

Distr.: General  
02 October 2017

English

---

## United Nations Economic Commission for Europe

### Conference of European Statisticians

#### Work Session on Migration Statistics

Geneva, Switzerland

30-31 October 2017

Item 3 of the provisional agenda

#### Data integration and administrative data

## Synthesis of administrative data and big data technologies for improved migration data

Note by Public Services Development Agency, Ministry of Justice of Georgia\*

### *Abstract*

Unified Migration Analytical System (UMAS) is an innovative IT solution to a long-standing challenge of migration data gaps. UMAS is linking migration-related administrative data collected by different state agencies involved in migration management using big data technology and an advanced analytical software. UMAS is created to enable analysts to uncover hidden patterns, correlations, and other insights invisible for traditional business intelligence solutions and summary statistics.

This is a mid-term progress report explaining UMAS nature, added-value, and challenges to the wider audience of migration professionals in Georgia and internationally who might be interested in the UMAS development path. The first half of the report explains the existing challenges and the need for innovation in the migration data collection field. Administrative data collection situation and present official migration statistics are reviewed.

The second half of the report details challenges, decisions, and results of the first phase of UMAS development in the 9 horizontal issue areas - UMAS Ownership, Inter-agency cooperation, System Architecture, Data Processing, Personal Data protection and Data Governance, UMAS output and BI Analysis, Communication and Usability of the Output, Data Quality and Standardization, and Human Capacity Building.

Final section of the report spells out tangible progress observed in the different areas of migration policy and management as a result of UMAS. More specifically, notable results are registered in terms of better understanding immigration flows and initiating policy discussion as a result; improved data sharing and cooperation among the data source state agencies, and generation of new, missing data.

\*Prepared by Ms. Nino Ghvinadze

## I. Missing Migration Data

1. In the age of big data and rapidly developing data gathering and analysis technologies, when all major business and policies are shaping their decisions and strategies on evidence backed by data, data on international migration is still scarce and fragmented. This problem has been acknowledged for decades now. The UN recommended more emphasis on improving international migration statistics back in 1998 (Recommendations on Statistics of International Migration, UN, 1998:1) and these recommendations are still on the table today. In February 2017, Mr. Peter Sutherland's report to the UN Secretary-General again recommended improvement of migration data for fact-based migration policies and accountability: "Better governance hinges on good data and a proper understanding of migration drivers and dynamics. Yet, we still struggle to establish basic facts, such as who migrants are, where they are, where they come from and when they moved." To address this problem, Mr. Sutherland recommended "compiling and releasing existing administrative data," "providing access to micro-data, not just tabulations", and clarification the potential of and capitalization on 'big data' as a complement to traditional data sources, among other things (Report of SRSB on Migration, 2017:25).

2. Georgian experience with migration data collection and official migration statistics is no different. Free movement of people, while significant benefit for international travellers, makes it challenging for migration administrative bodies to study the groups of international migrants in Georgia. The quality of statistics further suffers from the lack of experience of using administrative data for statistics and the restraints on sharing the micro-data.

3. These problems are widely recognized on policy-making level. Policy documents in different sectors point out the need to address the 'data deficit'. 2016-2020 Migration Strategy of Georgia states that the lack of reliable data is one of the main challenges in developing migration policy in the country and this situation adversely affects "the informed decision-making in the field" (2016-2020 Migration Strategy of Georgia: 41). One of the strategic goals of the State Strategy for the Labour Market Development in Georgia for 2015-2018 is to institutionalize the regular labour market studies and develop the labour market information system (Goal 3.4). Because "otherwise it is impossible to conduct comprehensive assessment of the situation [on the labour market] and make reliable forecasts on the national, regional, and sector levels" (2015-2018 Labour Market Development Strategy: 7). Similarly, one of the nine goals of the Georgia Tourism Strategy is the expansion and enhancement of "Georgia's ability to collect and analyse tourism data and measure industry performance" (Strategic goal #9, 2015-2025 Georgia Tourism strategy: 19). These are just a few of many examples of acknowledging the lack of quality data by decision-makers on highest political levels.

