Sector Review of Information Technology in the Namibia Statistics Agency (NSA)

Final Report
August 2022

Authors:

Volker Täube   (EFTA)
Rune Gløersen   (Statistics Norway)
Kristian Lønø   (Statistics Norway)
Mira Nikić   (Statistics Serbia)
Aleksandra Skoko   (Statistics Serbia)

Table of Contents

Preface

The European Free Trade Association (EFTA) undertook a Sector Review of Information Technology in Official Statistics in Namibia. The review responded to a request from the Namibia Statistics Agency (NSA). The aim of IT Sector Reviews is to check with IT experts on the appropriateness of IT infrastructure in National Statistical Offices in order to ensure that they are able to fulfil their mandate. This corresponds also to principle 3.1 of the European Statistics Code of Practice (ESCoP) on the adequacy of resources, stating that "Human, financial and technical resources, adequate both in magnitude and in quality, are available to meet statistical needs."

The Sector Review was conducted by a team consisting of Mr Rune Gløersen and Mr. Kristian Lønø (both Statistics Norway), Ms Mira Nikic and Ms Aleksandra Skoko (both Statistical Office of the Republic of Serbia), and Mr Volker Täube (EFTA). The review was conducted in cooperation with the managers and staff of the Namibia Statistics Agency (NSA).

The Sector Review findings are based on discussions and presentations during a mission of the review team to the NSA, which took place on 1-4 August 2022, in Windhoek. In addition to the review mission, the colleagues from Statistics Serbia also provided some "hands-on" guidance on the spot.

Prior to the mission, the NSA staff completed a self-assessment questionnaire, which served as a starting point for the Sector Review, and a basis for the detailed discussions. The collaboration between the review team and the staff of the NSA was very positive and constructive throughout all phases of the work. The international experts would like to thank the NSA management and staff, and particularly the staff of the IT and the Data Processing Departments, for this.

Executive Summary

After several attempts in previous years for organizing an IT Sector Review at the NSA, that had to be put on hold due to the Covid pandemic, an international team of experts convened by the European Free Trade Associations (EFTA), conducted a Sector Review of Information Technology in Official Statistics in Namibia in August 2022. This overview was undertaken at the request of, and in partnership with, the Namibia Statistics Agency (NSA). This report contains the observations and recommendations of the international experts and has been agreed with the management of NSA.

The international panel of experts very much appreciated the high degree of professionalism and motivation of the NSA staff to contribute to the IT Sector Review. Some of the existing solutions in the area of IT handling and data processing already implemented in the NSA are of good quality and appear to be appropriate for the existing needs in statistical production. However, the expert panel, in cooperation with the NSA staff, identified several areas in which improvements might be envisaged. It is the conviction of the expert panel that the production and dissemination of Official Statistics in the NSA would benefit to a great extent from the partial or full implementation of the related recommendations. These recommendations are set out in detail in this report, but the following over-arching themes should be mentioned here:

- It seems that currently no dedicated data base management is in place at NSA. This concerns amongst other things (but not exclusively) the Eurotrace database, used to compile and store trade data for statistical production.
- Overall IT structure at NSA is marked by a "silo" approach, providing several individual IT solutions to the different production areas, instead of trying to extend harmonized solutions to several production domains and reducing the number of particular IT tools.
- A new IT investment and development strategy, complemented by a strategy for training is needed in order to fade-out outdated hard- and software infrastructure and ensure trained staff in due time. To be successful, such a strategy will need dedicated funding.
- Because of a recent large stoppage of IT infrastructure at NSA, due to the use of outdated equipment, short-term financing was available that helped to get around the most urgently needed server needs. Currently NSA works with systems originally foreseen for back-up purposes. Still, some servers and operating systems are no-longer supported by manufacturers, and other hardware has reached the end of its expected lifespan. This means that the risk to NSA operations from hardware failure is still existing.
- Statistical production is overly dependent on desktop tools such as Microsoft Access and Excel. Most national statistical offices have moved, or are in the process of moving, from these tools to more standardized and centralized systems (including use of open- source software solutions such as R or Python), which are easier to maintain, and can improve consistency of the statistics produced.
- The exchange of statistical data between different government bodies is currently not done in a secure way. The new cross-government data exchange platform will hopefully address this.
  The IT staff in the NSA shows a high level of dedication despite their relatively low salaries. A review of financial and non-financial incentives could identify ways to maintain this level of motivation and reduce staff turnover.

