

Social IME - 辞書共有の問題点

2024-03-24

- Social IME サポート掲示板での投稿
- IPA に連絡
- 結果
- IPA に再度連絡

Social IME サポート掲示板での投稿

誤変換リスト ななし 2009/06/29 (Mon) 02:18:46

ひどい単語登録がいくつかあったので、削除してもらえないでしょうか?
★よみ 誤変換の単語
という内容で記述してます。

—————<以下、誤変換りすと>—————

60 60歳のおじさんですがいいですか?

はうすてんぼす つまらん

げー「

しろ ¥¥C[0]

しろ しる

しろ 占いC0 : 。

くろ 百計のクロ

しろ 占いC0 : ●

くろ 黒田 (ex.23163)

ひと 春夏秋冬

ひと ヒトモドキ

ゆきち 福沢祐麒

こいづみ マッガーレ↓

こいづみ ふんもっぷ

こいづみ ガチホモ

はとやま 日教組と共に日本をつくります

はとやま 在日が総理になつたら素晴らしいことです

はとやま 日本は日本人のものでない

ふくだ (ひ)

ふくだ 福田 洋司

ごみ 海岸ハングルゴミ

ごみ 朝鮮人

あか akaakamimi

かく ?Caution!!?

ちょうせん 朝癬

ちょうせん 犬を生きたまま皮剥ぎ

ちょうせん 日本海のハングルゴミ

ちょうせん 日本のEEZで糞と毒を撒き続ける

ちょうせん プロ被害者

ちょうせん 高麗棒子

にほん シナチョンの餌食

おんな 肉便器

たけしま 竹島 (日本の固有の領土)

ちゅうごく 犬を生きたまま皮剥ぎ

Re: 誤変換リスト - ぬるぼ 2009/06/29 (Mon) 19:38:20

別に誤変換じゃないと思うけど。

スラングが駄目だってなら「モナー→('∀')」なんかもアウトでしょ。

個人情報とかじやなきや問題無いと思うが。

Re: 誤変換リスト - hm 2009/07/04 (Sat) 15:18:54

げー「は合ってるだろ?ロシア文字のゲーなんだから

標準のIMEでもげーで「が出てくるだろ・・・ちょっとは調べてから書き込め
よ

Re: 誤変換リスト - 山本 誠字 2009/07/09 (Thu) 02:35:35

いい加減許可制で削除させてほしい

逆に使いにくくなってきてる

『やまもと』 変換で出てくる 『山本 誠字』 はいらんだろ

無題 - 素人 2009/07/09 (Thu) 04:11:52

よく使う単語が変換の上位に来る機能があつたらいいと思います。

使わない単語を削除する機能があれば不適切な変換にも少しは対応できると思います。

モナー→('∀')

これはOKでしょ。モナーという単語は2ちゃんねらでない人は使わないだろうし、これを見て不愉快になる人はあまりいないと思う。

でも誤変換リストに載ってるやつとか本当にそう使う人がいるの?

いたずら対策しないとそのうち誰にも使ってもらえなくなると思います。

Re: 誤変換リスト - えと 2009/08/08 (Sat) 04:40:53

別に誤変換じゃないと思うけど。

いつから「おんな」で「肉便器」を出すのが正しい変換になったんですか。:-/

スラングが駄目だってなら

誰もそんなこと言ってないですよ。

「モナー→('∀')」なんかもアウトでしょ。

アスキーアートと「日本海のハングルゴミ」などは別物でしょう。

「とあるスラングがアウトならモナーもアウト」とはとんだ俺様基準だ。そもそも「日本海のハングルゴミ」というスラングを聞いたことがない...

<http://www.google.co.jp/search?q=%E6%9C%A8%E6%9B%8B%EF%BC%88%EF%BC%89>

→検索結果 1 件

AAもスラングも程度って物があるので。

子供もSocial IMEを使う可能性があるわけだし。

「にほんのちょうせんがはじまつ」で
「シナチョンの餌食の犬を生きたまま皮剥ぎがはじまつ」と
変換されたらガッカリですよ。

Re: 誤変換リスト - えと 2009/08/08 (Sat) 17:59:09

お試し版を使ってみましたが、

<http://www.social-ime.com/shared-dictionary.html>

「あさひ」「あさひしんぶん」も

「まいにち」「まいにちしんぶん」も

予想通りの変換結果だなあ。

(「さんけい」「さんけいしんぶん」は問題なし)

