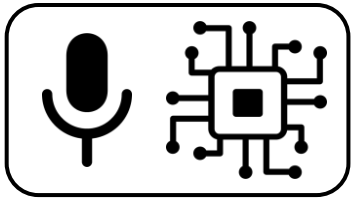
Computational Analysis of Sound and Music

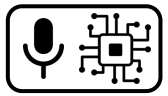


Introduction

Dr.-Ing. Jakob Abeßer

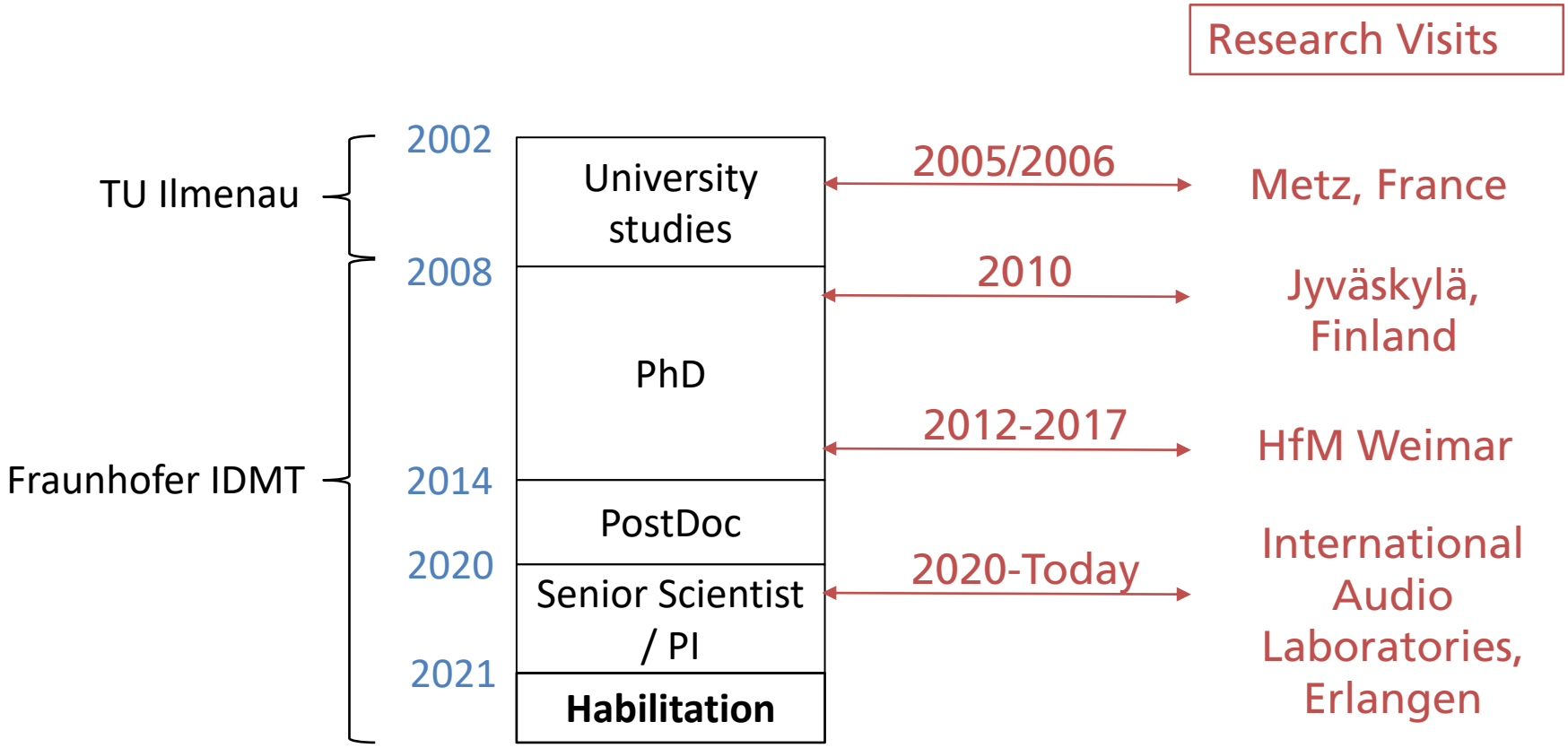
Fraunhofer IDMT

jakob.abesser@idmt.fraunhofer.de



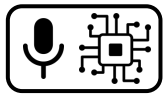
About me

- Dr.-Ing. Jakob Abeßer



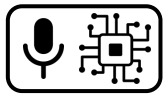
Lecture Structure

	Week	Date 1	Date 2
I. Foundations	1	Audio	Audio
	2	Audio	ML/DL
	3	ML/DL	ML/DL
II. Applications	4	Music Information Retrieval	
	5		
	6		
	7	Environmental Sound Analysis	
	8		
III. Research Project	9	Intro / Topics	Literature research
	10	Datasets	ML/DL pipeline
	11	Evaluation/metrics	Visualization/Paper writing
	12	Wrap-Up, Paper Deadline	Project presentation, Q/A



