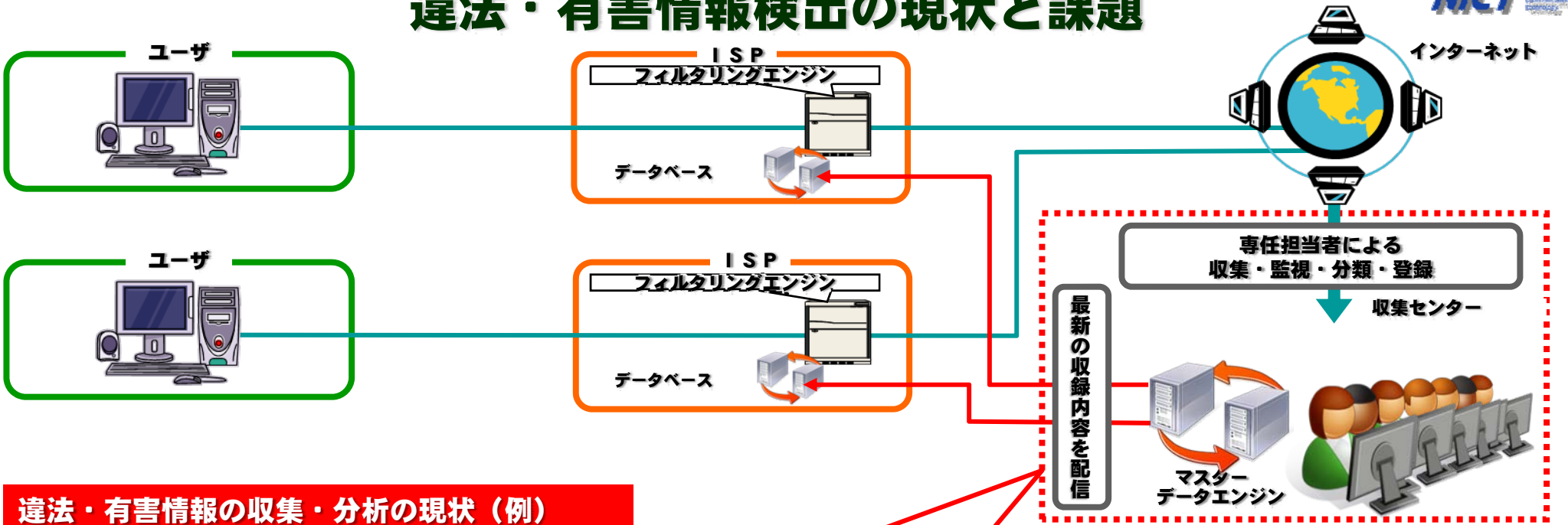
資料3

違法・有害情報を検出するための技術開発について

独立行政法人　情報通信研究機構
木俵　豊

2008年4月2日

違法・有害情報検出の現状と課題



違法・有害情報の収集・分析の現状（例）

違法・有害情報の削除等への対応の迅速化・効率化を図るためには以下の課題の解決が必要。

- ✓ 違法・有害情報の検出の迅速化
- ✓ 違法・有害情報の検出負担の軽減
- ✓ ネット上の情報量の急増に対する対応の限界の克服

違法・有害情報対策に資する研究開発を支援

インターネット上の違法・有害情報の検出を行うための技術開発

時々刻々新たに流通するコンテンツを監視するには、効果的なコンテンツ・チェック方式の高度化が不可欠。

現在は、単語レベルで一致したものを検出するのみで意味解析がされないため、不要なサイトまで検出する場合や違法・有害な情報が検出対象から除外される場合がある。

このため、自然言語技術を活用してWebコンテンツの構造分析を行い、違法・有害情報の検出精度を向上させるための研究開発に、インターネットサービスプロバイダー、コンテンツプロバイダー、関係団体、情報通信研究機構等が協力して取り組むことが必要。

