

誹謗中傷情報のフィルタリング結果を表示するインタフェース

石丸隆司

Ryuji Ishimaru

1 研究概要

1. 作りたいもの

インターネットにある誹謗中傷ワードをフィルタリングして、それ以外の情報やコメントを表示させるインタフェース。

2. 誰が使うか

インターネットを使う方々。主な対象として、子供にネットを使わせるのに対して不安がある親である。このシステムを使うことで意識的、無意識に見てしまう誹謗中傷ワードを見ないようにします。そうすれば、ネットをあまり知らない子供たちはネット上の誹謗中傷を目にすることなくインターネットを利用でき、親も安心する。また、誹謗中傷ワードをフィルタリングかけることで人に見られる機会が減り、ネット犯罪の防止にもなると考えられるので警察の方々にも良いシステムだと思う。

私たちのような若者やインターネットに慣れている方は、ネットの誹謗中傷のワードに左右されることはあまりないと思う。しかし、知りたい情報に対して関係ない誹謗中傷のコメントなどがあると見ていていい気分ではないと思う。それに、自分の情報収集に誹謗中傷ワードは関係ないものなので、それをフィルタリングすることで情報収集の効率上げることが出来る。

このように誹謗中傷ワードにフィルタリングをかけることで目的は違って多くの方が利用できるようにする。

3. どこで使うか

パソコンや携帯などのインターネットが使える環境でフィルタリングが使えるようにする。家などの場所は特に指定しないでどこでも使えるようにする。

4. いつ使うか

インターネットの閲覧時である。インターネットの情報は、誰かに見てもらいたいから書き込んでいるので誹謗中傷のワードがほかの人に見られないようにする。このフィルタリングが実用することで誹謗中傷のワードを書き込んだ側も自分の情報やコメントが見られることがないとわかれば誹謗中傷のワードを書き込むことがなくなっていくだろう。または、言い方を代えることで見ている人の気分を悪くはさせないと思う。

5. どのように使うか

誹謗中傷のワードの割合を抽出して誹謗中傷のワードの割合を表示させる。そして、誹謗中傷のワードを含むコメントを表示させず、それ以外のコメントだけを表示させる (Fig:1)。万が一、誹謗中傷ワードを含む文の中にも自分にとって有益な情報が含まれているかもしれないのでどの部分がフィルタリングによって消されたか表示させる (Fig:2)。これにより誹謗中傷のワードを見ずに閲覧することが出来るうえに、割合によってそのサイトが荒れているのが判断でき、それによってもその情報の信頼度が分かる。誹謗中傷ワードの割合が高ければそのサイトの情報量が分かるため自分の手間を省くことが出来る。しかし、ワードだけでは判断つかない場合もある。例えば「死ぬ」というワードでも「今日の練習きつすぎて死ぬかと思ったー」という文であれば悪い意味ではとらえることが出来ない。このような情報にも適切なフィルタリングが行えるように単語の前後の内容やコメントされた情報の内容によってそのワードが誹謗中傷であるか否かを判断しなければならない。ネットの書き込みは文の構造がうまく成り立っていない場合もあるのでそここのところも考慮して作れたらいいと思う。

6. 何を材料として作るか

誹謗中傷のワード

2 作りたいもののイメージ図

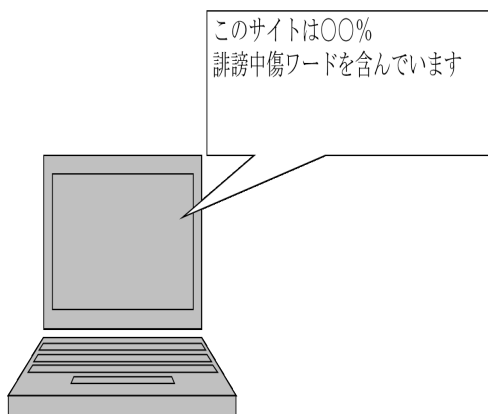


Fig.1 イメージ図

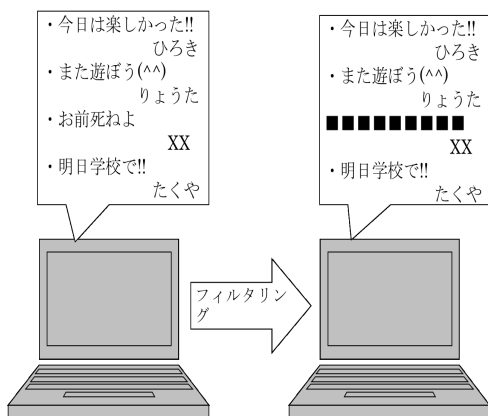


Fig.2 イメージ図

3 次回の発表までにすること

- サイトの情報量に対して誹謗中傷ワードがどれくらいあれば多いのか、少ないのかそのボーダーラインを調べる。
- 卒業生の岩佐さんの研究内容について調べる。