

itWikiCon  
2020  
online edition

# Community Health Metrics

**Cristian Consonni**, David Laniado, Pablo Aragón, Marc Miquel-Ribé  
Eurecat - Centre Tecnològic de Catalunya

2020-10-25 @ itWikiCon 2020

**eurecat**  
Centre Tecnològic de Catalunya

# Indice



- Perché?
- Cosa?
- Come?
- Quando?
- Chi?



itWikiCon  
2020  
online edition

# Perché

Perché occuparsi di “community health”?

# Motivazione



- Wikipedia è un progetto **collaborativo**
- Gli **editor** sono il cuore del progetto
- La partecipazione è influenzata dalle **interazioni** tra gli utenti
- **Capire** come funzionano queste interazioni è necessario
- Creare un **clima** migliore nella comunità e non perdere persone importanti.

Come attrarre  
nuove forze

Come evitare di  
perdere utenti  
esperti



itWikiCon  
2020  
online edition

# Cosa

Il "Drop-off"



1. Misurare il drop-off degli editor nelle varie versioni linguistiche
2. Identificare i fattori che possono spingere gli editor a lasciare il progetto
3. Capire le dinamiche del drop-off, specialmente per gruppi sottorappresentati (e. g. le donne)
4. Identificare gli spazi (pagine, gruppi di pagine) in cui le dinamiche tra utenti possono essere dannose
5. Aumentare la consapevolezza sulla salute della comunità

## 4 lingue + analisi automatiche



1. Analisi automatiche indipendenti dal linguaggio e visualizzate attraverso delle dashboard
  - **30 lingue**
  
2. Analisi approfondite (alcune dipendenti dal linguaggio)
  - **4 lingue: en, es, it, ca**



itWikiCon  
2020  
online edition

# Come

Metodi e ricerche correlate

# Metodi

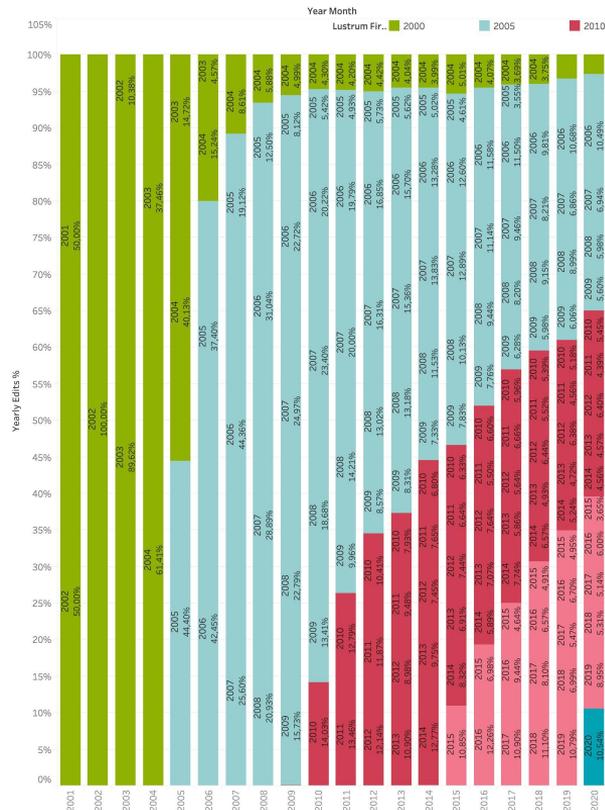
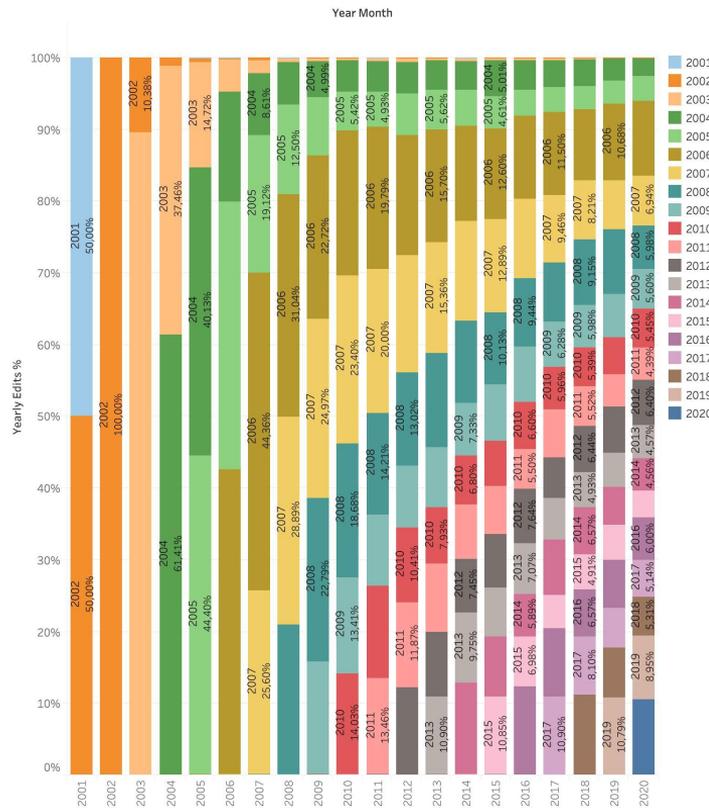


