

Big Sequence Management Scaling Up and Out

Karima Echihabi

Kostas Zoumpatianos

Themis Palpanas

*Mohammed VI
Polytechnic University*

*Snowflake Computing
work done while at
Harvard University &
University of Paris*

*University of Paris &
French University Institute (IUF)*

International Conference on Extending Data Base Technology (EDBT), March 2021



Questions This Tutorial Answers

how **important** are data series nowadays?

what does data series **analysis** involve?

how can we **speed up** such an analysis?

what are the different kinds of **similarity search**?

what are the state-of-the-art data series **indexes** for similarity search?

how can these indexes **parallelize/distribute** their operations?

can these indexes be used for **general high-d vector** similarity search?

what data series management **systems** exist?

what is special about them?

what are the **open research problems** in this area?

what are the connections to **deep learning**?