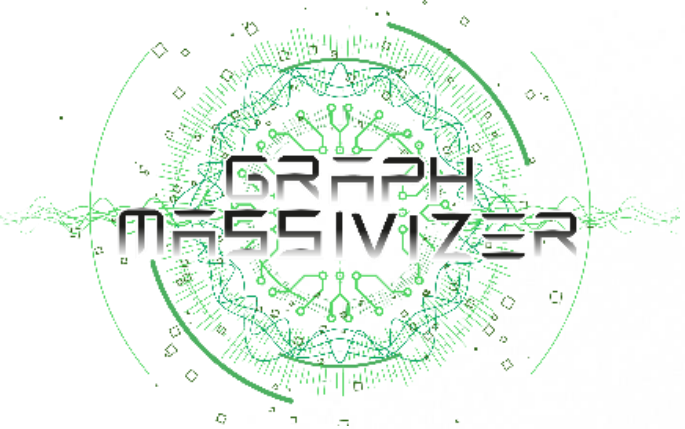


UNIVERSITÄT KLAGENFURT

Graph-Massivizer

Radu Prodan

Andrea Bartolini, Marko Grobelnik, Peter Haase,
Alexandru Iosup, Evgeny Kharlamov, Gregor Leban,
Dumitru Roman, Roberta Turra,
Laurențiu Vasiliu, Ana Lucia Vărbănescu

Graph-Massivizer

- Extreme and Sustainable Graph Processing for Urgent Societal Challenges in Europe

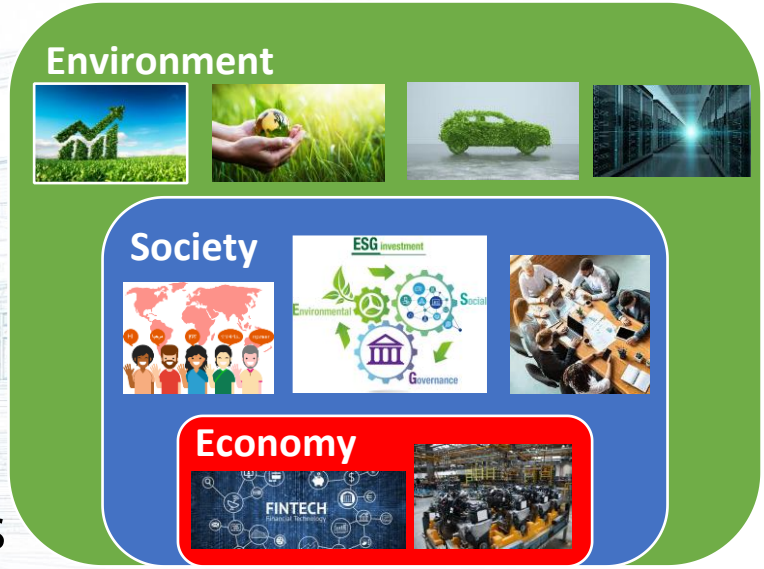
- HORIZON-CL4-2022-DATA-01-05
  - Extreme data mining, aggregation and analytics technologies and solutions (RIA)

- € 4,998,062.50

- 2023 – 2025

Goal

- High-performance and sustainable graph processing of extreme data
  - Holistic platform and integrated toolkit
  - Proper response for any need and any organization by 2030

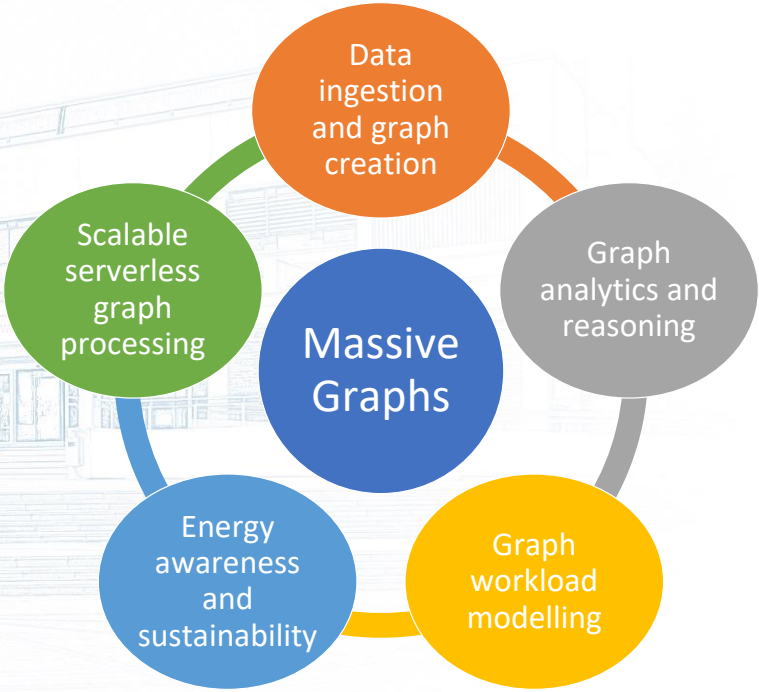


- Graphs
  - Universal mathematical abstractions
  - Capture, combine, model, analyze, and process knowledge about real and digital worlds