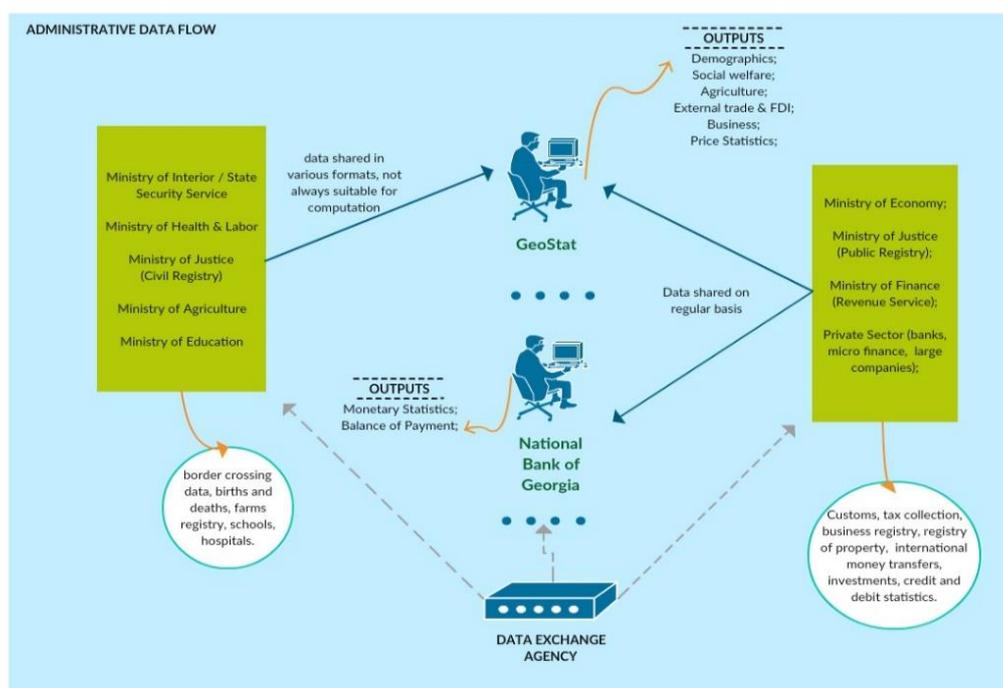
4. 2015 Medium Migration Profile of Georgia approved by the State Commission on Migration Issues (SCMI) of Georgia also states that the "availability of reliable national migration statistics remains a challenge, leading scholars and policy-makers to rely on various types of estimations rather than on solid statistical data" (MP 2015:65).

## II. Official Migration Statistics in Georgia

5. Georgia's National System of Official Statistics (NSOS) has two major pillars – the National Statistics Office of Georgia (GeoStat) and the National Bank of Georgia (NBG). Although GeoStat is a major statistics producer nationwide, NBG also collects data to produce monetary and balance of payment statistics. Further, GeoStat closely cooperates with the Ministry of Finance and the Ministry of Economic and Sustainable Development of Georgia, representatives of both ministries sitting in the governing board of the GeoStat.

6. The Data Exchange Agency (DEA), the Legal Entity of Public Law (LEPL) under the Ministry of Justice, is not part of the national statistics system, but according to its mandate, should be connected to it. DEA was established in 2010 to develop e-governance, data exchange, and information security in Georgia. Under the 2011 Law of Georgia on the Unified State Registry of Information, DEA is responsible for being a kind of 'mediator' in any form of electronic data exchange among different state agencies and with the statistical offices.

Illustration 1. Administrative data flow for production of official statistics



Source: Author's illustration

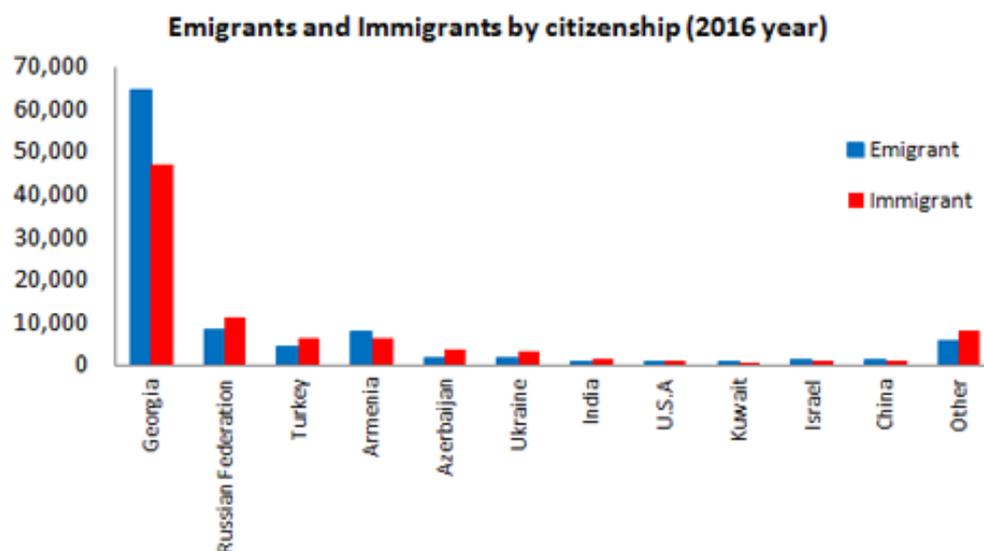
7. The Illustration 1 above provides a rather simplistic map of the major data sources and the data exchange with and among statistics producers. The sources mapped are by no means exhaustive and the flow of information is not as one-dimensional in practice.