The international experts would like to commend the professionalism of the NSA IT management and data processing staff, who make the best use of the currently available resources. The experts would

like to thank the management and staff of NSA for their full and active collaboration in the conduct of this Sector Review.

Chapter 1: Institutional Environment, Strategy and Resources

Institutional Environment

The Namibia Statistics Agency (NSA) is organized as a state-owned company (semi-public/private). Staff has therefore no functionary status. NSA has around 160 staff mainly located in the Headquarter (HQ) and the Data Processing Center (DPC) in Windhoek, plus 14 regional offices with only one employee each. The NSA is placed directly under government which gives it a sufficient independent position to pursue its mandate without interference from political influences. The IT department at NSA is split in two divisions: the IT-infrastructure division and the data processing division. Joint, both entities consist currently of 12 persons, with several vacant posts (approximately 10).

During the discussions with the expert panel, it became evident that the NSA has no dedicated database manager and network manager. Some of the issues reported by NSA staff, especially slow response of databases, is probably also linked to suboptimal database handling (e.g., poor database design, lack of primary keys, no indexing, poor performance queries, high table fragmentation, lack of archiving strategy etc.). There seems to be no budget for the recruitment of a database manager and network manager at NSA, who could help overcoming such shortfalls, the question should be raised if the financial budget for IT-staff is sufficient.

In addition to the above-mentioned observations and in the light of recent IT failures at NSA it became evident that replacement of hardware is urgently needed. A major upgrade will take place soon. However, a sustainable strategy to monitor hardware capacity and to plan for procurement and upgrades well in advance should be established.

Strategy

There is currently no single document covering all areas of IT strategy, such as IT training. To underpin the upcoming business-oriented IT-strategy, an IT maturity assessment should be conducted. There are several models and templates available to guide such an assessment. Rather than a big bang assessment, it could be split into prioritized areas. One important reason to take a stepwise approach, is to measure the current state of maturity in those areas where you prioritize to improve the capabilities of the office, to be able to estimate the impact of the efforts taken.

Resources

Currently, the IT staff at NSA marks a share of about 4% of the total staff. The total IT and DPC staff are 8,5% of the total staff, but the comparison with IT elsewhere may differ compared to the tasks and competence of the DPC. Anyway, this seems to be rather low, compared to other statistical agencies (average in Western Europe is around 10%) and an augmentation of resources seems to be justified.

Recommendations:

- Elaborate/finalize an IT strategy that supports the overall strategy of NSA
- Prioritize using IT staff for statistical production and related needs, avoid IT staff activities related to non-statistical business
- Consider appropriateness of IT budget and number of staff. Obviously, network and data base specialists are urgently needed
- Elaborate a strategy for capacity building and training needs, with focus on open-source software well suited for statistical needs, like R and Python

- Consider seeking funding to raise the share of IT staff to be able to meet the needs for IT systems development
- Establish an ICT steering committee to ensure common systems development and priorities and transparency
- Conduct a targeted IT maturity assessment
- Consider enabling participation of staff in international meetings in order to build networks with people doing similar work in other national statistical offices

Chapter 2: The National Statistical System and Other Cross-Government

Development of a data platform for the exchange of admin data is currently underway. Currently, data exchange with administration takes place mainly by e-mails, Excel files and USB devices, which puts a question mark on the data security. Access to different data sources (admin, private) is principally possible. Access to public administrative data is given by law. In addition to traditional surveys, web scraping is planned for price data.

Acquisition of software is normally handled for every public institution itself. NSA is entitled to provide paid services to private institutions.

