

13. Basics of Information Gathering (Reading, Screening, Strategy, Literature Analysis)

Lecturer: Dr. Sebastian Götz

Prof. Dr. Uwe Aßmann

Softwaretechnologie

Fakultät Informatik

Technische Universität Dresden

18.04.2017

<http://st.inf.tu-dresden.de/teaching/acse>

- 1) Data, Information, Knowledge
- 2) Reading Process RIK
 - 1) Checking Relevance of Texts
 - 2) Information Acquisition
 - 3) Knowledge Acquisition
- 3) Other Reading Methods
- 4) Methods of Recite
- 5) Information Gathering
- 6) Writing literature analysis papers



DRESDEN
concept
Exzellenz aus
Wissenschaft
und Kultur

Obligatory Literature

- ▶ Philip W.L. Fong. 2009. Reading a computer science research paper. SIGCSE Bull. 41, 2 (June 2009), 138-140. DOI=10.1145/1595453.1595493
<http://doi.acm.org/10.1145/1595453.1595493>
- ▶ Joseph D. Novak, Alberto J. Cañas. The Theory Underlying Concept Maps and How to Construct and Use Them. Technical Report. CmapTools 2006-01 Rev 01-2008, Florida Institute for Human and Machine Cognition (IHMC)
<http://cmap.ihmc.us/docs/theory-of-concept-maps>
- ▶ William G. Griswold: How to Read an Engineering Research Paper. Online:
<http://cseweb.ucsd.edu/users/wgg/CSE210/howtoread.html>



13.1. Data, Information, Knowledge Aquisition in Science

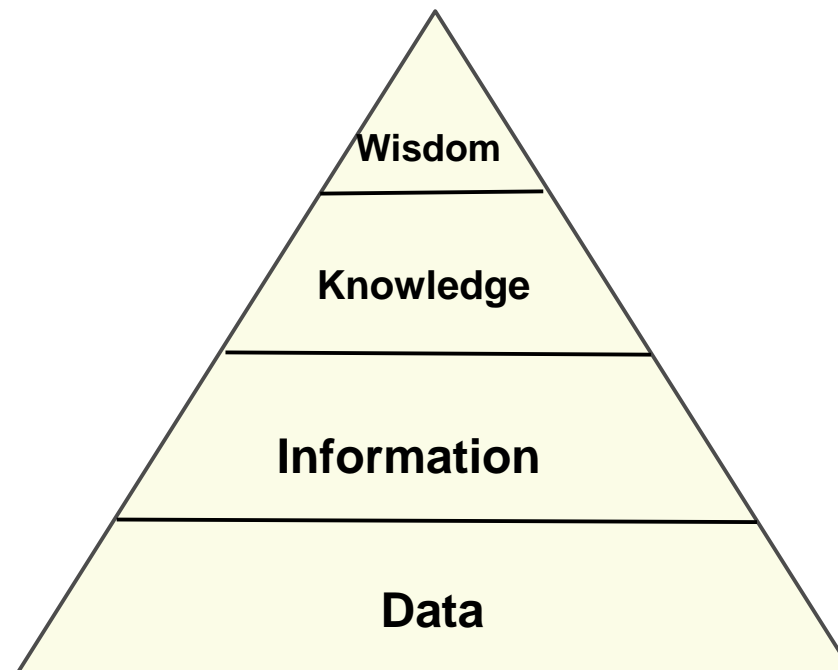


Science is about DIKW (Data, Information, Knowledge, Wisdom)

4

Academic Skills in Computer Science (ASICS)

- ▶ 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:

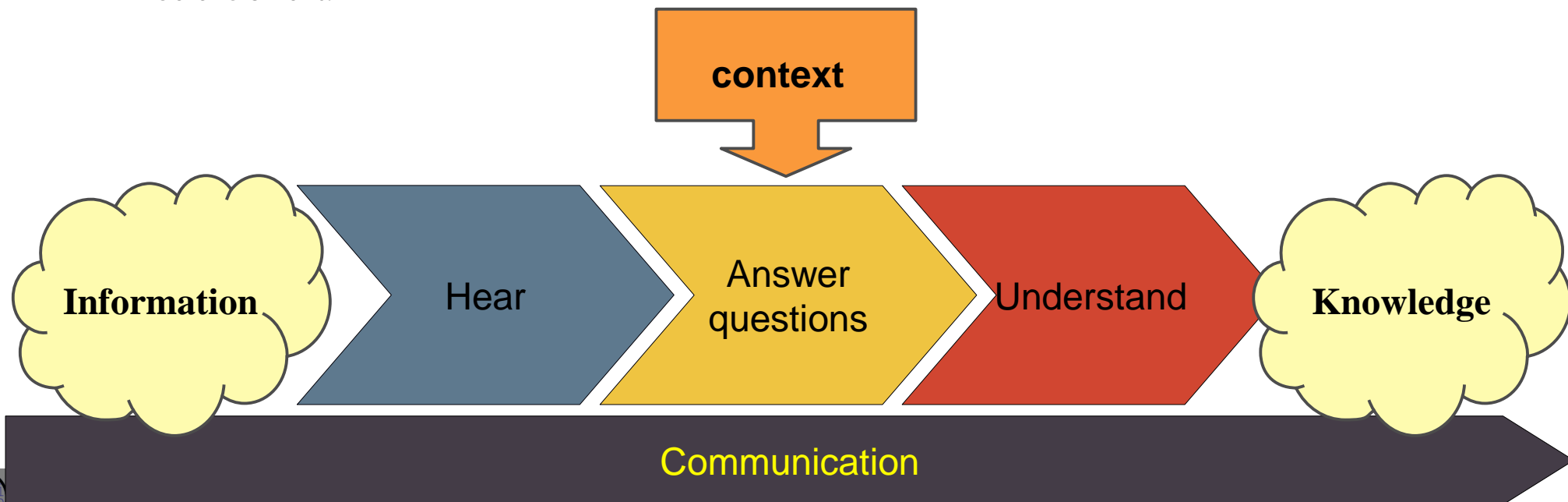


The Knowledge Aquisition Model from Spinner

5

Academic Skills in Computer Science (ASICS)

- ▶ 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:
 - “This tastes good.”
 - “This is interesting.”
 - “You idiot.”
 - “You are smart.”



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, superficial type, depressive,...)
- ▶ which channel has the sender used previously (facts, emotions, relations, etc.)?

About the receiver:

- ▶ What 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

7

Academic Skills in Computer Science (ASICS)

- ▶ How do you interpret the remarks
 - “This tastes good.”
 - “This is interesting.”
 - “You idiot.”
 - “You are smart.”
- ▶ from your partner? from your friend? from your mother?
- ▶ from your competitor?
- ▶ from your boss?

Knowledge is what remains after answering questions.



13.2. “Lazy” and Efficient Reading Process with Relevance Check, Information and Knowledge Aquisition (RIK)



Problems with Reading

9

Academic Skills in Computer Science (ASICS)

Motivation

- ▶ Bored, unfocusedness, tired

Good Steps

- ▶ Increase your width of focus
- ▶ Try to read fast
- ▶ Read slower, if text is hard to understand

Bad habits

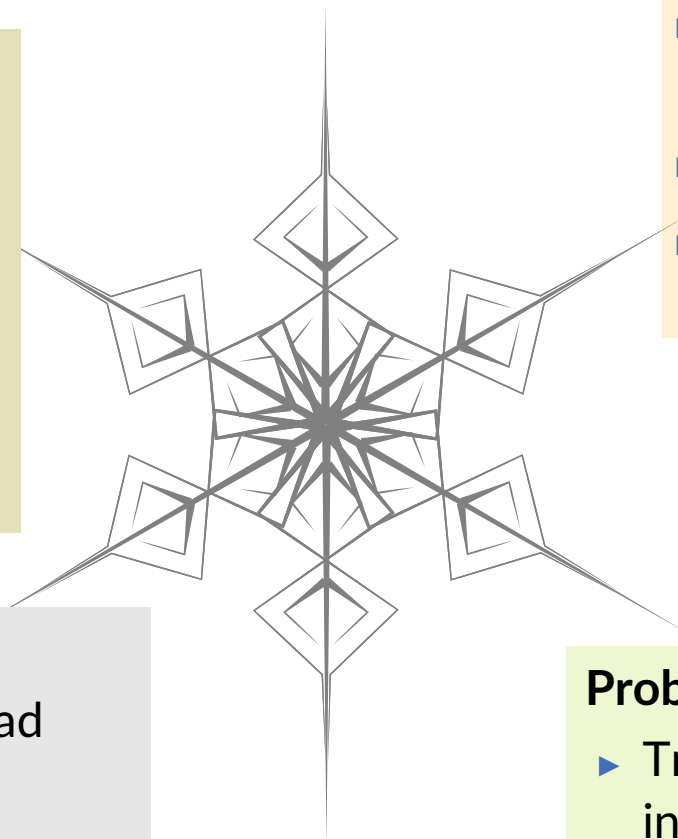
- ▶ Read word by word
- ▶ Jump back (regression)
- ▶ Talk while reading
- ▶ Listening to distracting music while reading

