

Vil vi bli millionær? - Botprodusert Wikipedia

Mattias Blomgren

Ordförande Wikimedia Sverige

Wikipedia Academy
Oslo 14 december 2013

Wikipedia

“Imagine a world in which every single person on the planet is given free access to the sum of all human knowledge. That's what we're doing.”

Jimmy Wales, 2004.

Alla kan läsa på Wikipedia, alla kan skriva.

Hundratals miljoner *läser*...

...men långt ifrån alla skriver.

”Sum of all human knowledge”

- Det är **mycket** kunskap
- Även med tusentals wikipediaskribenter, skulle vi enbart skrapa på ytan.
- Vi kommer att skriva om det som intresserar oss, vilket ger en ojämn täckning.
- Svenskspråkiga Wikipedia:
 - Fler än 100 figurer från Lord of the Rings.
 - Färre än 10 personer från Vietnamkriget.

Vad kan vi göra åt det?

Öka täckningen, och göra det mer representativt och mer tillgängligt!

Det kan göras på två sätt:

- Rekrytera miljoner skribenter med olika intressen och kunskaper.
- Använda verktyg för artikelskapande är effektivare än manuellt skapande.

Masskapande av artiklar

- Wikipedia är öppet för automatiskt artikelskrivande av programvara, ”botar”.
- Mycket rutinarbete görs av botar – stor andel av all redigering på Wikipedia (26%).
- Artikelskapande av botar är mer komplext, och mer kontroversiellt.
- **Flera projekt på olika språkversioner av Wikipedia under de senaste åren.**

Bot-skapande

- 7 miljoner bot-skapade artiklar (23,5% av alla)
- En handfull personer arbetar med botskapande.
- Lsjbot är den största bidragsgivaren (2,6 miljoner).
- Holländska, svenska, cebuano, waray är språkversioner som har omkring en miljon bot-skapade artiklar var.
 - **Inte** världens största språk.
- Djur, växter, samhällen, sjöar, ...
 - **Inte** de typiska intressena för de typiska wikipedianerna.

Absintmalmätare[redigera]

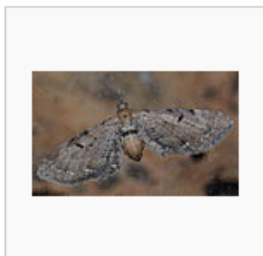


Denna artikel har skapats av boten Lsjbot och kan bland annat innehålla språkliga fel eller ett märkligt val av bilder.

Mallen kan tas bort efter kontroll av innehållet. Fler robotskapade artiklar med denna mall finns i kategorin robotskapade artiklar.

Absintmalmätare (*Eupithecia absinthiata*)^[1] är en fjärilsart som beskrevs av Carl Alexander Clerck 1759. Absintmalmätare ingår i släktet *Eupithecia*, och familjen mätare.^{[1][2][3]} Arten är reproducerande i Sverige.^[3] Inga underarter finns listade.^[1]

Bildgalleri [redigera | redigera wikitext]



Källor [redigera | redigera wikitext]

- ↑ [a b c] Bisby F.A., Roskov Y.R., Orrell T.M., Nicolson D., Paglinawan L.E., Bailly N., Kirk P.M., Bourgoin T., Baillargeon G., Ouvrard D. (red.) (2011). "Species 2000 & ITIS Catalogue of Life: 2011 Annual Checklist." ↗. Species 2000: Reading, UK.. Läst 24 september 2012.
- ↑ *LepIndex: The Global Lepidoptera Names Index*. Beccaloni G.W., Scoble M.J., Robinson G.S. & Pitkin B., 2005-06-15
- ↑ [a b] Dyntaxa *Eupithecia absinthiata* ↗

Externa länkar [redigera | redigera wikitext]

- Wikimedia Commons har media som rör Absintmalmätare.

Denna artikel om *mätare* är bara påbörjad. Du kan hjälpa till genom att *utöka den*.