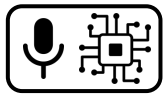
Lecture Structure

- Final grade
  - Written exam (90 min): 75 %
  - Research project: paper (2-3 pages) + presentation (5 min): 25 %
- Hybrid lecture-seminar structure (45 min – 45 min)
- Regular announcements on https://moodle.tu-ilmenau.de
- Requirements
  - Laptop + headphones
  - Access to https://colab.google/



Lecture Structure

- Teaching Approach
  - Task Introduction → Traditional Approaches → DL-based Approaches
- Course website (slides, notebooks, audio examples, demos)
  - https://machinelistening.github.io/casm
- Feedback → jakob.abesser@idmt.fraunhofer.de (Subject: "[CASM]: …")
  - Typos, mistakes, questions …
- Insights into projects & current research @ Fraunhofer IDMT
- Open student topics



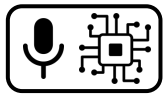
Lecture Structure

Research Project

- Learn “life-cycle” of a research project (essential skill for academia / industry R&D)
  - Literature & dataset research
  - Data processing → feature extraction → machine learning pipeline
  - Model evaluation
  - Data visualization / scientific writing

- Topics → Teams (~2 students)

- Outcome: scientific paper (2-3 pages) + presentation (5 min)



Resources

Books

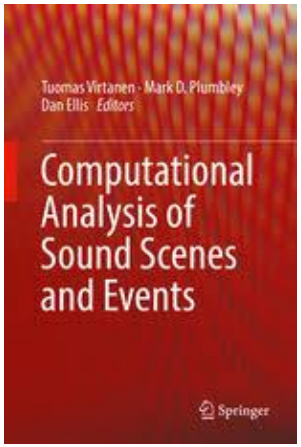


Fig. 1

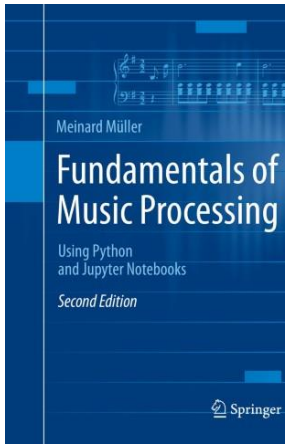


Fig. 2

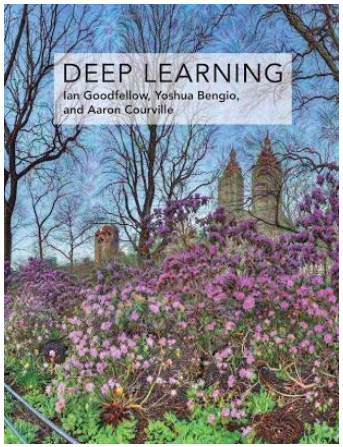


Fig. 3

- Virtanen, T., Plumbley, Mark D., and Ellis, Dan: *Computational Analysis of Sound Scenes and Events*, Springer, 2018.
- Müller, M.: *Fundamentals of Music Processing – Using Python and Jupyter Notebooks*, Springer, 2021.
- Goodfellow, I., Bengio, Y., and Courvill, A.: *Deep Learning*, The MIT Press, 2016.

Resources

Websites

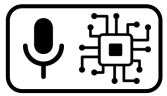
- Python
  - https://audiolabs-erlangen.de/PCP (Preparation Course Python)

- Digital Signal Processing
  - https://brianmcfee.net/dstbook-site/content/intro.html (Digital Signals Theory – online book)

Resources

Websites

- Machine Learning / Deep Learning
  - https://www.deeplearningbook.org/
  - https://machinelearningmastery.com
  - http://www.coursera.org (online courses)
  - http://www.udemy.com (online courses)



Resources

Websites

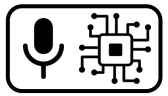
- Music Information Retrieval
  - https://www.audiolabs-erlangen.de/FMP (iPython notebooks)
  - https://musicinformationretrieval.com (iPython notebooks)

- Environmental Sound Recognition
  - http://dcase.community/ (DCASE challenges & workshop)

Resources

Websites

- Numerical computing / Statistics / Visualization
  - numpy, scipy, pandas, matplotlib, seaborn
- Machine Learning / Deep Learning
  - scikit-learn, tensorflow, pytorch
- Audio & Music Processing
  - librosa, soundfile, pysox
  - madmon
  - libfmp, synctoolbox, libtsm



References

Images

Fig. 1: https://media.springernature.com/w306/springer-static/cover-hires/book/978-3-319-63450-0

Fig. 2: https://media.springernature.com/w306/springer-static/cover-hires/book/978-3-030-69808-9

Fig. 3: https://mitpress.mit.edu/books/deep-learning