Massive Graphs

- **Massive graph (MG)** representation of **extreme data**, integrating patterns and storing interlinked descriptions of objects, events, situations, concepts, and semantics
  - General graphs
  - Knowledge graphs
  - Property graphs

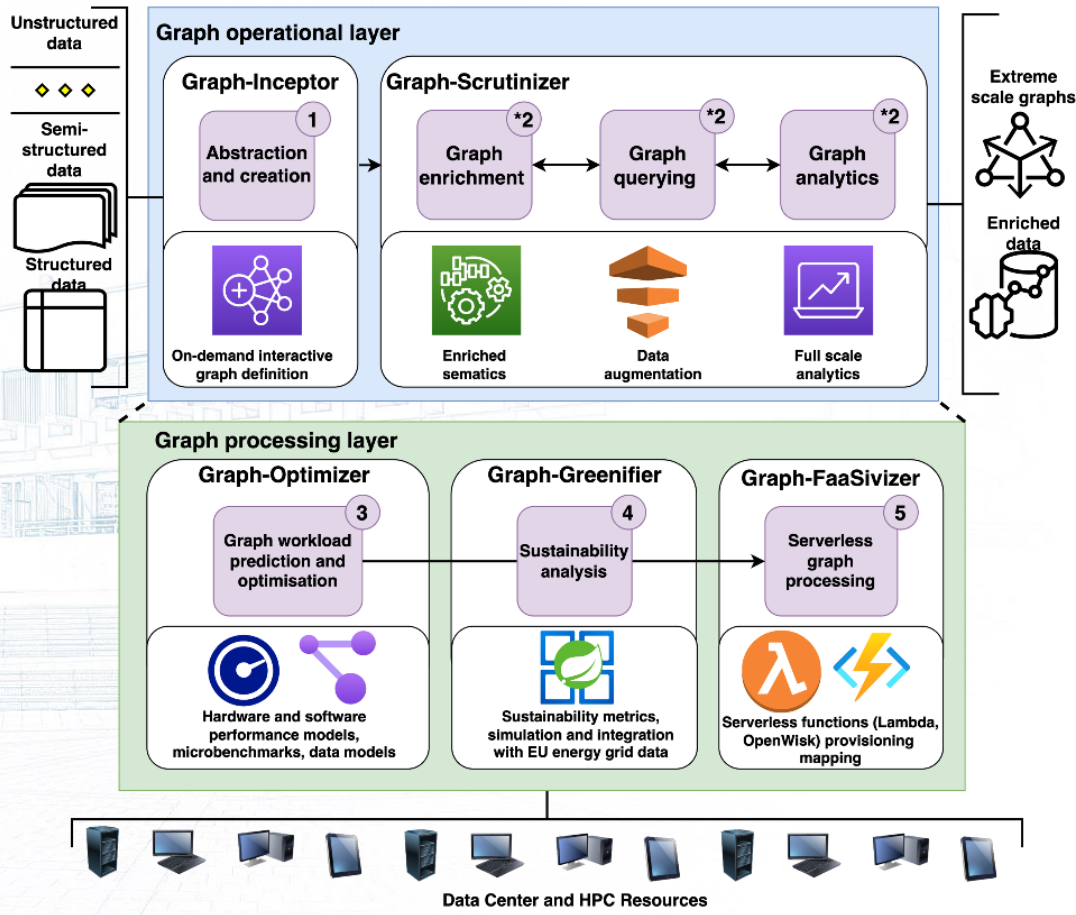
- **Extreme data challenge**, extending graph processing technology by orders of magnitude for several **"V"**-characteristics
  - **Volume** graph challenge by supporting up to billions of vertices and trillions of edges
  - **Velocity** graph challenge of dynamically changing topologies
  - **Viridescence** graph challenge for sustainable processing at scale



