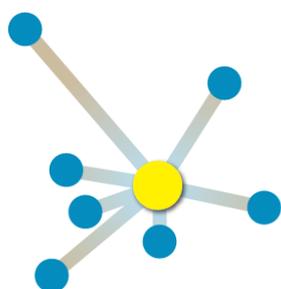


European Language Resource Coordination (ELRC) is a service contract operating under the EU's Connecting Europe Facility SMART 2015/1091 programme.



**European Language
Resource Coordination**
Connecting Europe Facility

Deliverable D3.2.2

Task 8

ELRC Workshop Report for Greece



Author(s): Maria Gavrilidou (ILSP)
Maria Giagkou (ILSP)
Kanella Pouli (ILSP)
Stelios Piperidis (ILSP)

Dissemination Level: Public

Version No.: <V1.0>

Date: November 30, 2017



Contents

<u>1</u>	<u>Executive summary</u>	<u>3</u>
<u>2</u>	<u>Workshop Agenda</u>	<u>4</u>
<u>3</u>	<u>Summary of Content of Sessions</u>	<u>6</u>
3.1	Welcoming session	6
3.2	Session 1: Connecting a multilingual Europe: European context and local needs	6
3.3	Session 2: Engage: hands-on data	8
<u>4</u>	<u>Synthesis of Workshop Discussions</u>	<u>10</u>
4.1	ELRC and the Greek Open Data Portal	11
4.2	Success stories and lessons learnt	11
<u>5</u>	<u>Workshop Presentation Materials</u>	<u>12</u>

1 Executive summary

The second ELRC Workshop took place on the 18th of October 2017 at the University of Athens, Amphitheater I. Drakopoulos.

The workshop was structured in two main sessions (after the initial welcoming session by the EC and the regional ELRC coordinator):

1. The first session focused on the European context (presented by EC officials) and on the local needs (presented by national officials and key stakeholders).

It became obvious that Greece has progressed in the last two years (since the first ELRC workshop in 2015) as regards digital policies, strategies for and implementation of digital services and transposition of the EC Regulations on Public Open Data and General Data Protection Regulation (GDPR). However, it was also openly recognized that a global change of culture of the Public Sector is deeply needed; linguistic and multilinguality issues need to be addressed for enhancing interoperability; extensive training of public sector employees on Open Data and Digital Services is already planned.

2. The second part was dedicated to presentations of ELRC: its goals and objectives, the achievements of the first phase of data collection in the country, detailed presentations of the ELRC-SHARE repository (“how-to” presentations for the preparation and sharing of data and new services offered), as well as presentation of the National and European legal framework for language data sharing.

The last part of the workshop was a very lively Questions & Answers session, during which crucial issues were raised and discussed:

- Access to MT@EC/eTranslation: rights, conditions and restrictions of use; request to granting of access also to freelance translators (subcontractors of the official Translation Service)
- Comparison of the performance of MT@EC with popular online services (Google, Bing) available to everyone
- Role of human translators in the loop: fear of extinction of human translation professionals, or transformation of translators to simple post editors
- Success rate of eTranslation in relation to domains and/or text types
- Some organizations expressed the reservation that their internal data might not be of high enough quality appropriate for CEF eTranslation
- Confusion of ELRC-SHARE (the repository) with MT@EC / eTranslation
- The internal administrative hierarchy of the public sector often hinders data collection process. A plea for help with this issue was expressed by the participants
- Country specific issues
 - need for the treatment of non-European languages, due to refugee populations
 - request for dissemination of the ELRC action to a broader audience within Public administrations
 - combine dissemination of ELRC with the dissemination concerning the obligation for the implementation of GDPR in all organizations of the public sector (by May 2018 in Greece).

Based on the number of participants, the interesting issues touched upon, both by speakers and by the audience, and the degree of involvement of the audience in the discussions, the second ELRC Workshop in Greece is considered very fruitful and insightful.

