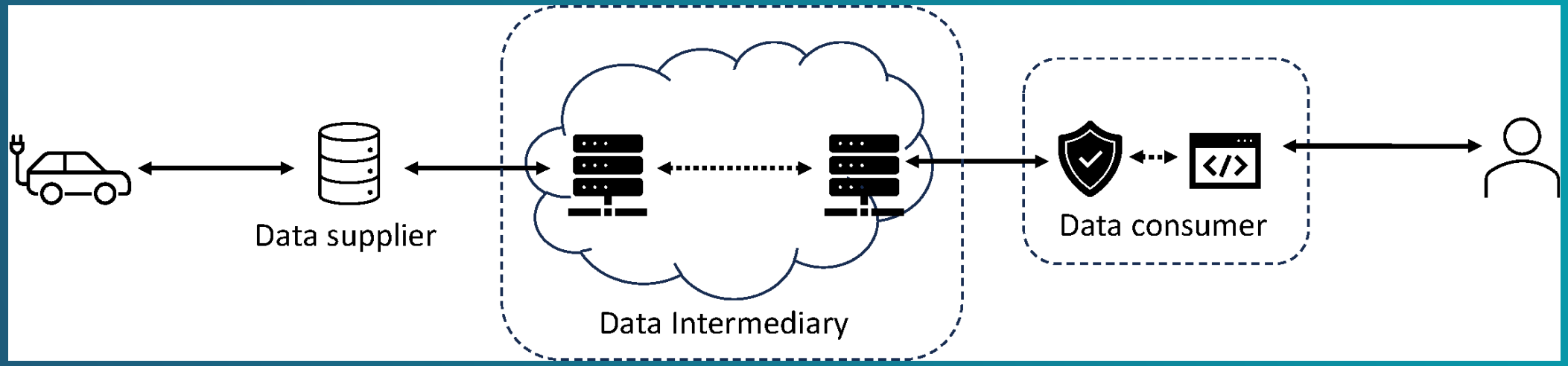
Data integrity protection for data spaces

Nikos Fotiou
ExcID

George Xylomenos, Yannis Thomas
Mobile Multimedia Laboratory, AUEB

https://mm.aueb.gr/projects/snds

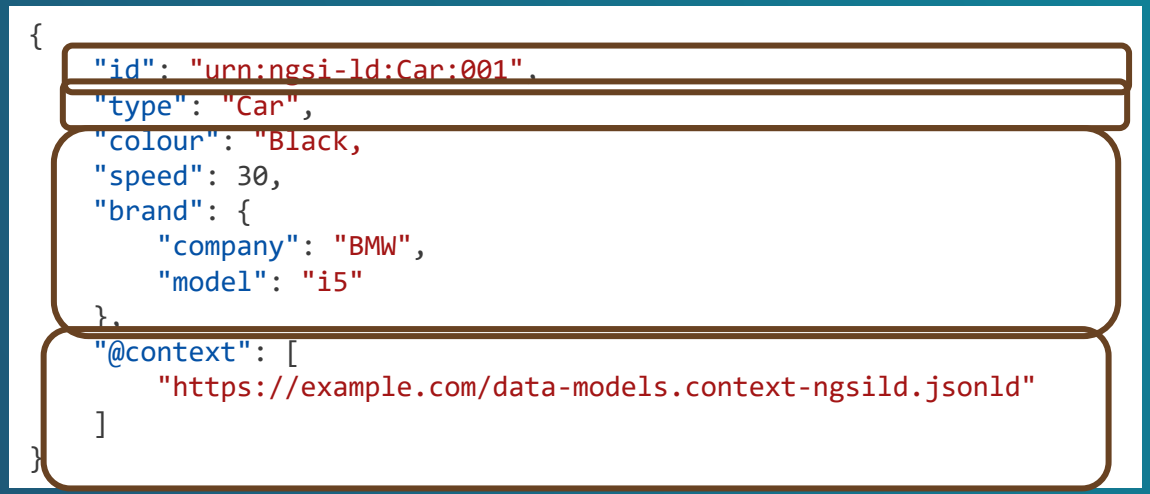
What is a data space?



