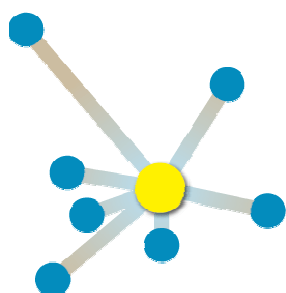


European Language Resource Coordination (ELRC) is a service contract operating under the EU's Connecting Europe Facility SMART 2014/1074 programme.



**European Language  
Resource Coordination**  
*Connecting Europe Facility*

## **Deliverable Task 6**

# **ELRC Workshop Report for Malta**



<b>Author(s):</b>	Michael Rosner (University of Malta)
<b>Dissemination Level:</b>	Public
<b>Version No.:</b>	V1
<b>Date:</b>	09.03.2016



## Contents

### Table des matières

1. Executive Summary.....	3
2. Workshop Agenda .....	4
3. Summary of the Content of Sessions .....	6
3.1. Session 1: Opening .....	6
3.2. Session 2: Local Welcome .....	6
3.3. Session 3. Objectives of the Workshop .....	6
3.4. Session 4. Europe and Multilingualism .....	7
3.5 Language and language technologies in Malta .....	7
Session 3.5a Language and language technologies in Malta .....	7
Session 3.5b The State and Status of Translation in Malta.....	8
3.7. Session 7: Machine translation: how does it work?.....	8
3.8. Session 8: How can Public Institutions benefit from the CEAF.AT Platform?.....	8
3.9/10. Sessions 9 and 10: The data for CEF-AT .....	9
3.11. Session 13: Data and language resources: Technical and practical aspects.....	9
3.12. Session 12: Wrap Up .....	9
4. Synthesis of Workshop Discussions .....	10
4.1. Panel 1: Public multilingual services in Malta.....	10
4.2 Panel 2: Data and language resources for Malta: where to find them .....	11
4.3 Interactive Session 12: How can we engage? .....	13
5. Workshop Presentation Materials.....	15
6. Appendix.....	16
6.1. Invitation letter (in Maltese) .....	16
6.2. Workshop Agenda (in Maltese) .....	17
6.3. Attendance Certificate (in English).....	19

## 1. Executive Summary

This document reports on the ELRC Workshop in Malta, which took place on the 26<sup>th</sup> of February 2016 in Valletta, at Dar I-Ewropa, the Representation of the European Commission in Malta. It includes the agenda of the event (section 2) and briefly informs about the content of each individual, interactive and panel workshop session (section 3).

The ELRC-Malta Workshop was attended by 37 participants, distributed as follows:

	No. of participants
<b>University</b>	13
<b>Maltese Government</b>	19
<b>EC</b>	3
<b>Industry</b>	1
<b>ELRC</b>	1

Government attendees came from a wide variety of offices and Ministries, as can be seen from the main participation list.

The workshop agenda was modelled mainly after the supplied template, with some minor alterations as follows: the sessions on language in Malta was split into to subsessions, the first dealing with language and related technologies, and the second with the translation scene. In the afternoon, when time begin to press, the sessions on the data needed and the legal framework for delivering it were combined

The dedicated event page can be found at <http://lr-coordination.eu/mt/malta>

## 2. Workshop Agenda

### Agenda for CEF.AT - ELRC Training Workshop (European Language Resource Coordination)

DAR L-EWROPA,

254, Triq San Pawl, Valletta VLT 1215, MALTA

26.02.2016

08:30 – 09:00	Registration
09:00 - 09:05	Opening <b>Mike Rosner</b> , ELRC National Anchor Point
09:05 – 09:20	Local Welcome <b>Elena Grech</b> , Head of EC Representation in Malta
09:20 – 09:30	Workshop objectives <b>Khalid Choukri</b> , ELRA/ELDA
09:30 – 09:50	Europe and multilingualism <b>Brian Buhagiar</b> , EC Representation in Malta
09:50 – 10:10	Language and language technologies in Malta <b>Ray Fabri</b> , Council for the Maltese Language
10:10 - 10:30	The State and Status of Translation in Malta <b>Joe Eynaud &amp; Clare Vassallo</b> , Dept Translation & Interpretation, University of Malta
10:20 – 11:00	Panel: Multilingual Public Services in Malta <b>Moderator:</b> Marc Bartolo, Malta IT Agency Panelists: <b>Thomas Pace</b> , Council for the Maltese Language <b>Joseph Azzopardi</b> , Malta IT Agency (Public Service Initiative) <b>Claudio Laferla</b> , Malta IT Agency, (Knowledge Management Centre) <b>Vanni Bruno</b> , Office of the Attorney General, (Justice Services)
11:10 – 11:40	<b>Coffee Break</b>
11:40 – 12:00	Machine translation: how does it work? <b>Mike Rosner</b> , University of Malta
12:00 – 12:30	How can Public Institutions benefit from the CEF.AT Platform? <b>Carmen Carabott</b> , EC, DG Translation, Head of Language Applications
12.30 – 13.30	<b>Lunch Break</b>