これにより、違法・有害情報の検出を迅速に行うとともに、検出の負担を軽減し、もってネット上の情報の適正化を推進することが可能。

サイト管理のための違法・有害情報検出フロー（例）

①特定のblogや掲示板サイトに対して30分ごとにアクセス。記事の増加量を計測

②'急激な書き込み増加、通常の増加曲線と異なる増加を検知

②NGワード（自殺、学校襲撃、殺人予告、襲撃予告など）の書き込み検知

③'急激な書き込み増加、通常の増加曲線と異なる増加に対して深い情報分析開始

③NGワード後の議論の流れなどについて情報分析開始

④各発信者の意見分析（否定・肯定）により、違法・有害情報を判定

（ISP, 掲示板管理者）

⑤削除、利用停止等の対応

対応結果に基づき判定条件等に反映

違法・有害情報の検出の現状と将来像

サイトから違法・有害情報を出さない

プロバイダによるフィルタリング

サイト運用者

違法・有害情報の削除

サイト管理者

インターネット

フィルタリング

ネット情報収集

違法・有害情報を主に人手により検出

違法・有害情報のフィルタリング情報に反映

当初はこちらでの利用をターゲットとして研究

違法・有害情報の所在情報を提供又は自らサイト内をチェック

情報分析技術を活用し検出の効率化を支援
・迅速な検出
・検出負担の軽減

現状

将来

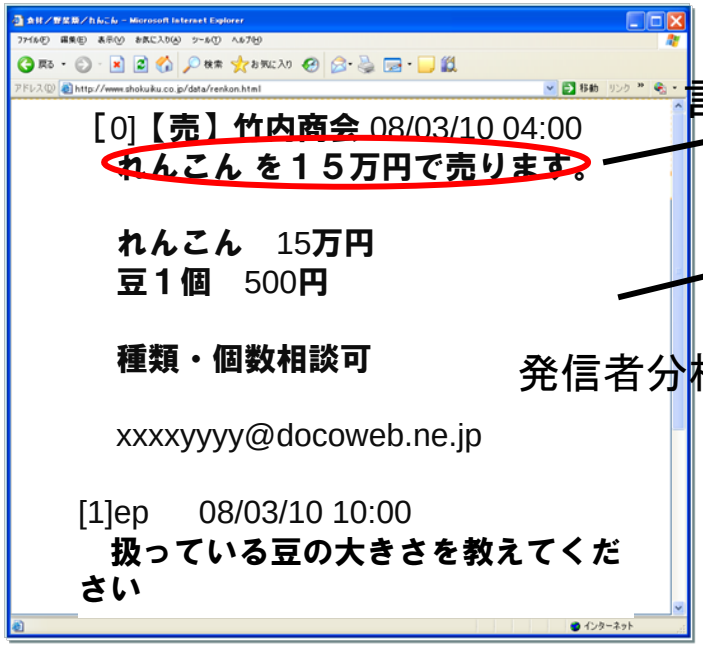
自然言語技術を活用するメリット

課題：単語「れんこん」だけでは、「食品」と「拳銃の隠語」とを区別ができない。

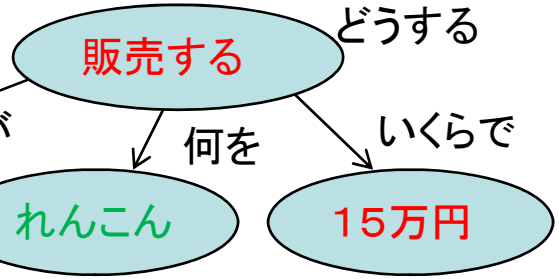
自然言語技術を活用して、「誰が何をどうする」を抽出することで、違法・有害ページの候補を迅速に高精度で判定が可能。

- 言語解析: 著者が事実を陳述
  「れんこんはビタミンが豊富である」
- 発信者分析／ページ外観分析
  発信者：株式会社　日本食品薬化
  掲載場所：企業サイト
  連絡先：　住所＋電話番号あり

発信者が企業で、
有害表現パターンを含まない
事実・意見のページであることを認識



言語解析



発信者：竹内商会：匿名(実世界の企業と認定できない)
掲載場所：掲示板
連絡先：電子メールのみ

発信者分析／ページ外観分析

知識データベースと照合することで
・匿名(実世界での連絡先がない)の人が高額で販売していること
・れんこんは隠語で拳銃を意味すること
から違法である可能性があることを認識

