DIT590 Research Methods & Technical Writing

Lecture 2: Designing Research

Richard Berntsson Svensson
http://www.rbsv.eu/courses/rmtw

Re-cap

Quantitative	Mixed	Qualitative
• Pre-determined • Instrument based questions • Performance data, attitude data, observational data. • Statistical analysis • Statistical interpretation	• Both pre-determined and emerging methods • Both open- and closed-ended questions. • Multiple forms of data drawing on all possibilities • Statistical and text analysis.	• Emerging methods • Open-ended questions • Interview data, observation data, document data and audio-visual data • Text and image analysis • Themes patterns and interpretation

Methodology	Primary objectvie	Primary data	Design
Survey	Descriptive	Quantitative	Fixed
Case study	Exploratory	Qualitative	Flexible
Experiment	Explanatory	Quantitative	Fixed
Action research	Improving	Qualitative	Flexible

How to design research?
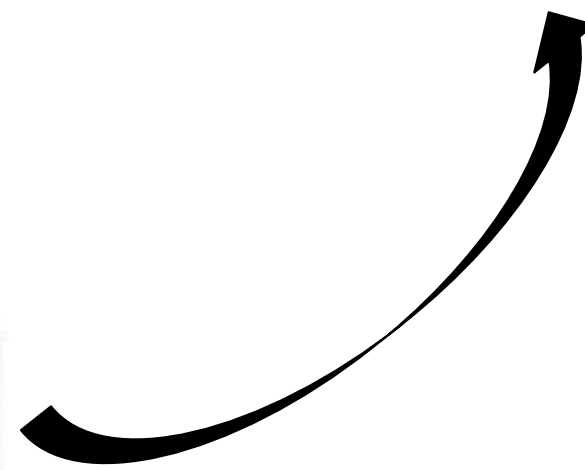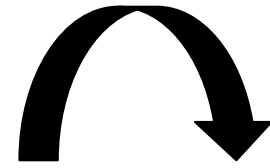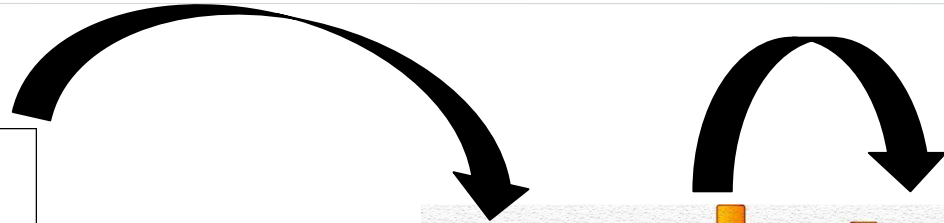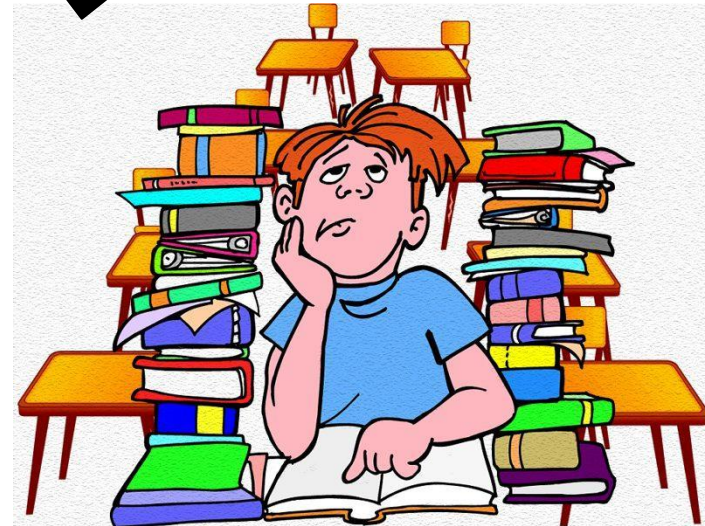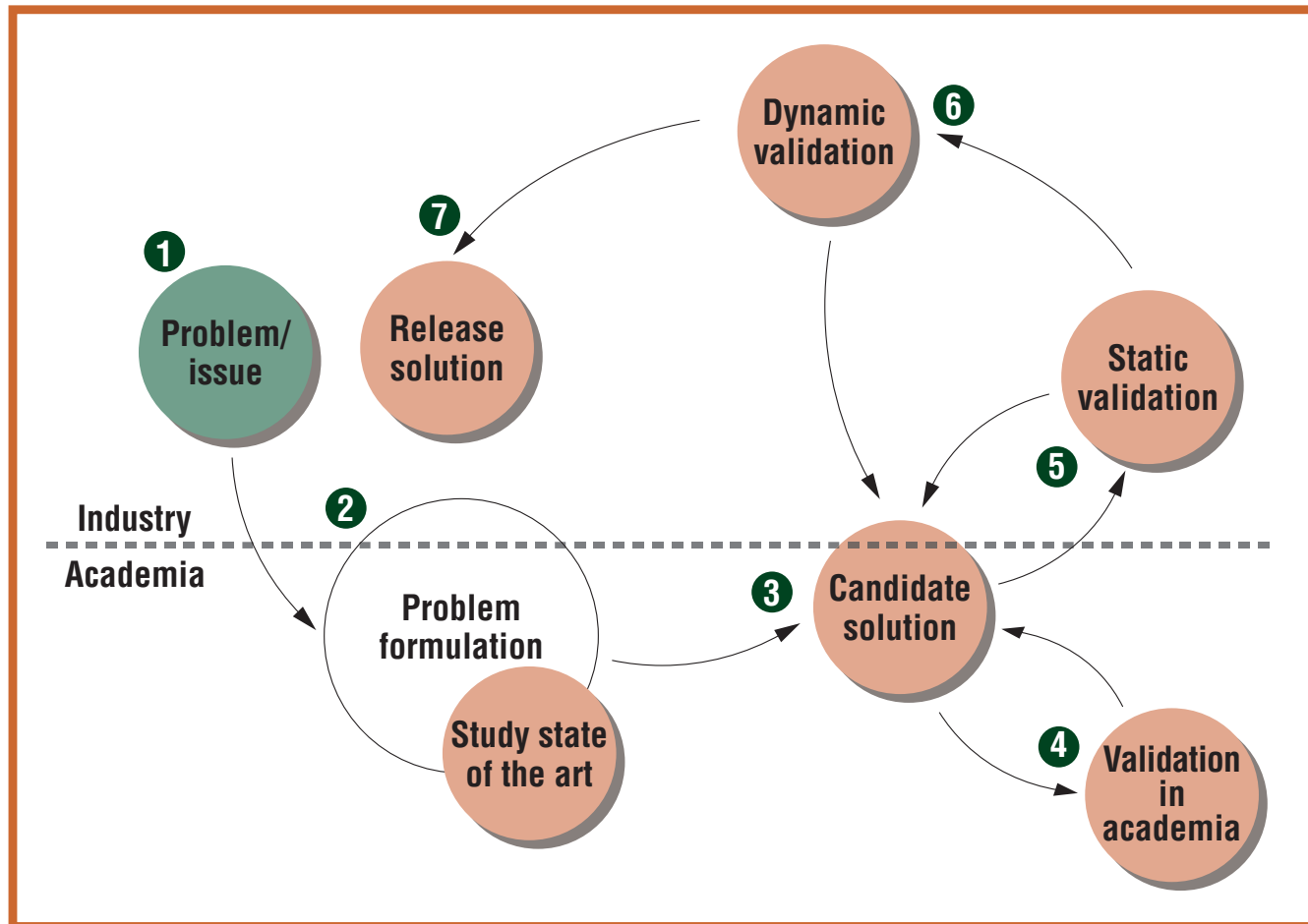
Technology Transfer in Practice

[TTiP]

Checklist for choice of thesis topic

- **Topic**
  - Why is it important?
  - Why urgent now?
  - Why can be solved/approached now?