- A way to liberate data from silos
  - Data supplier shares data with data consumers
  - An intermediary facilitates the exchange
  - Standardized storage & retrieval API at intermediary
    - Example: ETSI's NGSI-LD API

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

NGSI-LD Objects



```
{
    "id": "urn:ngsi-ld:Car:001",
    "type": "Car",
    "colour": "Black,
    "speed": 30,
    "brand": {
        "company": "BMW",
        "model": "i5"
    },
    "@context": [
        "https://example.com/data-models.context-ngsild.jsonld"
    ]
}
```

- Each entity is a uniquely identified JSON-LD encoded object
- Each entity has a "type"
- Each entity consists of name-value "attributes"
  - Which may be JSON-LD objects themselves
- The attributes of a type are defined in a "context"

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

NGSI-LD Queries - Retrieve Item

GET https://broker.com/ngsi-ld/v1/entities/urn:ngsi-ld:Car:001

```
{
    "id": "urn:ngsi-ld:Car:001",
    "type": "Car",
    "colour": "Black,
    "speed": 30,
    "brand": {
        "company": "BMW",
        "model": "i5"
    },
    "@context": [
        "https://example.com/data-models.context-ngsild.jsonld"
    ]
}
```

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

NGSI-LD Queries - Retrieve Attributes

GET https://broker.com/ngsi-ld/v1/entities/urn:ngsi-ld:Car:001?**attrs=brand%speed**


```
{
    "id": "urn:ngsi-ld:Car:001",
    "type": "Car",
    "speed": 30,
    "brand": {
        "company": "BMW",
        "model": "i5"
    },
    "@context": [
        "https://example.com/data-models.context-ngsild.jsonld"
    ]
}
```


Colour is hidden!

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Challenge

- Given that:
  - The data space stores complete items
  - Attribute hiding is implemented by the intermediary
  - Consumer may not be authorized to access hidden attributes
  - Consumers trust suppliers

➞ How can consumers verify the integrity of partially retrieved items?

✓ We propose two signing solutions that support selective disclosure

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Approach

NGI

ExcID

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

THE UNIVERSITY OF MEMPHIS

A "disclosures" construction mechanism

- A common construction shared by both solutions
- Step 1: Decompose an item into a list of "disclosures"
- Step 2: Reconstruct a "partial" item by selecting which "disclosures" to reveal

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Step 1: Decomposition

- Turn each attribute into a "disclosure"
  - The disclosure name is the attribute's "JSON Pointer"
  - The disclosure value is the attribute's value
- For each composite attribute repeat the process for all inner objects

```
{
    "id":"urn:ngsi-ld:Car:001",
    "type":"car",
    "colour":"black",
    "speed": 30,
    "brand":{
        "company":"bmw",
        "model": "i5"
    }
}
```

/id	urn:ngsi-ld:Car:001
/type	car
/colour	black
/speed	30
/brand	{'company': 'bmw', 'model': 'i5'}
/brand/company	bmw
/brand/model	i5

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Step 2: Reconstruction

- Select the disclosures to reveal
- Perform the reverse process
  - E.g., reveal 1st, 4th and 6th disclosures

/id	urn:ngsi-ld:Car:001
/type	car
/colour	black
/speed	30
/brand	{'company': 'bmw', 'model': 'i5'}
/brand/company	bmw
/brand/model	i5

```
{
    "id":"urn:ngsi-ld:Car:001",
    "speed": 30,
    "brand":{
        "company":"bmw",
    }
}
```

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Two non-solutions

- Sign the entire item
  - Cannot split it at the intermediary
  - Need to sign each and every possible attribute combination
- Sign each attribute separately
  - Does not preserve objects
  - No way to know if attributes belong to the same object

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Hash signature-based approach - Sign

