



Deliverable Task 2

Frequently Asked Questions (FAQ) – and their Answers



Author(s): Andrea Lösch (DFKI)
Maria Koutsombogera (ILSP)
Aivars Berzins (Tilde)

Dissemination Level: Public

Version No.: <V1.1>

Date: 2016-02-04



Contents

<u>1</u>	<u>Executive Summary</u>	<u>3</u>
<u>2</u>	<u>Provision of data: How? Why? What are our data used for?</u>	<u>4</u>
2.1	What is going to happen to the data we provide?	4
2.2	How can data be provided?	4
2.3	Why should we (public institutions) actually provide data?	4
2.4	We (public institutions) don't have any data for you!	4
2.5	We cannot just share our data with you – they are confidential!	4
<u>3</u>	<u>CEF AT, MT@EC and translation needs in the public administration</u>	<u>6</u>
3.1	What is MT@EC? What is CEF AT?	6
3.2	How can we access MT@EC?	6
3.3	Why would we need MT@EC? We have human translators!	6
3.4	Why should we support MT@EC / CEF AT – we can have our own national solution?	7
3.5	MT is opposed to young people should learning foreign languages.	7
3.6	Machine translation will never work for our languages.	7
<u>4</u>	<u>Managing and harvesting language data - why and how?</u>	<u>8</u>
4.1	Why should I care about translations and get hold of language data?	8
4.2	How should I manage my data and why? We don't have any resources!	8
<u>5</u>	<u>Best Practice for workshop organisation public sector collaboration</u>	<u>9</u>
5.1	Preparatory actions - Checklist	9
5.2	Whom to invite and how to invite them?	10
5.3	Drafting the agenda	10
5.4	Arranging the venue, catering and interpretation	11
5.5	PR and press releases	11
<u>6</u>	<u>Annex 1 – Sample Participant Lists</u>	<u>13</u>
6.1	Sample Participant List Greece	13
6.2	Sample Participant List Germany	15
6.3	Sample Participant List Latvia	16
<u>7</u>	<u>Annex 2 – National PSI Experts</u>	<u>17</u>

1 Executive Summary

The Frequently Asked Questions (FAQ) prepared in this document are based on the actual questions from workshop participants and National Anchor Points raised during the ELRC workshops and LRB meeting. This guide intends to give an overview of typical questions that may be raised during ELRC events and corresponding responses across the different themes:

- Provision of data in the framework of ELRC: How? Why? And what are the data used for?
- What is CEF AT and MT@EC? What are the translation needs of the public sector?
- Managing language resources: How and why?

Last but not least, this guide contains a section on best practice for the workshop organization and collaboration with the public sector.

2 Provision of data: How? Why? What are our data used for?

2.1 What is going to happen to the data we provide?

The data will go to the EC (DG Translate) to support the improvement of the machine translation system MT@EC.

2.2 How can data be provided?

Data can be donated through 2 channels: the ELRC channel and the DGT channel.

ELRC offers different options from which data donors can choose:

- Send data to ELRC via email: See email on the ELRC website (<http://www.lr-coordination.eu/resources>)
- Upload data directly into the ELRC repository: See access point on the ELRC website (<http://www.lr-coordination.eu/resources>)

For those providers wanting to deliver their data directly to the DGT, they can either send them by email (corresponding email address to be provided by the EC) or upload the data to a secure ftp server provided by DGT.

2.3 Why should we (public institutions) actually provide data?

Supporting your own language is supporting Europe and vice versa. Only with your help and with the provision of your language resources, CEF AT can be made fit to your needs. Within the CEF programme, CEF AT is available for free to public administrations in all EU member states and CEF affiliated countries (Iceland and Norway). So for your data, you receive a better service.

2.4 We (public institutions) don't have any data for you! We work only paper-based. We outsource our translations.

If translations are done externally, typically, you as contractor, can ask for the translation memories of your data. It is best to negotiate with the language service provider and make sure, that you obtain the translation memories and/or additional data relating to your translation. This is important to do in any case, because for subsequent translations, you may be able to negotiate a better price. In addition, there are several benefits associated with language data you hold (see point 4.2 below).

