Practical guide how to inventory and open public sector data

Milda Aksamitauskas

University of Wisconsin - Madison

Fulbright scholar, Lithuania

How to have ALL public sector data at your fingertips?

Most public sector organizations were not designed ~~for data analytics~~

Project framework

- Chief Data Officer function assigned to Data Agency
- Centralized technical platform – national data lake
- 70 FTE team
  - information managers*NEW*
  - programmers
  - data engineers
  - analysts
  - linguist*NEW*

- "Data audit" in ALL public sector institutions
  - Deep interview
  - Technical connection
  - Data integration
  - Open a data set

Project timeline

Year 0
dreams about
the project

2023 – Year 1

team build up (0->30 FTEs)

inventory of health and
social assistance
programs

100 inventoried information resources

100 open datasets

2024 – Year 2

scale up the team
(70 FTEs)

local government
inventory

2025 – Year 3

operational
optimization
(70 FTEs)

focus on data quality

100 data insights

2026 – Year 4

project
sustainability

maintenance mode

EU OPEN DATA DAYS

Legal approach

- **National law adjustment**
  - Data agency – controller of the data
  - Requirement to inventory and integrate data with the Data agency
  - 7 allowable data use cases

- **GDPR compliant**
  - processing is necessary for the performance of a task carried out in the **public interest** or in the exercise of official authority vested in the **controller**

EU OPEN DATA DAYS

LT open data eco-system overview

Data.gov.lt portal

- Infrastructure
- Familiarity
- Notifications
- Public requests

Data reuse

- Public transportation
- Tourism
- Issued permits

New focus

- Open at least one dataset from each inventoried system*NEW*
- Priority to open High Value Datasets
- Quality of data

Lithuania   94.2

7th place in EU Open Data Maturity ranking

Quality

Impact

EU OPEN DATA DAYS

Best practices for massive data inventory

- Speed and scale
- Team size and financing
- Tracking system
- Priorities
- Political support
- Communication
- Transparency
- Continuous process

Data governance

- Repeated questions
  - Who are you?
  - We don't have any data
  - We don't use data for analysis
  - Our data is too sensitive
- Low data maturity
- Outdated and incomplete agreements with technical vendors

opportunities
- Investment in data
- Leadership
- Center of excellence
- Standards

Successful communication strategies

- Same talking points for everyone
- Don't shy why away from difficult questions
- **Don't judge**
- In-person site visits
- Annual events and data ambassador awards
- Find spokespeople outside of the Data agency

Transferrable lessons

- Set a bold vision
- Identify the lead person
- Empower the team
- Don't overlook the basics
  - Composition and size of the team
  - Legal context
  - Communication plan
  - Agreements in writing
- Be prepared to be surprised

- Track everything – payout comes later
- Synergies of traditional statistics and on demand analysis
- Open data provide a taste of what data is available
- Not all data can be opened
- Restrictive data use cases

EU OPEN DATA DAYS

parengti metaduomenų aprašymus

sukurti jungtis

atverti

323
valstybės IS / registrus
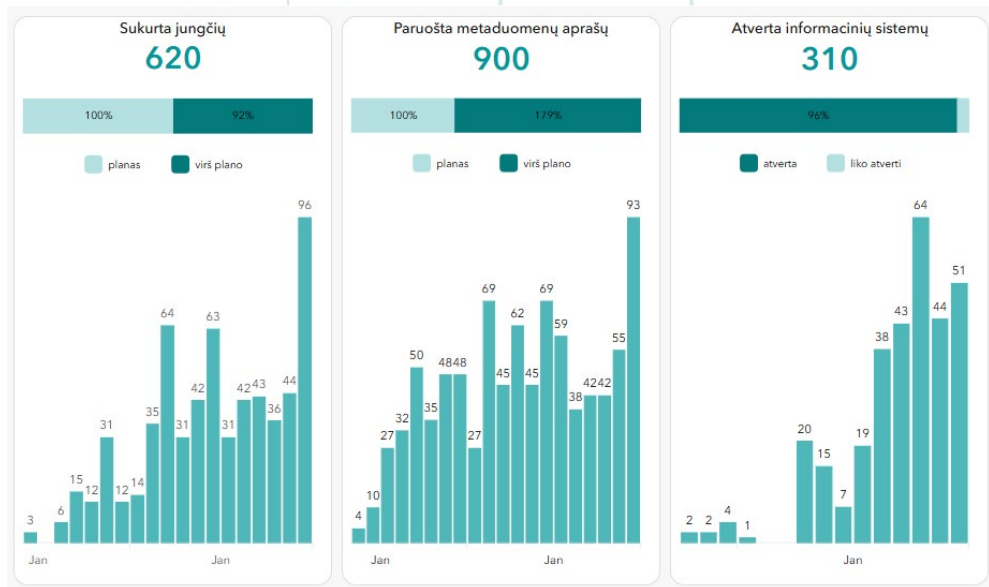
Pasiekti rezultatai:

467 įstaigos

1,677 aptikti šaltiniai

310 atvertos IS

494 atverti rinkiniai

Sukurta jungčių
620

100% 92%

planas virš plano

Paruošta metaduomenų aprašų
900

100% 179%

planas virš plano

Atverta informacinių sistemų
310

96%

atverta liko atverti

Projekto pradžia: 2022-06   Projekto pabaiga: 2026-06

...ašo, kad pamatytumėte daugiau informacijos apie jos duomenų

Institucija	Veiklos sritis	Įvertinimas
Nacionalinė švietimo agentūra	Švietimas	4
VĮ Registrų centras	Konkure…	4
VšĮ Statybos sektoriaus vystym…	Aplinka	4

Pasirinktos institucijos vertinimas
Žemiau pateikiama išsami informacija apie institucijos duomenų brandos vertinimo kriterijus ir gautus įverčius.

Nacionalinė švietimo agentūra
Vadybininkas (-ė): Justina Zadranavičienė
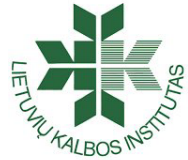
Bendras įvertinimas (maks. 5)

Įverčio detalizacija pagal kriterijus

Vertinimas atliekamas pagal aštuonis kriterijų blokus, apimančius institucijos ir informacinių sistemų lygmenis, siekiant įvertinti duomenų kokybę ir institucijos komunikaciją.

Bendro vertinimo formulė

Pastaba: užveskite pelės žymeklį ant bendro įvertinimo žuvų skalės, kad pamatytumėte detalų aritmetinį įvertinimo skaičiavimą.

1. Bendradarbiavimo įvertis	Komunikacijos vertinimas	2	+
2. Duomenų bazės brandos įvertis	Techninis vertinimas	3	+
3. Jungčių brandos įvertis	Techninis vertinimas	3	+
4. Duomenų bazės įvertis	Nevertinama	0	+
5. PDSA įvertis	Komunikacijos vertinimas	5	+

EU OPEN DATA DAYS
19 – 20 MARCH 2025

Thank you!
Ačiū

- Milda Aksamitauskas
- aksamitauska@wisc.edu
- https://www.linkedin.com/in/milda-aksamitauskas/