3 Assessment of the IT infrastructure in the State Employment Agency

The State Employment Agency in Latvia (SEA) recognises its IT backbone as a key prerequisite to deliver good services to its clients and has been able to develop an IT system that meets most of its current vital needs despite limited resources. The IT backbone supports the main tasks of the SEA, such as registering jobseekers, managing services and measures, registering vacancies, and matching jobseekers and vacancies. However, the system is not fully efficient and modern, for example in terms of solutions for data analytics, quality, exchange and protection. The IT architecture of the SEA is not entirely future-proof regarding the potential needs to develop new functionalities.

3.1. Introduction

This chapter reviews and assess the overall digital infrastructure and data management processes in the State Employment Agency (SEA) in Latvia, focusing on the architecture of its operational IT system, related processes and practices, and data analytics solutions. The specific applications and tools to support the core tasks of a public employment service (PES), like matching jobseekers and vacancies or profile jobseekers, are reviewed in the next chapter (Chapter 4).

The SEA recognises its IT backbone as a key enabler to deliver good services to its clients, but the scarce resources limit the modernisation and digitalisation of the SEA. In general, the overall IT backbone of the SEA is able to cater for most of the current needs of the SEA. The main operational IT system BURVIS supports the main tasks of the SEA, such as registering jobseekers, managing services and measures, registering vacancies, and matching jobseekers and vacancies. However, the IT system does not include sufficient data management functionality, particularly in terms of data quality management, data sharing solutions with external parties and data analytics. In addition, the architecture of the current IT system is not fully future-proof as the monolithic approach can add unnecessary complexity into the development of new functionalities.

This chapter starts by discussing the business needs of the SEA that have implications for the IT infrastructure, as well as the SEA's strategy and development processes to strengthen the IT backbone, followed by a review of the IT architecture (server environments, system interfaces, software), the data analytics solution, and processes and practices related to the IT infrastructure (backups, data quality management, system and data security).

3.2. The SEA recognises the importance of its IT infrastructure, but can assign only limited resources to its development

The IT infrastructure is only fulfilling its purpose if it meets the business needs of the organisation and helps to achieve the strategic objectives. This section first discusses the business needs that PES have and how these needs translate into the requirements for the IT infrastructure. Subsequently, the section reviews how the SEA has considered the implications for its IT system in its strategy, and which resources and capacities the SEA possesses to develop the IT backbone that would help to achieve the strategic objectives.

3.2.1. The IT infrastructure of the SEA needs to support jobseekers, employers, co-operation partners, as well as the SEA staff in different functions

Modern digital infrastructure is key for PES to meet the diverse needs of jobseekers, people at risk of job loss and employers effectively and efficiently (OECD, 2022[1]). An advanced and well-integrated IT backbone encompassing both back-office infrastructure and applications and interfaces for clients, partners and other stakeholders in the wider "PES ecosystem" enables to maximise the PES administrative capacity and the value-added of its services.

Digitalisation in PES is not a goal in itself but needs to support the business model of the PES as well as possible, facilitating the PES to fulfil its core objective of supporting jobseekers, people at risk of job loss and employers. It means that the main function of the IT infrastructure in the SEA is to enable its employment counsellors to fulfil such tasks like jobseeker registration, identification of support needs, development of individual action plans, case management, vacancy registration, matching jobseekers to vacancies, and the provision of different active labour market policies (ALMPs). In addition to these key elements of actual service provision, the IT infrastructure should support counsellors in managing their

tasks, such as monitoring and managing their portfolios of jobseekers or monitoring the process of ALMP provision.

The IT infrastructure in PES also needs to facilitate other business processes in the organisation, besides the delivery of services to jobseekers and employers. The management of the SEA needs to be able to monitor the processes and the performance of the organisation, including by different divisions and functions, to be able to make operational decisions on the organisation of service delivery, as well as design the SEA strategy. The SEA management needs to be supported by data analytics both by having systematic key information accessible via the IT infrastructure, and needs-based information presented by the statisticians and analysts in the SEA. The latter staff needs data analytics solutions to cater to the needs for evidence of the SEA management, as well as other internal and external stakeholders and the public more generally.

The SEA IT infrastructure also needs functionalities vis-à-vis external organisations and stakeholders. Functionalities for data exchange are necessary to receive data from other registers to be able to provide better services and generate more credible evidence, and share data to support other public sector organisations in providing their services and generating evidence. The data exchange solutions should ideally aim at the “once-only” data collection principle that is progressively being implemented in the OECD countries and for example encouraged by the European Commission in its Single Digital Gateway Regulation to support Digital Single Market in the context of cross-border data exchange.[1] In addition, secure data exchange channels are relevant to share data with external researchers.

For more efficiency and user-friendlier service provision, PES services need to be accessible for jobseekers, people at risk of unemployment and employers also digitally for independent use, i.e. without the intermediation of PES staff. The trend in client needs for digital self-service platforms was elevated particularly along with the containment measures and limits to in-person services put in place due to the emergence the COVID-19 pandemic (OECD, 2022[1]; 2022[2]; 2021[3]), and the demand for such digital services has remained higher than before the pandemic. Similarly, the higher need for digital communication channels between the PES and its clients has remained. Hence, the digital infrastructure of the SEA needs to facilitate receiving information from jobseekers and employers directly, sharing information from the SEA to jobseekers and employers and communicating with them in real time (e.g. receive counselling services digitally). Jobseekers should be able to, for example, initiate their registration independently, search for jobs in the SEA database, access digital resources to improve their application documents and employability, apply for jobs and ALMPs. Employers need to be able to access such functionalities like uploading vacancies, searching for suitable candidates in the SEA database and applying for the SEA services and measures. Similar digital self-services could be relevant for other stakeholders and the SEA partners, such as for training providers to apply be a partner organisation for the SEA.

Besides providing a general backbone for PES service provision, digitalisation has a potential to vastly enhance the effectiveness and efficiency of PES services. Such benefits can be reaped if the digital solutions aim to go past simply digitising otherwise non-digital processes, but aim to digitalise the processes, i.e. make the processes faster, leaner and smarter via automation, data linking, advanced data analytics and Artificial Intelligence (AI) algorithms. However, adopting new digital tools using advanced analytics and AI set additional requirements on the technological stack of the PES, so that these new tools could be integrated into the overall IT infrastructure seamlessly. The next sections and Chapter 4 discuss how the IT infrastructure in the SEA meets these (future) needs as well as the needs listed above.

