



HPI Research Symposium 2021 - Master Day
Welcome &
An Intro to Successful Systems Research

Prof. Tilmann Rabl
Data Engineering Systems
Hasso-Plattner-Institut

The Hasso Plattner Institute in Potsdam, Germany



HPI Research Symposium 2021 – Welcome!

Agenda

- **Monday**

- **13:00 – 16:00 – Master Day**



We are here!

- Tuesday

- 9:15 – 12:30 – Future SOC Lab and Data Lab Day

- Wednesday

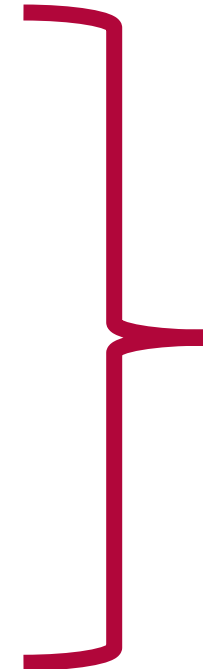
- 9:00 – 12:00 – Doctoral Symposium

- Thursday

- 15:00-18:00 - Symposium with Industry & Academia

- Friday

- 9:00-12:00 - Symposium with Industry & Academia



Also a very nice program!

HPI Research Symposium - Master Day

- Master projects
 - Mandatory semester projects
- Master theses
 - Final, researchy 6 month project
- Other projects
 - Seminar results are of publishable quality
- First steps into research
 - Some of this work is on par with PhD level research



Today's Agenda I

Time	Title & Speakers
13:00 – 13:30	Tilmann Rabl - <i>Successful Systems Research</i>
13:30 – 13:45	Julian Hugo, Spoorthi Kashyap, Nataniel Müller & Justus Zeinert (Master Project Renard): <i>Improving Network Integration Algorithms for Drug Predictions</i>
13:45 – 14:00	Finn Klessascheck, Tom Lichtenstein & Simon Siegert (Master Project Weske): <i>Process Mining in Personalized Medicine</i>
14:00 – 14:15	Henrik Wenck: <i>A Bayesian Analysis of the Effectiveness of Non-Pharmaceutical COVID19 Interventions</i>
14:15 – 14:30	Oliver Adameck, Lukas Fritzsche & Jonas Noack (Master Project Baudisch): <i>Super Fast Fabrication of 3D Models by Folding Laser-Cut Foam Core</i>
14:30 – 15:00	Break

Today's Agenda II

Time	Title & Speakers
14:30 – 15:00	Break
15:00 – 15:15	Lars Jonas Bollmeier & Björn Daase: <i>Maximizing Persistent Memory Bandwidth Utilization for OLAP Workloads</i>
15:15 – 15:30	Christian Flach: <i>Call Graphs for Live Programming - Implementing Call Tracing in Babylonian/S based on a Survey of Property Extraction Techniques for Dynamic Analysis</i>
15:30 – 15:45	Nicolas Klodt, Lars Seifert & Arthur Zahn (Master Project Friedrich): <i>A Color-blind 3-Approximation for Chromatic Correlation Clustering and Improved Heuristics</i>
15:45 – 16:00	Tobias Bredow, Jona Otholt & Emanuel Metzenthin (Master Project Naumann): <i>Multimodal Analysis for Cultural Data</i>

Please engage!

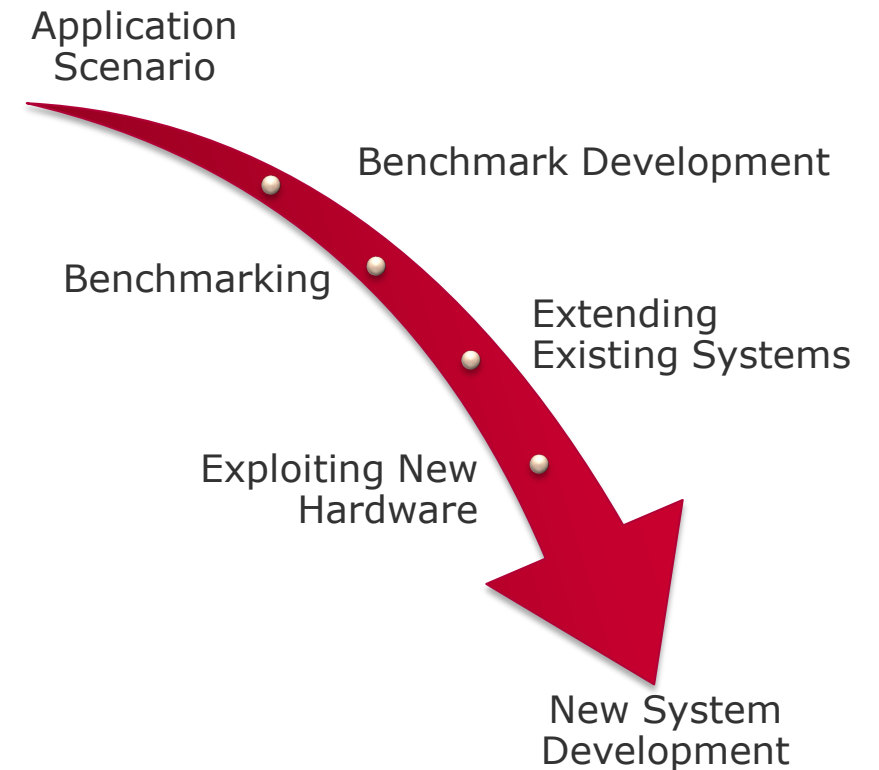
- Time for Q&A after every presentation
 - Unmute and ask
 - Type questions in chat
- Reach out to the presenters if you want to know more