Good steps

- ▶ Cover read lines or unread by paper sheet or ruler
- ▶ Jump-Stop movement
- ▶ Do Summaries
- ▶ Write Questions

Problems with Pertinence

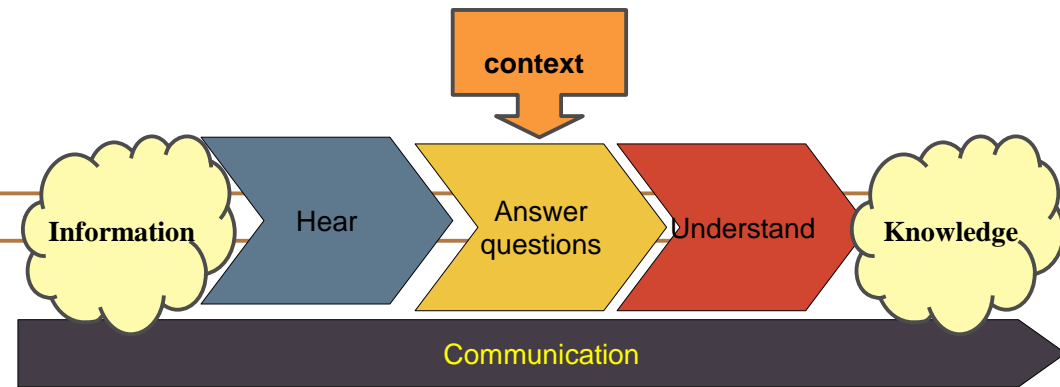
- ▶ Try to read everything, instead of finding the thesis statement of the paragraph
- ▶ Only read, never draw



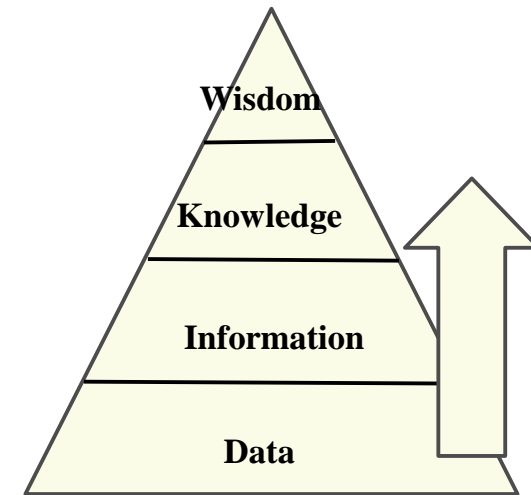
Lazy Reading Process RIK

11

Academic Skills in Computer Science (ASICS)



- ▶ RIK is a simple reading process allowing for stopping all the way and not wasting time
- ▶ The RIK process is structured along the DIKW pyramid and the Spinner IK acquisition process:
 - without questions no knowledge from information
 - without recitation no knowledge: no embedding of the information in your own knowledge (self-context)



Lazy Reading Process RIK (Survey)

12

Academic Skills in Computer Science (ASICS)



- ▶ **Analyse Paratext:** Table of Contents
 - Find out focus of work
 - Separate background from the author's work
 - Skip analysis: Find out chapters you know already and decide to skip them
- ▶ **Paging** through
 - How long are main chapters?
 - What is side material?

- ▶ **Analyze Eye Catchers**
 - Figures, tables
 - Central definitions
 - Other structuring aids

Lazy Reading Process RIK (Orient)

13

Academic Skills in Computer Science (ASICS)



- ▶ Analyze the **abstract** and the **introduction**
 - Research question? Research Hypothesis? Research Method? Research Validation?
 - Relevance, Positioning into the research landscape
 - Find out Assumptions of the paper
 - Find out Restrictions (Limits) of the paper
- ▶ 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



Use *Paratext* to Check for Relevance

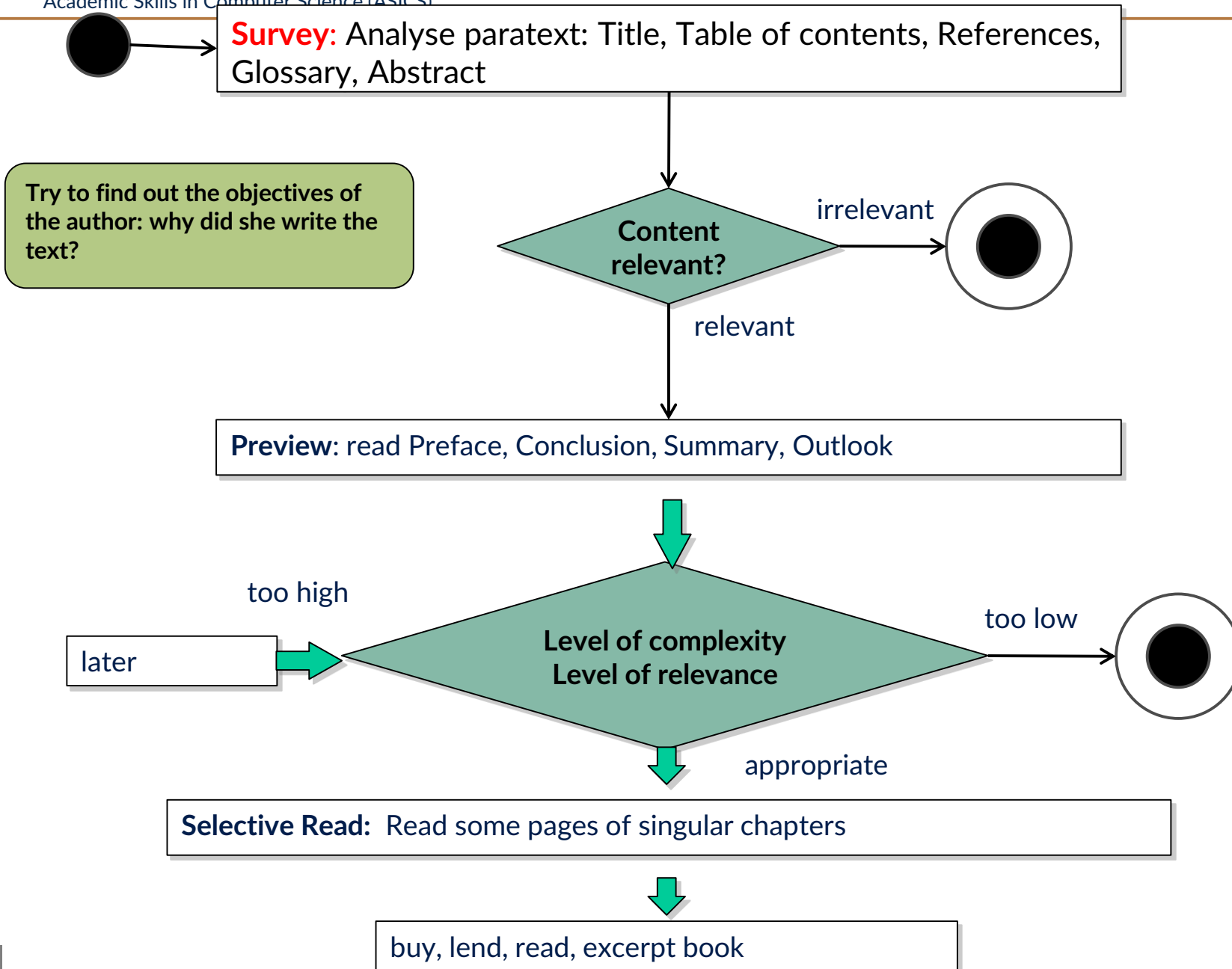
- ▶ Find out goal of the author and the relevance of research topic for you
- ▶ **Main text:** Title, preface, introduction, table of contents, summary, ...
- ▶ **Secondary text:** Envelope, Recensions, Amazon comments, ...
- ▶ **Computer reviews:** a journal with reviews of papers
- ▶ **Abstract:** read the abstract and analyze it
- ▶ **Search on the internet** about the paper or book
 - summaries, reviews