- 13:30 – 14:00 What data is needed for CEF-AT?  
**Khalid Choukri**, ELRA/ELDA
- 14:00 – 14:30 Legal framework for contributing data.  
**Khalid Choukri**, ELRA/ELDA
- 14:30 – 15:30 Panel: Data and language resources in Malta  
Moderator: **Mike Rosner**, University of Malta
- Panelists:
- Dr Albert Gatt**, Institute of Linguistics, University of Malta  
**Carl-Joseph Farrugia**, Dept of Information, Office of the Prime Minister  
**Chris Gruppetta**, Director of Publishing, Merlin Publishing Ltd  
**Victor Cauchi**, Office of the Attorney General, Ministry of Justice
- 15:30 – 16:00 **Coffee Break**
- 16:00 – 16:30 **Interactive Session: How can we engage?**  
**Chair: Dr Albert Gatt**, Institute of Linguistics, University of Malta  
Discussants:  
**Michael Spagnol**, Dept Maltese, University of Malta  
**Dr Odette Vassallo**, Centre for English Language Proficiency, University of Malta  
**Mark. S. Spiteri**, Malta Communications Authority  
**Carmen Carabott**, DG Translation, Head of Maltese Language Applications
- 16:30 – 17:00 Data and language resources: Technical and practical aspects.  
**Khalid Choukri**, ELRA/ELDA
- 17:00 – 17:15 Wrap-up and Conclusions.  
**Mike Rosner**, ELRC National Anchor Point

## 3. Summary of the Content of Sessions

### 3.1. Session 1: Opening

Mr Mike Rosner, in his role as national anchor point for the ELRC expressed his welcome message to the audience. After explaining the diverse and unusual constitution of the audience, which included participants from Government, University, and Industry, he thanked Dr Elena Grech, Head of the EC Representation in Malta, for hosting the meeting in Valletta and attending the workshop personally. He also expressed his gratitude to Mr. Brian Buhagiar, DGT, currently in residence at Dar L'Ewropa, and Mr. David Sant, MITA, for their unfailing support during the preparation phase. He encouraged participants to contribute with suggestions, needs and questions, and then introduced the Head of the EC Representation in Malta

### 3.2. Session 2: Local Welcome

Dr Elena Grech, Head of the EC Representation in Malta, welcomed the audience and thanked Khalid Choukri in particular for attending. She pointed out that although Maltese is the national language, Maltese and English are both official languages and that as a result, the education system is mixed. Maltese citizens are educated in both languages, that this brings advantages regarding job opportunities that are of benefit not only to Maltese, but to those seeking coming to Malta to work. in Malta. She insisted that being a part of the EU encourages interaction, from which it follows that today's EU citizens have to be able to interact with each other linguistically.

She mentioned the complexity of balancing the need to communicate efficiently with maintenance of linguistic and cultural diversity in Europe and expressed her belief in that language technology will help achieve the goals, referring to Umberto Eco's (1993) book "Search for the Perfect Language" set in a post Tower of Babel world. In Europe we seek the grammar of the language of Europe, and although we are unlikely to rediscover the perfect language, it is through technology like Machine Translation that we might hope to lower the language barriers. Machine Translation (MT) and language resource, as part of the DSM strategy, are playing an increasingly important role in generate opportunities for all in Europe

### 3.3. Session 3. Objectives of the Workshop

Khalid Choukri (ELDA/ELRA) thanked Dr. Grech, Brian Buhagiar (DGT local officer) and the workshop organizers for their support. Sharing linguistic resources is of key importance and is also the main activity for ELRA, his own organisation. He introduced the motivation for CEF.AT. He then turned to the objectives of the workshop, which mainly stand in relation to customising it to work with public services, administrations, ministries, etc. to find and share appropriate data to get the best results from CEF.AT for their particular needs and language. He invited everybody to participate and to share

their hopes and questions about the importance of language resources for language technology. He clearly stated the objectives of (i) sharing data, so they can be used to improve CEF.AT (ii) raising awareness and engage the participants (iii) help with legal and technical issues associated with the collection of data.

