# Trends and Challenges of Databases

M.Princz

University of Debrecen, Faculty of Engineering, Department of Basic Technical Studies, pmaria@unideb.hu

*Abstract. The database management, using relational databases, is part of curriculum in the Hungarian high schools. The aim of this paper is to present how we can show for students the challenges facing data processing, data retrieval, beyond the relational database management taught in high school.*

*Keywords: teaching databases, introduce XML, challenges of relational databases*

## Introduction

The University of Debrecen, Faculty of Engineering has been teaching high school students for years. The purpose of these lessons is to provide insights into the faculty's training through the subjects taught by the professors of the faculty, and to raise students' interest.

In the framework of EFOP-3.6.1-16-2016-00022, new topics were elaborated by professors of faculty, including Trends and Challenges of Databases.

The aim of this paper is to present how we can show for students the challenges facing data processing, data retrieval, beyond the relational database management taught in high school.

The informatics curriculum, what Hungarian students have to learn during high school, are

- The basics of Information Technology,
- Using Operating System
- Communication on the network
- Algorithms and Data
- Document creation with a computer
- Spreadsheets
- Database Management
- Library use

The spreadsheets management and database management are part of current curriculum of the Informatics for Engineers course at the University of Debrecen, Faculty of Engineering, but a decade ago, we taught text editing, too.

I have been teaching this subject for 20 years in this faculty, and my experience is that our students show the greatest shortcomings in the area of databases.

Therefore, I wanted to draw attention to databases and wanted to give more information to the high school students about this area in addition to the prescribed curriculum.

Relational Database Management Systems have been meeting the needs of the market for decades, but since the beginning of the 21st century, new demands have emerged that serve beyond the boundaries of the relational systems.

There are issues that I wanted to discuss with high school students.

# 1. Multimedia data

Multimedia data is a combination of a number of media objects (i.e., text, graphics, sound, animation, video, etc.) The efficient processing of today's huge data sets is a serious challenge.

Multimedia databases means larger data volume than previous data formats so the storage manager should be able to handle gigabytes of rows.

Information retrieval from multimedia data is not easy.

**Exercise: Search which Hungarian painters were born in Debrecen.**

In this case, we open a search engine (e.g. Google, Yahoo, Bing, etc.) and enter some keywords such as Debrecen, born, painter.

However, if our search engines does not know synonyms, we will not receive any documents such as:

XY's hometown is Debrecen, or XY's place of birth is Debrecen, but we will also receive documents such as

Margit Gréczi - Wikipedia
https://en.wikipedia.org/wiki/Margit_Gréczi ▾ Oldal lefordítása
Margit Gréczi (**born** 13 December 1941 in Gyöngyös) is a Hungarian **painter**. ... Born, 13 December
1941 ... **Debrecen** DOTE Mini Gallery, Hungary; 2009.

She was born in Gyöngyös and she had an exhibition in Debrecen. This is not, what we wanted.

If we could use a database query, it would be something like that:

place_of_birth="Debrecen" and profession="painter".

# 2. Limits of the relational data model

Until traditional databases contain structured data, multimedia databases are unstructured or semi-structured.

It is advisable to show what does it means. Here is a website, which contains different elements such as text, image, table, graphic, hyperlinks, etc.:

How do we store this in a relational database?

First of all let we see what does this page contain?



1.TEXT

2.IMAGE

3.TABLE

4.FORMATTING

5. GRAPHIC

6. HYPERLINK

We should know the internal structure of the web pages to store it in a relational database but the content of the web pages are changing and are not uniform. It would be hard to store this changing data in a table of a relational database.

What should we do then?

There are two possible approaches:

1. **Data mining, artificial intelligence**: By intelligent techniques, we try to recognize the information stored on the HTML page from the environment, formatting, and hyperlinks.