Relevance Check: Survey, Preview, and Selective Reading of Books

16

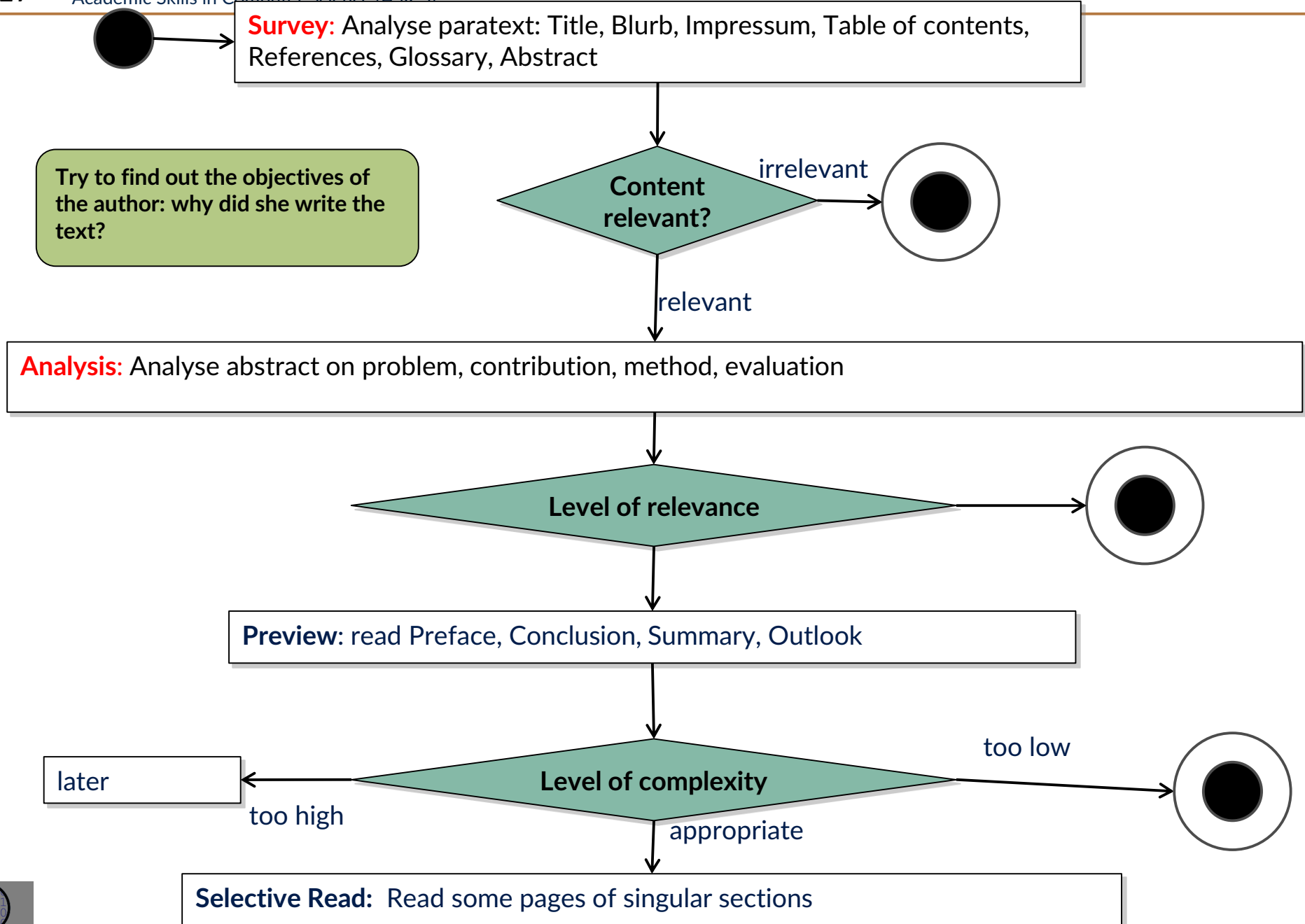
Academic Skills in Computer Science (ASICS)



Relevance Check Specific for Research Papers: Survey, Fong Analysis, Preview, and Selective Reading

17

Academic Skills in Computer Science (ASICS)



Abstract Analysis

18

Academic Skills in Computer Science (ASICS)

Analyze the abstract (and intro, conclusion) on

- ▶ Research Problem
- ▶ Research Contribution
- ▶ Research Method and Evaluation
 - Scientific paper: Proofs? Experiments? Empirical studies?
 - Essay: Opinions and arguments?
- ▶ Conclusions

- ▶ Problem and contribution together are often also called:
 - Research question – if formulated as a question
 - Research hypothesis – if formulated as hypothesis



13.2.1.2 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.



Different Classes of Research Results

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

- ▶ **Relevant research:** Somebody, the **research stakeholder**, needs the result.
 - **Significant problem?**
 - **Significant 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



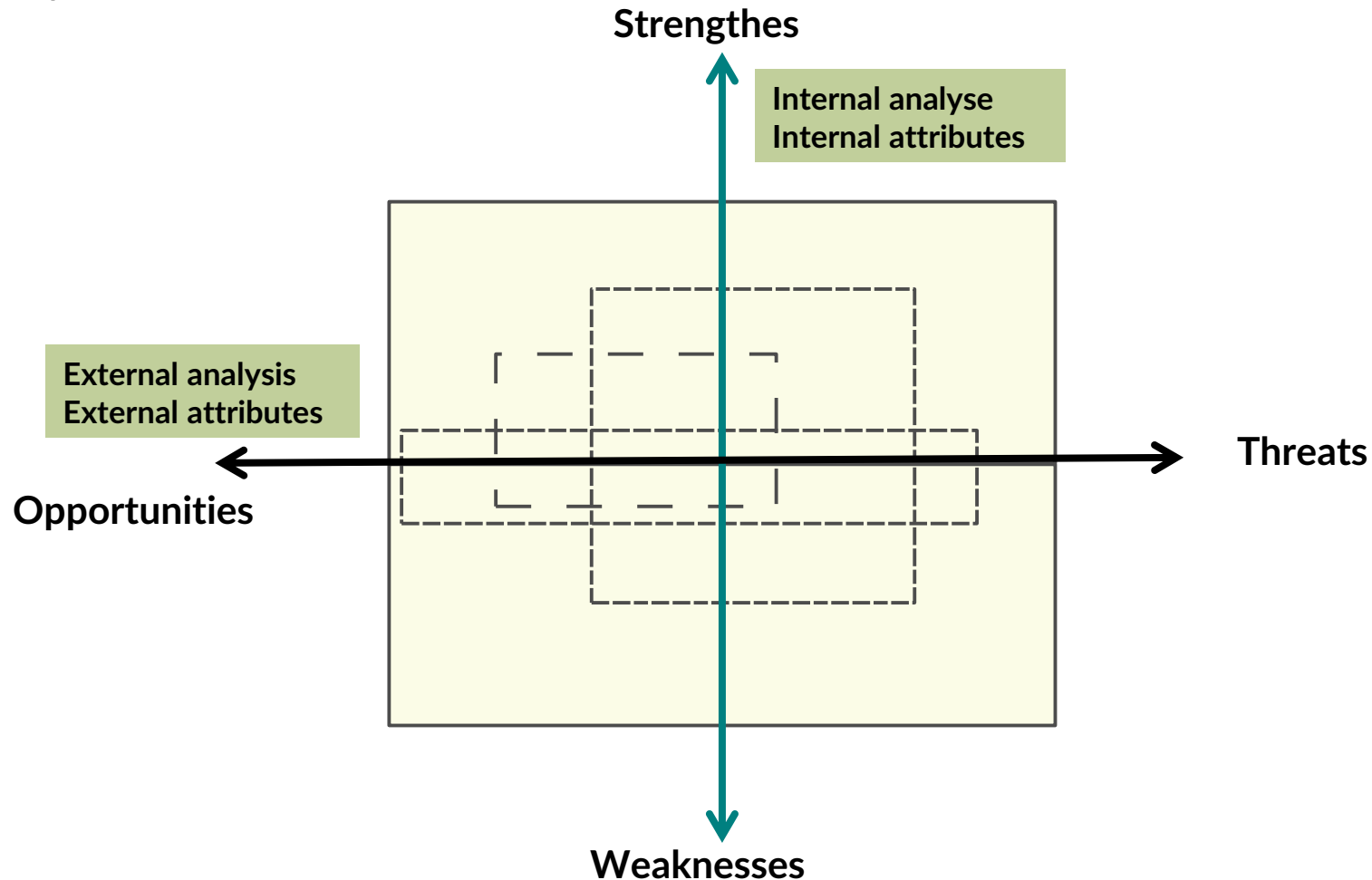
What are the Strategic Aspects of a Paper?