8. Georgia's NSOS is rather sound. Basic legislative and institutional framework is developed and the minimum national statistics are produced according to the predetermined calendar. Based on the World Bank (WB) Statistical Capacity Indicator (SCI) Georgia ranks above average compared to the Europe and Central

Asia region and the average of the International Bank for Reconstruction and Development (IBRD) (see the SCI Dashboard). The SCI measures country’s statistical capacity based on methodology, source data, and periodicity. It is noteworthy, however, that the SCI looks at predominantly economic data. Georgia’s good ranking in SCI has to do with the improved internal data collection systems at relevant state institutions as well as with the advanced methodology of economic statistics almost fully adjusted to the International Monetary Fund’s (IMF) standards. The IMF and the WB have been closely cooperating with the NBG and relevant line ministries to achieve this progress. That said, economic, financial, agricultural, social, and demographic statistics are all interlinked. It is impossible to have flawless financial statistics without proper demographics data. After the results of the 2014 National Census revealed that the estimates of Georgia’s population were off almost by 800,000 (see GeoStat’s annual population estimates), it became obvious that many economic indicators had to be revisited.

9. Production of official migration statistics as part of the demographics statistics is the GeoStat’s competence. It publishes annual statistics of migrant flows to and from the country and thus, net migration by year, as well as the citizenship, sex, and age of migrants.

Illustration 2. Emigrants and immigrants by citizenship



Source: Geostat

10. GeoStat’s net migration data has been fluctuating quite considerably during the last 10 years, partially caused by varying data collection and analysis methodologies applied by GeoStat during these years and partially, by changing migratory patterns. For preparing its annual estimates of migratory flows, GeoStat uses the registration system of actual border movements. A person is considered as emigrant or immigrant if s/he has crossed the border in or out and has not crossed back for the period of 6 months plus 1 day.

11. Another source of official migration statistics is the national census and the latest one was conducted in 2014. Census asked couple questions about international migration and reported number of immigrants, country of origin, sex, and age, and number of emigrants, country of their destination, sex and age. Census 2014 used

the UN definition of a migrant. Emigrant was defined as a person who since January 1, 2002 has left Georgia and resided in another country for 12 months or planned to stay abroad for more than 12 months. An immigrant was defined as the person who has ever lived abroad for more than 12 months (GeoStat 2016:13-14).

12. Finally, GeoStat's combined Labour Force and Household Income and Expenditure surveys ask a few questions on emigration and remittances from abroad and reports characteristics of remittance receiving households and their level of income from other sources. The estimates of this survey significantly differ from the numbers reported by the NBG (Hoffman 2014:17). The later reports money transfers to and from Georgia through electronic wire systems (western Union, Money Gram, Amelik, Unistream, etc.). The data is gathered from the monthly statistical reports of the commercial banks, including branches of non-resident banks and microfinance institutions in Georgia (NBG External Sector Statistics). Hence, it does not include transfers made by informal means or to the Georgian bank accounts.

13. The State Commission on Migration Issues publishes two types of Georgia Migration Profiles – medium and brief migration profiles. Medium Migration Profile (MMP) is a policy tool presenting an all-inclusive picture of migratory processes and trends. It utilizes primarily administrative sources of data, both Georgian and foreign (when covering emigration). MMP provides detailed but concise and critical review of migratory trends, the quality of the data available, and their interrelation with the existing migration policies. Brief Migration Profiles (BMP) are shorter reports providing in-depth coverage of one area of migration. For instance the two BMPs published in 2016 and 2017 focused on inflow and outflow of Remittances and Foreign Students in Georgia, respectively.

### III. Collection of Administrative Data in Georgia

14. Throughout the last decade Georgia has successfully reformed a number of state institutions. Noteworthy progress has been made towards improved public administration, public services, and the components of e-governance have been introduced. Improved data collection was an unintended side-effect of these developments. It was not the end goal of the reforms, but often served as the means to the end. One of the good examples was the reform of the civil and public registries. Back in 2004, when the reform started and later, in 2011 when the Public Service Halls (PSH) pooled together various state services, the focus of the leadership was cleaning the house from corruption and providing improved public services (IBRD 2012:63-74). However, this also meant accurate births and deaths registry, business registry, and many more. Combining services of different state agencies under the umbrella of the PSH created incentives for the partner agencies such as the Ministry of Labour, Health and Social Protection and the Ministry of Education and Science to tidy up their internal databases, at least the ones connected to the PSH services. In 2011 Ministry of Education started works to consolidate its databases and electronic systems under one umbrella – LEPL Education Management Information System (EMIS).