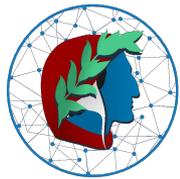
1. Edit (`mediawiki-history-dump` )
2. Discussioni e emozioni
3. Revert e controversie

# Edit versus anno di iscrizione



Percentuale degli edit in base all'anno di iscrizione degli utenti.





itWikiCon  
2020  
online edition

# Quando

Date e aggiornamenti sull'avanzamento del progetto



- **Durata del progetto:** Settembre 2020 / Settembre 2021
- Input dalla comunità: realizzazione al più presto di alcune dashboard di modo che la comunità possa dare del feedback
- Update del progetto su [Meta](#).



itWikiCon  
2020  
online edition

# Chi

La squadra

# Ricercatori



- Cristian Consonni (🔗) (it-N, en-5, fr-4, es-2, de-1)
- David Laniado (🔗) (it-N, es-5, ca-5, en-4, fr-2)
- Pablo Aragón (🔗) (es-N, en-4, ca-4, fr-2)
- Marc Miquel-Ribé (🔗) (ca-N, es-5, en-4, it-3, ro-2)
- + 2 studenti da UniTN (e altri in arrivo)



# Advisor



- Miriam Redi ([link](#))
- Aaron Halfaker ([link](#))
- Patricia Horrillo ([link](#))
- Margeigh Novotny ([link](#))

# We Need You



Vogliamo **coinvolgere**:

- la comunità (specialmente in en, es, it, ca)
- admins, checkusers, patrollers
- gli user groups e i gruppi affiliati

# Dalla lavagna



- Qual è la cosa peggiore che ti è capitata nella tua vita di wikipediano/a?
- Quale è la cosa più bella che ti è capitata nella tua vita di wikipediano/a?
- Hai mai pensato di abbandonare il progetto?
- Quali sono i comportamenti che ti danno più fastidio o ti mettono più a disagio su Wikipedia?
- Pensi che Wikipedia sia un luogo amichevole?