SWOT Analysis as a 4-D Analysis

21

Academic Skills in Computer Science (ASICS)

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

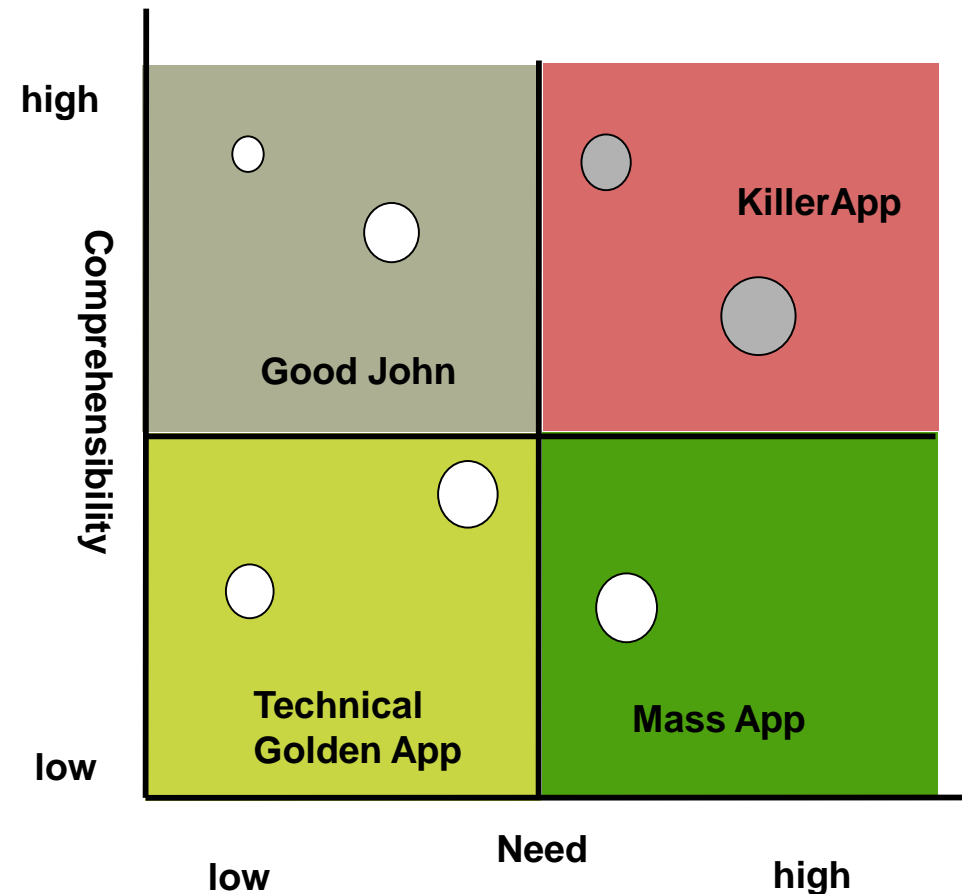


„KillerApp“ Analysis (Attractivity Portfolio)

22

Academic Skills in Computer Science (ASICS)

- ▶ **“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:
 - $\text{Attractivity} = \text{Need} * \text{Comprehensibility}$
- ▶ Most attractive papers or projects are “KillerApps”, because they are easy to comprehend and useful for many



13.2.2. Information Aquisition



Lazy Reading Process RIK - Information Aquisition

24

Academic Skills in Computer Science (ASICS)



- ▶ 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

25

Academic Skills in Computer Science (ASICS)



- ▶ 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
- ▶ **Structure** your ideas by a *semantic net* or *concept map*
 - finding out central concepts and their relationships

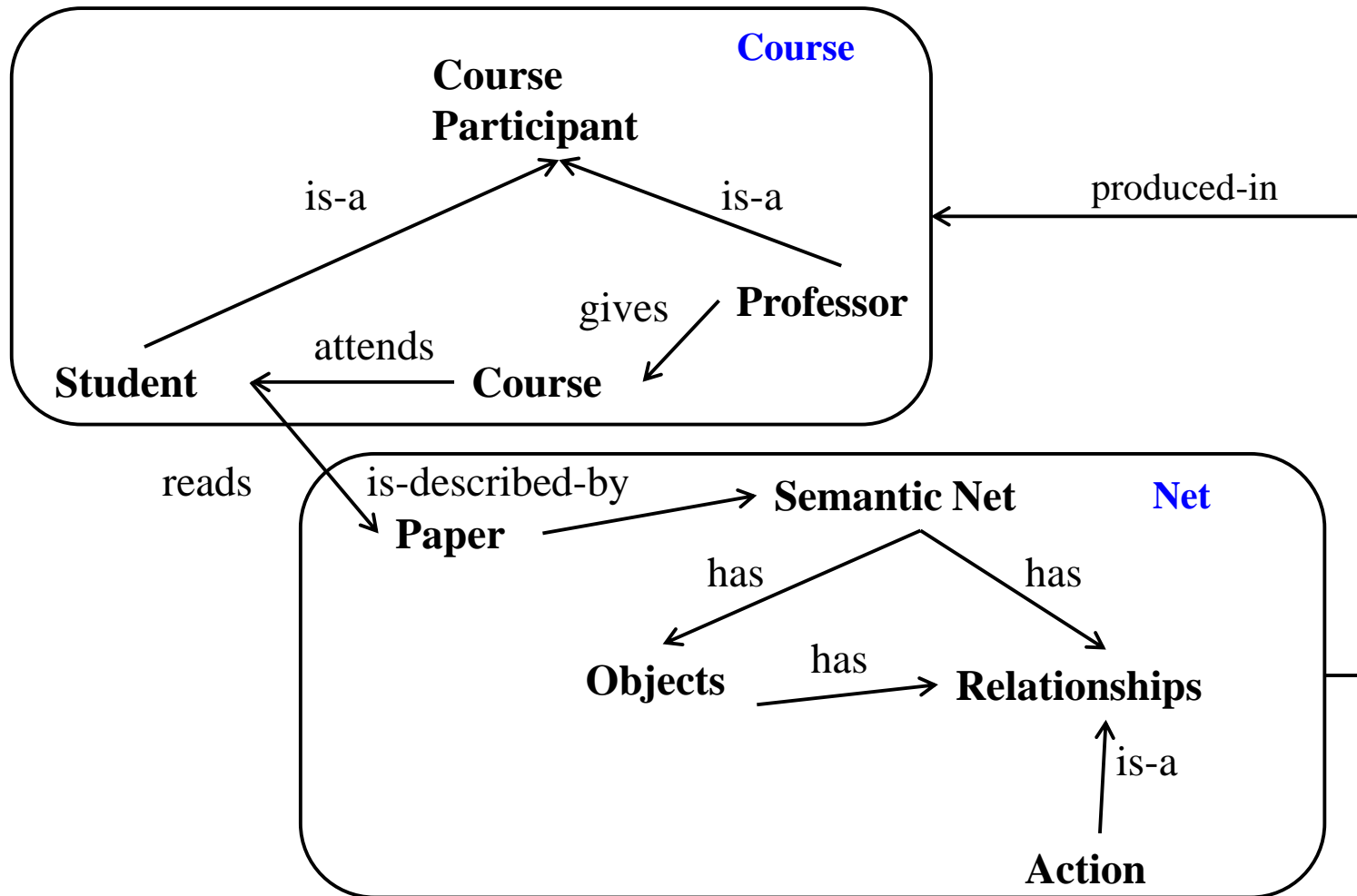


Grouping in Semantic Nets and Concept Maps

26

Academic Skills in Computer Science (ASICS)