	Scientifically	Industrially
1. Why is topic important?		
2. Why urgent now?		
3. Why can be solved/approached now?		

[R. Feldt, *Checklist for choice of thesis topic*]

Designing Research: The Purpose Statement

- **Qualitative**
  - Information about the central phenomenon explored in the study, the participants in the study, and the research site
  - *Purpose*, *intent*, *objective*
  - *Describe*, *understand*, *develop*, *examine the meaning of*, *discover*

- The purpose of this _____ study is to _____ the _____ for _____ at _____. At this stage in the research, the _____ will be generally defined as _____.

Designing Research: The Purpose Statement

- **Quantitative**
  - Includes the variables in the study and their relationships, the participants, and the research site.
  - *Purpose*, *intent*, *objective*
  - *The relationship between, comparison of*

- The purpose of this _____ study is to test the theory of _____ that _____ the _____ to ____, controlling for _____ for _____ at _____. The independent variable(s) _____ will be defined as _____. The dependent variable(s) ____ will be defined as _____, and the control and intervening variable(s), _____, will be defined as_____.

Designing Research: The Purpose Statement

- **Mixed methods**
  - Information of the overall intent, about both qualitative and quantitative strands of the study, a rationale of incorporating both strands to study the research problem.
  - *The purpose of, The intent of*
  - ***Sequential**, concurrent, transformational*
  - Discuss reasons for combining both quantitative and qualitative data.

- The intent of this two-phase, sequential mixed methods study is to _____. The first phase will be a qualitative exploration of a _____ by collecting _____ from _____ at _____. Findings from this qualitative phase will then be used to test _____ that relate/compare _____ with _____ for _____ at _____. The reason for collecting qualitative data initially is that _____

What are Research Questions?

Why are they important?

R. Feldt, *Guide to research questions*, available at: http://www.robertfeldt.net/advice/guide_to_creating_research_questions.pdf

Types of Research Questions

- Creating (better) solutions
- Creating (better) knowledge

- Knowledge focused RQ
  - RQs are mainly *Exploratory*
  - *Base-rate* RQs
  - *Relationship* RQs

Main Types of RQs	Answers
Exploratory	Answers give clearer understanding, better definitions of concepts, evidence that we can measure them validly
Base-rate	Answers describe the normal pattern of occurrence of the phenomena
Relationship	Answers describe if and how two phenomena are related
Solution-focused	Describes better ways to solve some problem or situation. Examples are: *"What is an effective way to achieve X?"*, *"Which strategies help achieve X?"*, *"How can we refined S to achieve X in a better way?"*

Designing Research: Research Questions

- **Qualitative**
  - States RQ, not objectives or hypotheses
  - Two forms
    - Central question
    - Associated sub-questions
  - *What*, *how*

- **Examples**
  - How do women in a psychology doctoral program describes their decision to return to school?
  - What are the current practices that support achieving REVV alignment?

Designing Research: RQ

- **Quantitative**
  - Research questions & hypotheses, sometimes objectives
  - Null hypotheses – predictions
  - Statistical procedures

- **Examples**
  - Hypothesis 1: Publicly traded firms will have higher growth rates than privately held firms.
  - RQ: What is the impact of the phenomenon of obsolete software requirements on the industry practice?

Designing Research: RQ

- **Mixed methods**
  - Mixed methods research question

- **Example**
  - How do the themes mentioned by the teachers help to explain why middle-school children score lower than the junior high students?

Several studies [7, 13, 16] observe that developers rely on informal and ad-hoc communication. Lack of, or problems in the informal communication channels may lead to increased development time [16]. Communication issues in GSE have been addressed by other studies; see e.g. [6, 11, 16], while communication in new product development has been addressed by, e.g., [8]. However, none of these have focused on comparing local peer-to-peer (e.g. face to face meetings), long-distance peer-to-peer (e.g. electronic chat, including instant messaging) and technical (e.g. architecture) communication tools in GSE, for their ability to provide information in a timely manner, with richness and with large volumes of information.

