

企業 Web サイト分析に基づいた問題生成を備えた企業研究支援システム

木村 風雅[†] 鷹野 孝典[†]

[†] 神奈川工科大学 情報学部 情報工学科

1. はじめに

就職活動において、業界・企業研究は大変重要である。業界・企業研究はインターネットの普及により企業 Web サイトやリクナビ、マイナビなどの就職活動支援サイトから情報を収集することができるが、企業情報を問題として出題することで、効率的な企業研究支援が行えると考えられる[1][2]。しかし、企業 Web サイトに記述されている情報を題材にして自動問題生成を行う場合、企業 Web サイト内の各ページを問題作成者の意図に応じて選別しないと、出題意図にそぐわない問題を生成してしまう可能性がある。本研究では企業研究を開始する前や開始したばかりの学生が利用することを想定した、企業 Web サイト分析に基づいた問題生成機能を備えた就職活動支援システムを提案する。

2. 提案システム

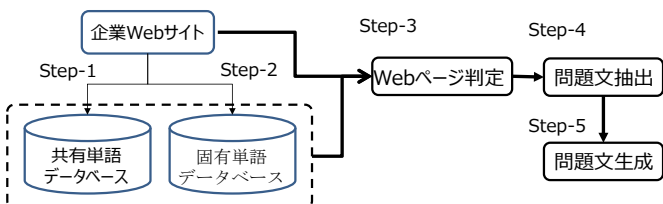


図1 提案システムの概要図

図1に提案システムの概要図を示す。企業 Web サイト分析に基づいた問題生成問題文生成は、下記の手順で実施する。

Step-1: 複数の企業 Web サイト分析による共有単語を、出現頻度に基づいて抽出する。「企業理念」と「求める人材像」に関する共有単語の例を表1に示す。

表1 共有単語の例

企業理念	出現頻度	求める人材像	出現頻度
会社	11	積極性	3
お客様	9	新しい	3
社員	9	チャレンジ	2

Step-2: 個々の企業 Web サイト分析による固有単語を選定する。出現頻度が低くても、重要な単語であれば抽出する。会社 A における「企業理念」と「求める人材像」に関する固有単語の例を表1に示す。

表2 会社 A における固有単語の例

健全, 最強集団, 困難, 価値観, 交友関係

Step-3: Step-1, Step-2 で抽出・選定した共有単語, 固有単語を用いて, Web サイト内の内容判定を行う。

Step-4: Step-3で内容が適合した Web ページから, 共有単語や固有単語を含む文章を抽出する。

Step-5: Step-4 で抽出した文章を, 共有単語, 固有単語に着目して問題文を生成する(表3)。

表3 問題例

(抽出文) お客様との親密性を最も大切にいたします。
↓
(問題) お客様との【 】を最も大切にいたします。
1. 親密性 2. 地域貢献 3. ノウハウ 4. 懸け橋

3. 実験結果

15社の企業 Web サイトから収集した企業理念や求める人材像に関する共有・固有単語を用いて, 新規の25社の企業 Web サイトを対象として適切な問題文を作成可能であるかを評価した。Web サイト分析を行わずに問題生成を行った場合と比較した結果を表4に示す。適合性の判定は人手で行った。

表4 生成問題の適合率の比較

トピック	分析なし	分析あり
企業理念	41%	71%
求める人材像	69%	76%

表5 「企業理念」について抽出した文章の例

適合性	文章
○	堅実性を失うことなく、アクティブに組織を活性化し、社会に貢献する企業を目指します。
×	これらの個人情報、採用活動及び、ご本人へのご連絡と情報提供を差し上げる目的に限定して利用いたします。

4. 実験考察と今後の課題

表4の結果より, Web サイト分析により, 「企業理念」や「求める人材像」といった出題意図に沿った問題生成の適合率が向上可能であることを確認できた。今後の課題として, Web ページの内容判定アルゴリズムの改善を図る事で, 問題生成の精度向上を行っていく予定である。

参考文献

- [1] 中桐齊之, 内平隆之, 米山寛二, “ゲーミフィケーション技術を用いた就職活動事前学習システムの開発,” 情報科学技術フォーラム講演論文集, vol.12, No.3, pp.631-634, 2013.
- [2] 谷口るり子, “情報リテラシー科目に対応した Web クイズの開発と授業への導入方法,” 教育システム情報学会誌,” vol.32, No.4, pp.257-262, 2015.