IT security in State Data Governance Information System Division of Statistics Lithuania

VALSTYBĖS
DUOMENŲ
AGENTŪRA

Marijus Bernotas
Senior developer
State Data Governance Information System division
Statistics Lithuania

Official Statistics and State Data Governance

On 1 January 2023, the Department of Statistics of the Republic of Lithuania became State Data Agency.

In accordance with Article 5, Part 1 of the Law on Official Statistics and State Data Management of the Republic of Lithuania, the State Data Agency is a government institution of the Republic of Lithuania that participates in forming state policy not only in the field of official statistics management entrusted to the Minister of Finance, but also in the field of state data management.

Needs and expectations

Our organization provides:

- Important indicators for our country and various data analysis
- Data solutions for governmental, health and science organizations
- Crisis management tools, including those related to COVID-19
- Open data
- Data for Eurostat
- IT environments for governmental organizations and the scientific community to make their own analyses using the data integrated into our system.

State Data Governance division

The success of our State Data Governance Information System was driven by these three core elements:

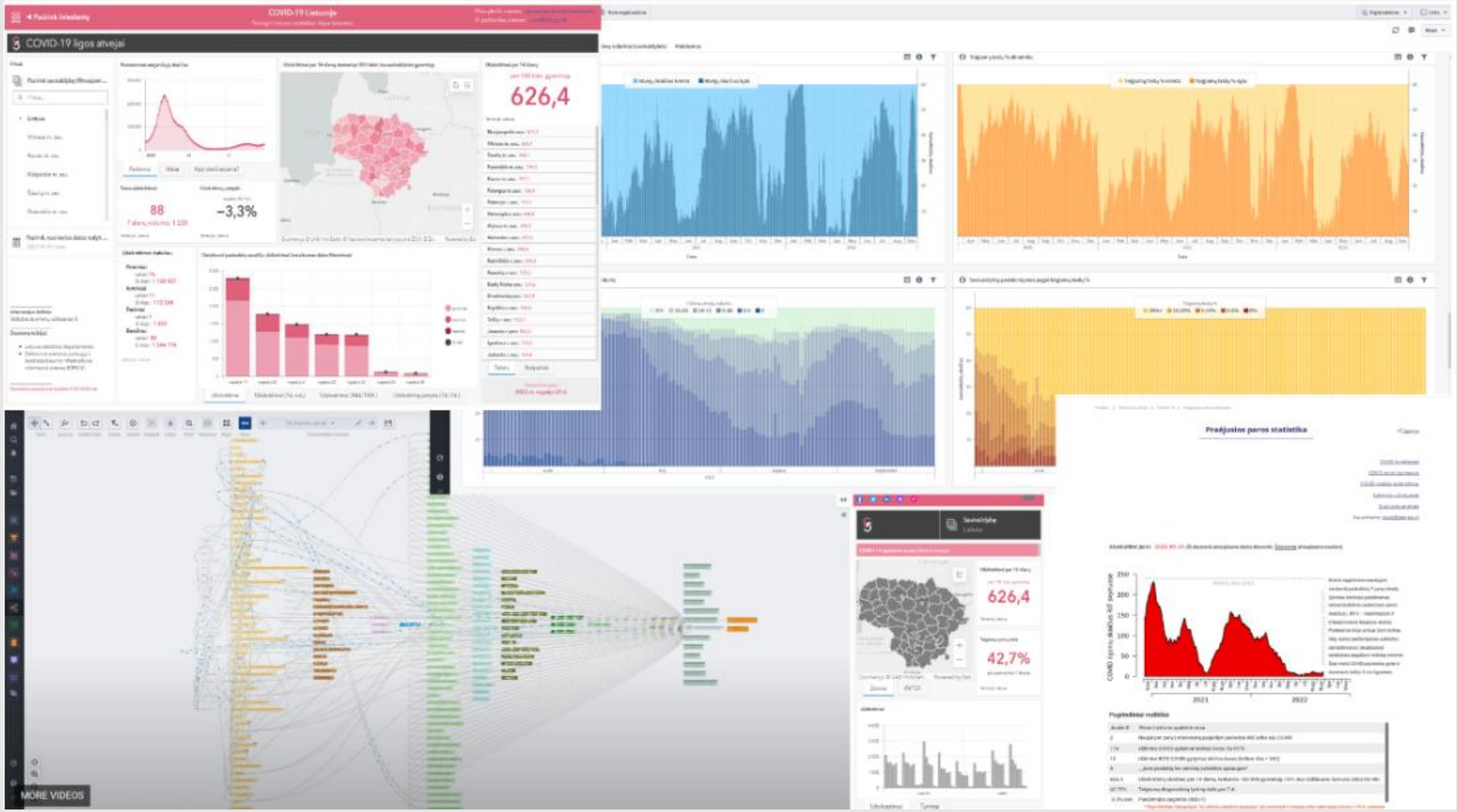
- Legal regulation
- Technical solutions
- Expertise/competencies