「ばいこくど」の変換結果に至っては

「民主党藤田幸久」「河野洋平」「電通」「小沢一郎」

ですよ。

研究・開発者 奥野陽

本名さらして放置するのは危険すぎる気が…

奥野さんのブログのコメント欄に投稿したけど、

<http://d.hatena.ne.jp/nokuno/20090806/1249563120#c>

その後もスルーしてブログもtwitterも更新されている。

<http://d.hatena.ne.jp/nokuno/20090808>

<http://twitter.com/nokuno/statuses/3190789495>

過去のコメントに対しては返信されているので、

<http://d.hatena.ne.jp/nokuno/20090706/1246783231#c1246970063>

コメントに気づかないというわけでもなさそうだし…

「放置が適当」というご判断なのかなあ。

Re: 誤変換リスト えと 2009/08/08 (Sat) 18:13:59

他の例

あべしんぞう 文鮮明のお友達

うつくしいくに 鬱苦死萎国

ちよん 誣告4千倍

すまっぷ 犯罪者グループ, メンバーの40%が犯罪者

奥野陽さんの知り合いのかたがいたら、一度連絡してあげたほうがいいかも…

ご自分で判断してこの辞書になっているのなら、もう何もいえないけど。

Re: 誤変換リスト えと 2009/08/09 (Sun) 00:19:27

他の例(2)

くさなぎつよし シンゴー, スマッパ, 犯罪者

ふくだやすお あなたとは違うんです

のむひょん 大酋長

かんなおと 韓直人, 菅工作員

わだあきこ 朝鮮人

ちょうせんじん リアルヤフー

にーと 仕事しろ, 日本の恥, 自宅警備員

うよく 右翼の9割は在日チョン

こうさくいん 金武貴

きょうさんとう 産業空洞化支援政党

そうか そーかそーか, カルト教団, 創価の献金術士, 在日

そうかがっかい カルト教団, 池田大作ファンクラブ

ざいにち 在日住民税免除, 在日住民税半額, 在日強姦, 在日強姦魔, 在日特権, 在日犯罪, 在日通名報道, 在日なりすまし, 在日エセ右翼, 在日カルト, 在日カルト宗教, 在日レイプ, 在日不法投棄

いけだだいさく 成太作, 池田犬作

じみんとう けいだんれんのポチ, 共産党を支持しましょう, 朝鮮民主よりも
しな自民党

わたなべつねお 球界のゴミ

きょじん ゴミ売り虚塵軍

じゃいあんつ 邪異暗津

おしおまなぶ LSD, MDMA

さかいのりこ 母子失踪

れいぶ 朝鮮

ろりこん 不治の病
てれびあさひ 捏造朝鮮テレビ
にってれ 売国奴, 捏テレ
NHK 犬HK
ちゅうごく 中国毒, 特亜
こうめいとう カルト教団, 創価学会, 在日, 売国奴集団
おおさか 春日歩, また大阪か
なごや 名古屋市西区又穂町xxxx
しんじゅく 心中苦
ばか オードリー春日
きちがい 松岡美羽
へんたい 変態毎日, 変態毎日新聞, 森近 霖之助, 毎日新聞
きむち <丶、丶'>
ぶた 青島
うえとあや 綾戸智恵
うんこ YUI
せくはら 死に誘うゲベル
ただの ただの一度も敗走はなく
こうのようへい 紅の傭兵, 売国奴
ろしあ 北方領土を返せ!

最初の指摘が2009/06/29だから、1ヶ月放置ですか…

[8月10日追記]

みんしゅとう ミンス, 傾国売国奴, 売国奴集団, 旧日本社会党左派, 民主党パチンコマネー, 民主党パチンコ献金, 民主党同和献金, 民主党在日パチンコマネー, 民主党在日パチンコ献金, 民主党在日献金, 民主党沖縄ビジョン, 民主党沖縄譲渡, 献金事件

みんしゅ 民主躍進, 国賊民主党藤田幸久, 売国奴, 売国奴の群, 日教組の犯罪で当選, 朝鮮民主躍進, 民主朝鮮

じみんとう 国賊

じみん 民主朝鮮よりましかも

Re: 誤変換リスト - nanasi 2009/08/09 (Sun) 03:41:55

えと氏
くさなぎつよし「裸の王様」「裸だったら何が悪い」
あめりか「とにかく胸でかい」「バスト平均Dカップ」
いたりあ「トマト箱の妖精」
かんこく「姦(奸)国」「世界一の誣告罪」
ちょうせんじん「リアルヤフー」
とかもあったな

IPA に連絡

Re: 誤変換リスト - えと 2009/08/09 (Sun) 21:31:57

「ななし」さんの投稿に奥野さんが気づいておられないのかなあと思って、奥野さんのブログのコメント欄に書き込んだりメールをお送りしたのですが、全部返答なしです。

「辞書が荒らされていることに気づいておられないのでは」と心配していたのですが、どうやら意図的にこういう辞書にされているみたい。ブログもtwitterも普通に書き込んでおられますしね。

というか今思うと、奥野さん自身がこれらの候補を足したのかも…
(「ぬるぽ」さんが奥野さんだったりして)

いちおうIPAのほうにも現状を報告しておきました。

<https://www.ipa.go.jp/about/inquiry/mailform01.html>

(Social IMEはIPAの援助を受けていたそうなので)

