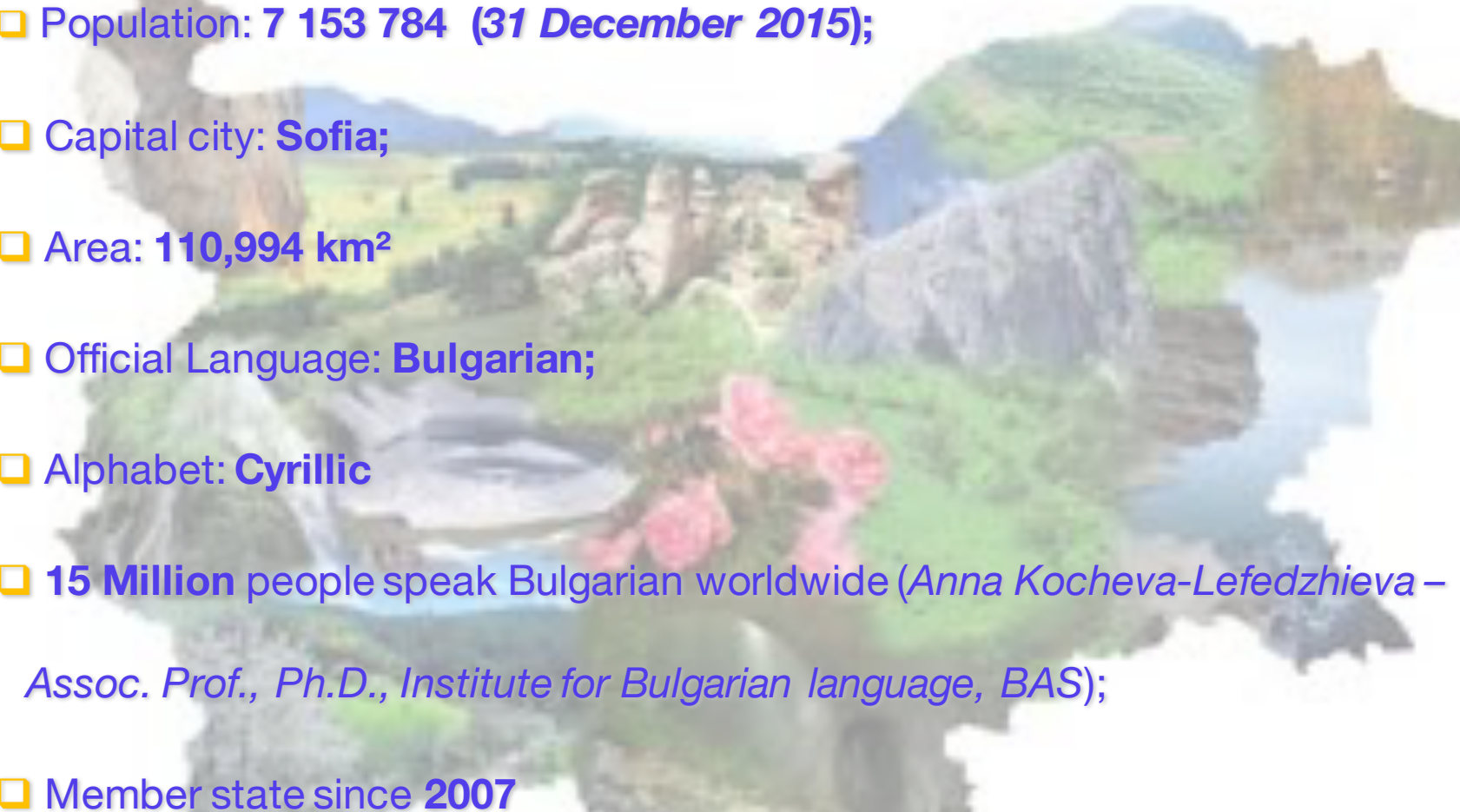


CONTEXT FOR DATA COLLECTING IN BULGARIA

Language Resources Board Meeting
Lisbon, Portugal
06.07.2016

HRISTINA DOBREVA
Chief Expert
Information Technology Directorate
Ministry of Transport, Information Technology and Communications

- 
- ❑ Population: **7 153 784** (*31 December 2015*);
 - ❑ Capital city: **Sofia**;
 - ❑ Area: **110,994 km²**
 - ❑ Official Language: **Bulgarian**;
 - ❑ Alphabet: **Cyrillic**
 - ❑ **15 Million** people speak Bulgarian worldwide (*Anna Kocheva-Lefedzhieva – Assoc. Prof., Ph.D., Institute for Bulgarian language, BAS*);
 - ❑ Member state since **2007**

