12. Basics of Information Gathering (Reading, Screening, Strategy)

Prof. Dr. Uwe Aßmann

Softwaretechnologie

Fakultät Informatik

Technische Universität Dresden

2013-1.1, 11/27/13

uwe.assmann@tu-dresden.de

http://st.inf.tu-dresden.de/asics

http://st.inf.tu-dresden.de/vba
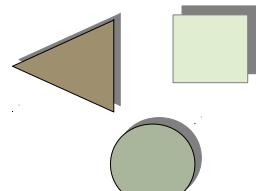
1) Data, Information, Knowledge
2) Reading Process RIIK
   1) Checking Relevance of Texts
   2) Information Aquisition
   3) Knowledge Aquisition
3) Other Reading Methods
4) Methods of Recite
5) Information gathering

---

Obligatorische Literatur

- Joachim Stary/Horst Kretschmer. Umgang mit wissenschaftlicher Literatur. Cornelsen.

- Christine Stickel-Wolf, Joachim Wolf: Wissenschaftliches Arbeiten und Lerntechniken. Erfolgreich studieren – gewusst wie! Gabler, 5., aktualisierte und überarbeitete Auflage 2009

- Reading methods are well described under http://www.teachsam.de/

- http://cseweb.ucsd.edu/users/wgg/CSE210/howtoread.html

13.1. Data, Information, Knowledge Aquisition in Science

- Repetition for ASICS

3
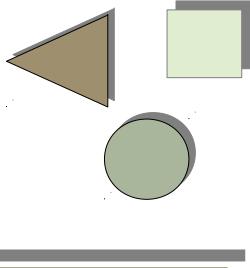
Science is about DIKW (Data, Information, Knowledge, Wisdom)

▲ Philosophy of Science quarrels about the right model for DIKW.

▲ The relationship of DIK and W is important for science, because
  - Natural science finds data in the world and has to interpret them to knowledge
  - Technical science should use knowledge to solve problems, but needs to be wise, because technology can be dangerous (e.g., see the use of nuclear energy)

▲ One DIKW model is the DIKW pyramid:

Wisdom

Knowledge

Information

Data

4
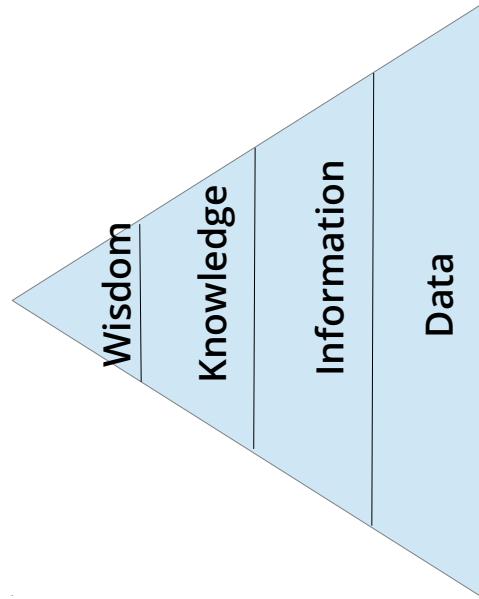
The Knowledge Aquisition Model from Spinner

- Knowledge is context-dependent and gained from information by interpretation [Prof. Helmut Spinner, Karlsruhe, Keynote at Fakultät Informatik, 1997]
- Every human being judges on a message immediately, answering 10-15 questions immediately
- Answering the questions creates knowledge
- What do I think about information such as:
  – "Das schmeckt gut."
  – "Das ist aber interessant"
  – "Du Idiot"
  – "Du bist ein Schlingel"
  – "Du bist aber schlau"

Information

Hear

Answer questions

context

Understand

Wissen (knowledge)