The main objective of this study is to assess the relative importance of ten commonly used communication mechanisms and practices from three different aspects, for their ability to transmit information quickly, transmit rich information, and to transmit large volumes of information across local and global development sites.

RQ1: What communication mechanisms are central to a distributed SPL environment in order to provide large volumes of rich information in a timely manner?

RQ1.1: What communication mechanisms are central to a distributed SPL environment in terms of speed?

RQ1.2: What communication mechanisms are central to a distributed SPL environment in terms of providing rich information?

Despite their importance, QR are often poorly understood, generally stated informally in a non-quantifiable manner, often contradicting, and difficult to validate [9], [22]. This is further aggravated in market–driven development, where the situation is even more complex [3], due to the large number of requirements stemming from multiple internal and external sources, and the continuous flow into the development organization [16], [18]. The challenges associated with QR have been addressed in part by other studies, see e.g. [25], [26], and [32]. However, none of these have primarily focused on QR.

This exploratory study can be seen as a study of state-of-practice in industry, but also an investigation as to what extent state-of-the-art in research, in terms of methods and tools, has penetrated industry practice.

This paper presents the results of an empirical study that includes data collected through in-depth interviews with twenty-two practitioners from eleven different companies in Sweden of which six are multinational. The study focuses on the elicitation, analysis and negotiation, management, and general handling of QR in industry.

RQ1: What QR are considered most important, and are there any distinguishably characteristics in relation to customer type (B2B vs. B2C)?

RQ3: How are cost estimations of QR performed, and what is the accuracy of these cost estimates?

RQ4: To what extent are QR dismissed from projects after project initiation?

The evaluation of this method showed that using the method can significantly improve the performance of the consolidation process as well as the number of correctly linked requirements, and that it can help to miss fewer requirements links (Natt och Dag et al. 2006). However, the *unsupported* method used in the original experiment was limited to a simple search functionality, while most currently available requirements management tools offer more advanced filtering and searching techniques.

This replication study has been designed to assess whether the tool with a linguistic analysis of the similarity between requirements can still perform better than currently available commercial requirements management tools in the task of requirements consolidation.

The objectives of the study are twofold: firstly to assess if a significant differences between the two methods tested in the original experiment can be confirmed in a replicated experiment setup and secondly to compare the results for the same methods between the two experimental sessions.

Q1: Can significant differences between the *assisted* and the *manual* methods that were achieved in the original experiment be confirmed in a replicated experiment where the original *manual* method is replaced with a keyword searching and filtering tool?

H_0: The *assisted* method results in the same number of requirements analyzed per minute as the *manual* method.

Much advice has been and is given regarding how software should be designed at many granularities, such as criteria for modules [17], common design idioms [10], and use of components [23]. With the increase in availability of open-source software, a number of empirical studies have been performed in recent years that give a picture of how software is really being written, and the results do not always appear consistent with received wisdom. We would like to know whether the perceived inconsistency is real, whether it is due to lack of understanding by developers, lack of belief that the design principles have the claimed benefits, or whether the design principles in fact do not have the claimed benefits. As the bulk of the empirical studies is on code written in object-oriented languages, in our study we focus on principles for object-oriented design.

There are three main goals of the our empirical study, of which parts are presented in this paper. The first is to gauge how the theory and best-practice put forward by experts is used and interpreted by practitioners. Second is to identify the actual preferences of practitioners when it comes to practical issues such as making design decisions. Third, but probably most important, we want to determine whether or not the practitioners are in agreement.

RQ1: How are object-oriented concepts understood and used in practice, and how does the use compare to recommended best-practice?