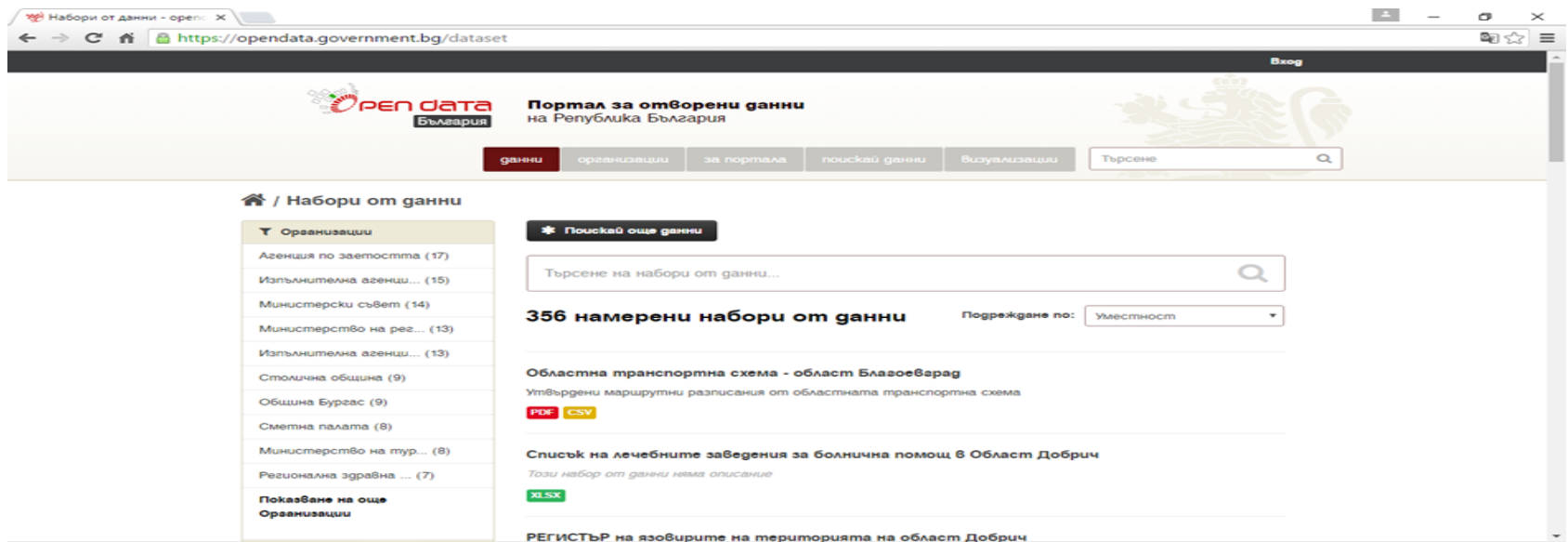


- ❑ **National Assembly;**
- ❑ **The Presidency of the Republic of Bulgaria;**
- ❑ **Administration of the Council of ministers and 17 Ministries;**
- ❑ **10 State agencies;**
- ❑ **National Statistical Institute;**
- ❑ **2 State Commissions;**
- ❑ **28 Agencies and Executive Agencies;**
- ❑ **32 Specialized Territorial Administrations;**
- ❑ **265 Municipal Administrations;**
- ❑ **35 Regional Administrations;**
- ❑ **3190 Mayoralties**





- ❑ Bulgarian Wikipedia: **218 376 articles (28 June 2016)**;
- ❑ Households who have internet access at home: **59.1% (2015)**;
- ❑ Of them - households with broadband connection: **99.5 % (2015)**;
- ❑ Public services provided from the state administrations: **2130**;
- ❑ Bulgarian open data portal: **356 data sets (28 June 2016)**