### 3.4. Session 4. Europe and Multilingualism

Brian Buhagiar (DGT Local Representative) described the European scene with its 24 official languages and 60 major regional/minority languages. He pointed out that all languages are considered equal in Europe and that Europe is committed to multilingualism. He described the EU's Digital Single Market strategy and the negative effects of language barriers upon private and public services. Pan-European public services face a multilingual challenge: 90% of EU web users prefer to use their own language in online services. Given that available language coverage by online translators is partial and not secure, and that human translation is not required for many use scenarios of public services, the solution is to build CEF-AT, aiming at making public digital services equally usable by all EU citizens and facilitating cross-border information exchange in public administration. The role of member states is therefore to take ownership of their own language and to make sure it is adequately supported in CEF-AT. Their involvement is essential.

### 3.5 Language and language technologies in Malta

#### Session 3.5a Language and language technologies in Malta

Prof. Ray Fabri (President, Il-Kunsill Nazzjonali tal-Ilsien Malti, Chairman, L-Istitut tal-Lingwistika / Institute of Linguistics, L-Università ta' Malta / University of Malta) began by providing a set of examples of Maltese words and sentences illustrating the mixed nature of Maltese, an amalgam of Arabic, Sicilian, Italian, English (Indo-European) elements. He then went on to paint a picture of the complex language situation in Malta, where the influence of English has given rise to various amalgams of the two languages ("Maltese English") and the extensive use of code switching. He elaborated on the legal status of Maltese, where it is enshrined as the only national language. It also shares the status of an *official* language with English, i.e. a language in which every citizen has the right to address the administration, and in which he may expect to receive an official answer. This is of course pertinent to the issue of translation. In practice, Maltese still tends to be the spoken language, whilst English the dominant written language (e.g. signage, etc.). After presenting some census statistics on the actual usage of the language, Fabri recounted developments. He singled out two particular events of importance that have had significant effects on the language itself and the way it is perceived: becoming an official EU language, and the founding of the National Council of the Maltese language. The latter remains under-resourced, he said. He went on to discuss the latest technological developments, which now include a substantial and searchable online corpus, POS tagger, and online dictionary, all of which are available to the public. He concluded that Maltese is a living, dynamic language faced with the challenges of balancing conservation and loss with modernisation and enrichment, and making sure that it survives in the digital world, in the context of rapid technological developments

### Session 3.5b The State and Status of Translation in Malta

Prof. Joe Eynaud (Head, Department of Translation and Interpretation) began with some historical information that led to the inception of the Department of Translation, Terminology and Interpreting Studies and two related Masters programmes in the area. He went on to provide statistics about the volume of translation now carried out at EC towards Maltese and the number of Maltese personnel employed in translation-related positions at EU institutions. He then turned to the local situation, stressing the need for regulation in the sector and the need for standardisation of qualifications.

Regarding material relevant for ELRC, he mentioned a number of terminology resources in the Health sector (at the Oncology Hospital) that had been produced by students and mentioned that he would be glad to discuss how these might be made available eventual incorporation into CEF.AT.

Clare Vassallo added a historical perspective on the role of the translator. The presentation focused on the status of translation and the translator, arguing that translation is now emerging from the shadows it was cast in the 19<sup>th</sup> century in Europe. From an 'invisible' act it is steadily drawing more attention to itself as a movement between systems, as a semiotic exchange of meaning, as a political tool of rewriting and manipulation with the potential for subversion, and as a cultural mediating force.

### 3.7. Session 7: Machine translation: how does it work?

Michael Rosner (National Anchor Point for Malta), from the Department of Intelligent Systems at the University of Malta, gave a brief overview the background and technical basics of automated translation. He covered four main questions: Why Machine Translation is Important; Why Translation is Hard for Machines and People; How Machine Translation Works; and finally, why MT with Maltese in particular is hard. First, he discussed the motivations to use MT, depending on the context, automated translation in addition to human translation and she gave an overview of the history of machine translation. The technical part of the presentation included both rule-based statistical MT approaches, through some basic illustrations, the data needed and the learning process. He concluded by highlighting the importance of abundant data, particularly in the case of Maltese.

### 3.8. Session 8: How can Public Institutions benefit from the CEF.AT Platform?

Carmen Carabott, Head of Language Applications of DGT, described in detail the MT@EC system in detail, i.e. the supported languages, the technologies upon which it is based, its user interface and other technical features, including input data format, delivery of results and security in document transfer. She continued by describing the CEF-AT platform, emphasizing that the focus will be the coverage of more domains (not just typical EU texts) and increasingly in-domain texts are needed to implement this move at sufficiently high quality. She pointed out that online services already connected to the generic MT@EC including IMI, an intranet for more than 7000 public administrations in the EU Member States, SOLVIT, an online problem-solving network concerning misapplication of internal market law by public authorities, nLex, a common gateway to



national Law, TED, an online information on tenders, CircaBC, a Communication and Information Resource Centre, Businesses and Citizens. Future customisation under CEF.AT would be aimed at Europeana, the digital European Library, ODP, the pan-European Open Data Portal for accessing open data infrastructures distributed over a EU and MS data repositories, EESSI, the Electronic Exchange of Social Security Information, a platform between 32 EU and EFTA countries, ODR, an Online Dispute Resolution platform, e-justice, a single point of access to law, enabling EU judicial cooperation, and SaferInternet, services to make Internet a trusted environment for children.