2 Workshop Agenda

Connecting Europe without language barriers: Greece in the European digital services landscape

The second European Language Resource Coordination (ELRC) workshop in Greece
Amphitheater I. Drakopoulos, University of Athens
Athens, 18 October 2017

08:00 – 09:00 Registration

09:00 – 09:10 Welcome and introduction
Stelios Piperidis, ILSP/"Athena" R.C., ELRC

09:10 – 09:15 Welcome by the EC
Panos Carvounis, Head of EC Representation in Greece

Session 1. Connecting a multilingual Europe: European context and local needs

09:15 – 09:35 Connecting public services across Europe: ambition and results so far
Aleksandra Wesolowska, Project Officer, Directorate-General Communications Networks, Content and Technology, EC (live video link, interpretation in Greek)

09:35 – 09:45 The position and prospects of Greece in a digitally connected multilingual Europe
Grigoris Theodorakis, Secretary General, Ministry of Administrative Reform

09:45 – 10.50 CEF in Greece: an outlook into current and future challenges – Panel session
Moderator:
Nancy Routzouni, Ministry of Administrative Reform, ELRC Public Services National Anchor Point
Panelists:
Leandros Maglaras, Director of Cybersecurity, General Secretariat for Digital Policy, Ministry of Digital Policy, Telecommunications and Media
Georgios Stefanopoulos, Board Member, eGovernment for Social Security (IDIKA S.A.), Special Advisor to the Minister of Health on eHealth matters
Ioannis Katopodis, Director, eGovernment Directorate of the Ministry of Education, Research and Religious Affairs
Georgios Papageorgiou, Appeal Judge, Legislative Initiative Office, Ministry of Justice, Transparency and Human Rights

Titika Dimitroulia, Director, National School of Public Administration and Local Government

Konstantinos Hambidis, City of Athens Chief Digital Officer

Michalis Vafopoulos, NCSR “DEMOKRITOS”, Open Data Institute - Athens node

10:50 – 11:10 The CEF eTranslation platform @ work

Markus Foti, MT@EC/eTranslation Project Manager, Directorate-General for Translation, European Commission (interpretation in Greek)

11.10 – 11:40 Coffee Break

Session 2. Engage: hands-on data

11:40 – 12:00 The European Language Resource Coordination (ELRC) action

Stelios Piperidis, ILSP/“Athena” R.C., ELRC

12:00 – 12:15 ELRC in Greece

Maria Gavriilidou, ILSP/“Athena” R.C., ELRC Technology National Anchor Point

12:15 – 12:45 Can language data be shared and how? – National and European legal framework

Prodromos Tsiavos, Law and policy advisor, UCL / Media Institute, “Athena” R.C.

12:45 – 13:45 Lunch Break

13:45 – 14:10 Preparing and sharing data with the ELRC repository

Maria Giagkou, ILSP/“Athena” R.C., ELRC

14:10 – 14:25 New services by the ELRC consortium

Prokopis Prokopidis, ILSP/“Athena” R.C., ELRC

14:25 – 15:00 Identifying and managing your data: Questions & Answers

Moderator: **Maria Gavriilidou**, ILSP/“Athena” R.C., ELRC Technology National Anchor Point

15:00 – 15:10 Discussion and Conclusions

Stelios Piperidis, ILSP/“Athena” R.C., ELRC

15:10 – 15:30 Coffee Break and networking

3 Summary of Content of Sessions

A video installed on a TV screen in the foyer of the amphitheater welcomed the participants to the second round of ELRC workshops. In this video, Rytis Martikonis, Director-General of the DGT and Claire Bury, Deputy Director-General of DG-CONNECT, gave a brief introduction on the MT@EC and elaborated on the use of eTranslation for the support of CEF digital services. This video played during the whole workshop, together with two additional videos, i.e. the “eTranslation: Making European digital public services multi-lingual” video, available at the CEF Digital website, [Media Library](#), and the [ELRC promotional video](#).

3.1 Welcoming session

The first session consisted of two welcoming speeches, the first by Stelios Piperidis, the local ELRC representative, who opened the event and the second one by Panos Carvounis, Head of the EC Representation in Greece.

Stelios Piperidis gave the context of the workshop, presented the overall framework of ELRC, emphasized the importance of data and specifically textual data and language resources and briefly guided the audience through the agenda.

Panos Carvounis stressed the importance of the CEF instrument in supporting the DSIs and, specifically, the investment in machine translation. He pointed out the potential for improvement offered to MT@EC provided the system is trained on national administrations texts.