Performed by the supplier

1. Generate a random salt value for each disclosure
2. Hash each disclosure + salt value
3. Digitally sign concatenation of all hashes
4. Store signature, hashes, disclosures and salt values at intermediary

H(	/id	urn:ngsi-ld:Car:001	<Salt>	)
H(	/type	car	<Salt>	)
H(	/colour	black	<Salt>	)
H(	/speed	30	<Salt>	)
H(	/brand	{'company': 'bmw', 'model': 'i5'}	<Salt>	)
H(	/brand/company	bmw	<Salt>	)
H(	/brand/model	i5	<Salt>	)

Sign()

Hash signature-based approach - Reveal

Performed by the intermediary

1. For each revealed attribute also reveal the salt
2. For the hidden attributes only return the hash
3. Reveal the signature
   - E.g., reveal 1st, 4th and 6th disclosures

	/id	urn:ngsi-ld:Car:001	<Salt>
H(	~~/type~~	~~car~~	~~<Salt>~~)
H(	~~/colour~~	~~black~~	~~<Salt>~~)
	/speed	30	<Salt>
H(	~~/brand~~	~~{'company': 'bmw', 'model': 'i5'}~~	~~<Salt>~~)
	/brand/company	bmw	<Salt>
H(	~~/brand/model~~	~~i5~~	~~<Salt>~~)

Hash signature-based approach - Verify

Performed by the consumer

1. For each revealed attribute calculate the hash
2. Concatenate the hashes
3. Verify the signature

H(	/id	urn:ngsi-ld:Car:001	<Salt>	)
H(	~~/type~~	~~car~~	~~<Salt>~~	)
H(	~~/colour~~	~~black~~	~~<Salt>~~	)
H(	/speed	30	<Salt>	)
H(	~~/brand~~	~~{'company': 'bmw', 'model': 'i5'}~~	~~<Salt>~~	)
H(	/brand/company	bmw	<Salt>	)
H(	~~/brand/model~~	~~i5~~	~~<Salt>~~	)

Verify()

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

BBS+ signature-based approach

- BBS+ signatures allow:
  - A signer to sign a "group" of messages
  - A third party to select a subset of the messages and calculate a ZKP to prove that "this subset of the messages is part of the set that matches the provided signature"
  - Size of the ZKP linear on hidden disclosures
- Straightforward application to the list of disclosures ➞ Treat each disclosure as a separate message

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Implementation*

- Part of Secure Named Data Sharing (SNDS) project
  - Data supplier stores disclosures & signatures at intermediary
  - Intermediary accessed via NGSI-LD API gateway
  - Gateway trusted (and operated) by consumer
- Gateway operation
  - Receives NGSI-LD API request from consumer
  - Retrieves disclosures from intermediary
  - Validates signatures and proofs
  - Responds with reconstructed partial item

