

# Seminar Informationssysteme

Prof. Roberto V. Zicari

Dr. Karsten Tolle, Todor Ivanov, Kim Hee

Sommersemester 2018



## Termine - Seminar

---

- Di. 17.4.2018**      Vorbesprechung und Themenausgabe  
14:15 Uhr, R501 Robert-Mayer-Str. 10
- Mo. 11.6.2018**      Abgabe der **Ausarbeitung**  
(23:59 Uhr, mit PDF **und** Quelldatei/Word/TeX/etc.)
- Mo. 25.6.2018**      Abgabe der **Präsentation**  
(23:59 Uhr, mit PDF **und** Quelldatei/Word/TeX/etc.)
- Mo. 02.7.2018**      Abgabe der überarbeiteten (**finalen**) **Ausarbeitung**  
(23:59 Uhr, mit PDF **und** Quelldatei/Word/TeX/etc.)

Alle Abgaben per Email an:

**[seminar@dbis.cs.uni-frankfurt.de](mailto:seminar@dbis.cs.uni-frankfurt.de)**

Das Nicht-Einhalten von Terminen führt zum Ausschluss!

# Präsentationstermine

---

## Präsentationstermine (Vorträge – möglichst auf English):

- Montag, 9.7.2018, 9:00 – 14:00 Uhr, R501 Robert-Mayer-Str. 10

## Hinweise zur Präsentation

---

- Bearbeitung in 2er-Teams.
- Der Umfang der Präsentation (ohne Diskussion) ist bei 2er-Teams 20 Minuten. ... auf Englisch!
- Pro Folie max. 5 Gliederungspunkte, Minimale Schriftgröße 18 pt.
- Vorab testen! Zeitaufwand für Folien ist unterschiedlich!
- Alle Teilnehmer haben an einem Präsentationstag **Anwesenheitspflicht** für alle Präsentationen (Klausuren sind nachzuweisen). Die Teilnahme an mindestens einem Präsentationstag ist nötig.

# Hilfreiche Links zu den Fragen „Wie Präsentiere ich? Worauf muss ich achten?“

- <http://karrierebibel.de/praesentationstechniken-so-bleibt-ihr-vortrag-in-erinnerung/>
- [https://www.deeds.informatik.tu-darmstadt.de/fileadmin/user\\_upload/GROUP\\_DEEDS/teaching/WS2014\\_2015/seminars/deeds\\_seminars\\_beamer\\_presentation.pdf](https://www.deeds.informatik.tu-darmstadt.de/fileadmin/user_upload/GROUP_DEEDS/teaching/WS2014_2015/seminars/deeds_seminars_beamer_presentation.pdf)
- <https://files.ifi.uzh.ch/rerg/arvo/ftp/wap/WAP-Praesentationstechnik.pdf>

# Hinweise zur Ausarbeitung

---

- Umfang der Seminar-Ausarbeitung (**auf English!!!**):
  - 8 bis max. 12 Seiten
- Die Ausarbeitung **muss** unter Verwendung der vorgegebenen Springer LNCS Vorlagen erstellt werden

LaTeX: <ftp://ftp.springer.de/pub/tex/latex/lncs/latex2e/lncs2e.zip>

Word:

[http://static.springer.com/sgw/documents/1124637/application/zip/CSProceedings\\_AuthorTools\\_Word\\_2003.zip](http://static.springer.com/sgw/documents/1124637/application/zip/CSProceedings_AuthorTools_Word_2003.zip)

## Hinweise zur Ausarbeitung

---

- Die Ausarbeitung darf keine reine Übersetzung des Original-Papers sein.
- Es muss erkennbar sein, dass der Text verstanden und aufgearbeitet wurde! Nutzen Sie dafür Beispiele zur Erläuterung, Sekundärliteratur zur Vertiefung und stellen Sie sich die Frage in welchem Zusammenhang der Text entstanden ist.
- Bilden Sie sich eine **eigene** kritische Meinung!
- Zitate sind in der Ausarbeitung zu kennzeichnen. Sollte der Betreuer feststellen, dass Teile der Arbeit aus existierenden Arbeiten kopiert wurden, gilt das Seminar als nicht bestanden („**Wer kopiert, fliegt!**“).

