



PAN 2010

Uncovering Plagiarism, Authorship, and Social Software Misuse

Bauhaus-Universität Weimar — Martin Potthast, Benno Stein
Andreas Eiselt, Teresa Holfeld

Universidad Politécnica de Valencia — Alberto Barrón-Cedeño, Paolo Rosso

University of the Aegean — Efstathios Stamatatos

Bar-Ilan University — Moshe Koppel

<http://pan.webis.de>

Who we are...



Benno Stein



Paolo Rosso



Efsthathios Stamatatos



Moshe Koppel



Martin Potthast



Alberto Barrón-Cedeño



Andreas Eiselt



Teresa Holfeld

PAN Overview

PAN Overview

Mission

- ❑ Plagiarism Detection
 - text plagiarism within and across languages, multimedia plagiarism
 - text reuse, paraphrasing, information flow, meme tracking
 - near-duplicates, high-similarity search, fingerprinting, hash-based search

PAN Overview

Mission

❑ Plagiarism Detection

- text plagiarism within and across languages, multimedia plagiarism
- text reuse, paraphrasing, information flow, meme tracking
- near-duplicates, high-similarity search, fingerprinting, hash-based search

❑ Authorship Identification

- models for authorship verification, authorship attribution, and writing style
- models to capture personal traits and sentiment
- text forensics, ghostwriting, intrinsic plagiarism detection

PAN Overview

Mission

- ❑ Plagiarism Detection
 - text plagiarism within and across languages, multimedia plagiarism
 - text reuse, paraphrasing, information flow, meme tracking
 - near-duplicates, high-similarity search, fingerprinting, hash-based search
- ❑ Authorship Identification
 - models for authorship verification, authorship attribution, and writing style
 - models to capture personal traits and sentiment
 - text forensics, ghostwriting, intrinsic plagiarism detection
- ❑ Social Software Misuse
 - serial sharing, lobbying, spam
 - trolling, stalking, Wikipedia vandalism
 - social trust, anonymity and de-anonymization

PAN Overview

Mission & Tasks

- ❑ Plagiarism Detection
 - text plagiarism within and across languages, multimedia plagiarism
 - text reuse, paraphrasing, information flow, meme tracking
 - near-duplicates, high-similarity search, fingerprinting, hash-based search
- ❑ Authorship Identification
 - models for authorship verification, authorship attribution, and writing style
 - models to capture personal traits and sentiment
 - text forensics, ghostwriting, intrinsic plagiarism detection
- ❑ Social Software Misuse
 - serial sharing, lobbying, spam
 - trolling, stalking, Wikipedia vandalism
 - social trust, anonymity and de-anonymization





Plagiarism is the practice of claiming, or implying, original authorship of someone else's written or creative work, in whole or in part, into one's own without adequate acknowledgment.



Plagiarism is the practice of claiming, or implying, original authorship of someone else's written or creative work, in whole or in part, into one's own without adequate acknowledgment.

[Wikipedia: Plagiarism, 2009]


... better technology nowadays ;-)


YAHOO!
ITALIA

Web Immagini Video Pagine Gialle Shopping Notizie Altro ▾


"my favorite topic" Cerca


Cerca: ☒ nel Web ☐ nei siti in italiano

 Search Pad

 Filtro Famiglia Disattivato

202.000 risultati per
"my favorite topic":

 Mostra tutti

 Tripadvisor

[Where Did My Favorite Topic Go? - READ ONLY](#)
You are here: Forums Future Vehicles Where Did **My Favorite Topic** Go? Where Did **My Favorite Topic** Go? - READ ONLY. 40 messages, Last post on Oct 04, 2005 at 5:45 AM
... ~ by kirstie_h (40 comments)
townhall-talk.edmunds.com/direct/view.ef15b84/0 - 51k - [Cache](#)