### **3.9/10. Sessions 9 and 10: The data for CEF-AT**

Since time was pressing, the next two sessions were combined. The first part addressed the kind of data being sought, and the second the legal framework under which it could be acquired. Both parts were given by Khalid Choukri

Khalid Choukri focussed on the data that needs to be collected for MT: texts (translations, aligned translations, collections of comparable texts), glossaries, terminological databases, dictionaries and lists of words in one or more languages. He discussed how different formats are useful to varying degrees and stressed the importance of metadata. He showed several examples of how language resources are produced from data. Finally, he encouraged the participants to contribute to the project by providing or identifying the required textual content.

The second part focused on the need for a clear and easy to follow regime for data reuse across the EU. He stressed the existence of legal and technical interoperability and simple redress mechanisms aiming to develop a single European market for innovative apps based on public data. He briefly presented the Public Sector Information (PSI) Directive and explained the structure of the data rights. He listed the stages for releasing the public data governed by the PSI Directive, and showed several case studies.

### **3.11. Session 13: Data and language resources: Technical and practical aspects**

Khalid Choukri presented in detail the workflow for the collection, processing and sharing of language resources. Most of the stages of this process, e.g. the identification of the data sources and datasets, the basic metadata documentation, data cleaning, privacy and ethics management are tasks in which the public sector providers will collaborate with ELRC. He then encouraged the audience to participate in these activities and work together with ELRC, and he showcased the mechanisms with which ELRC will fully support the providers throughout the whole process, i.e. the helpdesk and user forum mechanism, the ELRC repository and the ELRC website.

### **3.12. Session 12: Wrap Up**

Mike Rosner thanked all participants for having attended the workshop, speakers and contributors, Dr. Elena Grech, Dr. Khalid Choukri, Mr. Brian Buhagiar, Mr. David Sant, the student helpers Ms. Claudia Borg and Mr. Marc Tanti, and finally, the excellent

interpreters. He reassured those present that a feedback form would be sent without delay

## 4. Synthesis of Workshop Discussions

### 4.1. Panel 1: Public multilingual services in Malta

The Panel's moderator, Mr. Marc Bartolo, first explained the role of the Malta IT Agency (MITA), the central driver of Government's Information and Communications Technology (ICT) policy, programmes and initiatives in Malta. MITA's role is to deliver and implement the assigned programmes as set out in the Digital Malta National ICT Strategy 2014 - 2020, and to manage the implementation of IT programmes in Government to enhance public service delivery. MITA also provides the infrastructure needed to execute ICT services to Government.

He pointed out that MITA does not have translations of policies yet, since for technical matters priority is given to English. Typically budgets mention technologies but this doesn't include "presentation issues" particularly when this involves translation.

One of the problems in translated documents is code switching, which often occurs when an acceptable word in Maltese cannot be found. So simple translations for IT-related words such as "router", use English loanwords. Many other agencies within the Maltese public service dealing with other areas suffer the same problem, notably the Malta Financial Services Authority (MFSA) and the Malta Gaming Authority (MGA). Some standardisation in this regard would clearly be desirable.

VANNI Bruno, from the translation unit at the Office of the Attorney General, is involved with transposition of EU law. Accordingly, he was instrumental in translating fourteen thousand pages of the *Acquis Communautaire*. He indicated that generally, although most of this work of his unit concerns translation into both Maltese from English, a legal disposition is in place allowing transposition of EU law into English only when terminology is lacking as in e.g. aviation law, product safety legislation etc. Use of this disposition can be seen as an easy way out, but it must sometimes be invoked for practical reasons. He pointed out that the translation unit is poorly resourced. Although on paper there were two full timers and one part timer he spent much of the time working single-handed.

CLAUDIO Laferla is librarian at the Knowledge Management Centre, involved with research in educational materials for children. He pointed out that materials for the most part are EN texts found on internet, noting that it was exceedingly rare to find websites translated into MT. 12 yrs ago suggested the we gather all legends to have digital storytelling. With education you target two generations - parents and child.

Joseph Azzopardi, from MITA, has been involved with public sector information from a technical perspective. The idea of open data is a good one, so the PSI directive on open data is a good one because it forces government to make a conscious effort become more transparent about the data it holds. However, the concept of openness tends to be oversold within the political sphere. The essence of open data is linking - but most

politicians are more interested in the number of websites that have been opened rather than in the explicit connections between them. He concluded by stating that standardisation on the representation and quality of open data is crucial to the success of any effort to implement the directive on open data.