知識
データベース

自然言語技術を活用するメリット

人手で作成・チェックした知識(有害表現パターン、有害語)を元に、自然言語処理技術を用いて大規模Webデータから知識を自動拡張することで、知識データベースの半自動構築が可能

□知識データベース： 有害表現パターンのデータベースの例

「誰が」「何を」「いくらで」「どうする（販売する）」のパターンと有害度の関係を記述

誰が	何を	いくらで	どうする	有害度
匿名 (住所なし)	(拳銃) (覚せい剤) (口座)	高額[1万以上]	販売する	違法の可能性が極めて高い
匿名	*	高額[1万以上]	販売する	18才以下制限
匿名	*	*	販売する	15才以下制限
企業(住所あり)	(書籍)	*	販売する	12才以下制限
…	…	…	…	…

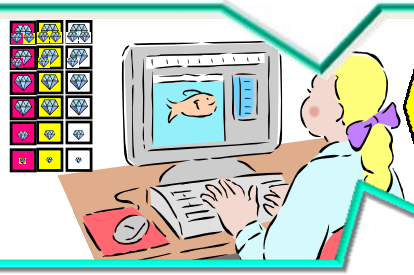
□知識データベース： 有害語の同意語辞書の例

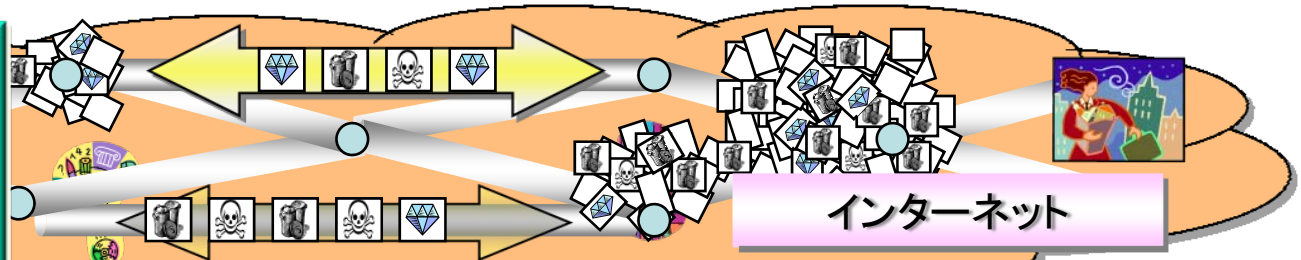
意味	同意語(隠語)				
拳銃	ちゃか	れんこん	レーニン	北京ダック	....
覚せい剤	シャブ	S	スピード	クリスタル	....
…					

情報分析技術の研究開発

【利用例】 アガリスクって本当に健康に良いのかしら？

★WISDOMを使って検索エンジンの検索リストをみると，リスト上位は販売会社ばかりよね。公的機関の内容を見る限りでは健康によいのかどうか疑問のようね。特に最近のWebページの評判は、販売サイトを覗くと90%以上否定的な意見だし、3つめの内容には、意見の矛盾があるみたい、4つめの効果を特に表現しているページの写真は、あるWebページの使い回しみたいだ。これらは信用できないなぁ。