Example - COVID-19 dashboards and reports

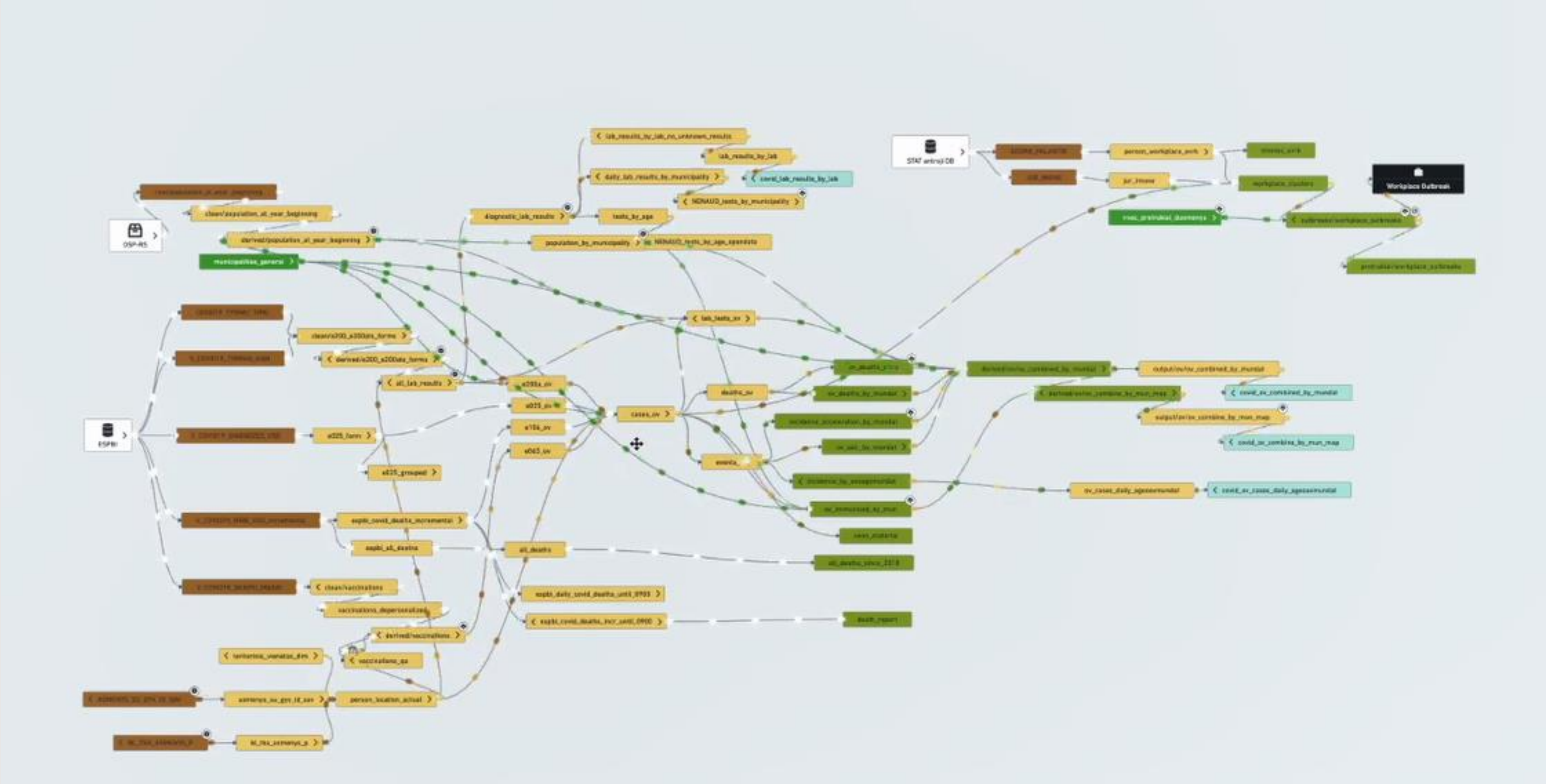
- Internal dashboards
- External dashboards