Absintmalmätare



Systematik

Rike	Djur
	Animalia
Stam	Leddjur
	Arthropoda
Understam	Sexfotingar
	Hexapoda
Klass	Egentliga insekter
	Insecta
Ordning	Fjärilar
	Lepidoptera
Överfamilj	Geometroidea
Familj	Mätare
	Geometridae
Släkte	<i>Eupithecia</i>
Art	Absintmalmätare
	<i>Eupithecia absinthiata</i>

Vetenskapligt namn

Eupithecia absinthiata   ↗

Auktor   Clerck 1759

Hitta fler artiklar om djur med

Djurportalen

Kännetecknen för bot-skapade artiklar

Fördelar:

- Grundläggande fakta presenteras koncist och konsekvent.
- Korrekt wiki-format.
- Infobox, mallar, källor, interwikilänkar...
- Inga stavfel.
- Problem kan ofta åtgärdas med en bot.

Nackdelar:

- Artikeltexten kort och standardmässig. Tråkig att läsa.
- Begränsat till standardiserade fakta som finns tillgängliga i maskinläsbart format.
- Följer källan blint.
- Hanterar specialfall och avvikelser dåligt.

Ämnen för bot-skapande

Krav:

- Stort antal liknande entiteter existerar, vilka är relevanta för Wikipedia-artiklar.
- Varje entitet kan på ett lämpligt sätt beskrivas utifrån en uppsättning standardiserade fakta.
- Grundläggande faktauppgifter bör helst vara språkoberoende.
- Databaser med komplett täckning av ämnet finns tillgängliga.

Exempel på lämpliga ämnen:

- Arter av djur och växter
- Geografiska objekt
- Administrativa enheter
- Astronomiska objekt
- Kemiska ämnen
- Gener och proteiner
- Böcker och författare

Några genomförda projekt

- Fågelarter, 8.000 artiklar:
 - Svenska, Lsjbot, Excel+Pythia+Google Translate
- Kommuner i Frankrike, 50.000:
 - Svenska, Naskobot, Excel+AWB
- Kemiska ämnen, 100
 - Cebuano, Lsjbot, C#
- Kommuner i Filippinerna, 1.300
 - Cebuano, Lsjbot, C#
- Sjöar i Sverige, 56.000
 - Svenska, Naskobot, Excel+GIS+AWB
- Amerikanska Counties
 - Svenska, Innocent Bot, ?
- Djurarter och svampar
 - Svenska+cebuano+waray, Lsjbot, C#, 3*1.000.000
- Djurarter på holländska och vietnamesiska
 - (Känner ej till detaljerna)

Fördelar och nackdelar med bot-skapande

Nackdelar:

- Tråkiga artiklar om obskyra ämnen.
- Fungerar bara för vissa ämnen.
- Kontroversiellt inom vissa Wikipedia-gemenskaper.
- Databasfel och programmeringsfel inträffar.

Fördelar:

- Effektivt för att få en fullständig täckning inom de områden där det fungerar.
- Korrekta och källbelagda artiklar.
- Täcker områden som få är intresserade av att skriva om.

Kan få oss väsentligt närmare målet att göra all kunskap tillgänglig för alla människor.

Erfarenheter från svenskspråkiga Wikipedia

Diskussioner

- Januari 2012: 250.000 artiklar om djurarter på **nl:wp**.
 - Lsj: Intressant idé
 - Invändningar
 - Tar bort det roliga med att skapa nya artiklar.
 - Intetsägande stubbartiklar som i Wikipedias barndom.
 - Blodigt tråkiga artiklar om amerikanska orter på en:wp
 - Behövs ingen konstgjord andning. Intressantare att jämföra storlek på wikipediaversioner om det gäller artiklar skrivna av människor.
 - Korta artiklar – påverkar statistiken.

Erfarenheter från svenskspråkiga Wikipedia

Diskussioner

- Januari 2012: 250.000 artiklar om djurarter på **nl:wp**. (forts.)
 - Synpunkter:
 - Låt robotarna göra det tråkiga jobbet.
 - Börja med faktaboxar.
 - Komplettera existerande artiklar med faktaboxar.
 - Lämpligare för kemiska ämnen med basfakta i siffror än datorspel, som är mer humaniora.
 - Acceptabelt utformade artiklar kan byggas ut senare.
 - Förslag: Använd IAAF:s databas över internationella friidrottare.
 - Diskussion om hur kategorisera artartiklar och namnsättning: vetenskapligt eller svenskt
 - Lsjbot startade i januari 2012

Erfarenheter från svenskspråkiga Wikipedia

Diskussioner

- Mars 2012: sv:wp nionde språkversionen
 - 100.000 kvar till pl:wp
 - Invändning: Inte längre relevant att räkna antalet artiklar – räkna antalet utvalda artiklar.
 - Positivt: De robotskapade artiklarna är bättre än många manuellt skapade.
 - Negativt: Inte alls informationsrika. Faktaruta och många källor.