itWikiCon  
2020  
online edition

# Grazie per l'attenzione

`cristian.consonni(at)eurecat(dot)org`

Utente:CristianCantoro

@CristianCantoro



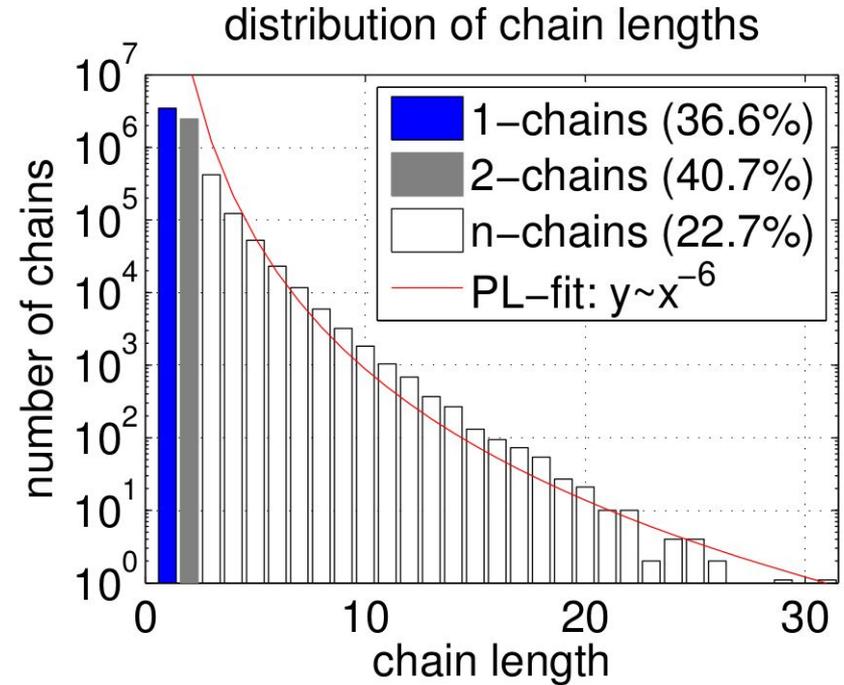
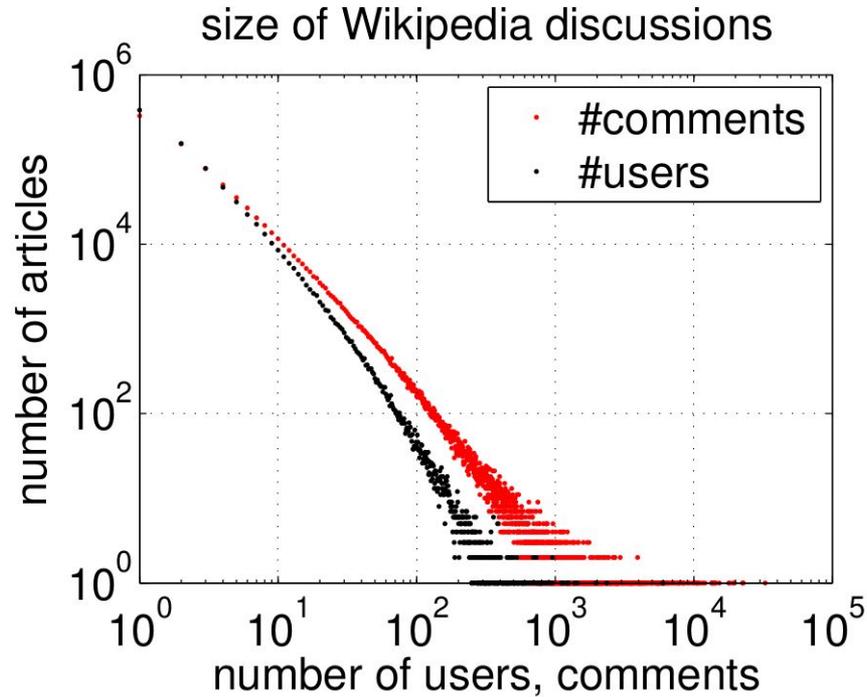
itWikiCon  
2020  
online edition