https://osp.stat.gov.lt/covid-dashboards

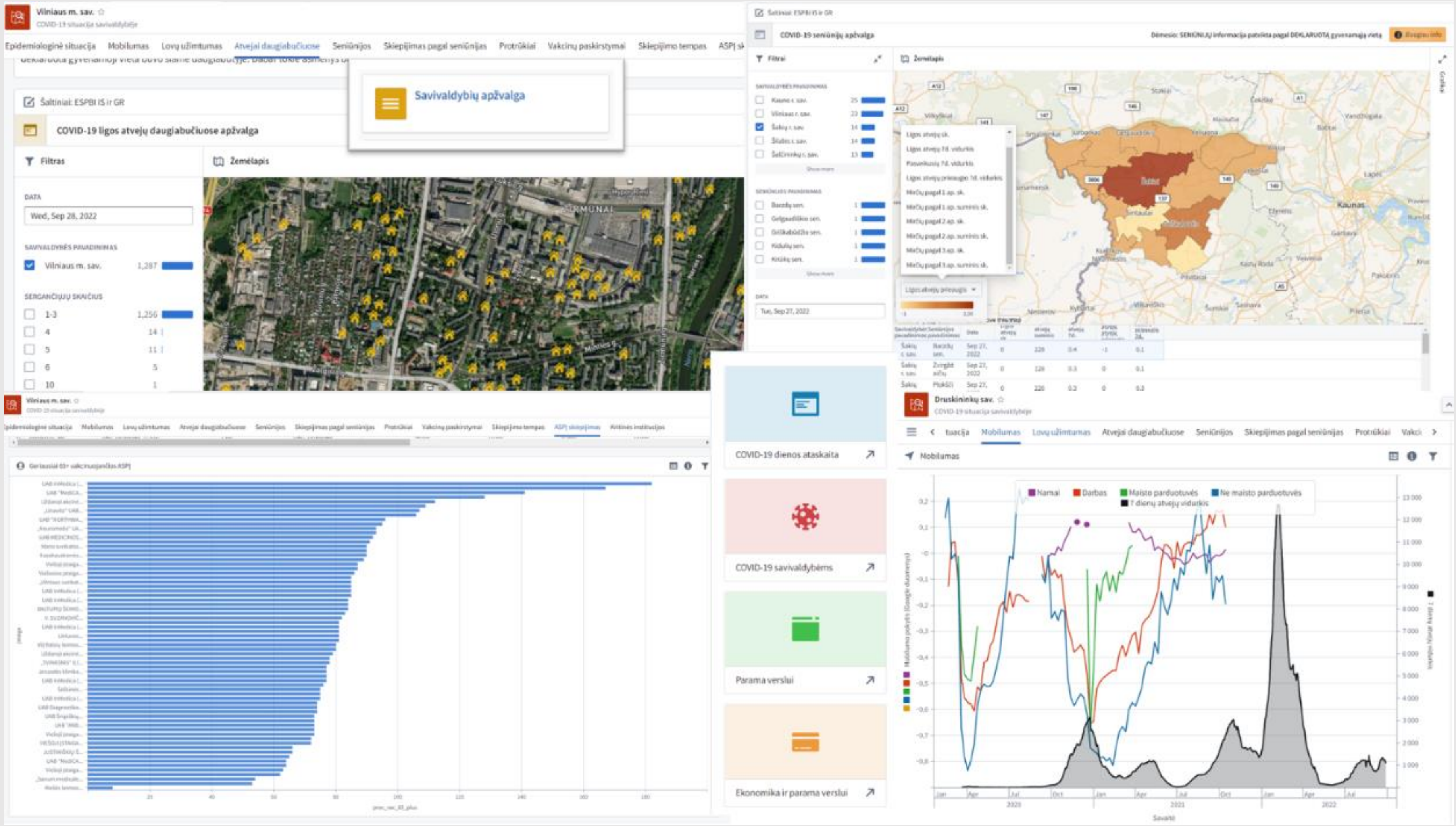
- Compatibile with mobile devices
- Daily report