- ▶ Remember important operations to create **knowledge** from information:
- ▶ Grouping, Hierarchizing, Re-drawing, Dualizing

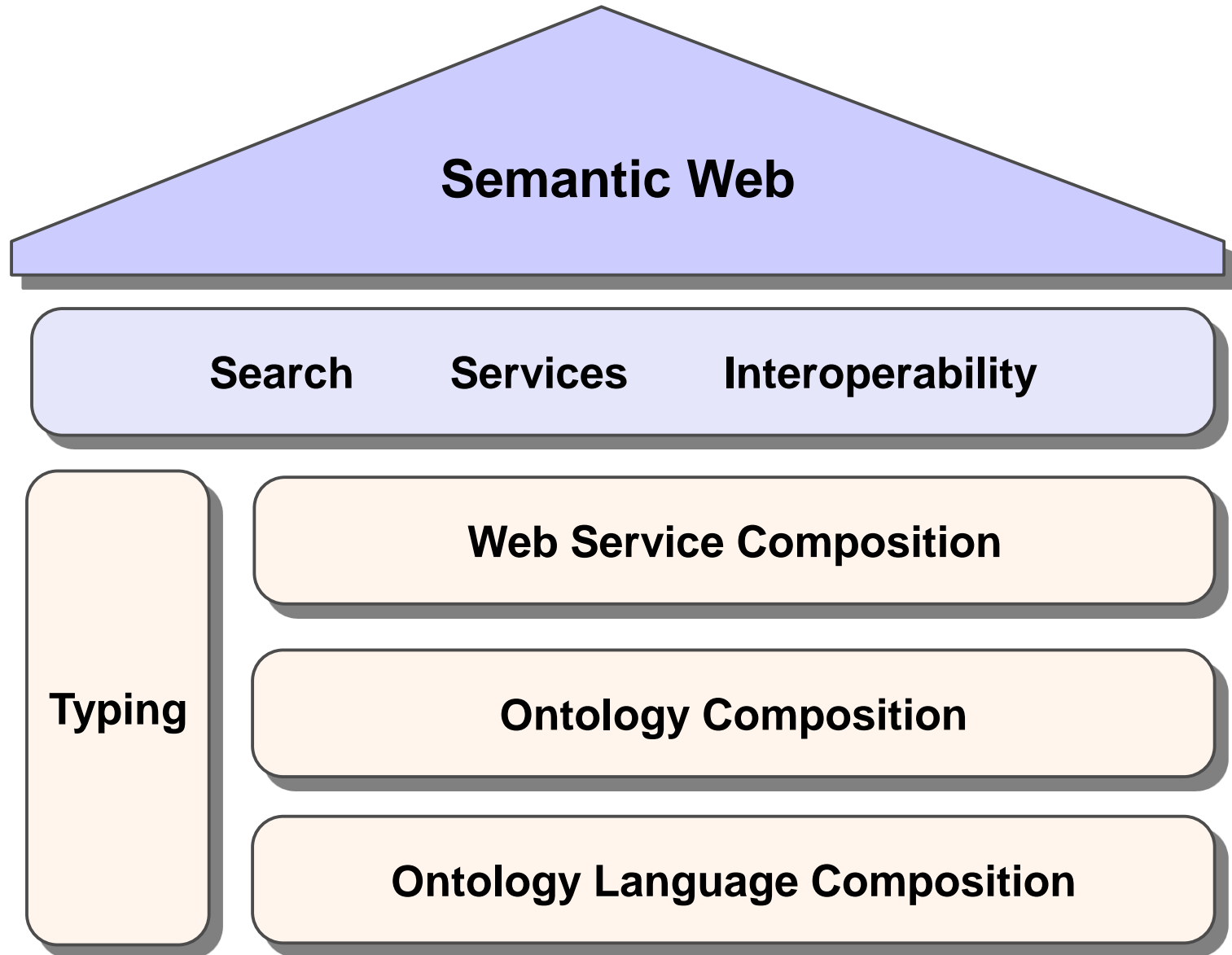


http://en.wikipedia.org/wiki/Semantic_network

<http://de.wikipedia.org/wiki/Concept-Map>



House Concept Map

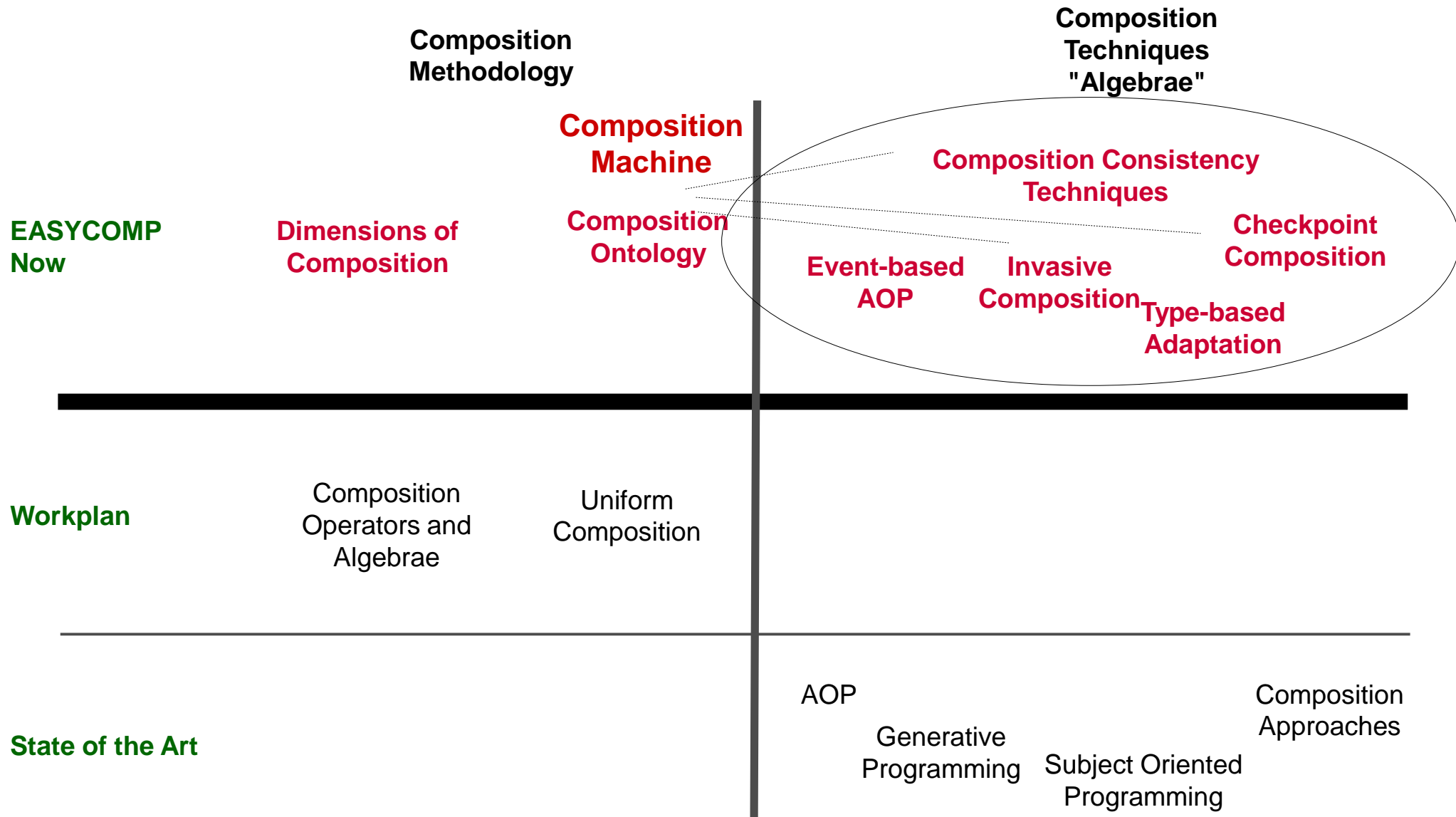


Abstract House Concept Map - Showing Own and Foreign Research Works

- ▶ An **Advance Map** is a concept map showing the advance of an approach, the gap to the state of the art, by visualizing 2 or 3 phases or layers
 - Comparing the advantage of approach B over approach A and C
- ▶ Clearly distinguished
 - Own and foreign research
 - State of the art and research agenda
 - Yesterday, today and tomorrow
- ▶ Advance maps are very useful for research papers and research proposals.