2.5 We cannot just share our data with you – they are confidential!

Most data held by the public sector actually is public data. Many ministries have – at minimum – various types of information available online for their citizens (e.g. news, legal texts, official communications, interviews, brochures, background information etc.) and this information is also typically provided at least in one foreign language. In Germany, for instance, the website of the national government is always at least tri-lingual, all information is provided in German, English, and French. Even more, in addition to several thousands of

Frequently Asked Questions (FAQ) and their Answers

tri-lingual news articles published by the federal government and ministries, there are more than 4.000 tri-lingual official brochures created by the different German federal ministries on all topics relevant to these ministries

(http://www.bundesregierung.de/SiteGlobals/Forms/Webs/Breg/Suche/EN/Infomaterial/Solr/Infomaterial_Startseite_Formular.html?nn=771722). The translation memories and original files of such data can all be shared without any worries.

3 CEF AT, MT@EC and translation needs in the public administration

3.1 What is MT@EC? What is CEF AT?

MT@EC is the current tool for machine translation used and provided by the EC. It exists already since 26th of June 2013. It has a web user interface in 24 languages for a human-to-machine use case, or can be used as a web service in a machine-to-machine scenario. It uses a highly secured protocol (sTESTA) coupled with the European identification ECAS which guarantees confidentiality of data. MT@EC can be used by any Member State administration free of charge. More detailed information is available online at http://ec.europa.eu/dgs/translation/translationresources/machine_translation/index_en.htm

CEF AT (Automated Translation platform of the Connecting Europe Facility CEF) is part of CEF Digital to provide automatic translation services with the goal of making digital services accessible to anyone everywhere from whatever language into the user's language. CEF AT should empower in particular European public online services such as Europeana (<http://www.europeana.eu/>), the Open Data Portal (<https://open-data.europa.eu/en/data/>), the Online Dispute Resolution Platform etc. More detailed information on CEF is available online at <http://ec.europa.eu/digital-agenda/en/connecting-europe-facility>

3.2 How can we access MT@EC?

MT@EC can be used by any Member State administration free of charge. It can be accessed as follows:

- Staff working for **EU institutions or agencies** can use MT@EC with their ECAS account credentials.
- Staff working for a **public administration** in an EU country should follow these steps:
 - Sign up for your personal ECAS account and password (using only your professional email address).
 - Send an email to DGT-MT@ec.europa.eu asking to have an ECAS account. Indicate what your job involves and which public administrative body you work for. Don't forget to include your full signature.
 - DGT will create your MT@EC account and notify you.

3.3 Why would we need MT@EC? We have human translators!

Machine translation / MT@EC is substantially helping to make the translation process more productive to help you translate more efficiently. Translators at the EC are responsible for translating content into all official EU languages. More than 2 000 translators are currently employed by DG Translation, while another 5 000 work at EU institutions. In 2014, they translated more than 2.3M pages.

MT@EC is used daily for French, Spanish, Portuguese and Italian to produce initial translations that are then post edited in a very efficient way. For some other languages (e.g. German) the quality of translation is not yet high enough for heavy usage. Even though in the last year, significant progress could be achieved here through domain specific engines. In particular for reports and texts (e.g. from economics), MT@EC is used successfully to

Frequently Asked Questions (FAQ) and their Answers

automatically translate these documents. In other cases, the tool is helpful for rapidly scanning long texts in a foreign language and deciding, e.g. which passages to focus on for human translation.

Overall, the quality of translation is directly linked to the availability of good data in this language: If the data for MT is good, then the MT system will be good as well. This is why we are here: To improve the MT@EC tool with your data so it is of better use for you.

3.4 Why should we support MT@EC / CEF AT – we can have our own national solution?

Typically, national solutions are targeted on particular range of topics. Hence, the scope of MT@EC is broader and more comprehensive. By supporting MT@EC, participants can expect to have access to a broader service.

3.5 Machine translation is directly opposed to our national policy that young people should learn foreign languages.

Not necessarily. Machine translation can actually provide a good basis for learning languages:

- Initially, it can be used to bridge the gap for people who are not able to speak a particular language, until they have acquired some initial language skills. For instance, at university level, machine translation is actually used to cover the original language gap of foreign students: Lectures are translated automatically and simultaneously into English, so that foreign students can get an idea what they are about and catch up for further language integration.
- For users who already have some proficiency in a foreign language, machine translation can be used to improve their language skills. E.g. when translating books or texts, users get some direct experience with the foreign language, they see possibly new sentence constructions or how things they did not know before could be translated. Like this, they learn by practice and by experience.