15. Accumulation of these data at different agencies on the one hand, and the lack of data analysis and evidence-based policy - on the other, asked for building the bridge between the two. To address this issue to some extent, several independent projects have been launched in 2014-2015 with the leadership of the Ministry of

Justice (MoJ). LEPL National Agency of Public Registry (NAPR) is coordinating development of inter-institutional geographic information system - Georgian National Spatial Data Infrastructure (NSDI).<sup>1</sup> DEA has launched a new online platform – Data.Gov.Ge – to create a unified space where all state agencies can publish datasets of public information they consider relevant.<sup>2</sup> Finally, The Unified Migration Analytical System (UMAS) is being developed by PSDA to tackle gaps in the international migration statistics. UMAS is innovative solution to a long-standing problem using the synthesis of administrative data and big data technologies. The concept, added value and already achieved milestones of UMAS are further detailed in the following sections of this report.

16. Migration-related administrative data are usually scattered across the different databases of different registering institutions. Below is the summary of major data sources by the area of migration administration.

- i. Authorized migration – visa and residence permits: PSDA of the Ministry of Justice is responsible for issuing residence permits on different grounds. It also receives applications on Georgia citizenship, studies the case and makes recommendations to the Office of the President of Georgia to make the final decision. MFA is issuing visas for short and longer stays in Georgia since September 1, 2014. Georgia introduced e-Visa<sup>3</sup> for short-term visitors in February 2015. E-Visas are also administered and registered by the MFA. Ministry of Education and Science keeps the registry of foreign students at Georgian public schools and at the higher education institutions. It is a complementary data source on top of residence permits and visas granted for education purposes.
- ii. Movement across the Georgia state border: registrations at the state border when going in or out of the country is one of the major source for migration statistics. Ministry of Internal Affairs is responsible for registering all travellers, including the daily commuters. MIA’s border crossing database and infrastructure are one of the most advanced ICT systems in the country. It stores data since 2009.
- iii. International protection: new database of the Ministry of Internally Displaced Persons from the Occupied Territories, Accommodation and Refugees of Georgia registers all asylum seekers applying to Georgian State as well as all their subsequent activities – type of accommodation, movement, and type of decision, appeal and final outcomes.
- iv. PSDA is responsible for identifying, registering and integrating (documenting) stateless persons in Georgia.
- v. Economic activities of immigrants in Georgia: establishment of new businesses and acquisition of real estate in Georgia is registered by the LEPL National Agency of Public Registry (NAPR). Subsequent activities of business – number of employees, taxes paid and

---

<sup>1</sup> See Georgian National Spatial Data Infrastructure (NSDI) at <http://nsdi.gov.ge/ge/> Seen on June 29, 2017.

<sup>2</sup> See Data.Gov.Ge at <http://data.gov.ge/Page/AboutUs> seen on June 29, 2017.

<sup>3</sup> See the e-Visa Portal at <https://www.evisa.gov.ge/GeoVisa/> seen on June 29, 2017.

active/passive status can be monitored through the registries of the Revenue Service (RS) of the Ministry of Finance.

#### **IV. Unified Migration Analytical System – Georgian Solution**

17. The concept of creating an analytical system which would combine all migration-related administrative data sources and allow further manipulations with the data was born in 2010. Up to 5 years were spent on studying experiences of other countries, searching similar systems for know-how and examining the baseline situation in terms of data collection in Georgia. After multiple study visits, close examination of Estonian, Austrian, Norwegian, and other models, consultancy of Georgian and European IT specialists, data scientists, and migration experts, the UMAS concept received a clearer shape.

18. With the support of the EU funding and partner international organizations (International Organization for Migration and International Centre for Migration Policy Development), a team of IT specialists, developers, data scientists, and lawyers were recruited in 2016. A trial version of the system was up and running in about one year. As of June 2017, after roughly a year and a half, first phase of UMAS development is complete – it is already processing real data secured in the databases of various administrative agencies involved in migration management. Although, naturally, UMAS is a work in progress and many challenges lie ahead.

19. For the sake of this report strategic decisions, technical solutions, key achievements and lessons learned in the first phase of UMAS development are summarized in the following 9 sections – (a) UMAS ownership, (b) inter-agency cooperation, (c) system architecture, (d) data processing, (e) personal data protection and data governance, (f) UMAS output and BI Analysis, (g) Communication and usability of the output, (h) data quality and standardization, (i) UMAS team and human capacity building.

