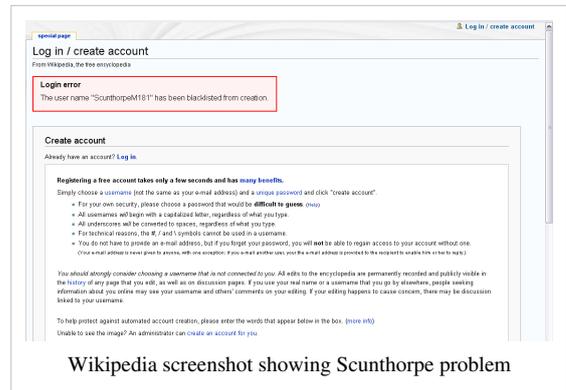


Scunthorpe problem

The **Scunthorpe problem** occurs when a spam filter or search engine blocks e-mails, forum posts or search results because their text contains a string of letters that are shared with an obscene word. While computers can easily identify strings of text within a document, broad blocking rules may result in false positives, causing innocent phrases to be blocked.

Origin and history

The problem was named after an incident in 1996 in which AOL's profanity filter prevented residents of the town of Scunthorpe, North Lincolnshire, England from creating accounts with AOL, because the town's name contains the substring *cunt*. Years later, Google's filters apparently made the same mistake, preventing residents from searching for local businesses that included *Scunthorpe* in their names.



Wikipedia screenshot showing Scunthorpe problem

Other examples

Mistaken decisions by obscenity filters include:

Refused web domain names and email addresses

- In April 1998, Jeff Gold attempted to register the domain name `shitakemushrooms.com`, but he was blocked by an InterNIC filter prohibiting the "seven dirty words" which was active between 1996 and the transfer of control to ICANN. Wikipedia:Manual of Style/Dates and numbers#Chronological items. (*Shitake* is from the Japanese name for the edible fungus *Lentinula edodes*.)
- In 2000, a Canadian television news story on web filtering software found that the website for the Montreal Urban Community (Communauté urbaine de Montréal, in French) was entirely blocked because its domain name was its French acronym *CUM* (`www.cum.qc.ca`); "cum" (among other meanings) is English-language slang for semen.
- In February 2004, in Scotland, Craig Cockburn reported that he was unable to use his surname (pronounced "Coburn") with Hotmail. Separately he had problems with his workplace email because of the name of a pharmaceutical, that was often the subject line used on spam or scam emails, being *cialis*, occurring within his job title of *software specialist*. He was told by Hotmail to spell his name C0ckburn (with a zero instead of the letter "o"); Hotmail later reversed the ban. In 2010 he had a similar problem registering on the BBC site where again the first four characters of his surname caused a problem for the content filter.
- In February 2006, Linda Callahan, a resident of Ashfield, Massachusetts, was initially prevented from registering her name with Yahoo! as an e-mail address as it contained the substring *Allah*. Yahoo! later reversed the ban.
- In July 2008, Dr. Herman I. Libshitz was initially unable to get the e-mail address he wanted from Verizon because it contained the substring *shit*. A spokesperson commented: "As a general rule (since 2005) Verizon doesn't allow questionable language in e-mail addresses, but we can, and do, make exceptions based on reasonable requests. The one from Dr. and Mrs. Libshitz certainly is reasonable and we regret the inconvenience and frustration they've been caused."

Blocked web searches

- In the months leading up to January 1996, some web searches for Super Bowl XXX were being filtered, because the Roman numeral for the game and the site (XXX) is also used to identify pornography.^[1]
- The filter of the free wireless service of the town of Whakatane in New Zealand blocked searches involving the town's own name, because the phonetic analysis used by the filter deemed the "whak" to sound like *fuck*. The town name is Maori, and in the Maori language "wh" is most commonly pronounced as "f".
- Gareth Roelofse noted in 2004, "We found many library Net stations, school networks and Internet cafes block sites with the word 'sex' in the domain name. This was a challenge for RomansInSussex.co.uk because its target audience is school children."
- In July 2011, web searches in China on the name Jiang were blocked following claims on the Sina Weibo microblogging site that former president Jiang Zemin had died. Since the word "Jiang" meaning "river" is written with the same Chinese character, searches related to rivers including the Yangtze (Cháng Jiāng) produced the message "According to the relevant laws, regulations and policies, the results of this search cannot be displayed."