3.2 Session 1: Connecting a multilingual Europe: European context and local needs

Aleksandra Wesolowska, Project Officer, participating via teleconference (with live video link), presented the European context; that is, the strategy behind CEF Digital and the Digital Service Infrastructures (DSIs) and the building blocks, with special emphasis on the eTranslation building block. Furthermore, she presented the results achieved so far in terms of the uptake of the CEF DSIs and building blocks in different countries and concluded with the need for the involvement and connection of the national public administrations with eTranslation and the respective benefits.

The two subsequent parts of the session focused on the local scene and the local needs.

The **Secretary General of the Ministry of Administrative Reform**, Grigoris Theodorakis, presented the position and prospects of Greece in a digitally connected multilingual Europe. In his talk, he highlighted Europe's multilinguality, and the need for its preservation through multilingual digital services available to all countries; he presented the Greek government's strategy and implemented initiatives as regards national public digital services (public open data, open Government data portal, eProcurement), and emphasized the need for interoperability between the public bodies' information systems. He concluded by expressly stating the commitment of the Ministry of Administrative Reform to support ELRC's action to improve eTranslation, recognizing the added value it will offer to the Greek public administration internally, but also in its cross-border communications with EU countries' administrations, businesses and citizens.

The second part of this session, titled **CEF in Greece: an outlook into current and future challenges**, was a panel moderated by Nancy Routzouni, Greek Public Sector NAP, featuring representatives of a broad range of public sector organizations, who elaborated upon the uptake of digital services in the country in their respective domains.

Concerning the issue of **Cybersecurity**, the Director of Cybersecurity, General Secretariat for Digital Policy, Ministry of Digital Policy, Telecommunications and Media made special reference to the EU regulations for security and privacy which are already or will soon become effective. The interconnection of systems renders them more vulnerable to new unforeseen threats; secure systems,

however, are a prerequisite to increasing users' trust in digital services. Given the importance of the human factor for the operation of security systems, the need for training on security becomes imperative. The Greek Digital Strategy includes security guidelines for the public sector, prepares an action plan which specifies the required activities towards this goal and the harmonization of the Greek legislation with NIS regulations is underway.

The country's initiatives on **eHealth** and **Social security** (EESSI) as described by the Special Advisor to the Minister of Health on eHealth matters and also representative of the eGovernment organization for Social Security (IDIKA S.A.), include the implementation of eIDAS within the Greek health care system, cross-border availability of patient files and e-prescription, for which the need for eTranslation is intense. An important side-effect of the need to adhere to CEF concerned interoperability issues: it triggered the need for internal organization and documentation practices to prepare for cross-border connectivity. The need for translation tools for non-EU languages was also stressed, to be used in the country for communication with non-EU citizens.

As regards the domain of **Education**, the Director of the eGovernment Directorate of the Ministry of Education, Research and Religious Affairs mentioned that 226 eGov services have been implemented in the last 3 years, e.g. e-submission of applications to universities, applications for housing allowance for students, connections with other public services, e.g. the social security number service and the taxation/income service. Open educational data have been made available by the Ministry of Education to teachers, students and the broad public, mainly for educational research/analytics purposes, on an open access basis. The National Educational Data Aggregator (<http://photodentro.edu.gr/aggregator/>) through a unified access point offers digital educational data (mainly Greek monolingual material), learning objects, search facilities and links to third repositories.

The representative of the Greek Ministry of Justice, Appeal Judge and member of the Legislative Initiative Office, Ministry of Justice, Transparency and Human Rights described the participation of the Greek Ministry of Justice in CEF **eJustice** and the interconnection of Greek courts with other EU countries, so that any application can be filed remotely. Automatic transcription and translation of court minutes is a long-standing demand; it is worth mentioning that the translation of published decisions (not minutes) of the Greek Supreme Civil and Criminal Court (<http://www.areiospagos.gr/>) at the moment is covered by Google translate!