##### **A. Technical Development of UMAS**

###### **1. UMAS ownership**

20. Business owner of UMAS is the State Commission on Migration Issues (SCMI), which chose Public Service Development Agency (PSDA) of the Ministry of Justice to implement the project. PSDA, which also houses SCMI Secretariat, is developing and administering UMAS. UMAS team of data scientists and developers, its servers, hardware and software are located at the PSDA.

21. UMAS development processes is closely monitored by the European Union Delegation to Georgia, which provides 50% of the funding of the project. The funding is provided through the implementing partners – Georgia offices of the International Organization for Migration (IOM) and International Centre for Migration Policy Development (ICMPD). The other 50% of the funding is provided by the Georgian state through PSDA budget. The process is also closely monitored by the Office of the Personal Data Protection Inspector to make sure all norms of personal data protection are protected according to the Georgian Law.

22. Other state agencies – members of the SCMI – are the sources of the raw data for the analytical system. They all have bilateral memoranda with the PSDA defining the scale and scope of data sharing, also the roles and responsibilities of both sides in data processing and security. These state agencies closely cooperate with the PSDA and UMAS team to make UMAS happen.

## 2. Inter-agency cooperation

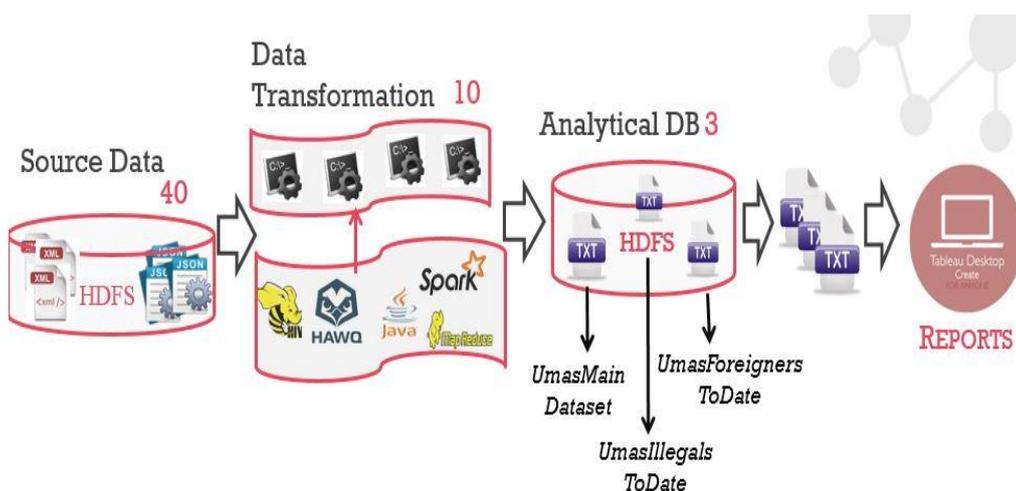
23. Major administrative registries, which supply data to UMAS are:
- i. The Public Service Development Agency (PSDA) of the Ministry of Justice – provides data on residence permits, citizenship, and civil acts (births, deaths, marriages and divorces);
  - ii. The National Agency of Public Registry (NAPR) of the Ministry of Justice – runs business and real estate registry;
  - iii. The Ministry of Internal Affairs (MIA) and the State Security Service of Georgia (SSG) – own border crossing database, visas issued at the border checkpoint, illegal migration, detention, and deportation data;
  - iv. The Ministry of Foreign Affairs (MFA) – stores data on visas issued by Georgian diplomatic representations and consular services;
  - v. The Ministry of Internally Displaced Persons from the Occupied Territories, Accommodation and Refugees (MRA) – registers applications for and decisions on international asylum in Georgia;
  - vi. The Revenue Service of the Ministry of Finance (RS) – stores information on the taxpayer legal entities registered in Georgia and customs registries;
  - vii. The Educational Management Information System (EMIS) and the National Centre for Educational Quality Enhancement – own data on foreign students in Georgia.
24. Other data sources to be included in the System on the later stages are the Ministry of Labour, Health, and Social Protection of Georgia, Courts, and the Office of Diaspora Relations (now being integrated with the Ministry of Foreign Affairs).
25. Coordination among the different data source agencies is taking place on several levels:
- i. Political decisions are taken by the State Commission on Migration Issues (SCMI) where state agencies are represented by deputy ministers and the Commission is chaired by the Minister of Justice;
  - ii. UMAS working group has two sub-groups – IT and database administrators' subgroup and case officers/business owners' subgroup.