これで「知らなかった」ということはなくなるでしょう。

結果

奥野さんは一部の登録を削除し始めました。IPAの担当者様が奥野さんに電話で連絡してくださったそうです。しかし危なっかしさは相変わらず。

Re: 要望のあった単語を削除しました - えと 2009/08/11(Tue) 01:00:49

「ななし」さんが最初に注意喚起されたのは「2009/06/29 (Mon) 02:18:46」

です。

なぜ1ヶ月以上も放置なさっていたのでしょうか。

毎日数百件、累計数万件もの新規登録があり、それらを毎日全て確認することはできません。

私なら1日分の新規登録を見ただけで、「やばい単語が多いなー。管理しきれないから登録は一時中止しよう」と思うことでしょう。奥野さんはそれをしなかった。要は奥野さんの許容範囲だったということでしょう。だからこそ注意喚起も放置できた。

それと数万件程度の登録であれば、適当に条件設定をして絞り込みをかけられると思います。私は「はてなダイアリーキーワード」(累計245,000件)を処理して自分の辞書に組み込んでいます。

要望のあった単語を削除しました。

「あべしんぞう」「はとやま」などはOKです。

こうのようへい 紅の傭兵, 売国奴

かんこく 日本のEEZで糞と毒を撒き続ける, 日本海のハングルゴミ, 奸国, 姦国, 世界一の誣告罪 <丶`丶`>

は変化なし(これは奥野さんのお考えなのかもしれません)。

まだまだ危うい登録は残っているようです(一部伏字にしました)。

おざわいちろう ブルドッグ, 売国奴

はんにち 河野洋平

でんわ 0120-444xxx

めるあと suexxx@

がっこう 薬師寺朋光

うんこ 松崎しげる, 崇康

こうのたろう 売国奴, 河野太郎改め河 太郎, 河野太郎日本人種解体

ねつぞう 朝日新聞, 每日新聞, 黄禹錫

なんきん 通州虐殺

さんご 朝日新聞

はんりゅう 寒流

はげ 中居正広, 小倉智昭, 棚原貴志

ぱちんこ 朝鮮玉入れ, パチンカス, ちんこ
しゃみんとう 北朝鮮

「きょう」で個人宅の住所も出てきますね。
特定の生徒を中傷するような登録もあるかもしれません。
ユーザから送られた登録語は一旦辞書から外すほうが良いと思います。
それと、送られた登録語は「はてなダイアリーキーワード」のように公開しておけば、みんながチェックしやすくなるのではないか。
<http://d.hatena.ne.jp/hatenadiary/20060922/1158908401>

今後について うにコード 2009/08/20 (Thu) 21:58:18

Social IMEを使い始め数ヶ月経ちますが月が経つ事にクセのある変換候補が多くなってきていると感じます。というのも、特に害はありませんが『まりさ → 魔理沙』を入力するのは日本人口の1万分の1もいるか否かな訳で、ブランド名とか社名とか一般人でも使う横文字すら変換できないことが変なのではないかと感じているんです。

当然これらは、職業などユーザーの環境にあっても異なるとは思いますが、Social IMEが今後どのようなソフトとして完成を目指すのかお聞きしたいです。

具体的に言うと、ネットユーザー向けを目指すのか、ATOKやIMEを超える変換精度を目指す(一般の人向け)なのか、もっと様々な環境(OS・ブラウザ)に対応させるのか、未踏事業の完遂が目的だからとりあえず変換候補データが言葉や単語として効率よく処理されれば中身はどうでもいい、でも構いません。

ただ何かと抜けている点が多すぎます。

IPA に再度連絡

2009/08/21

しばらくぶりにSocial IMEの無茶登録をチェックしましたが、あまり変化はありません。

ますこみ 変更捏造報道機関
ゆにせふ (天下り詐欺団体) 日本ユニセフ協会
あぐねす (天下り詐欺団体) 日本ユニセフ協会
あぐねすちゃん (天下り詐欺団体) 日本ユニセフ協会大使
はんぐる 日本人が教育した
ごうかん (強姦), 朝鮮
てれび マスゴミ
へんこう 偏向朝日新聞, 偏向毎日新聞, 朝日新聞、毎日新聞
さけい (左傾), 朝日新聞, 每日新聞
うけい (右傾), (なし)

■以前に報告したが変化なし
こうめいとう カルト教団, 在日, 売国奴集団
へんたい 変態毎日, 変態毎日新聞、毎日新聞

ユーザの登録語を

「朝日」「毎日」「偏向」「売国」「朝鮮」
「韓国」「中国」「捏造」「右」「左」
などで全文検索すればまとめてチェックできると思うのですが、
奥野さんは指摘された単語だけを渋々削除しているように見えます。

危ない登録はまだまだ残っていると思いますので、

- ・システムからユーザの登録語を一時的に切り離す
- ・ユーザからの登録語を「はてなダイアリーキーワード」のように公開して、誰でもチェックできるようにする

<http://d.hatena.ne.jp/hatenadiary/20060922/>

ということが必要だと思います。

主義主張は個人の自由で良いと思いますが、かな漢字変換でそれをやられると、ドン引きするし使いにくくなる一方です。

公的機関が支援した制作物としてもどうかと思います。

奥野さんの日記

<http://d.hatena.ne.jp/nokuno/20090815>

2009-08-15

夏コミに行ってきました

奥野さんの認識としては「急を要する部分の処理は終わった」ということなのでしょうね。

[HOME](#)