Communication

Typical Questions for Interpretation
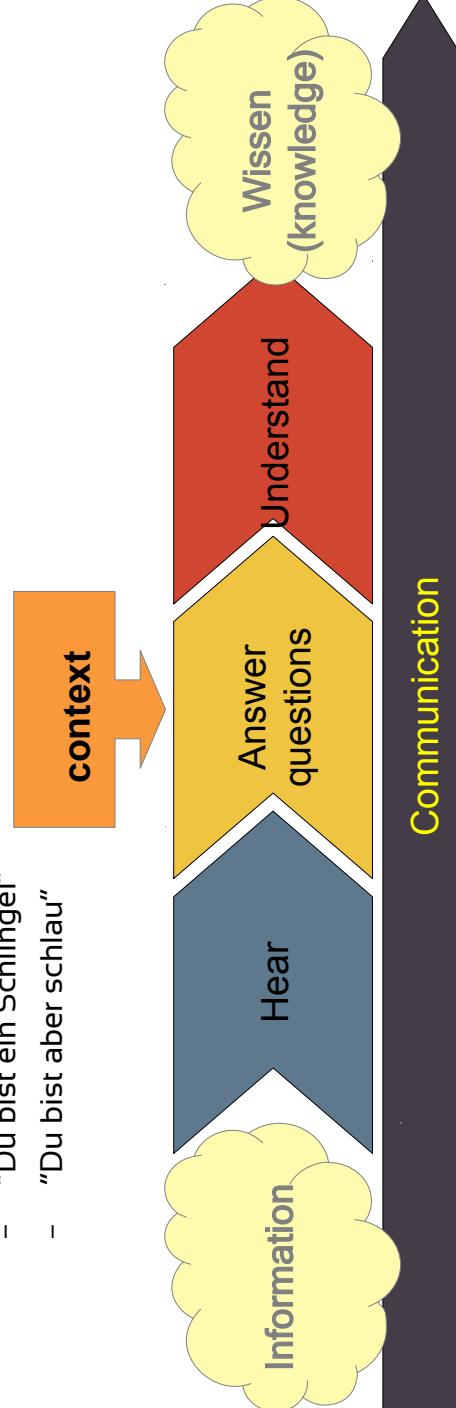
- About the sender:
- In which emotional state is the sender? (angry, sad, happy, joking, serious)
- Is the sender trustworthy? (unknown, friend, competitor, enemy, have I been disappointed by him already?)
- Which personality has the sender? (serious human being, funny, thinker, surperfical type, depressive,...)
- which channel has the sender used previously (facts, emotions, relations, etc.)?
- About the receiver:
- Which are my current expectations? Which channel do I expect?
- My emotional state
- About the context:
- In which state is the relationship (peace, quarrel, ..)
- the communication? (stress, hurry, joking, ..)

How Information Becomes Knowledge

- How do you interpret the remarks
  - "Das schmeckt gut."
  - "Das ist aber interessant"
  - "Du Idiot"
  - "Du bist ein Schlingel"
  - "Du bist aber schlau"
- from your partner? from your friend? from your mother?
- from your competitor?
- from your boss?

Knowledge is what remains after answering questions.

Knowledge is what remains in the scientist after answering questions of his value system.

13.2. "Lazy" and Efficient Reading Process with Revelance Check, Information and Knowledge Aquisition (RIK)
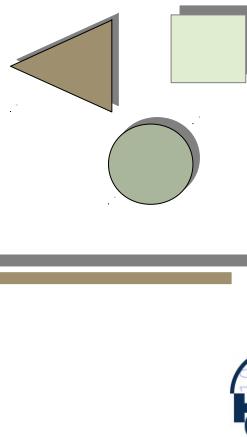
Problems with Reading
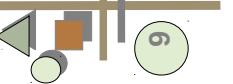
- ▲ Bored, unfocusedness, tired
- ▲ Bad habits
  - – Read word by word
  - – Jump back (regression)
  - – Talk while reading
  - – Listening to rock music while reading
- ▲ Good steps
  - – Cover read lines or unread by paper sheet or ruler
  - – Jump-Stop movement
  - – Summaries
  - – Questions

- ▲ Good steps
  - – Increase your width of focus
  - – Try to read fast
  - – Read slower, if text is hard to understand
- ▲ Try to find the *thesis statement* of the paragraph, forget the rest
 - – read only the skeleton

---

Reading: Objectives
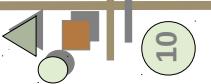
- ▲ Don't loose time: try to find out **relevant** things to read
  - – Be quick in reading and filtering
- ▲ Learn one or two main concepts
  - – Find out about important research questions
- ▲ Clearify your own position
- ▲ Find the position of your own work with regard to the paper you read (demarcation)
  - – Research landscape
  - – Qualitative analysis according to comparison criteria
- ▲ Prepare thesis writing
  - – Be able to write a section on what you read into your thesis report
  - – Be able to compare

Lazy Reading Process RIK
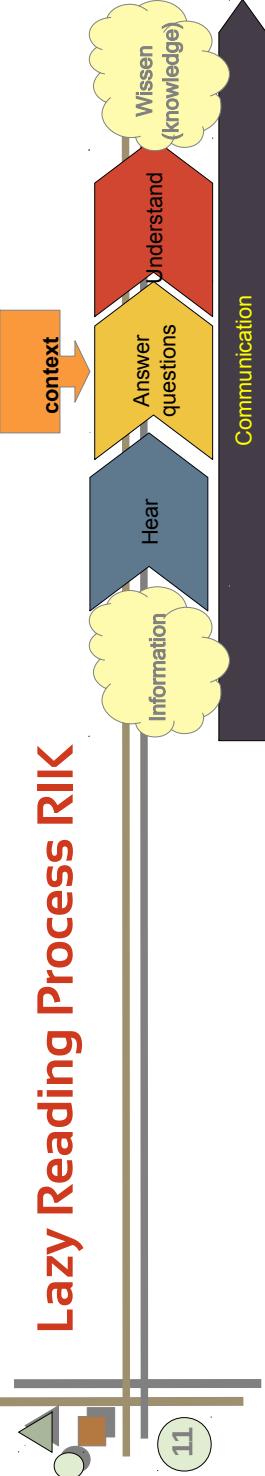
- RIK is a simple reading process allowing for stopping all the way and not wasting time
- Later, we will see more advanced reading processes
- The RIK process is structured along the DIKW pyramid and the Spinner IK knowledge acquisition process:
  - without questions no knowledge from information
  - without recitation no knowledge: no embedding of the information in your own knowledge (self-context)