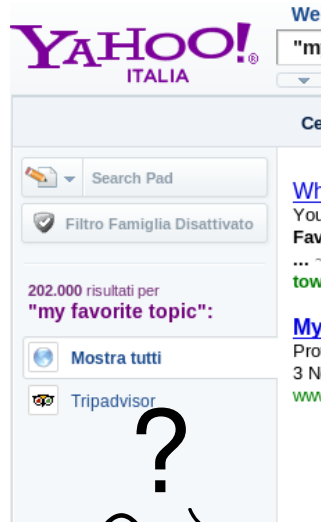
[My favorite topic..](#) - [Traduci](#)
Profile Log in to check your private messages Log in. **My favorite topic..** Goto page 1, 2, 3 Next ... 03 am Post subject: **My favorite topic..** Why i hate OldSpice ...
www.astrowars.com/forums/viewtopic.php?p=717474 - 65k - [Cache](#)

[My favorite topic..](#) - [Traduci](#)
Profile Log in to check your private messages Log in. **My favorite topic..** Goto page Previous 1, 2, 3 Next. Astro Wars Forum Index -> Off Topic Discussions ...
www.astrowars.com/forums/viewtopic.php?p=718035 - 68k - [Cache](#)

+



... better technology nowadays ;—)



Research Questions

- ❑ Is plagiarism a problem with respect to education?
- ❑ Is there a misunderstanding wrt. an evolving cultural technique?
- ❑ Can plagiarism be detected by humans?
- ❑ Can plagiarism be detected by machines?
- ❑ Should automatic plagiarism detection algorithms become standard?

Plagiarism Detection

Research Questions

- ❑ Is plagiarism a problem with respect to education?
- ❑ Is there a misunderstanding wrt. an evolving cultural technique?
- ❑ Can plagiarism be detected by humans?
- ❑ **Can plagiarism be detected by machines?**
- ❑ Should automatic plagiarism detection algorithms become standard?

Plagiarism Detection

Research Questions

- ❑ Is plagiarism a problem with respect to education?
- ❑ Is there a misunderstanding wrt. an evolving cultural technique?
- ❑ Can plagiarism be detected by humans?
- ❑ **Can plagiarism be detected by machines?**
- ❑ Should automatic plagiarism detection algorithms become standard?

For several reasons we should say “text reuse” rather than “plagiarism”.

Vandalism Detection

Vandalism Detection in Padua

WE ARE JUST TWO
LOST SOULS SWIMMING
IN A FISH BOWL



Vandalism Detection in Wikipedia

Bushido (rapper)

From Wikipedia, the free encyclopedia

[\(Difference between revisions\)](#)

Revision as of 07:33, 29 November 2009 (edit)

71.193.157.91 (talk)

[\(→Relationship with Bill Kaulitz\)](#)

[← Previous edit](#)

Revision as of 07:33, 29 November 2009 (edit) (undo)

71.193.157.91 (talk)

[\(→Lawsuits for assault and copyright infringement\)](#)

[Next edit →](#)

Line 36:

get rick rolld

- **===Lawsuits for assault and copyright infringement===**

- **More legal trouble found its way into his life following a party in Linz, Austria on July 30, 2005. Upon discovery that his tires had been slashed, he and two of his bodyguards were**

Line 36:

get rick rolld

+ **get rick rolld**

Vandalism Detection in Wikipedia

Bushido (rapper)

From Wikipedia, the free encyclopedia

([Difference between revisions](#))

Revision as of 07:33, 29 November 2009 ([edit](#))

71.193.157.91 ([talk](#))

([→Relationship with Bill Kaulitz](#))

[← Previous edit](#)

Revision as of 07:33, 29 November 2009 ([edit](#)) ([undo](#))


71.193.157.91 ([talk](#))

([→Lawsuits for assault and copyright infringement](#))

[Next edit →](#)

Line 36:

Line 36:

get rick rolld	get rick rolld
<div><div><div>===Law</div><div>More le</div><div>2005. U</div></div><div><div><div>Contents [hide]</div><div><div><div>1 Musical career</div><div>2 Controversy</div><div>2.1 Feud with Aggro Berlin</div><div>3 Discography</div><div>3.1 Studio albums</div><div>3.2 Underground releases</div><div>3.3 Compilations</div><div>3.4 Collaborations</div><div>4 Other releases</div><div>5 Bibliography</div><div>6 Filmography</div><div>7 Notes</div><div>8 External links</div></div></div></div></div></div>	<div><div><div><div></div><div><div>Background information</div><div><div><div>Birth name</div><div>Anis Mohamed Youss Ferchichi</div></div><div><div>Born</div><div>September 28, 1978 ([</div></div><div><div>Origin</div><div>Tempelhof ,Berlin, Ger</div></div><div><div>Genres</div><div>Hip-Hop, Gangsta Rap</div></div><div><div>Occupations</div><div>Rapper & Producer</div></div><div><div>Years active</div><div>1998-present</div></div><div><div>Labels</div><div>I Luv Money Records (1999-2001) Aggro Berlin (2001-2004) ersguterjunge (2004-[</div></div><div><div>Associated acts</div><div>Baba Saad, Chakuza, Fler</div></div><div><div>Website</div><div>www.kingbushido.de</div></div></div></div></div></div></div>
<div><div><div><div><div>Musical career</div><div>get rick rolld</div><div>get rick rolld</div><div>get rick rolld</div><div>get rick rolld</div></div></div></div></div>	

Vandalism Detection in Wikipedia

Henry Wadsworth Longfellow

From Wikipedia, the free encyclopedia

([Difference between revisions](#))

Revision as of 17:33, 1 December 2009 (edit)

[Snigbrook](#) (talk | [contribs](#))

m *(Reverted edits by 70.57.239.4 to last revision by Snigbrook (HG))*

[← Previous edit](#)

Line 89:

===Critical response===

[[Image:Sumner-Longfellow.jpg|thumb|right|Longfellow and his good friend [[United States Senate|Senator]] [[Charles Sumner]]]]

Longfellow's early collections, "Voices of the Night" and "Ballads and Other Poems", made him instantly popular. The "New Yorker" called him "one of the very few in our time who has

Revision as of 22:02, 1 December 2009 (edit) (undo)

[141.154.179.145](#) (talk)

(→Critical response)

[Next edit →](#)

Line 89:

===Critical response===

[[Image:Sumner-Longfellow.jpg|thumb|right|Longfellow and his good friend [[United States Senate|Senator]] [[Charles Sumner]]]]

Longfellow's early collections, **ugly**"Voices of the Night" and "Ballads and Other Poems" made him instantly popular. The "New Yorker" called him "one of the very few in our tim

Vandalism Detection in Wikipedia

Henry Wadsworth Longfellow

From Wikipedia, the free encyclopedia

([Difference between revisions](#))

Revision as of 17:33, 1 December 2009 (edit)

[Snigbrook](#) ([talk](#) | [contribs](#))

m (*Reverted edits by 70.57.239.4 to last revision by Snigbrook (HG)*)

[← Previous edit](#)

Revision as of 22:02, 1 December 2009 (edit) (undo)

[141.154.179.145](#) ([talk](#))

([→Critical response](#))

[Next edit →](#)

Line 89:

Line 89:

===Critical response===	===Critical response===
<div>[[Image:Senate Chamber, U.S. Capitol Building, 1878.jpg thumb right 200px The Senate Chamber, U.S. Capitol Building, 1878]]</div> <div>Longfellow's role with translations, Harvard established the Longfellow Institute in 1994, dedicated to literature written in the United States in languages other than English.^[102]</div>	<div>want a national literature altogether shaggy and unshorn, that shall shake the earth, like a herd of buffaloes thundering over the prairies.^[99]</div> <div>He was also important as a translator; his translation of Dante became a required possession for those who wanted to be a part of high culture.^[100] He also encouraged and supported other translators. In 1845, he published <i>The Poets and Poetry of Europe</i>, an 800-page compilation of translations made by other writers, including many by his friend and colleague Cornelius Conway Felton. Longfellow intended the anthology "to bring together, into a compact and convenient form, as large an amount as possible of those English translations which are scattered through many volumes, and are not accessible to the general reader. In honor of Longfellow's role with translations, Harvard established the Longfellow Institute in 1994, dedicated to literature written in the United States in languages other than English.^[102]</div> <div>In 1874, Longfellow oversaw a 31-volume anthology called <i>Poems of Places</i>, which collected poems representing several geographical locations, including European, Asian, and Arabian countries.^[103] Emerson was disappointed and reportedly told Longfellow: "The world is expecting better things of you than this. You are wasting time that should be bestowed upon original production".^[104] In preparing the volume, Longfellow hired Katherine Sherwood Bonner as an amanuensis.^[105]</div> <div><h3>Critical response</h3><p>Longfellow's early collections, <u>ugly</u> <i>Voices of the Night</i> and <i>Ballads and Other Poems</i>, made him instantly popular. The <i>New-Yorker</i> called him "one of the very few in our time who has successfully aimed in putting poetry to its best and sweetest uses".^[45] The <i>Southern Literary Messenger</i> immediately put Longfellow "among the first of our American poets".^[46] Poet John Greenleaf Whittier said that Longfellow's poetry illustrated "the careful moulding by which art attains the graceful ease and chaste simplicity of nature".^[106] Longfellow's friend Oliver Wendell Holmes, Sr. wrote of him as "our chief singer" and one who "wins and warms... kindles, softens, cheers [and] calms the wildest woe and stays the bitterest tears!"^[107]</p><p>The rapidity with which American readers embraced Longfellow was unparalleled in publishing history in the United States;^[108] by 1874, he was earning \$3,000 per poem.^[109] His popularity spread throughout Europe as well and his</p></div>



Vandalism Detection in Wikipedia

Jingle Bells/U Can't Touch This

From Wikipedia, the free encyclopedia
([Difference between revisions](#))

Revision as of 12:41, 28 August 2009 ([edit](#))

85.132.47.9 ([talk](#))
([→Single track listing](#))
[← Previous edit](#)

Revision as of 09:50, 2 December 2009 ([edit](#)) ([undo](#))

Ragger256 ([talk](#) | [contribs](#))
([→Jingle Bells](#))
[Next edit →](#)

Line 55:

===Jingle Bells===

- In some clips, his genitals are censored.

The video starts with the Crazy Frog playing in the snow with the bounty hunter robot from previous clips. It then shows flashbacks from clips of "[Axel F#Crazy Frog version[Axel F]]." a trip to a carnival, and the Crazy DJ clip, then more of the "Axel F" clip. The flashbacks end,

Line 55:

===Jingle Bells===

+ In some clips, his genitals are HUGE.

The video starts with the Crazy Frog playing in the snow with the bounty hunter robot from previous clips. It then shows flashbacks from clips of "[Axel F#Crazy Frog version[Axel F]]." a trip to a carnival, and the Crazy DJ clip, then more of the "Axel F" clip. The flashbacks e

Vandalism Detection in Wikipedia

Jingle Bells/U Can't Touch This

From Wikipedia, the free encyclopedia
([Difference between revisions](#))

Revision as of 12:41, 28 August 2009 ([edit](#))

[85.132.47.9](#) ([talk](#))
([→Single track listing](#))
[← Previous edit](#)

Revision as of 09:50, 2 December 2009 ([edit](#)) ([undo](#))

[Ragger256](#) ([talk](#) | [contribs](#))
([→Jingle Bells](#))
[Next edit →](#)

Line 55:

Line 55:

===Jin

- In som

The vi

previo

trip to

Music videos

"U Can't Touch This"

The video depicts Crazy Frog causes chaos at the underwater sealab of "The Boss".

Jingle Bells

In some clips, his genitals are HUGE.

The video starts with the Crazy Frog playing in the snow with the bounty hunter robot from previous clips. It then shows flashbacks from clips of "Axel F," a "Axel F" clip. The flashbacks end, and the bounty hunter robot begins to throw a snowball at the frog. But instead he kisses the bounty hunter robot, and h then make snow angels and a message reads "Have a ding dong Christmas, everyone!"

