

膨大な文章情報を意思決定に役立てる

膨大な文章から知識を引き出すシステム

農林水産業に関する情報は様々な形で蓄積されていきますが、最も一般的で量の多いものが各種報告書に代表される文書による情報です。文書による情報は出版された当時には流通・活用がなされませんが、多くの場合、そのまま「お蔵入り」になってしまいます。

そこで、本研究では、種々の文書を蓄積し、利用者の要求に応じて様々な観点で利用できるようにするしくみを開発しております。

■ 文書を分類・整理するしくみ

膨大な量の文書に埋もれている数種類の情報に着目し、それらをピンポイントで取り出して分類・整理するしくみを実現しました。例えば、産直サイトの分類・整理、漁海況の特異な現象の分類・整理などができます。

文書の分類・整理

重要な情報を抽出 →

抽出結果を直感的に分かり易く分類整理

質問応答

質問を入力すると答とその根拠が得られる

質問事項だけでも答えられるものもある

■ 同じような内容の文書を発見するしくみ

利用者がキーワードを与えるとそれらを含む文書を表示するのが通常の文書検索ですが、このしくみでは、それだけではなく、概念検索を実現しました。具体的には、質問文に内容として近い文書を見つけることができます。

■ 質問すると答えを探してくれるしくみ

質問文にズバリと答えしてくれる質問応答のしくみを実現しました。質問文に関連する文書を見つけるのではなく、答そのものを見つけることができます。

■ 貴重な文書情報の電子化

新しい情報を電子化文書として収集するとともに、紙面として保存されている古い情報を電子化することにより、各種システムで利用可能な知識源を拡充しております。

概念検索

「ホンシメジの人工栽培」で検索

良い検索結果があったので、その本文を使って概念検索

ホンシメジの情報はもろこの→他のきのこの類の情報も上位に検索される

システムを次のURLで公開中
<http://pc110.narc.affrc.go.jp/AgInfo/textmining.html>

お問い合わせ先

国立大学法人 横浜国立大学 大学院 環境情報研究院
 森辰則 (e-mail: mori@forest.eis.ynu.ac.jp URL: <http://pc110.narc.affrc.go.jp/AgInfo/textmining.html>)