Erfarenheter från svenskspråkiga Wikipedia

Diskussioner

- Januari 2013: Större än kinesiska Wikipedia
 - Sv:wp: 673.130 artiklar – zh:wp 668.264 artiklar
- April 2013: Tyska invändningar
 - Tysk användare hade uppmärksammat en negativ tysk diskussion om de svenska botartiklarna och ville veta mer om hur vi resonerat.
 - Tyskarna hade tidigare röstat nej till botskapande.
 - Förslag från tysk användare: Skapa en bot:wp språkversion. Förslaget avvisades.

Erfarenheter från svenskspråkiga Wikipedia

Diskussioner

- November 2013:
 - Workshop om överföring av institutioners databaser till Wikipedia-artiklar. Idéer:
 - 300.000 växter
 - 500.000 kommuner i hela världen
 - 1.200.000 böcker från Libris/Kungliga Biblioteket
 - Stjärnor och planeter
 - Invändningar:
 - Charmen ligger i att någon skrivit manuellt
 - Faktafel slinker igenom då den som botskapar inte har kunskap om all information (exempel: sjöar i fel socken).
 - Hur många artiklar kan vi underhålla?

Erfarenheter från svenskspråkiga Wikipedia

Diskussioner

- November 2013 (forts.):
 - Idé: Bot skapar på förfrågan en faktamall när en människa vill skriva en artikel om en stad/region.
 - Skepsis mot botartiklar om böcker (relevans?) – bättre skriva om författare.
 - Oro: Vi går från uppslagsverk till databas med data kopierad från andra databaser.
 - Sjöar har större allmänintresse än 5.939 obskyra hoppspindlar.
 - Teknokratiskt och användarfrånvänt intryck (**Matsbosjön (Hedemora socken, Dalarna, 668694-150841)**)



WIKIPEDIA
Den fria encyklopedin

- Huvudsida
- Skriv en ny artikel
- Deltagarportalen
- Bybrunnen
- Senaste ändringarna
- Slumpartikel
- Ladda upp filer
- Stöd Wikipedia
- Kontakta Wikipedia
- Hjälp

- ▼ Skriv ut/exportera
 - Skapa en bok
 - Hämta som PDF
 - Utskriftsvänlig version

- ▼ Verktyg
 - Sidor som länkar hit
 - Relaterade ändringar
 - Specialsidor
 - Permanent länk
 - Sidinformation
 - Citera denna artikel

- ▼ På andra språk
- Lägga till länkar

Artikel [Diskussion](#)

Visa

[Redigera](#) [Redigera wikitext](#)

[Visa historik](#)



Sök

Matsbosjön (Hedemora)

Koordinater: 60.30057°N 15.95722°Ö

(Omdirigerad från [Matsbosjön \(Hedemora socken, Dalarna, 668694-150841\)](#))

Matsbosjön är en *sjö* i *Matsbo*, *Hedemora kommun* i *Dalarna* och ingår i Dalälvens huvudavrinningsområde. Vid sjön finns en kommunal badplats.

Se även

[[redigera](#) | [redigera wikitext](#)]

- [Lista över insjöar i Hedemora kommun](#)

Källor

[[redigera](#) | [redigera wikitext](#)]

- SMHI: [Sjöar, vattendrag och avrinningsområden](#)

Externa länkar

[[redigera](#) | [redigera wikitext](#)]

- [Delavrinningsområden](#) på [SMHI Vattenwebb \(S-HYPE\)](#)
- [Ladda ner data från Svenskt Vattenarkiv](#)



