

知識情報のためのXMLによる電子文書データベース統合管理システムの構築

指導教員 皆川 勝
学生氏名 田村 郷司

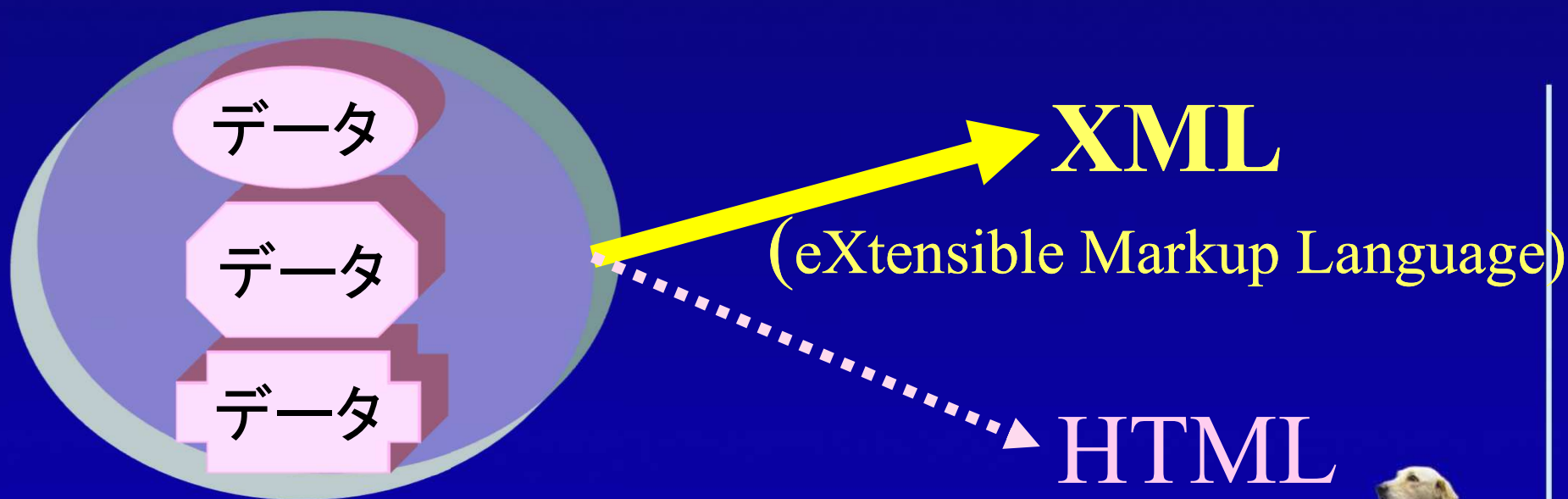


研究背景

情報を配布するための様々な手段が考案されてきている。そして、人類が到達したのは、デジタル化した形の情報の保存と配布である。



- ・ 更に、インターネットの急速な普及により Webで管理するデータの種類が増えているためHTMLでは情報を記述する手段としての限界に達してきている。



Webの世界



- ・建設業界においては、国土交通省(旧建設省)が成果品の電子納品の標準化に向けてXMLを採用した。

◆デジタル写真管理情報基準(案)

◆土木設計業務等の電子納品要領(案)

◆工事完成図書の電子納品要領(案)

◆CAD製図基準(案)

◆地質調査資料整理要領(案)



研究目的

- ・ 国土交通省の定めた**工事完成図書**の**電子納品要領(案)**（以下、**電子納品要領案**）により電子文書を管理，運用する効果を検討する。
- ・ 本研究で対象とする文書は、建設情報研究室の修士学位論文及び卒業論文の文書である。近年知識管理の重要性が叫ばれている中で，対象としては研究論文と並んで重要と考えた。



XMLとeXcelon

- ・ XMLはWorld Wide Web Consortium (W3C)の中に組織され1998年2月、「XML 1.0」が勧告として承認されている。
- ・ インターネット時代の汎用なデータ記述言語でインターネットのWebで情報を保存し、インターネットを介して情報を配布できるように設計されたマークアップ言語である。

