Secure and Efficient Data Spaces

Y. Thomas, I. Pittaras, F. Bistas, K. Mantaraki, George Xylomenos
MMLab, Athens University of Economics and Business

Christos Papadopoulos
University of Memphis

https://mm.aueb.gr/

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Motivation

- ETSI Data Space: a new form of digital platform
  - Liberates data from silos
  - Enables data-driven innovation
  - Shapes digital transformation
  - Offers NGSI-LD API for queries
- SNDS project (NGI Sargasso OC1)
  - Prototyped Data Space over NDN
  - Fully distributed and self-sovereign identity system
  - Outperformed traditional IP-based networks

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Goals

- SeEDS: Implement a full Data Space over NDN
  - Required + Optional data API operations
    - Content filtering based on conditions
    - Temporal queries over timestamped data
    - Subscriptions to data events
  - In-network security and privacy
    - Data integrity verification
    - Selective content revelation
  - Enhanced efficiency
    - Distributed data intermediaries
    - Migration of content filters

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Objectives

- Build an ETSI-compliant data space over NDN
  - Refactor & extend prototype from NGI Sargasso OC1 SNDS
- Full query-based content retrieval API
  - Multiple types of content filters to reduce traffic
- Security scheme based on self-sovereign DIDs
  - Privacy-preserving content and integrity protection
- Reuse of NDN codebase and validation in NDN testbed
  - Open-source data space implementation for NDN

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Clients – Users (Wearables)

- Who would be interested in SeEDS?
  - Anyone who has lots of sensitive data
  - But needs to share some of this data with others
- Wearable device vendors
  - Wearables generate private data
    - Medical readings, activities, location, interactions
  - SeEDS allows private data to be stored in untrusted clouds
    - Data are always stored encoded in the cloud
    - No need to trust the device vendor to store data
    - The user can selectively reveal data to doctor or hospital

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Clients – Users (Building IoT)

- Who would be interested in SeEDS?
  - Anyone who has lots of sensitive data
  - But needs to share some of this data with others
- Large building managers
  - Building IoT sensors generate private data
    - User identity, location, activity and behavior
  - SeEDS allows private data to be stored in untrusted clouds
    - Data are always stored encoded in the cloud
    - The building manager is not responsible for private data
    - The user can reveal energy profile to supplier for demand response

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Clients – Users (Federated Learning)

- Who would be interested in SeEDS?
  - Anyone who has lots of sensitive data
  - But needs to share some of this data with others
- Communities building federated learning models
  - Federated learning requires a lot of data exchange
    - Endpoints train model with local data, send parameters to global model
    - Update global model, send new model to endpoints
  - SeEDS provides data space with native multicast
    - Endpoints subscribe to model updates
    - Global model updates distributed via multicast

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Economic impact

- Incremental NDN deployment
  - CDNs with private backhauls exist (e.g., Facebook)
  - An open-source NDN solution would be ideal for them
- Benefits to SeEDS adopters
  - Open-source: no proprietary code, no vendor lock-in
  - Open API (NGSI-LD): can use different data intermediaries
  - Decentralized: relies on DIDs, no need for expensive PKIs
  - Secure: can check individual packets, access control via DIDs
  - Private: no exposure of content names, selective content revelation

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Environmental & social impact

- Content distribution greatly affects the environment
  - CDNs and social networks move data across large distances
  - SeEDS moves functionality inside the network
    - Rich API with temporal and content filtering capabilities
    - Content filter migration to the most appropriate place
- SeEDS hardens and extends NDN's security and privacy
  - Data Spaces address EU's privacy and security concerns
  - DIDs are critical to decentralization on the Internet at large
  - SeEDS gives control of data back to its creators/owners

NGI

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

Thank you

https://mm.aueb.gr/

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS