



誰もが安心・安全にWebコンテンツを利活用 できるための情報信憑性検証技術の実現

菱沼 宏之

総務省 総合通信基盤局 電気通信事業部 電気通信技術システム課

平成20年1月30日

情報の巨大集積化と利活用基盤技術開発連携群の活動

～情報爆発時代に果たす日本の役割と連携強化～

シンポジウム



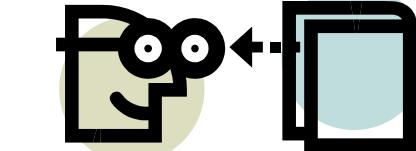
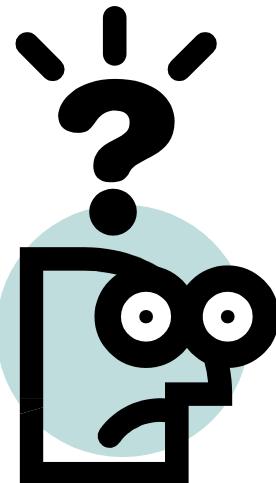


なぜ情報分析が必要なのか

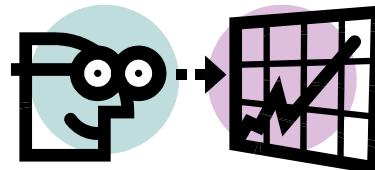
インターネットが一般的でなかった頃・・



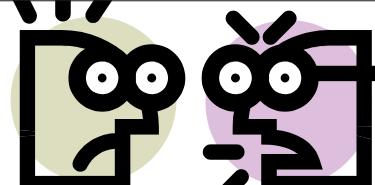
何かわからないことがあると…



信頼できそうな本を調べ



異なる分野の情報を調べ

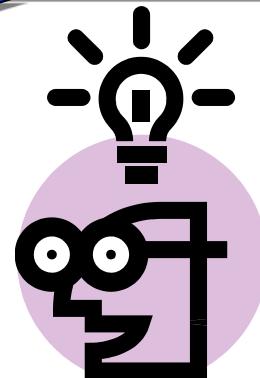


専門家の意見を聞き



他人の評判を聞き

発信元(源)の質に敏感となって
内容を総合的に判断した結果



調べた情報の質も理解した上で、
疑問に対する回答を自ら
導きだす。

現在は？

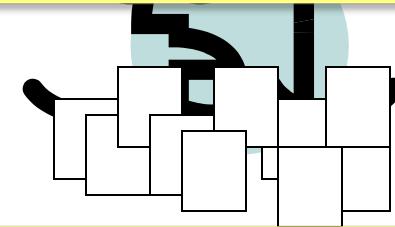
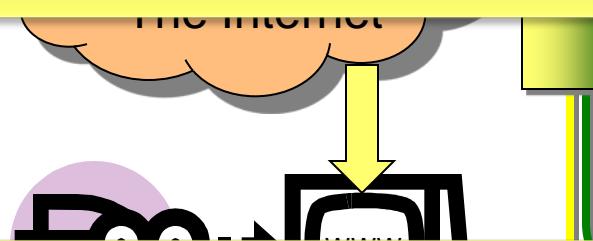


大量の情報のため情報の質
を判断する力も必要です。

既存の検索エンジンは、情報の質を考慮しているのか？



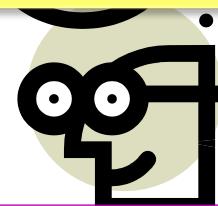
何かわからないことが
あります...



既存の検索エンジンは、ユーザの意図を理解してコンテンツを
検索しているのか？



思いついたキーワード
を検索エンジンに入力



既存の検索エンジンは、キーワードにヒットするコンテンツの存在を
明らかにするものにすぎなくなっている

したがって、

検索エンジンは、キーワードにヒットするコンテンツがあるかないかを調べるには有効である。

しかし、

- ★トップランクのコンテンツですら、いつも価値のある、また信頼できる情報を含んでいるとは限らない。
- ★ユーザの意図を理解してコンテンツを探しているとは限らない