Overview of Progress of a Project (EASYCOMP)



Progress 2nd Year of Project (EASYCOMP)

31

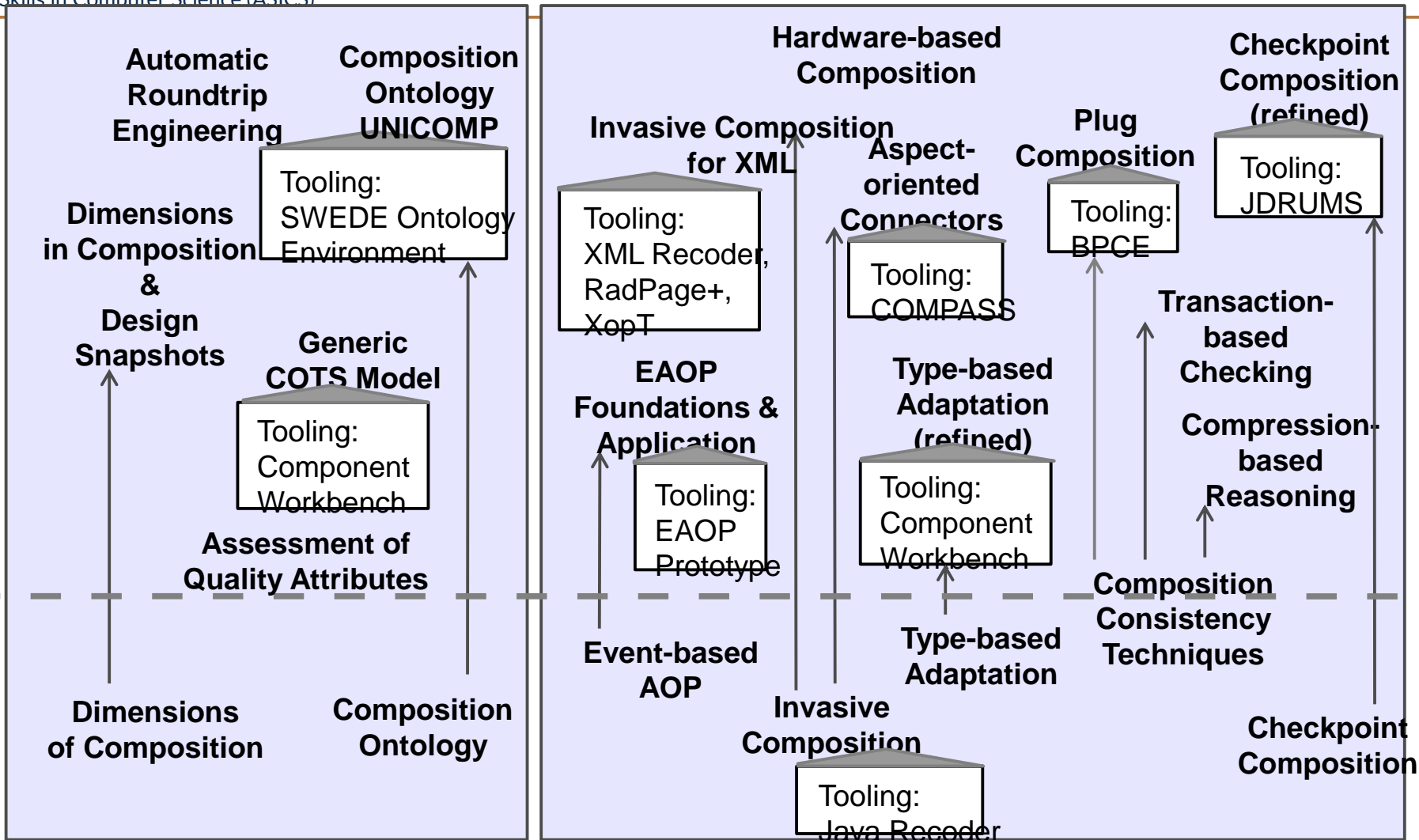
Academic Skills in Computer Science (ASICS)

Composition Methodology

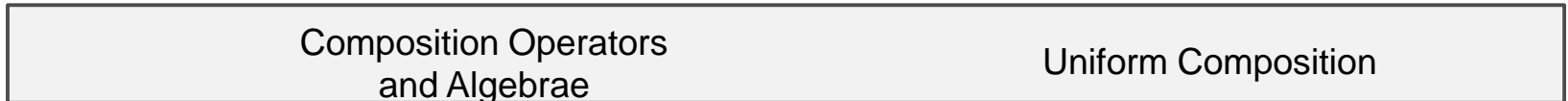
Composition Techniques

Inventions and Refinements in EASYCOMP Year 2:

Inventions in EASYCOMP Year 1 (proceeded in Year 2):



Workplan:



State of the Art:



13.2.3. Knowledge Aquisition



Lazy Reading Process RIK

33

Academic Skills in Computer Science (ASICS)



- ▶ **You must embed the new information into your (old) knowledge**
 - Look at your old summaries, record cards, semantic nets, mindmaps, concept maps – how to change them?
- ▶ **Write a text summary: Pose and answer questions**
 - What is the main thesis?
 - Rephrase the main results
 - What is the skeleton of the paper?
- ▶ **Talk about the paper to somebody else (your mate, your wife, your colleagues...)**
- ▶ **Repeat Information Acquisition in details, per section**
 - Structure tree per section, Record cards
 - Mindmap, semantic net, concept ma per section
- ▶ **Relation to own previous work**
 - What extends your knowledge? What contradicts your knowledge?
 - What is interesting?



Reading – the Lazy Process RIK (Rpt.)

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



13.3. Other Reading Methods



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?



Other Reading Methods

SQ3R Method	PQ4R Method	S2QAR Method
Robinson 1961	Thomas & Robinson 1972	Smith 1977
SQ3R is reflection-oriented. Uses mindmaps, concept maps 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



13.4. Methods to Recite for Knowledge Acquisition

and the Sustainability of Reading



DRESDEN
concept
Exzellenz aus
Wissenschaft
und Kultur

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
 - Content structure
 - Orientation based on paragraphs
- ▶ B. Logic (argumentative) outline
 - thesis statements, skeletons
 - Topic changes
 - Coherent sequences of paragraphs
 - Summaries



13.4.3. Excerpting

- ▶ Bottom-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, ...



13.5 Information Gathering



Search Machines

45

Academic Skills in Computer Science (ASICS)

- ▶ Google scholar <http://scholar.google.de>
- ▶ bib search engine <http://iinwww.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?