Thomas Pace, Secretary for Council for the Maltese Language, began by describing the mission of the Council. The council appoints various subcommittees whose function is to bring suggestions to main council in their designated areas of concern. Of particular interest to the present meeting is the translation committee which deals, amongst other things, with terminology in specialised areas e.g. geography. A book has been recently launched. A related issue is the visibility of Maltese in the public environment e.g. on signage in public buildings. Currently nearly all such signage is in English. Equally, slogans used by government are mostly English e.g. “one tablet per child”. The names of government offices and posts (e.g. “Permanent Secretary”) also tend to be in English, as are the names of rooms in buildings like the hospital. The general position of the Council is that such signs and names should be bilingual, if necessary (i.e. when the English name cannot easily be deduced from the Maltese one), with Maltese coming first, reflecting its status as a national language. He reported on some successes in implementing this philosophy e.g. at the new Parliament building

## 4.2 Panel 2: Data and language resources for Malta: where to find them

Mike Rosner explained that the purpose of the panel was to present language resources that actually exist with the aim of establishing potential future sources of benefit to CEF.AT.

Dr Albert Gatt, presented MLRS (Maltese Language Resource Server) which began life as a national project. When the MLRS portal opened in 2009 it was a manifestation of loose collaborations between a number of projects run by entities such as the Institute of Linguistics, Dept. of Intelligent Computer Systems and the Centre for English Language Proficiency. Since then, the collaboration has intensified and MLRS has acquired a status resembling that of a permanent programme, offering various resources and tools. Ġabra is an online, full-form Maltese-English dictionary, which sports a human-friendly web-based front end (see <http://maltesedictionary.org.mt>) thanks to a grant from the Vodafone Foundation to make it available as a national resource. MLRS also includes a collection of linguistic tools (POS tagger, a tokeniser, and a sentence splitter) as well as a number of searchable corpora, the latest and largest of which comprises 250 million words, with POS tagging, lemmatisation, and root annotation.

Dr Gatt mentioned a number of ongoing collaborations, notably with EU/DGT for the enhancement of corpora, through incorporation of EU documents, and also enhancement of lexical resources, through incorporation of terminology DBs etc. In addition, he outlined the ongoing collaboration with University partners like CELP / IoL collaboration on enhancing Learner Corpora and related resources, as well as the IoL Project on the construction of a digital, multimodal dictionary of Maltese Sign Language.

Dr Gatt was asked whether MLRS resources are available to users outside the University of Malta. Anybody interested can register to use the system. Dr. Gatt

responded by mentioning that MLRS enjoys a sizeable user-community with users of many kinds: researchers, students, translators, educators, writers, members of the general public. In principle the resources are available for anyone to use.

Carl-Joseph Farrugia, DOI, Office of the Prime Minister, presented the Government Gazette, the official publication of the Government in which every deed printed in the newspaper and signed by the competent authorities considered official is published. This includes Proclamations of the President, Police Notices, Government Notices, Notices to Seaman, Opportunities, Government offers and job announcements, court notices, etc. First published in 1813 in Italian, the Gazette was subsequently published in English and Italian, and in 1935, when Maltese was recognized as an official language, a Maltese version replaced the Italian one. After independence in 1964 the Maltese version took precedence over the English one. Until two years ago the newspaper was produced as a paper publication and sold to the public, but since 2003, an electronic version has been uploaded to the Government website at [www.doi.gov.mt](http://www.doi.gov.mt). Currently pdf versions in Maltese and English are available for download and a user-friendly web version is under development.

Chris Gruppetta, Director of Publishing at Merlin Publishing, gave a brief resume of the company which was set up in 1964. It has since grown into Malta's most innovative publisher, focusing primarily on children's books with a small but prestigious imprint in non-children's fiction. Home to most of Malta's key authors, Merlin prides itself on its backlist of Maltese literature, as well as on its nurturing of young new authors with their innovative styles and expression. Regarding language resources, Merlin has always been open to the possibility of making its publications available for linguistic research, provided that the works are protected by an appropriate licencing agreement.

Victor Cauchi, Office of the Attorney General, who was actively involved in translating the *Acquis Communautaire* into Maltese raised the issue of which kind of Maltese is the most suitable for legal translation: Semitic Maltese, in which our body of laws have been written since the 1940s and which was largely followed by law drafters throughout subsequent years, or the Italianized court language of lawyers, in which judgments are written. Laws in Malta are written both for legalistic persons, taking into account past uses and adaptations, and for the man-in-the-street alike, the latter having to understand laws in their native language to be able to abide by them. Despite these subtleties, the problem of volume and immediacy required by the EU, coupled with staff shortages, renders the editing and vetting of EU laws even more problematic to fine-tune the complex and continental language style of EU laws both into the legal mindset and terminology of English, the language in which our laws are first drafted, and thereafter into Maltese. He recalled that when the status of Maltese as an official language was being discussed, it would help offering Arab-speaking residents a reference language they could understand, naturally, but only if made up in great proportion of a Semitic element. Alas, this was not to happen.