Denna sjöartikel är maskinellt skapad och ett samarbete mellan [Wikipedia](#) och [VISS](#). Du kan se hur artikeln är gjord på sidan [Projekt Sveriges sjöar](#) och diskutera utformningen på [diskussionssidan](#). Fler maskinellt skapade sjöartiklar finns i kategorin [robotskapade sjöartiklar](#). När innehållet har kontrollerats kan mallen flyttas över till [diskussionssidan](#)

Matsbosjön Insjö



Geografiskt läge

Land	 Sverige
Län	Dalarnas län
Kommun	Hedemora kommun
Landskap	Dalarna
Socken	Hedemora socken
Koordinater	
WGS 84	 60.30057°N 15.95722°Ö
SWEREF 99 TM	6685270, 552905



Erfarenheter från svenskspråkiga Wikipedia

Veckans tävling

- Februari 2012: Robotskapade fågelartiklar
- Mars 2012: Puts av robotskapade hardjursartiklar
- Augusti 2012: Robotskapade fågelartiklar II
- December 2012: Robotskapade fågelartiklar III
- Maj 2013: Robotskapade däggdjursartiklar
- September 2013: Robotskapade ormartiklar
- November 2013: Puts av robotskapade sköldpaddartiklar

WIKIPEDIA
Den fria encyklopedin

Huvudsida
Skriv en ny artikel
Deltagarportalen
Bybrunnen
Senaste ändringarna
Slumpartikel
Ladda upp filer
Stöd Wikipedia
Kontakta Wikipedia
Hjälp

Skriv ut/exportera
Skapa en bok
Hämta som PDF
Utskriftsvänlig version

Verktyg
Sidor som länkar hit
Relaterade ändringar
Specialsidor
Permanent länk
Sidinformation
Dataobjekt

På andra språk

Kategori:Robotskapade artiklar

Den här kategorin innehåller artiklar som har skapats av robotar. Artiklar märkta med mallen {{robotskapad}} autokategoriseras här. Artiklarna bör kontrolleras av en människa, varefter mallen kan tas bort.

Underkategorier

Denna kategori har följande 24 underkategorier (av totalt 24).