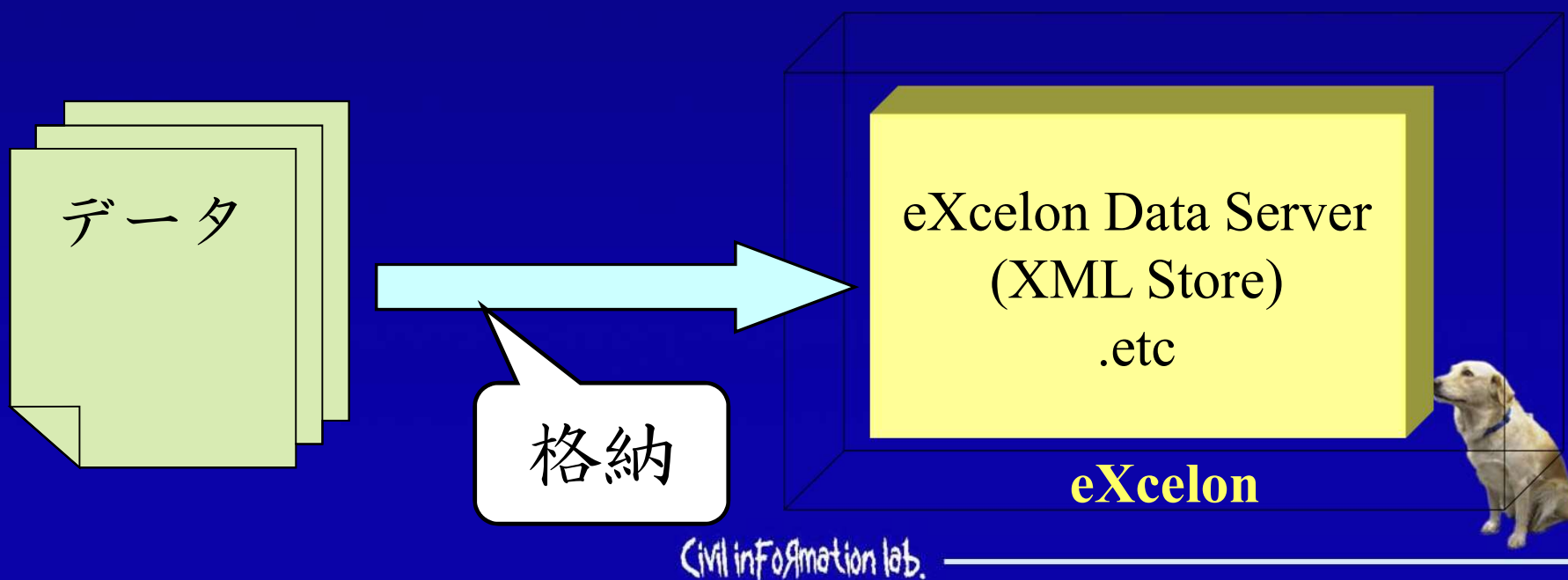


- ・ eXcelonとは、eXcelon社が開発したXMLアプリケーション開発環境である。
- ・ Data Server をベースに、XMLアプリケーション開発に必要な機能を統合している。
- ・ eXcelon Data Serverにより、XML情報の格納、管理、クエリー実行、及び配信が可能になる。