3.6 Machine translation will never work for our languages (e.g. Estonian, Finnish, Hungarian and other morphologically rich languages).

It is true that certain languages are more difficult for current MT systems to handle, because of their free morphology or their free constituent order. However, other methods are being explored. For example, in the Hungarian workshop the MT expert presented a new model (based on neural networks) that seems more suitable for languages like Hungarian, Finnish or Estonian. Moreover, the European Commission funds several actions (see e.g. <http://www.qt21.eu/>) to investigate and develop MT solutions for languages which currently receive only sub-optimal MT support.

What needs to be stressed is that, no matter what the methodology is, huge amounts of (preferably) parallel resources are needed for the systems implementation, since these systems rely on machine learning; and this need for data is primarily being addressed in the workshop objectives.

4 Managing and harvesting language data - why and how?

4.1 Why should I care about translations and get hold of/keep corresponding language data?

For a start, if you translate internally, human translation can be done more easily and faster if you share and can built on previous translations from the same domain. Moreover, if you outsource your translations, you can at least negotiate with your language service providers a better price in return for sharing data. Last but not least, keeping hold of your data and managing them adequately can have several additional benefits (see point 4.2 below).

4.2 How should I manage my data and why? We don't have any infrastructures or resources (especially small translation services)!

Especially in the public sector there is a great diversity in the organization (or absence of organization) of the management of translations and corresponding data flows: from digitized workflows with term lists and storage of translation memories up to almost purely paper-based workflows.

From an organizational point, much benefit can arise even from small changes in the organizational dealing with language data. Suggested actions that can be taken without major effort include:

- Analysis of all phases of data development
- Based on this, creation of a “data management plan” (DMP), even very basic one:
 - Which data is important?
 - Where is it stored?
 - Can it be further processed?
- Document all relevant data
- If possible, use the web as additional publication channel and reap benefits of linked data (see <http://www.w3.org/DesignIssues/LinkedData.html>)

(Check presentation “Best practice for the future: Capitalize on your valuable data”)

5 Best Practice for workshop organisation public sector collaboration

5.1 Preparatory actions - Checklist

The preparation phase of the workshop can take several weeks up (typically 2 months) and it involves the following steps:

- Step 1: Contact your local DGT Field Officer:
 - http://ec.europa.eu/dgs/translation/getintouch/localoffices/offices_en.htm
- Step 2: Get contacts to the national public administration from your local DGT field officer
 - Note: The public administrations - ministries, central authorities etc. - are the main target audience of your workshop!
- Step 3: Try to involve the country's experts and practitioners in the area of Open Data
 - E.g. local data.gov.x, EU Open Data Portal anchors, your national PSI experts (see Annex 2 or <http://ec.europa.eu/digital-agenda/en/news/public-sector-information-group-main-page> for current list)
- Step 4: Finalise the list of speakers and invitees as soon as possible, as latest **4-6 weeks before the event!**
- Step 5: Have the **invitations** signed by the EC and send it out to the participants no later than **4 weeks before the event.**
- Step 6: Adapt the workshop agenda to the local context dynamics **4 weeks prior to the event**
 - Central questions: Who are the speakers going to be, why, etc.
- Step 7: Identify a venue large enough to host all participants of your workshop which also provides the opportunity for simultaneous translation
 - Note: Check first with your DGT local field office / the EU representation in your country
- Step 8: Have the details of the workshop published on the ELRC website

It is important to note that if you do have any questions or if you are unsure about some steps or actions, the ELRC representative responsible for your country will help you. It is together with him that you organise this workshop and it is with him you need to agree on the principal modalities. The ELRC representatives are as follows:

- **Khalid Choukri** and his team from ELDA (choukri@elda.org, mazo@elda.org) is responsible for France, Spain, Portugal, Italy, Malta, Belgium, the Netherlands, Luxembourg, Germany, United Kingdom and Ireland.
- **Andrejs Vasiljevs** and his team from Tilde (andrejs@tilde.lv, aivars.berzins@tilde.lv) is responsible for Latvia, Lithuania, Estonia, Finland, Sweden, Norway, Denmark and Iceland.

