

インターネット上の違法・有害情報検出技術の研究開発

(1) 研究の目的

ISP やコンテンツ監視事業者が共通で利用可能な「監視業務支援システム」を研究開発し、監視業務のコストを削減すること

(2) 研究期間

平成 2 1 年度から平成 2 3 年度 (3 年間)

(3) 委託先企業

(株) K D D I 研究所

(4) 研究予算 (百万円)

平成 2 1 年度	1 7 9 (契約金額)
平成 2 2 年度	1 6 4 (契約金額)
平成 2 3 年度	1 5 8 (予定金額)

(5) 研究開発課題と担当

課題ア：モバイル対応型情報収集・解析結果提供技術

課題イ：違法・有害情報の大規模コーパス自動構築技術

課題ウ：違法・有害情報の高速高精度検出技術

課題エ：ISP やコンテンツ監視事業者等の実運用環境における実証実験

(6) これまでの主な研究成果

特許出願：国内出願	1 0 件	外国出願	0 件
外部発表：研究論文	3 件	その他研究発表	6 件
報道発表	4 件	展示会	0 件
		標準化提案	0 件

具体的な成果

- (1) 課題ア：違法・有害情報に関する PC サイト・携帯サイトのクローリング作業効率化技術に関する課題において、携帯電話端末と PC が連携した携帯クローラを開発し、年度目標を達成した。これにより、携帯サイトの収集・判定作業を効率化できる。今後、性能評価を実施する。
- (2) 課題イ：違法・有害情報の高精度な監視ツール構築のベースとなる、3 億

記事の大規模コーパスを構築する技術に関する課題において、機械学習向けの人手によるラベル付きのコーパス 3 種類を構築し、年度目標を達成した。

1. カテゴリラベル付与コーパス 2650 万記事(目標：1000 万記事)
2. 品詞ラベル付与コーパス 10 万記事(目標：10 万記事)
3. 品詞・格構造ラベル付与コーパス 1 万記事(目標：1 万記事)

- (3) 課題ウ：違法・有害文書検出手法 MDS(Multi Dynamic Scoring)方式において、H21 年度に課題イで収集したカテゴリラベルつき大規模コーパスを用いて学習を行うことで再現率を向上した。MDS 方式は小規模な学習データにおいては MSS(Multi Static Scoring)方式と比べ性能が劣るが、大規模データを用いての学習により、MSS 方式と比べて適合率を 5%以上向上し、再現率 85%、適合率 75%の年度目標を達成した。
- (4) 課題ウ：これまでの言語解析器で解析できなかった伏せ字やギャル文字といったネット特有の言語を高精度に解析するための手法を世界で始めて開発した。これにより、違法・有害な表現に頻繁に用いられる伏せ字を解析することができ、より多くの違法・有害文書を発見可能となる。研究成果を報道発表し、日経産業新聞第一面に掲載。NHK「おはよう日本」やテレビ東京「WBS(トレたまコーナー)」、リクルート社が運営する Web サイトおよびメルマガ「キーマンズネット」で大きく取り上げられた。
- (5) 課題ウ：Web サイトの背景色や画像数などの外形的な特徴に基づき、高適合率条件下(95%以上)において、違法・有害サイトを高速・高再現率で検出する手法を考案した。提案手法は Web サイトの HTML タグ部分などに注目し、違法・有害サイトに偏って出現するような外形的特徴を発見し、それらを SVM(Support Vector Machine)を用いて組み合わせて判定に利用することで、再現率 68%、適合率 97%といった極めて高い適合率の判定を可能とした。MDS などのテキストベースの違法・有害文書検出手法を利用する前に本手法を適用することで、処理時間を短縮すると共に、システム全体の適合率を高めることが可能となる。
- (6) 課題ウ：係り受け解析を用いることで違法・有害文書を高精度に検出する手法を考案した。違法・有害性の高い表現と係り受け関係にある文節の組を判定に用いることで、従来手法では判定を誤るような表現を正しく判定することが可能となった。小規模データで再現率と適合率がともに向上することを確認した。
- (7) 課題エ：CP および監視事業者向け監視ツールを試作した。今後は実際の事業者等の意見を反映した改良を行い、実証実験に活用する予定である。
- (8) 課題エ：平成 22 年度に予定している評価実験のための実験データを集め

るため、10 社程度に具体的に実験等について説明し、実際の監視事業者
や CP からデータを取得した。

(7) **研究開発イメージ図**
別紙 2 - 1 に記載