Wissen (knowledge) — Understand — Answer questions — Hear — Information — context

Communication

Wisdom — Knowledge — Information — Data

Knowledge Acquisition ← Information Aquisition ← Relevance Check

Recite ← Reading ← Questions ← Orient ← Survey

11

Lazy Reading Process RIK

Knowledge Acquisition ← Information Aquisition ← Relevance Check

Recite ← Reading ← Questions ← Orient ← Survey
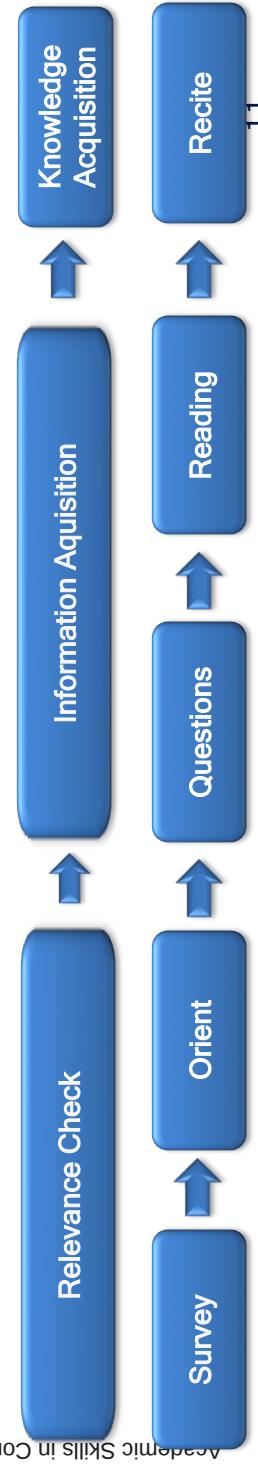
- Analyse Paratext: Table of Contents
  - Find out focus of work (Schwerpunkte)
  - Separate background from the author's work
  - Skip analysis: Find out chapters you know already and decide to skip them
- Paging through ("Durchblättern")
  - How long are main chapters?
  - What is side material?
- Analyse Eye Catchers
  - Figures, tables
  - Central definitions
  - Other structuring aids

Lazy Reading Process RIK

Knowledge Acquisition → Recite

Information Aquisition → Reading

Relevance Check → Questions

Orient → Survey
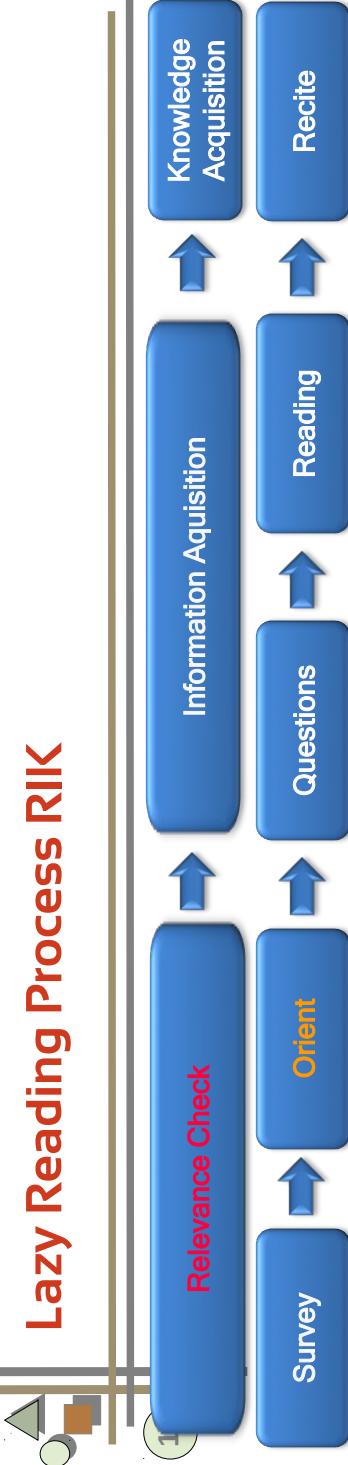
▲ Analyze the introduction
- Research question? Research Hypothesis? Research Method? Research Validation?
- Relevance
- Positioning into the research landscape

▲ Analyze summary or conclusion
- Central points, results
- What should be read more intensively? what are the main sections to be read?