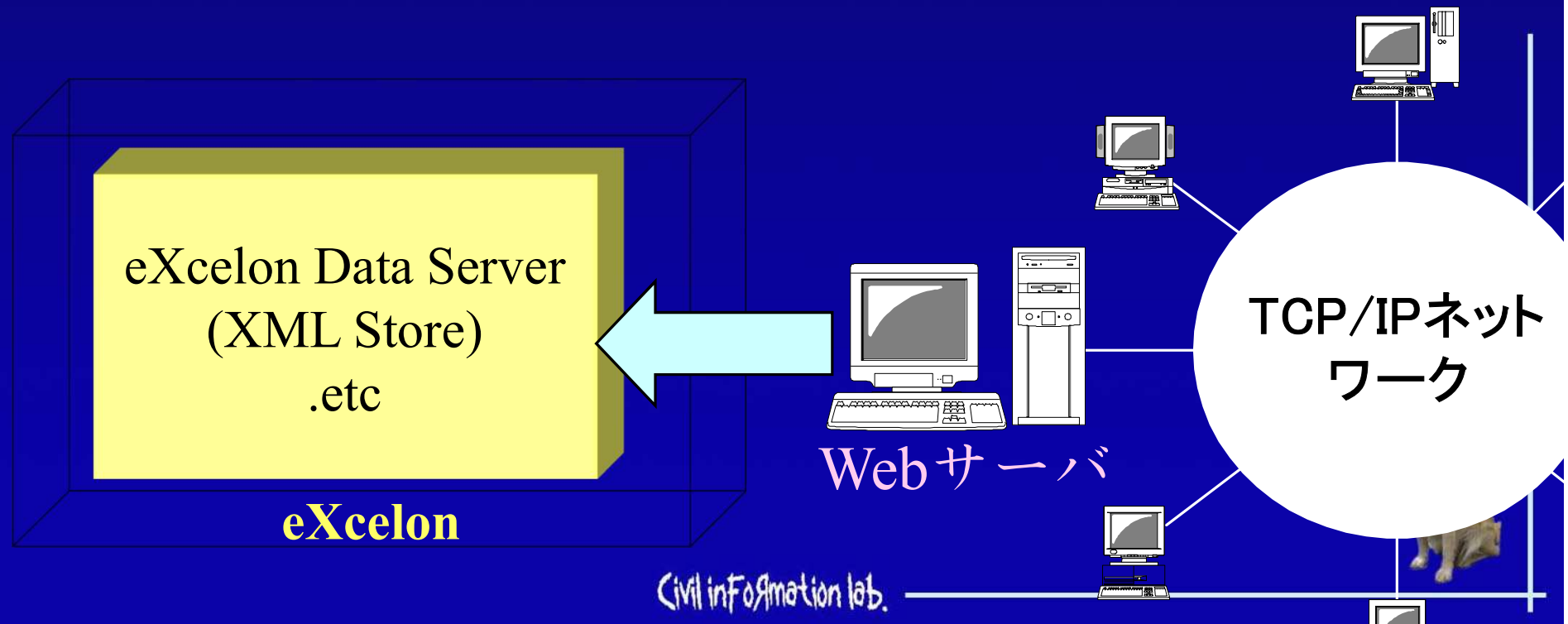


構築したシステム

- データをeXcelon Data ServerのファイルシステムルートであるXML Storeに格納した。



- ・ eXcelonのWebクライアントからXML Storeへのアクセスを提供するサーバエクステンションの機能を使用し， Webサーバを介してTCP/IPネットワークで閲覧， 管理などをできるようにした。



格納したデータ

- XML

XML文書はronbun.xmlと1991卒論概要.xml～2000卒論概要.xmlを作成した。

- XSL(eXtensible Markup Language)

XML文書は体裁情報がないのでスタイルシートを作成する必要がある。作成するXSLは全体表示させるronbun.xslと、卒業年・専攻別(情報, 構造解析, 構造実験)で検索できるkensaku.xslの2つのスタイルシートを作成した。