### 4.3 Interactive Session 12: How can we engage?

The session was animated by Dr Albert Gatt who oriented the session towards the following main themes:

1. Some initiatives have managed to motivate significant public engagement. What is the key to securing this level of public engagement? How feasible is it to harness such engagement in the development of language resources (e.g. via crowdsourcing)?
2. Projects – both within academia and industry – require collaboration and data sharing. Do we have a sufficiently well-rooted collaborative culture? Is there a case to be made for building a national infrastructure of language resources?
3. Are language resources only for researchers? Can they have a more practical purpose, for instance, in education, translation etc? Do people in these (and other) sectors feel the need for them?
4. Language resources are in fact being created in a lot of places. Can they be of use beyond their immediate function and for target users beyond the ones who create them? Do we need to talk of a "digital heritage under construction", with a potentially broad pool of users?

He then invited speakers to orient their remarks towards these themes

Mike Spagnol, (Department of Maltese, University of Malta) presented the Kelma Kelma Facebook page, which was set up with the objectives of getting people to engage with Maltese, and sharing his own knowledge of Maltese with Maltese speakers. The Kelma Kelma page has two main aspects: a repository of existing idioms and an environment for innovation aimed at the development of new idioms. The main observation was that the Facebook page had been and remains outstandingly successful - with a number of likes within the Maltese top 30. The page started off as a game - then people began to use it much more than expected, having become a reference point for many Maltese. An impressive level of input from the public is received. People do not simply like and forget - but often post every day. The material posted includes both material relevant to new words, and comments. When asked to explain this success, Mike Spagnol suggested the following aspects were key: (i) that every interaction with public was short and concise and (ii) humour played an important role (iii) the page supports two-way communication, being able to collect information as well as provide it (iv) that being hosted on Facebook integrates well with other activities that contributors might be involved with. Main problems seen were the need to filter input, and how to handle content in the long term, i.e outside the realm of the Facebook page.

Odette Vassallo presented the Centre for English Language Proficiency (CELP), where she is Director, strives to research the use of English in Malta. One of its foci is investigating how to address some of the gaps in learners' proficiency levels in order to design short courses and facilitate support in English. In collaboration with the Institute of Linguistics, it has developed a learner corpus of Maltese speakers of English. The Corpus of Learner English in Malta (CLEM) includes a collection of written examination scripts in English collected over a period of three years from learners aged 15 to 18 years. At present, this corpus serves two main functions: first, it is a database that researchers utilise in the field of second language acquisition, bilingualism, Maltese variety, learner errors, etc.; and second, it is being analysed with a view to creating a national

framework of English in Malta which will serve as the link to the Common European Framework of Reference (CEFR) for Languages. Its developers also plan to involve education stakeholders in order to encourage its use by teachers and learners. Unfortunately, this plan is not immediately attainable. As we outlined during our participation at the ELRC workshop, the online interface is not teacher friendly and a more flexible and intuitive interface or platform is needed to attract already busy teachers to utilise this important resource. The ideal interface should allow teachers to integrate it as a classroom tool that does not require lengthy preparation or a short course.

Mark Spiteri (Malta Communications Authority) presented the Dizzionarju tal-Malti - an online dictionary that originated from the collation of a singular, plural or collective'>number of existing lexical resources for Maltese, some of which were available online, while others were developed as a result of research projects by academics and postgraduate students. The online version of the dictionary is one of the outcomes of the Maltese Language Resource Server project (<http://mlrs.research.um.edu.mt>), a long-term initiative by the Institute of Linguistics and the Department of Intelligent Computer Systems at the University of Malta. The electronic version of the dictionary was initially developed by John J. Camilleri and subsequently fine-tuned and edited by the MLRS team. The dictionary was further developed and made public as a result of a strong collaboration between the Malta Communications Authority, the Institute of Linguistics, the Department of Intelligent Computer Systems at the University of Malta, as well as the Kunsill Nazzjonali tal-Ilsien Malti, which mediated the collaboration agreement. We have also made this dictionary available in a mobile-friendly version and as a downloadable application from Google Play to allow to use this service anytime, any place.

