707.009
Foundations of Knowledge Management
„Overview and Motivation“

Markus Strohmaier

Univ. Ass. / Assistant Professor
Knowledge Management Institute
Graz University of Technology, Austria

e-mail: markus.strohmaier@tugraz.at
web: http://www.kmi.tugraz.at/staff/markus
About me

Education:
• 2002 - 2004 PhD. Institute of Knowledge Management, TU Graz
• 1997 - 2002 M.Sc., Telematik, TU Graz

Background:
• July 2007 - present: Ass. Prof. (Univ.Ass.), TU Graz, Austria
• 2006 - 2007 Post-Doc, University of Toronto, Canada
• 2002 - 2006 Researcher, Know-Center, Austria
About me

Research Interests:
• Studying goals in the context of Social Media Analysis, Web Search and Knowledge Management

Interesting topics for projects, Bachelor/Master/PhD thesis:
• If you are interested in the topics of this course, it is likely that you are interested in doing a project / a thesis with me as well. Contact me to discuss opportunities.
Overview

Agenda

• Welcome & Introduction
• Course organization

• KM Motivation
• Course overview
• Course highlights
Teaching @ KMI

**Introduction to KM**
KM Orientation and Overview
Klaus Tochtermann
2. Semester, Softwareentwicklung-Wirtschaft
707.012, 2 VO

**Web-Science & WT**
Network Theory, Social Networks, Web Technologies
Markus Strohmaier
6. Semester, Softwareentwicklung-Wirtschaft
707.000, 2 VU

**Foundations of KM**
Theories of KM, KM Processes
Markus Strohmaier
7. Semester, Softwareentwicklung-Wirtschaft
707.009, 2 VO

**Knowledge Technologies**
Ontologies and Semantic Technologies
Michael Granitzer
7./9. Semester, Softwareentwicklung-Wirtschaft
707.011, 2 VU

**Applications of KM**
Selected Topics, Real-World Applications
Markus Strohmaier, Michael Granitzer
8. Semester, Softwareentwicklung-Wirtschaft
707.010, 2 VO

+ Multimedia Information Systems

---

Project Information Systems, Bachelor Thesis „Informatik“, Bachelor Thesis „Telematik“, Seminar/Project Knowledge Management, Master Project, Master Thesis

---

Markus Strohmaier
Course Context

• 707.009 Foundations of Knowledge Management
  – 3rd year as a „VL“
  – Has been held as a „VU“ before

• Part of studies „Software Engineering & Business“
  – Master studies, 7th semester
Goals I

At the end of the course, you should **know about and understand** different

- kinds of knowledge
- types of knowledge acquisition
- Knowledge acquisition methods
- types of knowledge organization
- types of knowledge bases

And you should **be able**

to distinguish between them, preliminarily assess their relevance for given contexts and apply them selectively
Non-Goals

In the research community, there is no broad consensus regarding the theoretical foundations of a „Scientific Discipline of Knowledge Management“

So the topics of this course are subjectively selective.

Instead of giving an authoritative account of KM, this course aims to give an overview of prominent, interesting and fundamental research results generated by the field of KM so far.
Course Organization and Logistics

- **Lectures**

November 2009 - January 2010,
Room HS Modul (Inffeldgasse 21a, Ground Floor)

Exceptions! (and updates)

**Final Exam**
Course Organization and Logistics

- **Language**: English/German


Enroll!

You need to enroll for this course via TUG online!

- **Slides** will be made available on the course website
- **Readings**
  
  Password to access protected documents on the course website:

Markus Strohmaier 2009
Grading

So how do you receive a grade in this course?

- There will be one written exam at the end of each semester (Standard). No limit on the number of students that can take the exam
- Two dates for oral exams (limited seats - First come, first serve)
  ➔ See policy on course website for more details

What will be part of the exam?

- **ATTENTION**: Adapted course contents from this semester on (Jan 2010)!
- Contents of slides **AND** weekly readings

How does class attendance affect your grade?

Indirectly:

- Type of Exam Questions
- In class, we will discuss issues from the weekly readings
- The papers discussed in the weekly readings will be part of the final exam
- Participating will likely increase your understanding of the subject
- Extra points to be earned during the course (for this semester's written exam)
Course Information

Please find **further information** about the course at

- Newsgroup available at tu-graz.lv.gwm
Questions?