RQ1.1: How is the concept of encapsulation understood and used in practice, and how does the use compare to recommended best-practice?
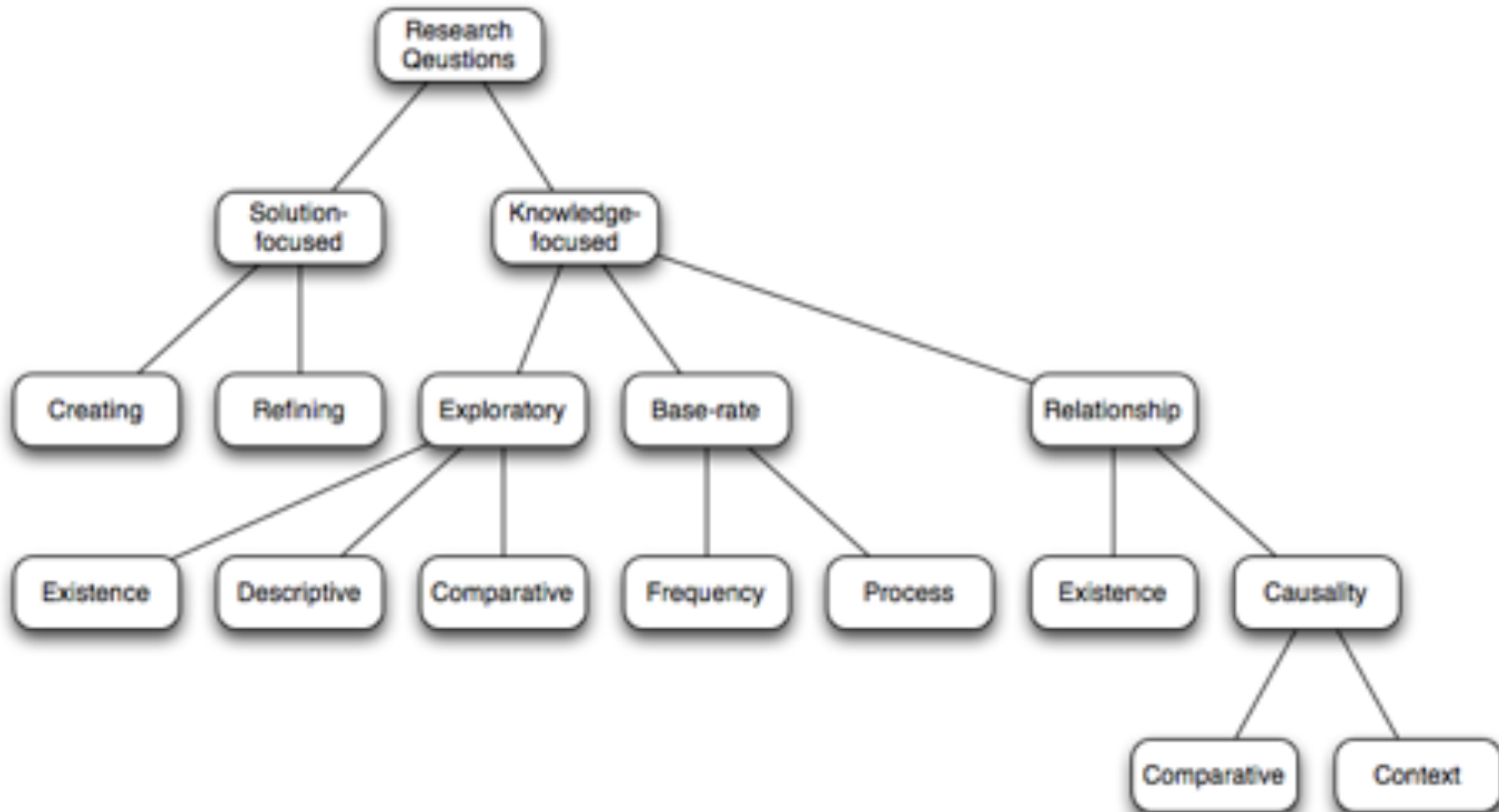
Creating research questions

- **Are you primarily focused on creating more and better understanding?**
  - "Not much is known, we need to explore." => Consider which types of Explorative RQs is most suitable.
  - "We now a lot about the phenomena but not how common it is, or how and when it occurs" => Consider primarily Base-rate type of RQs
  - "We have good descriptions about the phenomena and its occurrences but we do not understand how it relates to other phenomena or what causes it" => Consider Relationship type of RQs
- **Are you primarily focused on creating a better solution to a problem?**
  - Consider different Solution-focused RQs