Certifications

Country	Certification	Date	Sales certified
Australia ^[3]	Gold	2005	35,000

Charts

Chart (2006)	Peak position

End of year chart (2005)	Position
Belgian Albums Chart ^[8]	60

Vandalism Detection in Wikipedia

Groundhog Day

From Wikipedia, the free encyclopedia
([Difference between revisions](#))

Revision as of 16:29, 23 November 2009 ([edit](#))

[NawlinWiki](#) ([talk](#) | [contribs](#))

m (*Reverted edits by [Kappamikey36](#) ([talk](#)) to last version by [Alphageekpa](#)*)

[← Previous edit](#)

Line 18:

Groundhog Day received worldwide attention as a result of the 1993 film of the same name, [[Groundhog Day (film)|"Groundhog Day"]], which was set in Punxsutawney (though filmed primarily in Woodstock, Illinois) and featured Punxsutawney Phil.<ref>Yoder, pp. 14-15.</ref>

- **==History==**

Revision as of 16:04, 24 November 2009 ([edit](#)) ([undo](#))

[Kappamikey36](#) ([talk](#) | [contribs](#))

([→History](#))

[Next edit →](#)

Line 18:

Groundhog Day received worldwide attention as a result of the 1993 film of the same name, [[Groundhog Day (film)|"Groundhog Day"]], which was set in Punxsutawney (though filmed primarily in Woodstock, Illinois) and featured Punxsutawney Phil.<ref>Yoder, pp. 14-15.</ref>

+ **==History==**□□□ô□□□ô

Vandalism Detection in Wikipedia

Groundhog Day

From Wikipedia, the free encyclopedia
([Difference between revisions](#))

Revision as of 16:29, 23 November 2009 ([edit](#))

[NawlinWiki](#) ([talk](#) | [contribs](#))

m (*Reverted edits by [Kappamikey36](#) ([talk](#)) to last version by [AlphaGeekpa](#)*)

[← Previous edit](#)

Revision as of 16:04, 24 November 2009 ([edit](#)) ([undo](#))

[Kappamikey36](#) ([talk](#) | [contribs](#))

([→History](#))

[Next edit →](#)

Line 18:

Line 18:

Groundhog Day received worldwide attention as a result of the 1993 film of the same name.

Groundhog Day received worldwide attention as a result of the 1993 film of the same name.

[[Groundhog Day|**8 External links**]]

primar

</ref>

- ==Hist

History

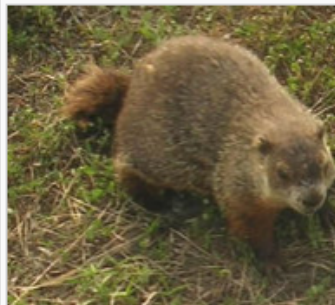
Historical origins

An early American reference to Groundhog Day can be found in a diary^[17] entry, dated [February 5, 1841](#), of [Berks County, Pennsylvania](#) storekeeper James Morris:

Last Tuesday, the 2nd, was [Candlemas](#) day, the day on which, according to the Germans,^[18] the Groundhog peeps out of his winter quarters and if he sees his shadow he pops back for another six weeks nap, but if the day be cloudy he remains out, as the weather is to be moderate."

In the United States the tradition may also derive from a Scottish poem:

As the light grows longer
The cold grows stronger
If Candlemas be fair and bright
Winter will have another flight
If Candlemas be cloud and snow
Winter will be gone and not come again
A farmer should on Candlemas day
Have half his corn and half his hay
On Candlemas day if thorns hang a drop
You can be sure of a good new crop



The [groundhog](#) (*Marmota monax*) is a rodent of the family [Sciuridae](#), belonging to a group of large [ground squirrels](#).