Example – COVID-19 data pipeline


DSP-RS
clean/population_at_year_beginning
derived/population_at_year_beginning
municipalities_general
ESPBI
clean/a200_a200ats_forms
derived/a200_a200ats_forms
all_lab_results
a025_form
a025_grouped
espbi_covid_deaths_incremental
espbi_all_deaths
clean/vaccinations
vaccinations_depersonalized
derived/vaccinations
vaccinations_qa
teritorinis_vienetas_dim
asmenys_su_gyv_id_sav
person_location_actual
iki_rta_asmenys_p
diagnostic_lab_results
lab_results_by_lab_no_unknown_results
lab_results_by_lab
daily_lab_results_by_municipality
covid_lab_results_by_lab
NENAUD_tests_by_municipality
tests_by_age
population_by_municipality
NENAUD_tests_by_age_aparatai
lab_tests_ov
a098a_ov
a025_ov
a104_ov
a063_ov
cases_ov
deaths_ov
events_
all_deaths
espbi_daily_covid_deaths_until_0900
espbi_covid_deaths_incr_until_0900
death_report
all_deaths_since_2018
cases_mutants
ov_immunised_by_mun
incidence_by_ageagemunicipal
ov_vacc_by_mundial
incidence_acceleration_by_mundial
ov_deaths_by_mundial
ov_deaths_since
STAT antroji DB
person_workplace_work
jur_imone
movies_work
workplace_clusters
Workplace Outbreak
mes_protrukiai_duomenys
outbreaks/workplace_outbreaks
protrukiai/workplace_outbreaks
derived/ov/ov_combined_by_mundial
output/ov/ov_combined_by_mundial
covid_ov_combined_by_mundial
derived/ov/ov_combine_by_mun_map
output/ov/ov_combine_by_mun_map
covid_ov_combine_by_mun_map
ov_cases_daily_ageasemundial
covid_ov_cases_daily_ageasemundial

Example – internal dasbhoards for municipalities



Vilniaus m. sav.