Frequently Asked Questions (FAQ) and their Answers

- **Stelios Piperidis** and his team from ILSP (spip@ilsp.gr, mkouts@ilsp.gr) is the ELRC representative responsible for Greece, Cyprus, Bulgaria, Romania, Croatia, Slovenia, Austria, Czech Republic, Slovakia, Hungary and Poland.

5.2 Whom to invite and how to invite them?

If you don't know who to look for in your national public administration and which participants might be the right one for an ELRC workshop, **Annex 1** provides **several sample lists of workshop participants** from past ELRC workshops.

Moreover, in terms of process, there are **several different steps** involved and to be considered when sending out invitations to your participants:

- Step 1: There is a master invitation letter in EN available to you
- Step 2: Translate/localize it in your language. The letter / email should contain the following pieces of information:
 - Date & venue
 - Contact person
 - A draft agenda (you may use the master agenda in EN if you do not have a draft one)
 - Online registration address from the ELRC website
 - Stress that the event is free-of-charge
- Step 3: Ask your ELRC representative to have the invitation letter signed by the European Commission (DGT Connect or DGT) – as the latest **5 weeks** prior to the event
- Step 4: Send out the official invitations no later than **4 weeks prior to the event**.
- Step 5: You may set up a local email account from which the email will be sent (e.g. we created elrc@ilsp.gr for the Athens event)

Further **hints and tips** regarding the invitation of participants from the public sector include:

- When sending out the invitations, request a read receipt. In this way, you can monitor registration (who did not read or chose not to send a read receipt, who has read but did not register and who read and registered)
- Keep close and careful track of this process (invitation → acceptance → registration) e.g. via a respective excel file
- Send reminders to invitees that did not react/register
- If possible, follow up with participants by phone
- Send thank you/confirmation emails 1 or 2 days before the event to remind registered people to appear!

5.3 Drafting the agenda

The typical sample agenda and master slides in English are all available via the NAP Information Package. However, it is important to note that the workshop programme will be **your (!) programme**, which needs to take into account the local conditions, frame and technicalities. As such, **in direct collaboration and agreement with your ELRC**

Frequently Asked Questions (FAQ) and their Answers

representative, you are free to adjust the agenda **according to your national needs**. Of course, all central themes must be covered – but the contents should be adapted according to the particular situation in your country and the knowledge of your audience.

Lessons learnt especially during the first workshops were as follows:

- A full day event might be too long. Shorter and easier to follow presentations are more efficient
- Time in general might present a problem if speakers overrun ⇔ Keeping the schedule is not always easy!
 - Pressing the speakers to hurry led to the speakers speaking faster
 - Which, in turn, made the presentations difficult to follow
 - And the interpretation task really challenging!
 - If you can, prep your speakers and maybe even make a dry-run
- Engagement of the audience within the panels: Try to take a local approach and look for strong moderators

5.4 Arranging the venue, catering and interpretation

- Workshop language: The official national language(s) in your country
- Arranging translation / interpretation services
 - The workshop should be mainly presented in the local national language
 - Slides should also be in the local national language (exception: external experts, e.g. from ELRC, who don't speak local language, may present in English)
 - Simultaneous interpretation must be provided for each workshop in local national language – EN
 - Workshop language: The official national language(s) in your country
- Arranging translation / interpretation services
 - The workshop should be mainly presented in the local national language
 - Slides should also be in the local national language (exception: external experts, e.g. from ELRC, who don't speak local language, may present in English)
 - Simultaneous interpretation must be provided for each workshop in local national language – EN

5.5 PR and press releases

For each workshop, there should be at least the following PR activities:

- ELRC Workshops Press Release to announce the upcoming workshop / to announce the results, outcomes or important traits of a past workshop:
 - The English templates are available via the NAP Information Package.
 - They need to be translated the draft press release into local language.
 - For some countries, the press release may seem too long. You may want to create a shorter version.