Vandalism Detection in Wikipedia

Television pilot

From Wikipedia, the free encyclopedia

([Difference between revisions](#))

Revision as of 17:47, 23 November 2009 (edit)

[Soc8675309](#) ([talk](#) | [contribs](#))

(→ *Retooled ideas: Add "Who's the Boss?" spinoffs*)

[← Previous edit](#)

Revision as of 01:34, 24 November 2009 (edit) (undo)

[209.17.173.177](#) ([talk](#))

(→ *Unintentional pilots*)

[Next edit →](#)

Line 145:

While, as listed above, there are many telemovies or episodes within series intended as pilots, there are often telemovies or episodes within other series which are so popular that they inspire later TV series. A popular example is "[[The Simpsons]]", which started as [[The Simpsons shorts|a set of shorts]] on "[[The Tracey Ullman Show]]". Another example is "[[South Park]]", which started as a cartoon with an extremely low budget which was created for a class at the University of Colorado, which the creators [[Trey Parker]] and [[Matt Stone]] were attending at the time.

Another use is the [[Larry shorts]] by [[Seth MacFarlane]] for "[[Family Guy]]": prototypes that where Larry was to later be transformed into the character [[Peter Griffin]] and Steve [[Brian

Line 145:

While, as listed above, there are many telemovies or episodes within series intended as pilots, there are often telemovies or episodes within other series which are so popular th they inspire later TV series. A popular example is "[[The Simpsons]]", which started as [[Simpsons shorts|a set of shorts]] on "[[The Tracey Ullman Show]]". Another example is "[[South Park]]", which started as a cartoon with an extremely low budget which was crea for a class at the University of Colorado, which the creators [[Trey Parker]] and [[Matt Sto were attending at the time.

+ THE FOLLOWING SECTION IS A TOTAL MESS AND NEEDS CLEANING UP

Another use is the [[Larry shorts]] by [[Seth MacFarlane]] for "[[Family Guy]]": prototypes where Larry was to later be transformed into the character [[Peter Griffin]] and Steve [[Br

Vandalism Detection in Wikipedia

Television pilot

From Wikipedia, the free encyclopedia

([Difference between revisions](#))

Revision as of 17:47, 23 November 2009 ([edit](#))

[Soc8675309](#) ([talk](#) | [contribs](#))

([→Retoold ideas: Add "Who's the Boss?" spinoffs](#))

[← Previous edit](#)

Revision as of 01:34, 24 November 2009 ([edit](#)) ([undo](#))

[209.17.173.177](#) ([talk](#))

([→Unintentional pilots](#))

[Next edit →](#)

Line 145:

While, as listed above, there are many telemovies or episodes within series intended as pilots, there are often telemovies or episodes within other series which are so popular that they inspire later TV series. A popular example is "[[The Simpsons]]", which started as [[The

Line 145:

While, as listed above, there are many telemovies or episodes within series intended as pilots, there are often telemovies or episodes within other series which are so popular th they inspire later TV series. A popular example is "[[The Simpsons]]", which started as [[

- British Cop Drama *The Bill* was originally an episode of the anthology series *Storyboard*^[1] [@](#) called "[Woodentop](#)".
- *Rumpole of the Bailey* first appeared on *Play for Today*.
- Popular British comedies *Steptoe and Son*, *Til Death Us Do Part*, *All Gas and Gaiters*, *The Liver Birds*, *Are You Being Served?*, and *Last of the Summer* all began as episodes of the *Comedy Playhouse* strand.
- The 2008 BBC series *Freezing* was expanded from the first episode (also titled *Freezing*) of the 2007 BBC comedy anthology series *Tight Spot*.^[12]

In some cases, a series is created specifically to showcase pilots.

- Both *Prisoner and Escort* (which led to *Porridge*) and *Open All Hours* first appeared as part of [Ronnie Barker's](#) *Seven of One* series.
- BBC2's series of comedy pilots which aired under the title *Comic Asides* spawned the series *The High Life*, *KYTV*, *Mornin' Sarge* and *Tygo Road*.