NSA is in the initial phases of implementing a COBIT-5 governance framework. COBIT (Control Objectives for Information and Related Technology) helps organizations meet business challenges in regulatory compliance, risk management and aligning IT strategy with organizational goal. Benefits of COBIT-5 are:

- Improve and maintain high-quality information to support business decisions.
- Use IT effectively to achieve business goals.
- Use technology to promote operational excellence.
- Ensure IT risk is managed effectively.
- Ensure organizations realize the value of their investments in IT; and
- Achieve compliance with laws, regulations and contractual agreements.

Recommendations:

- If web scraping will be implemented: ensure coordination with owners on either direct delivery or webpage access
- Proactively participate in the development of a secure governmental/state data exchange platform and establish a more secure system to exchange data with external bodies, avoiding the use of Excel sheets sent by e-mail, or USB devices or other portable devices.
- Try to obtain general government contract with software vendors in order to realize gains from discounts.
- Continue practice of providing paid services to other governmental agencies in order to generate additional budget for IT.

Chapter 3: Hardware Infrastructure and IT Security

The components of IT infrastructure in NSA are made up of interdependent elements, and the two core groups of components are hardware and software. Hardware uses software—like an operating system—to work. And likewise, an operating system manages system resources and hardware. Operating systems also make connections between software applications and physical resources using networking components.

In NSA hardware components include:

Desktop and laptop computers:
> 60% of the computers are outdated.

Tablets:
> 40% are outdated.

Servers:
> NSA runs VM Ware virtualized environment on traditional infrastructure. A traditional IT infrastructure is made up of the usual hardware and software components: facilities, data centres, servers, networking hardware desktop computers and NSA application software solutions. This infrastructure setup requires more power, physical space and money than other infrastructure types. A traditional infrastructure is installed on-premises, only for NSA use. All data are stored on-premises and the storage capacity is under pressure. NSA has experienced hardware failure incidents, some because of high temperature, causing unavailability of servers and services where some of the incidents has lasted for weeks.

Server storages:
> 1 volume storage server out of 6 is outdated.

Standalone servers:
> 3 out of 5 are outdated.

Network equipment:
> Network switches and routers are rather new.

Backup-power:
> Mainly outdated.

WiFi Network and LAN:
> It is very important to emphasize that WiFi is not separated from internal LAN. Cost of Internet is extremely high due to poor infrastructure in African region as compared with global rates.

Facilities:

The Data center serves HQ and the DPC. IT is currently strengthening the backup facilities, ensuring that the inhouse back-up site is separated from the data center for sufficient disaster recovery, and they are in process of implementing offsite backup for the same reason.

Physical security is good. NSA enforce strong access control policies in areas where hardware and physical equipment are housed. Hardware devices and peripherals are not left unattended in open areas, and employees take measures to secure devices.

Software components:

Operating systems – mostly Microsoft Windows 2016 OS and a few servers running Linux OS
Incremental and full backups are established between servers at the HQ and DPC (data processing centre)
Sufficient software is used for cyber security and anti-virus protection

Recommendations:

- In case of insufficient budget for ensuring server capacity, check for possibilities of renting server capacities (e.g., other government bodies)
- Change network switches to improve speed and security.
- Consider investing in more secure, faster and stable Wi-Fi
- Divide Wi-Fi from internal LAN
- Changing of obsolete standalone servers is must
- NSA should use the Census as opportunity to update equipment and infrastructure
- The vacant post of a senior network specialist should be filled

Chapter 4: Software Infrastructure

Software (application) infrastructure consists of all the components that are needed to deliver an application and its functions and services to the users. If an application infrastructure is flexible, reliable and secure, it can help an NSA meet its goals. Alternatively, if an IT infrastructure is not properly implemented, NSA can face connectivity, productivity and security issues—like system disruptions and breaches. NSA operates with commercial and non-commercial software.