Successful Systems Research

Disclaimer – Data Engineering Systems

- Database Systems (on Modern Hardware)
 - VLDBJ 18, SIGMOD 20, ICDE 21, PVLDB 21
- Stream Processing
 - PVLDB 19, SIGMOD 20, PVLDB 20, TODS 21
- Machine Learning Systems
 - PVLDB 17, SOCC 18, EDBT 19, SIGMOD 20
- Benchmarking
 - ICDE 18, TPCTC 19, PVLDB 20, SIGMOD 21

Research Approach

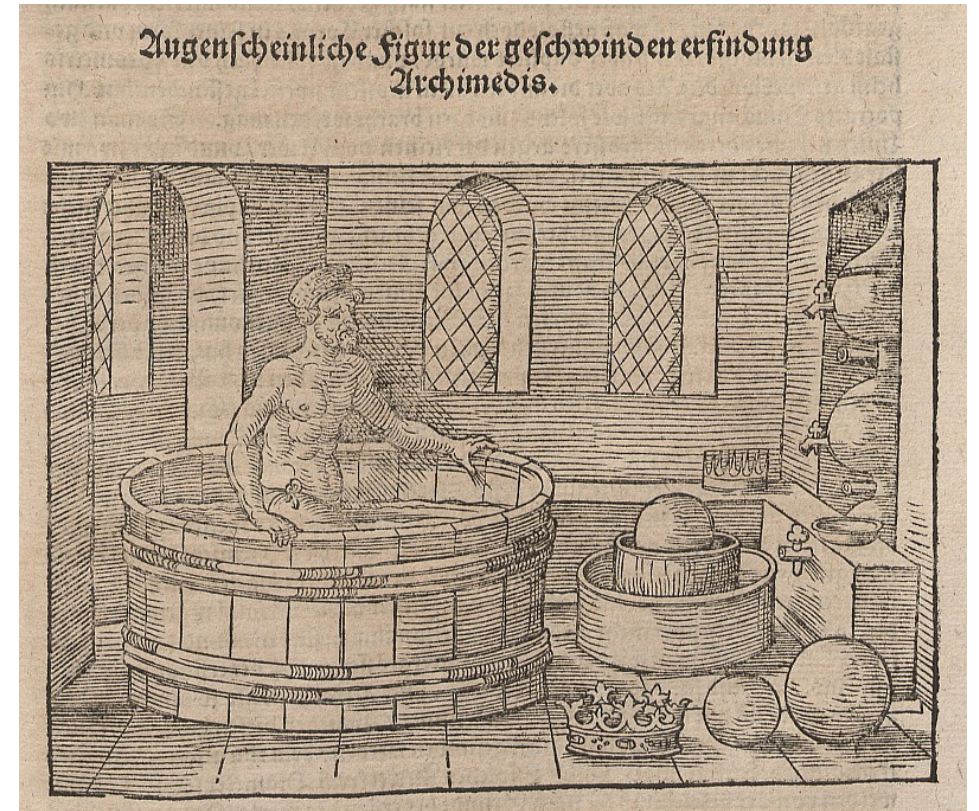


How to Start Successful Research?

- Look for an idea or a solution
 - E.g., combine recent buzz words
 - Blockchain, AI, green computing, IoT
- Problem: But why?