The Director of the National School of Public Administration focused on the need for **training of public sector officials**. The public sector needs a shift of mindset/culture, involving active engagement with openness and multilinguality. In this direction, a Digital Governance department has been founded, offering specialised courses to public servants on Economics and Development, Public Management and Public Administration, eGovernment and ICT, Digital Systems Security, Open Data, Privacy and relevant Legislation, Collaborative Tools, Translation tools (Translation Memories, CEF Automated Translation); furthermore two workshops are being organized on the issues of the EU regulation on data protection officers and on the use of CEF Automated Translation in public bodies.

Regional government was represented by the City of Athens Chief Digital Officer, who focused on the plans to support the citizens' and visitors' digital life and embrace innovation and knowledge, by implementing a strategy related to 5 policy sectors: *Infrastructures*, *Digital services for the citizens* (digital signature is adopted by 80% of the city services, interconnection with the tax office is in place, as well as issuing of digital documents), *content* (should be accessible in many languages), *education* (current action "open schools" offers digital literacy classes to adults) and *innovation* (activities such as Innovathens and Athens Digital Lab). A further challenge for the city of Athens is the recent change of its anthropology, due to the incoming of populations from the East: tools and applications that were designed mainly for the western world need adaptation to the new users' needs.

Finally, an actual use case of **exploitation of open public data** was presented, referring to the comparison of open governmental data (e.g. public spending) from different countries. The project

YourDataStories (<https://yourdatastories.eu/>) which undertook this task, faced the language barrier while trying to understand public procurement and spending in different countries (<https://yourdatastories.eu/about-your-data-stories/use-cases/>). Multilinguality is far from being the norm in governmental organisations' portals of the EU member states and even in national open data portals. In some of these portals there is no real multilingual content, just translations of metadata at best.

The panel session showed great improvement of the Greek public sector: the previous workshop of 2015 had attested great diversity as regards the degree of digital literacy and digital services offered by various organizations of the public sector. The present workshop, however, showed great progress towards the digital competence of the Greek public sector (fact openly acknowledged by the audience).

The last presentation of the first session concerned **The CEF eTranslation platform @ work**. Markus Foti, MT@EC/eTranslation Project Manager presented the eTranslation platform, its principles and mode of operation, its intended and current users (in terms of DSIs), the steps required to connect with the platform and the benefits of using MT@EC/eTranslation.

The presentation was followed by a demo of the system at work, which displayed the system's capabilities and flaws by applying it on a broad range of domains and text types, in EU documents and Greek texts produced by other sources; the presentation concluded with plans for future improvements of the eTranslation platform.

The audience vividly expressed their interest for the eTranslation platform, and their appreciation for the demo. The questions can be classified in the following broad categories:

- Access issues: who has the rights to use the system, under what conditions, what are the restrictions.
- The request for use of MT@EC/eTranslation by freelance translators (subcontractors of the official Translation Service).
- Comparison of the performance of MT@EC with popular online services (Google, Bing) which are available to everyone.
- Role of human translators in the loop: fear of extinction of human translation professionals, or transformation of translators to simple post editors.
- Success rate in relation to domains and/or text types.

3.3 Session 2: Engage: hands-on data

The second session covered a broad range of aspects of ELRC, namely, the ELRC action in general and in the country, the legal framework of data sharing, the procedure for contributing data to the ELRC repository and new services offered by ELRC, and closed with a final session of discussion with the audience, during which questions were addressed to all speakers of the event.

The opening presentation of this session (Stelios Piperidis) focused on **the European Language Resource Coordination action**, presented the consortium and its goals, its activities and the current situation as regards data collection at the European level, the repository developed and the services offered by the helpdesks to data contributors and users.

The next presentation (Maria Gavriilidou) zoomed in on Greece, and presented **ELRC achievements at the national level** as regards the very successful data collection activity of the first phase, the efforts related to the introduction of a change of culture in the Greek public sector as regards open data, the procedure followed for the mobilisation of the public sector and the attempts made to train public sector contributors on the concept of data and metadata curation. The obstacles faced (whether successfully overcome or not yet) gave fruitful input for actions in the current phase.

In presenting the **legal framework for data sharing** Prodromos Tsiavos stressed the importance of the modification of the PSI Directive in order to accommodate the re-use of public data and its transposition to the Greek legislation, whose scope for re-use of public data has been extended to cover Museums, Archives and Libraries; furthermore, it defines "openness by default", introduces the use of standard open licenses, and establishes the Public Sector Open Data Registry. The presentation continued with a list of the necessary steps for the opening of public sector data, protecting at the same time the contributors from violating any existing rights.