## Bewertung (grob)

Was	Bemerkung	Anteil
Ausarbeitung (erste Fassung)	muss mind. mit „ausreichend“ bewertet werden	30%
Präsentation und Folien	muss mind. mit „ausreichend“ bewertet werden	30%
Mitarbeit (bei Diskussionen nach Präsentationen)		10%
Ausarbeitung (finale Fassung)	muss mind. mit „ausreichend“ bewertet werden	30%



# Hinweise zur Ausarbeitung Struktur

---

1. Abstract (150-300 words)
2. Introduction (up to 1 page)
3. Background (up to 3 pages) – explain related terminology and technologies used in the benchmarks that you review.
4. Benchmark Reviews (up to 1.5 pages per Benchmark) – use the benchmark features listed below to describe each benchmark:
  - Description
  - Benchmark type and domain
  - Workload
  - Data type and generation
  - Metrics
  - Implementation and technology stack
  - Reported results and usage
  - Reference papers
5. Conclusions/Summary (up to 1 page) – compare the benchmarks (advantages and disadvantages) and include your opinion of the comparison.

# Master Themen: Benchmark Kategorie

## 1. Machine Learning (1)

- **Sanzu** - <http://bigdata.cs.unb.ca/projects/sanzu/>
- **Penn machine learning benchmark (PMLB)** - <https://biodatamining.biomedcentral.com/articles/10.1186/s13040-017-0154-4>
- **OpenML benchmark suites** - <https://arxiv.org/pdf/1708.03731.pdf>
- **Minimal ML benchmark** - <https://github.com/szilard/benchmark-ml>
- **Criteo 1 TiB benchmark** - <https://github.com/rambler-digital-solutions/criteo-1tb-benchmark>
- **HPOLib2** - <https://automl.github.io/HPOLib2/stable/index.html>
- **MLbench** - <https://github.com/mlbench/mlbench>

## 2. Deep Learning (1)

- **DeepBench** - <https://github.com/baidu-research/DeepBench>
- **DeepMark** - <https://github.com/soumith/convnet-benchmarks/issues/101>
- **TensorFlow Performance** - <https://www.tensorflow.org/performance/benchmarks>

## 3. Deep Learning (2)

- **Fathom** - <https://arxiv.org/abs/1608.06581>
- **DAWNBench** - <https://dawn.cs.stanford.edu/benchmark/>

## 4. Streaming

- **Linear Road** - <https://pdfs.semanticscholar.org/7e3e/0563696d14c57c8339f03cf2c4bcca912063.pdf>
- **AIM Benchmark** - <https://db.in.tum.de/~kipf/papers/fastdata.pdf>
- **Senska** - [https://link.springer.com/chapter/10.1007/978-3-319-72401-0\\_3](https://link.springer.com/chapter/10.1007/978-3-319-72401-0_3)

## 5. Data Pipeline/Exploration

- **LinkBench** - <http://people.cs.uchicago.edu/~tga/pubs/sigmod-linkbench-2013.pdf>
- **AdBench** - [https://link.springer.com/chapter/10.1007/978-3-319-54334-5\\_8](https://link.springer.com/chapter/10.1007/978-3-319-54334-5_8)
- **IDEBench** - <https://idebench.github.io>

## 6. Big Data and SQL

- **TPC-H** - <http://www.tpc.org/tpch/>
- **TPC-DS v2** - <http://www.tpc.org/tpcds/default.asp>
- **TPCx-HS v2** - <http://www.tpc.org/tpcx-hs/default.asp>
- **BigFUN** - <https://asterix.ics.uci.edu/pub/BigFUNPaper.pdf>

## 7. Graph

- **Graphalytics** - <http://www.vldb.org/pvldb/vol9/p1317-iosup.pdf>
- **GARDENIA** - <https://arxiv.org/pdf/1708.04567.pdf>
- **WatDiv** - <http://dsg.uwaterloo.ca/watdiv/>