Licensed Software (Statistical and GIS):
- SPSS Data Analysis
- STATA Data Analysis
- ARCGIS geographic & Map design Software
- Survey123 Spatial data collection
- Digital Namibia dissemination spatial data dissemination

Other Licensed Software):
- Pastel & Payroll Finance, Accounting & HR Software
- Mfiles Document Archiving Software
- SAGE300 Accounting & HR Software
- Microsoft volume license E3 All Microsoft Office package
- VMware Virtualization & cloud computing software
- Beachhead simply secure Data Encryption and MDM
- Kaspersky Antivirus
- IP Guard Network Monitoring
- Fortinet Firewall software
- Veeam Backup Software
- Readsoft Optical character recognition

Open-Source Software (Statistical and GIS):
- R Data Analysis Software
- R Studio Data Analysis Software
- Python Programming Software
- CSPRO Data collection tools
- Survey Solution Data collection tools
- Quantum GIS Geographic Information System Software

Other Open-Source Software:
- Zabbix Networking Monitoring tool
- MySQL Database Software
- Postgres Database Software
- Apache and IIS webhosting

For the National account system, NSA uses NADABAS and Excel software for data processing. NSA also uses a Microsoft Excel based system for the compilation of national accounts, but without a database system. NADABAS provides a solution where the national accounts compilation remains in Excel but are more robust and reliable by replacing links between workbooks with a database. In doing so, it provides a tool to design a structured system to compile national accounts at low cost that will ultimately lead to improved quality of national accounts.

Excel is also used for sampling and collecting data for CPI (Consumer Price Index) purposes. Excel based questionnaires are sent by email, index calculations are processed in Excel. Prices division has plans to develop index calculation (CPI) in R language. R is a programming language and software environment for statistical analysis, graphics representation and reporting. R is currently considered as a strategic direction for data processing. Some developments have been done in Python as well. Decision on future programming tools has not yet been made. Mainly, NSA statistical divisions use SPSS and STATA software for data processing.

External trade division uses Eurotrace. Eurotrace is an application for the collection, compilation and dissemination of external trade data at national and regional level. It is a generic and open system, but its final release is from 2016. NSA migrated Eurotrace data from MS Access to MS SQL Server. The main obstacle is slowness of the application and limited number of reporting columns (up to 6). The main reason for application slowness lies in poor database architecture and missing indexation. For example, backend Eurotrace database has lack of primary keys in all database objects. Hardware infrastructure is not the reason for slowness. Limited number of reporting columns is system inherent, e.g., NSA uses version of Eurotrace software which has limitation on number of displayed reports columns.

A Labour Market Information System (LMIS) is in development phase. Work is supported by ILO. The overall compatibility with NSA's software architecture has not yet been evaluated.

NSA has developed GIS Portal using ARCGiS with ESRI maps, which uses PostGres as backend database and Geofind as metadata browser. NSA is in charge of storing and maintaining all state geo maps. Therefore, storage and maintenance in this area is a burden to an already stretched IT capacity.

Statistical metadata (structural, operational and referential) are not in place.

On premise MS Exchange server is used as mail server. NSA has all necessary licenses for MS SharePoint, but has not put this system into production. SharePoint is primarily a document management and storage system, but the product is highly configurable.

Recommendations:

- NSA should prepare inventory of all "silo" software solutions currently in use
- As part of the software strategy, NSA should formulate clear standards for more unified application solutions
- NSA should elaborate a strategy for implementing open-source software like R and/or Python and a plan for migrating (phasing out) from excel based production. The use of Stata and SPSS and other tailor-made solutions should be identified and considered replaced over time by open-source software
- NSA should provide trainings in R and/or Python
- Eurotrace software version should be upgraded to newest version
- NSA should consider creating reports outside of Eurotrace (for instance in R or Python) in order to overcome the system's current limitations
- MS SharePoint should be put in production for document management, ticketing system and sharing internal documents
- Government must provide additional budget for maintaining mapping capacities by GIS department

Chapter 5: Data Architecture, Storage and Management

Over time, several different software systems, tailormade for different statistical products have been implemented. In addition, Excel is used for a wide variety of tasks when analysing data and producing statistics. All in all, this is reflected in a lack of an overall data architecture. Data are stored and maintained in different ways, often locked in by the different systems. Sufficient metadata is partly missing, or not harmonized. This has led to multiple copies of the same data being stored in different places. In addition to the total volume of storage needed, back-up capacity etc, the challenge to keep control of data becomes a problem.

