

平成21年度「インターネット上の違法・有害情報検出技術の研究開発」の開発成果について

1. 施策の目標

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

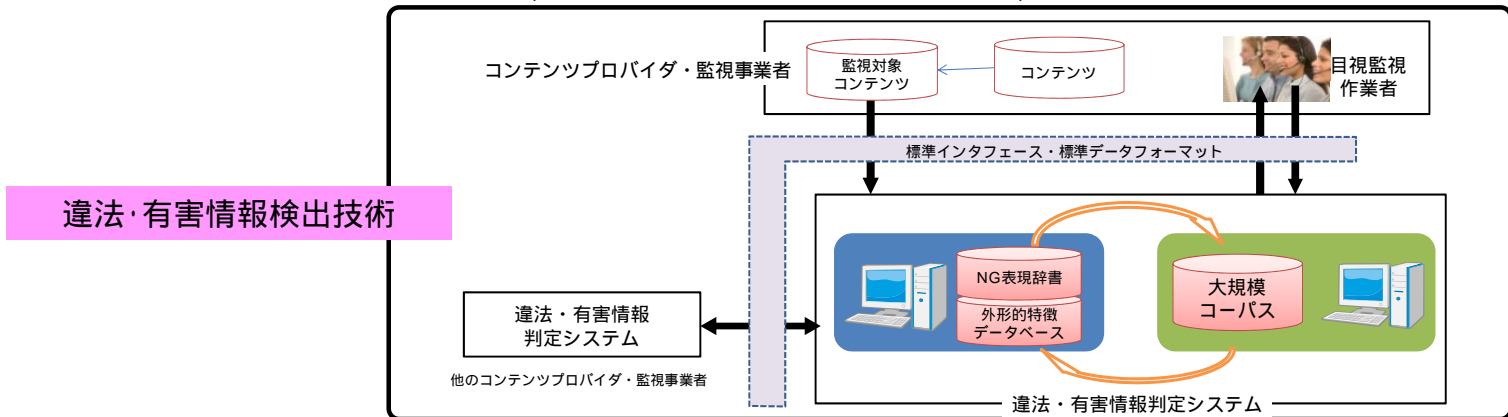
2. 研究開発の背景

インターネットが国民生活に必要不可欠となっている一方で、違法・有害情報による犯罪の助長や自殺の誘因が社会的に大きな影響を与えており、近年では、ISPやコンテンツ監視事業者などによる監視作業が進められている。しかしながら、今後も爆発的に情報が増大すること、およびその内容が多岐に渡ることを勘案すると、それらの監視対象に対して現在実用化されている技術だけでは限界が生じることは推測に難くない。そのため、システムが監視業務を支援する割合をこれまで以上に増やし、人手による業務を軽減するための新たな技術が必要である。

3. 研究開発の概要と期待される効果

以下の3つの要件を満たす監視業務を支援する違法・有害情報検出技術の研究開発する。これにより、コンテンツ事業者や監視事業者が低コストで違法・有害監視作業を実施できるようになり、ISPの自治の促進や安価なASP型違法・有害判定サービスの実現が可能となる。

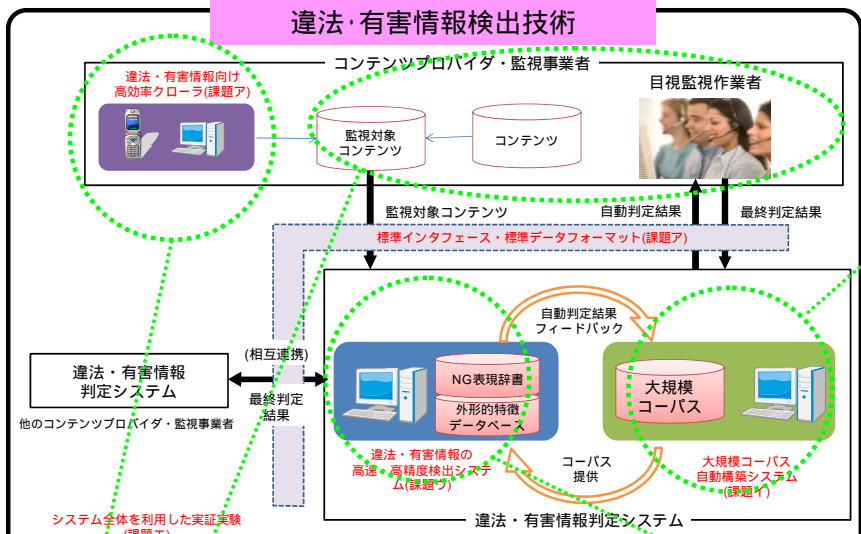
- ・各事業者が共通で利用できること(共通ポリシー)
- ・各事業者の監視ポリシーを反映できること(事業者個別ポリシー)
- ・監視コストを大幅に下げることができること(コスト:人件費、設備費、維持費など)



4. 研究開発の期間及び体制

平成21年度～平成23年度(3年間)
NICT委託研究(株式会社KDDI研究所)

違法・有害情報検出技術の主な成果



- 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

課題ア: モバイル対応型情報収集・解析結果提供技術

- 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の効率化技術に関する研究
- 携帯電話端末とPCが連携した携帯クローラを開発

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

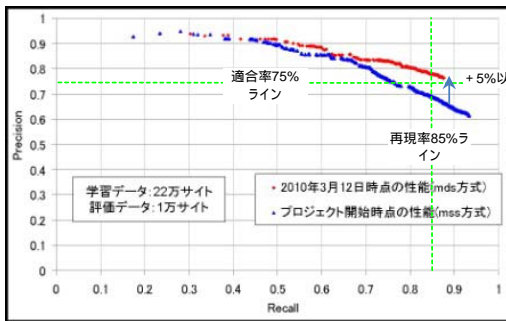
- 本開発の実証実験の実施に向けた環境構築
- CPおよび監視事業者向け監視ツールを試作
- 評価実験のための実験データを集めるため、10社程度に具体的に実験等について説明し、実際の監視事業者やCPからデータを取得

課題イ: 違法・有害情報の大規模コーパス自動構築技術

- 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研究
- 機械学習向けの人手によるラベル付きの3種類のコーパスを構築し、初年度目標を達成
 - カテゴリラベル付与コーパス2650万記事(目標: 1000万記事)
 - 品詞ラベル付与コーパス10万記事(目標: 10万記事)
 - 品詞・格構造ラベル付与コーパス1万記事(目標: 1万記事)

課題ウ: 違法・有害情報の高速高精度検出技術

- 違法・有害情報の監視業務を効率化するため、違法・有害情報の頑健かつ高速高精度な自動検出技術に関する研究
- 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初めて開発
- 今年度収集したカテゴリラベルつき大規模コーパスを用いることで、適合率を5%以上向上
- 年度目標である再現率85%、適合率75%を達成



ギャル文字自動翻訳【日経産業新聞第一面 2009年12月3日】他、NHK「おはよう日本」、テレビ東京「WBS(トレたまコーナー)」にて紹介

1. これまで得られた研究成果(特許出願や論文発表等)

| | 国内出願 | 外国出願 | 研究論文 | その他研究発表 | 報道発表 | 展示会 | 標準化提案 |
|---------------------------|------|------|------|---------|------|-----|-------|
| インターネット上の違法・有害情報検出技術の研究開発 | 10 | 0 | 3 | 6 | 4 | 0 | 0 |

2. 特筆すべき成果

- ・くだけた表現の修正技術について報道発表により研究成果を幅広くアピール
日経産業新聞1面に掲載
TBS番組「World Business サテライト」の「トレたま」コーナーで紹介
NHK番組「おはよう日本」で紹介
リクルート社Webページ「キーマンズネット」に掲載
記事紹介のメールマガジンのタイトルに採用された

さらに、上記発表を基にネット新聞等で数多く取り上げられた。

Yahoo News、livedoor News、BIGLOBENews、msnデジタルライフ、RBB Today、2ちゃんねる、日本漫画新聞、等