Šaltiniai: ESPBI IS ir GR

COVID-19 ligos atvejų daugiabučiuose apžvalga

DATA: Wed, Sep 28, 2022

SAVIVALDYBĖS PAVADINIMAS

☑	Vilniaus m. sav.	1,287

SERGANČIŲJŲ SKAIČIUS

□	1-3	1,256
□	4	14
□	5	11
□	6	5
□	10	1

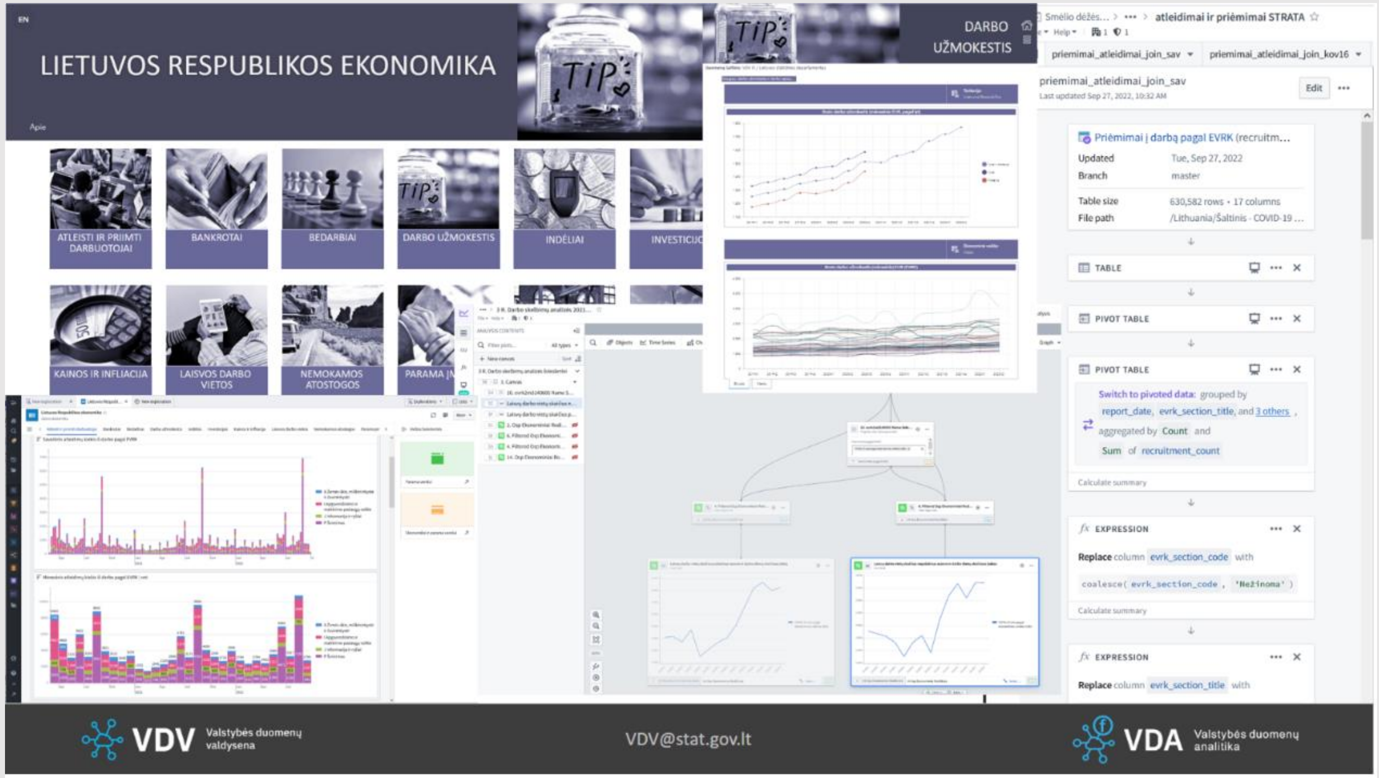
COVID-19 seniūnijų apžvalga

DATA: Tue, Sep 27, 2022

Druskininkų sav.

Example – Economy and Business in the Republic of Lithuania (the dashboard)

- Integrating data from various institutions and preparing indicators
- Internal dashboards
- Public application
  - https://osp.stat.gov.lt/ev-dashboards