3.2.2. The importance of the IT backbone is highlighted in the SEA strategy

The SEA does not have a dedicated strategy or overall concept for its digital infrastructure and services. Nevertheless, the IT infrastructure is addressed extensively in the SEA’s general strategy for 2021-23 (Nodarbinātības valsts aģentūra, 2021[4]). One of the three key objectives that the SEA has set to itself, is

to promote the SEA as the leading partner in personnel recruitment in Latvia,[2] under which the strategy highlights the sub-goals of promoting the SEA's portal for matching jobseekers and vacancies (CVVP portal), improving the services for employers e.g. by a forecasting platform for labour supply and demand and improvements in digital services, and improving labour market information system for the public (labour market indicators, labour market forecasts, training opportunities, skills and occupations in demand, career guidance information, etc.).

The SEA's three-year-strategy also rightfully recognises the SEA's IT infrastructure as a key prerequisite to achieve the set objectives, and increase the effectiveness, efficiency, accessibility and user-friendliness of the SEA's services. To maintain and improve the IT backbone, the SEA strategy highlights the needs to implement more modern and harmonised approaches within its key systems and approaches, such as in terms of a uniform approach to document management, data storage, data security, and software and hardware management. Furthermore, the strategy includes ambitious intentions for the future in terms of strengthening the SEA's analytical capacity via modern data analytics and visualisation tools, adopting AI solutions to be able to process Big Data in the SEA's digital tools, and making more of the SEA's data available for the clients and public as open data.

The SEA's strategy justly acknowledges that developments in the IT infrastructure are not separate processes but need to serve the purposes of the SEA's core business processes and services, and consider the needs of end-users, such as the SEA staff. The SEA staff is expected to have sufficient digital skills and improve them regularly, including on cyber security to reduce security risks. Furthermore, the SEA aims to facilitate the working modalities that have considerably changed since the COVID-19 pandemic, comprehending the needs to ensure IT equipment and appropriate technical solutions for service provision both via face-to-face and remote channels.

3.2.3. Tight resources limit strategic planning

Although the SEA's strategy has identified the medium-term needs for improvements in the IT infrastructure, the budget to implement modernisation is limited and often project-based similarly to the overall SEA budget (see Chapter 2). The IT team in the SEA is very small, aiming to manage day-to-day maintenance issues and co-ordinate and manage IT developments that are implemented by external providers. The funding for these outsourced IT developments from the state budget is low and covers the urgent maintenance, updating and improvement needs, but does not enable major developments. The additional funding from the European Union (EU) resources (such as European Social Fund, Recovery and Resilience Fund or European Regional Development Fund) help cover additional needs for IT developments, but these projects are then with a limited scope and do not cover longer-term maintenance costs of these newly developed IT solutions.

While combining the resources from the state budget and the EU funding has enabled the SEA to develop an IT infrastructure that covers most of its business needs (see the next sections), this funding model is posing serious risks for a more sustainable modernisation. The scarce funding has been one of the main reasons the SEA has not been able to implement all necessary data exchanges with the external registers (Section 3.3.3) or adopt modern and efficient data analytics (Section 3.4) and data quality management solutions (Section 3.5.2). Furthermore, the volatile funding model poses a threat of ending up with a patchwork of digital solutions that are not fully integrated, compatible or aim at the same strategic concept for the digital backbone.

In addition to the insufficient financial resources, the SEA faces challenges in adopting and implementing new digital tools and technologies due to a lack of skills to develop and implement modern solutions, some resistance from the SEA staff to take up such solutions, and difficulties in understanding how the technology works or explaining how it works to staff and jobseekers (responses from Latvia to the OECD questionnaire on PES digitalisation launched in March 2023). These challenges have for example contributed to the changes in the digital tool to profile jobseekers to segment them into appropriate service

streams, which since 2019 significantly simplified and turned into a manual process to be conducted by the employment counsellors (see Chapter 4).

The challenges related to staff skills, comprehension of and resistance to advanced digital solutions is potentially connected to low and constrained wage structures for the SEA staff (see Chapter 2). As such, the low wages do not hinder only service provision to the SEA's customers by the front-line staff, but also the high-level and strategic changes in the overall IT backbone of the SEA, compromising the key underlying prerequisite for efficient and effective services.

3.2.4. End-user needs are not systematically considered in the IT developments

The major decisions on adopting new or adjusting existing digital solutions are taken jointly by the Ministry of Welfare (MoW) and the SEA (i.e. decisions going beyond general upgrading, bug elimination or small-scale fine-tuning). This is an appropriate decision-making process, as MoW is responsible for the high-level design of services for jobseekers, people at risk of unemployment and employers, as well as needs to drive the process of achieving a sufficient budget for the IT developments from the state budget and external funding sources. In addition to the high-level management, the SEA also involves its IT department and the departments in charge of the relevant services and measures in the decision-making to take on board the SEA's business needs, as well as IT capabilities.

Within the process of making decisions on adopting or adjusting digital tools, the potential end-users are sometimes consulted to ensure that the digital tools meet the users' needs. For example, to make decisions on the objective and design of the new skills profiling tool (see Chapter 4), the SEA management consulted its front-line staff to understand which kind of insights from the new tool would help them best to provide career management advice to jobseekers. Nevertheless, the consultations with end-users have not taken place systematically regarding all end-users (such as employers or partner organisations) or all new digital developments. The end-users are also not involved in the later stages of the IT development processes. Only the SEA front-line staff is involved more systematically in the testing stage.

The key issues considered during the decision-making processes for new IT developments concern the value-added and objectives of the new solutions, data protection issues, impact on the end-users (including the SEA staff), and integration with the existing IT infrastructure and service provision processes. Yet, monitoring of these different aspects after a new solution is adopted, takes place only ad hoc if at all. The SEA has monitored so far for some of its new solutions such indicators like the take-up rates by users, efficiency gains in processes and some aspects of service quality. The SEA has no experience yet in monitoring and evaluating the impact on labour market outcomes (e.g. the effects on jobseeker employment rates), end-user experience and satisfaction or cost-savings of its digital solutions. The technical support provided by the OECD and the European Commission to MoW and the SEA also includes a component to build capacity in MoW and the SEA to monitor and evaluate ALMPs, including digital tools used for ALMP provision (OECD, 2022[5]).