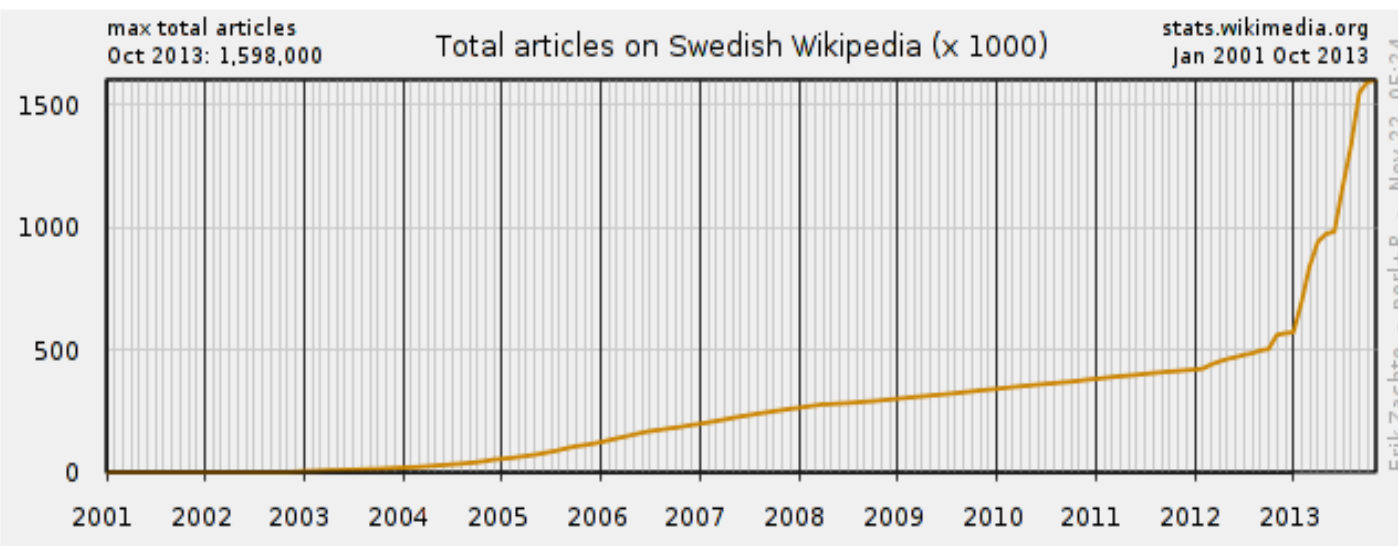
- | | | |
|---|--|---|
| <p>—</p> <ul style="list-style-type: none"> ▶ Robotskapade artiklar utan datum (tom) <p>R</p> <ul style="list-style-type: none"> ▶ Robotskapade artiklar 2012-08 (1344 sidor) ▶ Robotskapade artiklar 2012-10 (54302 sidor) ▶ Robotskapade artiklar 2012-11 (2532 sidor) ▶ Robotskapade artiklar 2012-12 (480 sidor) ▶ Robotskapade artiklar 2013-01 (124095 sidor) ▶ Robotskapade artiklar 2013-02 (132377 sidor) ▶ Robotskapade artiklar 2013-03 (96895 sidor) | <p>R forts.</p> <ul style="list-style-type: none"> ▶ Robotskapade artiklar 2013-04 (28997 sidor) ▶ Robotskapade artiklar 2013-05 (6748 sidor) ▶ Robotskapade artiklar 2013-06 (174806 sidor) ▶ Robotskapade artiklar 2013-07 (166677 sidor) ▶ Robotskapade artiklar 2013-08 (216810 sidor) ▶ Robotskapade artiklar 2013-09 (40264 sidor) ▶ Robotskapade artiklar 2013-10 (4062 sidor) ▶ Robotskapade djurartiklar (24 kategorier, 4528 sidor) | <p>R forts.</p> <ul style="list-style-type: none"> ▶ Robotskapade filippinska kommunartiklar (1344 sidor) ▶ Robotskapade förgreningssidor (3131 sidor) ▶ Robotskapade sjöartiklar (56884 sidor) ▶ Robotskapade sportartiklar (162 sidor) ▶ Robotskapade stubbar om djur (31 kategorier, 45 sidor) ▶ Robotskapade svampartiklar (1 kategori, 55721 sidor) ▶ Robotskapade växtartiklar (10 kategorier, 8 sidor) <p>S</p> <ul style="list-style-type: none"> ▶ Sidor för översättning från diskussion (305767 sidor) |
|---|--|---|

Swedish Wikipedia



Swedish Wikipedia at a glance *October 2013*

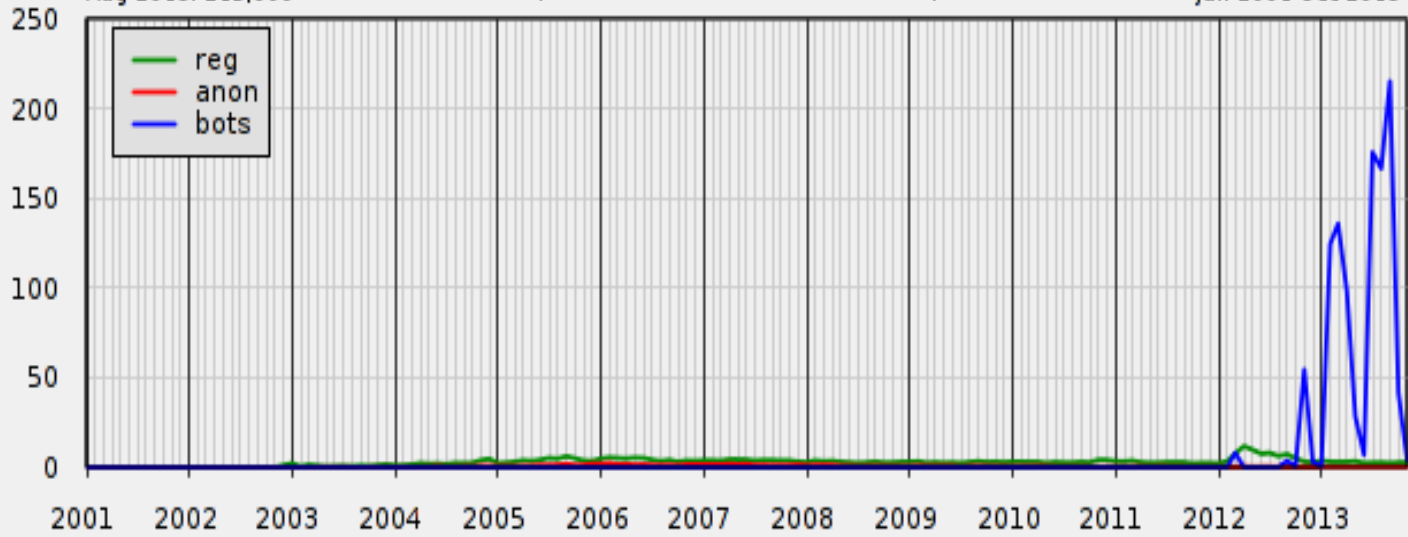
top ⇐	Data	Yearly change	Monthly change
Page Views per Month	157,973,400	--	--
Article Count	1,598,300	+185%	+0%
New Articles per Day	230	--	--
Edits per Month	198,411	-43%	+48%
Active Editors	755	-12%	+5%
Very Active Editors	115	-11%	+14%
New Editors	130	-16%	+25%
Speakers	10,000,000	--	--
Editors per Million Speakers	76	--	--



