


SNSからの耳より情報抽出手法の提案



甲南大学大学院 自然科学研究科 服部 祐基
甲南大学 知能情報学部 灘本 明代

背景

SNSの発達により、インターネット上には多くのコミュニティが存在

- ◆個人の行動や経験談などの情報発信
- ◆ユーザ同士での情報交換



公式サイトや一般のWebページに掲載されていない有益な情報が含まれている。

例えば: 諏訪湖花火大会

神奈川方面からなら諏訪インターの一つ前の諏訪南インターから下道が良いかと。
やはり自動車よりも電車で行かれることをお勧めしますが...



コミュニティから有益な情報を収集することが重要となっている。

膨大な情報が存在するコミュニティからユーザの欲しい情報を取得するのはとても困難である。

目的

イベントを対象

コミュニティのコメントから

閲覧者の知らない、信憑性のある有益な情報を抽出

耳より情報



耳より情報を優先的に提示することにより、ユーザは効率的に有益な情報を取得することが可能となり、新たな知識の獲得が期待できる。



耳より情報とは

- 実際に経験したことに基づいたコメントはある程度信憑性がある.
- 多くの閲覧者が有益と感じた情報の中には共通のキーワードが含まれている.

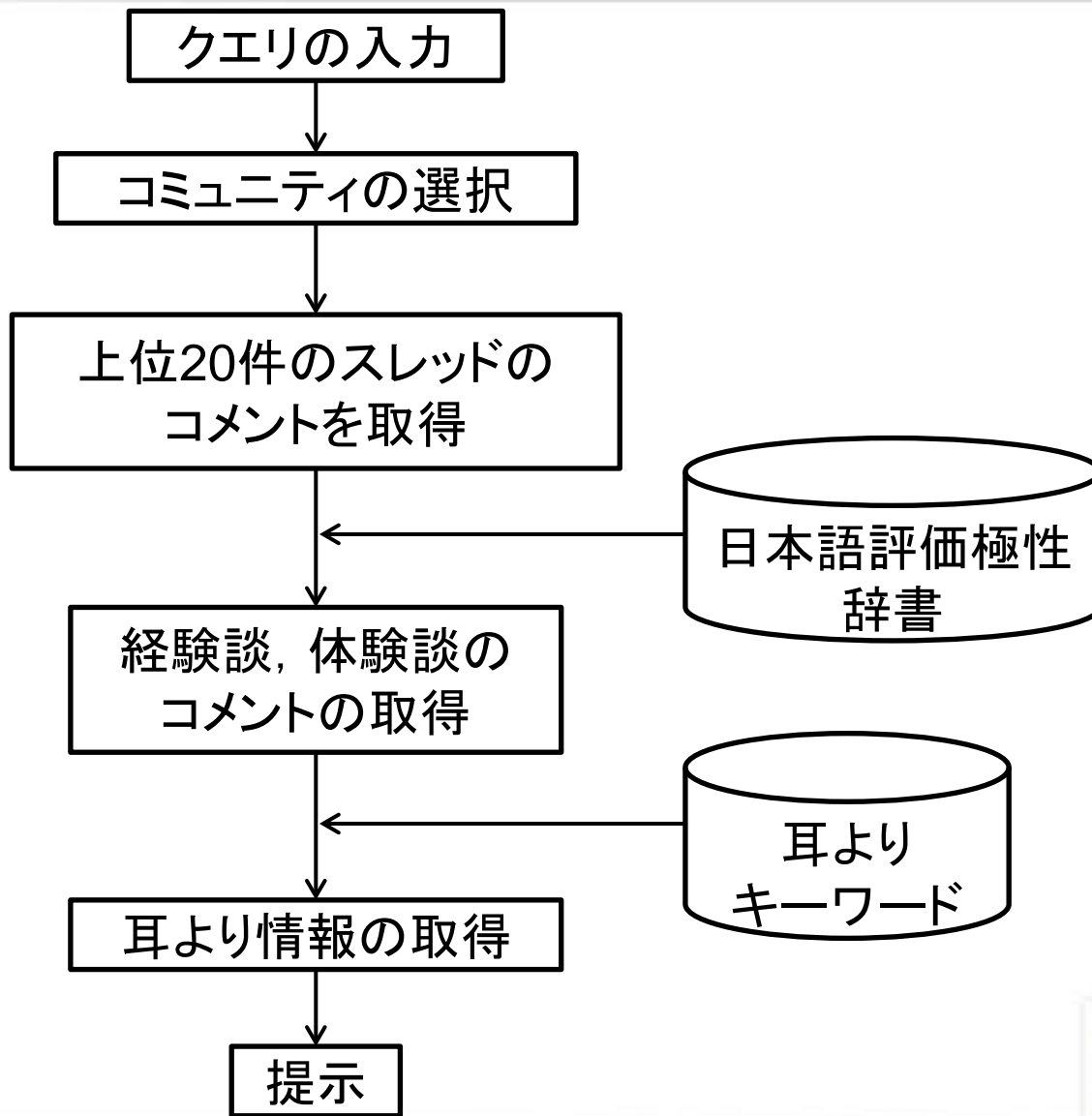
閲覧者が知らない, ある程度信憑性があり、有益と思える情報

経験談に基づいて
書かれているコメントの抽出

被験者実験より
耳よりキーワードの抽出



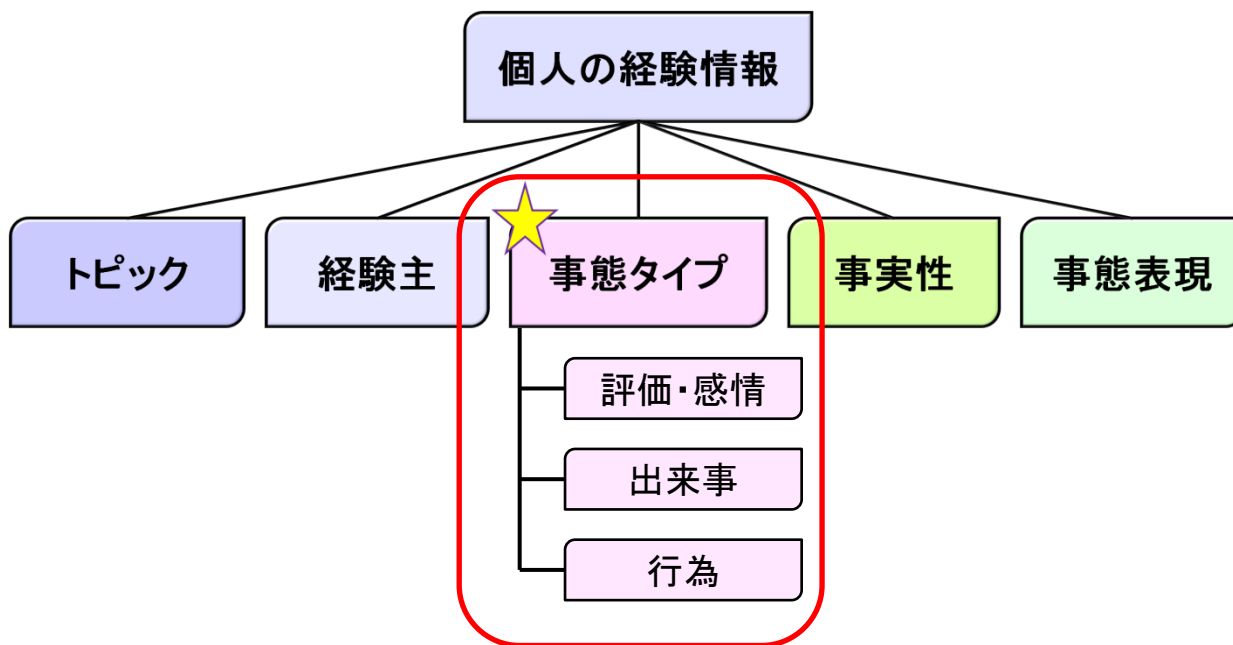
システムの流れ