3.2.5. IT developments are tied to the capacity of the development partner

The IT developments of the SEA are fully outsourced to external providers, while the SEA IT department is driving the procurement and development processes. The IT department is involving the relevant departments in the SEA in charge of the respective business processes needs-based, above all in identifying and describing the business needs, taking decisions related to functionality and design, as well as testing. As such, the business units in the SEA are not clearly assigned to be the drivers for the digital improvements in the services they should generally be in charge of. Furthermore, smaller-scale changes in the IT infrastructure are decided internally in the IT department of the SEA. In general, the development practices have some features of agile development methodology, but these are not fully aimed to be or implemented as such.

The developments in the SEA's IT infrastructure are implemented mostly by the same external provider in Latvia (UNISO). This provider is continuously winning the tenders as it complies well with the requirements that the SEA has set in the procurement documents. The co-operation with UNISO has turned into a long-lasting co-operation, and this provider has been the single developer regarding the key parts of the IT infrastructure (BURVIS, see more in the next sections).

As the SEA is essentially working with one provider and the funding for the developments is project-based, it has been difficult to achieve a continuous improvement process for the IT infrastructure. The project-based funding leads to short development processes, which have led to compromised development quality. Working with a single provider also poses capacity constraints, as the resources that can be mobilised in short notice are limited. As of early 2023, the SEA had some 40 urgent improvement requests in the development list (e.g. requests for changes from the SEA internal users, including those mediated by the internal users, but coming from the external users, such as a request from external partners to publish their vacancies in the SEA portals) that were expected to take still some time to be implemented. As a solution, the SEA is analysing the feasibility and possibilities to work with additional IT providers.

3.3. The technology stack in the SEA does not exhibit major challenges in supporting service provision

The key elements of the IT backbone, such as the high-level architecture of the infrastructure (the set-up of databases, user interfaces, interfaces for data exchange between internal databases and external registers), the server environments and the programming languages used for the individual components define the capacity, speed, security, flexibility for adjustments and the potential functionality of the overall IT infrastructure. This section reviews the key elements of the IT backbone in the SEA.

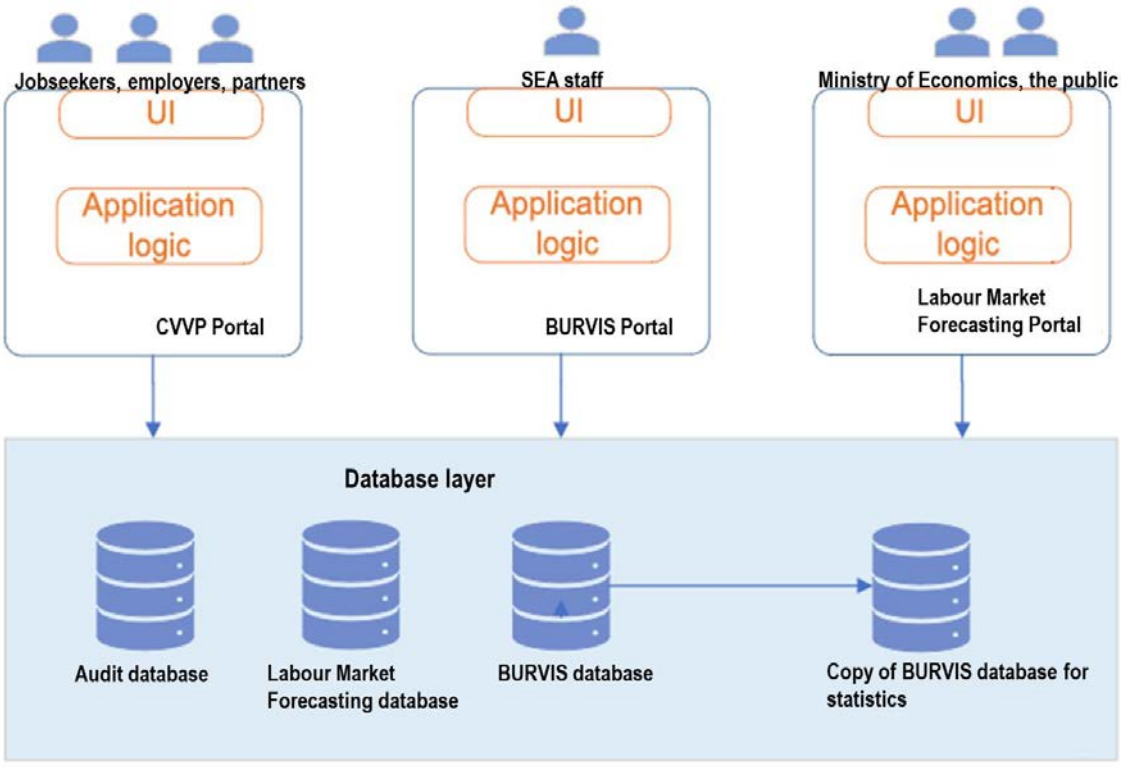
3.3.1. The high-level IT architecture generally meets the operative needs of the SEA

The high-level IT architecture in the SEA contains three portals, each of them having a dedicated interface and application logic, as each of them serves ad different purpose and supports a specific group of users (Figure 3.2):

- BURVIS portal – a portal for the SEA internal staff to provide services for jobseekers and employers. It is the main portal linked to the BURVIS operational database which is the official register of unemployed and vacancies.[3]
- CV and Vacancy Portal (CVVP) – a public self-service portal for the SEA clients (jobseekers, employers) and co-operation partners available at https://cvvp.nva.gov.lv/ to make some of the SEA services digitally available for independent use, e.g. for jobseekers to find suitable vacancies. The CVVP is also linked to the BURVIS operational database.
- Labour Market Forecasting Portal – a dedicated tool developed in co-operation with the Ministry of Economics to disseminate labour market forecasts to the public, available at https://prognozes.nva.gov.lv. This tool is not used in the process of day-to-day service provision and is not linked to the BURVIS operational database, but its own separate database containing only aggregate statistics from the BURVIS database and aggregate data from external sources (i.e. no personal data). As this tool is less linked to the operational IT system of the SEA, it is discussed separately in Section 3.3.4.