max new articles
Aug 2013: 215,000

New articles per month on Swedish Wikipedia (x 1000)

stats.wikimedia.org
Jan 2001 Oct 2013

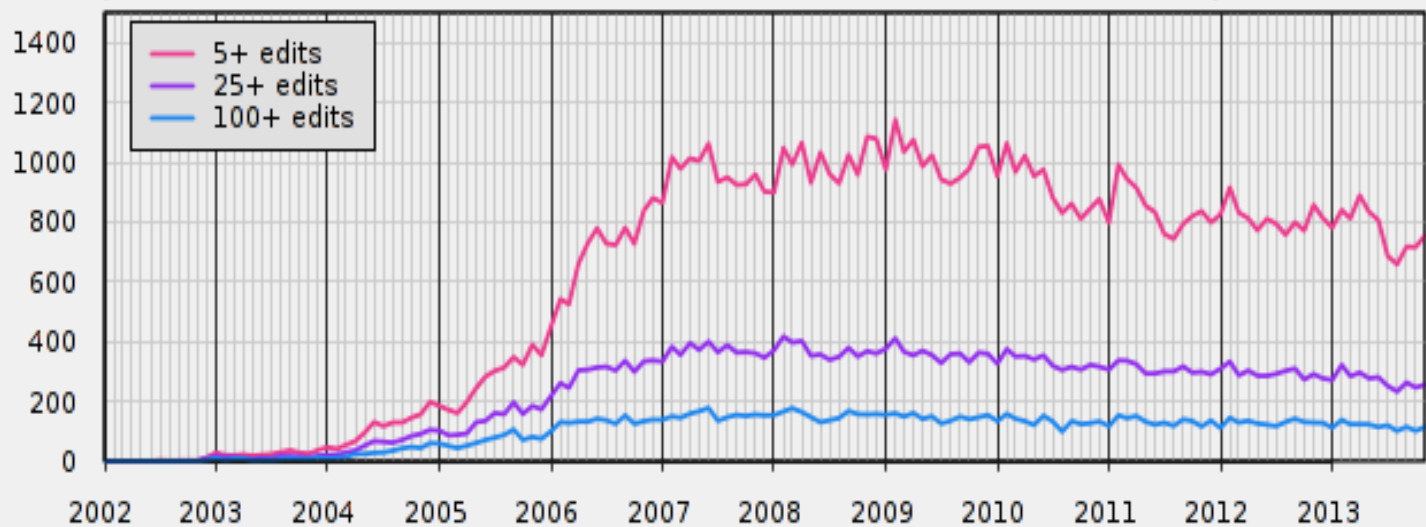


Erik Zachte - perl+R - Nov 22, 05:24

max editors (5+ edits) in
Jan 2009: 1,142

Active Editors on Swedish Wikipedia

stats.wikimedia.org
Jan 2002 Oct 2013



Erik Zachte - perl+R - Nov 22, 05:24