既存の検索エンジンでインターネットから価値のある情報を取り出すためには、

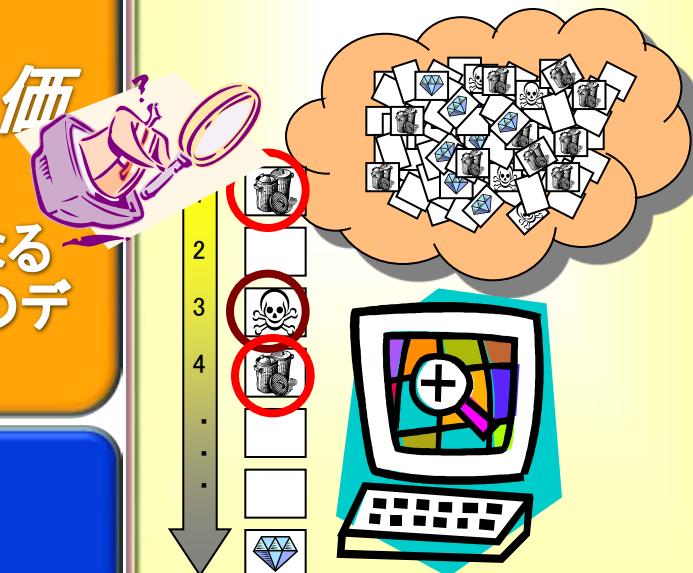
- ★ユーザは(検索意図を反映するために)何度も異なるキーワードで検索し、(質を判断するために)複数のデジタルコンテンツを閲覧して比較するしかない。

情報検索技術に加えて情報分析技術が必要となる。

既存の検索エンジン

現在の検索エンジンはメジャーなコンテンツを高く評価する。

サーチエンジンはキーワードで検索して、検索リストを出力する。(ユーザはトップランクのページに価値があると信じている。)

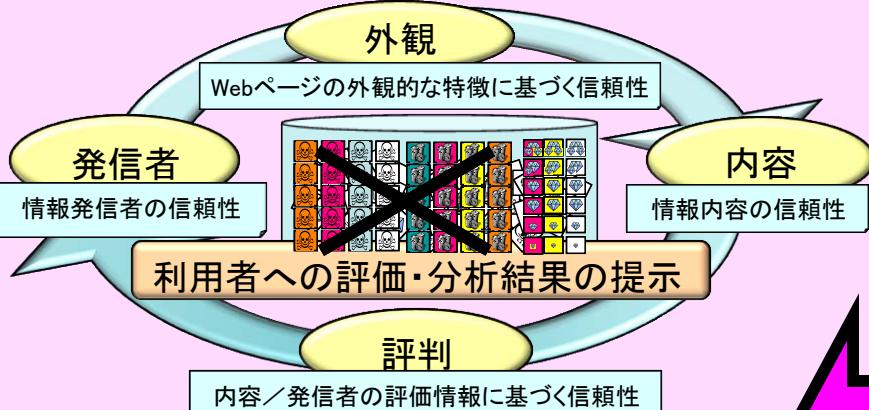


情報信憑性検証技術実現のための研究体制

情報通信研究機構(NICT)自ら実施する研究と、委託研究により、強力に推進。

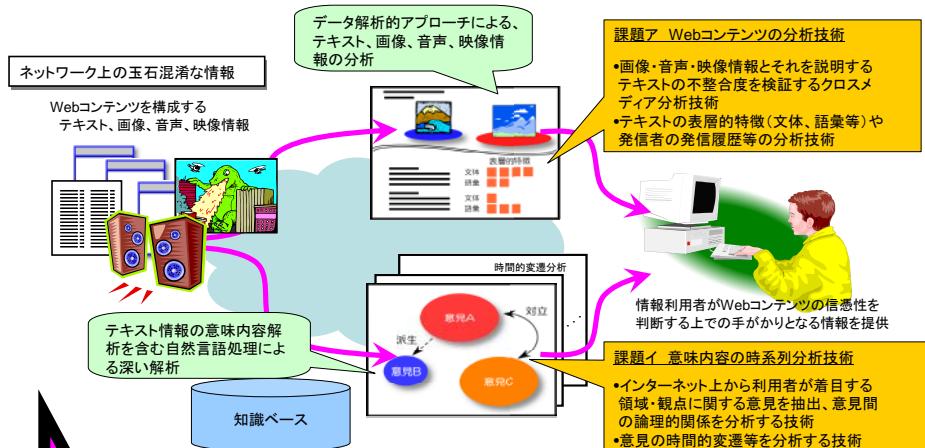
NICT研究

情報信頼性検証(情報の評価・分析・集約) 情報信頼性の評価基準



- 大量のWebコンテンツから自然言語処理技術に基づく4種類の情報分析手法で、ユーザが信頼性を判断する基準となる情報を提供

委託研究



- データ解析的アプローチによるWebコンテンツを構成するテキスト、画像、音声、映像情報の分析技術
- 自然言語処理による深い解析による情報相互の論理的関係性、時間的変遷の分析技術