- ▶ <http://ercim-news.ercim.eu/>
- ▶ Augmented Reality (AR)
- ▶ <http://ercim-news.ercim.eu/en103/special/augmented-reality-introduction-to-the-special-theme>
- ▶ About the IoT
 - <http://ercim-news.ercim.eu/en101?view=featured>
 - <http://ercim-news.ercim.eu/en101/special/compose-an-open-source-cloud-based-scalable-iot-services-platform>
 - <http://ercim-news.ercim.eu/en101/special/3d-web-visualization-for-real-time-maintenance-of-smart-buildings>
- ▶ Fraunhofer Magazin:
 - <https://www.fraunhofer.de/de/publikationen/fraunhofer-magazin/archiv.html>



(Collaborative) Literature Management and Search Tools

47

Academic Skills in Computer Science (ASICS)

- ▶ Saving Bibliographic Meta-Data
 - From structured sources (BibTex)
 - From webpages (collaborative parsers for important sites: CiteSeer, GoogleScholar, ACM, SLUB Dresden ...)
- ▶ Organizing your References
 - Tags, Folders
- ▶ Sharing References with others
- ▶ Adding Notes
- ▶ Exporting BibTex (e.g. for LaTeX)

- ▶ Examples...

zotero

 RefWorks

 MENDELEY



- ▶ Free to use (but limited free storage)
- ▶ Firefox Add-On/Standalone/Web-based
- ▶ OpenSource
- ▶ <http://www.zotero.org/>



Commercial Tools

49 Academic Skills in Computer Science (ASICS)



- Commercial
- Desktop application
- TU Dresden recommends it and has a license:

–<http://www.slub-dresden.de/service/schreiben-publizieren/refworks/>

- (We recommend Zotero)

•<http://www.refworks-cos.com/refworks/>



- Commercial, Closed source, but API
- Desktop/Web
- <http://www.mendeley.com/>
- Often used (good advertising)

- ▶ <http://www.dagstuhl.de/>
- ▶ Schloss Dagstuhl is the German meeting centre for computer scientists.
- ▶ It organizes very interesting seminars on diverse topics in software engineering
 - All abstracts and many papers are online
- ▶ Extremely valuable to understand the State of the Art in an area!
- ▶ List of 2014:
 - http://www.dagstuhl.de/de/programm/kalender/?dag_type=12&dag_year=2014
- ▶ Unifying Product and Software Configuration
 - <http://www.dagstuhl.de/de/programm/kalender/semhp/?semnr=14172>
- ▶ The Future of Refactoring
 - <http://www.dagstuhl.de/de/programm/kalender/semhp/?semnr=14211>
- ▶ Scientific Visualization
 - <http://www.dagstuhl.de/de/programm/kalender/semhp/?semnr=14231>



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



13.6 Writing Literature Analysis Papers

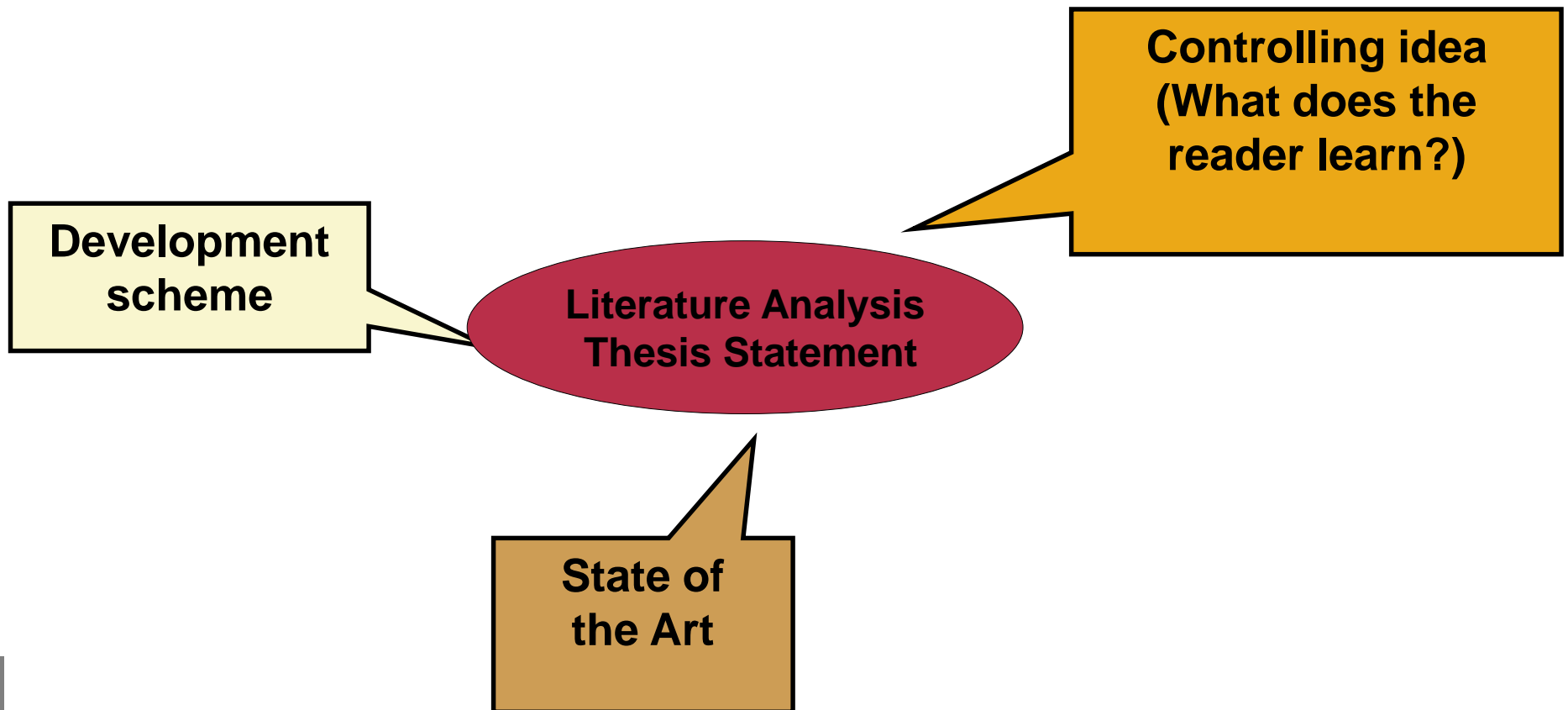
Literature analysis papers can be written:

- standalone for overview journals such as ACM computing surveys
- as Background chapter for your Master or PhD thesis