*https://github.com/mmlab-aueb/selective-disclosure

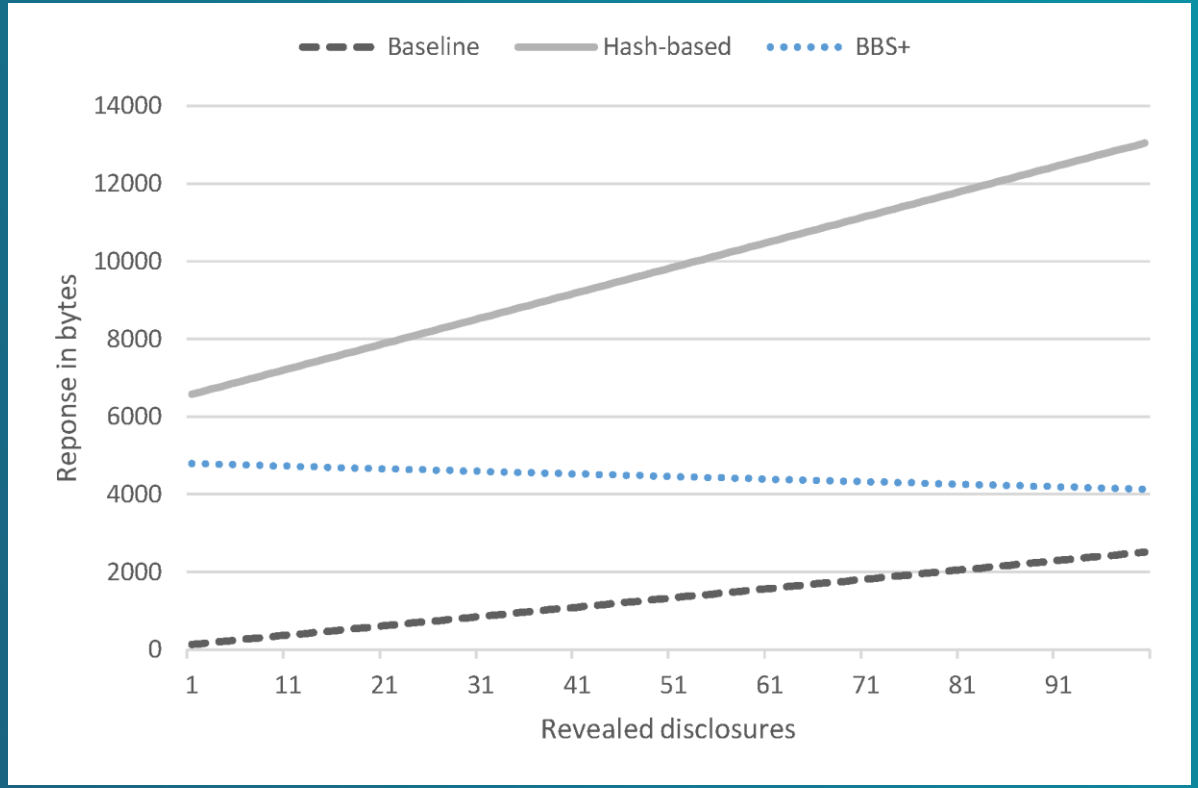
NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Evaluation

- Baseline scenario
  - Intermediary is fully trusted
  - Constructs partial content item
  - Signed with ECDSA with the P-256 curve
- Hash-based scenario
  - 128-bit random salts
  - Hashes calculated using SHA-256
  - Signed with ECDSA with the P-256 curve
- BBS+-based scenario
  - MATTR's BBS+ implementation
  - Group signatures using the BLS12-381 pairing-friendly elliptic curve
  - Must also create/send ZKP based on revealed disclosures

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Communication overhead



- Size of responses returned
  - Against disclosures revealed

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Computing overhead

	Supplier (sign original items)	Intermediary (sign disclosed items)	Consumer (verify signatures)
Baseline	None	<0.1ms	<0.1ms
Hash	<0.1ms	None	<0.1ms
BBS+	22ms	70ms	60ms

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Security properties

	Hash-based	BBS+
Integrity protection	✅	✅
Indistinguishability of hidden disclosures (Let D1 and D2 be the same attribute for two items, if D1 and/or D2 is hidden, the gateway should not be able to say if D1==D2)	✅	✅
Untraceability of disclosures (Let D1 be a hidden attribute of an item and D1' the same attribute after some time, the gateway should not be able to say if D1==D1')	❌	✅
Hide the total number of disclosures	❌	❌

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS

Summary

- Selective revelation in data spaces
  - Data integrity protection
  - Content only signed by supplier
  - Content selectively revealed by intermediary
  - Hash based solution: higher communication overhead
  - BBS+ based solution: higher computing overhead
- Future work
  - Disallow hiding of some attributes (e.g. object ID)
  - Integrate with FIWARE Orion context broker
  - Integrate proofs directly into NGSI-LD objects

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS UofM THE UNIVERSITY OF MEMPHIS

Thank you

https://mm.aueb.gr/projects/snds

NGI ExcID ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS THE UNIVERSITY OF MEMPHIS