The BURVIS and CVVP portals store data in and fetch data from the BURVIS operational database, which is the main operational database in the SEA to record processes to support jobseekers, people at risk of job loss and employers. In addition, the SEA IT infrastructure includes an audit database that logs all actions ("clicks") in the BURVIS and CVVP user interfaces for data and system security management, and a copy of the BURVIS database that is used above all for data analytics (see discussion in Section 3.4 on the data analytics solution).

Figure 3.1. The high-level architecture of the SEA includes three main portals to serve the different user groups and purposes



Note: UI – user interface. SEA – State Employment Agency. In addition to the IT infrastructure depicted on the graph, the SEA has another system for document management to support BURVIS system and accounting, called Horizon.
Source: Authors' work based on inputs from the State Employment Agency.

The overall architecture of the operational IT system (the BURVIS portal, CVVP portal and BURVIS database) generally meets the current operative needs of the SEA. This architecture also allows meeting the potential future development needs, although it is not completely modular and thus not as flexible as it could be. It means that the current architecture does not prevent developing new functionalities, but these developments can be slightly more difficult to implement than in a more modular set-up. Nevertheless, modifying the system architecture that has been developed over a longer time period is usually more costly than the value-added from the changes. Yet, also the more granular architecture of the system (particularly the BURVIS portal and database) might not be sufficiently flexible to accommodate even some of the minor changes in its functionality or these changes can be unnecessarily cumbersome, which has also been one of the reasons for the long list of improvement needs to accumulate and strain the capacity of the external development partner (see Section 3.2.5).

3.3.2. The server capacity and set-up are sufficient

Production environment uses a single-server approach

All applications (i.e. portals, databases) in the production environment of the SEA's IT infrastructure (i.e. live environment where the different applications are used by the SEA staff and clients for their intended purposes) are running in their dedicated virtual servers:

- The application server for the BURVIS portal
- The application server for the CVVP portal
- The BURVIS database server
- The audit database server (logs)
- The web frontend server
- The server for the Network File System (previously used to upload CVs to the CVVP portal) and an old database (old logs)
- The server for the copy of the BURVIS database (a nightly database dump for data analytics and backup)
- The database server for bug tracking / ticketing system

The set-up of eight servers for the production environment is likely able to meet the current needs of the SEA, considering also that the new physical servers were taken into use only two years ago. Nevertheless, the SEA staff sees some processes in the user interfaces as being too slow (e.g. when using the jobseeker and vacancy matching services), which can refer to the low capacity of the server as one possible underlying reason. In addition, the current set-up might not be sufficient in the future in case additional interfaces/applications would need to be adopted, the volume of the databases would continue growing over the years or if the number of users of the system would grow substantially.

Indeed, the volume of the SEA databases is growing as historic data are not deleted or archived. The data volume is growing particularly fast in the audit database, which is about 400 gigabytes as of March 2023. The full details of log data are retained in the audit database without anonymising or dropping information that would be less relevant after many years have passed. Regarding the BURVIS database, the regulation foresees a required time to keep the data in the system (different deadlines depending on the type of data), prescribing anonymisation or deletion of data afterwards. Yet, these processes have not been implemented in the BURVIS database either.

It will be possible to extend the capacity of the current set-up by adding more capacity to the servers. However, this set-up cannot be extended beyond the capabilities of single server(s) as there is currently no multi-server capability in the SEA which would allow adding more servers with load balancers (mechanism to redistribute network traffic between servers to maximise speed and capacity). Yet, moving from a single-server set-up per application to a multi-server approach would usually also require changes in the applications themselves, so that these would be able to function in a "stateless" mode. Similar to a multi-server approach with load balancers, a modern system could be operated in containers (such as Kubernetes, which is an open-source container orchestration system for automating software deployment, scaling, and management). A system with containers requires similar application functionality as the multi-server approach, and additionally knowledge of container configuration and management.

Test environments enable thorough testing before adoption

In addition to the production environment, the SEA has two test environments that enable comprehensive testing before deployment (releasing updates in the applications to the production environment). One of the testing environments is a local test environment without connections to external registers and systems, and the other one is a test environment with full services with authentications and connections relevant to

communicate with external systems (i.e. enables testing the full functionality of the SEA applications). The two testing environments are hosted in five additional servers (a single-server solution similarly to the production environment).

Having separate environments for production and additionally for the different stages of testing is generally sufficient for the current needs of the SEA. A more granular approach for the testing stages, such as separating “development testing” and “acceptance testing” could further enhance testing before production.

The SEA IT experts deploy new software versions in the production environment manually. In the test environments, full CI/CD pipelines exist, which are automatic process that drive software development through a path of building, testing and deploying code. Yet, the SEA has decided not to use a similar automated process in the production environment to mitigate risks. A trade-off between the costs and benefits of automation of deployment processes is key in deciding whether to automatise or not. The manual deployment is a good practice in case new software versions are infrequent, such as once in a few months and which is also currently the case in the SEA. If new versions have to be deployed frequently, such as daily, the deployment process should be automatic.

The servers of the SEA are hosted in external data centres

All the SEA’s servers are hosted in the common data centre maintained by the State Social Insurance Agency under MoW (i.e. it is a physical storage of servers), and since 2023 partially in the data centre of the Latvian State Radio and Television Centre. The staff in of these external data centres maintains the physical servers and VMware does the virtualisation of the servers, so that the SEA IT staff could maintain the virtual machines. This set-up meets the current needs of the SEA as the approach of centralised data servers reduces the effort required from the SEA to maintain servers, as well as the number of people needing to access the servers, which in turn improves security.