## 3. System architecture

26. UMAS uses a combination of new technologies widely used today in the big data world, but a novelty for the public sector. UMAS architecture is built based on Hadoop ecosystem. It has four core components – Hortonworks/HDP, HAWQ/HDB,

Talend, and Tableau. The System architecture is shown in the illustration 3. The system maintains two data storage databases, one for the raw data and the other for operational data, i.e. processed and depersonalized data. The system also keeps track of analytical models, person identification and other algorithms used over time. Hence, the modelling history is available in case analysts need to compare several models or want to revisit old analytical reports. The system also requires a common knowledge pool where information about the relevant legislation and regulatory changes are catalogued.

Illustration 3. UMAS Architecture



Source: Author's illustration

27. A special attention is paid to building UMAS knowledge base to make sure the System functions efficiently and more importantly, to build institutional memory and the ability to replicate the system.

#### 4. Methodology

28. Data integration and processing is one of the critical aspect of UMAS architecture. At present, different state registries collect data for their own specific needs and according to their internal standards. This, of course, creates a compatibility challenge when merging different datasets and identifying a unique person across the datasets. To address the compatibility and data quality issues, raw datasets received from the source state agencies undergo multiple types of technical and logical processing. Data is being cleaned for errors and duplications. Errors are either removed or corrected and filled in from other datasets, where possible.

29. Major part of the logical processing is person identification. Currently UMAS team applies deterministic person identification algorithm based on the four standard identifiers – name, surname, ID, and date of birth. However, non-deterministic algorithms will also be developed to be used in cases when the standard identifiers are missing.

30. Each time new data is added to the warehouse, it automatically goes through the standard checks; if the person is successfully identified, it is linked to the existing profile in the database, or a new profile is created. In cases of unsuccessful

identification, further checks are applied to correct the errors or to remove flawed data. After the person identification is complete, person identifiers are deleted from the high performance operational data storage. Business analysts have access to the depersonalized datasets only.

31. More in-depth analysis of received datasets and integration of new data sources may show up other challenges and needs. Hence, it is expected that the data processing methodology will be further refined over time.

## **5. Personal data protection and data governance**

32. After long contemplations it was decided that UMAS receives personal and special categories of data and stores them for unlimited time as the improvement or the revision of analytical models may require to go back to the original data and re-run person identification algorithms. Storage of personal and special categories of data further increased the pressure on UMAS team to pay special attention to data protection and data governance in the system design.

33. In Georgia personal data protection is regulated by the Georgian Law on Personal Data Protection from 2011, which has been amended twice to allow UMAS to receive personal and later, special categories of data. The Law is being enforced by the Office of the Personal Data Protection Inspector. In addition, UMAS team closely examined European data protection regulations to make sure UMAS practices were compatible with the European standards.

34. PSDA itself is one of the largest personal data depository in the country. It maintains registry of civil acts and issues national identification and travel documents, among other services. Throughout the years PSDA's relevant units have developed extensive data protection and data governance policies and practices, which will be applied to UMAS as well.

## **B. UMAS Results and Usability**

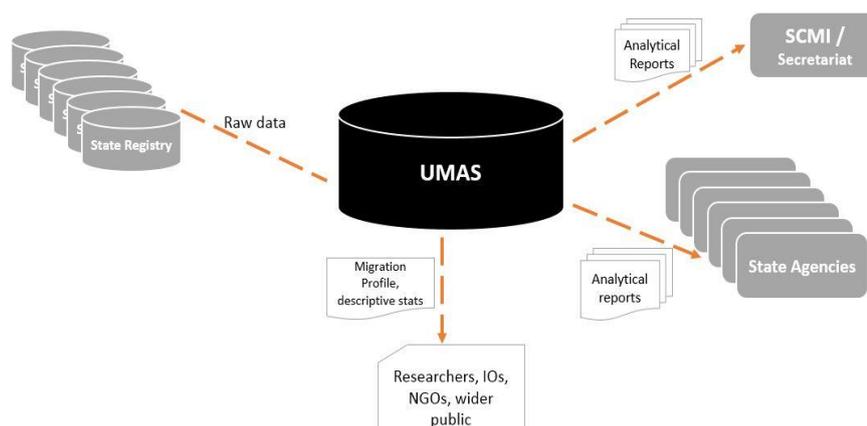
### **6. UMAS output and BI analysis**

35. Initial scope of UMAS is immigration. It is collecting data only on foreign citizens who have crossed Georgian state border. When fully operational (expected by the end of 2018 – beginning of 2019) UMAS should be able to generate real time migration descriptive statistics, analysis of immigration trends and immigrant behaviours. UMAS will not be used for individual case management; it collects data only for analytical purposes. And while it generates migration statistics in areas where statistics based on traditional administrative data is missing, its primary function is not production of official statistics. Using person identification algorithms UMAS is created to enable analysts to uncover hidden patterns, correlations, and other insights invisible for traditional business intelligence solutions and summary statistics. In the medium-term future, analytical reports produced by UMAS should shape migration policy, contribute to the improvement of migration management and migration data collection nation-wide.