Graph-Massivizer Toolkit

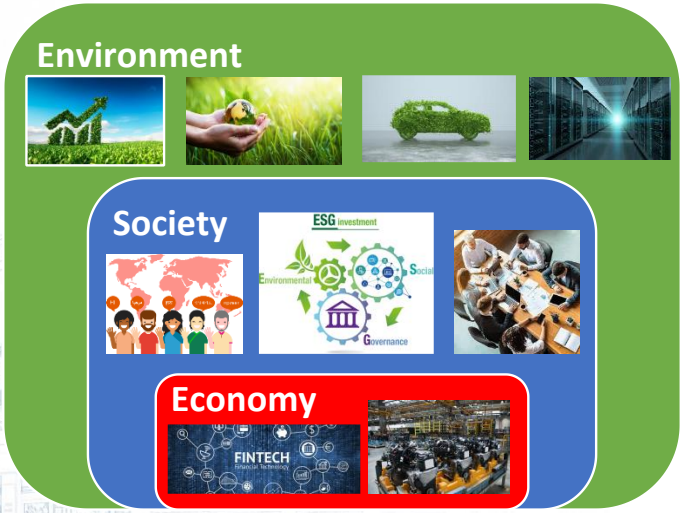
- Graph operational layer
  - **Graph-Inceptor**: extreme data ingestion, MG creation and storage
  - **Graph-Scrutinizer**: MG analytic and reasoning

- Graph processing layer
  - **Graph-Optimizer**: workload modelling with performance and energy guarantees
  - **Graph-Greenifier**: sustainable and energy-aware MG processing
  - **Graph-Serverlizer**: scalable serverless MG analytics over a codesigned computing continuum



Use Cases

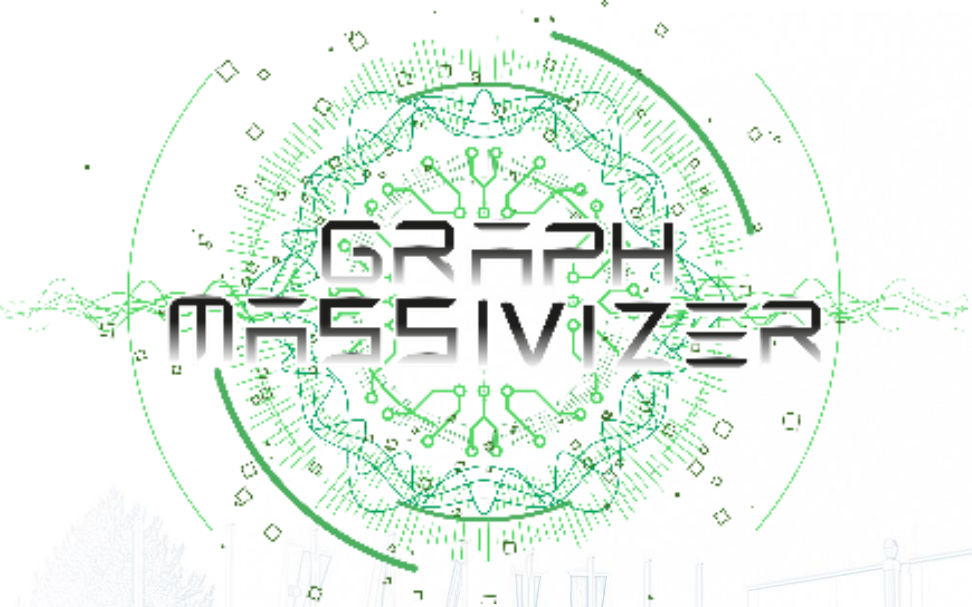
- UC-1: Green and sustainable finance
- UC-2: Global foresight for environment protection
- UC-3: Green AI for sustainable automotive industry
- UC-4: Data center digital twin for sustainable exascale computing



	Volume	Velocity	Value	Veracity	Variety	Viscosity	Viridescence
UC-1	✓		✓	✓			✓
UC-2	✓				✓	✓	✓
UC-3		✓		✓	✓		✓
UC-4	✓	✓			✓		✓

Key Performance Indicators

- < 500 milliseconds streaming ingestion latency for 95% of the data at 1,000s of new edges per second
- 500 giga edges per second sampling and analytics throughput
- 3 million triples per second streaming-graph reasoning speed
- 80% accurate performance and energy consumption models for BGOs
- 2x higher data center energy efficiency and >25% lower GHG emissions
- 70% faster graph analytics than AliGraph or other OSS solutions
- 30% less energy for ETL storage than Amazon Redshift and OSS
- 40% faster runtime deployment than AWS CloudFormation



UNIVERSITÄT KLAGENFURT

IDC

Peracton

Jožef Stefan Institute

SINTEF

UNIVERSITY OF TWENTE.

metaphactory

VU VRIJE UNIVERSITEIT AMSTERDAM

CINECA

THANK YOU !

European Commission
Horizon Europe 2021-2027

EVENTREGISTRY

BOSCH