相互補完

誰もが安心・安全にWebコンテンツを利活用できるための情報信憑性検証技術を実現

信頼できる情報を見つけるには？

NICT研究

The screenshot shows a Google search results page for 'アガリクス' in Microsoft Internet Explorer. The results include various links related to Agaricus products, such as '返金保証付きのアガリクス' and 'アガリクスをお探しなら'. There are also sponsored links from companies like '高品質低価格アガリクス' and 'ノン・クロレラアガリクス'. The bottom of the page features a navigation bar with links to 'Yahoo!検索', 'JWordサーチ', 'Google検索', '画像比較', '買い物比較', 'その他比較', 'おすすめ', and 'インターネット'.

検索語: アガリクス

検索結果数: 約 1,670,000 件中 1 - 10 件目 (0.21 秒)

ウェブ

返金保証付きのアガリクス
www.aicplus.com 飲みやすい返金保証付のアガリクス お試しセット好評。小冊子無料進呈

アガリクスをお探しなら
www.aqua-japan.com/ 全国のお客様に愛されるアガリクスのトップブランド。お喜びの声も掲載。

天仙液と代替医療について
www.new-ganchiryo.com/daitaiyou. 天仙液と代替医療について理想的なガン治療情報を提供しています。

アガリクス
アガリクス 協和の仙生露だけの「協和アガリクス販」濃厚エキスです。エキスタイル センセイロ エキス新ロイヤル 新発売！只今キャンペーン実施中！
www.keysoft.jp/abmk/-14k- キャッシュ - 関連ページ

