

Data Systems Group

Presentation to ENS Paris-Sarclay

Dr. Laurent Bindschaedler | Research Group Leader

✉ bindsch@mpi-sws.org



MAX PLANCK INSTITUTE
FOR SOFTWARE SYSTEMS

My Research: Next-Gen Big Data Systems



The Data Revolution puts Big Data systems under strain!



Hard to support new requirements and use cases

Data Volume



Scalability



Maximize
Performance / Utilization

Data Velocity



Interactivity



Provide
Low-Latency Results

Workload & Data Variety



Autonomy



Automate
Performance Tuning

Benchmarking “Learned” Systems

Learned Systems use Machine Learning (ML) to synthesize new components (e.g., index, query optimizer) => major performance gains

Problem: how can we trust that a learned system works?

- ▶ Benchmarking (show improvements, avoid regressions)
- ▶ Adversarial workloads that will “break” the ML models



Synthesize data&workloads from real-world deployments



Build a benchmark that incorporates new workloads and metrics



Automatically generate adversarial workloads

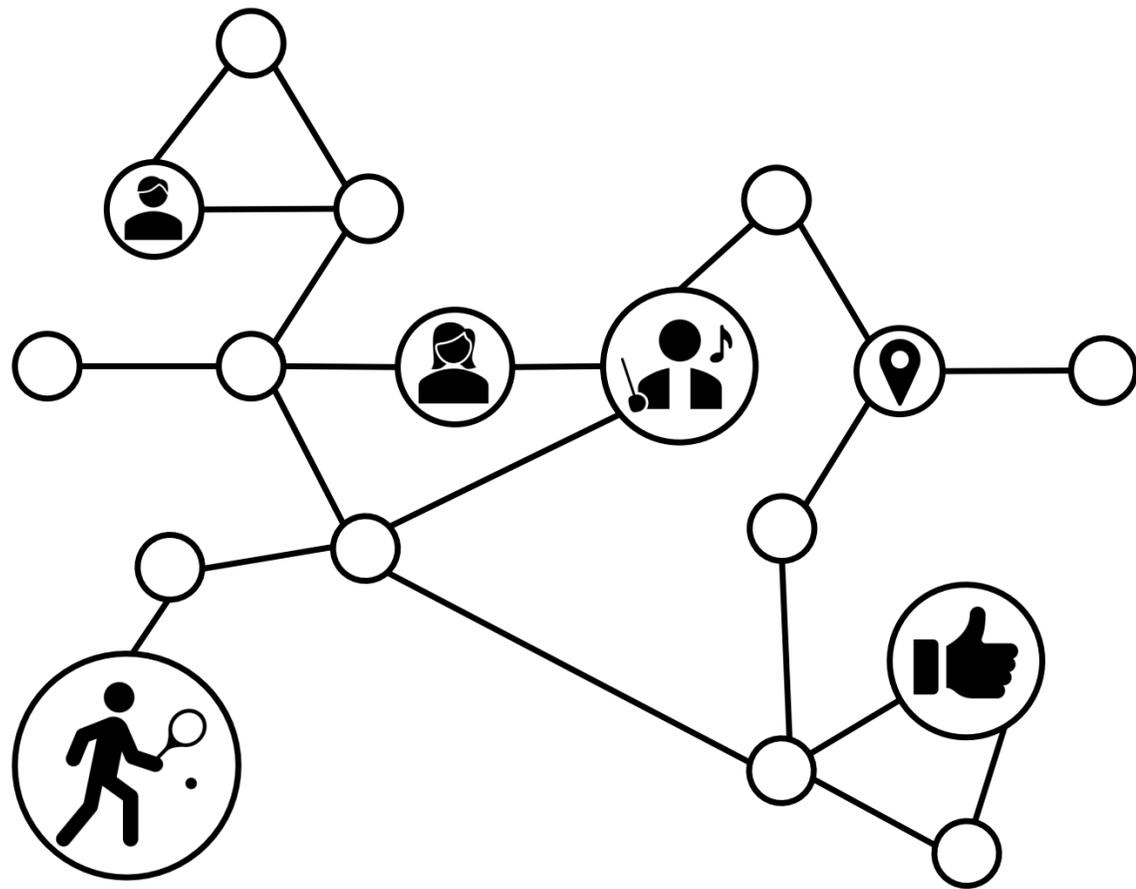


Anomaly Detection in Dynamic Graph Datasets

SCALABILITY

INTERACTIVITY

Find outliers (e.g., surprising patterns)
in fast-changing graphs



Applications

Network intrusion detection
Financial fraud detection
Failure troubleshooting
Law/customs enforcement
Event detection in IoT sensors
“Fake” news identification

- ? What system architecture given millions of updates per second?
- ? How to define what is “normal” and “abnormal”? (API)
- ? How to deal with noise and present usable results?

