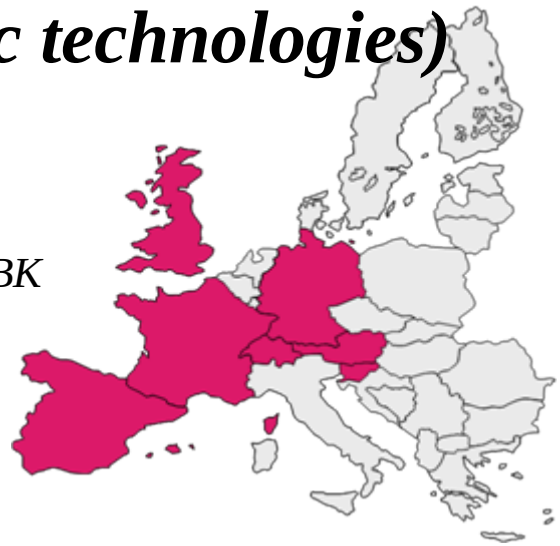
Costs, benefits, incentives (of semantic technologies)

Tutorial

Elena Simperl, Igor Popov, Tobias Bürger, Katharina Siorpaes, UIBK



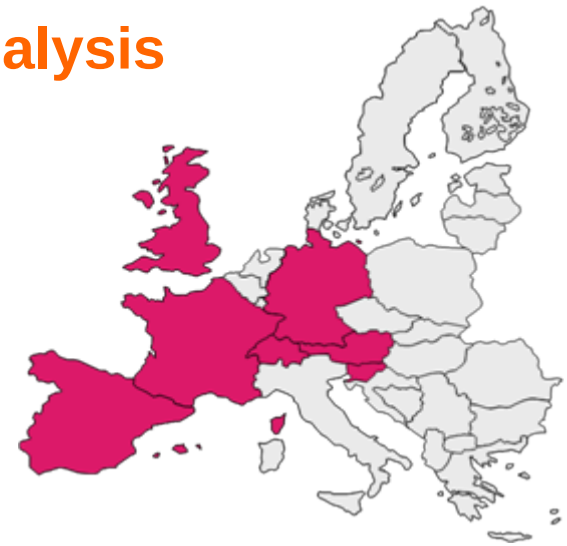
- CIOs are starting to acknowledge the technical value of semantic technologies for enterprises. In the last five years early adopting players have been increasingly using them in various application settings ranging from content management to enterprise integration platforms.

- Despite this promising position, it is still difficult to argue in favor of semantic technologies in front of the CFOs because of the lack of convincing measurable benefits.

- Semantic technologies are not designed for large-scale user participation.
- They rather aim at a complete (or at least partial) automation of the tasks, as a means to lower costs and improve productivity.
- Whilst the quality of such (fully) automated approaches has constantly improved, it is still far from outweighing the manual effort savings achieved, especially when it comes to the creation of meta-data for non-textual sources or the development of a widely accepted ontology, tasks which are human-driven through their very nature.
- To bring the humans back into the loop we have to look into incentives and motivation models for humans to use semantic technologies.

ACTIVE
KNOWLEDGE - POWERED ENTERPRISE

Cost estimation and benefit analysis



Outline

- Motivation
- Challenges of cost/benefit analysis in semantic technologies
- Cost estimation methods
- Benefit analysis
- Applicability issues
- Example: ONTOCOM

Motivation

- Assessing economic value is a key requirement for moving semantic technologies from the realm of academia to industry
- A popular and common economic metric for value in technology investments is ROI (Return of Investment)
- Cost estimation is usually carried out by using one or more methods to estimate the development effort in person-months
- Benefits analysis tries to assess the created value created by the technology investment in appropriate terms

Cost/benefit framework

- Envisioned is a framework to assess costs and benefits of using semantic technologies within enterprizes applicable to existing IT infrastructure exteded into semantics as well as to newly builts semantic systems.