# Backup

# Idee per loghi



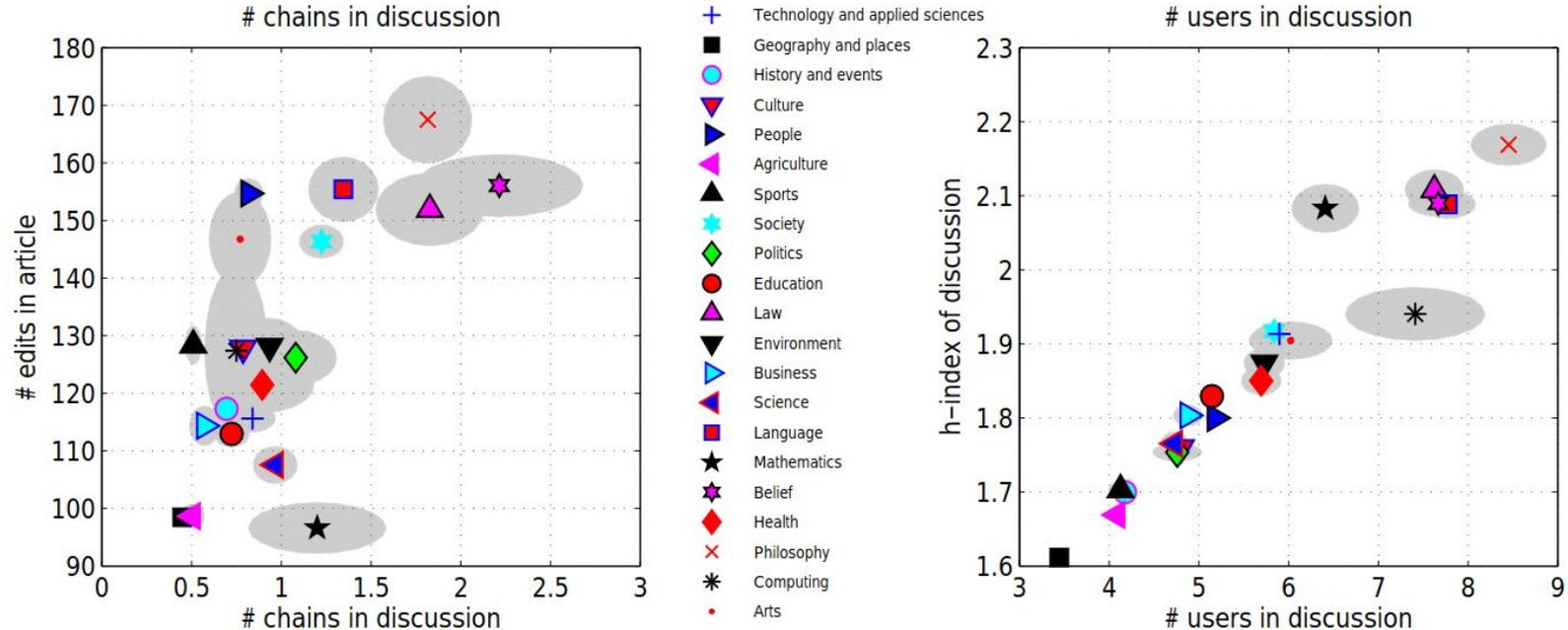
# Discussioni (1)



# Discussioni (1)



## “When the Wikipedians talk” (Laniado et al., 2011)



Laniado, David, et al. "When the Wikipedians talk: Network and tree structure of Wikipedia discussion pages." *Proceedings of the Fifth International Conference on Weblogs and Social Media (ICWSM-11); 2011 July 17-21; Barcelona, Spain. California: The AAAI Press,[2011]. p. 177-84.. Association for the Advancement of Artificial Intelligence (AAAI), 2011. <https://www.slideshare.net/sdivad/when-the-wikipedians-talk>*

# Discussioni (2)



## **“When the Wikipedians talk” (Laniado et al., 2011)**

- Pattern nelle interazioni tra Wikipediani
  - Discussioni tra il “centro” e la “periferia” della comunità (admin ↔ newbies)
- Metriche per caratterizzare le discussioni:
  - h-index → profondità delle discussioni e partecipazione
  - catene di risposte → argomenti controversi
- Indicazioni di differenze strutturali nelle discussioni a seconda del (macro) categorie

# Emozioni nelle discussioni



## “Emotions under discussion: Gender, status and communication in online collaboration” (Iosub et al. 2014)

- Amministratori e utenti esperti hanno un ruolo di collegamento:
  - interagiscono con utenti meno esperti
  - promuovono un ambiente positivo e impersonale
- Differenze di stili di comunicazioni tra uomini e donne:
  - le donne hanno un tono più positivo
  - tendenza a interagire di più con persone dello stesso genere (omofilia)
  - le donne hanno uno stile comunicativo più “orientato alla relazione” e non sono più insicure rispetto agli uomini
  - diverso stile di leadership tra admin donne e uomini



# Controversie



## Pandemia di COVID-19 del 2019-2020 :: layer view



① Currently viewing revision 115700033 from 2020-09-25 00:14:50, with controversy scores calculated since 2020-01-22 11:39:52

## Pandemia di COVID-19 del 2019-2020

**⚠** Le informazioni riportate non sono consigli medici e potrebbero non essere accurate. I contenuti hanno solo fine illustrativo e non sostituiscono il parere medico: [leggi le avvertenze](#).