Blocked emails

- In 2001, Yahoo! Mail erroneously changed words, including *medireview* in place of *medieval*. That year Yahoo! introduced an email filter which automatically replaced Javascript-related strings with alternate versions, to prevent the possibility of JavaScript viruses in HTML email. The filter would hyphenate the terms "Javascript", "Jscript", "Vbscript" and "Livescript", and replaced "eval", "mocha" and "expression" with the similar but not quite synonymous terms "review", "espresso" and "statement", respectively. Assumptions were involved in the writing of the filters: no attempts were made to limit these string replacements to `script` sections and attributes, or to respect word boundaries, in case this would leave some loopholes open.
- In October 2004, it was reported that the Horniman Museum in London was failing to receive some of its e-mail because filters mistakenly decided that its name was a version of the words *horny man*.
- Problems can occur with the words *socialism*, *socialist*, and *specialist* because they contain the substring *Cialis*, the brand name for an erectile dysfunction medication commonly advertised in spam e-mails. Blocking of the word *specialist* is liable to block emailed résumés and curricula vitae and other material including job descriptions.

Blocked for word with two meanings

- In May 2006, Ray Kennedy from Manchester in the UK found that e-mails that he had written to his local council to complain about a planning application had been blocked as they contained the word *erection* when referring to a structure.^[2]
- In October 2004, e-mails advertising the pantomime Dick Whittington sent by a teacher from Norwich in the UK were blocked by school computers because of the use of the word *Dick*, sometimes used as slang for *penis*.^[3]
- In February 2003, members of Parliament at the British House of Commons found that a new spam filter was blocking e-mails to them. It blocked e-mails containing references to the Sexual Offences Bill then under debate, and some messages relating to a Liberal Democrat consultation paper on censorship.^[4] It also blocked e-mails sent in Welsh because it did not recognise the language.^[5]
- Resumes of *magna cum laude* graduates have been blocked by spam filters because of inclusion of the word *cum*, which is Latin for *with* (in this usage), but is sometimes used as slang for *semen* in English usage. Wikipedia:Link rot
- Blocked e-mails and web searches relating to *The Beaver* (based in Winnipeg) caused the publisher to change its name to *Canada's History* after 89 years of publication. Publisher Deborah Morrison commented: "Back in 1920, *The Beaver* was a perfectly appropriate name. And while its other meaning is nothing new, its ambiguity began to pose a whole new challenge with the advance of the Internet. The name became an impediment to our growth".

- A councillor in Dudley found an email flagged for profanity by his council's security software after mentioning the Black Country dish faggots.
- In 2007, the Royal Society for the Protection of Birds blocked ornithological terms such as *cock* (male bird) and *tit*, Shag and Booby from its discussion forums.
- Residents of Penistone in South Yorkshire.
- Lightwater in Surrey suffered similarly because its name contains the substring *twat*.
- Residents of Clitheroe (Lancashire, England) have been repeatedly inconvenienced because their town's name includes the substring *clit*, which (among other meanings) is short for "clitoris".

News articles damaged

- In June 2008, a news site run by the American Family Association filtered an Associated Press article on sprinter Tyson Gay, replacing instances of "gay" with "homosexual", thus rendering his name as "Tyson Homosexual".
- The word "ass" may be replaced by "butt", resulting in "clbuttic" for "classic" and "buttbuttinat" for "assassinate".
- In December 2011, it was reported that software used by Virgin Media had filtered words including "Arsenal" (for "arse"), and "Canal" (for "anal").

Blocked pages

- In January 2014, files used in the online game League of Legends were reported as being blocked by the ISP filters due to the names containing the letters used in the word "sex".

