELON>.

Recibido: 3 de diciembre de 2007
Aceptado: 23 de diciembre de 2007

- 26 Para un resumen, el artículo de la *Wikipedia*: <http://es.wikipedia.org/wiki/Carnivore>.
- 27 *ASCII table and description*: <http://www.asciitable.com/>.
- 28 Véase el artículo de la *Wikipedia* "Bulletin Board System", http://es.wikipedia.org/wiki/Bulletin_Board_System.
- 29 Véase el artículo de la *Wikipedia* "Leet speak" (http://es.wikipedia.org/wiki/Leet_speak). Es también el procedimiento mediante el que los correos *spam* ocultan determinados nombres de marca al *software* censor, pero no al humano inquisitivo: "\!/\\6R\\/".
- 30 <http://www.imdb.com/title/tt0073802/>.
- 31 P I, Capítulo I: <http://www.cervantesvirtual.com/servlet/SirveObras/cerv/12371067559018288532624/ima0027.htm>.
- 32 Traduzco del texto que presenta Niall W. Slater (2002, 118-119).
- 33 <http://books.google.es/>. Observo que estoy citando muchos productos de Google, aunque esto sólo demuestra una cosa: que esta compañía está en cabeza en el desarrollo y aprovechamiento comercial de las máquinas lectoras...
- 34 Tesseract OCR, el mismo que usa Google (<http://asv.aso.ecei.tohoku.ac.jp/tesseract/>).
- 35 En palabras del impresor del XVII Alonso Víctor de Paredes. Sobre las relaciones entre continente y contenido en el libro véase mi artículo "Libro: el sarcófago del texto", por aparecer en *Trama y texturas* (Madrid), n.º 3.
- 36 <http://books.google.es/books?id=VTQ7IXB-ZmIC&printsec=titlepage#PPA226,M1>.
- 37 http://books.google.es/books?qtid=cb7e7834&id=VTQ7IXB-ZmIC&as_brr=0. Se trata del recurso conocido como "Pasajes populares" de Google Libros.
- 38 <http://es.wikipedia.org/wiki/Captcha>.
- 39 Por cierto: la versión para disminuidos visuales consiste en un audio que presenta, en medio de ruido de conversaciones, una serie de cifras. He aquí un ejemplo: <https://www.blogger.com/captcha?type=AUDIO&captchaKey=1ql96hbcw21t0>.
- 40 <http://recaptcha.net/learnmore.html>.
- 41 <http://www.archive.org/details/texts>.

BIBLIOGRAFÍA

- Gimeno Blay, Francisco M. (2006): *Escribir, reinar. La experiencia gráfico-textual de Pedro IV el Ceremonioso (1336-1387)*, Madrid, Abada editores.
- Millán, José Antonio (2000): "El libro de mil millones de páginas. La ecología lingüística de la Web", en *Revista de Libros (Madrid)*, n.º 45, septiembre, y versión web en <http://jamillan.com/ecoling.htm>.
- Saenger, Paul (1997): *Space Between Words. The Origins of Silent Reading*, California, Stanford University Press.
- Slater, Niall W. (2002): "Dancing the Alphabet. Performative Literacy on the Attic Stage", en Ian Worthington y John Miles Foley (eds.), *Epea and Grammata: Oral and Written Communication in Ancient Greece*, Leiden/Boston, Brill.
- Starr, Raymond J. (1991): "Reading Aloud: Lectores and Roman Reading", en *The Classical Journal*, vol. 86, n.º 4, April-May 1991, pp. 337-343.
- Turing, Alan M. (1950): "Computing Machinery and Intelligence", *Mind*, 49, pp. 433-460, <http://cogprintsorg/499/0/turing.html>.