- Start with a problem or a question!
 - Current or future
 - There are problems everywhere
 - Often similar to ideas but with a concrete, **measurable** goal



Quelle: Deutsche Fotothek

What is a Good (System) Research Question

- Relevant
 - Advances research
 - Safes lives, the environment, ...
 - Makes money
 - Brings pleasure

- Measurable
 - You can track if you are making progress / when you are done
 - Comparable to other solutions
 - Science, not magic (i.e., I know how it works)

- If you like some research, try making it ...
 - faster, more accurate, incremental, more efficient, ...
 - **simpler**

Ethical Considerations

- With great power comes great responsibility
- Not everything is a tool
 - True for systems, algorithms, data
- Everybody in/on the pipeline needs to be responsible, careful, and thoughtful
 - Mistakes, misuse, manipulation
- Examples
 - Scoring / rating
 - Targeting
 - Optimization goal



VS.



Approach to Research

- On our way to find out if there is gold in a far away mountain

- We need to cross a jungle full of obstacles

- How should we go about it?
 1. Construct a road and bring mining equipment!

 2. Take our whip and walk through the jungle!

What is Gold?

- In general could be anything of value
 - Publication, fame, startups, money, fun, ...
 - *Relevancy*

- For academics: publications

- What can be published (in a good conference)?
 - In my field: something that is measurably better than previous work

 - Better ...
 - Performance, accuracy, functionality, utility, usability, ...

Understanding System Performance

- Modeling

- Back of the envelope calculation
- Analytical model

- Measurement

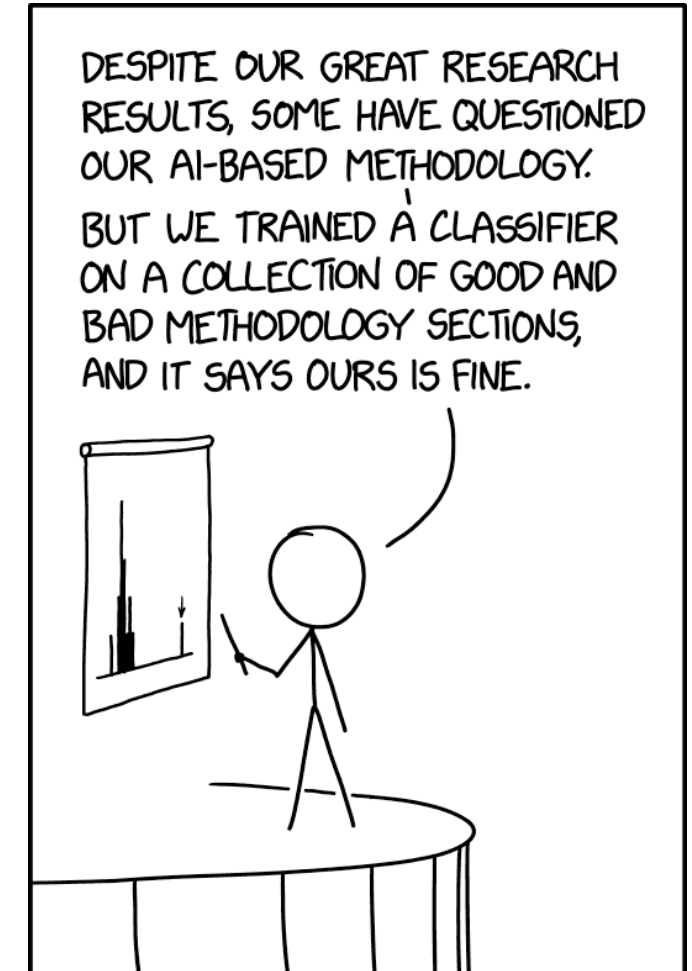
- Experimental design
- Benchmarks

- Simulation

- Emulation
- Trace-driven
- ...

Rule of Validation:

- Do not trust result of a single technique but validate with another
- Often: validate measurements with model



Source: <https://xkcd.com/2451/>

What to do with your Research?

- Publish!
 - Paper - journal / conference
 - Website
 - Presentation
 - Open-source repo

- Without some form of publication your work is lost

- The better the publication the better the visibility
 - Better venue, but also better writing!

Summary

- Successful (System) Research
 - Solution to a problem
 - Relevant
 - Measurable

- Verify, justify assumptions
 - Back of the envelope calculation
 - Rule of validation

- Publish your results or they are lost

- Join the research school! 😊

Thank you!

Contact: tilmann.rabl@hpi.de

Studi Pausenexpress Special!



- <https://mediaup.uni-potsdam.de/Play/23145>