For Metadata, NADA (National Data Archiving) central catalog, made by US Census Bureau and the African Development Bank is implemented. There are plans to develop a new, more comprehensive metadata system in-house.

NSA have no generic system for classifications and correspondence tables in place. For Data archiving, a system that defines common rules and recommendations like naming conventions, versioning of data, corresponding metadata etc. is not in place. In some areas, parts of Data Archiving needs have been implemented. Data storage is also a concern.

The development of a Generic Business register started last year. It is developed by a single external consultant and will be finalized in March 2023. It will be handed over to NSA with manuals, and training will be given at the handover. Then it will be NSA responsibility to maintain and further develop the system. The support of the system by the developer will end in 2024.

Recommendations:

- In connection with the IT strategy, elaborate a strategy to achieve more coherent data and metadata management, including central metadata handling, classification server with correspondence tables and an integrated archiving system.
- Establish routines for version control of data, metadata and programs
- Make sure that there will be enough competence and resources in NSA to be able to maintain the Business register and other externally developed production systems to reduce vulnerability
- Redesign and maintenance of databases is needed to avoid data redundancies and to improve overall performance
- Vacant roles within data and database management, through a database specialist should be filled.

Chapter 6: Data Collection and Other Inward Data Flows

NSA has dislocated data processing centre (DPC) from HQ. About half of IT staff is situated in the DPC. Data Processing Division is in charge of:

- Design, development and testing of data collection and cleaning tools applications
- Data management – process, clean, check data completeness, perform structural edits and analysis, and store data
- Database Management
- End User training

Data collection methods in use are:

- Pen-and-Paper Personal Interview (PAPI)
- Computer-Assisted Personal Interviews (CAPI) to conduct face-to-face data collection using tablets which improves data accuracy and timeliness of data transmission

Sampling is done with Excel. NSA wants to make a more coherent sampling system in Stata. Currently, there is an attempt to come up with an automated sampling module in Python but due to shortage of staff (Geospatial Technologist assisting in this activity) and is occupied with a lot of work (GIS), the activity is partly done.

For CAPI purposes, the Census and Survey Processing System (CSPro) software is used. CSPro is a public domain software package used for entering, editing, tabulating, and disseminating census and survey data. CSPro is user-friendly and powerful enough to handle the most complex applications. It can be used by a wide range of people, from non-technical staff assistants to senior demographers and programmers. CSPro supports data collection on android devices (phones and tablets). The CSEntry Android App works in collaboration with the desktop version of CSPro. Data transfer from Android devices is done by using Bluetooth connection. NSA is in the initial stages of exploring non-face-to-face data collection methods: Computer Assisted Telephonic interviews (CATI) and Computer Assisted Web interviews (CAWI). For non-face-to-face data collection methods, NSA explores other non-commercial software Survey Solutions. Survey Solutions is a survey management and data collection system developed by the World Bank.

ODK is also used to some extent for data collection on mobile phones and tablets as an alternative the CSPro system.

Administrative data are exchanged, mainly, over email using Excel files and via USB devices. A complete and secure platform for collecting administrative is still in the process of being developed. For monitoring Census fieldwork Dashboard has been developed with support from Office of the National Statistics UK (ONS) and United National Economic Commission for Africa (ECA). This dashboard can be used for other surveys fieldwork as well.

Recommendations:

- Prioritize the ongoing work to establish one single integrated system for all survey data collection work comprising of a dashboard to follow up the field work, pay-roll data etc.
- Consider using R or Python instead of Stata whenever a new sampling system is to be made
- The survey data collection system should handle multi-mode data collection (CAPI, CATI, CAWI) in an efficient way

- NSA should establish a secure data exchange system for administrative data and data from other sources
- IT department should consider obtaining higher data integration (e.g. to use centralized and unified databases on servers)
- There is urgent need for staff training on database management, database design and using SQL language for data manipulation and maintenance.

Chapter 7: Data Processing and Analysis

SPSS is used by statisticians for their data processing and analysis. Stata is also used; however the usage is little and decreasing. R and Python are used in a limited capacity by data processing.

Excel is widely used. A common experience is that Excel based systems tends to be difficult to maintain in the long run.