アガリクス
アガリクス 販、確かな品質！
ソジニアリングの販を、世界基準の健
www.keysoft.jp

アガリクス
アガリクス
技術
ミネラル
www.suncreda.com

アガリクス
アガリクス
アガリクスアガリクス
www.agaricusya.com/ - 22k - キャッシュ

アガリクス 選りすぐり商品【ケンコーコム】
アガリクスの通販ならケンコーコム。厳選されたアガリクス商品を取り揃えております。アガリクスの通販(通信販売)。

既存の検索システム(グーグル等)
とは異なる
情報分析システムが必要！

- ◆妥当な内容が書かれてるの？
- ◆どんな人・組織が書いてるの？
- ◆連絡先や情報源は明示されている？
- ◆社会的にはどう見られている？

「アガリクス」の検索結果

健康によいという宣伝ばかり！
でも、本当に効くの？

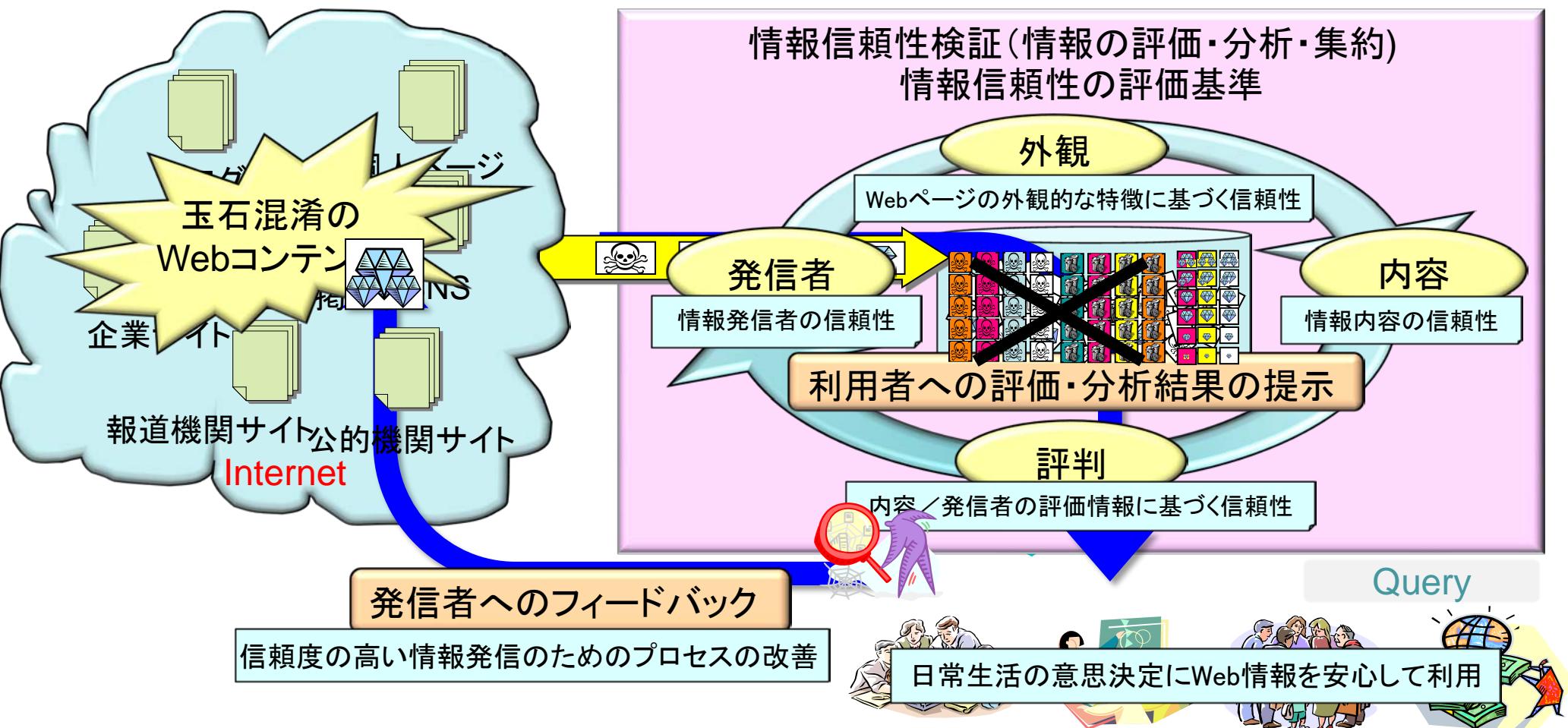
信用するためには
いろんな面から見たい！

自然言語処理技術に基づく4種類の情報分析手法で解析

NICT研究

- 情報通信研究機構(NICT)自ら研究開発するWISDOM*(情報信頼性検証システム)は、大量のWebコンテンツから自然言語処理技術に基づく4種類の情報分析手法で、ユーザが信頼性を判断する基準となる情報を提供する。

【*WISDOM:Web Information Sensibly and Discreetly Ordered and Marshaled】



• WISDOMの画面例

情報発信者の分析

分析方法
(ここでは発信者分類を選択)

