Trusted Data Sharing

Day 1 | 13:30 – 14:30
1. **Why** share data?
2. **What** data should we (not) share?
3. **How** should we share data?
4. **Enablers** of data sharing
Why share data?

Use cases

G2G
Policy planning, verifying eligibility, administrative efficiency

G2P
Open data, government transparency, accountability

B2B
Targeted ads, supply chain coordination, open banking

Climate resilience
Early warning, disaster response, resource monitoring

Innovation
New business models, improved service quality
AI use case

Generative AI

Train

Discover

Share

Store

Process

Create

Image credit: DALL-E 2
Many diverse requirements
What data should we share?

**Types of data**
- Personal data
- Sensitive data
- Secret data (classified)
- Anonymized data
- Aggregated data
- Open data (public)
- Financial data
- Metadata

**Lawful bases**
- Consent
- Contractual necessity
- Legal obligation
- Vital interests
- Legitimate interests
- Public interest
- Public task

**Purpose specification**
- Relevance
- Risk assessment
- Data minimization
- Selective disclosure
- Anonymization
- Granularity
- Aggregation
- Least privileges
- Retention
How much data should we share?
Making sense of diverse requirements

Risk of undersharing
- Availability
- Discoverability
- Automation
- Scalability

Risk of oversharing
- Confidentiality
- Access Controls
- Data minimization
- User control

General
- Data integrity
- Data provenance
- Accountability
- Regulatory compliance
Data sharing methods

Direct
Transactional (systems integration)
Batch (file-based, ETL workflows)
Push (broadcasts, event-driven, IoT)

Intermediated
Data brokers, marketplaces
Data fiduciaries

User-centric
Credentials (paper and digital)
Digital wallets

Aggregation
Data warehouses, data lakes
Open data platforms
Method strengths

- Direct
  - Provenance ✓
  - Scalability ✓

- User-centric
  - User control ✓
  - Natural consent ✓

- Aggregation
  - Availability ✓
  - Discoverability ✓
Method weaknesses

**Direct**
- User controls hard to implement
- Low transparency

**User-centric**
- Low availability, discoverability
- Hard to scale, automate
- Always consent based
- Only for personal data

**Aggregation**
- Privacy
- Increased risks with AI
Data sharing enablers

Enablers and controls to prevent over- and under-sharing, to optimize the amount of data shared.

**Standardization**
- Semantic interoperability
- Data standards
- Standard interfaces, protocols
- Discoverability

**Technology enablers**
- Authentication, digital ID
- Public key infrastructure
- Encryption
- Monitoring, logs, audit trails
- Systems integration (API, ESB...)
- Consent receipts, e-signature

**Institutional enablers**
- Data governance
- Legal sanctions and penalties
- Legislation (data protection, open data, e-signature...)
- Policies (access control, authorized use, retention...)