- Such a framework will comprise
  - methods to estimate the cost of introducing semantic technologies into enterprise environments, including the changes triggered by this adoption at process and organizational level, and the need for training and additional know-how
  - methods to anticipate the cost savings achievable through semantic technologies
  - methods to estimate the option of investing in semantic technologies in terms of their potential business value.
  - methods to measure benefits of semantic technologies in enterprise IT systems
  - evaluation criteria and methods to assess the quality and compare alternative technological solutions.
  - the way the usage of semantics achieves efficiency gains instruments to derive and estimate the value of semantic technologies from quantitative and qualitative criteria, and to visualize the effect on overall costs and revenues according to the economic value added principle

Challenges of cost/benefit analysis

- Estimating the cost related to developing, deploying and maintaining semantic systems requires a empirically tested cost model which exploit the results from related fields (e.g. software engineering)
- Cost estimation depends on the structure of the development process, quality and quantity of data from previously similar projects
- Benefits are hard to pinpoint and quantify because often technologies investments acquire value when used in collaboration with other resources
- The nature of benefits cannot always yield countable results
  - tangible benefits (measurements which directly influence the performance)
  - Intangible benefits (benefits which can't be easily measured in financial terms)

Cost estimation methods

- Expert judgment or Delphi method
- Analogy method
- Decomposition method
- Parametric/algorithmic method

Expert judgment or Delphi method

- A structured process for gathering knowledge from a group of human experts
- Expert forecast costs on predefined cost drivers based on their experience
- Using well formed questionnaires with controlled opinion feedback
- Experts can answer questionnaires in one or more rounds
  - After each round, a facilitator can provide feedback to experts and allow experts to revise their earlier judgments
- Critique point: difficulties in explicitly stating the decision criteria used by contributing experts

Analogy method

- Use available data from similar projects to estimate costs of the proposed project

- Data from other projects are subject to:
  - Availability
  - Accuracy in establishing real differences between completed and current projects

Decomposition method

- Break a product in smaller components or into activities an task to produce lower-level, more detailed descriptions of the product/project
- Result: More accurate cost estimates?
- Success criteria: Availability of the necessary information related to the work breakdown structure

Parametric/algorithmic method

- Use mathematical model which combines input form expert and historical data to produce an estimate
- Allows analyses of cost drivers from specific class of projects and their interdependences
- Uses statistical techniques to refine and calibrate the model
- Main challenge is the availability and reliability of data

Top-down vs. bottom-up

- Top-Down method relies on the overall project parameters
  - The project is partitioned into lower-level components and life-cycle phases
  - Method is applicable only in the early stages when global properties are known