入力キーワード
アガリクス

発信者分類

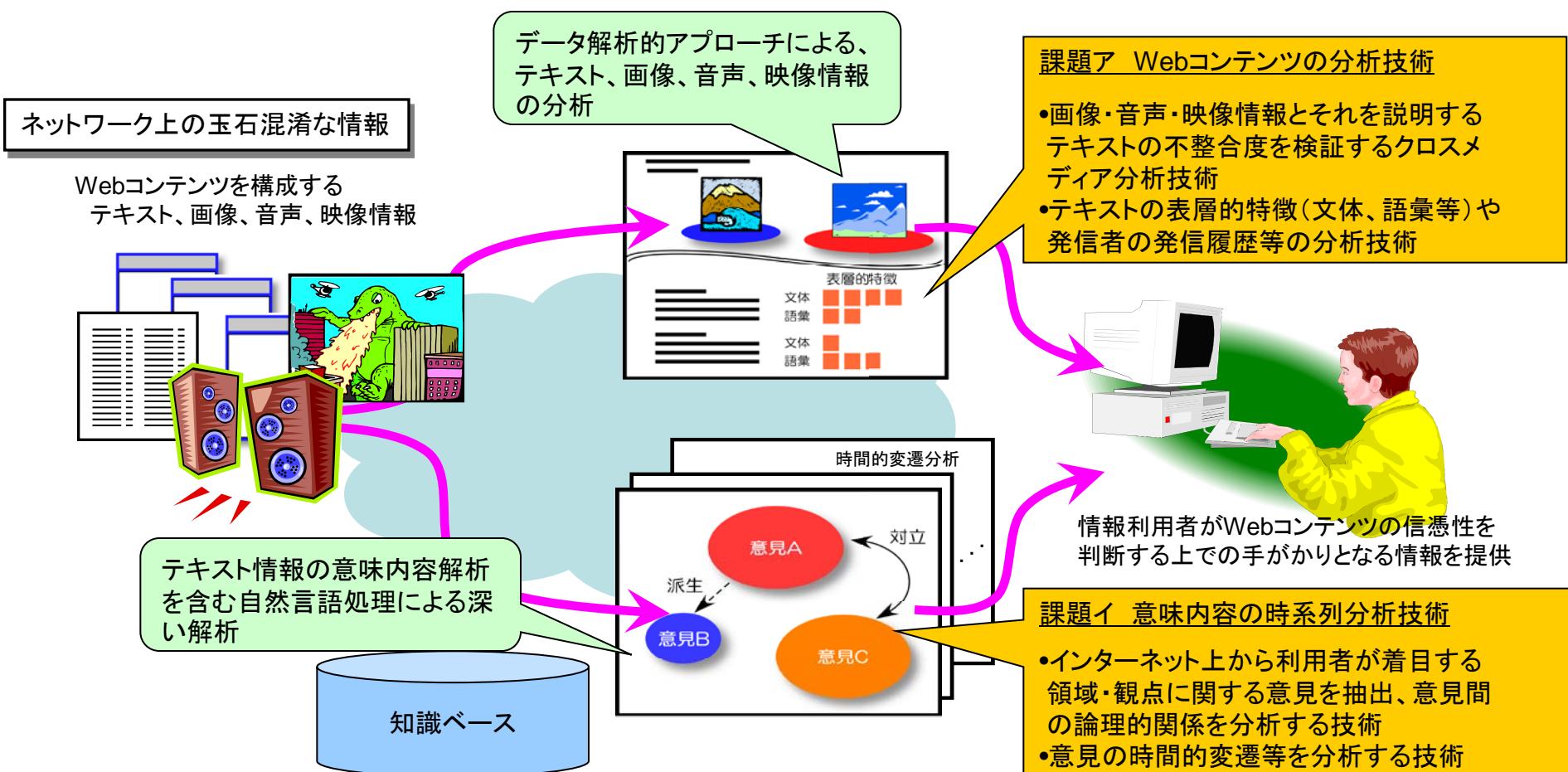
現在の分析対象ページ数(100)
個人(36)
有識者・専門家・著名人(11)
一般人(4)
匿名・ハンドルネーム(21)
団体(62)
官公署(45)
企業(45)
業界団体(0)
非営利団体(6)
行政(3)
公益法人等(2)
大学(0)
学会(0)
任意団体(1)
報道機関(11)
▶新聞社(11)
雑誌(0)
テレビ・ラジオ(0)
個人の集合(1)
参加者は実名(0)
参加者は匿名(1)
その他(0)
不明(1)

選択した分類項目のページリスト
アガリクス、プロポリスの抗がん効果と安全性 キャッシュ
「アガリクス」発がん促進 キリン子会社4品を回収 キャッシュ
アガリクスで発がん促進 キャッシュ
健康食品 2社分は安全 キャッシュ