Raise them **NOW**!

Or ask them later:

- At the end of each class
- Via e-mail: markus.strohmaier @ tugraz.at

(now would be a good time though!)
Let‘s start!
- Knowledge Management –

(it‘s all about questions today)
How much information is being produced?

In 2003

- The World Wide Web contains about 170 terabytes of information on its surface; in volume this is seventeen times the size of the Library of Congress print collections.
- Instant messaging generates five billion messages a day (750GB), or 274 Terabytes a year.
- Email generates about 400,000 terabytes of new information each year worldwide.
- P2P file exchange on the Internet is growing rapidly. Seven percent of users provide files for sharing, while 93% of P2P users only download files. The largest files exchanged are video files larger than 100 MB, but the most frequently exchanged files contain music (MP3 files).
How much information is being produced?

http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/

Table 1.2: Worldwide production of original information, if stored digitally, in terabytes circa 2002. Upper estimates assume information is digitally scanned, lower estimates assume digital content has been compressed.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper</td>
<td>1,634</td>
<td>327</td>
<td>1,200</td>
<td>240</td>
<td>36%</td>
</tr>
<tr>
<td>Film</td>
<td>420,254</td>
<td>76,69</td>
<td>431,690</td>
<td>58,209</td>
<td>-3%</td>
</tr>
<tr>
<td>Magnetic</td>
<td>5187130</td>
<td>3,416,230</td>
<td>2,779,760</td>
<td>2,073,760</td>
<td>87%</td>
</tr>
<tr>
<td>Optical</td>
<td>103</td>
<td>51</td>
<td>81</td>
<td>29</td>
<td>28%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5,609,121</td>
<td>3,416,281</td>
<td>3,212,731</td>
<td>2,132,238</td>
<td>74.5%</td>
</tr>
</tbody>
</table>

Source: How much information 2003
How much information is being produced?

http://www2.sims.berkeley.edu/research/proiects/how-much-info-2003/

<table>
<thead>
<tr>
<th>Table 1.1: How Big is an Exabyte?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kilobyte (KB)</td>
</tr>
<tr>
<td>1,000 bytes OR $10^3$ bytes</td>
</tr>
<tr>
<td>2 Kilobytes: A Typewritten page.</td>
</tr>
<tr>
<td>100 Kilobytes: A low-resolution photograph.</td>
</tr>
</tbody>
</table>

| Megabyte (MB)                     |
| 1,000,000 bytes OR $10^6$ bytes  |
| 1 Megabyte: A small novel OR a 3.5 inch floppy disk. |
| 2 Megabytes: A high-resolution photograph. |
| 5 Megabytes: The complete works of Shakespeare. |
| 10 Megabytes: A minute of high-fidelity sound. |
| 100 Megabytes: 1 meter of shelved books. |
| 500 Megabytes: A CD-ROM. |

| Gigabyte (GB)                     |
| 1,000,000,000 bytes OR $10^9$ bytes |
| 1 Gigabyte: a pickup truck filled with books. |
| 20 Gigabytes: A good collection of the works of Beethoven. |
| 100 Gigabytes: A library floor of academic journals. |

| Terabyte (TB)                     |
| 1,000,000,000,000 bytes OR $10^{12}$ bytes |
| 1 Terabyte: 50000 trees made into paper and printed. |
| 2 Terabytes: An academic research library. |
| 10 Terabytes: The print collections of the U.S. Library of Congress. |
| 100 Terabytes: National Climactic Data Center (NOAA) database. |

| Petabyte (PB)                     |
| 1,000,000,000,000,000 bytes OR $10^{16}$ bytes |
| 1 Petabyte: 3 years of EOS data (2001). |
| 2 Petabytes: All U.S. academic research libraries. |
| 200 Petabytes: All printed material. |

| Exabyte (EB)                      |
| 1,000,000,000,000,000,000 bytes OR $10^{18}$ bytes |
| 2 Exabytes: Total volume of information generated in 1999. |
| 5 Exabytes: All words ever spoken by human beings. |

Source: Many of these examples were taken from PV Williams’ Data Powers of Ten? web page at Caltech.
Knowledge Work
[Drucker 1999]

- Knowledge about the task
- Autonomy
  - Knowledge workers have to manage themselves, they have autonomy
- Continuing innovation
- Continuous learning and teaching
- Quality at least as important as quantity
- Motivation
Why YOU should you learn about KM [Yu 2007]

Most knowledge intensive activities presuppose technology support.