▲ Read the skeleton (the first sentence of each section)
- Relevance check of the sections
- "Points" of the sections
- Get an overview about the argumentative structure of the paper

---

13.2.1. Relevance Check
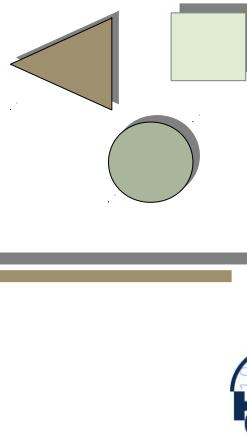
- Find out goal of the author
- Relevance of research topic
- **Main text**: Title, preface, introduction, table of contents, summary, …
- **Nebentext**: Envelope, Blurb, Recensions, Amazon comments...
- **Computer reviews**: a journal with reviews of papers
- **Google, google, google**

- Ex.: **Blurb (Klappentext)** aus Wikipedia, der freien Enzyklopädie
- „Als Klappentext (auch unter der Bezeichnung *Waschzettel* bekannt) wird ein auf den Einschlagklappen eines Schutzumschlags stehender Text bezeichnet. Der Verleger Robert Langewiesche gilt als "der Erfinder" des Klappentextes. Üblich sind eine kurze, werbende Zusammenfassung des Buchinhalts (meist auf der vorderen Einschlagklappe), eine Autorennotiz (meist auf der hinteren Einschlagklappe) und gegebenenfalls Hinweise auf weitere Bücher des Verlags."

Relevance Check: Survey, Preview, and Selective Reading

Survey: Analyse paratext: Title, Blurb, Impressum, Table of contents, References, Glossary

Try to find out the objectives of the author: why did she write the text?

Content relevant?
- irrelevant → stop
- relevant

Preview: read Preface, Conclusion, Summary, Outlook

Level of complexity / Level of relevance
- too low → stop
- too high → later
- appropriate

Selective Read: Read some pages of singular chapters

buy, lend, read, excerpt book

Training Unit: Writing Paratext

- ▲ What is important when you write a paratext as an author?

12.2.1.2 Relevance Analyses for Research

- Not all research is *relevant*
- Often, decisions have to be made about which way to go in research. Several general analysis for strategy can be used.
- [more material in course "Software Management (summer)"]
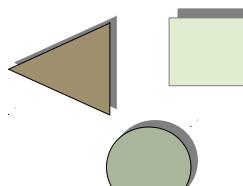
Different Classes of Research Results

Try to classify every paper you read according to the following classes:

- **Relevant research**: Somebody, the **research stakeholder,** needs the result.
- **High innovation depth**: research result lies much beyond the state of the art
- **Narrow result:** the research result will not influence many applications, products, or markets, nor other research.
- **Disruptive result:** The research result will change many technologies, products, markets, value chains.
- **Epsilon-result:** The research result is not far away from the state of the art, but contains a definite improvement.
- **"low hanging fruit"-result:** the research result is quite easy to achieve or to document

---

# What are the Strategic Aspects of a Paper? SWOT Analysis as a 4-D Analysis

- SWOT is a 4-dimensional attribute analysis for of the development of a strategy for of a project [Albert Humphrey]
- For strategic decisions of your thesis and your research
- Try to combine with the 6 honest men!

Internal analyse
Internal attributes

External analysis
External attributes

Strengthes

Weaknesses

Threats

Opportunities
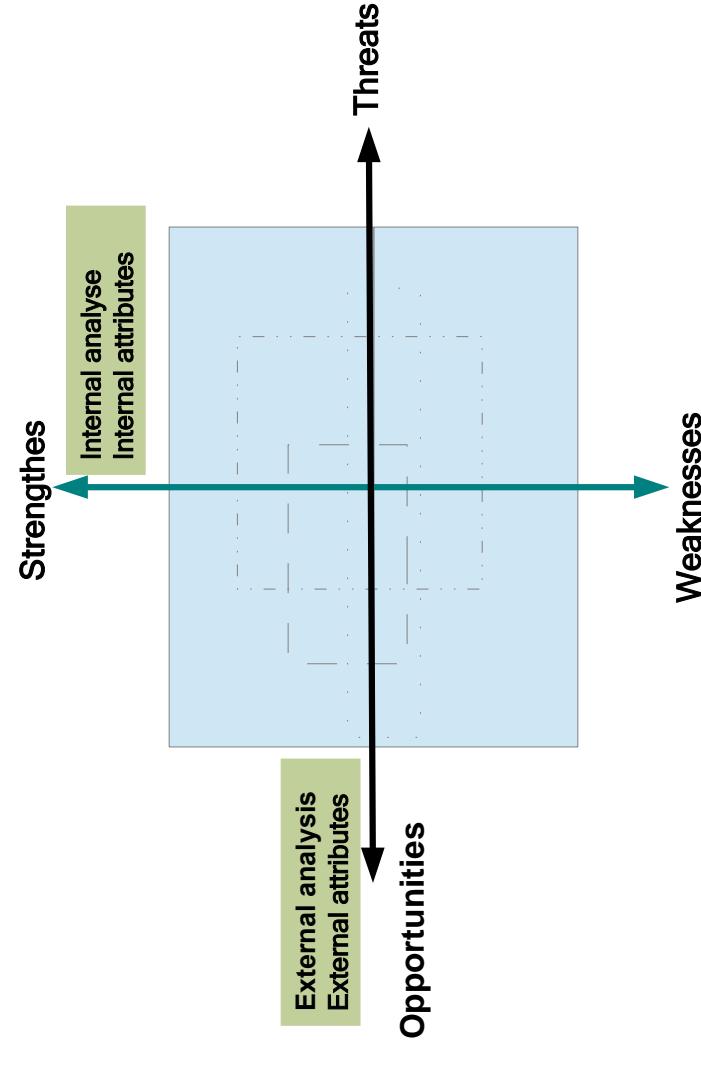
„KillerApp" Analysis (Attractivity Portfolio)