Ingestion & Storage of Satellite Imagery

SCALABILITY

INTERACTIVITY

AUTONOMY

Image Streams

Cleaning & Pre-processing

Applications

- Agriculture
- Climate
- Navigation
- Disasters
- Weather
- Infrastructure
- Military
- ...

Image Recognition

Indexing

Storage Engine

Multi-modal Data

OSINT Integration

Queries

?

Caching

?

Materialization

?

Compression

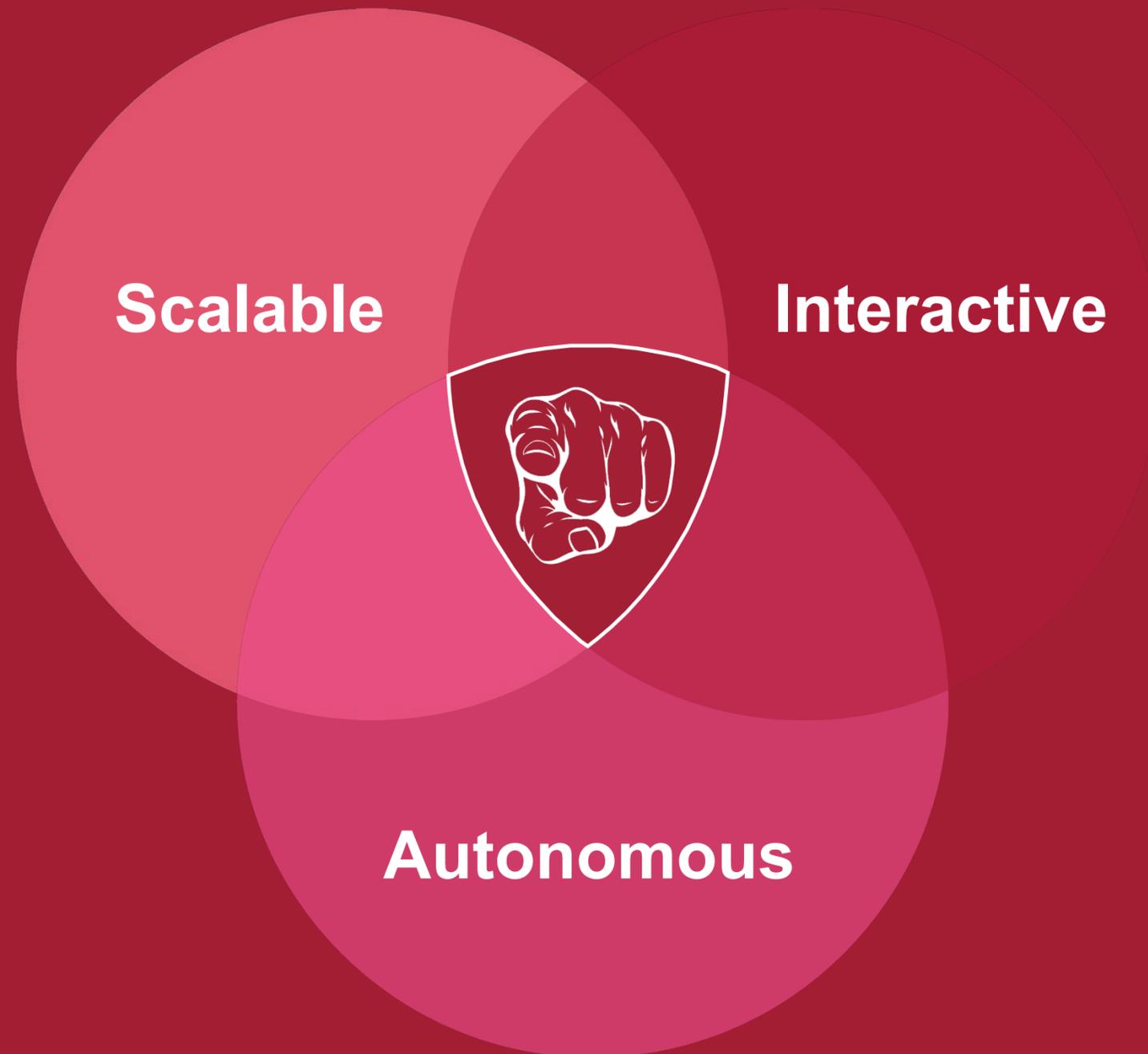
?

Archival

Collaborators



I am looking for motivated interns/students!



Please reach out!

bindsch@mpi-sws.org



JE PARLE
Français