①玉石混淆の情報資源を収集

玉石混淆のデジタル情報

意味内容時系列分析

Webコンテンツの分析技術（クロスメディア分析）

委託研究

強化補完

WISDOM:情報信頼性検証システム（NICT）

情報信頼性の評価基準

外観

Webページの外観的な特徴に基づく分析

内容

情報内容の分析

発信者

情報発信者の分析

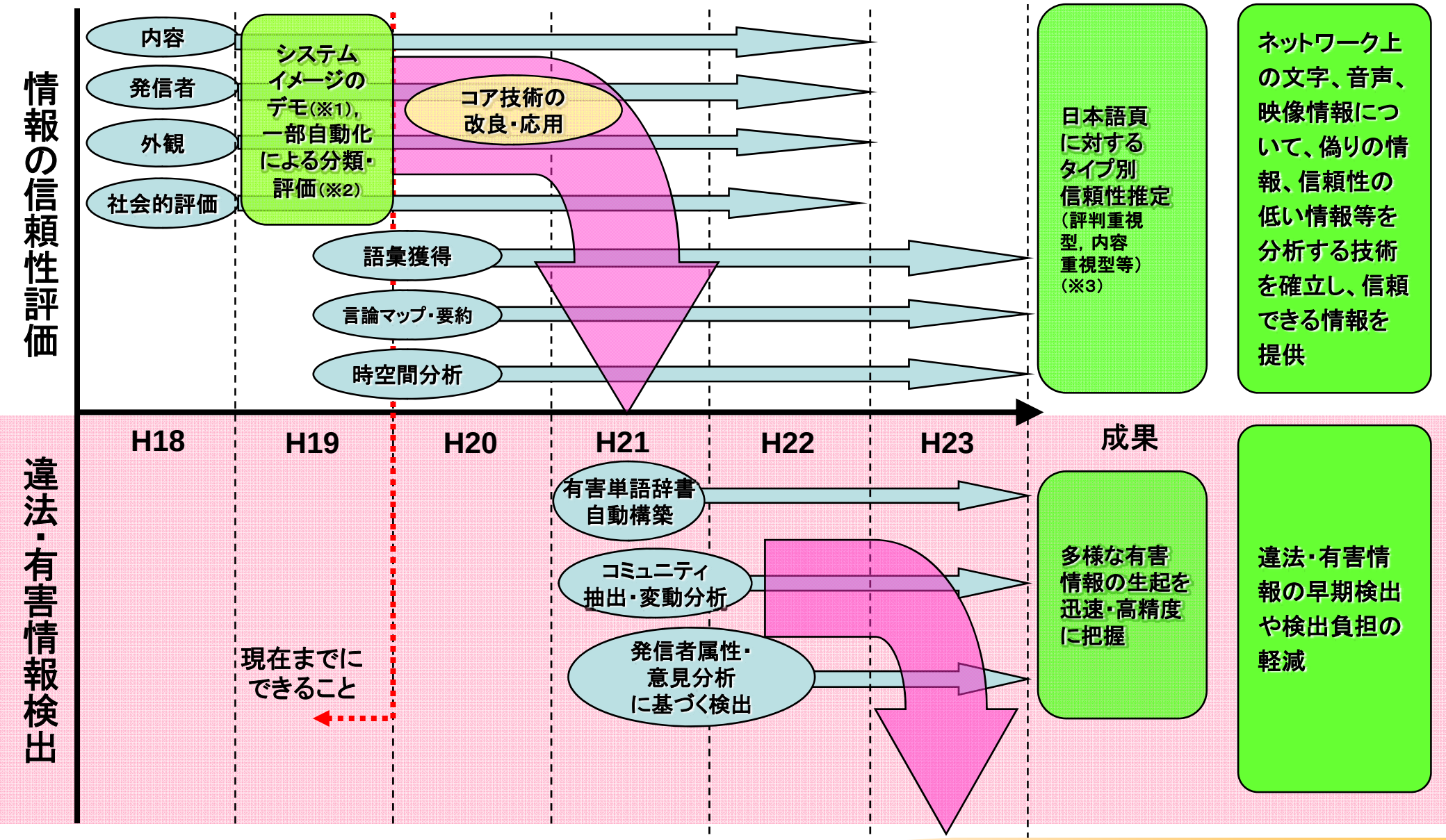
評判

内容／発信者の評価情報に基づく分析

③分析結果を提示，ユーザにインターネット上の価値のある知識情報を発見させる。

②玉石混淆の情報資源を分析

情報分析研究の現状と今後



研究成果は順次提供

※1:日本語2千頁規模,基盤データ(人手)を使用
※2:日本語1億頁規模,発信者・社会的評価の一部自動分析を実装
※3:日本語5億頁規模,内容・発信者・外観・社会的評価の各観点での自動分析を実装