The portfolio diagram shows:
- Y-axis: Comprehensibility (high to low)
- X-axis: Need (low to high)
- Top-left (red): KillerApp
- Top-right (green): Mass App
- Bottom-left (grey): Good John
- Bottom-right (yellow): Technical Golden App
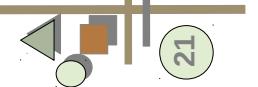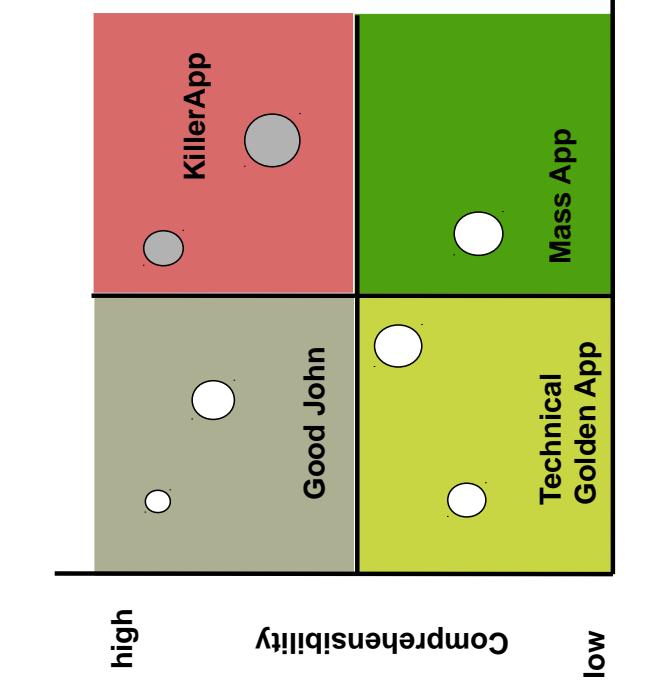
▲ **"KillerApp"-Analysis** investigates for a product or a research paper
- whether it is needed
- whether it is comprehensible

▲ the **Attractivity Product** is a Utility-utility-product:
- **Attractivity** = Need * Comprehensibility

▲ Most attractive papers or projects are "KillerApps", because they are easy to comprehend and useful for many


Home Work

▲ Conduct a "low hanging fruit" analysis for the topic "Smart Grid"

▲ Conduct a "Riller app analysis" for the topic Smart Grid
- find a Riller ap
- find a golden technical app

13.2.2. Information Aquisition

Lazy Reading Process RIK - Information Aquisition

Information Aquisition

| Relevance Check	Information Aquisition	Knowledge Acquisition
Survey → Orient	Questions → Reading	Recite
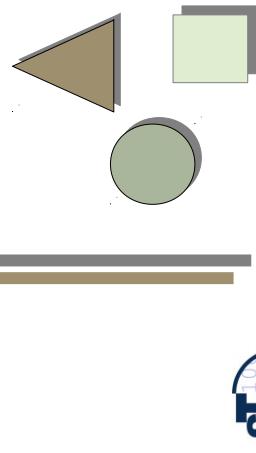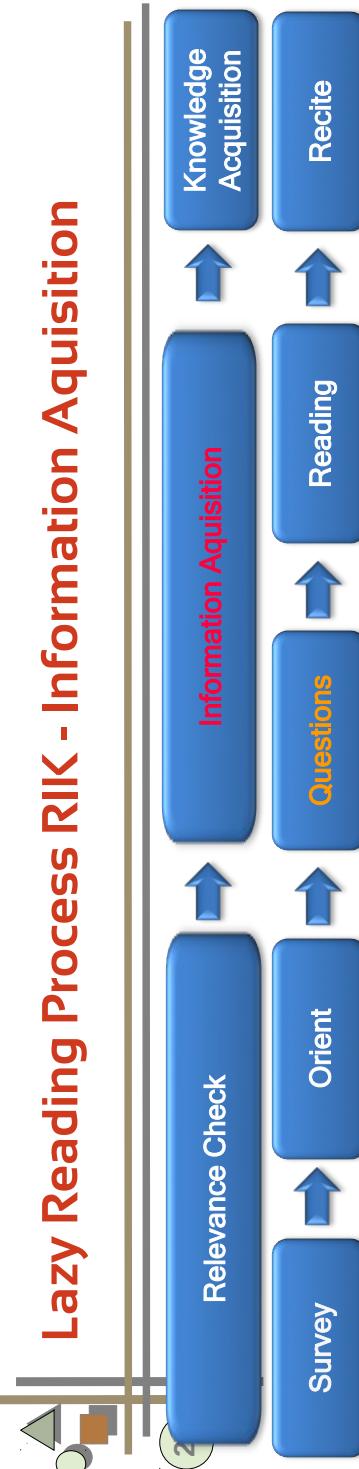
- Overview about preexisting knowledge
  - What do I know already about the subject? Important other related papers?
- Formulate questions **before** reading
  - Use the 7W questions to find valid good questions