La **pandemia di COVID-19 del 2019-2020** è la **Pandemia** attualmente in corso della cosiddetta "malattia da nuovo coronavirus".

I primi casi conosciuti hanno coinvolto principalmente lavoratori del **mercato umido** di **Wuhan**, in **Cina**, in cui si vendevano pesce e altri animali, anche vivi. Nelle prime settimane di gennaio 2020 gli scienziati hanno individuato in questi soggetti strane polmoniti causate da un nuovo **coronavirus**, designato **SARS-CoV-2** (*coronavirus 2 da sindrome respiratoria acuta grave*), risultato essere simile almeno al 70% della sua **sequenza genica** a quella del **SARS-CoV**.<sup>[9][10][11]</sup> Alla fine del mese di gennaio 2020 non erano ancora state ben determinate le caratteristiche del virus, sebbene fosse accertata la capacità di trasmettersi da persona a persona, e permanevano incertezze sulle esatte modalità di trasmissione e sulla **patogenicità** (la capacità di creare danno).<sup>[12]</sup> La malattia associata è stata riconosciuta con il nome di **COVID-19**.<sup>[13]</sup>

La prima segnalazione attribuibile al nuovo virus è avvenuta il 31 dicembre 2019.<sup>[14]</sup> ma già l'8 dicembre sono comparsi i primi **pazienti** con malattia sintomatica.<sup>[15]</sup> Il 1º gennaio 2020 le autorità hanno disposto la chiusura del mercato e l'isolamento di coloro che presentavano **segni e sintomi** dell'infezione.<sup>[14]</sup> Il primo decesso confermato risale al 9 gennaio 2020.<sup>[16]</sup> Al 26 gennaio 2020 si registravano più di 4 600 casi di contagio confermati in molti Paesi del mondo e 106 decessi mentre il 13 febbraio tali dati erano già saliti a 49 053 casi e 1 381 decessi. A partire dal 23 gennaio 2020, Wuhan è stata messa in **quarantena** con la sospensione di tutti i trasporti pubblici in entrata e in uscita dalla città,<sup>[17]</sup> misure che il giorno seguente sono state estese alle città limitrofe di **Huanggang**, **Ezhou**, **Chibi**, **Jingzhou** e **Zhijiang**.<sup>[18]</sup> Ulteriori limitazioni e controlli sono stati adottati in molte zone del mondo.

I pazienti accusano **sintomi** simili all'influenza come dermatiti, febbre, tosse secca, stanchezza, **difficoltà di respiro**. Nei casi più gravi, spesso riscontrati in soggetti già gravati da precedenti patologie, si sviluppa **pneumite**, **insufficienza renale acuta**, fino ad arrivare anche al **decesso**. I pazienti presentano anche **leucopenia** e **linfocitopenia**. Dalla metà di gennaio 2020 è disponibile un test per effettuare la **diagnosi** di infezione da SARS-CoV-2, ma al momento sono in fase di ricerca e sperimentazione cure e **vacchi** specifici. Le guarigioni sono principalmente volti a gestire i sintomi e a supportare le **funzioni vitali** anche se sono stati testati alcuni **farmaci antivirali** già utilizzati per contrastare altre infezioni.