The main conclusions of the discussion were

- The importance of collaboration with bodies, especially government agencies, such as MCA
- Resources are useful for a variety of potentially interested parties, not just the usual suspects. E.g. corpora for teaching/pedagogy (CLEM as an example)
- Collaboration between university and DGT of the European Commission is of key importance for obtaining language resources.
- Resource generation can take place via crowdsourcing, but success often requires an engaging manner of presentation (kelma kelma is a good example)
- When we asked whether we need a "national framework" for curation of language resources, a counter-argument was that it is better to seek to align our resources to existing international frameworks

## 5. Workshop Presentation Materials

The workshop presentations can be accessed at the event webpage, at [http://lr-coordination.eu/malta\\_agenda](http://lr-coordination.eu/malta_agenda)

And a few pictures of the workshop.



## 6. Appendix

### 6.1. Invitation letter (in Maltese)

Għażiż Sinjur/a,

Għandna pjaċir nistidnuk għall-edizzjoni Maltija tal-workshop tal-Koordinazzjoni tar-Riżorsi tal-Lingwi Ewropej (European Language Resource Coordination - ELRC) li se jsir fis-26 ta' Frar 2016 f'Dar l-Ewropa, 254, Triq San Pawl, il-Belt VLT 1215.

Dan il-workshop huwa bla ħlas u jinkludi ftit ikel u xorb ħafif kif ukoll buffet f'nofsinhar. Madankollu l-postijiet huma limitati u għalhekk nitolbuk biex ġentilment tikkonferma l-partecipazzjoni tiegħek fuq

[http://www.lr-coordination.eu/malta\\_registration](http://www.lr-coordination.eu/malta_registration)

Il-lingwi u d-diversità lingwistika huma fil-qalba tal-kultura, il-kummerċ, in-negozju u l-kollaborazzjoni Ewropea. Sabiex tiffaċilita l-komunikazzjoni multilingwi u l-iskambju tad-dokumenti bejn l-amministrazzjonijiet pubbliċi nazzjonali fl-Ewropa, il-Kummissjoni Ewropea tipprowdi pjattaforma ta' Traduzzjoni Awtomatika fi ħdan il-Connecting Europe Facility (CEF.AT). Il-pjattaforma hija bla ħlas għall-istituzzjonijiet pubbliċi kollha. Hija tibni fuq is-sistema ta' traduzzjoni mekkanika MT@EC li kienet żviluppata mid-Direttorat Ġenerali tat-Traduzzjoni tal-Kummissjoni Ewropea u li kienet offruta lill-amministrazzjonijiet pubbliċi tal-Istati Membri. Il-Programm CEF se jtejjeb il-kopertura u l-kwalità tas-sistema tat-traduzzjoni, u se jagħmilha lesta biex tiġi integrata kompletament għas-servizzi pubbliċi pan-Ewropej fuq l-internet.

Fl-ELRC workshop ta' Malta, nixtiequ li nidhlu fi djalogu miftuħ miegħek biex nifhmu aħjar il-ħtiġijiet ta' kuljum tal-amministrazzjonijiet pubbliċi Maltin fl-ambjent multilingwi tax-xogħol tal-lum u biex nagħmlu s-CEF.AT xieraq għal dawn il-ħtiġijiet. Il-programm tal-workshop qiegħed mehmuż ma' din l-ittra elettronika. Esperti ddedikati mill-Kummissjoni Ewropea, tekniċi tal-lingwi, fornituri tas-servizz tal-lingwi u amministrazzjonijiet pubbliċi nazzjonali u korpi governattivi se jaħdmu flimkien biex jidentifikaw dak li għandek bżonn mil-lat ta' traduzzjoni awtomatika, se jidentifikaw b'mod kongunt ir-riżorsi multilingwi, u se jiddiskutu kwalunkwe kwistjoni teknika u legali involute fl-użu tad-dejta għat-traduzzjonijiet awtomatiċi. Għalhekk ir-reazzjoni tiegħek u l-partecipazzjoni tiegħek huma kruċjali biex nagħmlu s-CEF.AT xierqa għalik.

Jekk jinteressak li tuża s-CEF.AT jew jekk tixtiequ li taġġorna ruħek dwar l-iżviluppi attwali u possibbli fir-rigward tat-traduzzjoni awtomatika, aħna nkunu ferħanin li nipprovdulek id-dettalji neċessarji waqt dan il-workshop.

Għal aktar informazzjoni jekk jogħġbok ikkuntattja l-organizzatur lokali tal-workshop, Mike Rosner, [mike.rosner@um.edu.mt](mailto:mike.rosner@um.edu.mt), +356 2340 2542.

Nistennew bil-herqa li nisimgħu mingħandek.