References

- [1] "E-Rate And Filtering: A Review Of The Children's Internet Protection Act" (<http://www.gpo.gov/fdsys/pkg/CHRG-107hhrg72836/pdf/CHRG-107hhrg72836.pdf>). Congressional Hearings. General. Energy and Commerce, Subcommittee on Telecommunications and the Internet. April 4, 2001.
- [2] BBC E-mail filter blocks 'erection' (http://news.bbc.co.uk/2/hi/uk_news/england/manchester/5031156.stm) 30 May 2006
- [3] Sam Jones Panto email falls foul of filth filter (http://www.guardian.co.uk/uk_news/story/0,3604,1326590,00.html) *The Guardian* 14 October 2004
- [4] BBC E-mail vetting blocks MPs' sex debate (http://news.bbc.co.uk/1/hi/uk_politics/2723851.stm) 4 February 2003
- [5] BBC Software blocks MPs' Welsh e-mail (<http://news.bbc.co.uk/1/hi/wales/2727133.stm>) 5 February 2003

External links

- Article on autocorrecting (<http://www.guardian.co.uk/technology/blog/2008/jun/30/computerautocorrectssurname>) from *The Guardian*

Article Sources and Contributors

Scunthorpe problem *Source:* <http://en.wikipedia.org/w/index.php?oldid=591778208> *Contributors:* Adrian J. Hunter, Akaihola, AlmostReadytoFly, Angmering, AnonMoos, Anthony Appleyard, Arwel Parry, Auric, BD2412, BDD, Before My Ken, Belowenter, Biscuittin, Blenheimears, Canadian Skeptic, CecilWard, Cenarium, Chovain, Cnilep, Counsell, Crasshopper, Curly Turkey, Czech is Cyrillized, Dale Arnett, Damian Yerrick, Davidhorman, Dgies, Dhartung, DouglasCalvert, Drmies, Duncan.Hull, EamonnPKeane, Estoy Aquí, Explicit, Eyesnore, Father Goose, Florian Sening, Fubar Obfusco, Furrykef, Gadfium, Graham87, Ground Zero, Haeinous, Harodotus, Hvn0413, Ianmacm, IceCreamAntisocial, IllaZilla, ImperatorExercitus, Imroy, Inanygivenhole, IstvanWolf, Izn0, JasonAQuest, Jeremyleader, Jim Henry, John, John of Reading, K6ka, Kay Dekker, Kevin, Lemnaminor, Leuce, Lisapollison, Loco830, Lotje, ManekiNeko, MartinPoulter, Martinevans123, Medinoc, Mogism, Mortense, Neitherday, NetRoller 3D, Nick Cooper, NuclearWarfare, Nunquam Dormio, Ohconfucius, Owain, Paul Magnussen, Peyre, Philip Cross, Philipwhiuk, Phlegat, Pperos, Prari, ProhibitOnions, Quiddity, RegenerateThis, Reinyday, Rich Farmbrough, Richardphx, Rickyrab, Rjanag, RockMFR, Roscelese, Saudade7, Scott Martin, Secretlondon, Sfan00 IMG, Siliconglen, Sillyfolkboy, SlamDiego, Slashme, Snigbrook, Solomonfromfinland, Sumnjam, The Anome, The Storm Surfer, Theis101, Tony Sidaway, TonyW, Triangular, Tyvole, Uncle G, UnicornTapestry, Urutapu, Vegaswikian, Vexorg, Viciarg, ViperSnake151, WadeSimMiser, XP1, 94 anonymous edits

Image Sources, Licenses and Contributors

File:Scunthorpe problem.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Scunthorpe_problem.png *License:* GNU General Public License *Contributors:* User:Snigbrook; Wikipedia editors (see and [MediaWiki developers](#))

License

Creative Commons Attribution-Share Alike 3.0
[//creativecommons.org/licenses/by-sa/3.0/](http://creativecommons.org/licenses/by-sa/3.0/)