経験談, 体験談のコメントの取得

経験マイニング
(乾ら)

Web上に大量にあるテキストから個人の経験情報を、構造化情報として抽出することを目的



経験談, 体験談のコメントの取得

経験談、体験談

行為

+

評価・感情

出来事

日本語評価
極性辞書

日本語評価極性辞書

動詞、形容詞、名詞に対してポジティブとネガティブの評価極性を付与している辞書

単語	P/N	単語	タイプ
あか抜ける	P	1番	状態
あかく	N	あいまい	評価・感情
あきらめる	N	あく抜け	出来事
あきれる	N	いさかい	存在・性質

行為を表す単語の取得

入手を表す「買う」、利用を表す「食べる」、決定を表す「選択」など

→ イベント関連のコミュニティ30000コメントから行為を表す名詞と動詞各上位50単語を取得

行為

動詞		名詞	
行く	買う	参戦	利用
見る	並ぶ	参加	入場
できる	探す	乾杯	移動



経験情報から耳より情報の抽出

経験情報から有益な情報を取得するために

□実際の耳より情報から耳より情報の要因となるキーワードを抽出。

予備実験

被験者: 5名

データセット: mixiのイベント関連5つのコミュニティ2000コメント

➤5人中4人が耳より情報であると感じたコメントを収集

収集されたコメントから人手により耳より情報の要因となるキーワードを抽出

耳よりキーワード		
おすすめ	の方がよい	いかが
すべき	是非どうぞ	イチオシ
狙い目	穴場	便利
ガラガラ	渋滞	混雑

経験情報を含むコメントに対して、耳よりキーワードが含まれている場合には、そのコメントを耳より情報とし抽出し、ユーザに提示する。