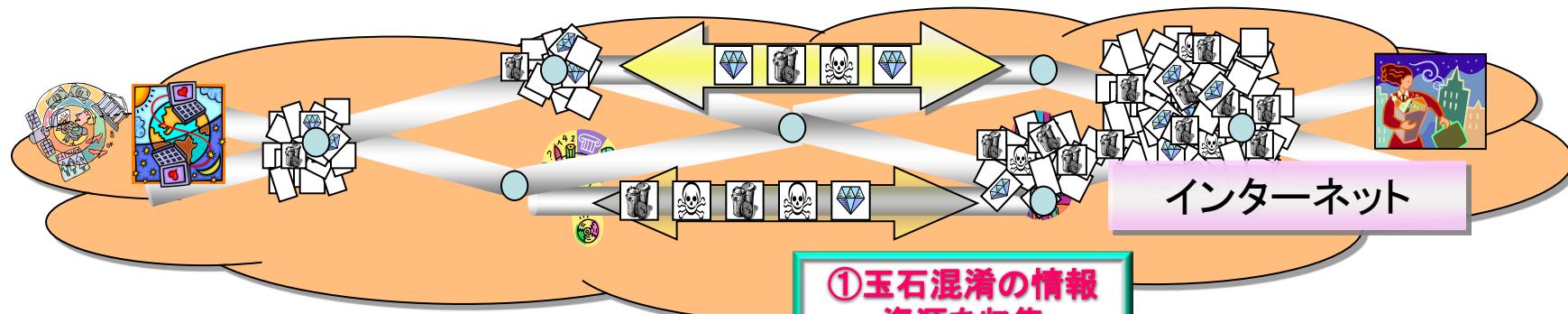
情報の外観と評判の分析

アガリクスで発がん促進 キャッシュ
Webページの要約
中島久美子@医療情報部@株式会社読売新聞
新規
情報外観
連絡先-実世界
サイト内検索
PrivacyPolicy
複数言語
レイアウト
サイト内ナビ
更新- 3日以内
広告-やや多い
このページに含まれる意見対象・属性
健康食品のマイナス情報(否定)
外観の特徴
意見対象
アガリクスの場合、製品を使ったがん患者らが肝臓障害で死亡した例や、発がん作用を示す動物実験結果も過去の学会で報告されているが、こうした健康食品のマイナス情報は利用者に伝わりにくい。有効性や安全性を客観的に検証する第三者機関や、その情報を消費者にわかりやすく伝える仕組みが求められる。
否定 肯定
ページ中の意見と極性

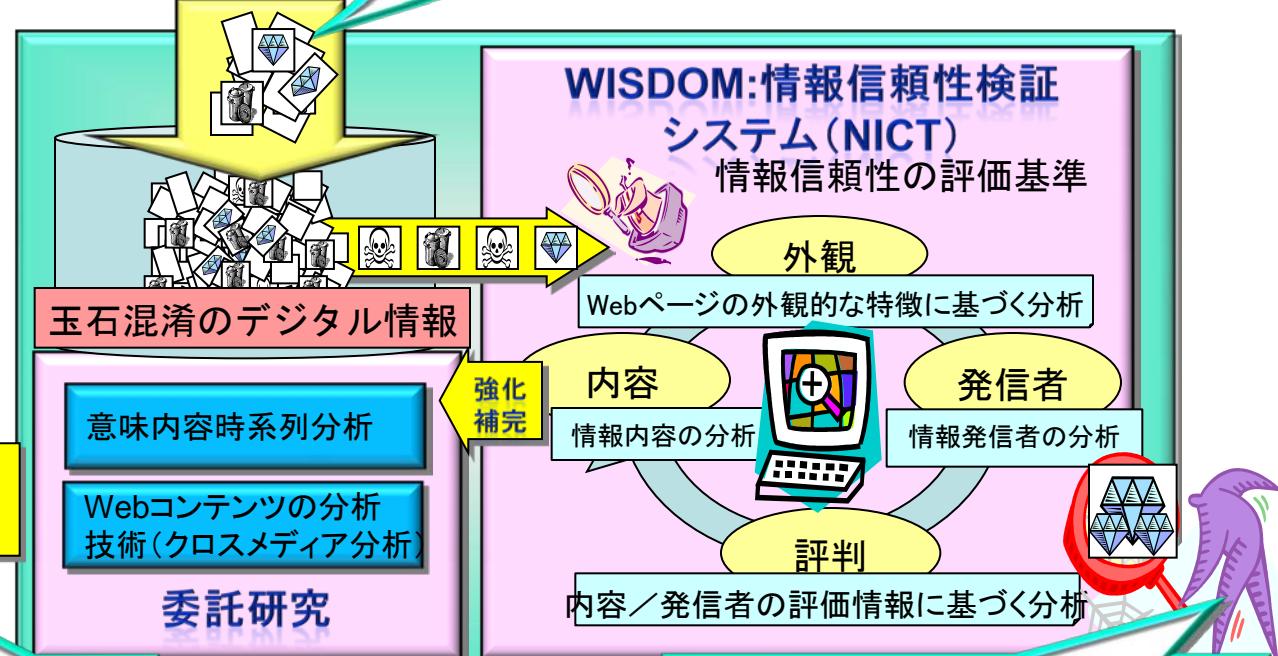
- データ解析的アプローチによるWebコンテンツを構成するテキスト、画像、音声、映像情報の分析技術
- 自然言語処理による深い解析による情報相互の論理的関係性、時間的変遷の分析技術



利用イメージ



【利用例】アガリクス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。アガリクスに関して多面的に検索をしてみよう。