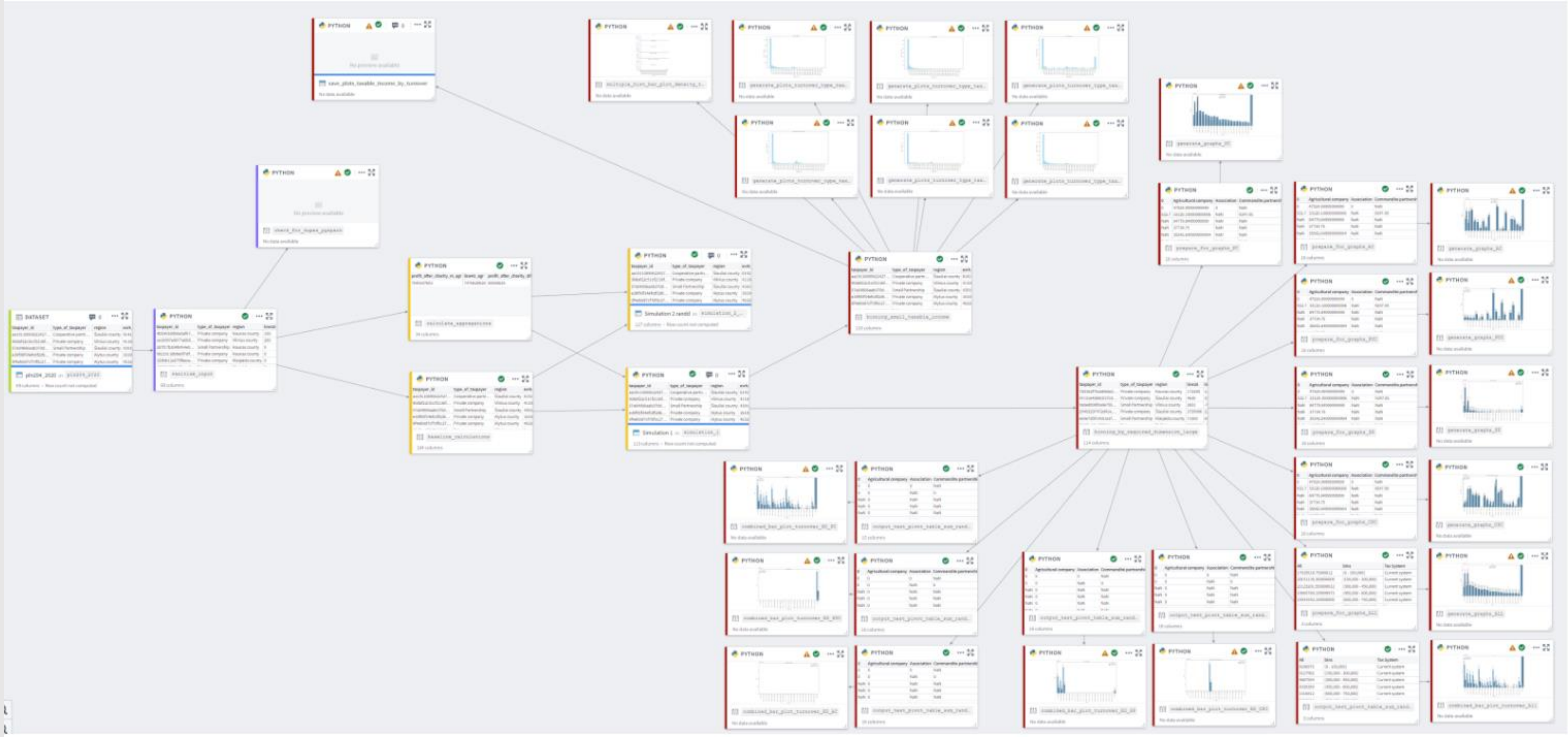


EN
LIETUVOS RESPUBLIKOS EKONOMIKA
Apie
DARBO UŽMOKESTIS
ATLEISTI IR PRIIMTI DARBUOTOJAI
BANKROTAI
BEDARBIAI
DARBO UŽMOKESTIS
INDĖLIAI
INVESTICIJ
KAINOS IR INFLIACIJA
LAISVOS DARBO VIETOS
NEMOKAMOS ATOSTOGOS
PARAMA J
Smėlio dėžės... › atleidimai ir priėmimai STRATA
priemimai_atleidimai_join_sav
priemimai_atleidimai_join_kov16
priemimai_atleidimai_join_sav
Last updated Sep 27, 2022, 10:32 AM
Edit
Priėmimai į darbą pagal EVRK (recruitm...
Updated Tue, Sep 27, 2022
Branch master
Table size 630,582 rows • 17 columns
File path /Lithuania/Šaltinis - COVID-19 ...
TABLE
PIVOT TABLE
PIVOT TABLE
Switch to pivoted data: grouped by report_date, evrk_section_title, and 3 others, aggregated by Count and Sum of recruitment_count
Calculate summary
EXPRESSION
Replace column evrk_section_code with
coalesce(evrk_section_code , 'Nežinoma')
Calculate summary
EXPRESSION
Replace column evrk_section_title with
VDV Valstybės duomenų valdysena
VDV@stat.gov.lt
VDA Valstybės duomenų analitika

Example - Analytical modules for assessing the taxable environment for self-employed individuals and small businesses.

Our information system provides environments, where R and Python code can be used to make various data analysis.

Apache Spark is used as a main database engine.



Security context in the State Data Governance Information System

- In the last two years, an established and functioning state data ecosystem has been created, with integration with over **40** state information **systems**/resources.
- Approximately **300 more systems**/resources will be integrated in the coming years.
- Our division also develops internal applications for operational data and external public dashboards.
- Currently, our information system has nearly **2000** internal and external user accounts, with number of roles, and responsibilities
- Our developers and analysts use *git* repositories, code workbooks, various analysis tools. We use **Palantir Foundry** platform, which allows us to us Python and R programming languages, GIT repositories, Apache Spark data engine.