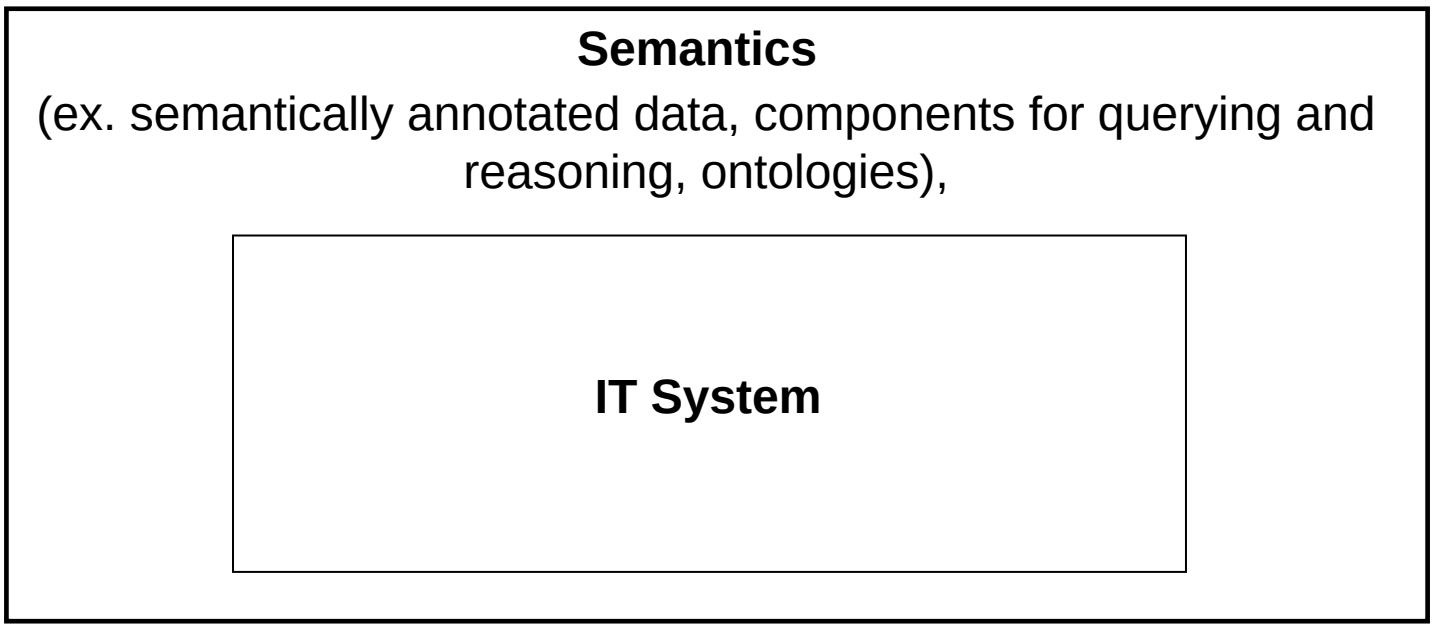
- Bottom-Up method involves identifying and estimating cost of individual project components separately
  - It cannot be applied early in the life cycle of the process because of the lack of information related to the project components
  - It is more likely to produce more accurate results

Methods and approaches to cost estimation

	Bottom-up estimation	Top-down estimation
Expert Judgment	Experts estimate the costs of low-level components or activities	Experts estimate the total costs of a product or a project
Analogy Method	Costs are calculated using analogies between low-level or activities	Cost are estimated using a global similarity function for products or projects
Decomposition Method	Costs are calculated as an average sum of the costs of lower-level units, whose development are known in advance	
Parametric Method	Costs are calculated using a statistic model which predicts the costs of lower-level units on the basis of historical data about the costs of developing such units	Costs are calculated using a statistic model which is calibrated using historical data and predicts the current value of the total development costs

