

---

# Stream Data Mining and Anytime Algorithms

Thomas Seidl

Department of Computer Science 9 (Data Management and Data Exploration)  
RWTH Aachen University, 52056 Aachen, [seidl@informatik.rwth-aachen.de](mailto:seidl@informatik.rwth-aachen.de)

**Abstract.** Sensors pervade all areas of personal, environmental and industrial domains, and nearly all applications in engineering, telecommunication, business, and life sciences produce tremendously increasing amounts of data. Though the availability of storage space grows at decreasing prices, many of the data require immediate analysis as they cannot be stored for reasons of their huge size or the fast reaction they require. In contrast to static data mining algorithms, stream data mining techniques follow the data and, as an additional challenge, the evolution of their concepts. In contrast to real-time algorithms which strictly obey fixed time budgets, anytime algorithms are designed to exploit the available time between the arrival of objects in a stream even for varying stream rates. Recently, various anytime classification techniques as well as anytime clustering algorithms have been proposed which were integrated into the MOA framework.

## References

- KRANEN P., ASSENT I., BALDAUF C., SEIDL T. (2011): The ClusTree: Indexing Micro-Clusters for Anytime Stream Mining. *Knowl Inf Syst*, 29(2), 249–272.
- KRANEN P., KREMER H., JANSEN T., SEIDL T., BIFET A., HOLMES G., PFAHRINGER B., READ J. (2012): Stream Data Mining using the MOA Framework. (Demo) *DASFAA*, Springer, 309–313.
- KRANEN P., SEIDL T. (2009): Harnessing the Strengths of Anytime Algorithms for Constant Data Streams. *Data Min Knowl Disc*, 19(2), 245–260.
- SEIDL T., ASSENT I., KRANEN P., KRIEGER R., HERRMANN J. (2009): Indexing Density Models for Incremental Learning and Anytime Classification on Data Streams. *EDBT/ICDT*, 311–322.

## Keywords

Data Mining, Stream Data Analysis, Clustering, Anytime Algorithms