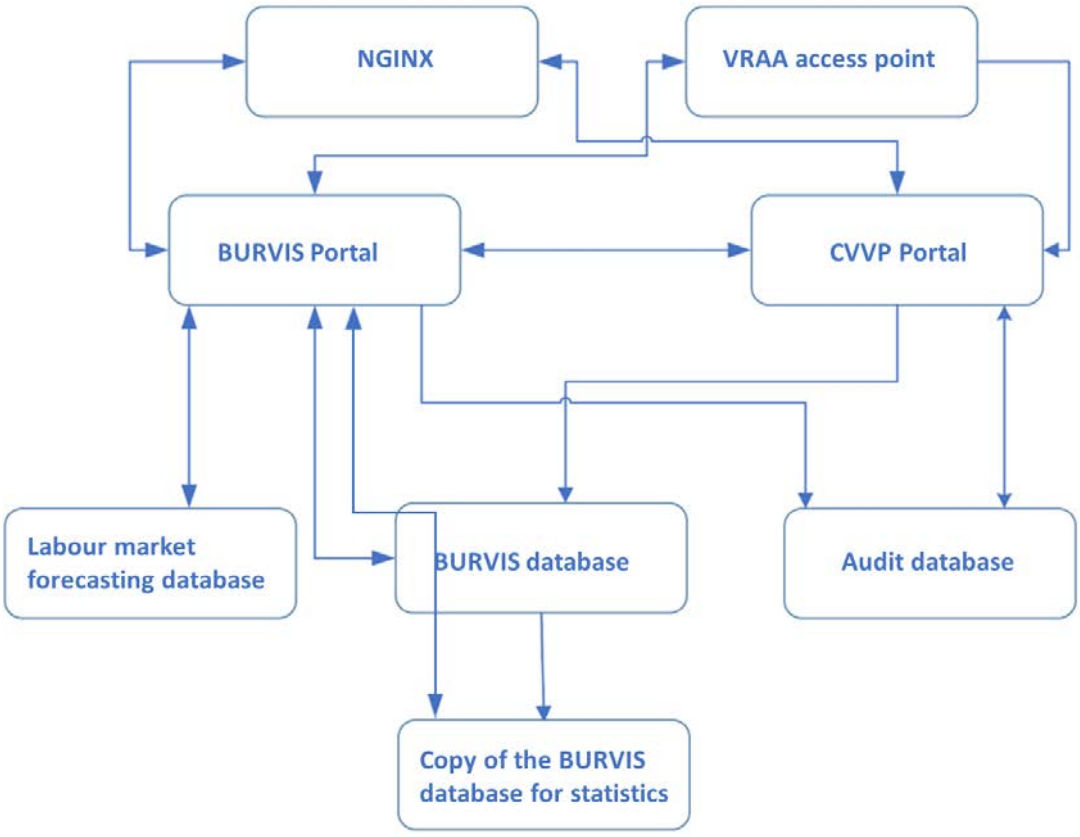
SEA does not currently use cloud services as it has no explicit necessity to host its software in the cloud. Yet, using the cloud services could offer some benefits beyond just hosting applications or servers. For example, some data analytics tools are readily available in the cloud without needing immediate investment in licenses or hardware. The (public) cloud also offers more possibilities for scaling up applications in case this would be relevant for the SEA in the future.

All the SEA servers run in the same network as there is no separation at the network level, which can pose a security risk. If an attacker gains access to any of the server in the network, the attacker also has easier access to the other servers in the network than in the case of having network level separation. Systems with more divided networks (sub-networks) enable to limit such risks.

3.3.3. Data exchanges are not fully secure, automatic or meeting the needs of service provision

Figure 3.2 is depicting the connections between databases and system components in the SEA. The BURVIS and CVVP portals are well connected to most other components in the IT infrastructure of the SEA. The connections to external systems are implemented either directly via the BURVIS or CVVP portals, or via the VRAA access point (service provided by the State Regional Development Agency for the providers of public services in Latvia). In total, the core IT infrastructure depicted in Figure 3.2 have in total 20 interfaces to exchange data, some of which are with external registers or IT systems (user authorisation, addresses database, EURES for vacancy mediation across EU, data from other registers on jobseekers and employers used for service provision), and some with additional internal systems of the SEA (bug tracking, mail, document and accounting register Horizon, website of the forecasting tool), seven external classification systems, and three file sharing solutions.

Figure 3.2. The BURVIS and CVVP portals are well connected to most other components in the IT infrastructure of the SEA



Note: DB – database. NGINX – an open source web server in the State Employment Agency. VRAA – the State Regional Development Agency.
Source: Authors' work based on inputs from the State Employment Agency.

Most of the interfaces to enable connections between the different systems are separated to a specific API (Application Programming Interface) layer, but some interfaces access databases directly. Such direct access should in general be avoided as it impacts system security and development. In terms of system security, using a direct access to a database reveals the inner structure of the database, which can disclose information to potential attackers and provide them a direct access to actual data. In addition, such direct connections tie the database structure to the current implementation, making it difficult to change the implementation without affecting the users of the interface. These types of problems are usually mitigated by an API layer (preferably through API Management) which provides an additional layer of security and hides the implementation details, enabling the underlying implementation to be changed if needed while keeping the API the same for the external partners.

Some of the data exchanges are fully automatic as background processes during service provision while some need manual initiation or are launched only daily, monthly or ad hoc. Yet, more automatic processes could make service provision more efficient. For example, during the jobseeker registration process in the BURVIS portal, some data on the jobseeker are automatically taken from external registers via the access point as a background process, while for some information the employment counsellor has to launch the dedicated query. Nevertheless, some of the manual processes have been implemented for data protection reasons, e.g. the query on disabilities is launched only if the jobseeker first themselves tell the counsellor that they have such status.

The SEA has been able to establish data exchanges with many key registers in Latvia to increase the effectiveness and efficiency of service provision. For example, the SEA receives data from the State Revenue Service (data on enterprises, employment data), the State Social Insurance Agency (benefit information, employment data), the Social Assistance Administration Information System (data on social services by municipalities), the Citizenship and Migration Affairs, the Register of Enterprises, the State Land Service, the State Labour Inspectorate, Road Traffic Safety Directorate (data on driver's licences).

The SEA staff has identified further needs for data exchange that have not been possible to implement yet. For example, information on education in the BURVIS database is currently based on what the jobseeker says or provides documents for, but not on the register data that could be received from the Ministry of Education. The register of State Revenue Service includes occupation data, but these are not available for the SEA and thus, again, the counsellors need to rely on information from jobseekers themselves for information on their previous occupations, which makes the process less efficient and the data potentially less reliable.

Establishing new data exchanges is often cumbersome as each new data exchange needs a legal base, a contract between the register owners, and some technical capacity in both sides to implement the data exchange. The latter step in establishing new data exchanges is particularly challenging for the SEA as its development capacity is tied to the limited capacity of the development partner (see Subsection 3.2.5). As the IT infrastructure in the SEA faces a list of improvement needs to increase the efficiency and effectiveness of the SEA services, needs for new data exchanges need to compete with other urgent needs. For example, the data exchange with the Ministry of Education has been already agreed on, but the implementation has not taken place yet due to the limited resources.