Sub-Types of RQs	Examples
Exploratory/Existence	*"Does X exist?",* *"Is Y something that software engineers really do?"*
Exploratory/Descriptive	*"What is X like?",* *"What are its properties/attributes?",* *"How can we categorize/measure X?",* *"What are the components of X?"*
Exploratory/Comparative	*"How does X differ from Y?"*
Base-rate/Frequency	"How often does X occur?", "What is an average amount of X?"
Base-rate/Process	*"How does X normally work?",* *"What is the process by which X happens?",* *"In what sequence does the events of X occur?"*
Relationship/Existence	*"Are X and Y related?",* *"Do occurrences of X correlate with Y?"* *"What correlates with X?"*
Relationship/Causality	*"What causes X?",* *"Does X cause Y?",* *"Does X prevent Y?",*
Causality/Comparative	*"Does X cause more Y than Z does?",* *"Is X better at preventing Y than Z is?"*
Causality/Context	*"Does X cause more Y under one condition than others?"*

Plan for Research Study

- **Base data about study**
  - Name, acronym, driving researcher, participating researcher(s)
  - Partners
  - Target forum
  - Suitability for target
- **Main idea for study**
  - 2-6 sentences
- **Background**
  - Short summary of previous results
- **RQ/hypotheses**
  - List one main and possibly 1-4 sub-questions/hypotheses

Plan for Research Study

- **Research strategy/methodology & empirical work**
  - Overall research strategy
  - Data collection methods
  - Empirical work
  - Development work
- **Time plan**
  - Idea/planning
  - Study design
  - Preparation for data collection
  - Data collection
  - Analysis
  - First draft
  - Ready to be submitted

Assignment 1: Survey

- Compose a possible survey study (where questionnaire is the data collection method) based on a SE topic that interest you. Describe the purpose of your survey, population and size of population, instrument used, content areas addressed, scales, variables in the study and validity issues

- Create questions and create an online questionnaire (e.g. google doc, soSci…), write report.

- Peer-review other groups' assignments, give feedback (both oral and written)

Assignment 1

1. Introduction

- Statement of the problem
- Existing literature of the problem
- Why it is important to conduct research about the problem
- For the above three items, a minimum of 5 trustworthy references (in total) must be used when describing the problem, why it is important etc.

2. Purpose of the study

- Purpose statement (concise and clear) leading into RQ(s)
  - i.e. describe the purpose of your survey/questionnaire

2.1 Research question

- State a clear and concise main research question that will be investigated
- State another one or two sub research questions related to the main RQ.

Assignment 1

3. Methodology

- Describe and explain the research strategy (i.e. Survey/Questionnaire) and investigation methods that will be used for the study. Motivate your choice of survey/questionnaire using relevant references.
- Describe data collection procedures, sampling method(s), number of subjects, response rate etc.
- Create questions for your survey (include the questions in your report)
- Describe the population and size of population
- Describe and motive method for analyzing the data
- Discuss validity threats

4. References

- Use one referencing style in the report. Be consistent with how you use it. Recommended to use the style in the IEEE template

Acknowledgement

- Write who is responsible for what part

Assignment 1

- Peer-review and discussion of assignment 1. Be prepared to read and give constructive feedback to each others' work. **Maximum 1 page** (given during seminaries)

- Criteria for judging the assignment:

  - **Purpose statement** - is the central purpose of focus of the study clear?

  - **Relevance** - is the stated research questions relevant? Does it address a real problem?

  - **Suitable research design** - is the selected research design suitable to answer the RQ? IS the research design justified?

To do ...

- Register for the course!
  - Contact student office
  - student_office.cse@chalmers.se

- Find someone to work with
  - Send email to richard@cse.gu.se

- Start with Assignment 1

- Theory: Read CRES: 8

- Attend Monday's lecture about Survey