

単語親密度って？

| 単語親密度の例 | |
|---------|-------|
| 青 | 6.344 |
| 愛着 | 5.719 |
| 青刈り | 3.844 |
| 青黄粉 | 1.938 |

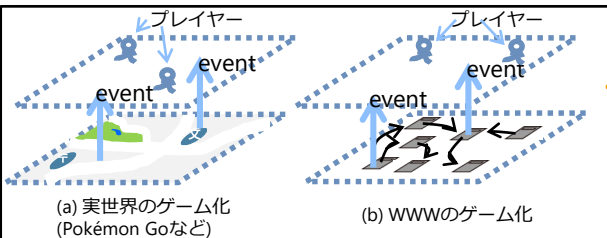
単語に対する馴染みの度合いを表す数値指標だ。1~7の数値で表され幼少期から知っている簡単な単語ほど単語親密度が高く、あまり知られていない難しい単語ほど単語親密度が低くなる傾向があるよ。語彙力チェックや失語症の回復訓練研究などに利用されている。



NTTでは長年単語親密度の調査をしているが、最近生み出された単語にはまだ単語親密度が付けられていない。WWW上にはこのような単語がどんどん出現する。実はテキストモンスターでは、プレイヤーにそのような単語を探して、単語親密度を付けてもらっているんだ。



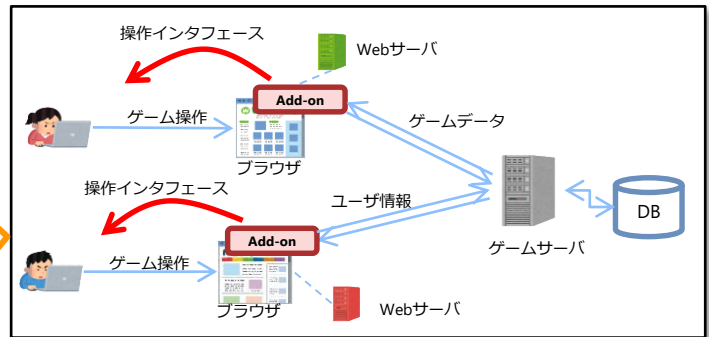
WWWゲーム化技術とは？



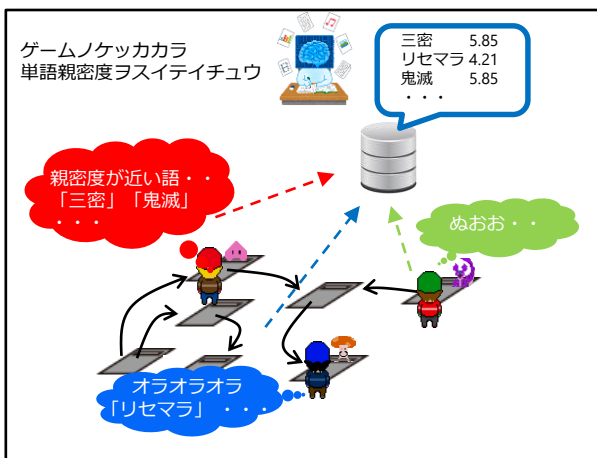
WWW上に仮想的なゲームレイヤを重畳してゲームの舞台を作ってしまう技術だ。この技術を使うと、Webサイト全体、WWW全体を使って、マルチプレイヤーゲームが作れてしまう。



仮想的なゲームレイヤの重畳はブラウザのアドオンを用いて実現している。頑張ればWWWを舞台としたRPGやWebサイトを奪い合うSLGなんかも作れるぞ！



WWWをゲーム化すると何がうれしいの？



WWW上には大量かついろいろな資源がある。それをゲームを通して集めて加工してもらうヒューマンコンピュータシオンが実現できる。テキストモンスターでは、ゲームを通して、ページ上の単語に単語親密度のラベリングをしてもらっているよ。例えば、親密度に近い単語を入力してもらうゲームでは、「単語親密度が割り振られていない単語をプレイヤーが入力した際に、一定の条件を満たせばテキモンの語の親密度を割り振る」といったことをしているよ。



| テキモン語 | プレイヤーが入力した語 | | | | |
|-------|-------------|------|-------|-------|-------|
| 上限 | 膨大 | 機械学習 | 観点 | 公共 | 遺伝子 |
| 5.375 | 5.469 | — | 5.000 | 5.531 | 5.941 |

その他にも、ゲームによって特定のページに誘導する集客効果、ゲームをしながらWWWのコンテンツを利用して学習する教育効果なども期待できる。

「上限」テキモンと近い親密度の単語を入力できているから「機械学習」の親密度は、5.375で大丈夫かな？