Una grande risposta, sia in Cina sia a livello globale, ha seguito un aumento dei casi a metà gennaio 2020, portando restrizioni di viaggio, **quarantene** e coprifuoco. Si nota ad esempio la quarantena della nave da crociera *Diamond Princess* nelle acque giapponesi; il coprifuoco di oltre 780 milioni di persone in Cina; un coprifuoco volontario a Taegu, in Corea del Sud, nonché il coprifuoco di una dozzina di città con oltre 50 000 persone nelle regioni italiane della **Lombardia** e del **Veneto**.<sup>[19][20][21]</sup> L'epidemia è stata dichiarata un'emergenza sanitaria pubblica di interesse internazionale (PHEIC) dall'**Organizzazione Mondiale della Sanità** (OMS) il 30 gennaio.<sup>[22][23]</sup> Aeroporti e stazioni ferroviarie hanno effettuato controlli della **temperatura corporea**, rilasciato dichiarazioni sanitarie e pubblicato segnali di informazione nel tentativo di identificare i vettori del virus.<sup>[24]</sup> Un certo numero di Paesi ha emesso avvertimenti contro i viaggi verso la sola Wuhan, Hubei o la Cina in generale.<sup>[25][26]</sup> Un altro consiglio è che i viaggiatori che tornano da determinate regioni, in particolare Cina, Iran e Italia settentrionale, si isolino da soli.<sup>[27][28]</sup>

Le conseguenze più ampie dell'epidemia comprendono preoccupazioni sull'instabilità economica. La crisi politica ha incluso anche il licenziamento di diversi capi locali del **Partito Comunista Cinese** per la loro scarsa risposta allo scoppio dell'epidemia.<sup>[29]</sup> In diversi Paesi sono stati segnalati episodi di xenofobia e razzismo contro persone di origine cinese e dell'Asia orientale.<sup>[30][31][32][33]</sup> La diffusione di disinformazione e **fake news** sul virus, principalmente online, è stata descritta dall'OMS come una "**infodemia**".<sup>[34]</sup>



## Contropedia

Quali sono gli elementi più controversi in una pagina di Wikipedia?

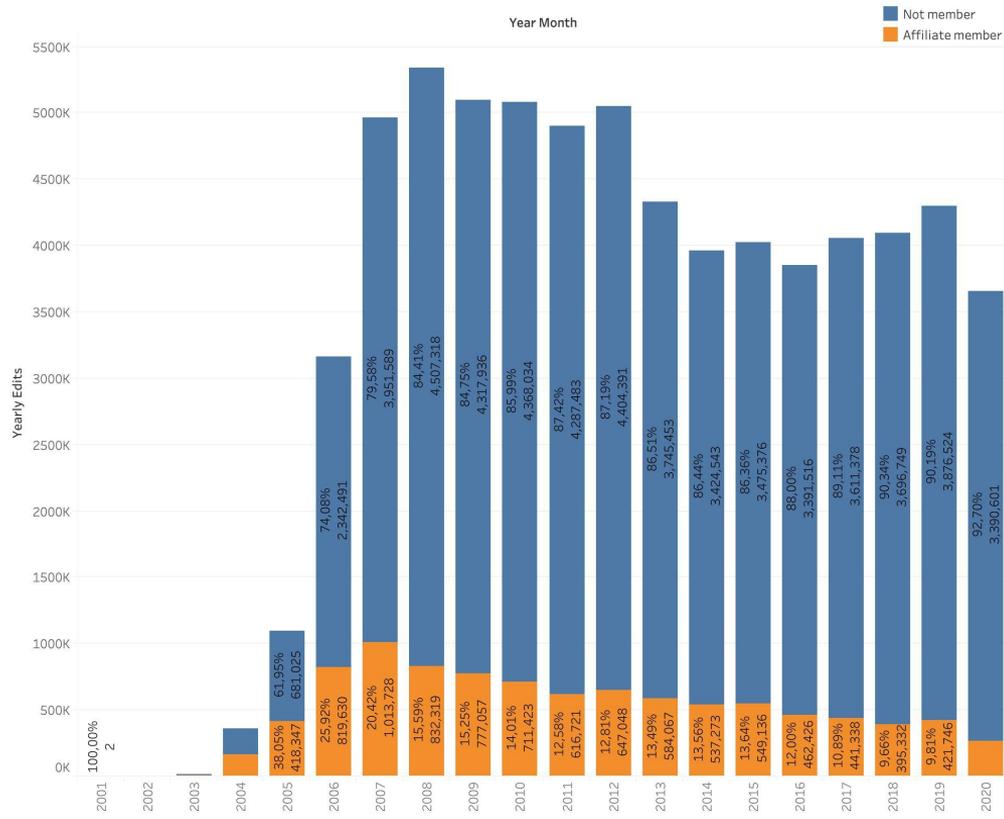
Colori più caldi: più controversia

(più edit di sostanza e disaccordo)