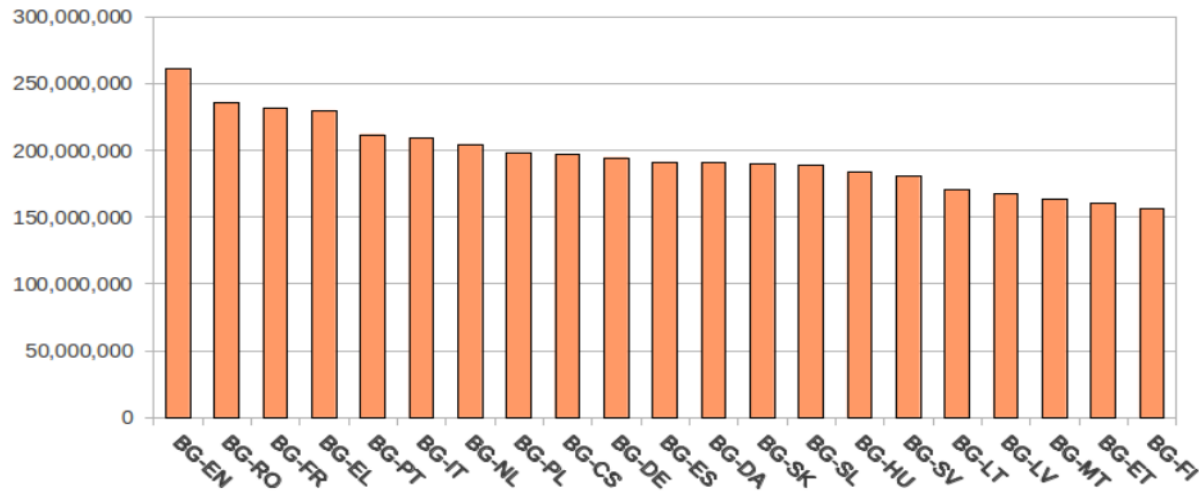


The screenshot shows the Bulgarian Open Data Portal interface. The browser address bar displays <https://opendata.government.bg/dataset>. The page header includes the 'open data България' logo and the text 'Портал за отворени данни на Република България'. A navigation menu contains 'данни', 'организации', 'за портала', 'поисков данни', and 'визуализации'. A search bar is present with the text 'Търсене'. The main content area is titled 'Набори от данни' and features a sidebar with a list of organizations and their respective dataset counts: 'Агенция по заетостта (17)', 'Изпълнителна агенци... (15)', 'Министерски съвет (14)', 'Министерство на рег... (13)', 'Изпълнителна агенци... (13)', 'Сполучна община (9)', 'Община Бургас (9)', 'Сметна палата (8)', 'Министерство на тур... (8)', and 'Регионална здравна ... (7)'. The main content area shows a search bar with the text 'Търсене на набори от данни...' and a search icon. Below the search bar, it displays '356 намерени набори от данни' and a dropdown menu for sorting by 'Уместност'. The first result is 'Областна транспортна схема - област Благоевград' with a description 'Утвърдени маршрутни разписания от областната транспортна схема' and file format icons for PDF and CSV. The second result is 'Списък на лечебните заведения за болнична помощ в Област Добрич' with a description 'Този набор от данни няма описание' and a file format icon for XLSX. At the bottom, there is a link for 'РЕГИСТЪР на азобирите на територията на област Добрич'.



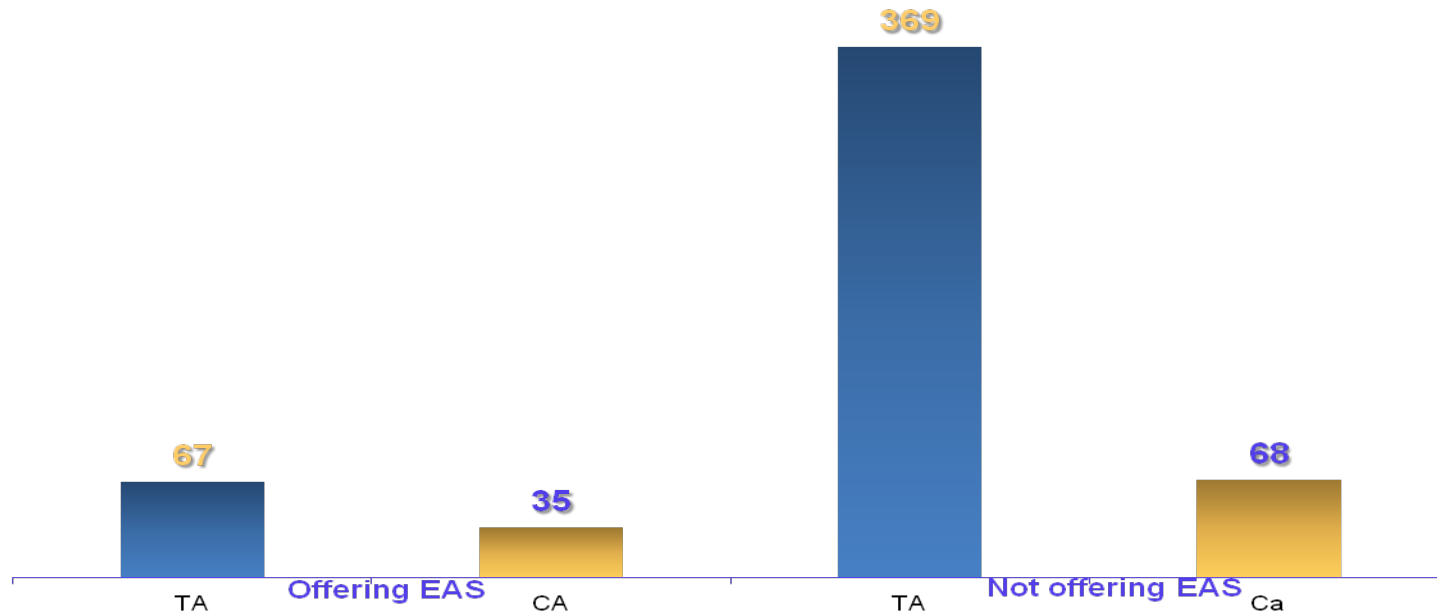
- ❑ Created at the Institute for Bulgarian Language, Bulgarian Academy of Sciences.
- ❑ Multilingual corpus which includes:
 - Over 1.2 billion words and about 240,000 texts;
 - parallel corpora of various size for 47 foreign languages (about 4.2 billion words);
 - The overall size of the corpus exceeds 5.4 billion words;
 - Texts in English, German, French, more of Slavic and Balkan languages etc.
 - Free access: <http://dcl.bas.bg/bulnc/>