Dejjem tiegħek,

Kimmo Rossi, European Commission, DG CONNECT, Unit G3 – Data Value Chain



## 6.2. Workshop Agenda (in Maltese)

### Agenda għall-Workshop tal-ELRC, 26 ta' Frar 2016

Dar l-Ewropa, 254, Triq San Pawl

Valletta VLT 1215, MALTA

08:00 – 09:00 Registrazzjoni

09:00 – 09:05 Ftuħ (Mike Rosner, ELRC National Anchor Point)

09:05 – 09:20 Merħba lokali (il-Kap tar-Rappreżentanza tal-KE, Dott. Elena Grech)

09:20 – 09:30 Miri u għanijiet tal-Workshop (Dr Khalid Choukri, ELRC)

09:30 – 09:50 L-Ewropa u l-Multilingwiżmu (Brian Buhagiar, Direttorat Ġenerali tat-Traduzzjoni KE)

09:50 – 10:10 Il-Qagħda tal-Lingwa f'Malta (Prof. Ray Fabri, Kunsill tal-Ilsien Malti)

10:10 – 10:30 L-Istat u l-Istatus tat-Traduzzjoni f'Malta (Prof. Joe Eynaud, Prof. Clare Vassallo, Università)

10:30 – 11:10 **Panel A: Servizzi Pubbliċi Multilingwi f'Malta (Jippresjedi: Mark Bartolo, MITA)**

- Thomas Pace, Kunsill tal-Ilsien Malti
- Dott. Joseph Azzopardi, MITA (PSI, Open Data)
- Dott. Claudio Laferla, Knowledge Management Centre, MITA
- Dott. Vanni Bruno, Unità tas-Servizzi tal-Ġustizzja, Uffiċċju tal-Avukat Ġenerali

**11:10 – 11:40 Waqfa għall-kafè u biex insiru nafu 'l xulxin**

11:40 – 12:00 Traduzzjoni Awtomatika: Kif taħdem? (Mike Rosner, Università ta' Malta)

12:00 – 12:30 Kif jistgħu l-Istituzzjonijiet Pubbliċi jgawdu mill-Pjattaforma CEF.AT? (Carmen Carabott, Direttorat Ġenerali tat-Traduzzjoni, KE)

**12:30 – 13:30 Waqfa għall-lkel**

13.30 – 14.00 X'dejta hija meħtieġa? (Dr Khalid Choukri, ELRC)

14:00 – 14:30 Il-Qafas Legali għall-Kontribuzzjoni tad-Dejta (Dr Khalid Choukri, ELRC)

14:30 – 15:10 **Panel B: Riżorsi tad-Dejta u l-Lingwa f' Malta: fejn tista' ssibhom (Jippresjedi: Mike Rosner)**

- Albert Gatt, MLRS, Università ta' Malta
- Carl-Joseph Farrugia, DOI, Uffiċċju tal-Prim Ministru (Gazzetta tal-Gvern)
- Chris Gruppetta, Direttur tal-Pubblikazzjoni, Merlin Publishing
- Victor Cauchi, Uffiċċju tal-Avukat Ġenerali

**15:10 – 15:40 Waqfa għall-kafè u biex insiru nafu 'l xulxin**

15:40 – 16:00 Id-Dejta u r-Riżorsi Lingwistiċi: Aspetti Tekniċi u Prattiċi (ELRC)

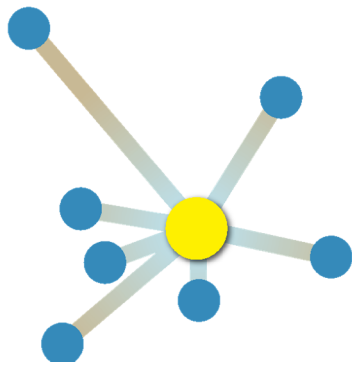
16:00 – 16:30 **Sessjoni Interattiva: Kif Inżidu l-Involviment (Jippresjedi: Dr Albert Gatt, Università ta' Malta)**

- Dott. Michael Spagnol, Kelma-Kelma
- Dott. Odette Vassallo, Ċentru għall-Profiċenza fl-Ingliż, Università ta' Malta
- Mark S. Spiteri, Awtorità Maltija tal-Komunikazzjoni (MCA)
- Carmen Carabott, Direttorat Ġenerali tat-Traduzzjoni, KE

16:30 – 17:00 Diskussjoni: L-Aqwa Prassi għall-Futur – Nisfruttaw u Nkabbru l-Valur tad-Dejta Tiegħek (Jippresjedi: l-ELRC)

17:00 – 17:15 Għeluq: Konkluzjonijiet u Impenji Immedjati (Mike Rosner, Università ta' Malta)

### 6.3. Attendance Certificate (in English)



# European Language Resource Coordination

*Connecting Europe Facility*

This is to certify that

---

attended the ELRC Workshop, Malta at Dar L'Ewropa, 254 St Paul's Street, Valletta,  
on 26th February 2016

Michael Rosner  
National Anchor Point  
University of Malta