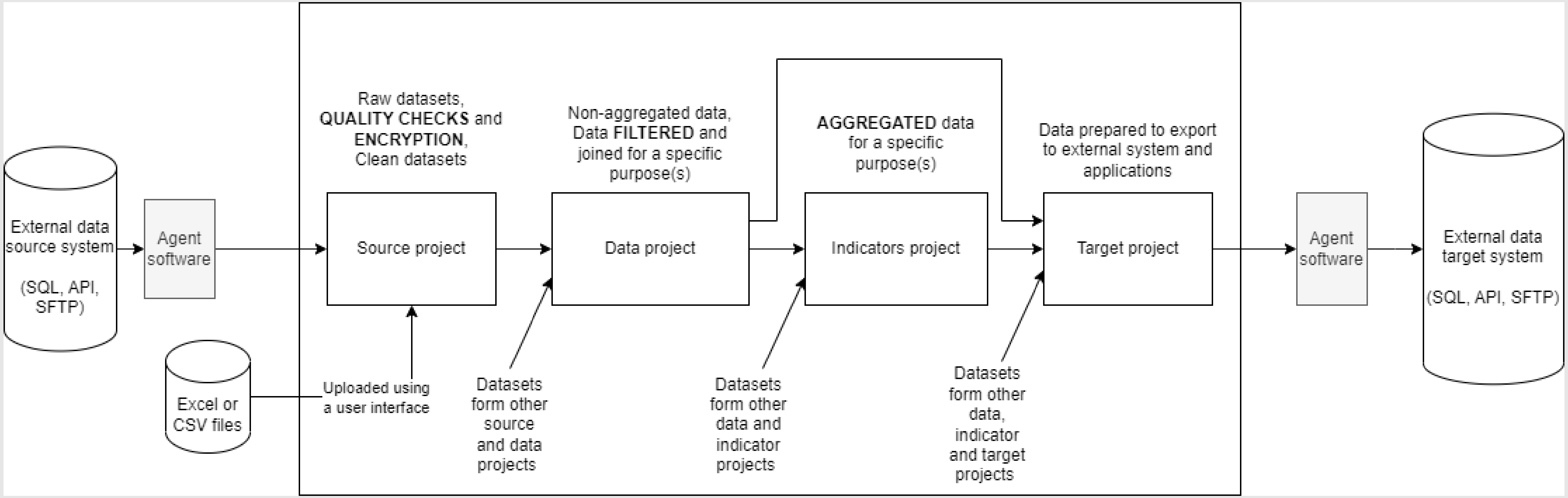
Security context in the State Data Governance Information System

- Our Information System ingests and exports data from a large number of datasets from other organizations.
- Our data processing solutions help organizations, such as municipalities or hospitals, to make data-driven decisions.
- The data from our system is published on the web.
- Many datasets are transformed to open data.
- Our IS also features sandbox projects where other organizations can make their own analysis using the integrated data and tools available within the system.

The scope of Information Security Management System (ISMS) in State Data Governance division

- There is a constant need of robust IT security solutions to reduce the risks related to:
  - large number of heterogenous systems
  - sensitive personal and health data
  - relatively large number of projects
  - relatively large number of users
  - complexity of data processing
  - data processing expenses due to large amounts of data
  - other risks, related to IT technologies (web security, viruses, social engineering, physical security, etc.)

Typical data pipeline in State Data Governance Information System



External data source system

(SQL, API, SFTP)
Agent software
Excel or CSV files
Uploaded using a user interface
Raw datasets, **QUALITY CHECKS** and **ENCRYPTION**, Clean datasets
Source project
Datasets form other source and data projects
Non-aggregated data, Data **FILTERED** and joined for a specific purpose(s)
Data project
Datasets form other data and indicator projects
AGGREGATED data for a specific purpose(s)
Indicators project
Datasets form other data, indicator and target projects
Data prepared to export to external system and applications
Target project
Agent software
External data target system

(SQL, API, SFTP)

Leadership and commitment in our division

- Employees in our division are actively committed to developing and maintaining secure processes for our information system, to provide the highest quality services to interested parties.
  - We have weekly division meetings where leading managers and specialists discuss various topics, including topics related to system security and robustness, handling of personal information, and legal topics.
  - Our division‘s technical leads and management together with dedicated security specialists develop and continually improve our IS security.
  - Our division improves IT security awarenes, by participating in training activities and security related processes, writing documentation, reporting security issues and participating in discussions on IT security topics.