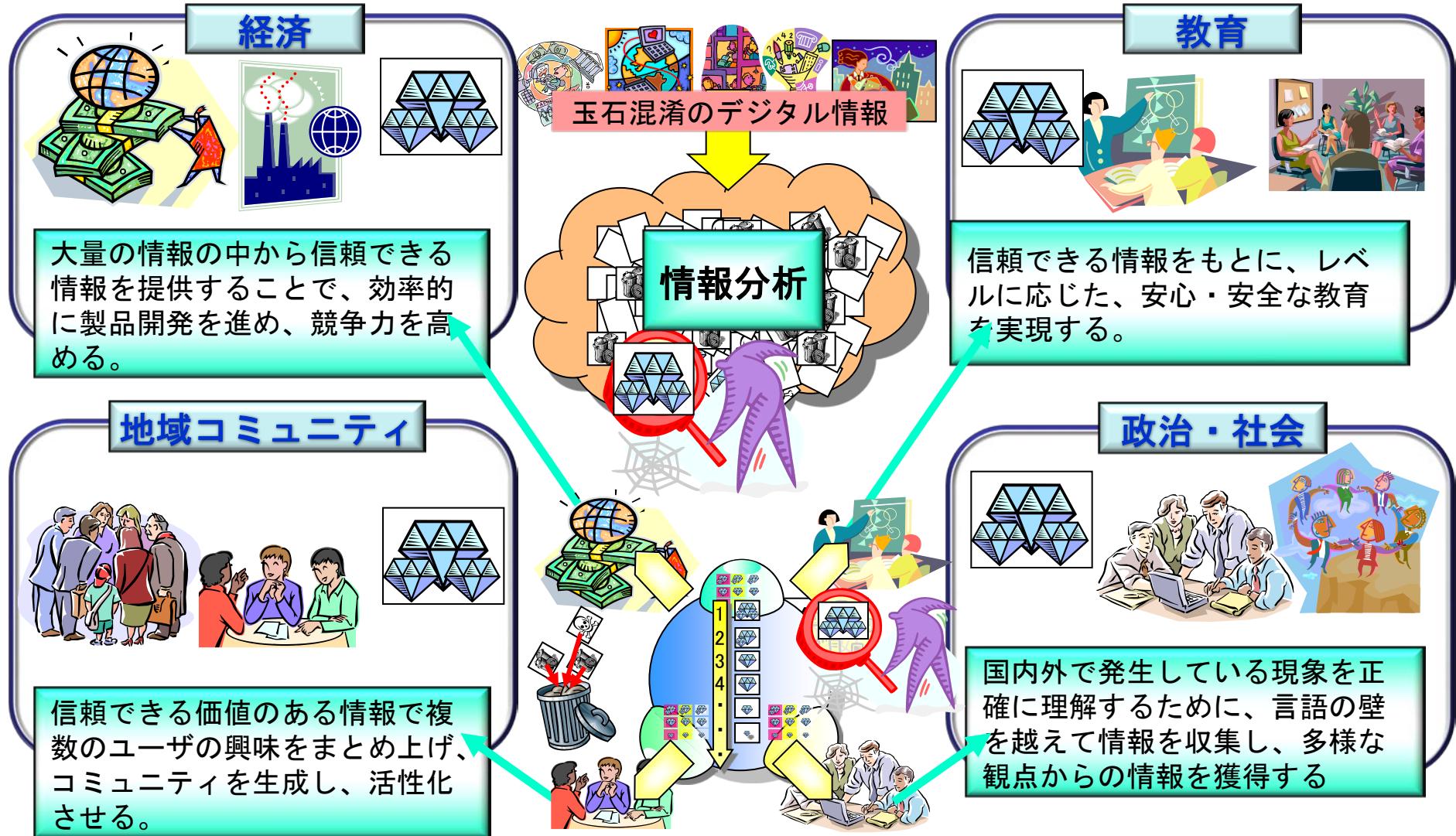
③分析結果を提示、ユーザにインターネット上の価値のある知識情報を発見させる。

②玉石混淆の情報資源を分析

10

アウトカム

信頼できる情報を正しく利活用できる情報通信基盤を構築し、ユビキタスネット社会への革新を加速させる。



- **現在のWeb**
 - インターネットによる情報の集合
- **次世代のWebとは？**
 - 信頼できる情報の利活用によるユビキタスネット上の「知の還流」を実現
- 以下の技術の研究を通して次世代Webへの変革を目指す。
 - 電気通信サービスにおける情報信憑性検証技術
 - 信頼できる情報を発見するための情報分析技術

ご静聴ありがとうございました