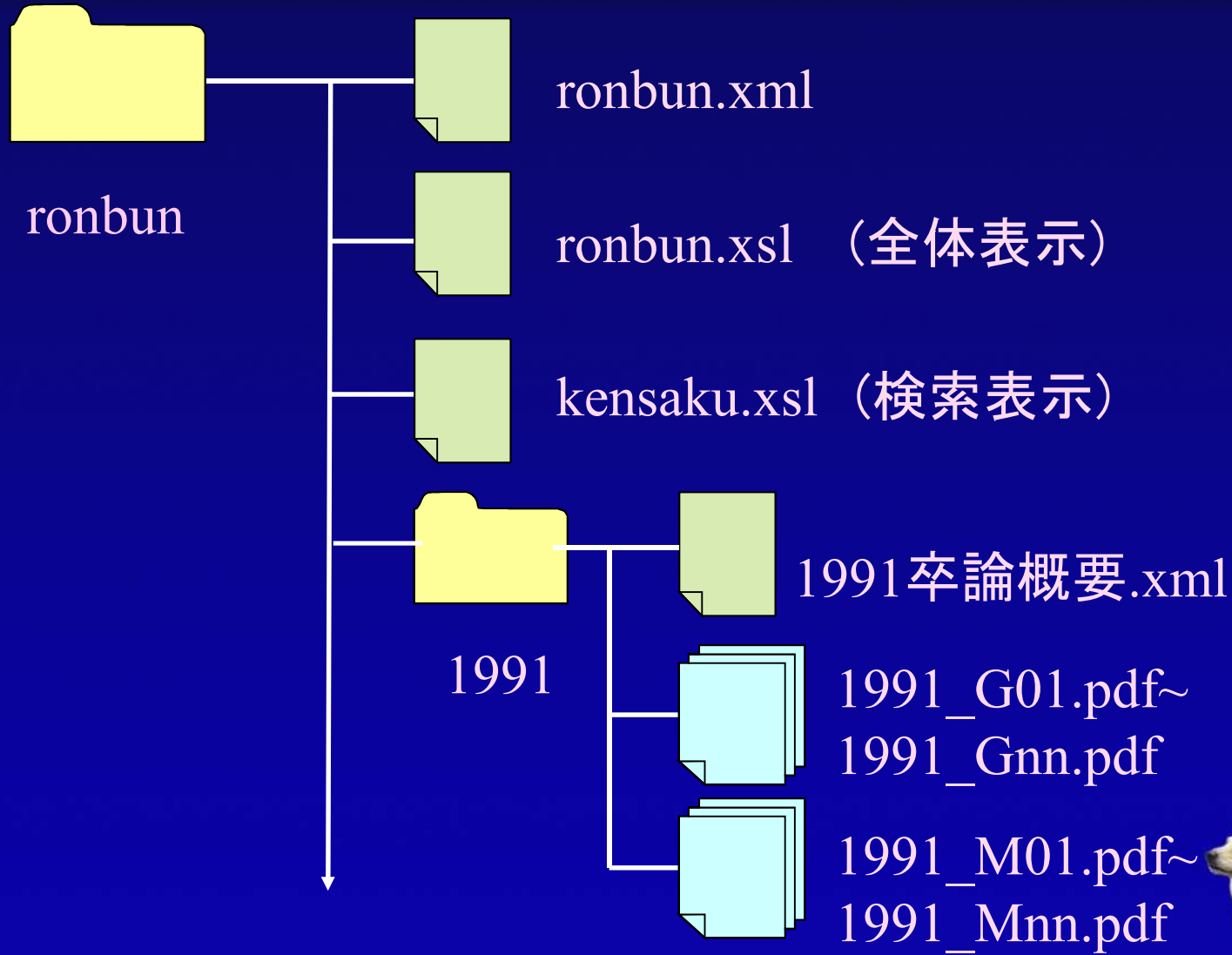


- ・ PDF

Microsoft Wordや紙データで存在する論文概要はアドビシステムズ社のAdobe Acrobatを使用しPDFで格納することとした。



フォルダ構成



生じた問題と対策結果

- ・ 問題点 1

kensaku.xslにronbun.xmlを適用し検索結果を表示させようとしたが、エラーが出て表示されないという問題が生じた。

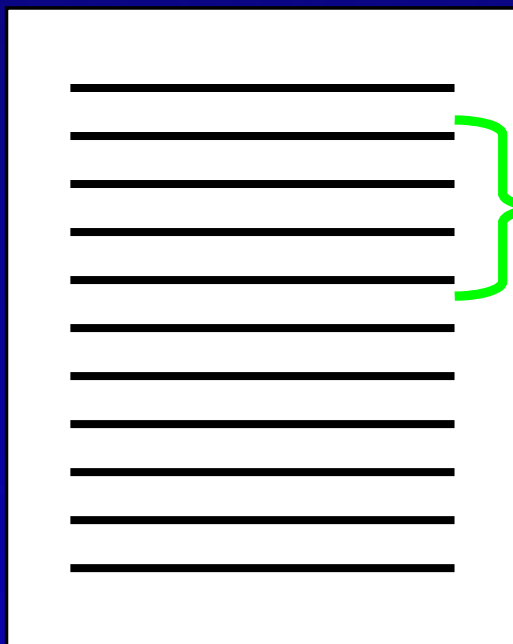


選択

kensaku.xslにronbun.xmlを
適用した画面

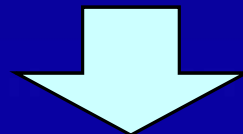


- ・ 1991卒論概要.xml～2000卒論概要.xmlのXML文書が全角であるのに問題があるのではないかと考え、検索をする部分のXML文書のタグ要素及びタグの内容をすべて半角英数に書き換えた。その結果正常に検索され表示された。



検索する部分

