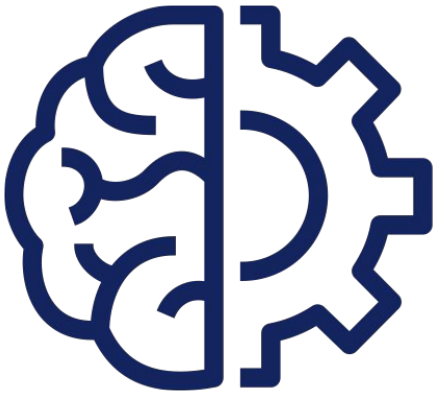

VIENNA DATA SCIENCE TOOLS

NOVOMATIC

https://reinforceconf.com/

PyWren

http://pywren.io/

PyWren

- Run your existing python code at massive scale via **AWS Lambda** and **S3**
- PyWren supports the stuff you would use in your normal Python workflow:
  - Jupyter
  - NumPy
  - scikit-learn
- Learn more at
  - https://www.oreilly.com/ideas/building-accessible-tools-for-large-scale-computation-and-machine-learning
- numpywren
  - Designed to deal with large matrices by storing them on Amazon S3 instead of on RAM on live EC2 instances.
  - Decouples memory and computation for parallel algorithm design.
  - https://github.com/Vaishaal/numpywren

Example - Trivial Parallelization with PyWren

```
def my_function(b):
  x = np.random.normal(0, b, 1024)
  A = np.random.normal(0, b, (1024, 1024))
  return np.dot(A, x)

pwex = pywren.default_executor()
res = pwex.map(my_function, np.linspace(0.1, 100, 1000))
```

TimescaleDB

https://www.timescale.com/

TimescaleDB

- Open-source time-series database powered by PostgreSQL
- Why TimescaleDB?
  - Full SQL. Not SQL-ish
  - Looks like PostgreSQL on the outside, architected for time-series on the inside.
  - JOIN time-series with relational metadata in the database, not the application.
  - Write millions of data points per second.
  - Store 100s of billions of rows and 10s of terabytes in TimescaleDB.
- Transparent, automatic partitioning
  - TimescaleDB automatically partitions data into time-based chunks to support faster performance at scale. Users interact with a single table that transparently manages chunks.
- Optimized query engine
  - We modified the PostgreSQL insert path, execution engine, and query planner to intelligently process queries across chunks.
- Time-series functionalities
  - Supports efficient data lifecycle management operations like data retention and archiving. It also comes with time-series specific analytical functions for easier data manipulation.
- Learn more at
  - https://www.oreilly.com/ideas/a-scalable-time-series-database-that-supports-sql

Thank you!

bpirvu@novomatic.com
jwilms@novomatic.com

NOVOMATIC