# Edit (soci WM-IT versus non soci)



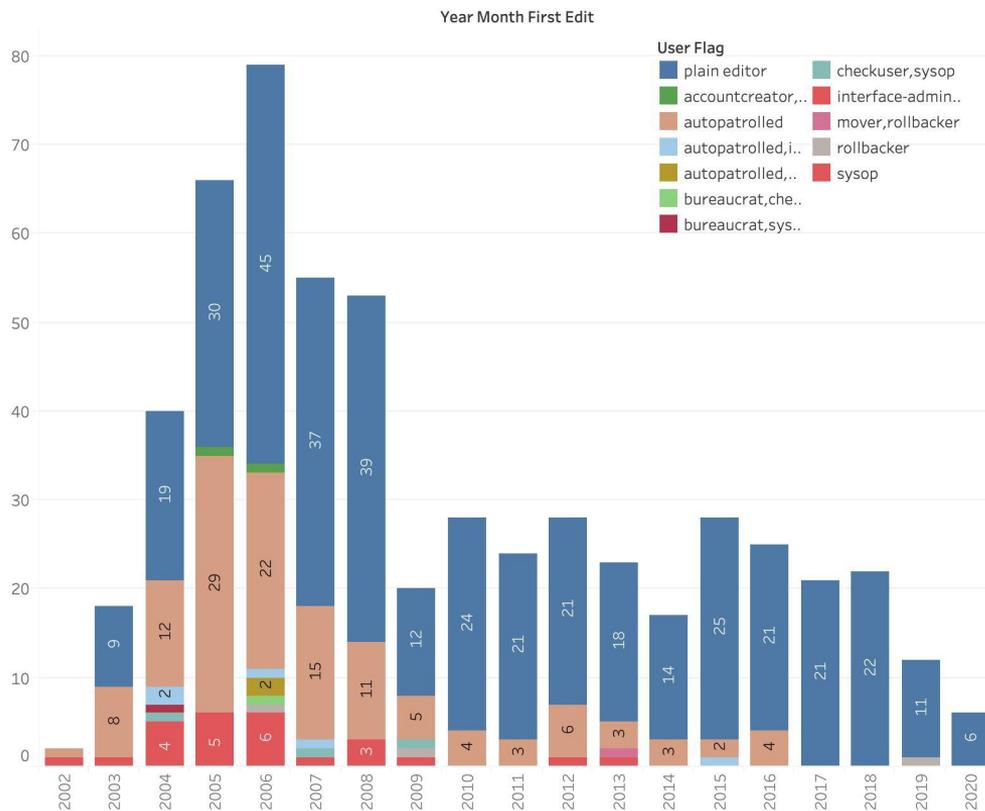
- Il numero di edit è in generale in diminuzione dal 2008, anche se non sembra un trend definitivo.
- In arancione si vede la proporzione di editti fatti da soci di Wikimedia Italia, che sta scendendo anch'esso con il tempo in modo più pronunciato del trend generale.



# Flag (soci Wikimedia Italia)



- Numero di editor rispetto all'anno del primo edit e alla flag utente o combinazioni di flag utente.
- La maggior parte degli editori hanno iniziato tra il 2005 e il 2009.
- Questi stessi utenti sono la maggioranza degli editor flaggati.





## “Societal controversies in Wikipedia Articles” (Borra et al., 2015)

- Conteggio degli edit nelle frasi che includono un wikilink
- Si considerano solo gli edit sostanziali e in disaccordo:
  - sostanziali -> si escludono i vandalismi or anti-vandalismi
  - in disaccordo -> eliminazione di contenuto
- Approccio indipendente dalla lingua, basato solo sull'attività degli utenti
- Classificazione degli elementi più controversi in una voce
- Timeline che mostra quando ogni elemento era controverso