The following presentation (Maria Giagkou) focused on the technical aspects **of preparing and sharing data through the ELRC repository**: types of data needed by eTranslation, appropriate data formats and preferred subject domains were explained and exemplified, tips for good and bad practices were given as regards data preparation and management. The presentation also included a detailed, step-by-step guided tour of the procedures of registration to the repository and contribution of data.

The focus of the last presentation of this session (Prokopis Prokopidis) was on the **new services offered by the ELRC consortium**, aiming to facilitate public sector data contributors with all stages of data extraction, conversion to appropriate formats, data clean-up, alignment, anonymization, and metadata curation and validation. Last but not least, the presentation explained the services of on-site assistance and the help desks' function.

The last part of the workshop was formulated as a **Questions and Answers session**, introduced by a presentation (Maria Gavriilidou) of the imminent procedures to be implemented by the national governments, concerning data management and data protection (implementation of the General Data Protection Regulation, GDPR), thus connecting the contribution requested by ELRC from the public sector with the broader scene of Data Management in view of improved multilingual digital services.

In this interactive session, the audience raised points for discussion and posed questions to all speakers/panellists. Some of the questions had already been posed to the respective sessions, but were repeated in the final round; this is an indication of the importance these issues have for the audience. The following section provides a concise synthesis of the workshop discussions.

4 Synthesis of Workshop Discussions

The main points raised not only during the panel and the QA sessions but also after the various presentations are:

1. Access to MT@EC / eTranslation
 - a. Public sector administrative and technical obstacles: not all public organizations create official governmental email accounts for all their employees, especially those on short contracts. The use of personal email accounts is frequent, leading to problems of access to MT@EC.
 - b. Freelance translators with renewable (not permanent) contracts with Public bodies: in this intermediate situation translators are subcontractors with the public sector being their sole client; as they are not officially public sector employees, they do not have official governmental email accounts and cannot have access to MT@EC. However, they may aggregate their translated data (translation memories and term lists generated out of their different public sector assignments) and need incentives (e.g. access to MT@EC among others) in order to share their data.
 - c. State owned companies: their status is not clear as regards access rights.
2. During the final session of Q&A it became obvious that some participants had confused ELRC-SHARE (the repository) and MT@EC / eTranslation: they believed that by registering to the repository they gained access to the platform. The difference should be stressed immediately at the beginning of the next ELRC workshops to avoid misunderstandings.
3. There was a recurring issue of various stumbling blocks posed by the internal administrative hierarchy of the public sector; these include
 - a. The need for official decision for the appointment of the person responsible within the organization;
 - b. The need for permission to abandon other tasks in order to undertake ELRC data collection task;
 - c. The approval of the data selected by the responsible person by a higher-rank official.

A plea for help with this issue was expressed by the participants.

4. The comparison of MT@EC/eTranslation with available online services and the role of human translators in the loop (human translators reduced to post editors of automatically produced translations or even replaced by well-trained MT systems) were issues already raised after the presentation of the eTranslation platform, but they were discussed again during the closing session.
5. The quality issue was often raised, from two perspectives, namely,
 - a. data quality: organisations that have and produce data potentially useful to CEF eTranslation are sceptical about their quality and therefore are reluctant to deliver them and
 - b. translation quality: the need for translation quality validation is considered crucial, especially for official documents.
6. Country specific issues
 - a. In Greece there is a need for the treatment of non-European languages, due to the need to facilitate communication with refugee populations, their contact with the country's administrative services, and education of refugee children and adults.
 - b. Public sector employees who were interested but who were not reached by the circular sent out by the Secretary General recommended to ELRC NAPs in Greece to request a third round of the circular
 - i. to a broader audience, i.e. not just the political hierarchy (Ministers and General Secretaries), but to the Organizations' hierarchies, e.g. General

Directors, Heads of Administrations, to core and peripheral public bodies, to regional organizations etc.

- ii. to combine and coordinate this circular with the dissemination concerning the implementation of the General Data Protection Regulation (GDPR) in all organizations of the public sector (by May 2018 in Greece).