Applicability

- A system that uses semantic technologies will have many new components

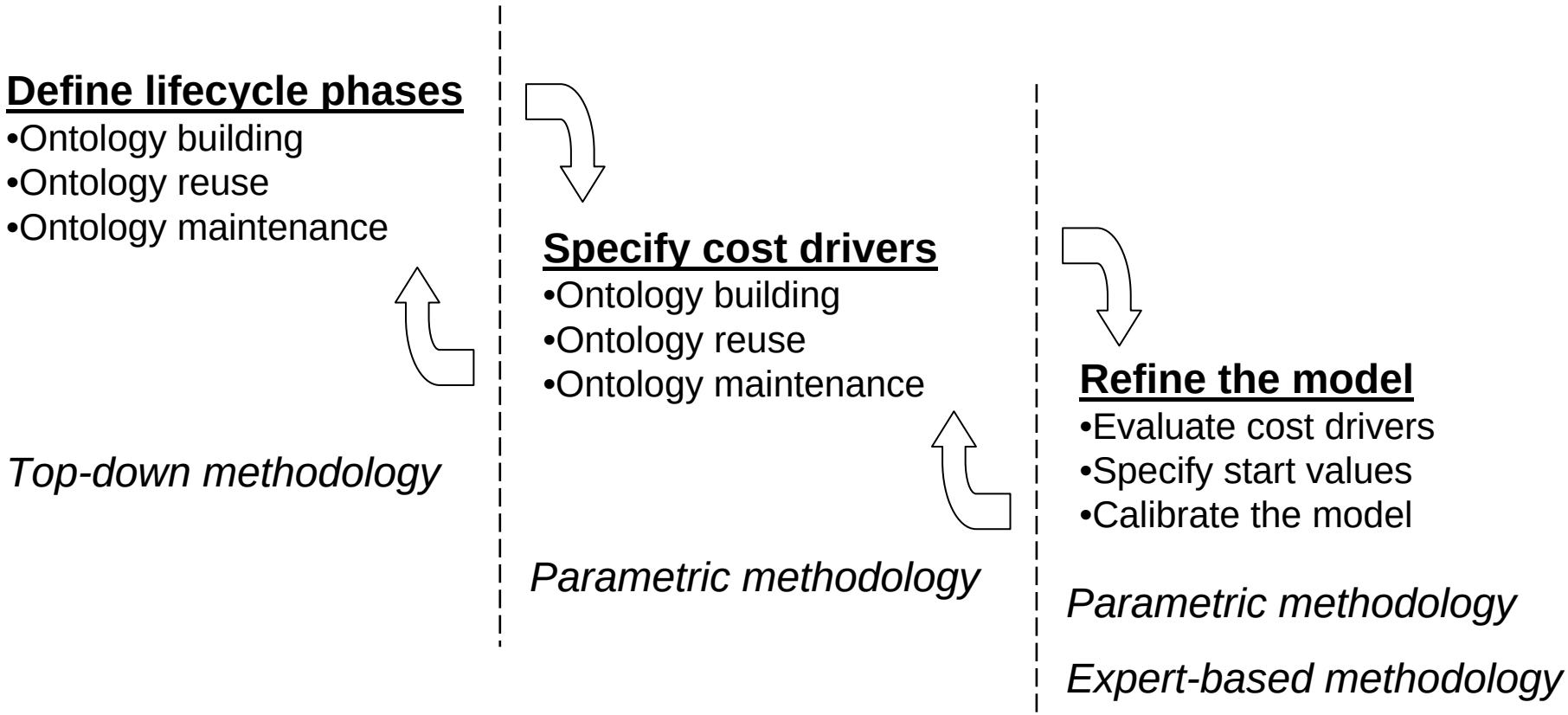


- Adding semantics to systems will mean calculating new costs
- **Challenge: finding which of the methods gives a sound basis for constructing a cost estimation model**