Major parallel corpora in Bulgarian National Corpus





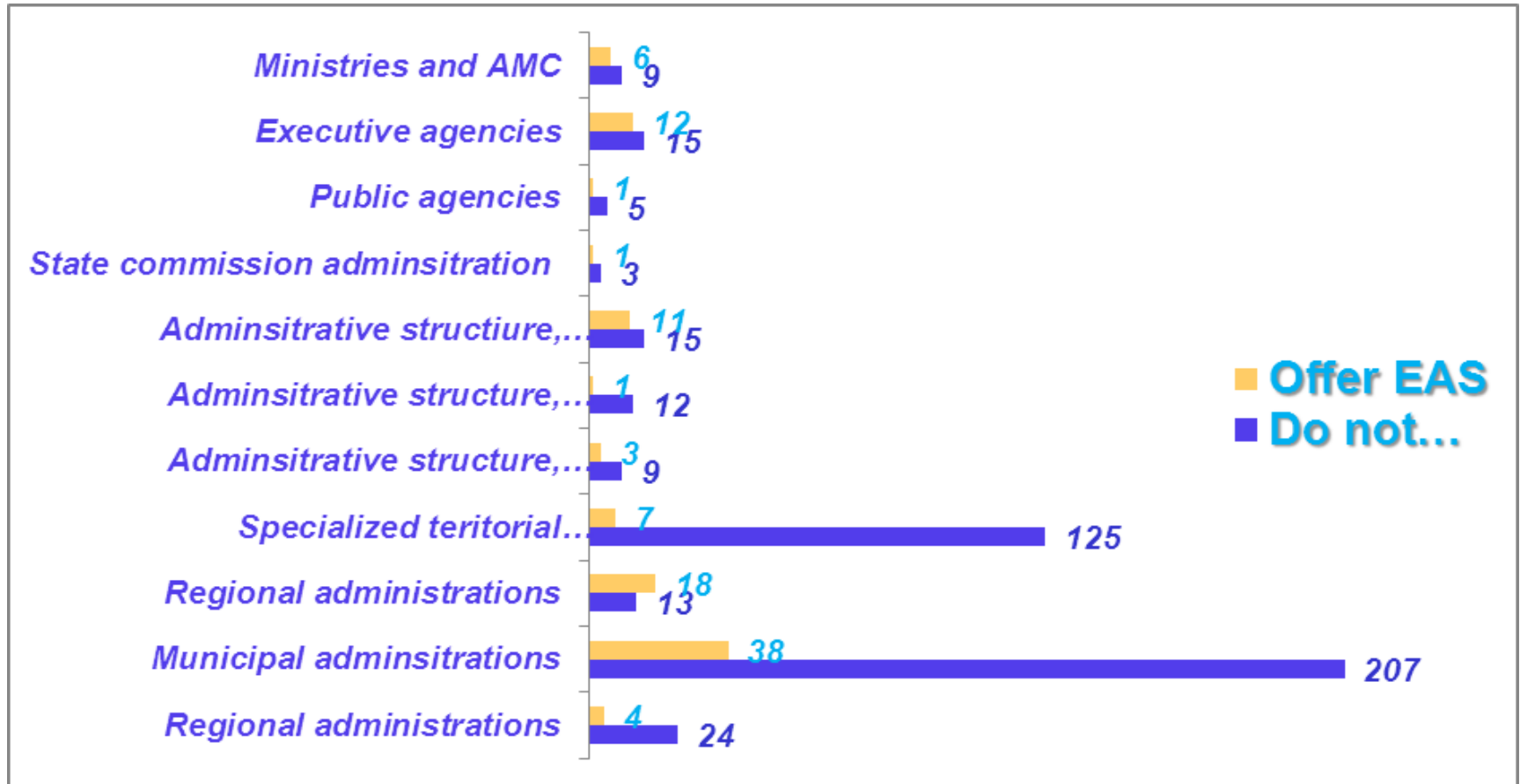
According to the Integrated Information System of the State Administration, **102** administrations provide electronic administrative services - **35** of them from the Central Administration and **67** from the Territorial Administrations.