Lazy Reading Process RIK

Relevance Check	Information Aquisition	Knowledge Acquisition
Survey → Orient	Questions → Reading	Recite
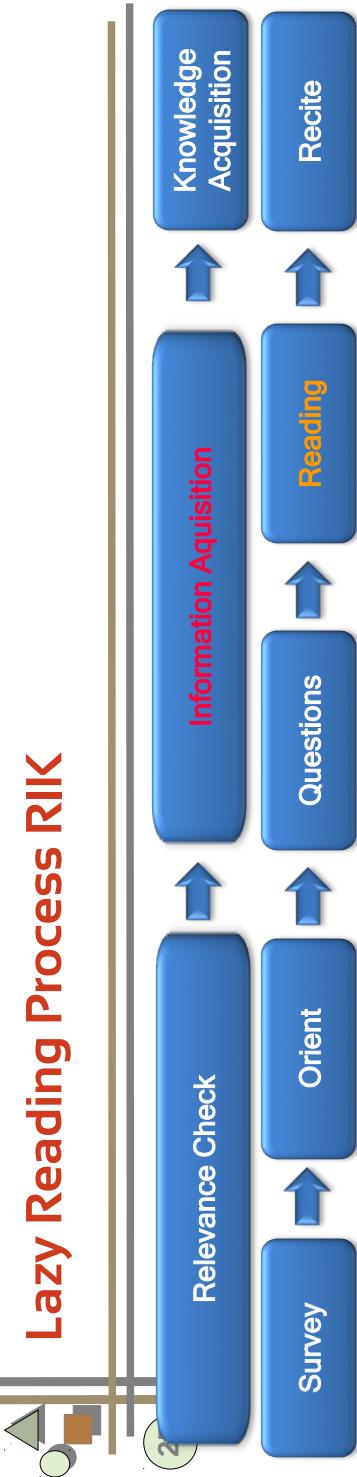
- Work yourself systematically through the text
  - Focus on most important sections
- Mark up central terms and paragraphs
  - Find out the **main point (main thesis)** of the paper
  - Mark it up, excerpt it: memory aid for later
  - Relate (by arrows) different important sections and topics
- Formulate questions **while** reading
  - Note the questions on first page of the paper
- Record your ideas
  - Remarks, critical comments, ideas into the bibtex-entry or citation database
  - Write the central main point on top of the paper

13.2.3. Knowledge Aquisition

Lazy Reading Process RIK

Survey → **Relevance Check** → **Information Aquisition** → **Knowledge Aquisition**

Orient → **Questions** → **Reading** → **Recite**
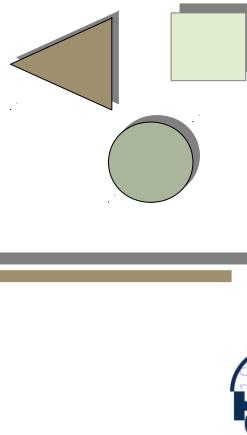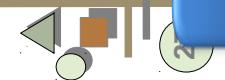
- **You must embed the new information into your (old) knowledge**
 – Look at your old summaries, record cards, mindmaps – how to change them?
- Text summary
  – **Pose and answer questions**
  – rephrase the main results
  – What is the main thesis of the section?
  – What is the skeleton of the paper?
- Work out per section
  – Mindmapping per section
  – Structure tree per section
  – Record cards (Karteikarten)
- Relation to own previous work
  – What extends your knowledge?
  – What contradicts your knowledge?
  – What is interesting?

27

Reading – the Lazy Process RIK (Rpt.)