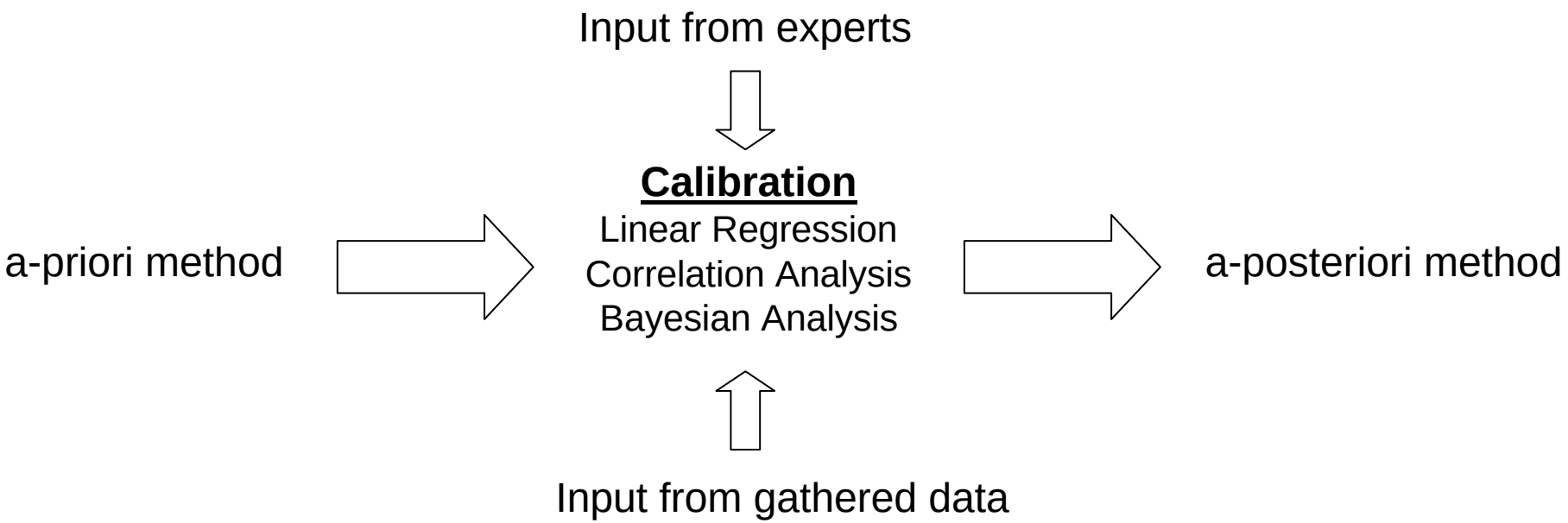
An example: ONTOCOM

- **ONTOCOM –** A cost estimation model for building ontologies
- ONTOCOM uses *top-down, parametric and expert-based* methods to form its basis for cost estimation of ontology building
- ONTOCOM is realized in three steps:
  - A top-down work breakdown structure for ontology
    - identify the cost-intensive sub-tasks of ontology development processes
  - Make a statistical prediction model (i.e. a parameterized mathematical formula)
  - Calibration of the a-priori method based on previous project data to create a valid (more accurate) a-posteriori model
    - identify cost drivers of the calculation model
    - initialize the calculation model

- How ONTOCOM works:



- ONTOCOM Model Calibration



Benefit analysis

- The nature of benefits can be
  - Tangible - directly influence the performance of the firm and as such potentially reduces costs☐
  - Intangible - influence the overall behavior and circumstances of a system indirectly

- Step one towards benefits form a certain technology is identifying all the possible benefits from it.

- Example: A list of suggested benefits from adoption of ontologies
  - Interoperability
  - Browsing / searching (automatic inferring of implicit facts)
  - Reuse
  - Structuring
  - Automation / code generation
  - Disambiguation (unique identification)
  - Knowledge transfer (by excluding unwanted interpretations through informal semantics)
  - Spotting logical inconsistencies

Benefit analysis (cont)

- All the benefits listed are intangible (they cannot be directly and easily measured), except for automation/code generation
- The diversity of different types of benefits demands a variety of applications of benefit analysis

Classifying benefits and methods

- Classifying benefits; each benefit falls into one of these categories

Investment purpose	Investment type	Evaluate/measure
Business survival	Must do	Continue/Discontinue business
Improving efficiency	Vital/core	Cost benefit
Improving effectiveness	Critical/core	Business analysis
Competitive leap	Strategic/prestige	Strategic analysis
Infrastructure	Architecture/must do/corn seed	Very broad terms

Classifying benefits and methods (cont)

- Generic approaches to measurement
  - Physical counting
  - Assessment by ordering, ranking, scoring
- Counting not always possible
- Measuring intangible benefits, several suggestions [Remenyi et al.,1995]
  - Conceptualize the chain of cause-and-effect events
  - Identify how it will be possible to establish the changes that are likely to occur as a result of the introduction of the information system. Here the focus is on the direction of the changes, i. will the inventories rise or fall?
  - Consider how the size of the change may be measured
  - Where the effect of the system is clear, the analyst may proceed with the next two steps
  - Measure the size of the change
  - Put a monetary value on the changes that have been observed. Use techniques such as payback, Return-Of-Investment Net Presence Value, etc.

Classifying benefits and methods (cont)

- Methods for assessment can be grouped according to their output:
  - Financial methods
  - Quantitative methods
  - Qualitative methods

- Selection of methods should be selected based on the use case to which they are applied

Example: User Information Satisfaction

- Estimating User Information Satisfaction (UIS) from use of ontologies
  - It measures intangible benefits
  - It will not have a financial output
  - It will produce a quantitative output
- Choosing a method(s) from an defined set of methods [Remenyi et. al.] based on the use case
  - Single Gap vs. Multiple Gap and factor analysis
  - Using a questionnaire
- Choosing a method is on a case-by-case basis: UIS for ontologies use Single Gap using questionnaires (Tobias Bürger, SALERO)

