



REPUBLIC OF ESTONIA  
INFORMATION SYSTEM AUTHORITY

# Estonian e-government

## RIHA & data architecture

**Andres Kütt**

Information System Authority, architect

July 11, 2015

## **In this session**

- Data architecture of Estonia
- Information System Registry in the country context
- Key challenges of system governance

**Goal: explain (data) governance of Estonian information system**

# **Data architecture of Estonia**

# Key principles

- Once only: you should not have to give the government data it already has
  - Basically “thou shalt not create a master data problem”
  - Also cost and bureaucracy reduction
- You collect it, you protect it
  - The organisation collecting a data item has full responsibility
  - For example, data access cannot be granted by a third party
- All data is public by default
  - Explicit protection exists for key areas
  - Meant to foster open data
- The citizen owns the data
  - And can see, who has accessed it and why
  - This creates strange issues with medical information

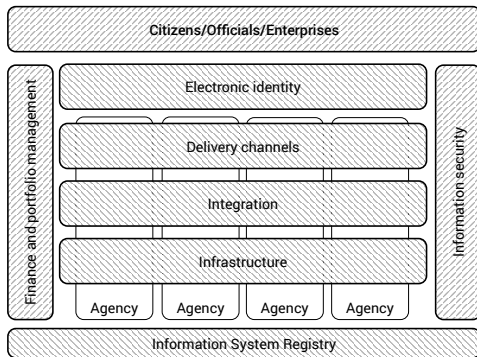
# Key challenges

- Marrying x-road and open data
  - Most data collected is sensitive
  - How can we form a coherent data model out of this?
  - Why would I access diluted public data when I can access the rich private dataset?
- Relationships between data and information
  - A SQL dump is a really poor source of documentation
  - Information is emergent property of data
  - Master vs. secondary data
  - Semantics questions

# **Information System Registry in the country context**

# Why do we need another registry?

- Service discovery
- Data discovery
- Competence discovery
- Organisation discovery
- Regulatory enforcement



**You can't govern what you don't measure**

# **Challenges of system governance**



# Service discovery

- Using services assumes you know about them
  - 1000+ services is a lot to go through manually
  - Both x-road and channels layers depend on the services
- Using services assumes you know how to use them
  - Documentation
  - WSDL files
- Using services assumes you can get access
  - Who'm specifically do I contact to get access?
  - Where do I ask questions?

**Key question: what services do exist and whom do I talk to about access?**

# Data discovery

- What data could I **potentially** use?
  - What organisations collect what data?
  - What is the semantics of the data collected?
  - Whom do I talk to about building a service to access it?
- Service descriptions are not necessarily useful
  - Movement towards service oriented APIs
  - Primary and secondary data elements (Marriages and ages)

**Key question: what data is there?**

# Competence discovery

- Not all data is alike
  - Geoinformation
  - Marriage (the Y2Gay problem)
  - Foreigners with their funny names and documents
  - Weather, health etc.
- Who has the problem I have but on a larger scale?
  - They must have figured out a solution
  - Or at least have some ideas

**Key question: who knows how to deal with an issue?**

# Organisation discovery

- It's all about people, folks
  - Even in Estonia, not everybody knows everybody
  - Who can give me access?
  - Who has data?
  - Who can help me?
- **Beware:** conflict discovery
  - You can only fight someone if you know about them!
  - Who else is seeking to be the source of truth in this field?

**Key question: whom do I talk to?**

# Regulatory enforcement

- Enforcement of policies assumes information about the proceedings
  - What data is stored and how?
  - Have the systems been audited?
  - Are you storing duplicates of data?
  - Is your registry even legal?
- General sense of maturity levels
  - When was the documentation last updated?
  - Does it look like the person wrote it knows what they are doing?

**Key question: is it legal?**

## How to collect the data?

- Having people feed another registry does not work very well
  - High benefit disparity: providing information is costly and yields little value, consuming it is cheap and adds a lot of value
  - Documentation standards are a challenge
  - Documenting the documentation of code: what could possibly go wrong?
- What about making the data available?
  - Providing access instead of providing data
  - “Don’t call us we’ll call you” approach
  - Machine readability gives open data for free and enables a lot of interesting things



REPUBLIC OF ESTONIA  
INFORMATION SYSTEM AUTHORITY

# Thank you!

**Andres Kütt**

andres.kutt@ria.ee