Unintentional pilots

While, as listed above, there are many telemovies or episodes within series intended as pilots, there are often telemovies or episodes within other series wh so popular that they inspire later TV series. A popular example is *The Simpsons*, which started as [a set of shorts](#) on *The Tracey Ullman Show*. Another ex is *South Park*, which started as a cartoon with an extremely low budget which was created for a class at the University of Colorado, which the creators [Trey Parker](#) and [Matt Stone](#) were attending at the time.

THE FOLLOWING SECTION IS A TOTAL MESS AND NEEDS CLEANING UP Another use is the *Larry shorts* by [Seth MacFarlane](#) for *Family Guy*: prototy where Larry was to later be transformed into the character [Peter Griffin](#) and Steve [Brian Griffin](#). Two of his earlier cartoons, called "Life with Larry" (made in Rhode Island College) and another called "Larry & Steve" (a sequel to "Life with Larry" (made once MacFarlane had been hired by Hanna-Barbera in 1996), was aired for [Cartoon Network](#) as a part of the *What a Cartoon!* show, led to [Fox Broadcasting Company](#) to offer MacFarlane a chance to develop them into show. Coincidentally Larry and Steve included a Fight with a chicken and a woman named Cindy who vaguely resembled Lois.

Vandalism Detection in Wikipedia

Mainframe computer

From Wikipedia, the free encyclopedia

([Difference between revisions](#))

Revision as of 15:57, 17 November 2009 (edit)

[TXiKiBoT](#) (talk | [contribs](#))

m (robot *Adding: bg:Мейнфрейм компютър*)

[← Previous edit](#)

Revision as of 14:48, 24 November 2009 (edit) (undo)

[74.202.102.94](#) (talk)

[Next edit →](#)

Line 12:

Today in practice, the term usually refers[[Citation needed|date=November 2009]] to computers compatible with the [[IBM System/360]] line, first introduced in 1965. ([[IBM System z10]] is the latest incarnation.) Otherwise, large systems that are not based on the System/360 but are used for similar tasks are usually referred to as [[computer server|servers]] or even [[supercomputer]]s. However, "[[server]]", "[[supercomputer]]" and "mainframe" are not synonymous (see [[client-server]]).

Some non-System/360-compatible systems derived from or compatible with older (pre-Web) server technology may also be considered mainframes. These include the [[Burroughs large systems]], the [[UNIVAC 1100/2200 series]] systems, and the pre-System/360 [[IBM 700/7000 series]]. Most large-scale computer system architectures were firmly established in the 1960s and most large computers were based on architecture established during that era up until the advent of Web servers in the 1990s. (Interestingly, the first Web server running anywhere outside Switzerland ran on an IBM mainframe at Stanford University as early as 1990. See [[History of the World Wide Web]] for details.)

Line 12:

Today in practice, the term usually refers[[Citation needed|date=November 2009]] to computers compatible with the [[IBM System/360]] line, first introduced in 1965. ([[IBM System z10]] is the latest incarnation.) Otherwise, large systems that are not based on the System/360 but are used for similar tasks are usually referred to as [[computer server|serv or even [[supercomputer]]s. However, "[[server]]", "[[supercomputer]]" and "mainframe" a not synonymous (see [[client-server]]).

Some non-System/360-compatible systems derived from or compatible with older (pre-W server technology may also be considered mainframes. These include the [[Burroughs la systems]], the [[UNIVAC 1100/2200 series]] systems, and the pre-System/360 [[IBM 700/7 series]]. Most large-scale computer system architectures were firmly established in the 1960s and most large computers were based on architecture established during that era up until the advent of Web servers in the 1990s. (Interestingly, the first Web server running anyw outside Switzerland ran on an IBM mainframe at Stanford University as early as 1990. **Harris H**Acks the Mainframe all the time. See [[History of the World Wide Web]] for details.)