Source: Council of Ministers Administration, Integrated Information System of the State Administration

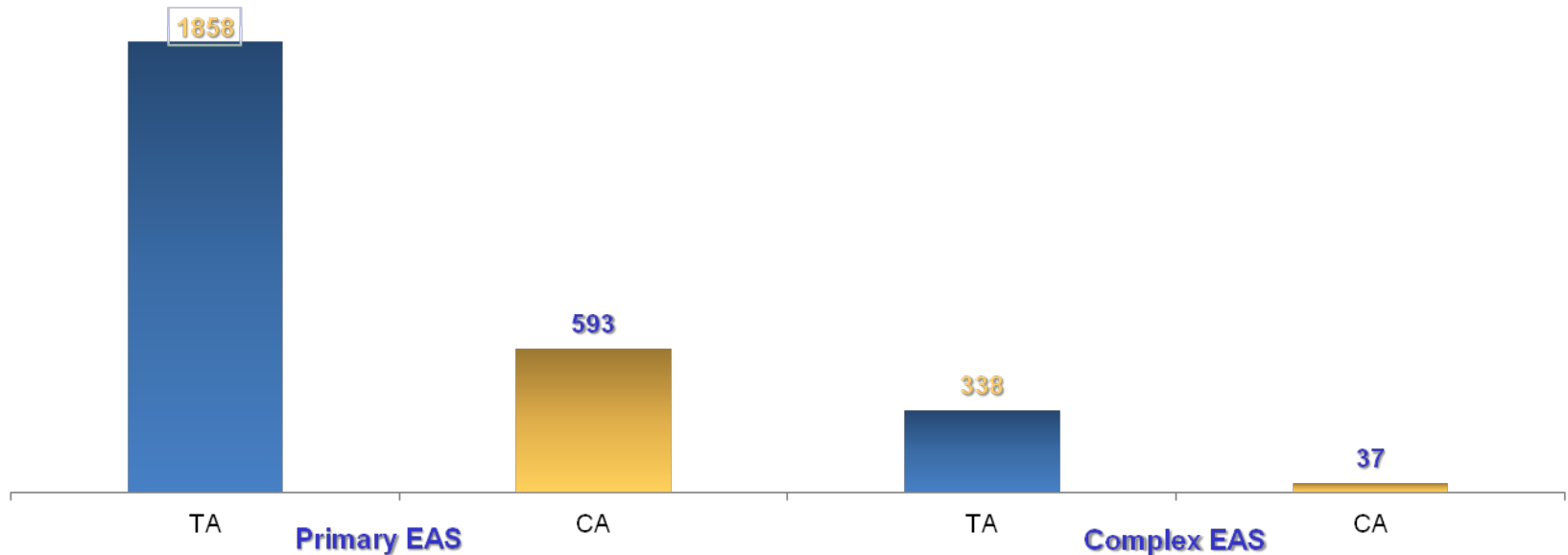


Electronic Administrative Services (EAS)/ Central and local administrations:





- The total number of the offered EAS is **2 826** (compared to the **2 525** EAS recorded in 2014).
2 451 of these are primary services and **375** - complex services.
- 437** administrations reported that they do not offer e-services yet.