- You (computer scientists) are expect to provide it.
- Do we have the right methods and tools?
- This course provides you with a set of selected tools for knowledge acquisition and organization.

Software engineering work is itself knowledge-intensive.

- We should apply KM concepts to analyze and improve software engineering – software development, deployment, evolution, etc.

Often, software projects do not fail because of technological reasons

- What kind of knowledge plays a role in software engineering? What are the implications of mis-management?
Categorization: The DMOZ Project

How effective is this for navigation?
Library Catalogs

How does this compare to other catalogs?
Large Image Collections

What is wrong about this list?
Microblogging

What is wrong with this?

Realtime results for "Thierry Henry" OR Henry

Irasocot: @jmwenz The Henry "fair play moment" opportunity is a great lesson in changing morality, sportsmanship, responsibility, nationalism.

Half a minute ago from Seesmic - Reply - View Tweet


Less than a minute ago from TweetDeck - Reply - View Tweet

KarenSadler: ...the moments that stand out...are the moments that you have done things in the spirit of love. Henry Drummond #quote #aging

Less than a minute ago from TweetDeck - Reply - View Tweet

Izzylizze: Thierry Henry...you slimeball of a cheater!!!

Less than a minute ago from web - Reply - View Tweet

Diamond: L LN: William Henry Harrison died one month after delivering the longest inaugural speech at 1 hour 45 minutes.

Less than a minute ago from API - Reply - View Tweet

Trending topics:

#Ihottosayonfirstdate
New Moon
FAA
#justbecause
Thierry Henry
#funnybunnotcool
#openwebawards
Thanksgiving
#amiright
Goodmorning
Geographical/Social Knowledge
Tupalo: An Austrian Startup Company

Why is this difficult to achieve?
Wikipedia vs. Britannica

Why are people maintaining it?
What is the difference in quality?
What is the difference in categorization?
What is the difference in contribution?
What would happen if the people “caring” for Wikipedia/Britannica were to stop caring?
Note!

• This course **does not focus on what we can‘t do** in knowledge management (and there‘s a lot we can‘t do)

• **but rather focuses on what can be done** (at least technically).
KM methods and techniques

- Knowledge Organization
- Broad Knowledge Bases
- Categorization and Formal Concept Analysis
- Categorization and Latent Semantic Indexing
- Probabilistic Topic Models
- Association Rule Learning
- Participative Knowledge Acquisition Methods
Preliminary Schedule I

Preliminary Schedule II
Some Course Highlights
Explicit or Implicit? [Kirsh 1990]

1. Is 5 as the solution to $\sqrt{125}$ explicit in $\sqrt{125}$?

2. Is the $200^{100}$ digit of $\pi$ explicit?

3. Is 3 explicit in $A: \{1,5,3,7,4,4\}$?

4. Is the cardinality of $A$ explicit in $A: \{1,5,3,7,4,4\}$?

5. Is $(6754, 9629)$ in a matrix of $10,000 \times 10,000$ explicit?

6. Is the answer to „Why does the pop star *P!nk* perform 4 Non Blondes songs at her concert“ explicit on the web?
Some Course Highlights

Categorization

Taste: Sweet/Sour, Shape: Round/Long/, Color: Red/Yellow/..., Texture: Smooth/Bumpy,

```

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>Green</th>
<th>Yellow</th>
<th>Round</th>
<th>Long</th>
<th>Sweet</th>
<th>Sour</th>
<th>Smooth</th>
<th>Bumpy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Lerruri</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Banana</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Strawberries</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grapes</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Pear</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
```
Some Course Highlights
Commonsense Knowledge

Common Sense Knowledge
ConceptNet, MIT, 2007
Some Course Highlights

Game with a Purpose

Markus Strohmaier
2009
Check

- Is there anything else you want to know w.r.t. this course?
- What aspects are you most interested in?
- Anything else?
Any further questions?

Have a good start in the new semester!
- See you next week