4.1 ELRC and the Greek Open Data Portal

The Greek Open Data Portal (www.data.gov.gr) is the central catalog of the Greek public sector data, offering access to the open data of the Greek Government's organizations. It implements the country's open data policy which adheres to the relevant national legislation following the EC Directive 2013/37/EE.

Its objective is to facilitate and increase access to digitally readable datasets, by offering services for cataloguing, indexing, storing, maintaining and searching of public sector data, as well as web services for citizens and third party information systems. The Greek Open Data Portal is the responsibility of the Ministry of Administrative Reform <http://www.ydmed.gov.gr/>.

It currently contains 5,9k datasets from 260 Public Organizations from the central government and the regional authorities, the State Archives, the Hellenic Statistics Authority etc.; the majority of the data (45%) is in the form of spreadsheets (typically .xls), while the rest data formats are html pages (12%), pdf (10%), doc (3%) and txt (0,9%) files. The thematic domains of the datasets include, indicatively, finance, statistics, administration, elections, health, transport, labor as well as geospatial data.

ELRC repeatedly stressed that public language data (i.e. public documents, monolingual or translated, term lists and thesauri, etc.) do fall within the scope of the PSI directive and as such should be made openly available.

4.2 Success stories and lessons learnt

Success stories identified during the first phase of ELRC data collection in the country as presented at the workshop refer to the following issues:

4.2.1 How to contact the Public Administration: a good practice

The procedure followed by the Greek public administration was recognized as good practice: setting up the procedure was based on the need for the motivation of the public administration by a high-level official, positioned in the appropriate Ministry. This was achieved by the issuing of a circular from the Ministry of Administrative Reform (so-called the "Ministry of Ministries") to all other Ministries, which specified in detail the steps of the procedure. The role of the Ministry of Administrative Reform was crucial for the motivation of the public administration. The procedure adopted was the following:

- the Greek ELRC coordinator and the T-NAP formally informed the Ministry about the ELRC initiative
- the P-NAP informed the Secretary General responsible for Digital Governance about ELRC and its objectives and convinced him about its significance
- the T-NAP (with the advice of the P-NAP as regards the public administration procedures) drafted the circular to be issued by the Secretary General to all other Ministries with the request for broad dissemination to all supervised organisations

- The circular explained the ELRC initiative and its objectives and focused on the request for the collection of public administration language resources. Furthermore, it specified the procedure to be followed within each Ministry, whereby
 - each organization appoints a contact point, responsible for data collection and communicates his/her details to ILSP (the Greek ELRC team)
 - each contact point registers to ELRC-SHARE repository
 - after consultation with his/her service, the contact point collects the data and uploads them to the repository.
- The contact points from all Ministries formed a network, set-up and closely-monitored by the Greek ELRC team with meetings and teleconferences, and continuously supported via the ELRC helpdesks, telephone and email. The formation of the network greatly aided the data collection procedure, as it created a community of people sharing a common goal.
- Establishing a friendly and collaborative relation with the public administration contact points was a crucial success factor.

4.2.2 Unified procedure for data collection

The implementation of a unified procedure for the identification, documentation and uploading of relevant data greatly boosted data collection.

4.2.3 Textual data of the public sector are indeed open data

The recognition (by the public sector) of the fact that textual data also fall in the scope of the PSI Directive (and not only statistical, geospatial etc. data) was also a success.

4.2.4 Valuable lessons learnt, in a nutshell

- The circular issued by the Ministry of Administrative Reform proved effective in unlocking data in many cases. However, since changes of roles in the public bodies' hierarchies are not infrequent, such circulars need to be re-issued and a more stable network of data providers within the public bodies should be established.
- Public sector employees are in need for training, initially on digital literacy and subsequently on the value of open data and the use of eTranslation.
- Open textual data should be translatable across the EU to facilitate cross-border reuse for a wide spectrum of use cases and applications (e.g. public spending comparability as mentioned in section 4.2 above).

5 Workshop Presentation Materials

The presentations are published on the Greek workshop agenda webpage (http://lr-coordination.eu/el/L2Greece_agenda and <http://lr-coordination.eu/node/665>).