3.3.4. Programming languages and other technologies do not include legacy systems

The technology stack of the SEA includes some technologies that are somewhat dated as of Q2 2023. However, as these versions are still supported, the current set-up does not (yet) cause any challenges for the system.

More specifically, the BURVIS and CVVP portals are built in Scala and Java (mostly Java 11, while an interface to external registers via the VRAA access point uses Java 8 for which the release date dates back almost ten years, and the newest version in Q2 2023 is Java 20). The BURVIS database, the copy of it and the audit database are PostgreSQL 13 databases, i.e. use a quite recent version (released in 2020, the latest version of PostgreSQL as of Q3 2023 is PostgreSQL 15). In general, PostgreSQL is a fitting choice for the needs of the SEA and it does not pose any specific risks compared to other database alternatives.

Other technologies in the technology stack of the SEA include: tresql, querease, mojoz, sbt, nginx, angularjs, bootstrap, less, pug, akka, akka-actor, akka-http, akka-stream, coffeescript, jasperreports, bower, npm, grunt, jasmine, node.js, javascript, and bash. All of these are common software development tools and do not create any risk for further development.

3.4. Data analytics solution is not supporting efficient knowledge dissemination

Operational IT systems are generally not fit to support data analytics functionality and developing an additional technology stack for that purpose is necessary. This section reviews the technology and practices concerning the data analytics solution in the SEA.

3.4.1. The statistics database is not restructured to support data queries

As data needs for operational functions and data analytics usually differ, data in the operational database are generally not fully fit for data analytics purposes. As such, modern IT systems include additional layers for data analytics purposes, which as a minimum include an ETL or ELT processes (Extract, Transform, Load) and a Data Warehouse or Data Lake solution (a data repository for data analytics purposes), although additional layers might be sometimes relevant for better data management. These solutions enable to prepare the data for analytics purposes via automatic processes (e.g. restructuring, cleaning, processing duplicates, reclassifying and coding, and addressing other data quality issues, as well as linking data from different data sources), making the production of statistics, analysis and research more efficient and effective.

The SEA does not have a dedicated IT solution to prepare the operational data for data analytics, as an exact copy of the BURVIS operational database is used to produce statistics and query data for analytical and research purposes. The copy of the BURVIS operational database is made every night without any modifications, including no changes to the data structure to support queries better.

The current solution does not enable including or processing data from additional sources in the statistics database, as the content of the database is overwritten nightly. Yet, additional data could be relevant for analytical purposes, even if these are not relevant for direct service provision and do not need to be included in the BURVIS operational database. Additional data could for example benefit monitoring and evaluation of ALMPs in terms of using a wider range of data on the labour market outcomes for the jobseekers.

Updating data by overwriting these in both the BURVIS operational database and the statistics database means that it is not possible to observe possible changes in some of the characteristics of specific jobseekers in time within the statistics database. For example, over time a jobseeker can attain a higher education level, they can change their geographic location and the SEA branch they visit, they can learn new languages and gain new IT certificates, etc. Yet, in the operational and statistics database, only the last value of these variables is available, which can limit or compromise some types of analyses (e.g. when the characteristics of the jobseeker at a certain point of time are relevant, and not only at the most recent values for these, such as when conducting counterfactual impact evaluations of ALMPs). The changes in data are recorded in the log data in the audit database, but this is not available for the statistics department and would be even less fit for statistics purpose than the BURVIS database.

The statistics department has worked around the issue of overwritten data by saving the monthly key outputs of personal level queries in Excel files externally of the statistics database. This enables the statisticians link the different Excel files to see changes in jobseekers' characteristics. However, the practice raises additional concerns regarding data protection (processing personalised data and saving these externally from the dedicated database, i.e. also the access to the respective server and folders need to be at least as restricted as the main statistics database). In addition, the changes are not always traceable in case the ID code of the jobseeker changes. For example, a change in the regulation in 2017 enabled Latvian citizens to receive a new ID code that would not contain their birthdate (Office of Citizenship and Migration Affairs, 2020[6]).

Despite the shortcomings of the statistics database, the current solution is better than querying data directly from the operational database, as currently statistics production does not interfere with the capacity and functionality of the operational BURVIS system, while still enables to use up to date data for statistics. Furthermore, building such a system has incurred less expenditures for the SEA compared to a fully-fledged modern data analytics system.

3.4.2. The solution for querying data for statistics and analysis is limited and inflexible

Modern solutions for data analytics include Business Intelligence (BI) tools that enable retrieving data from the Data Warehouse or Data Lake solutions, as well as analyse, transform, present and report data, including via data visualisation. BI tools facilitate automating routine data processing for statistics and analysis, as well as disseminate data via interactive tailored dashboards for different user groups.

The SEA has not adopted a BI solution for statistics purposes, but has developed SQL queries within the BURVIS portal that staff in the statistics department can launch. The staff has limited flexibility to adjust the criteria of the built-in queries (above all the dates to establish the time interval for data to be queried) and has developed some capacity to create elementary new queries, although the external developing partner is mainly developing and adjusting the built-in queries. The capacity of the external developer is limited (also as other types of developments often take priority), but some adjustments of the queries have been possible over the years to accommodate changed needs for statistics.

The built-in queries in the BURVIS portal return individual level data tables in (large) MS Excel files, i.e. the queries do not enable aggregate statistics reports or data visualisation. All data processing to produce statistics and analysis takes place thus externally from the SEA IT infrastructure, in MS Excel. The statistics department applies manually essentially the same methods to derive the regular statistics products each month, quarter and year. Thus, the current solution is inefficient due to low level of automation, ineffective due to low flexibility to query data, and error prone due to high level of manual processing.

As the data in the statistics database are not optimised for data analytics purposes (see previous section), all relevant computations have to be made during the query process. This set-up makes the queries slow, and it would be even more problematic in case aggregates and visualisations would be implemented within the same set-up. In such case, many calculations would need to be made repeatedly, even for more historic data. As the volume of BURVIS database increases, these calculations would get slower over time and require more computing power.