Conclusions

- Cost/benefit analysis is a hard (but necessary) thing to predict/measure
- Cost methods depend on the availability and quality of data
- Benefit analysis methods can not always be countable
- Cost/benefit methods are regularly refined and adapted for use for specific areas (like ontologies)

ACTIVE
KNOWLEDGE - POWERED ENTERPRISE

Incentives

Some observations

- Lack of semantic content
- Lack of user involvement
- However, not all the tasks on the Semantic Web can be automated

  - Building ontologies,
  - Annotating content, and
  - Aligning ontologies

at least partly require human intelligence.

- **How do we motivate people to contribute to semantic content authoring? Web 2.0 has done this very successfully: some examples.**

Powerful Web 2.0: A Selection

XING

facebook

del.icio.us
your bookmarks

LinkedIn

last·fm
the social music revolution

skype

YouTube
Broadcast Yourself™

myspace.com
a place for friends

BT

iSOCO

AIFB

"Jožef Stefan" Institute

fit

Microsoft
>
accenture

cadence

EURESCOM

Powerful Web 2.0: Some examples

- File Sharing:
  - Flickr
  - YouTube (Videos)
  - Wikipedia
  - Blogs
  - Open Source Community (Linux)
- File management: from file hierarchies to tagging

- Social Portals:
  - Facebook
  - LastFM
  - Skype
  - LinkedIn, Xing
- Open Systems: APIs, open source allow further development

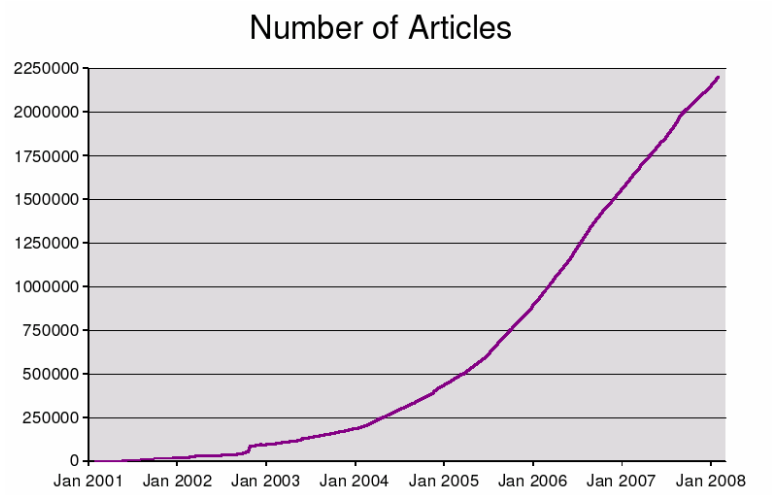
Platform for social networking

- Founded in 2004
- **64 Million** active members
- **250,000 new registrations** daily
- More than half of members are not in college anymore
- More than **65 Milliarden page views** a month
- More than half of members use Facebook **daily**
- Avg. duration 20 minutes
- **15 Billion Dollar**

Wikipedia

- **2,214,717 Articles** (english)
- **6,383,758 users**

- High quality
- Open and uncontrolled



- Within 1month: amount of videos to **6.1 Million**
- 45 Terabyte Videos
- **1.73 Billion Video** Views

- Google bought YouTube for **1.6 Billion Dollar**

Web 2.0 Incentives

- **Altruism**
- **Reciprocity** (Tags: Organisation, Reuse)
- **Reputation**
- **Competition**
- **Belonging to a community, a common goal**
- **Autonomy, freedom**
- **Attracting attention**
- **Self Portaits** (Facebook)
- **Social Component**

(Kuznetsov, 2004; Marlow et al., 2006)

Summary

- Web 2.0 generates a huge amount of data and many people contribute
- Each application implements an incentive
- We have to investigate those incentives
- And find out, how we can apply them to the Semantic Web
- In order to generate more semantic content.
- Examples for this:
  - Semantic MediaWiki
  - OntoGame
  - myOntology

Thank you for your attention