36. UMAS team now operates based on the preliminary map of analytical tasks covering following 6 broad areas: (1) stocks of long- and short-term immigrants and

visitors, (2) immigration dynamics and immigrant behaviour, (3) International Protection dynamics and behaviour, (4) illegal migration trends, (5) immigrants' integration and their socio-economic impact, (6) migration risk analysis. By the type of data analysis needed, analytical reports prepared can be descriptive statistics, dynamics, factor analysis, or predictive analysis. Existing map of analytical tasks is a work in progress as is the UMAS itself. It is expected to be developed and adjusted multiple times as seen fit for the System and its end-users.

Illustration 4. UMAS users



Source: Author's illustration

37. Working version of analytical tasks have been designed keeping in mind the regular needs and interests of the UMAS end-users. After the system is fully operational, primary users of the system will be the state agencies which are members of the State Commission on Migration Issues (SCMI) and are the sources of the raw registry data to the System.

## 7. Communication and usability

38. The true value of large IT systems such as UMAS is its usability for the end-users. Sound system architecture and fast processing is important only if there is a demand on analytical reports generated. Generally speaking, such demand is anticipated from the data source state agencies. However, special measures are being taken by PSDA to properly communicate UMAS abilities and possible benefits to the users to further boost the demand.

39. To make sure the communication between UMAS end-users, business owner (PSDA) and UMAS BI analysts generating needed reports is smooth, a working group has been set up under the SCMI. A special communication scheme is designed to lay out the channels and forms of communication. Smooth communication in this context has two indicators:

- i. End-users are able to clearly formulate needed analytical tasks given that they have a good understanding of the available data and technical capacities of UMAS;

- ii. Analytical reports generated by UMAS are correctly interpreted by the end-users given that they have a good understanding of the methodology and algorithms and the quality of the data used.

## **8. Data quality and standardization**

40. Unlike classic big data, quality of administrative data is less problematic and easier to be addressed in the big data environment. However, administrative data also has its issues. Data collection and storing quality at the source data agencies as well as regulatory aspects impacting data collection have to be addressed.

41. In the development phase of UMAS, SCMI member state agencies showed willingness to cooperate and share their datasets to improve migration data collection and allow for more in-depth analysis. But further success of UMAS requires more commitments. Data source agencies will have to work on their internal databases to standardize data collection across registries and solve data quality problems on the level of data registration. In certain cases this may require collection of new data, in others – changing the way the data is registered and stored may be needed. In all cases, deeper commitment and investment of time, human, and/or financial resources will be required.

## **9. Human Capacity Building**

42. In analysis and especially, in decision-making individual intellectual capabilities as well as collective knowledge are crucial. Often, people with diverse expert knowledge need to work together towards a meaningful interpretation of the generated data. Therefore, one of the targets of UMAS is to effectively accommodate input from multiple human experts.

43. For this to happen, human resources have to be developed at PSDA as well as among the end-users of UMAS products. A set of trainings on migration data, different sub-topics of migration, BI analysis, visualization and other topics are planned for the UMAS working group members and UMAS team of analysts. Counterpart data analysts at the data source and user state agencies are going to be retrained and kept in the loop as UMAS develops.

## **V. In lieu of a conclusion – some of the tangible results achieved**

44. Even though UMAS has just started to process real data and reports are developed for the purpose of methodology and data quality testing only, some of the positive side-effects of UMAS already stand out vividly.

45. Since its establishment in 2010, institutional set-up of the State Commission on Migration Issues (SCMI) has effectively resolved the miscommunication and duplication issues among the state agencies involved in migration management. However, data source agencies have remained reluctant to share their micro-data to certain extent. Even GeoStat does not have full access to the micro-data for its manipulations (Hoffmann 2014:8). In these circumstances UMAS/PSDA managed to build functional bilateral agreements with all major data source agencies. Efficient communication mechanism is in place among the IT specialist and case officers of respective agencies. They have already accommodated multiple requests and

recommendations from the UMAS/PSDA side to share, correct or start collecting new types of data needed for the analysis. At this stage most agencies demonstrate readiness to work with the UMAS team along the way. This close partnership is being built upon the understanding that UMAS is developed to benefit first and foremost the data source agencies.