<専攻>情報 </専攻>



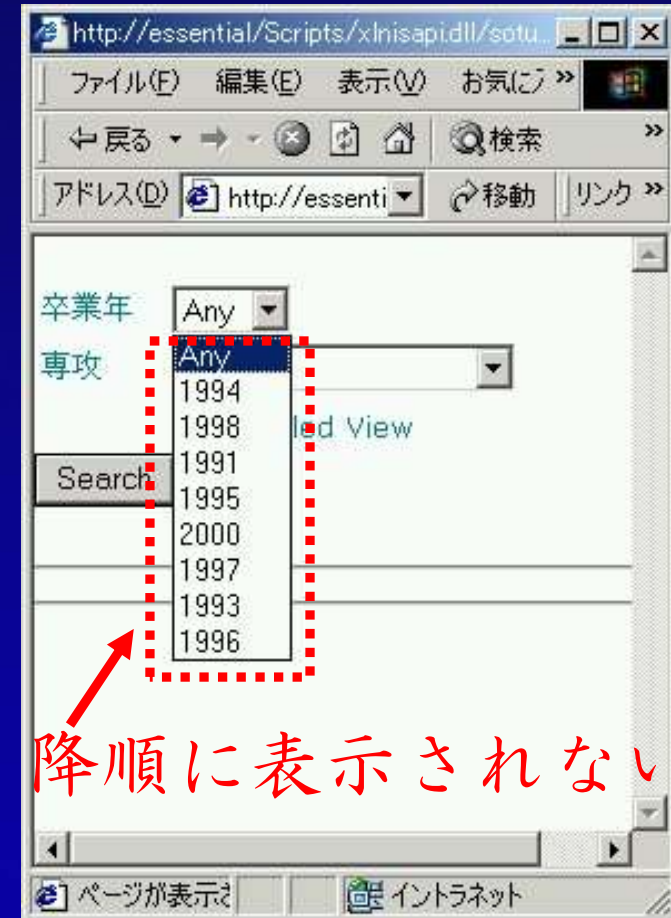
<senkou>Information </senkou>



XML文書

- ・ 問題点 2

kensaku.xslを
ronbun.xmlに適用
した画面で年代別の
プルダウンメニュー
が降順に表示されな
いという問題が生じ
た。



- ・ XSL中にXML文書中のデータを並べ替える
ソート(sort)を使用した
- ・ ソートにはW3C仕様とMicrosoft社独自仕様
の2つがある。W3C仕様はIE5.0x標準インス
トール環境下では機能しないためMicrosoft社
独自仕様を使用した
- ・ しかし改善されなかった。

ronbun.xslより抜粋

```
<xsl:for-each select="$qry"  
order-by="-number(graduate-y)">
```

降順の並べ替えを指定

指定されたタグの内容を繰り返す要素

Civil Information Lab.