29

RIK Method

1	**Relevance Check** –**Survey**: Structure (table of contents, paratext) **Orientation/Preview**: Abstract, Intro, Conclusion Selective Reading
2	Questions
3	Reading – headlines, main theses, bold text parts, definitions, graphics
	Recite - mind mapping - summary writing

13.3. Other Reading Methods

30
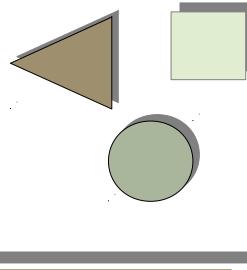
Before-Reading and After-Reading Questions

Before-Reading

- What do I know already? (previous knowledge)
- What would I like to know?
- What do I know about the author?
- What is my goal?
- Apply the 6+1 honest serving men

After-Reading

Impression:

- [PMI-Method of de Bono]
- What was positive (P)?
- What was negative, minus? (M)
- What disappointed me?
- What surprised me? (I)

Content:

- What was the main thesis?
- Supporting points?

31
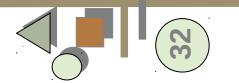
|	SQ3R Method	PQ4R Method	S2QAR Method
	Robinson 1961	Thomas & Robinson 1972	Smith 1977
	SQ3R is reflection-oriented. Uses mindmaps and summaries in the „recite phase"	improved SQ3R (could be called SQ4R) with additional phase „reflect"	oriented towards active answering of questions
	Survey, Questions, Read, Recite and Review	**Preview, Questions, Read, Reflect, Recite and Review**	**Survey, Summary, Questions, Answer and Review**

Comparison of Reading Methods

		SQ3R-Methode	PQ4R-Methode	S2QAR Methode (Smith)
1 R	a	Survey/Relevanzprüfung Preview/Vorprüfung/ •Paratext analysis (table of contents, Klappentext, summary) •Abstract, Introduction and Conclusion •Selective inspection of chapters	Preview/Vorprüfung/ Survey/Relevanzprüfung •Paratext analysis (table of contents, Klappentext, summary) •Abstract, Introduction and Conclusion •Selective inspection of chapters	Survey/Relevanzprüfung – do this as fast as possible
	b			Summary/Zusammenfassung – what is the content of the text? What is the point of the text? the controlling idea?
2 I	c	Questions/Fragen – Before-Reading Questions.	Questions/Fragen – Before-Reading Questions.	Questions/Fragen – Before-Reading Questions. Do not look into text
	d	Reading/Lesen – Answer questions - find main points	Reading/Lesen Answer questions - find main points	no not read!
3 K	e		Reflect/Nachdenken – Beispiele, Text auf vorhandenes Wissen über das Beschriebene beziehen	Answer/Antworten •– Do not look into text
	f	Recite/Rekapitulieren – Mindmaps per section - bullet minutes, summary writing – Explain text Answer questions	Recite/Rekapitulieren •– Mindmaps per section •- bullet minutes, summary writing •– Explain text •Answer questions	
	g	Review/Repetieren – Look at recitings again –Answer questions again	Review/Repetieren •– Look at recitings again •–Answer questions again	Reading/Überprüfung der Antworten •– Read text and answer questions again

13.4. Methods to Recite for Knowledge Aquisition

and the Sustainability of Reading
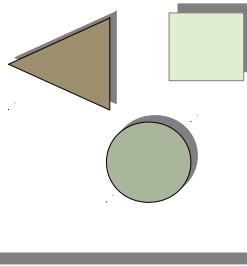
How to Recite and Summarize

- Objectives of Recite phase:
  - ▪ Personalize the information (-->data to knowledge)
  - ▪ Learn actively by reformulation
  - ▪ Abstract from unnecessary details
  - ▪ Easy way to find focused information again
  - ▲
- Methods for summary
  - 1) Underlining
  - 2) Margin notes
  - 3) Excerpting
  - 4) Mindmapping
  - 5) Structure Trees
  - 6) Cracking Sentences
  - ▲

13.4.1. Underlining

▲ Goals:
- Use a personal color scheme
- Underline for later re-reading
- Underline for comprehension

▲ What:
- Underline *main theses* of text (.. skeleton..)
- Underline *research results*
- Underline surprising things

▲ Good to read the text passage first, then underline

13.4.2. Margin Notes

▲ Put your own outline on the margin of the paper you read
▲ A. Topic outline
- Erschließung der inhaltlichen Struktur durch Randbegriffe
- Orientierung anhand von Absätzen

▲ B. Logic (argumentative) outline
- thesis statements, skeletons
- Topic changes
- Coherent sequences of paragraphs
- Summaries

13.4.3. Excerpting

▲ Bottum-up process

▲ Excerpting can have a *specific* or *global* question in mind

▲ Step 1: Orientation
  – Overview to understand the structure of the paper

▲ Step 2: Excerpting all paragraphs
  – What are the topics? theses?