Vandalism Detection in Wikipedia

Mainframe computer

From Wikipedia, the free encyclopedia

([Difference between revisions](#))

Revision as of 15:57, 17 November 2009 (edit)

[TXiKiBoT](#) (talk | [contribs](#))

‎m (robot Adding: *bg:Мейнфрейм компютър*)

[← Previous edit](#)

Revision as of 14:48, 24 November 2009 (edit) (undo)

[74.202.102.94](#) (talk)

[Next edit →](#)

Line 12:

Today in practice, the term usually refers[Citation needed][date=November 2009] to computers compatible with the [IBM System/360] line, first introduced in 1965. ([IBM System z10] is the latest incarnation.) Otherwise, large systems that are not based on the System/360 but are used for similar tasks are usually referred to as [computer server|servers] or even [supercomputer]s. However, "[server]", "[supercomputer]" and "mainframe" are not synonymous (see [client-server]).

Line 12:

Today in practice, the term usually refers[Citation needed][date=November 2009] to computers compatible with the [IBM System/360] line, first introduced in 1965. ([IBM System z10] is the latest incarnation.) Otherwise, large systems that are not based on the System/360 but are used for similar tasks are usually referred to as [computer server|serv or even [supercomputer]s. However, "[server]", "[supercomputer]" and "mainframe" a not synonymous (see [client-server]).

Mainframes (often colloquially referred to as **Big Iron**^[1]) are computers used mainly by large organizations for critical applications, typically bulk data processing such as census, industry and consumer statistics, enterprise resource planning, and financial transaction processing.

The term probably had originated from the early mainframes, as they were housed in enormous, room-sized metal boxes or frames.^[2] Later the term was used to distinguish high-end commercial machines from less powerful units.

Today in practice, the term usually refers[citation needed] to computers compatible with the [IBM System/360] line, first introduced in 1965. ([IBM System z10] is the latest incarnation.) Otherwise, large systems that are not based on the System/360 but are used for similar tasks are usually referred to as [servers] or even [supercomputers]. However, "[server]", "[supercomputer]" and "mainframe" are not synonymous (see [client-server]).



An IBM 704 mainframe

Some non-System/360-compatible systems derived from or compatible with older (pre-Web) server technology may also be considered mainframes. These the [Burroughs large systems], the [UNIVAC 1100/2200 series] systems, and the pre-System/360 [IBM 700/7000 series]. Most large-scale computer system architectures were firmly established in the 1960s and most large computers were based on architecture established during that era up until the advent of V servers in the 1990s. (Interestingly, the first Web server running anywhere outside Switzerland ran on an IBM mainframe at Stanford University as early as ' Harris **H**ACKs the Mainframe all the time. See [History of the World Wide Web] for details.)

Vandalism Detection in Wikipedia

The Machine Learning Perspective

Every edit on Wikipedia has to be double-checked for integrity—even if it affects just one char.

The task is to discriminate between regular edits and vandalism edits.

The achievements of ML enfold their full power in discrimination situations.

→ **Feature engineering** plays an outstanding role.

PAN Overview Cont'd

Facts and Stats

- ❑ Previous workshops at SIGIR'07 and ECAI'08;
previous PAN plagiarism detection competition at SEPLN'09.
- ❑ Sponsorship by **YAHOO!** Research (2009, 2010).
- ❑ Media coverage on **Slashdot** (2009, 2010), among others.

	2009	2010	
	plagiarism	plagiarism	vandalism
Corpus size (GB)	5 GB	3.4 GB	8.2 GB
Corpus size (cases)	94 000	68 000	32 000
Registrations	21	38	15
Countries	17	24	11
Run submissions	14	18	9
Notebook submissions	11	17	5
Followers (mailing list)	78	151	

PAN Overview Cont'd

Program

Sessions

- ❑ Wednesday, 9:00. PAN Task 1 - Plagiarism Detection
- ❑ Wednesday, 11:00. PAN Task 2 - Wikipedia Vandalism Detection
- ❑ Wednesday, 18:00. Poster Session
- ❑ Thursday, 9:00. Reports from the Labs

Web

- ❑ <http://pan.webis.de>
- ❑ pan@webis.de