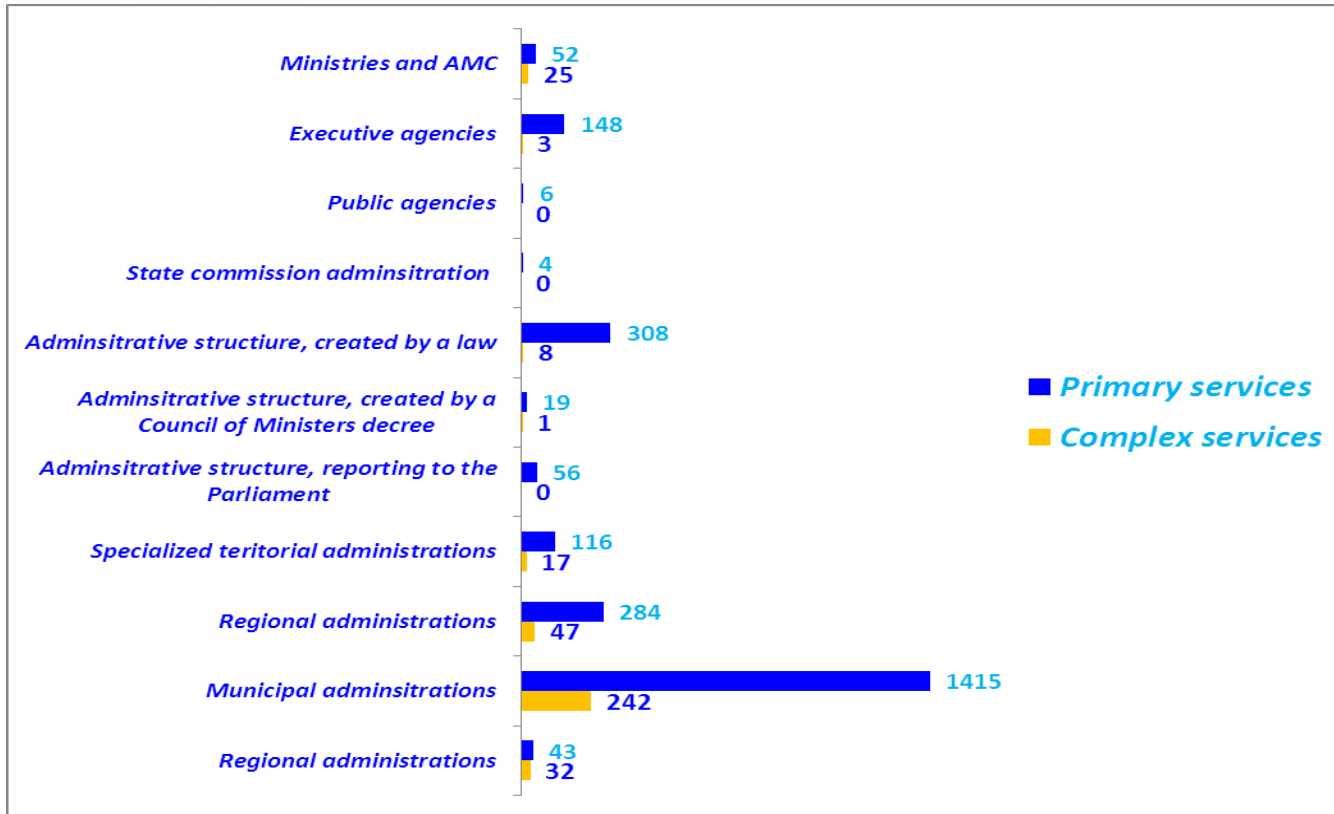


Electronic Administrative Services (EAS) 4/4



□ **Eighteen (18)** territorial administrations and **8** central level administrations have reported that they offer complex EAS.

□ Of the central administration structures, the **National Revenue Agency** offers the largest number of primary EAS - **110**, followed by the **Geodesy, Cartography and Cadastre Agency** with **70** EAS, the **"Customs" Agency** with **56** EAS and the **Directorate General "Civil Aviation Administration"** with **55** EAS.





- ❑ Held on **18 March 2016**, at the Representation of the European Commission in Bulgaria;
- ❑ Number of participants: **130**;
- ❑ **Institutions:** European Commission; Bulgarian Academy of Sciences; Ministry of Transport, Information Technologies and Communications; Ministry of Justice; Ministry of Finance; Ministry of Environment and Water (and representatives from nine other ministries of the Republic of Bulgaria); the Presidency of the Republic of Bulgaria; the Council of Ministers; National Institute of Justice; National Revenue Agency; Bulgarian Food Safety Agency; Supreme Judicial Council; National Statistical Institute; National Customs Agency; Executive Agency of Maritime Administration; Executive Agency "Automobile Administration"; Sofia University;
- ❑ **Participants** from the public administration and the private sector, including representatives of translation and interpretation agencies, NGOs, IT companies, scientists, etc.