Leadership and commitment in our division

- Our IT department has a 24/7 support team
- The IS developer and service provider actively informs us of important security issues and constantly improves and updates the system. 24/7 support is also available.
- At the moment, at our division we have 7 teams that hold internal sync meetings at least a few times per week (or every day when using scrum project management) and weekly team leader meetings. In those meetings, team members can discuss major problems, including IT security issues.

Actions to adress risks and oportunities

- Risks associated with our division's information system are addressed by responsible employees within our division and, if necessary, escalated to the IT security team within our department and to our Information System (IS) service provider.
- The IT division within our department assesses both risks and opportunities, analyzes reports, and actively participates in planning IT security solutions, such as software and hardware purchases and new security procedures.
- Many suggestions related to IT security are provided by our IS service developer and provider. These suggestions are discussed and incorporated into plans for IT security if necessary.

Information security objectives and planning to achieve them

- Our goal is to create secure and reliable solutions for state data governance and official statistics
- IT security is a top priority in our organization, and we consistently incorporate it into our plans.
- Our department has planning meetings where our tech. leads, IT security specialists and management includes security improvements in our plans.

Resources involved

- IT security is ensured by various parties, including:
  - management
  - lawyers
  - IS administrators
  - data engineers and analysts
  - IT support team, including dedicated IT technology and communication security specialists
  - State Data Governance Information System service developer and provider
- Organizational instruments:
  - Legal documents
  - Internal documentation and procedures
  - Publicaly available information on data security policy and security requirements, provided on our organization‘s website
- Technical tools:
  - Regular IT tools, such as firewalls, antiviruses, tools for automatic software updates, server and personal computer security managements tools, personal computer and mobile phone security tools
  - Dedicated applications for project and data access rights management (developed by our engineers)
  - Incident reporting systems (Axence nVision, Palantir Foundry)
  - Tools for log analysis (ELK stack, FortiSIEM, etc.)
  - Automatic data scanning tools to identify sensitive data (provided by Palantir Foundry)

Support - competences

- Understanding the importance of security
- Using best practices in secure coding and data processing
- Following general IT security guidelines
- Understanding physical security (such as clean desktop policy, physical access to the building, etc.)

Support - awareness

- Security awareness is achieved by informing our employees and users on various security topics related to our systems, and information technologies in general.
- Our employees get e-mail alerts from our IT department and information system.
- National Cybersecurity Center provides our organization with training material, such as interactive training courses

Support - communication

- Our department has tools to report IT security issues to responsible resources in our IT division for immediate further action (Axence nVision)
- State Data Governance IS has tools to report issues to IS service provider for immediate further action (Palantir Foundry Amplify ticket system)
- IT division and IS service provider constantly notifies our employees on various IT security topics via e-mail messages containing security recommendations and/or reports, such as new security vulnerabilities
- Data pipeline developers can create alerts based on data health checks (Palantir Foundry)
- In urgent cases, IT security team or IS service provider can be contacted 24/7

Support – documented information

- IS documentation created by our specialists covers topics related to data and source code security such as user roles, data security markings, data health checks, encrypting and exporting data, data pipeline robustness, and other topics. This documentation is constantly improved
- IT security guidelines (e.g. personal data safety guidelines, secure coding guide, etc.) are provided by IS service provider in the product documentation (Palantir Foundry)
- Interactive online training courses, security news and yearly IT security reports are provided by National Cyber Security Center on the web
- Our employees sign legal security documents when onboarding

Internal audits and evaluations

- Our department does random, periodic and constant IT security audits and evaluations (e.g., active checking of project data and source code access restrictions, making code reviews, etc.)
- Our division has dedicated employees for division’s information system audit log analysis

- General security audits are made by IT department

Performance evaluation

- Results of internal security audits and any security incidents are reviewed by responsible IT security specialists, tech. leads, management and IS service provider
- We constantly plan and take appropriate actions to reduce the number of security incidents and reduce security risks in our department

VALSTYBĖS
DUOMENŲ
AGENTŪRA

We are grateful for the opportunity to participate in the Twinning project, and we hope that this mission will enhance IT security in our organizations.