Frequently Asked Questions (FAQ) and their Answers

- What we found efficient is to use the shorter press release as email body in the message circulated, with the long (official) version attached. In this way, the press will use the short version and whoever is interested will read the long one.
- Once published, monitor press release impact with respect to the registrations
- Further hints and tips:
 - The human interest angle is key: Illustrate the impact and importance of your news for the target audience
 - Provide answers to the following questions:
 - **Who?** Who are the key players — ELRC, other national bodies, local industry, EC? Who does your news affect/who does it benefit?
 - **What?** What is new or newsworthy about your information?
 - **Why?** Why is this important news? What makes it different or outstanding?
 - **Where?** Where is this happening? Does geography or location matter?
 - **When?** Is the timing important? Is it part of a bigger action?
 - **How?** How did this come about?
 - Important: Pick an outstanding feature of the specific workshop and focus on that in the two short paragraphs.
 - Use brief but positive quotes of relevant stakeholders (e.g. on the importance of the workshop to be held, on the value of the outcomes of the workshop etc.)
- Examples of Headlines for Workshop Press Releases:
 - ELRC holds First Groundbreaking Workshop in Berlin
 - Second ELRC Workshop in Athens, Greece Another Huge Success
 - ELRC Workshop Series Continues Positive Trend in Riga, Latvia

6 Annex 1 – Sample Participant Lists

6.1 Sample Participant List Greece

Organisation
Bank of Greece
Civil Aviation Authority
Democritus University of Thrace
Dep.of History, Ionian University, Corfu
DFKI
EC Representation in Greece
EC- DGT Athens
ELDA
European Commission
Freelance translator
General Secretariat for Industry
Greek Free/Open Source Software Society
Hellenic Army
Hellenic Ministry of National Defence/General Directorate of Defence Policy & International Relations
Hellenic Police
Hellenic Single Public Procurement Authority
Hellenic Statistical Authority
Hellenic Telecommunications Organization
ILSP/RC Athena
IMIS/RC Athena
Institute of Eastern Mediterranean
Institute of National Relations
Managing Authority "Reform of the Public Sector"
Ministry of Economy
Ministry of Foreign Affairs
Ministry of Health
Ministry of Interior and Administrative Reform
Ministry of Justice
Ministry of Labour
Municipality of Acharnai
National and Kapodistrian University of Athens
National Documentation Centre
National Museum of Contemporary Art
National School of Public Administration
NGO of the Archbishopric of Athens
Special Secretariat for Water

Frequently Asked Questions (FAQ) and their Answers

TAUS
Technological Educational Institute of Athens
The Media Institute, UCL
The National Centre for Public Administration and Local Government
TMServe
Union of the Graduates of the School of Public Administration (ENAP)

6.2 Sample Participant List Germany

Auswärtiges Amt	3
Bundesinstitut für Berufsbildung	1
Bundesministerium der Verteidigung	2
Bundesministerium des Innern	2
Bundesministerium für Arbeit und Soziales	1
Bundesministerium für Finanzen	1
Bundesministerium für Gesundheit	1
Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit	1
Bundesministerium für Verkehr und digitale Infrastruktur	1
Bundesministerium für Wirtschaft und Energie	2
Bundesnetzagentur	1
Bundessprachenamt	3
Bundesverband der Dolmetscher und Übersetzer	1
Deutsche Bahn	2
Deutscher Bundestag	1
DFKI GmbH	6
DGT Local Field Office Berlin	1
DIN e.V.	1
Europäische Kommission	1
European Language Resources Association	2
Fraunhofer-Gesellschaft	1
GlobalSprachTeam	2
Institut für Deutsche Sprache	1
Institute for Language and Speech Processing, "Athena" RC	1
iRights	1
Karlsruhe Institute of Technology	2
Ministerium für Wirtschaft und Energie Brandenburg	1
Rat für Sozial- und Wirtschaftsdaten	1
Senatskanzlei Berlin	1
Staatskanzlei des Saarlandes	1
Volkswagen AG	1
ZF Friedrichshafen AG	2

Frequently Asked Questions (FAQ) and their Answers**6.3 Sample Participant List Latvia**