The data processing is often conducted within a heterogenous environment of subject matter tailormade systems, often implemented by external consultants and/or donor organizations, which has led to a silo approach. Foreign trade statistics (which uses Eurotrace), National Accounts (which uses Nadabas) and part of the systems within the labour statistics (which uses LMIS) are examples of this approach

If the variety of software decreases, the ability to provide sufficient training capacity for the software used will increase.

Recommendations:

- Prepare an inventory of software used; and consider mapping the usage of the different systems according to the GSBPM model.
- Ensure that any introduction of a production system is in line with the software strategy and that there are sufficient resources within IT to maintain the systems
- Provide sufficient training capacity and follow-up for the approved software used
- Make sure that external technical assistance contributes to the IT capacity and is in line with the software strategy
- Elaborate a process template that is based on best practices, to provide for a recognizable and effective data editing, processing and analysis of survey data

Chapter 8: Dissemination

Dissemination activities at NSA are mainly web based. The frequency of hits or requests made on the NSA website is monitored by Google analytics. As in other countries, the NSA is in charge of compiling and disseminating data on the UN Sustainable Development Goals (SDGs). NSA has developed in-house SDG-portal.

Press releases are announced by email alert; there is a comprehensive user mailing list. Social media (Twitter, Facebook) are widely used by NSA communication unit. Official NSA web site https://nsa.org.na offers a lot of information and data:

- Namibia data portal
- Covid 19 dashboard
- Namibia Geo Portal and "Digital Namibia", a platform that allows to view and perform basic spatial data analysis held by different government bodies
- Namibia trade forum
- Namibia Census 2011
- Nada micro data

FAQ page and a release calendar are published on the web site. The Statistics Act, strategic/business plan, as well as an annual work plan are also on the NSA website.

Recommendations:

- Enhance the digitalization skill set of the communications staff
- In their five-year plan, NSA should consider developing a statistical dissemination data base, as part of the web site which will enable quick access to data, search of data, downloading of dynamically made reports in various formats, simple data visualization and B2B connection using APIs.
- NSA should find a way to slightly automate process of release monthly reports and press releases in order to release bottleneck pressure in communication department
- Consider new ways to disseminate data, like infographics. Infographics can be put on web site, Twitter and Facebook
- Consider putting popular statistics on web (like most common names in country)
- Use Google analytics data in order to identify users interests on statistical areas and put more focus on what is mostly demanded by users
- Open a YouTube channel for promoting statistical office as a whole. YouTube channels can contain various videos promoting press releases, importance of official statistics, simple trainings in order to achieve users' data literacy etc.

Chapter 9: Non-Statistical IT

NSA non-statistical IT is used in the following departments:

- Legal department
- Financial department
- Audit division
- HR department

A data management system also for non-statistical IT is missing. The Financial department uses outdated Pastel software which is not integrated with other NSA systems. In current software version there is no control reports, no audit possibilities and biggest issue is with payroll for interviewers. NSA has no additional budget for replacement Pastel software. There is no centralized procurement system, and it is planned to put in place in two years from now. Audit division is using exclusively Excel files. All reports are done manually, there is lack of contract management software and lack of special audit software. HR department also uses Pastel software. The biggest challenge is enumerators recruitment. NSA uses national employment information system for recruitment. Performance management (evaluation of co-workers) is in place. A 360-degree review system should be acquired. During the Covid19 pandemic teleworking was not possible for non-statistical departments.

Recommendations:

- Separate non-statistical IT from Statistical IT
- Use ready-made integrated systems for administrative tasks
- NSA should implement SharePoint platform (licenses are already bought by NSA), where all statistical and non-statistical documentation can be stored and shared. SharePoint is web-based collaborative platform that integrates natively with MS Office files (Word, Excel, etc.). Share Point can be used as a document management and storage system, but the product can be also used for variety purposes as well (ticketing system, helpdesk system etc.)
- NSA should be informed whether IT office under government has plans for developing unified HR and/or accounting platform at state level. In that case NSA should try to have influence in development of that platform and should proactively be involved in testing of that potentially new cross government application to ensure that application meet NSA's needs.