▲ Step 3: Excerpt all sections
  – Do summaries for sections

▲ Step 4: Excerpt a summary sentence for whole text

Use Mindmapping, Structure Trees, …

13.4.6. Refactoring Sentences for Comprehension: Subclause Reordering

▲ A **uphill sentence** starts with a subclause followed by the main clause:

subclause

main clause

„On that day when it was dark and the sun was hidden behind the clouds, John Silver left his home forever."

▲ If the main clause comes too late, it is hard to understand

▲ Change sub and main clause

„John Silver left his home forever, on that day when it was dark and the sun was hidden behind the clouds. "

Refactoring Sentences: Verbalisation

▲ Verbalisierung a sentence helps to understand it

- Bsp. „ Dass ein Leser Wort- und Satzsinn versteht, kann in Aussagen über den Text zum Ausdruck kommen.“

- Zum Ausdruck kommen => ausdrücken

  „ Dass ein Leser Wort- und Satzsinn versteht, kann sich in Aussagen über den Text ausdrücken.“

13.5 Information Gathering
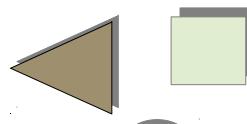
Search Machines

- Google scholar http://scholar.google.de
- bib search engine http://liinwww.ira.uka.de/bibliography/index.html
- DBLP at Trier University
  – http://www.informatik.uni-trier.de/~ley/db/index.html
- Springer LNCS
- ACM Digital Library www.acm.org/dl
  – ACM Journals
  – ACM Conferences
- IEEE explore
- Research Gate

Who tried what already? Experiences?

Information Gathering (Recherche)

- Most often, literature is found today on the internet.
  – Google scholar
  – Research gate
  – bib-Server in Karlsruhe
  – dblp search engine
- Use the SLUB license to find papers with Springer, ACM, IEEE.
- For non-licensed papers, use the SLUB search engine
  – http://www.slub-dresden.de/
  – Go and lend a paper copy
- Saxony stores most of its Master's thesis and PhD theses on "Quality Content of Saxony", our permanent pdf server
  – http://www.qucosa.de/
  – Here you can find most of the Master's theses of the chair of Software Engineering.

What to Cite

▲ Distinguish primary from secondary sources
- Read and cite primary sources!
- If you found a good pedagogic overview article interesting for others, too, you may also cite this secondary source

▲ Important journals in Software Engineering
- ACM Transactions on Software Engineering and Methodology (TOSEM)
- ACM Transactions on Programming Languages and System (TOPLAS)
- IEEE Software
- Springer Software and Systems Journal (SoSym)

▲ Overview journals or bibliographies for certain topics
- ACM Computing Surveys

Homework: Look for interesting Journals and Conferences in Software Engineering

▲ Surf ACM, IEEE, Elsevier, Wiley, Springer

Screening the World

▲ Screening analyzes trends to find interesting subjects for research.

▲ Internal screening group: everybody presents something cool

  – Motivation: Discovering disruptive technologies early is very important

  – Track important web sites

  – Collect new ideas for demonstrators

  – Collect interesting labcast videos and web sites on an inspiration site

▲ Research Blog

▲ Paper reading group

Fix Interesting Web Sites for News

▲ Finding the newest news is very important for research.

▲ Interesting addresses:

  – Www.mozillalabs.com

  – Google labs

  – MIT media labs

  – Microsoft research

  – Berkeley CPS lab

  – Golem.de

  – Technology Review

  – Heise Newsticker

The End - Homework

Many slides are courtesy to Birgit Grammel, Dr. Birgit Demuth

---
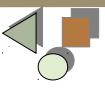
Homework – The Weekly Schmidt

▲ Write an outline of an essay - "Why the Euro should not be given up".
  – Background: European debt crisis, Greeks almost bankrupt etc.

▲ Read first Helmut Schmidt's essay "Sechs Gründe, warum der Euro nicht scheitern darf" from "Einmischungen", S. 169. Apply PQ4R.
  – Write down questions
  – Summarize the article with a mind-map or structure tree
  – Recite it loud to your friend

▲ Try to find one other article on the web sites of FAZ, Süddeutsche, Zeit or similar, on the subject, and select some of your arguments from the material. Read with PQ4R.

▲ For the outline, use your recited material.

▲ After you have written a clear argumentative outline, write an introduction and a conclusion.

Homework – The Weekly Churchill

- Write an outline of an essay - "Why Germany should belong to Europe".
- Read first Winston Churchill's speech "Council of Europe". Apply PQ4R.
  - https://www.winstonchurchill.org/learn/speeches/speeches-of-winston-churchill/1946-1963-elder-statesman/lll-the-council-of-europe
    - Write down questions
    - Summarize the article with a mind-map or structure tree
    - Recite it loud to your friend
- Try to find one other article on the web sites of FAZ, Süddeutsche, Zeit or similar, on the subject, and select some of your arguments from the material. Read with PQ4R.
- For the outline, use your recited material.
- After you have written a clear argumentative outline, write an introduction and a conclusion.