Academic Information Centre	2
Culture Information System Centre (KISC)	1
DFKI	1
ELRC	1
European Commission, DGT Translation	1
European Commission, Representation in Latvia	1
Institute of Mathematics and Computer Science, Univ. of Latvia	2
Latvian Academy of Sciences Terminology Commission	1
Latvian Association of Local and Regional Governments	1
Latvian Geospatial Information Agency	4
Latvian Parliament Saeima	3
Ministry of Education and Science	1
Ministry of Environmental Protection and Regional Development of Latvia, Terminology Commission	1
Ministry of Health	1
Riga Technical University	1
Sportacentrs (sports news website)	1
State Education Development Agency	1
State Employment Agency	1
State Regional Development Agency	1
The Latvian Institute	1
The Latvian Language Agency	2
The Latvian National Standardisation Body Latvian Standard (LVS)	1
The Latvian State Language Center	5
The Ministry of Environmental Protection and Regional Development	1
The National Archives of Latvia	2
The Office of Citizenship and Migration Affairs	3
The State Chancellery	1
Tilde	2
Tilde	1
Tilde/ELRC	2
Translate 24/7	1
University of Latvia Agency "Latvian Language Institute of the University of Latvia	2

7 Annex 2 – National PSI Experts

LIST

PSI and sub group governmental experts

Update January 2016

No	Name	Country	Signature
		Member States	
1	Christian Muller	Austria	Christian.Mueller@bmwfw.gv.at
2	Jean-Charles Quertinmont	Belgium	Jean-Charles.Quertinmont@premier.fed.be
3	Krasimira Dimitrova	Bulgaria	kdimitrova@mtitc.government.bg
4	Anelia Dimova	Bulgaria	adimova@mtitc.government.bg
5	Tomas Kroupa	Czech Republic	tomas.kroupa@mvcz.cz
6	Aristos Malaos	Cyprus	AMalaos@dits.mof.gov.cy
7	Anamarija Musa	Croatia	povjerenica@pristupinfo.hr
8	Zoran Lusa	Croatia	zoran.lusa@uprava.hr
9	Adam Arndt	Denmark	arndt@digst.dk
10	Antti Eskola	Finland	antti.eskola@tem.fi
11	Perica Sucevic	France	Perica.SUCEVIC@modernisation.gouv.fr
12	Aet Rahe	Estonia	Aet.Rahe@mkm.ee
13	Allar.Viik	Estonia	Allar.Viik@mkm.ee
14	Christian Horn	Germany	Christian.Horn@fb.hamburg.de
15	Kalliopi Angeletopoulou	Greece	k.ageleto@ypes.gov.gr

Frequently Asked Questions (FAQ) and their Answers

No	Name	Country	Signature
16	Brian Costello	Ireland	Brian.Costello@per.gov.ie
17	James Daly	Ireland	James.Daly@per.gov.ie
18	Gabriele Ciasullo	Italy	ciasullo@agid.gov.it
19	Toms Celmillers	Latvia	toms.celmillers@varam.gov.lv
20	Kestutis Andrijauskas	Lithuania	kestutis.andrijauskas@ivpk.lt
21	Anna Komer	Luxembourg	Anne-Julie.Kommer@smc.etat.lu
22	Patrick Weber	Luxembourg	Patrick.Weber.Ext@act.etat.lu
23	Azzopardi Joseph	Malta	joseph.s.azzopardi@gov.mt
24	Hayo Schreijer	Netherlands	Hayo.Schreijer@koop.overheid.nl
25	Groń Maciej	Poland	Maciej.Gron@mac.gov.pl
26	Cláudia Gonçalves Barroso	Portugal	claudia.s.goncalves@ama.pt
27	Andrei Nicoara	Romania	andrei.nicoara@gov.ro
28	Anna Andrejsinova	Slovakia	anna.andrejsinova@justice.sk
29	Robert Dobrovodsky	Slovakia	robert.dobrovodsky@justice.sk
30	Renata Zatler	Slovenia	Renata.Zatler@gov.si
31	Ales Versic	Slovenia	AVersic@gov.si
32	Manuel Ruiz del Corral	Spain	manuel.ruiz1@seap.minhap.es

No	Name	Country	Signature
33	Karina Aldén	Sweden	karina.alden@regeringskansliet.se
34	Erik Borålv	Sweden	erik.boralv@vinnova.se
35	Malcolm Todd	United Kingdom	Malcolm.Todd@nationalarchives.gsi.gov.uk
		EFTA	
36	Heather Broomfield	Norway	heather.broomfield@difi.no
37	Jean-Luc Cochard	Switzerland	Jean-Luc.Cochard@bar.admin.ch