- a) To spread the information about the initiative for collection of the language resources which will be used in the development and improvement of automated translation systems of the CEF.AT platform;
- b) To identify owners of relevant data;
- c) To present the benefits of collecting and sharing language resources;
- d) To bridge the gap between the actual needs and requirements of the public administration in Bulgaria and the potential of the machine translation systems currently offered by the European Commission;
- e) To support Bulgarian language as one of the EU languages in its efforts to gain recognition and equal status;
- f) To discuss opportunities and obstacles for data sharing etc.

National Seminar / Reasons for success



- ❑ Good collaboration between NAPs and representative of the European commission, the Directorate-General for Translation;
- ❑ Good time planning;
- ❑ Perfect organization - Institute for Bulgarian Language;
- ❑ Invitations via email to high number of potential participants – Institute for Bulgarian Language;
- ❑ Official invitations from the Ministry of Transport, Information Technology and Communications sent to all public administrations (incl. information about the project, MT@EC and CEF.AT);
- ❑ Sent reminder for the event;
- ❑ Held many personal conversations with an invitation to make a presentation and inclusion in panel discussions;
- ❑ Held a preliminary meeting with the presenters and participants in the panel discussions.



Language combinations

English↔Bulgarian

French↔ Bulgarian

German↔ Bulgarian

Spanish↔ Bulgarian

Bulgarian -Italian

Bulgarian –Russian

Macedonian- Bulgarian;

Croatian- Bulgarian;

Turkish- Bulgarian;

Japanese - Bulgarian;

Catalan- Bulgarian



Terminology: Domestic and foreign policy, national and international law, diverse depending on the field

Type of documents:

- ❑ Letters, official correspondence, invitations, greetings, thank addresses to various Bulgarian, foreign and international institutions;
- ❑ Reports on the meetings results;
- ❑ Agendas and materials for events, roundtables, workshops, conferences, official visits of delegations;
- ❑ Agreements, declarations, contracts, statements, reports;
- ❑ Statements by members of parliament, members of delegations
- ❑ Memorandums, resolutions;
- ❑ Strategic plans.

Language combinations

English↔Bulgarian

English – Russian

Type of documents (examples):

- ❑ Report on the annual consumption of fish and aquaculture households and their production in Bulgaria (Bulgarian/English)
- ❑ Letters and correspondence between the institution and the European Commission, Directorate-General for Maritime Affairs and Fisheries
- ❑ Materials for training and seminars in Bulgaria by foreign lecturers
- ❑ Investment strategies for the Operational Programmes (OP);
- ❑ Report on the implementation and actions taken under the OP, changes in legal basis concerning the governance of the OP and the sector development;
- ❑ Annual and audit reports on the OP implementation;
- ❑ Procedures relating to the implementation of projects financed under the OP;
- ❑ Commission decisions.

Terminology: fisheries, aquaculture, names of fish, legislation in the sector



Language combinations

English↔Bulgarian

English – Russian

French↔Bulgarian

Type of documents:

Materials for trainings, forums and seminars held by foreign lecturers, cases and references for a preliminary ruling.

Practical vocabulary of humanitarian law and human rights from French to Bulgarian.

Terminology: law in EU and human rights

Parallel texts (in Bulgarian and mostly in English) in the following areas:

- ❑ Tax and social security legislation (BG, EN) — generally available, basic laws translated from Apis, Ciela, and certain regulations, translated by external translation agencies;
- ❑ Bilateral agreements, including DTAs (double taxation agreements), which are international and cooperation agreements between foreign institutions and NRA;
- ❑ Correspondence of the NRA's management staff (official letters, greeting addresses, information on upcoming events);
- ❑ Information on the membership of the NRA in Intra-European Organization of Tax Administrations (IOTA), the Charter, Strategy, official letters — may also be shared after consultation with the organization.



The rules for public sector information re-use are on the top of the other legal regimes and provide as easily as possible re-use of the available public data.

Directive 2013/37 / EU amending Directive on the re-use of public sector information is focused on:

- open data;
- clearer rules for charging;
- inclusion of cultural institutions (museums, libraries and archives) within the scope of the rules on the re-use of information;
- a clear regime of exclusive rights for the re-use of information;
- standard machine-readable authorizations for re-use.

National rules transposing the Directive:

- national standards;
- conditions for re-use;
- national procedure for the provision of information;
- national open data access policy;
- short time update - 3 working days unless the law provides otherwise;
- a separate section on the websites of organizations, which gathers all rules on how to obtain information.