Due to the inflexible set-up of the data analytics solution and its functionality to only produce MS Excel tables with individualised data, the solution is accessible mostly for the staff in the statistics department and not all user groups in the SEA that would need analytical information. The data visualisations and reports compiled in MS Excel might suffice the key needs for information for the SEA management, but not the needs of the SEA branch offices that need more granular and up to date information. For example, the SEA offices would like to see the workload of their office by counsellors to manage the distribution of workload actively. The counsellors would like to have the up-to-date overviews of their portfolios of jobseekers. Currently, the counsellors can only query their clients (jobseekers, employers) to look up their details in the BURVIS operational database, so even compiling lists of jobseekers that could be invited to an event of the SEA have to be prepared by the SEA statistics unit.

3.4.3. The SEA's statistics department makes a lot of effort to share data and statistics with external parties

To share statistics and data with external stakeholders, the only automatic data exchange is currently set up between the BURVIS operational database and the database of the Labour Market Forecasting Portal, enabling sharing aggregate data via an API solution (see Section 3.4.4). All other automatic data exchanges within the SEA IT infrastructure serve the purposes of operational needs, either of the SEA or the external partner.

As such, all needs for statistics and data of external stakeholders are catered by the statistics department in the SEA manually, by querying data from the statistics database (copy of BURVIS operational database), exporting the data to MS Excel, processing the data as relevant, and sharing the output with the respective stakeholder. The same process is performed for example for both regular and ad hoc needs of MoW, the Ministry of Finance, as well as to share data with researchers. The same approach is used also regardless

of whether the data need to be on personal level (with Latvian ID codes or anonymised) or aggregates, raising data security concerns in addition to inefficiencies.

The SEA's statistics department also sends the data manually to the data analytics system called the Unified Welfare Information System (LabIS), owned by MoW. LabIS entails a database that includes personal data from different registers under MoW (welfare, social insurance, data on jobseekers and vacancies from the SEA, etc) and enables producing statistics based on data linked across the registers via the Latvian ID codes using a BI tool (SAP BI). Although the data are linked using the Latvian ID codes during the data loading process, the ID codes are replaced by unique identifiers in an additional step and not made available to the users in LabIS.

In total, the SEA's statistics department shares every month 27 different MS Excel files with MoW to be used in LabIS. After extracting the relevant data from the copy of the BURVIS database, the statistics department processes the data to make these fit for the needs of LabIS. For example, the BURVIS database does not include data fields for the region or aggregate level education codes (only more granular address data and education data respectively), and the statistics department adds the corresponding information manually. This procedure is inefficient, and more error prone compared to conducting such data processing automatically within a solution containing a data warehouse and a BI tool, for example in the already existing LabIS solution by MoW.

Although the SEA is an important contributor to LabIS, the data in this analytics solution is broadly not usable for the SEA and only one statistician from the SEA has access to LabIS currently. The data content is generally too historic for the statistics purposes of the SEA (essentially with a lag of one month), largely because of the manual processes by the individual registers to deliver the data to MoW, which is time consuming.

The SEA is also obliged to share additional aggregate statistics with MoW weekly and monthly in MS Excel files, regardless that the SEA makes these statistics available publicly in its website, as well as shares the underlying individual level data for LabIS (which thus could enable any relevant aggregate statistics through LabIS). These files contain tables that include filters and drop-down boxes, as well as built-in figures to increase interactivity and user-friendliness of the files.

Although the current IT infrastructures of the SEA and MoW do not support the statistics department of the SEA sharing data and statistics with external partners sufficiently to maximise efficiency and security, the statistics department has been successful in finding solutions to work around the limitations and disseminate knowledge. For example, the SEA was quick to establish new relevant statistics products during the outbreak of the COVID-19 pandemic, as well as in the beginning of the refugee crisis caused by Russia's war of aggression against Ukraine. Also, the SEA has effectively shared data for the evaluations of ALMPs that have been conducted by the Ministry of Finance, as well as the OECD (2019[7]).

3.4.4. The Labour Market Forecasting Portal enables disseminating knowledge interactively

Although the IT infrastructure of the SEA does not include modern data analytics solutions to support the production and dissemination of statistics in general, a dedicated interactive digital tool has been developed over 2016-22 to disseminate labour market forecasts – the Labour Market Forecasting Portal https://prognozes.nva.gov.lv (LMFP). The LMFP was developed by the SEA in co-operation with the Ministry of Economics to share the short-term labour market forecasts of the former, and the medium-term and long-term forecasts of the latter. The development of this digital tool was possible as it received funding from the European Social Fund and the European Regional Development Fund and was prioritised in the SEA strategy (see Section 3.2.2).

The LMFP aims to support anybody needing to take decisions that require insights on labour market developments, such as jobseekers, students and schoolchildren that have to make career choices or

employers making decisions on human resource planning. Furthermore, the portal supports ministries and other public sector agencies by providing the labour market information for policy making and policy implementation, including for the staff in the SEA to be able to advise jobseekers and employers.

The LMFP focuses on key information relevant for its wide audience, i.e. above all long-term forecasts of general labour demand and supply, and shorter-term forecasts of labour demand by occupations and skills. The information is visualised on a few main dashboards, enabling the user to apply a few basic filters, order information and search for specific information (find information for a specific type of skill or occupation). As such, the interface allows some interactivity for the public users.

The database of the LMFP contains only aggregate data on different labour market indicators. Some of the aggregate data from BURVIS are exported monthly to the LMFP via an interface. Most of the data for the database of the LMFP are prepared in MS Excel, i.e. the forecast values for the different labour market indicators are prepared externally from the LMFP platform using econometric modelling and subsequently uploaded to the database. In addition to statistics from BURVIS and the forecasts by the Ministry of Economics and the SEA, the LMFP benefits from complementary information, such as statistics on unemployment and statistics on salaries from the State Revenue Service.

3.5. The processes and practices in the SEA do not ensure high data quality and security

This section discusses the key practices and processes in the SEA to ensure that the IT systems are resilient and secure, and that data processing complies with data protection requirements and supports the SEA's service provision and statistics production with high data quality.

3.5.1. Backup management is partly a side-product of statistics production

The backups of applications (stored in servers) are made using Veeam, which is a commonly used and accepted virtual server backup system. Using Veeam is a good approach for the SEA, as it can create snapshots which are easy to restore, for example when implementing system updates or installing updates to underlying operating systems.

