

2005.01.26

情報処理軽井沢セミナー報告書

上智大学 電子計算機センター

峰内 暁世

1. 課題

語学学習に役立てるために OAI-PMH でニュースサイトからハーベストしたメタデータを Namazu で検索できるようにする。

2. 1 の概要

ニュースサイトから OAI-PMH でメタデータをハーベストし、データを記事ごとに分ける。分けたデータを Namazu で検索できるようにする。

3. 演習とその成果, 何を計画し, 何を実装して, 何ができたか

(1) 演習第 1 日

OAI-PMH でメタデータをハーベストできるニュースサイトの検索

(2) 演習第 2 日

スクリプトにより、ハーベストしたメタデータを記事ごとに分割し、1 行の XML ファイルとした。

(3) 演習第 3 日

Namazu で検索できるようにインデックスを作成し、検索可能にした。

4. 研修で学んだ技術および知識

OAI-PMH によるデータのハーベスティング。Namazu によるデータ検索システムの構築。

<http://www10.plala.or.jp/miyazawa/namazu/>

Namazu による検索システム構築について

<http://www.ki.nu/software/namazu/tutorial/purpose.html>

Namazu による検索システム構築について

5. 事前準備として必要と思われるもの

XML の処理に関する基礎的な知識

OS のセットアップに必要な情報

6. 今後の課題 (職場でさらに調査する必要があるもの等)

7. 今後の計画 (実際のポータル構築計画等)

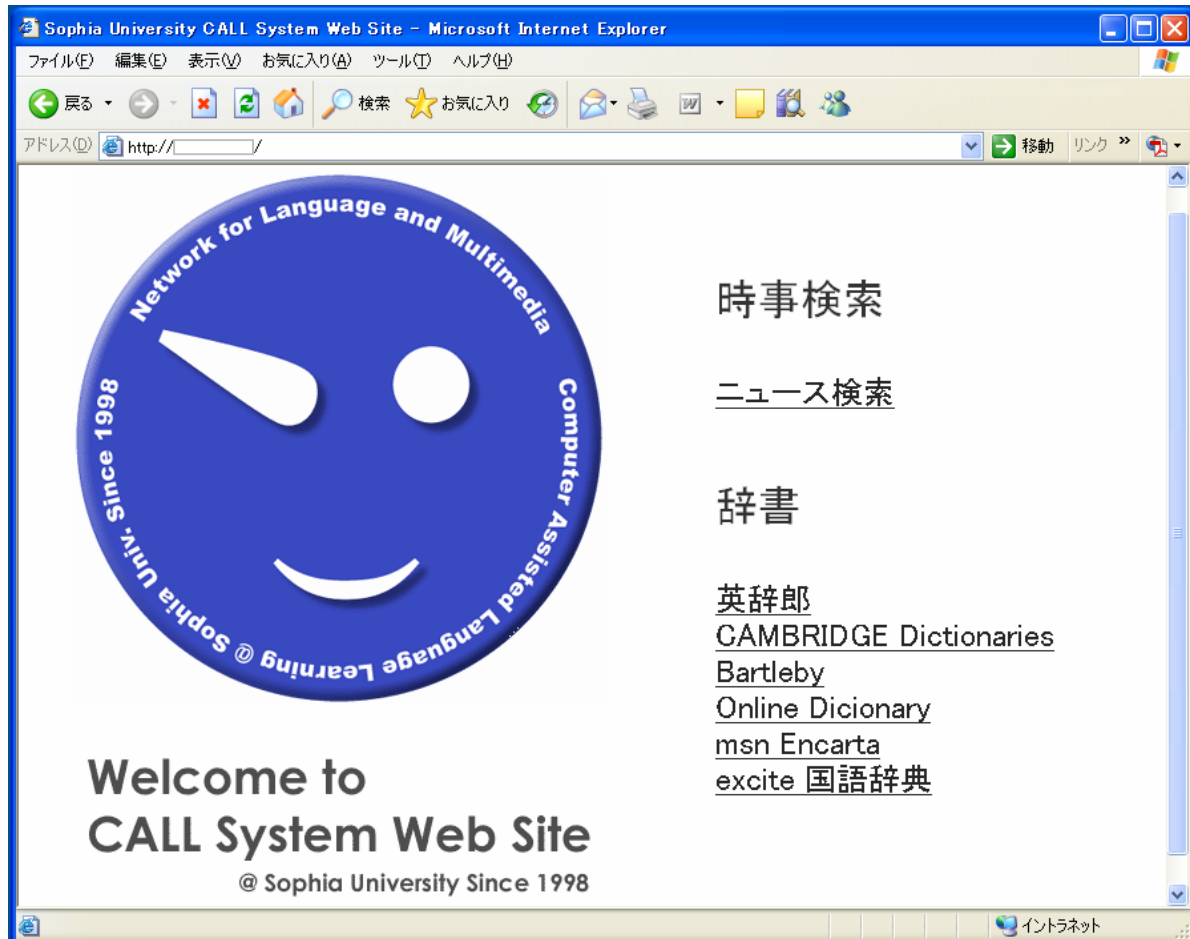
CALL システムの使い方および FAQ の Namazu, GETA による検索

8. 演習の感想

私は、図書館勤務の経験が無かったので、申し込み時に想像していたセミナー内容とだいぶ異なる

り、事前レポートや軽井沢での第1日目、2日目にはどのように自分でセミナーの課題をこなすべきかの迷いがあった。課題別にチーム分けをした後からはなんとなく自分なりにこなすべき課題が見つかった。自分にとっては、図書館では目録を作るのは当然の業務としてあり、それをどのようにサポートしてゆくかも含め新鮮な発見の連続であり、有意義な時間が過ごせた。

<< スクリーンショット >>





Namazu: a Full-Text Search Engine: <Bush> - Microsoft Internet Explorer

ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)

戻る 検索 お気に入り

アドレス http://NAMAZU/NAMAZU/namazu.cgi?query=Bush&max=20&result=normal&sort=score 移動 リンク

Namazu: a Full-Text Search Engine

This index contains 78 documents and 916 keywords.

Last modified: 2004-12-20

Query: [[How to search](#)]

Display: Description: Sort:

Results:

References: [Bush: 2]

Total 2 documents matching your query.

- [Bush warns Iran and Syria on Iraq](#) (score: 18)
Author: *unknown*
Date: *Thu, 16 Dec 2004 17:13:30*
<?xml version="1.0" encoding="UTF-8" ?> US President George W **Bush** says Iran and Syria should not interfere in Iraq in advance of January's elections. <http://news.bbc.co.uk/go/click/rss/0.91/public/-http://133.12.31.38/output/output1.xml> (320 bytes)
- [Bush's economic agenda](#) (score: 17)
Author: *unknown*
Date: *Thu, 16 Dec 2004 17:13:36*

ページが表示されました

イントラネット