Amendments in the Access to Public Information act (16.02.2016):

- ❑ Implement Directive 2013/37/ EU amending Directive 2003/98 / EU on the re-use of public sector information.
- ❑ Improving the active publication of information.
- ❑ Creating a legal framework for open data portal.
- ❑ Annually planned phased publication on the internet of information resources in an open format by each organization.
- ❑ Improvements have been made in the classic mode of access to public information - can request and receive access to information electronically and without further formalities.

Supporting documents:

- Methodology approved by the Council of Ministers with objective, transparent and verifiable criteria for calculating the general fees depending on the categories and the quantities of data provided for re-use.
- Ordinance on standard terms and conditions for reuse of information and its publication in an open format (24.06.2016).



Public sector organisations, obliged to provide information:

- ☐ State bodies, the local units thereof or the bodies of local self-government in the Republic of Bulgaria;
 - National Assembly;
- ☐ Judiciary - investigation, prosecution, court;
- ☐ President of the Republic;
- ☐ Constitutional Court;
- ☐ Other public legal entities, to which by law are assigned administrative powers– National Health Insurance Fund, National Social Security Institute, Central Election Commission, Council for Electronic Media, Bulgarian National Bank, Ombudsman etc;
- ☐ Public legal organizations.



Information mandatorily published on the Internet:

- ❑ Acts - legal acts, information on public consultation on new legislation;
- ❑ information on the issuing of general administrative acts; list of issuing individual administrative acts;
- ❑ Information for administrative services to citizens – powers, functions and responsibilities, rules of organizations, internal regulations pertinent to the provision of administrative services to the public, information about the exercise of the right of access to public information, the terms and procedure for re-use of information, the fees and the formats in which such information is maintained;
- ❑ Information on the activities of the administration – strategies, plans, programmes and activities reports;
- ❑ Information in view of transparency– budget and financial statements of the administration, public procurement procedures;
- ❑ Information on the Access to Public Information Act ;
- ❑ Information for staff;
- ❑ Information of great public concern.



- ❑ Appropriate language resources that could be shared exist;
- ❑ To start the data sharing process a permission from the relevant political cabinets is necessary;
- ❑ A letter of support from the European Commission is needed;
- ❑ There are concerns about the legal aspects of data sharing;
- ❑ Translation memories, owned by the respective administration are kept in translation agencies and they could submit them to the relevant departments;
- ❑ The machine translation tool of the European Commission MT @ EU is not sufficiently known (before the seminar only 15 Bulgarian users have been registered, after it they exceed to 50).
- ❑ Translations in most of the institutions are assigned to external translation agencies by conducting public procurement;
- ❑ The quality of the translation agencies is not always satisfactory. Large volumes of texts have to be translated in a short term;
- ❑ There are no special units for translation in most institutions, mainly the international departments deal with bilingual and multilingual documents;
- ❑ The country has not developed a specific policy on multilingualism;
- ❑ Most institutional sites are bilingual (BG, EN).

Before the accession of Bulgaria to the European Union since 2001 had existed the Translation and Revision Centre (TRC) which was an administrative structure subordinated by the Minister of State Administration and Administrative Reform and was officially and institutionally committed to prepare an uniform, in terms of the terminology, translation of the *acquis communautaire*.

The center had a database which contained terms, term combinations, keywords and key expressions from all 20 fields of the *acquis communautaire* in Bulgarian, English, French and German. The database was created in cooperation with the Institute for Bulgarian Language at the Bulgarian Academy of Sciences but is not available in the moment.



A letter sent by Prof. dr Svetla Koeva to Deputy Prime Minister on Coalition Policy and Public Administration Ms Rumiana Bachvarova with information for the initiative and request for a meeting.

Meeting held on **28.06.2016**

Participants: NAPs, Mr. Atanas Slavov - Advisor to the Cabinet of the Deputy Prime Minister, Mr. Krasimir Bojanov - Director of "Modernization of the administration" directorate and Ms Nusha Ivanova expert in the same directorate.

Results:

- ❑ Main objectives of the national seminar were presented;
- ❑ Support for encouraging the sharing of language resources was requested;
- ❑ After receiving the letter from the European Commission, Ms Buchvarova will send a letter to all public administrations to launch the provision of data;
National anchor points will be listed as contacts for support at the national level;
- ❑ In operational terms, the letter will be submitted for review and amendments of NAPs;
- ❑ The visibility and the access to the resources provided in the platform were discussed.
- ❑ The initiative of the European Commission and ELRC were positively assessed.

THANK YOU FOR THE ATTENTION!