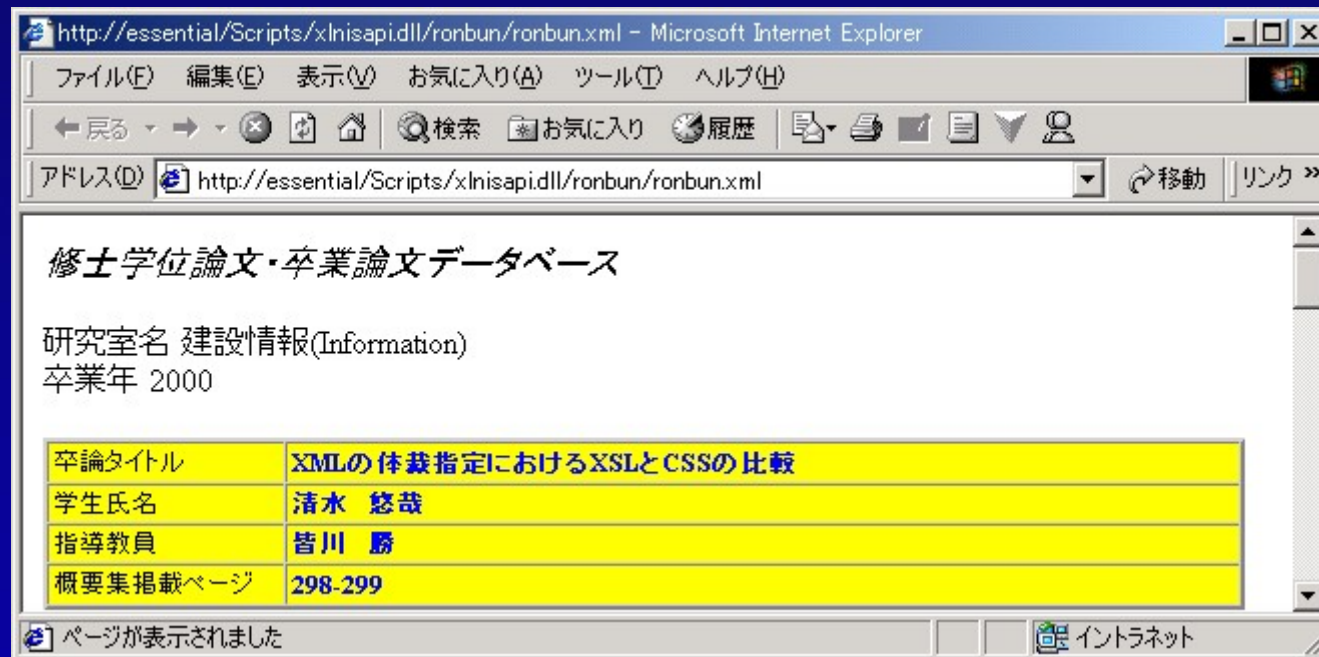


考察及び結論

- ・ (問題点1)により検索をするXML文書を部分的に半角英数に書き換えたため、ローマ字の部分と日本語の部分に分かれてしまった。
eXcelonがXML Storeに文字コードをUTF-8で格納することに問題があると考えられる。
- ・ (問題点2)に関しては、Microsoft社独自仕様なのでeXcelonと関係して正常に表示されないと考えられる。



- ・ 以上の問題点があったものの、eXcelon
を利用し検索結果を表示することが
できた。



卒業年=2000, 専攻=Information, の
検索結果の表示画面



- ・ 修士学位論文・卒業論文データベースの構築と運用ができたことは本研究の最大の成果である。
- ・ この成果により重要な文書を劣化や汚損から守り，情報の再利用性を向上させ，データに永続性，完全性を持たせることができ，当初の目的を達成できたと考える。



おわり