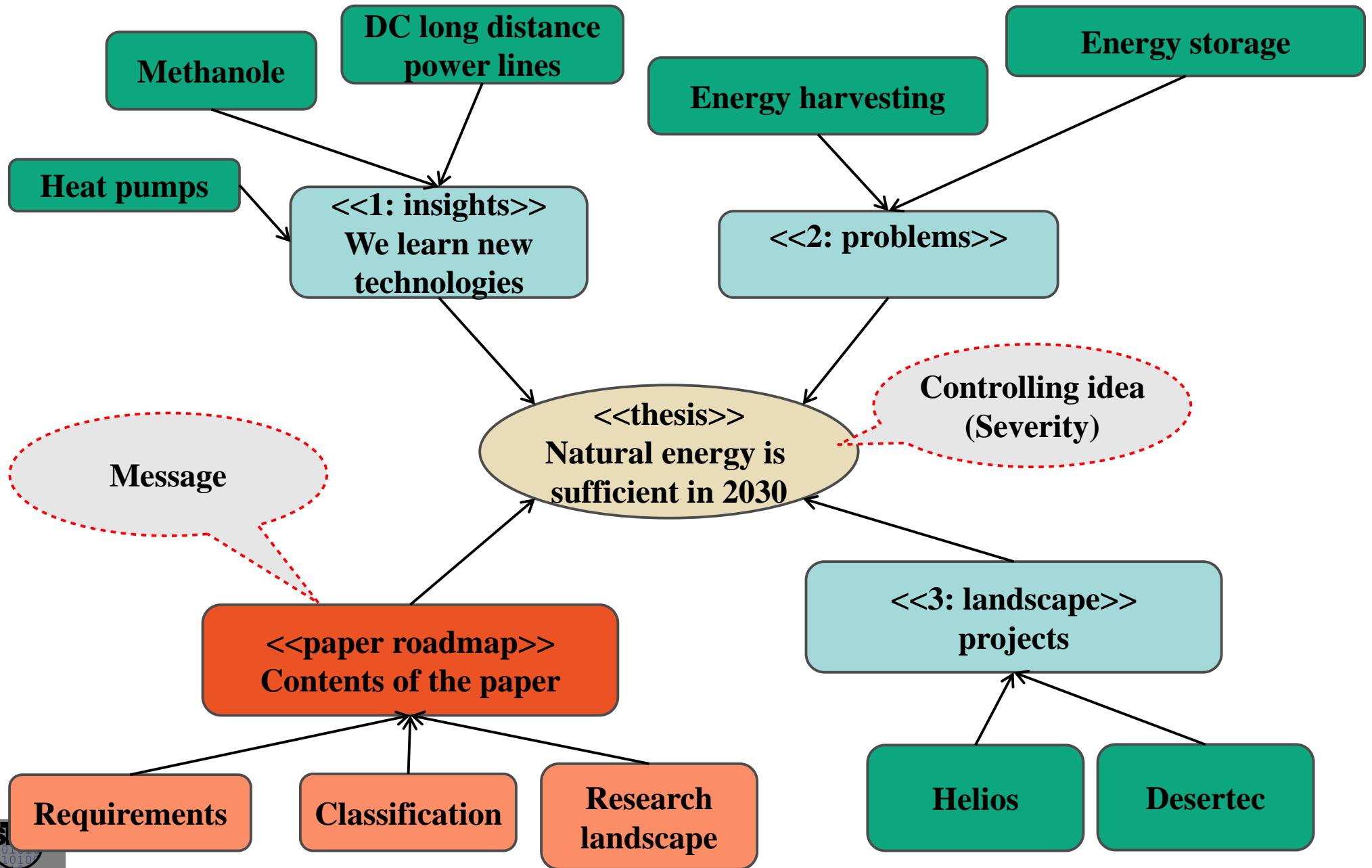


Thesis Statements about Literature Analysis

- ▶ A **literature analysis thesis statement** is a thesis statement showing the state of the art of the literature with regard to a certain area of knowledge.
- ▶ **Classification Thesis: State-of-the-Art-in-Area + Controlling idea (what does the reader learn?) + Development Scheme**

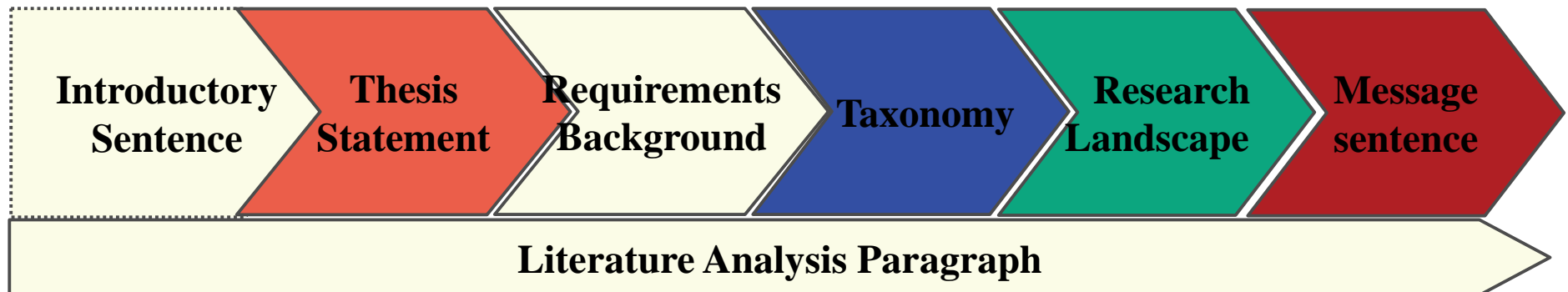


Example: Mindmap for Introductory Paragraph of Literature Analysis Paper



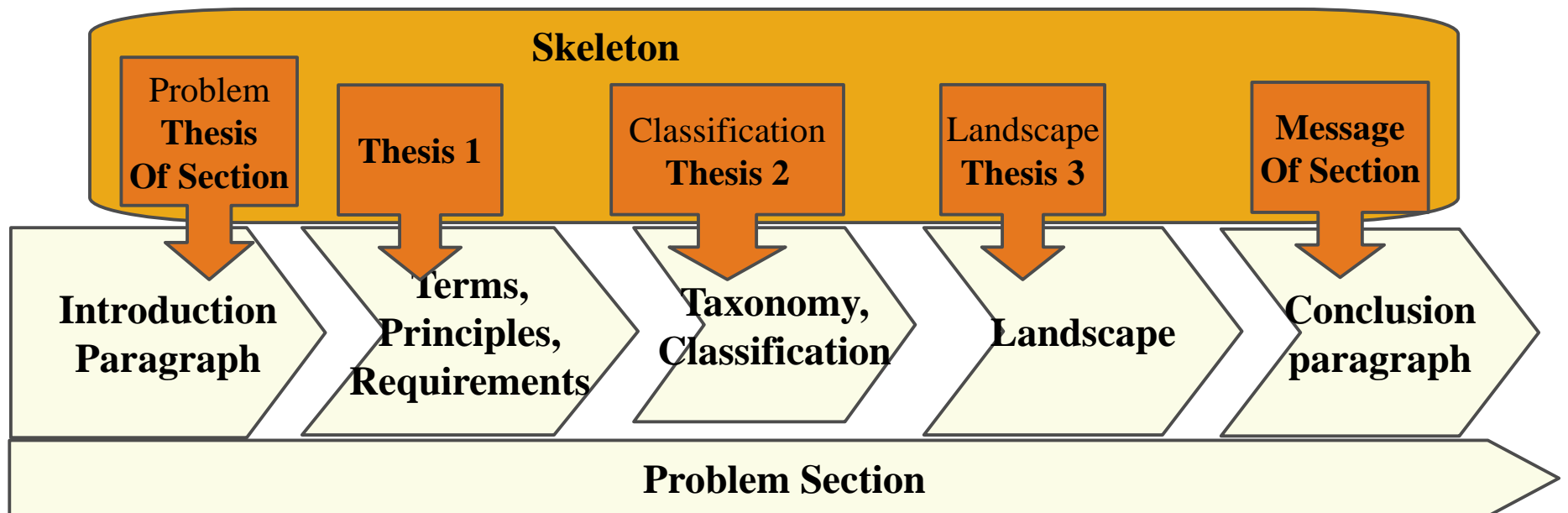
Literature Analysis Paragraphs in 1-3-1 Structure (May be as an introduction to a paper)

- ▶ We live in interesting times.
- ▶ As recent works show, technical progress in harvesting energy is so fast that we might be able to earn all our energy demands from natural sources in 2030.
- ▶ We are learning at the moment how to store wind energy in methane gas and methanole, we have a political union with countries like Greece who do not have a stable economy, and could buy their natural energy and transport it by long-distance DC-powerlines to central and northern Europe. And since recently, we can build cheap and efficient multi-stage heat pumps to even heat old houses with natural energy.
- ▶ All these techniques must solve two problems: energy harvesting and energy storage. The former techniques are usually based on solar, wind, and heat harvesting. The latter show different degrees of efficiency; storing energy with power2gas methods seems to be most profitable.
- ▶ Large projects such as Desertec (D) and Helios (GB) develop challenging agendas for changing our economy from an oil-based to one based on natural energy.
- ▶ This paper will give an overview on requirements for the natural energy economy, will classify the energy techniques, and will show the research landscape of projects in Europe. May also you know that we live in interesting times.



Skeletons of Literature Analysis Paper

- ▶ The **skeleton** of a literature analysis paper (or section) is the sequence of all thesis statements of all paragraphs. [Salehie] suggests:
 - Terms, principles, requirements
 - Taxonomy, facet classification, multi-criteria comparison
 - Landscape (projects, papers), with concept maps, portfolios and Kiviati graphs
- ▶ Additional elements: Advance map, discussion of major approaches, past-present-future
- ▶ Literature analysis sections may be positioned as section 2 or 3 of a paper



Why Do We Need Literature Analysis and Information Gathering

58

Academic Skills in Computer Science (ASICS)

- ▶ Research has to bring **novelty**, and novelty must be demarcated to the state of the art in the literature
- ▶ Know where you stand! (know your competitors)



The End

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