The backup of BURVIS operational database is created every night by taking a database dump and restoring it immediately as a new database for statistics purposes (Section 3.4.1). The process of copying the BURVIS database also serves the purpose of a backup verification process. In general, very often database (and other) backups are not tested regularly, but in the case of the SEA at least a rudimentary restore testing is carried out every night. However, other databases of the SEA are backed up irregularly, which may cause information loss. To make a decision on the usefulness of implementing more frequent backups of the other databases, the risks of information loss need to be compared with the cost of more frequent backups.

3.5.2. Data quality management is mostly manual

The BURVIS interface includes somewhat limited data validity and integrity checks implemented directly in the system (e.g. classifications and built-in controls for data fields), and a substantial part of data quality management takes thus place manually. A data quality administrator in the statistics department is the main staff member to identify data quality issues, as the statistics department accesses the functionality to query personal level data, regardless that the query would not be for statistics and research purposes. Each week, the data quality administrator queries a different set of personal data to search for errors and missing data using MS Excel. Also statisticians run data quality checks in case they have temporary capacity to do so. The identified data issues are then shared with the respective branch offices of the SEA

together with instructions to address the data issues. This process contributes to a higher data quality and integrity for operational purposes, as well for statistics and research purposes. Furthermore, the statistics unit checks the (past erroneous) data again, and resends the errors to the branch offices if necessary, before using these data for statistics publications.

The volume and diversity of errors identified in the manual data quality checks are substantial, and involve all different sections in the BURVIS system – jobseeker registration, vacancy registration, provision of ALMPs. In addition to simply missing values in the data fields, some values do not match the content or legal criteria. For example, people in retirement age are not eligible to be registered as unemployed, yet have appeared in the quality checks. In terms of vacancies, the current approach requires to use ESCO codes, but the previously used classification appears in the data. The ending dates of ALMPs do not appear automatically in the process and need to be manually inserted by the counsellors, leading to often missing dates. Errors and missing data appear also as some of the data (updates) from external registers need to be manually queried, but the counsellor has failed to launch the queries. All these examples of errors could be avoided by implementing built-in controls in the BURVIS portal, thus not requiring manual quality checks or the data correction processes.

While the SEA has recognised the inefficiencies of the current data quality management, implementing better data quality checks within the BURVIS portal has not taken place due to limited resources and capacity of the development partner (Section 3.5.2). As adding additional features in the BURVIS portal was assigned higher priority in the list of necessary IT developments in the SEA, the manual process involving staff in the statistics department was chosen.

Although it is good that the SEA had the possibility to work around the capacity constraints of IT development and still implement data quality control, the solution also raises data protection issues and process effectiveness questions in addition to efficiency concerns. In the current set-up, staff in the statistics department accesses personalised data for quality checks, which could be avoided in automatic built-in quality controls. The manual quality checks might also not be able to detect all issues that an automatic process could, and as fixing the identified issues is also manual, the process is still error prone.

3.5.3. System and data security is prioritised, but not sufficiently

System and data security are rightfully considered important in the SEA. The SEA organises system security auditing annually via an external assessor, which identifies potential issues and bugs, and reports these back to the SEA. While most of the issues tend to be minor, the process has enabled to also identify some critical security issues in the past years. The IT development partner of the SEA is subsequently tasked to eliminate the identified issues.

In terms of compliance with the General Data Protection Regulation (GDPR) of the EU, the last audit from 2021 identified 16 issues in the BURVIS system that would need attention from the SEA. The main issues concern ensuring clearer and leaner roles, purposes and processes of processing personal data, anonymising and archiving historic data (these processes are missing fully, see Section 3.3.2), and providing the data subjects (clients) with full information regarding the processing of their personal data. However, the SEA aims to address only about half of the identified issues partially or fully in the near future (or has already done by Q2 2023), as several of the recommendations need more substantial changes in the BURVIS system in the context of very limited resources for such developments.

In terms of high-level security management, the SEA manages staff access to personal data and functionality in BURVIS via a central directory of users (Active Directory) to ensure that user rights correspond to the work tasks. In addition, improvements in data security and data protection are highlighted in the SEA's strategy for 2021-23 (Nodarbinātības valsts aģentūra, 2021[4]). As one of the measures to mitigate the system security risks, the SEA foresees regular staff training on cyber security in its strategy.

References

Nodarbinātības valsts aģentūra (2021), *Nodarbinātības valsts aģentūras darbības stratēģija 2021. – 2023.gadam [State Employment Agency's operational strategy 2021-2023]*, https://www.nva.gov.lv/lv/media/12401/download?attachment. [4]

OECD (2022), *Harnessing digitalisation in Public Employment Services to connect people with jobs*, OECD, Paris, http://oecd.org/els/emp/Harnessing_digitalisation_in_Public_Employment_Services_to_connect_people_with_jobs.pdf. [1]

OECD (2022), *Modernizing digital systems of the Latvian Public Employment Services*, https://www.oecd.org/els/emp/latviadigitalpes.htm. [5]

OECD (2022), *OECD Employment Outlook 2022: Building Back More Inclusive Labour Markets*, OECD Publishing, Paris, https://doi.org/10.1787/1bb305a6-en. [2]

OECD (2021), *OECD Employment Outlook 2021: Navigating the COVID-19 Crisis and Recovery*, OECD Publishing, Paris, https://doi.org/10.1787/5a700c4b-en. [3]

OECD (2019), *Evaluating Latvia's Active Labour Market Policies*, Connecting People with Jobs, OECD Publishing, Paris, https://doi.org/10.1787/6037200a-en. [7]

Office of Citizenship and Migration Affairs (2020), *Change of the personal identity number*, Office of Citizenship and Migration, https://www.pmlp.gov.lv/en/change-personal-identity-number?utm_source=https%3A%2F%2Fwww.google.com%2F. [6]

Notes

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1724&from=EN.

[2] The other two main objectives concern the integration of persons with disabilities into the labour market, and the implementation of the SEA's human resource policy in accordance with contemporary trends in human resource management.

[3] BURVIS information system is regulated by the Regulation of the Cabinet of Ministers No. 172 from 28 March 2017 "Regulation of the information system for the registration of unemployed persons and registered vacancies".

From:

Modernising Latvia's Public Employment Service through Digitalisation

Access the complete publication at:
https://doi.org/10.1787/07e830f5-en

Please cite this chapter as:

OECD (2024), "Assessment of the IT infrastructure in the State Employment Agency", in *Modernising Latvia's Public Employment Service through Digitalisation*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/003356e5-en