46. As mentioned above, to make sense of such vast data collective knowledge is needed. Methodological discussions that started in the narrow circles of UMAS team and working group are now expanding to include wider audience of academics and statisticians in Georgia and internationally. In the nearest future roundtable discussions are planned to operationalize the term immigrant in the context of UMAS and determine who is regarded as an immigrant in Georgia. Another set of discussions is being organized to identify possible indicators of socio-economic impact of immigration in Georgia and determine where and how UMAS data could come into play. These are the discussions that are taking place until the end of 2017. Many more will follow as other issues come to light.

47. New data available for more in-depth analysis. Even at this early stages of data processing UMAS is able to produce initial versions of analytical reports. Of course, these are test reports and are not fit for using as evidence for policy discussions yet. Still, these are good enough reports for illustrating certain trends. One of the initial reports developed attempted to break down foreign visitors into the groups of regular visitors, tourists, and transit travellers and analyse annual tendencies by their lengths of stay, border crossing, origin and other criteria.

48. New skills and expertise are acquired. PSDA has already acquired new IT and analytical skills which can be applied to migration analysis as well as to other areas of PSDA's mandate. In fact, PSDA is already studying possibilities of modernizing and improving its data warehouses using big data technology. This collective knowledge and expertise will naturally grow as UMAS continues to progress and will inevitably spill over in other state agencies involved.

---

## VI. References

- GeoStat (2016), 2014 General Population Census, Main Results, pp. 13-14.
- IBRD (2012), Fighting Corruption in Public Services. Chronicling Georgia's reforms, pp. 63-74.
- SCMI (2016), Development Guidelines: Medium Migration Profile and Brief Migration Profile, Tbilisi, Georgia.
- UNDESA (1998), Recommendations on Statistics of International Migration, Revision 1, Statistical Papers, New York, Chapter 1, pp. 1-3.
- UN (2017), Report of the Special Representative of the Secretary-General on Migration, Advanced Unedited Version, UN General Assembly, p. 25.
- Bosikashvili, Z. (2016), Analytical System of the Migration Data Georgia, High Level Solution Architecture, Tbilisi, Georgia, p.7.
- Hoffmann, E., Reichstadter, P. (2014), Assessment of Current Migration Related Data Collection Systems in Georgia. Methodological, Institutional and

- Technical Perspectives on Possibilities for (Improved) International Migration Statistics. Reinforcing the capacities of the Government of Georgia in Border and Migration Management, pp. 17-18.
- Zwick, M. (2017), EMOS Webinar Big Data II: Examples from National Statistical Institutes, Federal Statistical Office Germany, Institute for Research and Development in Official Statistics.
- Zwick, M. (2017a), Introduction to Big Data in Official Statistics, Federal Statistical Office Germany, Institute for Research and Development in Official Statistics.
- European Statistical System (2014), 22<sup>nd</sup> meeting of the European Statistical System Committee, ESS Big Data Action Plan and Roadmap 1.0., Riga, Latvia. [https://ec.europa.eu/eurostat/cros/system/files/ESSC%20doc%2022\\_8\\_2014\\_EN\\_Final%20with%20ESSC%20opinion.pdf](https://ec.europa.eu/eurostat/cros/system/files/ESSC%20doc%2022_8_2014_EN_Final%20with%20ESSC%20opinion.pdf) seen on June 29, 2017
- Georgia Tourism Strategy 2015-2025, Ministry of Economy and Sustainable Development of Georgia, p.19.
- Migration Strategy of Georgia 2016-2020, State Commission on Migration Issues, p. 41.
- [http://www.migration.commission.ge/files/migration\\_strategy\\_2016-2020\\_eng\\_final\\_-\\_amended.pdf](http://www.migration.commission.ge/files/migration_strategy_2016-2020_eng_final_-_amended.pdf) Seen on June 29, 2017
- 2015-2018 Action Plan of the State Strategy for the Labour Market Development in Georgia, Government of Georgia, p. 7.
- 2015 Migration Profile of Georgia, State Commission on Migration Issues, Tbilisi, Georgia, P. 65.
- [http://www.migration.commission.ge/files/migration\\_profile\\_2015\\_30.11.15.pdf](http://www.migration.commission.ge/files/migration_profile_2015_30.11.15.pdf) seen on June 29, 2017.
- Geostat population statistics, [http://www.geostat.ge/index.php?action=page&p\\_id=152&lang=eng](http://www.geostat.ge/index.php?action=page&p_id=152&lang=eng) seen on June 29, 2017.
- National Bank of Georgia, External Sector Statistics, Methodological notes. <https://www.nbg.gov.ge/index.php?m=306&lng=eng#external> seen on June 29, 2017.
- UNECE Statistics Wikis <https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data> seen on June 29, 2017.
- World Bank Statistical Capacity Indicator dashboard. <http://datatopics.worldbank.org/statisticalcapacity/SCIdashboard.aspx> Seen on June 29, 2017.