2. **Database Approach:** Instead of using relational database, the data itself describes their content and formatting. We use XML to describe the content.
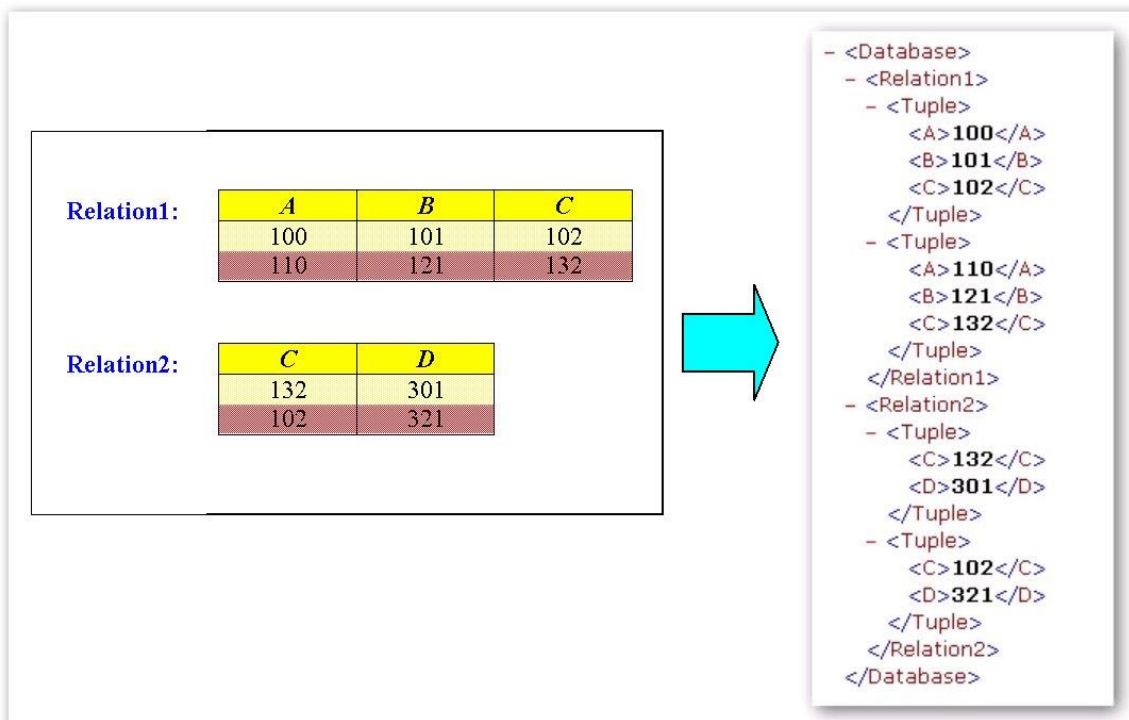
# 3. XML and relational databases

A descriptive language is required for automated machine processing to make the logical structure of the document. This is the XML (eXtensible Markup Language).

XML was designed to store and transport data.

XML was designed to be both human- and machine-readable

Any relational database can be represented as an XML document:



Any web page can be represented as an XML document:

```xml
<?xml version="1.0" encoding="UTF-8"?>
<breakfast_menu>
<food>
    <name>Belgian Waffles</name>
    <price>$5.95</price>
    <description>
   Two of our famous Belgian Waffles with plenty of real maple syrup
    </description>
    <calories>650</calories>
</food>
<food>
    <name>Strawberry Belgian Waffles</name>
    <price>$7.95</price>
    <description>
    Light Belgian waffles covered with strawberries and whipped cream
    </description>
    <calories>900</calories>
</food>
</breakfast_menu>
```

XML uses metatags to add structural and semantic information to documents to describe the content. XML allows a free choice of markup vocabulary. Thus, the web becomes a semi-structured XML database from which information can be extracted easily with XML query languages (Xpath, XQuery).

## Summary

The aim of this paper is to present how we can show for students the challenges facing data processing, data retrieval from multimedia data. With this problem, students can encounter their everyday web searches. A description of a possible solution goes beyond the school curriculum but can attract students' interest in databases and data retrieval.

## References

[1]     Oktatáskutató és Fejlesztő Intézet, Kerettantervek
        http://kerettanterv.ofi.hu/03_melleklet_9-12/index_4_gimn.html

[2]     Attila Kiss, Félig-strukturált adatmodell,
        http://people.inf.elte.hu/sila/eduXML/XML_Kiss_habitEA.pdf